SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(94th Edition)

OCTOBER 2020

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING)*
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

PREFACE

This is the **Ninety Fourth** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **October, 2020.** The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A novel approach to design optimal 2-D digital differentiator using vortex search optimization algorithm, Suman Yadav, Richa Yadav, Ashwni Kumar and ***3.Manjeet Kumar***, Electronics, DTU

2. A Systematic Review on Firefly Algorithm: Past, Present, and Future, Vijay Kumar and ***3.Dinesh Kumar***, Electronics, DTU

3. An Authenticated Enrolment Scheme of Nodes using Blockchain and Prevention of Collaborative Blackhole Attack in WSN, ***3.R K Yadav*** and ***6.Rashmi Mishra***, CSE, DTU

4. An overview of conventional and non-conventional hydrogen production methods, ***6.Ravindra Kumar***, ***3.Anil Kumar*** and ***3.Amit Pal***, Mechanical, DTU

5. Analysis of a novel Metamaterial Absorber using equivalent circuit model operating at 3.5 GHz, ***6.1.Priyanka Garg*** and ***3.Priyanka Jain***, Electronics, DTU

6. Application of industry 4.0 technologies in SMEs for ethical and sustainable operations: Analysis of challenges, ***6.Ravinder Kumar***, Rajesh Kr. Singh and Yogesh Kr. Dwivedi, Mechanical, DTU

7. Artificial neural network models for global solar energy and photovoltaic power forecasting over India, ***6.Gulnar Perveen***, ***3.M. Rizwan***, Nidhi Goela and Priyanka Anand, Electronics and Electrical, DTU

8. ATT: Attention-based Timbre Transfer, Deepak Kumar Jain, ***3.Akshi Kumar***, Linqin Caix, ***8.Siddharth Singhal*** and ***8.Vaibhav Kumar***, CSE, DTU

9. BOHR RADIUS FOR SOME CLASSES OF HARMONIC MAPPINGS, ***3.S. SIVAPRASAD KUMAR*** AND ***6.KAMALJEET GANGANIA***, Mathematics, DTU

10. Comparative Study of different Missile shapes using Computational Fluid dynamics, ***8.Gaurav Kumar***, ***8.Rohit Kumar***, ***8.Sagar Dagar*** and ***3.Raj Kumar Singh***, Mechanical, DTU

11. Crowd Emotion Analysis Using 2D ConvNets, ***6.Gaurav Tripathi***[1], ***3.Kuldeep Singh***[2] and ***3.Dinesh Kumar Vishwakarma***[3], [1]Electrical, [2]Electronics and [3]IT, DTU

12. Data Aggregation Algorithms in IoT: An Organized Evaluation of The Literature, ***3.Rajesh Kumar Yadav*** and ***6.Monica Gupta***, CSE, DTU

13. Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease, ***8.Surendrabikram Thapa***, Priyanka Singh, Deepak Kumar Jain, Neha Bharill, Akshansh Gupta and Mukesh Prasad, CSE, DTU

14. Deep NeuralNet For Violence Detection Using Motion Features From Dynamic mages, ***7.Aayush Jain*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

15. Deep Transfer with Minority Data Augmentation for Imbalanced Breast Cancer Dataset, ***6.Manisha Saini*** and ***3.Seba Susan***, IT, DTU

16. Diode Based Reconfigurable Microwave Filters for Cognitive Radio Applications: A Review, HASHINUR ISLAM, ***7.1.SAUMYA DAS***, TANUSHREE BOSE AND TANWEER ALI, Electronics, DTU

17. End to End Binarized Neural Networks for Text Classification, Harshil Jain, ***8.Akshat Agarwal***, Kumar Shridhar and Denis Kleyko, Electrical, DTU

18. Evolutionary Relationship of Polycyclic Aromatic Hydrocarbons Degrading Bacteria with Strains Isolated from Petroleum Contaminated Soil Based on 16S rRNA Diversity, ***6.Sakshi***, ***3.S. K. Singh*** and ***3.A. K. Haritash***, Environmental, DTU

19. Experimental and Theoretical Interpretations of Spectral Behavior of 6-Methoxyflavone, Nisha Fatma, ***3.Mohan Singh Mehata***, Nupur Pandey and Sanjay Pant, Applied Physics, DTU

20. Explicit finite element analysis of single pass rolling using ABAQUS CAE, ***3.Sanjay Kumar*** and ***8.Pranjal Shiva***, Mechanical, DTU

21. Fabricating and Assembling Acoustic Metamaterials and Phononic Crystals, Christabel Choi, ***8.1.Shubhi Bansal***, Niko Münzenrieder and Sriram Subramanian, Physics, DTU

22. Gate Oxide Variability Analysis of a Novel 3 nm Truncated Fin–FinFET for High Circuitry Performance, ***8.Mridul Prakash Kashyap*** and ***3.Rishu Chaujar***, Physics, DTU

23. Improved ant colony optimization for achieving self-balancing and position control for balancer systems, ***6.Rupam Singh*** and ***3.Bharat Bhushan***, Electrical, DTU

24. Improvement of Throughput using Dual Connectivity in Non-Standalone 5G NR Networks, ***7.Sakshi Dubey*** and ***3.Jasraj Meena***, IT, DTU

25. IPDroid: Android Malware Detection using Intents and Permissions, ***8.Kartik Khariwal***, ***8.Jatin Singh*** and ***3.Anshul Arora***, Mathematics, DTU

26. Managing supply chains for sustainable operations in the era of industry 4.0 and circular economy: Analysis of barriers, ***3.Pravin Kumar***, Rajesh Kr Singh and Vikas Kumar, Mechanical, DTU

27. Mechanical and wear behaviour of Friction stir processed surface composite through Self-Assembled Monolayer Technique, ***6.Ravi Butola***, ***3.Ranganath M Singari*** and ***3.Qasim Murtaza***, Mechanical, DTU

28. Modeling information diffusion in online social networks using a modified forest-fire model, ***3.Sanjay Kumar***, Muskan Saini, Muskan Goel and B. S. Panda, CSE, DTU

29. Multimodal Multi-Task Financial Risk Forecasting, Ramit Sawhney, Puneet Mathur, Ayush Mangal, ***8.Piyush Khanna***, Rajiv Ratn Shah and Roger Zimmermann, CSE, DTU

30. Multi-objective Reinforcement Learning based approach for User-Centric Power Optimization in Smart Home Environments, Saurabh Gupta, ***8.Siddhant Bhambri***, Karan Dhingra, Arun Balaji Buduru and Ponnurangam Kumaraguru, CSE,DTU

31. Performance Evaluation of a Pulse Detonation Engine using Vortex Generators as obstacles in the combustion chamber, ***8.Mukesh Yadav***, ***8.Avhishek Jha***, ***8.Manish Bista*** and ***3.B.B Arora***, Mechanical, DTU

32. Review – polymeric materials for energy harvesting and storage applications, ***6.Sarita S Nair***, Sujeet K Mishra, and ***3.D. Kumar***, Applied Chemistry, DTU

33. The Effect of Variable Electrolyte Thickness of LMO vs Li Metal Foil Lithium-Ion Coin Cell using Simulation, ***8.M Aditya***, ***8.Naveen Kumar***, ***8.Chetan Kumar*** and ***3.Amrish K. Panwar***, Applied Physics, DTU

34. Weakly Supervised Metric Learning with Majority Classes for Large Imbalanced Image Dataset, ***3.Seba Susan*** and ***7.Ashu Kaushik***, IT, DTU

35. Web Security in IoT Networks using Deep Learning Model, ***3.Sanjay Patidar*** and ***7.Inderpreet Singh Bains***, CSE, DTU

1. *Chancellor*

2. *Pro Vice Chancellor* ***2.1.*** *Ex Pro Vice Chancellor*

3. *Faculty* ***3.1.*** *Ex Faculty*

4. *Teaching-cum-Research Fellow* ***4.1.*** *Asst. Librarian*

5. *Alumni*

6. *Research Scholar* ***6.1.*** *Ex Research Scholar*

7. *PG Scholar* ***7.1.*** *Ex PG Scholar*

8. *Undergraduate Student* ***8.1.*** *Ex Undergraduate Student*

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-10012-6

A novel approach to design optimal 2-D digital differentiator using vortex search optimization algorithm

Suman Yadav [1,2] · Richa Yadav [2] · Ashwni Kumar [2] · Manjeet Kumar [3]

Received: 22 June 2020 / Revised: 8 September 2020 / Accepted: 15 September 2020
Published online: 10 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

The designing of 2-D digital differentiator is multimodal and high dimensional problem which requires large number of differentiator coefficients to be optimized, hence conventional design techniques will not lead to optimal 2-D differentiator design. Metaheuristic approaches is a good approach to handle multimodal and high dimensional problem under consideration. In this paper, a novel method for the design of FIR two-dimensional (2-D) digital differentiator with quadrantally odd symmetric properties is proposed. The coefficients of 2-D digital differentiator are computed and optimized using vortex search optimization (VSO) algorithm with the support of L_1 error objective function. The unique combination of VSO and L_1 error fitness function aids in achieving flatter magnitude response in designing of 2-D differentiator. Further, some comparative analysis with the well-known established techniques like cuckoo search algorithm (CSA), particle swarm optimization (PSO) and real coded genetic algorithm (RCGA) methodology is reported over full frequency bands, in order to demonstrate the advantage of the proposed novel 2-D digital differentiator design approach. To measure the performance of proposed system, some parameters are selected for illustration such as absolute magnitude error, mean absolute error, standard deviation and execution time. The excellence of the proposed approach is observed in approximating the ideal response of 2-D differentiator in contrast to other algorithms and to provide global optimal solution.

Keywords Digital FIR 2-D differentiator · Vortex search algorithm · Optimization · Metaheuristic algorithm

1 Introduction

Digital filters are extensively accepted in many applications due to their superior performance, less area requirement during implementation, ease in implementation with adaptive

✉ Manjeet Kumar
manjeetchhillar@gmail.com

Extended author information available on the last page of the article

programmability and exact reproduction capability. In past few decades extensive research is reported by practitioners in the area of designing and implementation of digital differentiator and integrator, due to their applications in various fields like signal processing for linear and nonlinear system identification [18, 35], digital image processing to extract edges in an image [25, 26, 32], biomedical signal processing to detect R-peaks in electrocardiogram (ECG) signal [19–21], control engineering for disturbance rejection [13], radar to measure position and acceleration [29] and many more. The main function of the differentiator is to determine the time derivative of a signal to extract information regarding rapid transients. For differentiator applications in real time and their practical implementation, it is desirable that they have a wide band frequency response and minimum group delay with low order to save hardware cost. Digital differentiators are mainly of two types FIR and IIR and further FIR differentiators can be either type III or type IV. Type III are antisymmetric having odd length impulse response while type IV are symmetric with even length response. Type IV are considered better because their frequency response is defined over full range of frequency and meet design requirement well compare to type III [34]. Differentiators act like high pass filter, their magnitude response amplifies high frequency component present in input signal. Approximation and interpolation-based methods are the two conventional approaches to design 1-D and 2-D integrator and differentiators. Taylor series approximation [17] have been used for differentiator design, and it approaches the ideal frequency response closely but with increased filter order. Tseng et al. [36, 37] designed 2-D digital filter and considered 16 type of cases as per symmetry and anti-symmetry in both directions as per the length of filter, used fractional derivative constraint and gradient search optimization. This technique used Lagrange multiplier and least square method to determine odd symmetric filter coefficients. They also proposed an idea of using hyperbolic sine and cosine functions to design digital log differentiator, with an advantage of less storage space requirement for storing filter coefficients than traditional approaches. Still there were challenges in approximating the ideal frequency response of the differentiator as close as possible.

Many real time problems are solved by natures inspired meta-heuristic optimization algorithms [10, 14, 15, 24, 41–44] published in last two decades. Initially wide research was done in designing 1-D digital differentiator and integrators [6, 11] later shifted to the designing and implementation of 2-dimensional systems. Upadhyay et al. [38] designed recursive wideband differentiator and integrator by optimizing the location of poles and zeroes. Kumar et al. [22, 23] in 2014 designed fractional order differentiator using radial basis function and window to truncate coefficients to approximate the fractional order derivative for a digital signal. Error fitness function was reduced with improved frequency response characteristics. Interior search algorithm was also proposed with integral square error fitness function to design digital differentiator with only one control parameter and higher convergence rate was achieved. Yadav et al. [12] used particle swarm optimization (PSO), genetic algorithm (GA) and hybrid PSO-GA optimization to design fractional order differentiator by reducing error fitness function amplitude. In 2016 Aggarwal et al. [1, 3–5] designed 2-D digital differentiator using cuckoo search algorithm and L_1 error fitness function with quadrantally odd symmetric properties and further also used fractional derivative constraint to design optimal 2-D FIR filter. Chang [7] implemented the design of 2-D fractional differentiator using differential evolution algorithm with an excellent searching ability. Wulf et al. [39] designed efficient low pass FIR differentiator suitable for biomedical signal processing with reduced filter order using Fourier series of triangular function. Mahata et al. [27, 28] implemented the design of wideband digital FIR differentiator and integrator using Harmony search algorithm and later

also applied hybrid flower pollination algorithm and particle swarm optimization to design digital wide band differentiator and integrator. It improved the quality of solution, convergence rate and reduce the time to optimize the system. Mohan et al. [30] designed fractional digital differentiator with hybrid shuffled frog leaping algorithm and the proposed technique efficiently removed the noise from ECG signal with good stability and convergence speed. Suman Yadav et al. [40] proposed vortex search optimization to design 2-D FIR filters and reduced error fitness function value and execution time to optimize the filter. Aggarwal et al. [2] utilized hybrid combination of Bat algorithm and Particle Swarm optimization algorithm to design 5th-, 7th- and 11th-order FIR differentiator using the L_1-method and contributed superior simulation results in contrast with traditional approaches of designing digital differentiator.

These meta-heuristic techniques as discussed above have drawbacks such as premature convergence, requires a lot of parameter to be tuned with complex nature, higher execution time and not able to achieve global optimum solution due to stagnation in local optima. So, to overcome the above disadvantages, proposed research designed digital differentiator that employ vortex search optimization (VSO) [8, 31]. Its capability to solve these critical problems depends on its searching behaviour that includes an exclusive step size adjustment strategy which helps it to maintain effective balance in exploration and exploitation.

VSO is based on vortical flow of stirred fluid, excellent in exploration and exploitation of the search space. The unique technique of step size adjustment strategy helps it to reach a global optimum solution without stucking in multiple local optimal solution with least execution time. This algorithm has been also used for the optimization of analog filter group delay, in wireless sensor network by deploying the sensor nodes in their best optimal position and for optimization of the fore part of KCS container ship [9, 16, 33]. In these problems, it outperforms the other algorithm such as Particle swarm optimization, artificial bee colony and harmony search algorithms. It do not require a lot of parameters except the number of iteration, number of candidate solution and the upper and lower limit of the problem identified. Filter coefficients are optimized by going through various trials by setting different values of control parameters. To prove its superiority, it has been compared with various other optimization techniques such as cuckoo search optimization, particle swarm optimization and genetic algorithm and offer these advantages as mentioned below:

- VSO algorithm provides least overshoot at the edges in frequency response.
- Lower execution time in obtaining global optimum solution.
- Least magnitude error with sharp transition from pass band to stop band.
- VSO is simple to implement, computationally efficient as compare to population based optimization hence it is a good choice for the real time life problems.

In this paper a digital differentiator of order 14*14 is designed using optimization algorithms VSO, by evaluating the impulse response of 2-D digital differentiator in such a way that it approximates the ideal magnitude response of 2-D digital differentiator as much as possible. Its simulation results that includes magnitude and phase response are compared to cuckoo search (CS) algorithm, particle swarm optimization (PSO) and real coded genetic algorithm (RCGA) over full frequency bands. It is observed from the experimental result that the VSO based 2-D differentiator results in least magnitude error and execution time compared to other techniques.

Rest of the paper is organized as follows: Section 2 describes vortex search optimization algorithm, its mathematical equations and strategy followed to avoid local optimal solution. It

also includes the design formulation of 2-D FIR differentiator. Section 3 includes results simulation and analysis. Section 4 concludes the paper and section 5 presents references.

2 Vortex search algorithm

Vortex search optimization technique is recently proposed meta-heuristic algorithm and based on vortical flow of stirred fluids. For a 2-D optimization problem, a vortex pattern can be created using a number of nested circles with reduced radius. Every possible solution to the problem is said to be individual solution whereas combination of solutions is termed as population. To achieve the global optimum output, solutions are updated in every iteration. To design 2-D differentiator, system coefficients are assumed to be candidate solution C_s in algorithm, best differentiator coefficients are the optimized candidate solution with minimum error fitness function. Flow chart to design a FIR 2-D digital differentiator based on VSO is given in Fig. 1.

2.1 Mathematical description of VSO algorithm

Step 1. Outer most circle of the vortex is centered in the search space with an initial centre and radius defined below in Eqs. (1) and (2):

$$initial\ centre\ (\mu) = \frac{upper_{limit} + lower_{limit}}{2} \tag{1}$$

$$initial\ radius\ (\sigma) = \frac{upper_{limit} - lower_{limit}}{2} \tag{2}$$

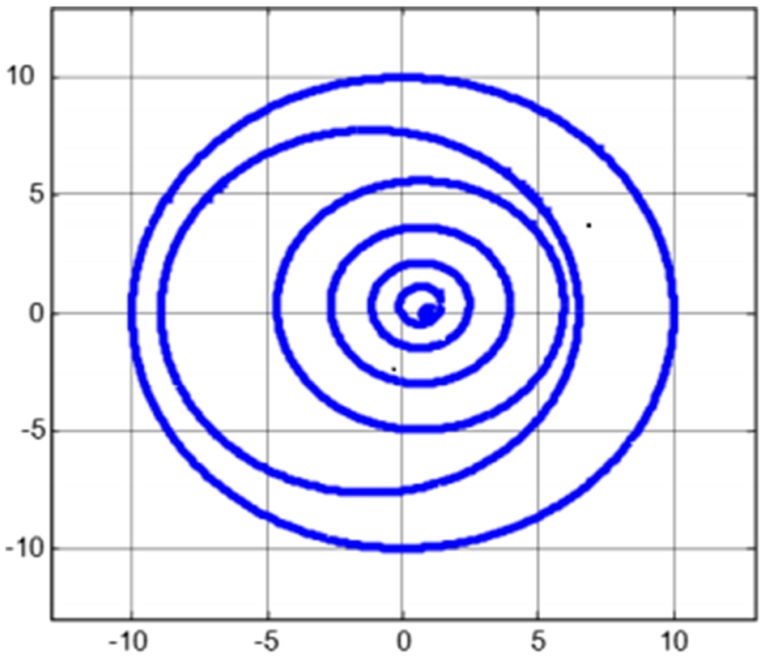


Fig. 1 Searching pattern of vortex search algorithm [8]

$upper_{limit}$, $lower_{limit}$ are the limits of solutions within range [−1 1] in two dimensional space. Outer circle is centered in such a way so as to maximize the coverage of search space with *initial radius* (σ) and *initial centre* (μ)

Step 2. All the candidate solutions are distributed in two-dimensional search space of the vortex using Gaussian distribution function.

Step 3. Candidate solutions are not allowed to move outside the range specified in step 1. Error fitness value is computed for each candidate solution. A solution with least error fitness function value is selected and taken as best solution.

Step 4. Best solution from previous step is taken as the centre to create second vortex and its radius will be reduced using inverse incomplete gamma function in Eq. (3) given below:

$$r_t = r_0 . \frac{1}{x} . gammaincinv(x, b_t) \tag{3}$$

$$\gamma(x, b) = \int_0^x e^{-t}\, t^{b-1}\, dt \quad b > 0 \tag{4}$$

Here $x = 0.1$, $b \epsilon [0\ 1]$. The parameter b of the VSO algorithm in Eq. (4) determines the resolution of the search space, its value can be equally sampled within the defined range and therefore resolution can be adjusted. At each iteration the value of b can be taken as:

$$b_t = b_0 - \frac{t}{Max\ Iteration}$$

$b_0 = 1$ to ensure the coverage of the search space, t is the current iteration and *Max Iteration* are the total number of iteration taken.

Step 5. Now the process is again repeated to introduce candidate solution in this second next inner vortex circle. Best candidate solution is selected with minimum error fitness function, which replaces the current best solution taken as the centre of next circle with reduced radius and this process comes to an end in two cases either the searching criteria is fulfilled or the maximum number of iterations N is over.

Step 6. Radius reducing process is known as step size adjustment technique. It initiates with exploration and then transforms into exploitation. As value of radius decreases in each step, it needs to be tuned to the searching style.

Step 7. Optimum solution provided by VSO is the centre of innermost circle with minimum radius shown in Fig. 1.

2.2 Two-dimensional fir differentiator design

2.2.1 Frequency response of 2-D digtal differentiator

Linear phase FIR differentiator transfer function is given by:

$$A(z_1, z_2) = z_1^{-L_1}\ z_2^{-L_2}\ \sum_{l_1=-L_1}^{L_1}\sum_{l_2=-L_2}^{L_2} a(l_1, l_2) z_1^{-l_1}\ z_2^{-l_2} \tag{5}$$

$A(z_1, z_2)$ defines the transfer function of a two-dimensional linear phase FIR system in Eq. (5). L_1 and L_2 are the length in both the directions. $a(l_1, l_2)$ is a matrix that represents the impulse response of the system. The objective is to determine the value of $a(l_1, l_2)$ and approximating the ideal magnitude response of the designed FIR system. Whereas $z_1 = e^{j\omega_1}$, $z_2 = e^{j\omega_2}$ and ω_1, ω_2 are the frequency in both the directions.

$$-\pi \leq \omega_1, \omega_2 \leq \pi$$

Impulse response needs to follow certain condition to be quadrantally odd symmetric [39] given below in Eq. (6).

$$a(l_1, l_2) = -a(-l_1, l_2) = -a(l_1, -l_2) = a(-l_1, -l_2) \tag{6}$$

The impulse response defined in Eq. (1) is a matrix of order $(2L_1 + 1) \times (2L_2 + 1)$ with a condition in Eq. (7)

$$a(l_1, 0) = 0;\ a(0, l_2) = 0 \tag{7}$$

Frequency response of 2-D FIR differentiator is given by Eq. (8):

$$A\left(e^{j\omega_1}, e^{j\omega_2}\right) = e^{-j\omega_1 L_1} e^{-j\omega_1 L_2} (j)^2 \widehat{A}(\omega_1, \omega_2) \tag{8}$$

$\widehat{A}(\omega_1\ , \omega_2)$ is the amplitude response given by Eq. (9)

$$A\ (\omega_1, \omega_2) = \sum_{l_1=-L_1}^{L_1}\sum_{l_2=-L_2}^{L_2} b(l_1, l_2) \sin l_1\omega_1 \sin l_2\omega_2$$
$$b(l_1, l_2) = 4a(l_1, l_2) \tag{9}$$

Frequency response of Ideal 2-D digital differentiator having linear phase represented in Eq. (10):

$$P\left(e^{j\omega_1}, e^{j\omega_2}\right) = (j)^2 \omega_1, \omega_2$$
$$-\pi \leq \omega_1, \omega_2 \leq \pi \tag{10}$$

2-D FIR differentiator designing is considered to be an optimization problem. In computing coefficients a, such that difference between the $\widehat{A}(\omega_1\ , \omega_2)$ and the ideal response $P(\omega_1, \omega_2)$ is minimized using optimization. Finite impulse response of 2-D digital differentiator follows anti-symmetric property.

The error fitness function is defined as:

$$\widetilde{E} = \sum_{\omega_1} \sum_{\omega_2} \left| \widehat{A}(\omega_1, \omega_2) - P(\omega_1, \omega_2) \right| \tag{11}$$

Here, L_1 error fitness function is selected instead of least square and minimax to design the 2-D differentiator as it gives a flat response in pass band and stop band of magnitude response.

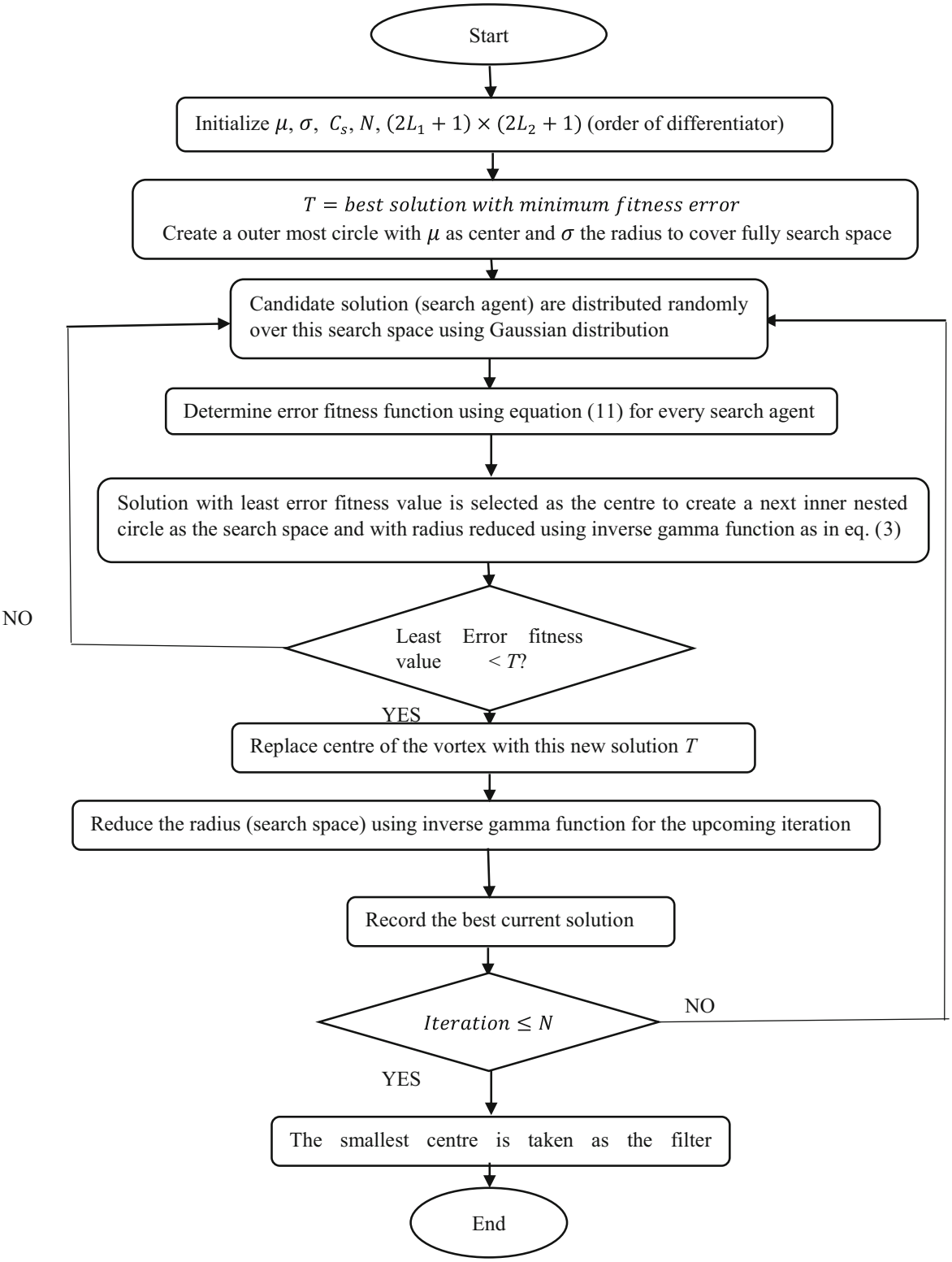


Fig. 2 Flow chart for 2-D differentiator design based on VSO

Also it yields minimum overshoot at the band edges with sharp transition. Eq (11) represents the absolute difference between the magnitude response of designed and the ideal differentiator. As the value of error reduces, more the frequency response of the system approaches the ideal response.

2.3 Methodology to design 2-d differentiator using VSO

2.3.1 Steps to design 2-D differentiator using employed algorithm

Step 1. Set the initial value of center and radius of outermost vortex as μ and σ.
Step 2. Randomly generate candidate solutions as C_s, initialize maximum number of iterations N, dimension of differentiator coefficient matrix as $(2L_1 + 1) \times (2L_2 + 1)$. Each candidate solution represents a matrix having dimension of differentiator coefficients.
Step 3. Create a vortex over search space and candidate solutions are randomly distributed over this vortex using gaussian distribution.
Step 4. Evaluate error fitness function in Eq. (11) for each candidate solution. Select the candidate solution with minimum error fitness value. Draw the next circle with this candidate solution as the center and radius is reduced inverse incomplete gamma function as in Eq. (3).
Step 5. Repeat step 3 and 4 till a candidate solution is obtained with minimum fitness function value, or the maximum number of iterations is over.
Step 6. This particular candidate solution with minimum fitness function value represents the optimal 2-D differentiator coefficient.

3 Results and simulations

In this section simulated results are presented and analyzed to design a two-dimensional differentiator based on VSO as shown in Fig. 2 and same are compared with recently published algorithms such as CSA, PSO and RCGA. In order to keep fairness in comparison, common parameters for all the algorithms are taken same such as number of iteration, lower and upper limit of the solution coefficients, order of the filter and the range of frequency. Solution is taken after a minimum fitness value is achieved The range for the 2-D frequency components ω_1, ω_2 lies between $[-\pi, \pi]$. VSO requires least numbers of parameters to be controlled or tuned as given in Table 1. The upper and lower limits for the differentiator coefficients are taken as −1 and 1 and length of coefficient matrix is taken as 15×15. The

Table 1 Controlling parameters defined for various algorithm

Algorithms Parameters	VSO	RCGA [5]	PSO [5]	CSA [5]
Population size	55	100	90	25
Iterations cycle	1000	1000	1000	1000
Lower Limit	−1	-1	-1	-1
Upper Limit	1	1	1	1
Tolerance	–	10^{-6}	10^{-6}	10^{-6}
Crossover rate	–	0.8, 1.2	–	–
Mutation rate	–	0.01	–	–
Learning Parameter	–	–	2,2	–
Particle velocity	–	–	0.01, 1	–
Inertia weight	–	–	0.4	–
Alien eggs probability	–	0.9	–	0.25

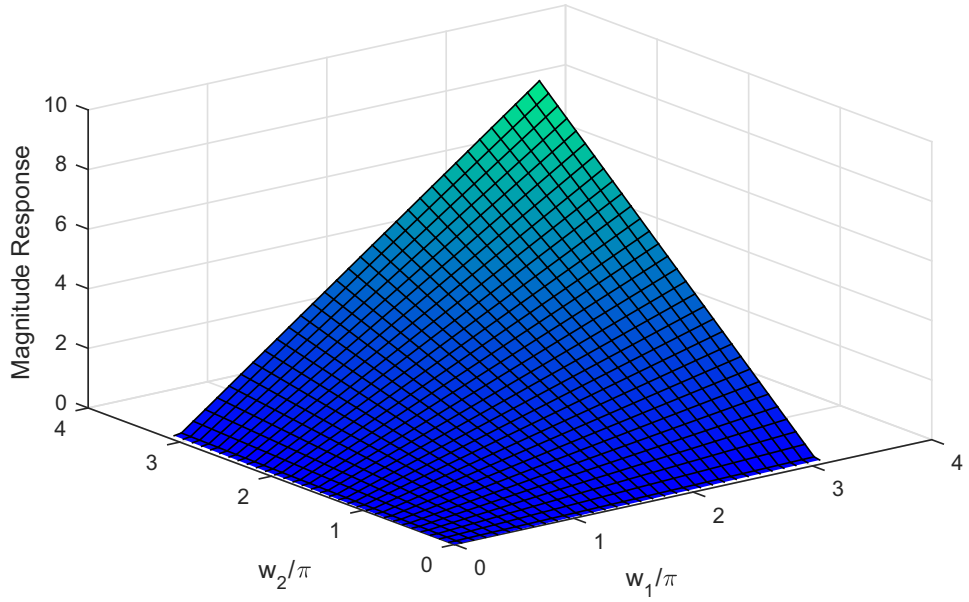


Fig. 3 Magnitude response of 2-D Ideal digital FIR differentiator

magnitude and phase response for the ideal differentiator is depicted in Figs. 3 and 4. For an ideal differentiator, frequency response is directly proportional to the frequency, while the phase response is constant for all the frequency as shown in Fig. 4. Number of possible solutions are taken as 55, its value is different for all algorithms to which VSO is compared. Parameters need to be tuned to get global optimized solution, selected after many trials by making variation in their magnitude but within specified limits, are mentioned in Table 1. There is no method available in literature to choose exact parameters value, their value can vary with different set of problem.

Optimized coefficients are determined using VSO algorithm is reported in Table 2, on the basis of these coefficients, magnitude response and phase response obtained for 2-D

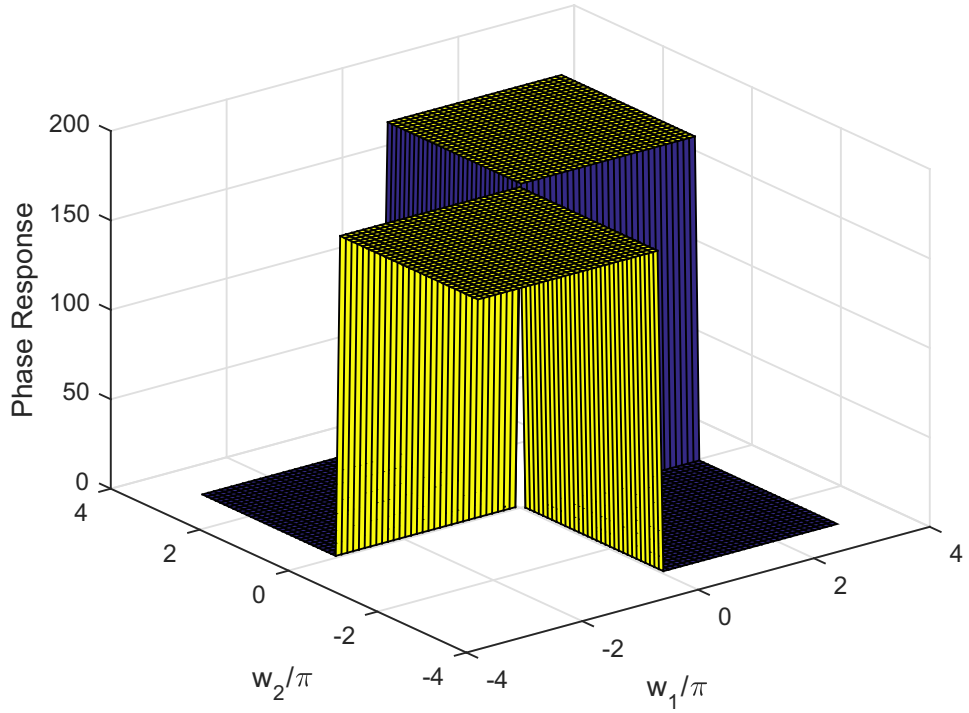


Fig. 4 Phase response of 2-D Ideal digital FIR differentiator

Table 2 Optimized Coefficient 2-D FIR DD with order 15 × 15

Optimized Coefficients				VSO
$a(0,0)$	$=-a(0,14)$	$=-a(14,0)$	$=a(14,14)$	−0.004168160
$a(0,1)$	$=-a(0,13)$	$=-a(14,1)$	$=a(14,13)$	−0.002177658
$a(0,2)$	$=-a(0,12)$	$=-a(14,2)$	$=a(14,12)$	0.004305658
$a(0,3)$	$=-a(0,11)$	$=-a(14,3)$	$=a(14,11)$	0.007895813
$a(0,4)$	$=-a(0,10)$	$=-a(14,4)$	$=a(14,10)$	0.008467236
$a(0,5)$	$=-a(0,9)$	$=-a(14,5)$	$=a(14,9)$	0.008563646
$a(0,6)$	$=-a(0,8)$	$=-a(14,6)$	$=a(14,8)$	0.000000000
$a(1,0)$	$=-a(1,14)$	$=-a(13,0)$	$=a(13,14)$	−0.002638470
$a(1,1)$	$=-a(1,13)$	$=-a(13,1)$	$=a(13,13)$	0.003955356
$a(1,2)$	$=-a(1,12)$	$=-a(13,2)$	$=a(13,12)$	0.009709308
$a(1,3)$	$=-a(1,11)$	$=-a(13,3)$	$=a(13,11)$	0.005183379
$a(1,4)$	$=-a(1,10)$	$=-a(13,4)$	$=a(13,10)$	−0.000942506
$a(1,5)$	$=-a(1,9)$	$=-a(13,5)$	$=a(13,9)$	−0.008563646
$a(1,6)$	$=-a(1,8)$	$=-a(13,6)$	$=a(13,8)$	0.000000000
$a(2,0)$	$=-a(2,14)$	$=-a(12,0)$	$=a(12,14)$	0.004163103
$a(2,1)$	$=-a(2,13)$	$=-a(12,1)$	$=a(12,13)$	0.010792100
$a(2,2)$	$=-a(2,12)$	$=-a(12,2)$	$=a(12,12)$	0.001407317
$a(2,3)$	$=-a(2,11)$	$=-a(12,3)$	$=a(12,11)$	−0.018146610
$a(2,4)$	$=-a(2,10)$	$=-a(12,4)$	$=a(12,10)$	−0.024165060
$a(2,5)$	$=-a(2,9)$	$=-a(12,5)$	$=a(12,9)$	−0.021612745
$a(2,6)$	$=-a(2,10)$	$=-a(12,6)$	$=a(12,8)$	0.000000000
$a(3,0)$	$=-a(3,14)$	$=-a(11,0)$	$=a(11,14)$	0.008559292
$a(3,1)$	$=-a(3,13)$	$=-a(11,1)$	$=a(11,13)$	0.008143449
$a(3,2)$	$=-a(3,12)$	$=-a(11,2)$	$=a(11,12)$	−0.015611443
$a(3,3)$	$=-a(3,11)$	$=-a(11,3)$	$=a(11,11)$	−0.026525839
$a(3,4)$	$=-a(3,10)$	$=-a(11,4)$	$=a(11,10)$	0.004240427
$a(3,5)$	$=-a(3,9)$	$=-a(11,5)$	$=a(11,9)$	0.029064393
$a(3,6)$	$=-a(3,8)$	$=-a(11,6)$	$=a(11,8)$	0.00000000
$a(4,0)$	$=-a(4,14)$	$=-a(10,0)$	$=a(10,14)$	0.008378050
$a(4,1)$	$=-a(4,13)$	$=-a(10,1)$	$=a(10,13)$	0.000196434
$a(4,2)$	$=-a(4,12)$	$=-a(10,2)$	$=a(10,12)$	−0.023916892
$a(4,3)$	$=-a(4,11)$	$=-a(10,3)$	$=a(10,11)$	0.001845555
$a(4,4)$	$=-a(4,10)$	$=-a(10,4)$	$=a(10,10)$	0.096989214
$a(4,5)$	$=-a(4,9)$	$=-a(10,5)$	$=a(10,9)$	0.155381941
$a(4,6)$	$=-a(4,8)$	$=-a(10,6)$	$=a(10,8)$	0.000000000
$a(5,0)$	$=-a(5,14)$	$=-a(9,0)$	$=a(9,14)$	0.007460971
$a(5,1)$	$=-a(5,13)$	$=-a(9,1)$	$=a(9,13)$	−0.003693997
$a(5,2)$	$=-a(5,12)$	$=-a(9,2)$	$=a(9,12)$	−0.024292654
$a(5,3)$	$=-a(5,11)$	$=-a(9,3)$	$=a(9,11)$	0.023628792
$a(5,4)$	$=-a(5,10)$	$=-a(9,4)$	$=a(9,10)$	0.152403568
$a(5,5)$	$=-a(5,9)$	$=-a(9,5)$	$=a(9,9)$	0.227609386
$a(5,6)$	$=-a(4,8)$	$=-a(9,6)$	$=a(9,8)$	0.000000000
$a(6,0)$	$=-a(6,14)$	$=-a(8,0)$	$=a(8,14)$	0.000000000
$a(6,1)$	$=-a(6,13)$	$=-a(8,1)$	$=a(8,13)$	0.000000000
$a(6,2)$	$=-a(6,12)$	$=-a(8,2)$	$=a(8,12)$	0.000000000
$a(6,3)$	$=-a(6,11)$	$=-a(8,3)$	$=a(8,11)$	0.000000000
$a(6,4)$	$=-a(6,10)$	$=-a(8,4)$	$=a(8,10)$	0.000000000
$a(6,5)$	$=-a(6,9)$	$=-a(8,5)$	$=a(8,9)$	0.000000000
$a(6,6)$	$=-a(6,8)$	$=-a(8,6)$	$=a(8,8)$	0.000000000

differentiator is shown in Figs. 5 and 6, respectively. These simulations are done using MATLAB with Intel core i5 7th Generation, 2.70 Ghz, 8GB RAM. The purpose is to get a matrix of order $(2L_1 + 1) \times (2L_2 + 1)$ having quadrantally odd symmetric coefficients to design a linear phase FIR 2-D differentiator. The optimized coefficients of the matrix depend on the

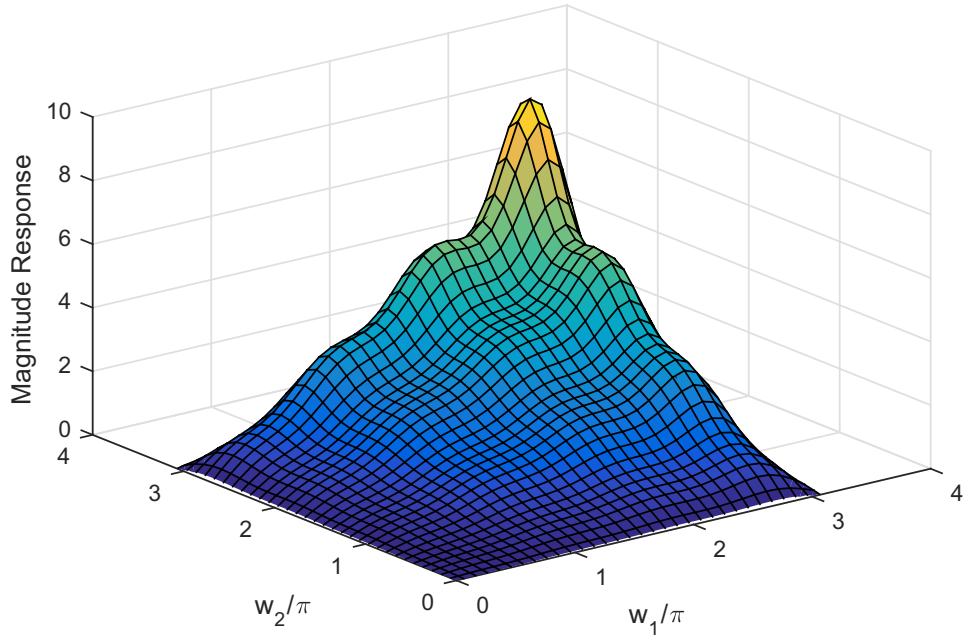


Fig. 5 Magnitude Response of 2-D FIR differentiator designed based on VSO

exploration and exploitation ability in the search space, as well as parameters selected for optimization technique. To compare the results obtained, three more optimization algorithms are used such as CSA, PSO and RCGA.

Due to linear phase characteristics of the FIR differentiator, it provides the constant delay to the input signal that is best suitable for the real time application.

Figure 7 represents the magnitude response of the digital 2-D differentiators obtained using CSA. CSA is based on the parasitic behaviour of cuckoo birds which lay their eggs in host bird nest. It uses the levy flight concept to enhance its searching ability with random walk to attain global optimum solution.

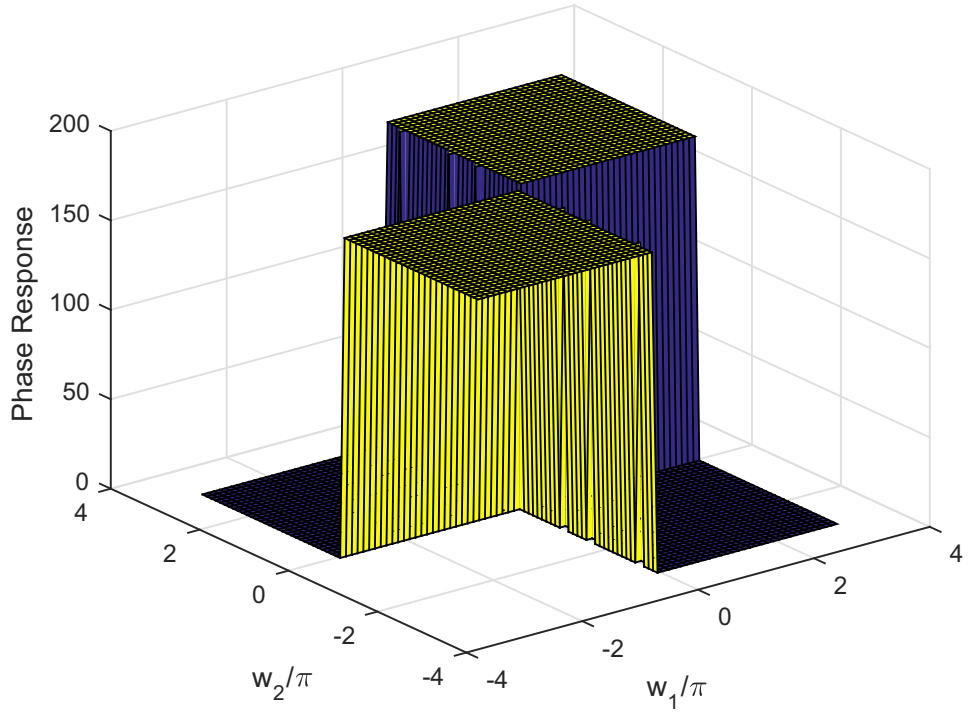


Fig. 6 Phase response of 2-D FIR differentiator designed based on VSO

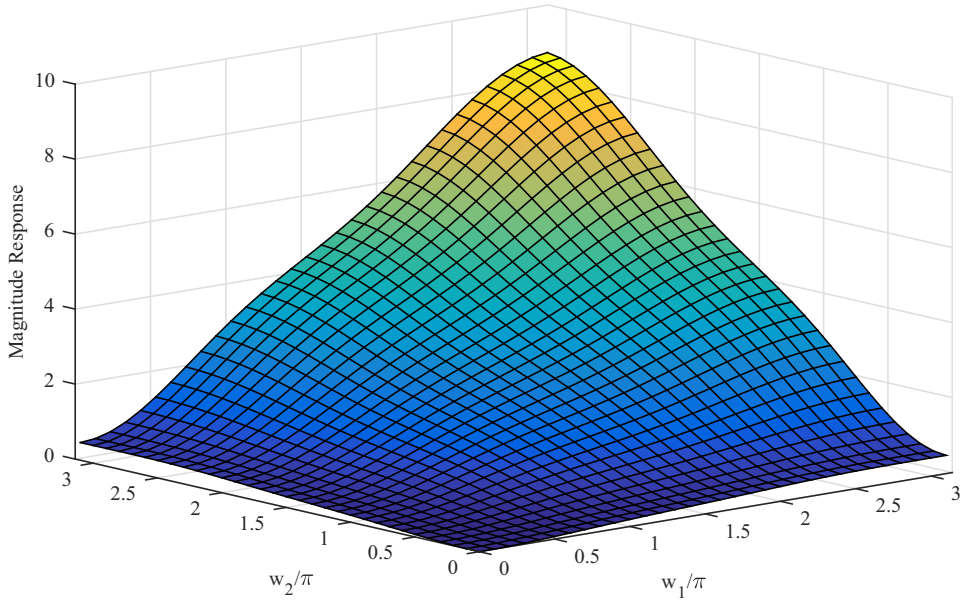


Fig. 7 Magnitude Response of 2-D FIR differentiator designed based on CSA [5]

Figure 8 shows the frequency response of 2-D digital FIR differentiator obtained using PSO. PSO basically works by moving candidate solutions (particles) in search space guided by simple mathematical formulas. Initially it has its own best position and velocity of the swarm. Particles are iteratively updated toward better positiontill they attain a global solution.

Magnitude response of 2-D differentiator shown in Fig. 9 is obtained using RCGA. RCGA implements selection, crossover and mutation operators in series framework. It follows survival to the fittest and achieve optimal solution. Analysis of the result for all the algorithms obtained from their magnitude response is further divided in to two parts:

1. Comparison with the existing methods
2. Execution Time

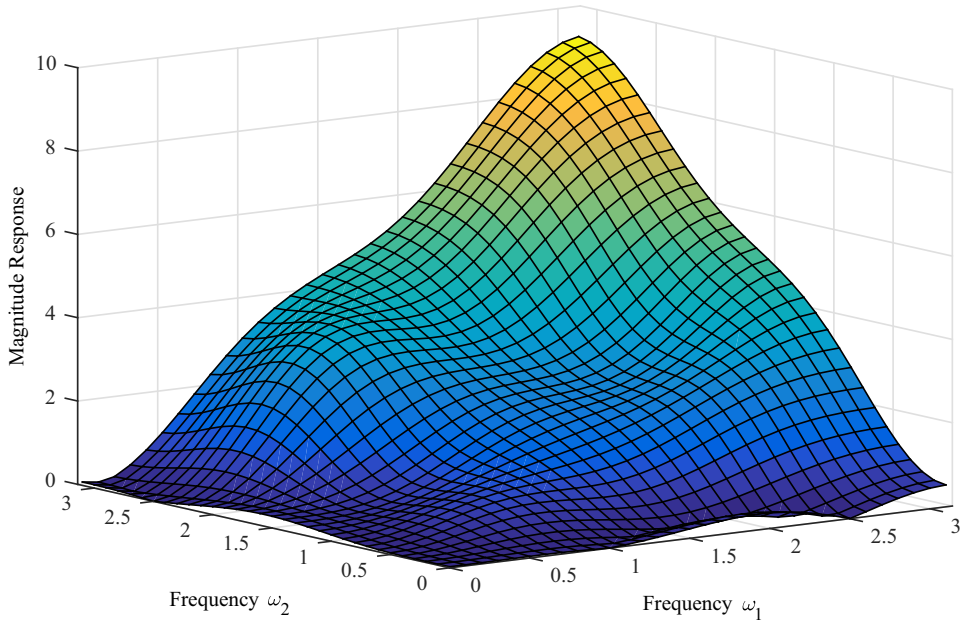


Fig. 8 Magnitude Response of 2-D FIR differentiator designed based on PSO [5]

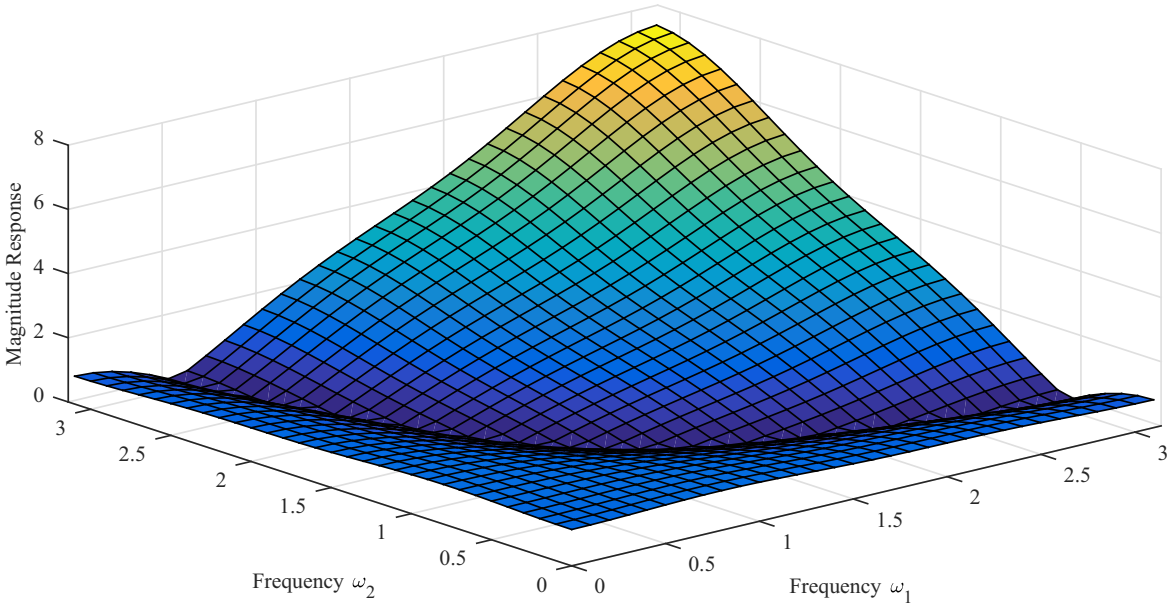


Fig. 9 Magnitude Response of 2-D FIR differentiator designed based on RCGA [5]

It is clear from the Table 3. that VSO based designed 2-D digital FIR 2-D differentiator has got least magnitude error, that is 0.1456 which approximates the ideal magnitude response of ideal 2-D digital differentiator in an efficient manner as compare to CSA, PSO and RCGA algorithms. It can be seen that CSA has magnitude error 0.1678, PSO got 0.9561 and RCGA has 1.3169. VSO is superior to other algorithms in terms of mean of the absolute error for all iterations. It has got least value which is a desirable characteristic for designing. VSO converges to a lower value of magnitude error as compare to other algorithms. Performance depends on the accuracy of the selection of initial parameters for all techniques.

The value of mean absolute error and standard deviation are finalized after taking 100 simulation trials. It can be clearly concluded that VSO has got least mean absolute error 1.3612, whereas CSA has been reported with 1.7238, PSO based 2-D differentiator has 2.2148 and RCGA got 2.7625. The inference is that employed algorithm yields the best result.

From Table 4, it can be depicted that time taken to execute the code or to attain the optimized coefficients for designing is minimum for VSO. This proves its fast convergence speed with excellent exploration and exploitation searching ability. Time taken to execute the code for VSO based differentiator is 19.85, whereas for CSA it is 24.44, PSO 28.54 and RCGA 35.76. The lesser time taken by VSO to run the code also proves that it requires less number of calculations to be performed. Hence it is computationally efficient. So, it can be concluded that VSO is preferable choice to design 2-D digital differentiator.

Table 3 Statistical analysis of 2-D digital differentiator design

Optimization technique	Error fitness value	Mean absolute error	Standard Deviation
VSO	0.1456	1.3612	0.0076
CSA [5]	0.1678	1.7238	0.0084
PSO [5]	0.9561	2.2148	0.0247
RCGA [5]	1.3169	2.7625	0.0549

Table 4 Execution time for different design techniques

Optimization technique	Execution time
VSO	19.85
CSA [5]	24.44
PSO [5]	28.54
RCGA [5]	35.76

3.1 Average error rate

It is defined as the sum of difference between optimal value at particular instance and the average of optimal values obtained for all iterations as in Eq. (12).

$$Error\ rate = \sum_{i=1}^{n} \frac{optimal\ fitness\ value(i) - average\ fittness\ value}{optimal\ fitness\ value(i)} \tag{12}$$

Here n is the total number of iteration. It is measured for VSO algorithm in designing 2-D differentiator, its value is 0.0040. Average error rate is very low for this algorithm. Hence it proves to be an efficient technique.

4 Conclusion

This paper proposed a FIR 2D digital differentiator design methodology, an approach based on vortex search optimization (VSO) algorithm. The proposed FIR 2D digital differentiator design method effectively employ the quadrantally odd symmetric properties of the differentiator response. L_1 error objective function is articulated to solve the FIR 2D digital differentiator design problem. The VSO based design is turn out as simple to implement to achieve the desired differentiator design specifications due to a smaller number of parameters tuning. Furthermore, the performance of the proposed FIR 2D digital differentiator is compared to cuckoo search algorithm (CSA), particle swarm optimization (PSO) and real coded genetic algorithm (RCGA) methodology. As a result, this approach design 2D digital differentiator with better quality of coefficients, exploration strategy and execution time than the other reported methods. In future the same algorithm can be used to design 2-D Hilbert transformer.

References

1. Aggarwal A, Kumar M, Rawat TK (2019) Design of 2-D FIR filters with Quadrantally symmetric properties using the 2-D L1-method. IET Signal Processing 13(3):262–272
2. Aggarwal A, Kumar M, Rawat TK (2019) Design of Digital Differentiator Using the L_1-method and swarm intelligence-based optimization algorithms. Arab J Sci Eng 44(3):1917–1931
3. Aggarwal A, Kumar M, Rawat TK, Upadhyay DK (2016) Optimal design of 2-D FIR filters with quadrantally symmetric properties using fractional derivative constraints. Circuits, Systems and Signal Processing, (Springer) 35(6):2213–2257

4. Aggarwal A, Kumar M, Rawat T K, Upadhyay D K (2016). Design of optimal 2-D FIR differentiators with quadrantally symmetric properties using the L1-method. International conference on signal processing and communication systems, ICSPCS 19-21
5. Aggarwal A, Kumar M, Rawat TK, Upadhyay DK (2017) Optimal design of 2-D FIR digital differentiator using L1-norm based cuckoo-search algorithm. Multidim Syst Sign Process 28(4):1569–1587
6. Ali TA, Xiao Z, Sun J, Mirjalili S, Havyarimana V, Jiang H (2019) Optimal design of IIR wideband digital differentiators and integrators using salp swarm algorithm. Knowl-Based Syst 182:104834
7. Chang WD (2009) Two-dimensional fractional-order digital differentiator design by using differential evolution algorithm. Digital Signal Processing 19(4):660–667
8. Dogan B (2016). A modified vortex search algorithm for numerical function optimization. International journal of Artificial intelligence and application (IJAIA) 7(3)
9. Doğan B, Ölmez T (2015) Vortex search algorithm for the analog active filter component selection problem. AEU-International Journal of Electronics and Communications 69(9):1243–1253
10. Fan GF, Peng LL, Hong WC, Sun F (2016) Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. Neurocomputing 173:958–970
11. Goswami O P, Rawat T K, Upadhyay D K (2019) A novel approach for the Design of Optimum IIR differentiators using fractional interpolation. Circuits, Systems, and Signal Processing 1–11
12. Gupta M, Yadav R (2014) New improved fractional order differentiator models based on optimized digital differentiators. Sci World J
13. Han J, From PID Active disturbance rejection control. IEEE Trans Ind Electron 56(3):900–906
14. Hong WC (2011) Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm. Energy 36(9):5568–5578
15. Hong WC, Dong Y, Zhang WY, Chen LY, Panigrahi BK (2013) Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm. Int J Electr Power Energy Syst 44(1):604–614
16. Huang Y, Hou G, Cheng X, Feng B, Gao L, Xiao M (2017) A new vortex search algorithm with gradient-based approximation for optimization of the fore part of KCS container ship. J Mar Sci Technol 22(3):403–413
17. Khan IR, Ohba R (1999) New design of full band differentiators based on Taylor series. IEE proceedings-Vision, image and Signal Processing 146(4):185–189
18. Kumar M, Aggarwal A, Rawat T K and Parthasarathy H (2017) Optimal nonlinear system identification using fractional delay second-order Volterra system IEEE/CAA Journal of Automatica Sinica 1–17
19. Kumar A, Komaragiri R, Kumar M (2018) Design of Wavelet Transform Based Electrocardiogram Monitoring System. ISA Trans 80:381–398
20. Kumar A, Komaragiri R, Kumar M (2018) Heart rate monitoring and therapeutic devices: a wavelet transform based approach for the modeling and characterization of congestive heart failure. ISA Trans 79:239–250
21. Kumar A, Komaragiri R, Kumar M (2019) Design of Efficient Fractional Operator for ECG signal detection in implantable cardiac pacemaker systems. International Journal of Circuit Theory and Applications 47(9): 459–1476
22. Kumar M, Rawat TK (2014) On the designing of fractional order FIR differentiator using radial basis function and window. WSEAS Transactions on Signal Processing 10:538–543
23. Kumar M, Rawat TK, Jain A, Singh AA, Mittal A (2015) Design of digital differentiators using interior search algorithm. Procedia Computer Science 57:368–376
24. Li MW, Geng J, Hong WC, Zhang LD (2019) Periodogram estimation based on LSSVR-CCPSO compensation for forecasting ship motion. Nonlinear Dynamics 97(4):2579–2594
25. Liao X, Li K, Zhu X, Liu K R (2020) Robust detection of image operator chain with two-stream convolutional neural network. IEEE Journal of Selected Topics in Signal Processing
26. Liao X, Yin J, Chen M, Qin Z (2020) Adaptive payload distribution in multiple images steganography based on image texture features. IEEE Transactions on Dependable and Secure Computing
27. Mahata S, Saha SK, Kar R, Mandal D (2017) Optimal design of wideband digital integrators and differentiators using harmony search algorithm. Int J Numer Modell Electron Networks Devices Fields 30(5):e2203
28. Mahata S, Saha SK, Kar R, Mandal D (2018) Optimal design of wideband digital integrators and differentiators using hybrid flower pollination algorithm. Soft Comput 22(11):3757–3783
29. Mcewan TE (1994) Ultra-wideband radar motion sensor. U.S. Patent 5:361–370
30. Mohan G K, Srinivasa R Y (2019) An efficient design of fractional order differentiator using hybrid shuffled frog leaping algorithm for handling noisy electrocardiograms. International journal of computers and applications 1–7
31. Montoya OD, Gil-González W, Grisales-Noreña LF (2019) Vortex search algorithm for optimal power flow analysis in DC resistive networks with CPLs. Express Briefs, IEEE Transactions on Circuits and Systems II

32. Muthukrishnan R, Radha M (2011) Edge detection techniques for image segmentation. International Journal of Computer Science & Information Technology 3(6):259–267
33. Rajeswari M, Amudhavel J, Dhavachelvan P (2017) Vortex search algorithm for solving set covering problem in wireless sensor network. Advances and application in Mathematical Sciences 17(1):95–117
34. Rawat T K (2010). Signals and systems. Oxford University Press
35. Singh S, Ashok A, Kumar M, Rawat TK (2019) Adaptive infinite impulse response system identification using teacher learner based optimization algorithm. Appl Intell 49(5):1785–1802
36. Tseng C C, Lee S L (2011). Design of variable fractional order differentiator using infinite product expansion. 20th European conference on circuit theory and design (ECCTD) IEEE 17-20
37. Tseng CC, Lee SL (2013) Designs of two-dimensional linear phase FIR filters using fractional derivative constraints. Signal Process 93(5):1141–1151
38. Upadhyay DK (2017) Design of second order recursive digital integrators with matching phase and magnitude response. Radio engineering 26(1):377
39. Wulf M, Staude G, Knopp A, Felderhoff T (2016) Efficient design of FIR filter based low-pass differentiators for biomedical signal processing. Current Directions in Biomedical Engineering 2(1):215–219
40. Yadav S, Yadav R, Kumar A, Kumar M (2020) Design of Optimal two-Dimensional FIR filters with Quadrantally symmetric properties using vortex search algorithm. Journal of Circuits, Systems, and Computers 29(10):2050155
41. Yadav S, Yadav R, Kumar A, Kumar M (2020) A novel approach for optimal design of digital FIR filter using grasshopper optimization algorithm. ISA Trans https://doi.org/10.1016/j.isatra.2020.08.032
42. Zhang Z, Ding S, Jia W (2019) A hybrid optimization algorithm based on cuckoo search and differential evolution for solving constrained engineering problems. Eng Appl Artif Intell 85:254–268
43. Zhang Z, Ding S, Sun Y (2020) A support vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task. Neurocomputing
44. Zhang Z, Hong WC (2019) Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. Nonlinear Dynamics 98(2):1107–1136

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Suman Yadav [1,2] · Richa Yadav [2] · Ashwni Kumar [2] · Manjeet Kumar [3]

[1] Department of Electronics and Communication Engineering, Bharati Vidyapeeth College of Engineering, New Delhi 110063, India

[2] Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for women (IGDTUW), New Delhi 110006, India

[3] Department of Electronics and Communication Engineering, Delhi Technological University (DTU), Rohini, Delhi 110042, India

Archives of Computational Methods in Engineering
https://doi.org/10.1007/s11831-020-09498-y

ORIGINAL PAPER

A Systematic Review on Firefly Algorithm: Past, Present, and Future

Vijay Kumar[1] · **Dinesh Kumar**[2]

Received: 18 February 2020 / Accepted: 19 September 2020
© CIMNE, Barcelona, Spain 2020

Abstract
Firefly Algorithm (FA) is one of the popular algorithm of Swarm Intelligence domain that can be used in most of the areas of optimization. FA and its variants are simple to implement and easily understood. These can be used to successfully solve the problems of different areas. Modification in original FA or hybrid FA algorithms are required to solve diverse range of engineering problems. In this paper, a comprehensive review of firefly algorithm is presented and various characteristics are discussed. The various variant of FA such as binary, multi-objective and hybrid with other meta-heuristics are discussed. The applications and performance evolution metric are presented. This paper provides the possible future research direction of FA.

1 Introduction

Soft computing have various techniques for computation (as shown in Fig. 1) such as Neural system, Fuzzy, evolutionary computing, chaos theory and many more. Swarm Intelligence (SI) is one of the very promising domain of AI (Artificial Intelligence) as shown in Fig. 2, become drastically important and popular in last few years [1]. The study of behaviour of ants, fireflies, bees, worms, termites, group of birds and fishes inspires the further research done in the field of SI [2], as shown in Figs. 3 and 4. The coordinated behaviour of these groups direct them to achieve their desired goal. The simple interaction between unsophisticated individuals leads group towards a self organizing behaviour for the whole group with their collective intelligence. The simple interaction in this multi agent system forms this self organizing and coordinated behaviour. Termites and worms interacts to other individuals for construction of their nests. Ants and bees forms a collective behaviour to find out their food [3]. To get shortest way to the place where food has found from their colony, ants use chemical pheromone trails while interact to each other. The informer scout bees have responsibility to find new origins of food and these bees use waggle dance to communicate or direct the other bees. During the search of new source of food, bees trade off the work of collection of new information (exploration) and the information (exploitation) [3]. The bee colony is pretty much aware about when to use the food and when to search the food, to maximize intake and minimize foraging effort.

These above mentioned swarm utilize their behaviour for making decisions such as: reproduction, foraging, finding new home, division of task among each other, etc. [3]. The beauty of this approach is, information collected by interaction of individuals in decentralized manner.

Xin-She Yang inspired from the behaviour of fireflies and patterns formed by them, and develop a novel meta-heuristic algorithm named as Firefly Algorithm (FA) [4]. It is stochastic, in nature. It is capable of solving NP-Hard optimization problems. Some randomization is followed in this scheme after searching the set to available solutions. Figure 5 shows basic framework of firefly algorithm. This approach finds the solution using trial and error methodology, which defines Heuristic nature. However, it does not guarantee the optimal solution in some application [4].

Exploration and exploitation are two most important balancing factors in FA [5].

Where, Exploration is the process of finding diverse range of solution with in search space. Exploitation is searching with in best available coverage, while targeting or using the discovered information.

✉ Vijay Kumar
vijaykumarchahar@gmail.com

Dinesh Kumar
dineshkumar@dtu.ac.in

[1] Computer Science and Engineering Department, National Institute of Technology, Hamirpur, Himachal Pradesh 177005, India

[2] Electronics and Communication Engineering Department, Delhi Technological University, Delhi 110042, India

Published online: 29 September 2020

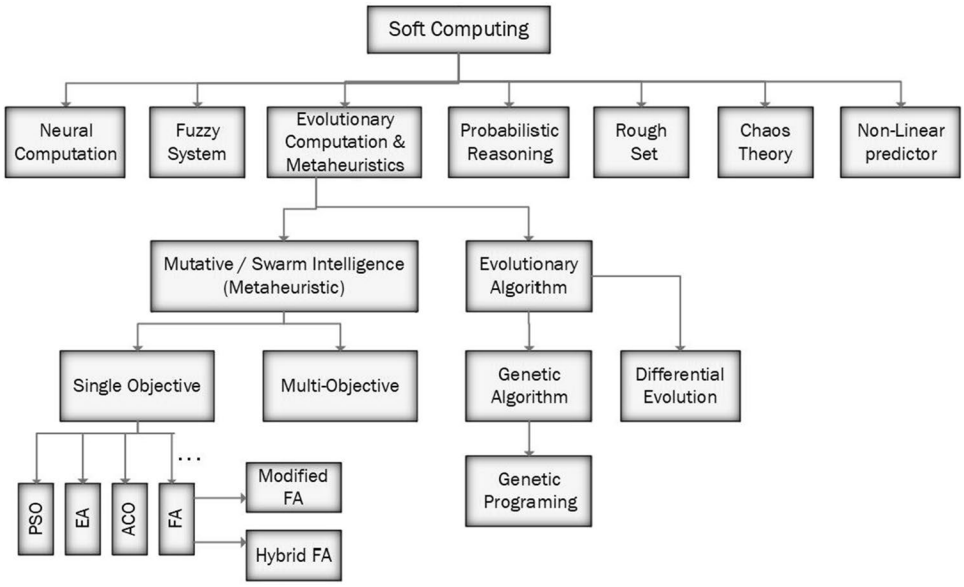


Fig. 1 Evolution of Firefly algorithm

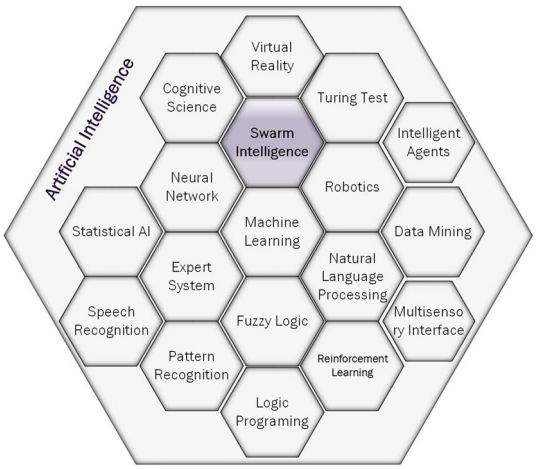


Fig. 2 Domains of artificial intelligence

Firefly algorithm is a population based algorithm. FA is advantageous in comparison of single point search algorithms such as

- Crossover process put together building block of different solutions.
- Cross over influence the focused search. I means that off-springs have same value if the parents are similar in value of variables.
- Distraction in a landscape is minimized or ignored by a low pass filter.
- For initializing position or making decision, it hedges to bad luck.
- The algorithm is opportunistic to tune good values of parameter to imbalance exploration and exploitation.

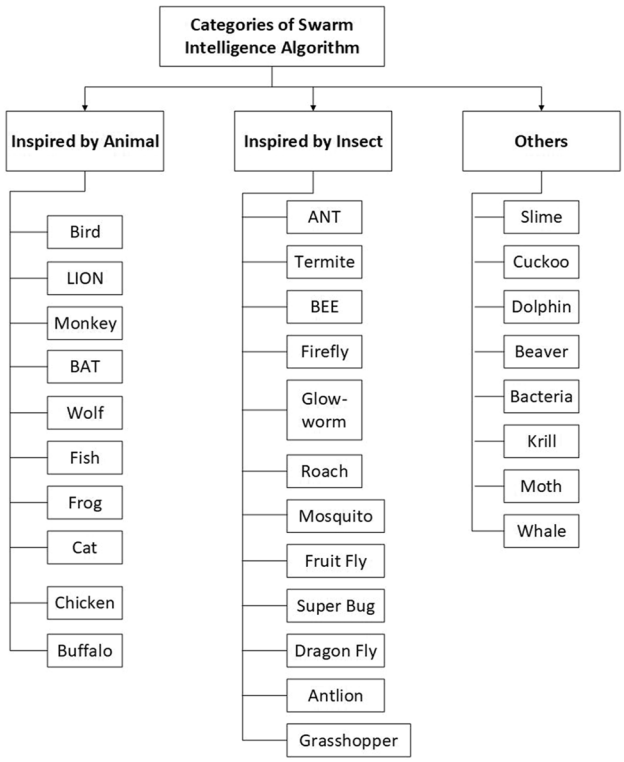


Fig. 3 Different swarm intelligence algorithm inspiration

1.1 Characteristics of Fireflies

Fireflies has a flash light as a major characteristic. It can be used to attract possible mates and to warn from potential

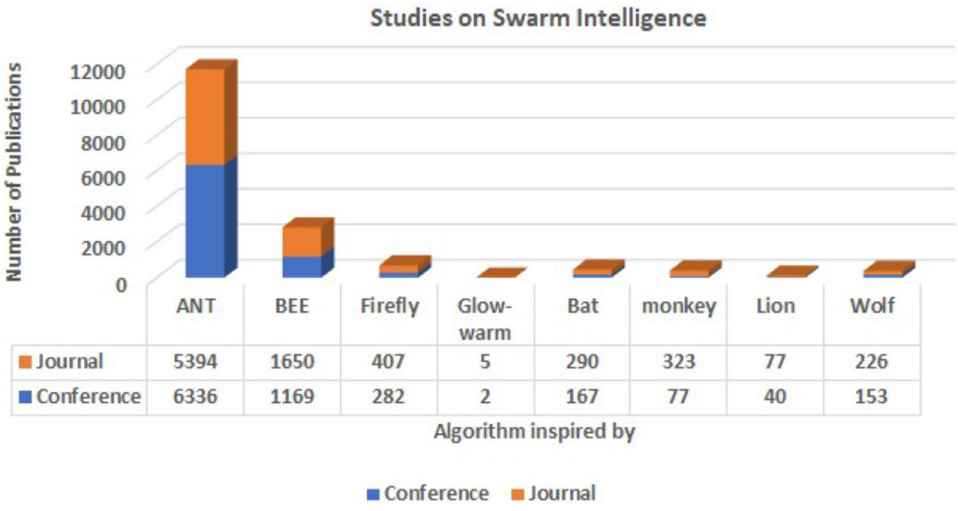


Fig. 4 Publications of different swarm intelligence approaches [2]

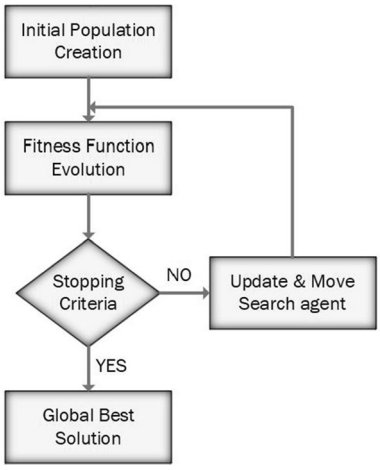


Fig. 5 General framework of Firefly algorithm

threats or enemies. This working of flash light follow the physics's rule.

- With increase in distance, flash light intensity decrease according to law of physics

$$I\alpha 1/r^2 \tag{1}$$

where r is distance and I is intensity of light.

- Secondly, the firefly blink their flash light at regular time instances using $\theta = 2\pi$, which is again a law of physics [6].
- Whenever a firefly comes in neighbourhood, they come together and form a couple. This behaviour inspired from graph coloring problem [7].

The remaining structure of the paper is organized as follows. Section 2 consist the inspiration, working flow, algorithm with complexity, controlling parameters, characteristics and application of classical firefly algorithm. However, Classification and analysis of FA is present in Sect. 3. Furthermore, various significant application of firefly algorithm, categorically described in Sect. 4. and final conclusion of the paper is elaborated in Sect. 5.

2 Classical Firefly Algorithm

In 2008, Yang et al. [4] got inspiration from fireflies about their light emitting behaviour. They studied the pattern and intensity of emitting light. On the basis of that study, they developed the firefly algorithm . They found out some natural meta-heuristics in that algorithm. To determine the performance, they designed a 2D function having 4 peaks and the results shows a multi- characteristics of algorithms. This algorithm can simultaneously find more optimal results in the search space.

A newly developed firefly algorithm specially for multi-optimization problems introduced by yang et al. [8]. They found out that fireflies can subdivide them in to subgroups due to stronger neighbour attraction over long distance attractiveness. For experimental prove, the test functions of many multi- problems compared with genetic algorithms and swarm intelligence claiming global optimum result, and observed that FA provide better success rate and efficiency. Based on these fats some researchers said that FA is a powerful algorithm to solve even some of NP-hard Problems.

Considering a continuous optimization problem, Yang et al. [9] again use classical FA and shows experimented results to solve non-linear pressure vessel optimized designing . Firstly they used some specific test function to validate

algorithm. the results shows pretty better results than PSO(particle swarm optimization).

Yang also reviewed some natural meta-heuristic algorithms like ACO, Cuckkoo search, genetic algorithm, differential evolution, harmony search, simulated annealing, PSO,and firefly algorithm [10]. Based on the common properties, a new algorithm 'GEWA(Generalized Evolutionary Walk Algorithm)' proposed by yang. This algorithm was a generic meta-heuristic kind of optimization algorithm.

The GEWA has three major components:

- Global Exploration using randomization
- Use of random walk for intensive local search
- Use of elitism for 'Selection of Best'

This algorithm use problem dependent parameter to imbalance the exploration and exploitation.

A survey on Swarm intelligence inspirations, shows the inspiring factors for researchers to develop new nature inspired swarm intelligence algorithm. Parpinelli et al. [11] found out that these inspiring factors are: Bacterial foraging, life cycle of slime moulds, Bio-luminescence of fireflies, Cockroach infestation, Bats echolocation, Bee's algorithms based on Bee's foraging, Bee's mating, Host seeking of mosquitoes. They also determined other applications of using meta-heuristic features.

Zang et al. [12] did a systematic, analyzed reviewing of ACO, algorithms based on Bee's characteristics, firefly algorithm, genetic algorithm . Discussion on applications and principle behind original algorithms has done mainly. Figure 6 shows the major steps involved in firefly algorithm [4].

2.1 Inspiration Behind Firefly

Fireflies get active in summer night as they prefer warm environment.

Bio-luminescence is a biochemical process that causing the flashing of light in fireflies. It is used for signal of courtship for the mate and also used for warning signal if any potential predators is nearby. Some of the firefly spices use pheromone like ants. The organ that produces light using bio-luminescent is known as lanterns. Most of the adult firefly can control bio-luminescence process to emit low or high flash light. This flashing is initialized by the signal originating by central nerve system of firefly.

Firefly start signaling to attract other firefly in response of this signal the other one flash. The mating couple of firefly generate a unique pattern of flash, that is a very precisely timed on the basis of information like sex or identity. The attraction depends upon the behavioural difference of signalling. Mostly the brighter flashing mate attracts the other one.

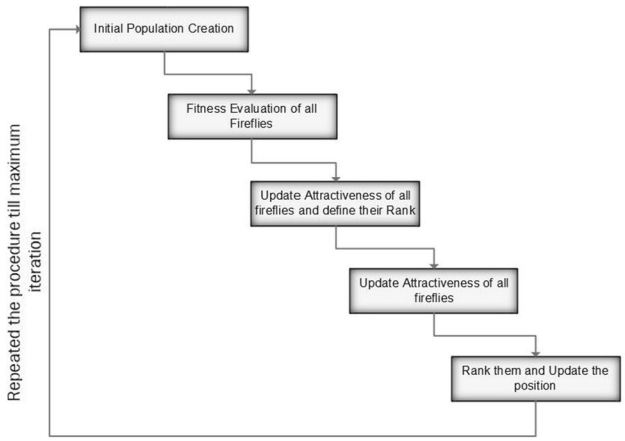


Fig. 6 Working of FA

As the flashing intensity gets weaker with distance. Some of the firefly spices does not discriminate between dull and bright flash signal.

The flash signal is pretty much visible so it can be used to discourage a variety of potential predator. On the basis of survival of fittest, the flashing is evolved and used as mechanism for defence. Therefore it can be used for warning to predator.

Decentralized decision making and self organizing behaviour are two characteristics that are used in firefly algorithm. To live in a common place interactive communication is a necessity. Individuals living in the group should adapt the goal of group. They can not behave autonomously. This social life is dedicated to reproduction of fireflies and foraging.

The collective decision of group is closely related to the flashing pattern of fireflies which is most important biological foundation for developing firefly algorithm.

Three ideal rules are followed by FA [6].

- Fireflies attract each other regardless their preference of Sex, because fireflies are unisex.
- This attraction depend on flash light brightness. Attraction start decreasing when flash brightness decreases, when the between distance increases. So, the brighter one attracts the less bright firefly. If no bright flash firefly is near then they move randomly.
- The objective function's landscape defines the brightness of flash.

The proportionality of attractiveness depends on the intensity of light seen by other firefly. So the attractiveness variation 'β' on 'r' distance expressed as [6]

$$\beta = \beta_0 e^{-\gamma r^2} \tag{2}$$

Here 'r' equals to 0 and β_0 denotes the attractiveness.

Let us assume that there are two firefly x_i and x_j. Movement of *i*th firefly towards *j*th firefly due to more brightness is defined as

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}(x_j^t - x_i^t) + \alpha_t \epsilon_i^t \tag{3}$$

Where,

- Second part of Eq. 3 explains the attraction.
- 3rd part of Eq. 3 is for randomization with α_t.
- ϵ_i^t defines the vector of Random numbers generated by either Gaussian distribution or it can be introduced by uniform distribution with time 't'.

Simple random walk is accompanied, when β_0 is zero. FA become a reduced mutant version of PSO (Particle Swarm optimisation) when γ is zero [4]. Levy flights and others, randomization of ϵ_i^t is easily extended some other distributions, etc. [4].

'Figrue 7' shows the working of firefly algorithm. Let there are N initial solutions and i denotes the number of iteration, with 'max_itr' maximum iteration number. 'I' is the intensity or brightness of flash.

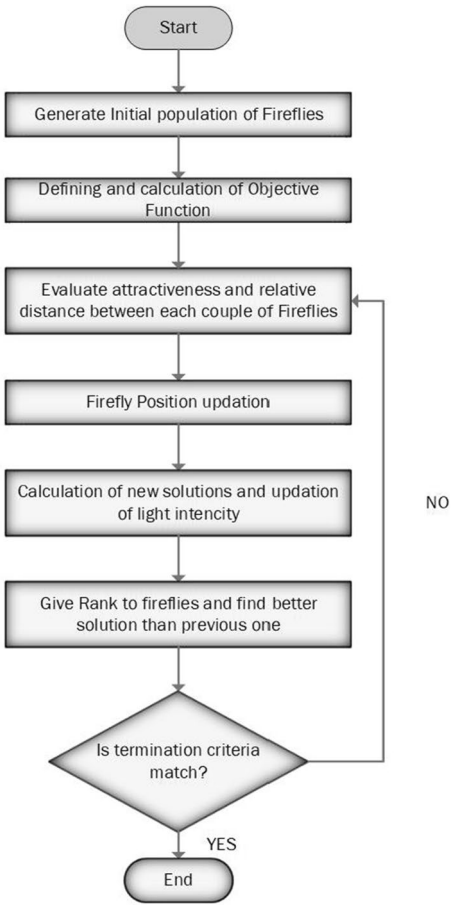


Fig. 7 Flowchart of FA

Algorithm 1: Steps of FA

Result: Best Solution Available

Let there are N initial solutions and i denotes the number of iteration, with 'max-itr' maximum iteration number. 'I' is the intensity or brightness of flash;

1. Generate initial solution 'N' randomly.;
2. **for** $i \leftarrow 1$ **to** *max-itr* **do**
   - Calculate Brightness I;
   - sort I such as $I_j >= I_{j-1}$, for each j;
   - **for** $j \leftarrow 1$ **to** $N - 1$ **do**
     - **for** $k \leftarrow j + 1$ **to** N **do**
       - **if** $I_k > I_j$ **then**
         - Firefly 'j' move in the direction of Firefly 'k';
       - **end**
     - **end**
   - **end**
   - Move firefly 'N' in random manner;

   end
3. Store the best solution

2.2 FA Characteristics

Firefly algorithm have been suffered from two important issue.

- Flash intensity variance
- Evaluation of attractiveness

The researchers must address these issues to solve the problem. In the algorithm, the flash intensity 'I' determines a solution 'S', in proportion to the fitness value 'F' such as I(S)αF(S). But we know that this intensity of light changes with distance according to following relation:

$$I(r) = I_0 e^{-\gamma r^2} \tag{4}$$

where, I_0 is Source Light intensity. γ is Fixed light absorption coefficient. r is Distance.

At r=0, i/r^2 avoid the singularity due to combination of inverse square law and Gaussian absorption. Equation 2 defines the attraction due to emitted lighted. The intensity is the absolute measure of light while attractiveness relatively measure to emitted light,judged by other fireflies.

Euclidean distance is computed between two fireflies(s_i, s_j), expressed as:

$$r_{ij} = ||s_i - sj|| = \sqrt{\sum_{k=1}^{k=2}(s_{ik} - s_{jk})^2} \tag{5}$$

where dimensions of the problem are represented by 'n'. The attraction of *i*th firefly towards more brighter *j*th firefly is given as:

$$s_i = s_i + \beta_0 e^{-\gamma r_{ij}^2}(s_i - s_j) + \alpha\epsilon_i \tag{6}$$

Where, ϵ denotes the randomness extracted from Gaussian distribution [6].

Firefly movement made up of three main component such as:

- *i*th firefly's current position.
- attraction toward more bright light emitting firefly.
- and Random walk, where α is randomize parameter and random generate number rang from zero to one.

Movement of fireflies depends only on random walk, when β_0=0. γ is a main factor of convergence speed. Accepting range of γ is [0 to ∞] but depending on the optimization problem it may varies from 0.1 to 10 [6].

Three factors,

- α as randomization parameter
- β as Attractiveness
- γ as absorption coefficient

control Firefly algorithm

Two asymptotic behaviours of firefly algorithm has been observed as per parametric setting of γ, when γ leads from 0 and 1.

At $\gamma \to 0$, β become β_0(A special variant of Particle Swarm Optimization with constant Attractiveness at any place of search space)

At $\gamma \to \infty$, The second term of Equation (7) become zero and the movement of firefly follows random walk (and become parallel simulated annealing).

Any implementing of FA is become either of any these two asymptotic behaviour.

3 Classification of Firefly Algorithm

Various versions of FA have been proposed by different researchers. Multi-Objective optimization version of firefly algorithm is also available. Chaos can be incorporated in FA for performance improvement. Hybrid variants of firefly also are developed for performance enhancement. The general classification of firefly algorithm can be done in two most effective manner such as parameter tuning and parameter control.

Parameter tuning is finding out the good values of used parameters before executing algorithm and fixing them during course of iterations [13]. Parameter control is altering the value of parameter during the execution process [14].

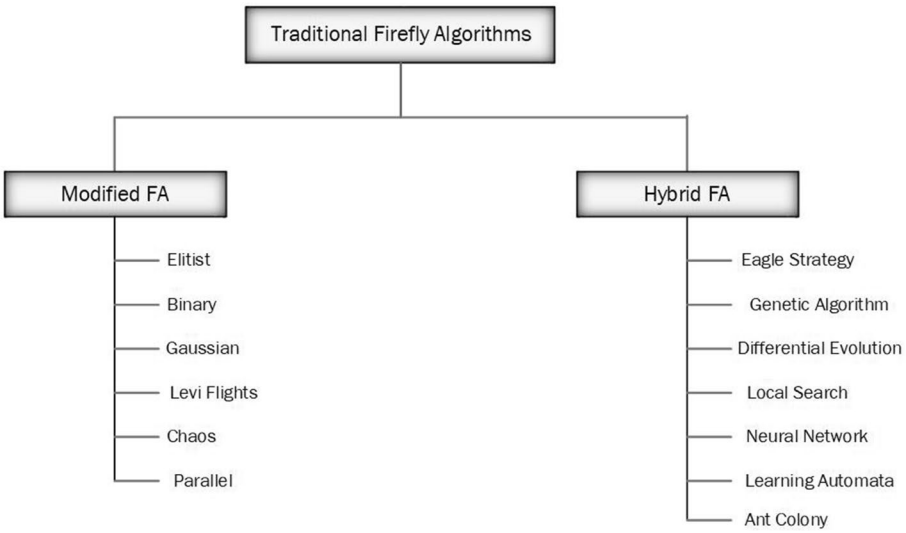
The behaviour of firefly algorithm depends not only to the proper parameter value but also to the features of the components involved in the algorithm. Therefore, the classification aspects of FA should considered the following points:

- What is altered or Modified?
- How these changes are done?
- What is the scope of alteration?

The first aspect is based on the following factors:

- Firefly representation(Either Binary of real Value)
- Scheme of Population (Swarm or Multi-Swarm)
- Fitness function Growth
- Best Solution calculation (Elitism or non-Elitism)
- Movement of Fireflies can be Uniform, Levy Flights, Gaussian, chaos distribution)

Fig. 8 Classification of FA



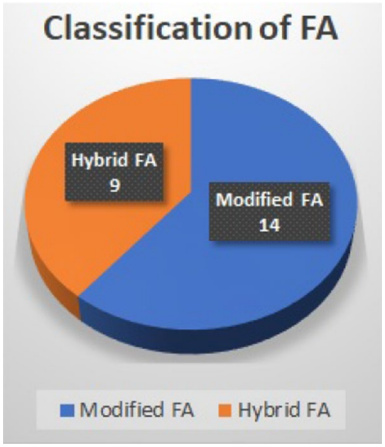


Fig. 9 Classification of Firefly algorithm

The Second classification aspect is parametric control of FA such as Deterministic, Adaptive, Semi-Adaptive. The third aspect affects FA as:

- Element of firefly
- Firefly as whole
- and, the population of Fireflies

In the start, FA seems like global optimization problem. It is not able to find optimized solution for some of the optimization problems. To get rid off the above mentioned problem, FA has been hybridized with other meta-heuristics technique. Machine learning, heuristic nature, hybridization and many other techniques were incorporated in firefly algorithm to solve many optimization problem. The main components of FA like initializing process, evaluation of fitness function and movement calculation were modified to propose many hybrid FA.

Figure 8 shows the classification of FA.

Continuous optimization problems use the classical firefly algorithm. Inclusion of hybridization and several types of changes done in classical FA that make this algorithm possible to provide optimal results. After such modification, FA is influenced to develop Elitist, Binary, Gaussian, Levi Flight, Chaos, parallel algorithms. Figure 9 represent the graph of such categorical studies of firefly algorithm.

3.1 Modified FA

The two significant factors for improvement of FA are : Light variation and Attraction. These two factors mainly lead researchers to develop the other algorithms.

Surafel et al. [15] changed the random movement of brightest firefly and find out that brightness of light gets decreased if the current best position stop improving further. They tried to enhance the position of brightest firefly which follows m-uniform random walk and move in best position direction.

A binary FA is proposed by Palit et al [16]. for cryptanalysis deciphering using Merkle-hellman knapsack cipher algorithm . Almost all the components needs to modify for new binary firefly algorithm. Another researcher Falcon et al. [17] did binary encoding on candidate solution with adaptive light absorption coefficient to improve search and problem specific facts to deal with infeasible solutions.

when System level fault detection combinatorial optimization problem is empirically analyzed, the results outperforms AIS (artificial immune system) and PSO algorithms on the basis of memory requirements and convergence speed [18, 19]. K. chandrasekaran [20] developed a binary version of FA to solve reliability constrained unit commitment problem.

Farahani [21] proposed a new variant of FA. He used Gaussian distribution to increase convergence speed for getting global optimum move of all fireflies in every iteration. He used five different standard functions to test the algorithm but the concluded experimental results are very promising and accurate than classical FA.

Yang [22], overviewed inefficiency and convergence study of Meta-heuristics and provide a framework to analyze these. He also analyzed the impact of Gaussian random walk, levy flight on meta-heuristic results. He concluded on the basis of results that to deal newly developed meta-heuristics, global and local exploitation imbalance trade-off must be considered.

Inspired by Levy flight strategy for movement, a new meta-heuristic firefly algorithm was developed [23] that represent better success rate and efficiency in comparison of GA and PSO. Efficiency and convergence associated Meta-heuristic such as cuckoo search, FA, Swarm intelligence, levy flight and random walk are considered for analysis in this paper [24]. In this analysis author tries to formulate some mathematical for meta-heuristic behaviour. He found out that mathematical modeling was still a mystery even though the new natural meta-heuristic algorithms work well.

Dos et al. [25] used chaotic maps with a combination of Firefly algorithm to get better convergence of classic FA. Chaos sequence helped to escape local optima effectively. Light absorption coefficient and randomize parameter of Eq. (7) were tuned using chaos sequence. The power of chaotic map FA is defined by a benchmark of reliability-redundancy optimization. Another researcher used chaos in FA to enhance global search mobility [26]. The tuning of absorption coefficient and attractiveness was done using chaotic maps.

Subutic et al. [27] proposed parallelized firefly algorithm for unconstrained optimization problem and tested it on standard benchmark function. A pretty fast and less time consuming algorithm has been obtained. But it is limited only to the scenario when more than one population generation has been considered.

Husselmann et al. [28], firefly algorithm is used on GPU (parallel graphical processing unit) and compared with classical FA on standard benchmark function basis. The proposed algorithm provide fast and much accurate result than original FA. It is observed that original FA is suitable for uni-model optimization problem because less number of fireflies are required which reduced the computation time drastically.

3.2 Hybrid FA

Yang [29] introduce the first hybrid FA, which use eagle strategy as a meta-heuristic of search method. This eagle strategy combines levy flight in FA, which shows the foraging behaviour of eagle. It is similar to levy flight because like this eagle also randomly fly in their vicinity. Eagle catch the pray most efficiently.

Eagle Search algorithmic characteristic, the eagle strategy has two components such as

- Intensive Local search
- Random search by levy flight

Luthra et al. [30] used hybrid FA for mono alphabetic substitution cipher crypt-analysis. They use crossover and mutation operation inspiring from genetic algorithm. They used dominant gene crossover operator and permutation mutation.

HEFA (Hybrid Evolutionary Firefly Algorithm) combined the classic FA with differential evolution to improve the accuracy of search and information sharing between fireflies [31]. after calculating the fitness value this algorithm divide the total population in two sub groups. the first group of fireflies is applied with classical FA while second group adopt the differential evolution. the estimation of biological is done in this proposal. the result of HEKA are more accurate and performing speed was excellent as compared to GA, PSO classical FA and Evolutionary programming.

A well known optimization problem, 'Graph three-coloring' use hybrid classical FA with local search heuristic [32]. when a graph of 500 vertices generated by culberson graph generator, is applied with this proposed MFA (memetic FA), HEA (Hybrid evolutionary algorithm), tabucol,EA-SAW algorithms for graph coloring. the FA outperforms rest of algorithms. it state that FA could successfully solve various combinatorial optimization problems.

In a paper, published by Hassanzadeh et al. [33], A proposal of using FA for speech recognition by training SEFNN (structure equivalent fuzzy neural network) parameter. the generalization ability of fuzzy neural network is improved by firefly algorithm. the result shows that Speech recognition done by hybrid algorithm has higher recognition rate than classical fuzzy neural network trained by PSO.

Nandy et al. [34] use Meta-heuristic FA with back propagation to train FFNN (feed forward neural network). back propagation involved in FA to get faster and better convergence rate while training a feed forward neural network. The developed algorithm take few iterations to converge local optima, while it is tested on some standard datasets. Less convergence time and better convergence rate is provided by proposed algorithm as compared to GA applied on same problem.

Another use of hybrid FA is for cellular learning automata. the meta heuristic 'Cellular leaning automata' is used to find diverse solution of firefly algorithm. local search is improved by FA. Results were tested on five benchmark function and better exploration rate of standard FA and global optima has been found out [35].

Table 1 Performance and implementation study of various variants of Firefly algorithm

References	Implemented on	Performance
[15]	Simulated on uni and multi model, continuous and discontinuous, stochastic benchmark problems	Outperform standard FA
[16]	Merkle–Hellman algorithm	Perform better than GA
[17]	Programmed in java, processor:Intel Core2 Duo E4500 2.2GHz, 2GB RAM, OS:Windows Vista, validated on t-diagnosable systems	Improved convergence, less memory consumption, better than AIS and PSO(worst)
[18]	Programmed in C++, processor: Pentium IV 2.4GHz, validated on t-diagnosable systems	Correct identification of faulty processors, take small amount of time to diagnose faulty situation with few generations
[19]	JDK1.6, processor:Intel Core2 Duo E4500 2.2GHz, 2GB RAM, validated on t-diagnosable systems	Outperform AIS
[20]	MATLAB 7.01, processor: Pentium IV 3.2 Ghz, 1GB RAM, validated on 10, 26 unit system and IEEE RTS24 bus system	Outperformed GA and standard PSO
[21]	Validated by simulation using uni-model function (Rosenbrock and Sphere), Multi-model(Rastrigin, Griewank, Ackley)	Better than Standard FA and PSO
[23]	MATLAB, Validated on Ackley function	Performed better than GA and PSO
[25]	MATLAB, processor: Pentium IV 3.2 Ghz, OS: Windows XP, validated on benchmark reliability-redundancy problem of optimization	Improved performance than standard FA
[26]	12 chaotic maps for parameter tuning, used sphere function, Griewank's function	Better than standard FA
[27]	Used java on Lukasik and Zak software of Matlab, processor: Intel I7- 2600k, RAM: 8GB, OS: Windows7 x64, IDE: NetBeans 7.0.1, tested on unconstrained benchmark functions (Rastrigin, Griewank, Ackley, Sphere)	Improved performance than origional FA
[28]	processor: Intel Core i7 3.4Ghz, GPU: 2 NVidia GTX 590, validated on Rosenbrock function (unimodal), Schwefel function (multi-model)	Better than Ackley's Path function
[29]	Simulated on MATLAB, Validated on Ackley function	Improved success rate and better efficiency than PSO and GA
[30]	Programmed in C++, processor: Intel Core 2 Duo 2.0GHz, RAM: 3GB	Works better for large cipher input, for smaller cipher large number of generations are required
[31]	Used general kinetic model of maillard reaction	More accurate and faster than PSO, GA, EP
[32]	Used DSatur traditional heuristic, tested on test suite	Outperformed SAW-EA, Tabucol, and HEA
[33]	Used SEFNN	Better than FNN with PSO,
[34]	Programmed in python, used matplotlib tool, CPU: Pentium-IV core2duo 1.66Ghz, RAM: 512MB , validated on iris, wine and liver data set	Outperformed GABPNN (back propagation neural network using genetic algorithm)
[35]	Tested on 5 standard function (sphere, ackly, rastrigin, step, and XS yang)	outperform standard FA
[36]	Validated on sphere, ackly, rosenbrock, rastrigin, and griewank functions	competitive with standard FA and PSO.
[37]	MATLAB used Wisconsin Prognostic breast cancer (WPBC) data set, tested using mean absolute percentage error (MAPE)	Better efficiency and successrate than EPSO

A proposal of categorizing classical FA into three classes based on improved performance is published by farahani et al. [36]. Adaptive randomization parameters and absorption coefficient was achieved using learning autometa in first class. while in second class, balanced exploitation and exploration properties of meta heuristic was achieved using firefly algorithm with hybrid genetic algorithm. Gaussian distribution inspired random walk was used for movement of fireflies over search space done in thirst class. result displayed that a very competitive algorithm is proposed while comparing to PSO and classical FA, after testing it on five standard benchmark functions.

ACO was used to predict cancer by developing a FNT(flexible neural tree) for micro-array data [37]. FA tunes neural tree parameter. model achieved optimal solution with faster convergence and low error rate. the comparative study of FA and EPSO(extensive PSO) found superior success rate and efficiency with FA than EPSO.

Table 1 list performance and implementation details of different variants of firefly algorithm.

3.3 Reason of Efficiency

As observed from above literature review, almost all of the proposal shows better and improved results. one must can think about why Firefly algorithm is efficient in mostly cases. after analysing classical or standard FA few main characteristics has been found out for its success and efficiency.such as:

- Firefly algorithm is aware about local attraction is superior than attraction at long distance. This is why FA divide its population into different smaller subgroups. which makes efficient and it suitable for multi-model, non-linear problems of optimization.
- Like PSO, Firefly algorithm does not have drawback of premature convergence because it does not make decisions using explicit global best and historic individual best. Ignoring velocity does not degrade performance the FA.
- FA control scaling parameter to make it self adaptable to the problem landscape. FA can control it's modality.

Firefly algorithm is a generalised form of SA (simulated annealing), PSO, Differential evolution. The changes in Eq. 6 describes these different versions.

- When β_0 =0 it a type of parallel simulated annealing where α control the annealing schedule.
- Similarly, at γ =0 and β_0 = Unif(0,1). It become a type of DE (differential evolution) without mutation where β_0 controls crossover rate.
- It become a PSO variant at γ =0 and current global best is used in place of s_j.

This means, firefly algorithm have advantages of SA, PSO, DE which supports the efficient performance of FA.

4 Controlling Parameters

As we know that α_t controlling the randomness and the solution diversity (till some extent). So α_t can be used to change 't' used for counting the iteration. 'α_t' expressed as,

$$\alpha_t = \alpha_0 \delta^t, Zero < \delta < 1. \tag{7}$$

Where,

- 'α_0' is a scale factor of Initial randomness
- 'δ' used for Necessary cooling factor.
  mostly 'δ' is in between 0.95 to 0.97 [4].

Firefly algorithm gives promising and efficient results when 'α_0' is used with scaling of designing values of variables.

α_0 can be initialized as 0.01L where L express problem average scale. Factor of 0.01 is used because random walk need few steps for target with balanced local exploitation, and without covering large distance with few steps [8, 38].

- Attractiveness is controlled by β. β equals to one is suggested for most application.
- Parameter 'γ' is related to problem scaling L. Generally, $\gamma = 1/\sqrt{L}$. γ is O(1) if for insufficient variation is there.
- Best population range is 25 to 40 but range of 15 to 100 population size is used mostly [4, 8].

Complexity of firefly algorithm is explained in following steps:

1. FA has two loops of population 'n' inside a loop of iteration 't'. which deduce its time complexity as O(n^2t).
2. for Relatively larger population(n), one inner loop of sorted attractiveness rank or brightness of fireflies is recommended.if this is the case then time complexity of FA become O(ntlog(n)).
3. Computation cost is small with small population and large iteration and gives some sort of linear time complexity. objective function evaluation increases the computational cost for external black box objectives.this consideration has effect on most of the meta-heuristic algorithm that is why objective function of all optimization problem is computationally extensive.

Similar to other meta-heuristic algorithm, the parameters of Eq. 3 are user dependent and affect the performance of algorithm by controlling exploitation and exploration. This parametric adjustment and adaptation are the reason behind some modified Firefly Algorithms by changing α, γ,and 'r'. Modification in α affect random movement, and degree of attraction is modified by changing γ of 'r'. Brightness adjustment is done by modifying β_0.

4.1 Random Movement Parameter Modification

Shafaati et al. [39] modified FA for IIR (infinite impulse response) and non-linear system parameter identification. The modified random movement was shown in Eq. 8, where α_0 was initial step length and α_∞ was final step length.

$$\alpha = \alpha_\infty + (\alpha_0 - \alpha_\infty)e^{-Itr} \tag{8}$$

An additional step i.e. $\alpha\epsilon(x_i - x_b)$ was added for updating process, where x_b represents the most brightest firefly. This added term provide resemblance and behave like PSO. The user provide initial and final step length. Randomized

parameter exponentially decrease and became null within few iteration. The added term took x_i far away from x_b with $\alpha\epsilon$ step length, while in current best solution both are same. They used two parameters (i.e. α_0 and α_∞) in place of α.

FA with adaptive randomized movement parameter(α) was presented [40]. It was modified according to Eq. 9.

$$\alpha^{(Itr)} : \alpha^{(Itr-1)}(1/(2Itr_{Max}))^{1/(Itr_{Max})} \tag{9}$$

Based on randomly selected 3 solutions, other 2 new solutions were generated. Solution with improved brightness supplied to the next iteration after updating x_i. This modification was also used in optimal capacitor placement problem [41], and in [42] with crossover and mutation.

Yu et al. [43] evaluated randomized parameter using number of iteration. $0.4/(1+e^{0.005(Itr-Itr_{Max})})$ was used for this calculation. Historical information along with current situation was used to update step length. Equation 10 used for calculating α, Where Eq. 11 was for $f_b^{Itr} = f(x_b)$, $f_i^{Itr} = f(x_i)$ after Itr iterations. $f_{i_{best}}^{(Itr-1)}$ and $f_{i_{best}}^{(Itr-2)}$ gives best performing solutions till Itr-1 and Itr-2 iterations. Different simulated tests shows the improved performance of this modified FA. Controlling random parameter based on previous and current performance of solution benefits in a way such that whenever solution reach near to the brightest firefly, it's step length decreased.

$$\alpha_i^{(Itr+1)} = 1 - \frac{1}{\sqrt{f_b^{(Itr)} - \left(f_i^{(Itr)}\right)^2 + \left(f_i^{(Itr)}\right)^2 + 1}} \tag{10}$$

$$h_i^{Itr} = \frac{1}{\sqrt{\left(f_{i_{best}}^{(Itr-1)} - f_{i_{best}}^{(Itr-2)}\right)^2 + 1}} \tag{11}$$

Adaptive random movement parameter was used for multi-objective optimization [44]. It was modified as $\alpha = \alpha_0 0.9^{Itr}$, to decrease step length faster than linear.

Solution's historic performance was used to modify random movement parameter [45]. Best position and global best solution till current iteration were evaluated using below equation.

$$\alpha_i^{(Itr+1)} = \alpha_i^{(Itr)} - \left(\alpha_i^{(Itr)} - \alpha_{\min}\right) e^{-\frac{Itr|x_{gbest} - x_{ibest}|}{MaxGen}} \tag{12}$$

4.2 Attraction Parameter Modification

Lin et al. [46] introduce concept of virtual distance to overcome the drastic change in attraction while there was a small decrease in distance between two fireflies. Virtual distance 'r' (between 0 to 1) was defined as

$$r' = \frac{r - r_{\min}}{r_{\max} - r_{\min}} \tag{13}$$

where $r_{\min} = 0$ and $r_{\max} = \sqrt{\sum_{i=1}^{d} \left(x_{\max}(i) - x_{\min}(i)\right)^2}$ and $\beta = \beta_0\gamma\left(1 - r'\right)$

Then swarm start converge to optimal solution. When virtual distance 'r', r_{max} also decrease with it. However, faster decrease rate of 'r' than r_{max} slightly increase β. To handle dominating attraction, they proposed an updating Eq. 14 for later iteration. This updation contradict random attraction and firefly only move towards brighter firefly with $\beta\alpha\epsilon$ step length.

$$x_i = x_i + \beta\left(x_j - x_i\right)\alpha\varepsilon \tag{14}$$

Tilahan et al. [15] proposed a light intensity function as $\beta_0 = e^{I_{0,j} - I_{0,i}}$ for attraction. They suggested, Firefly's random movement might degrade the brightness, so selecting a direction for improved brightness from 'm' random direction could be beneficial. But introducing new parameter 'm' affect complexity.

Gandomi et al. [26] used chaos function to modify β, γ. They suggested that chaos's non-repetition and ergodicity could carry searches at higher rate. Later, this proposal is used by many researcher. This approach was used for MRI brain tissue segmentation using chebyshev chaos mapping [47], Gaussian's mapping used for heart disease prediction [48], reliability redundancy optimization [25], and solving definite integral problems [49]

Khalil et al. [50] used chaotic mapping for updating β, γ. Based on solution's intensity, they decreased α using below equation.

$$\alpha = \alpha_{\max} - \left(\alpha_{\max} - \alpha_{\min}\right)\frac{I_{\max} - I_{mean}}{I_{\max} - I_{\min}} \tag{15}$$

β was calculated using $\beta = \left(\beta_{\max} - \beta_{\min}\right)e^{-\gamma r^2} + \beta_{\min}$ [51, 52], where β_{min} and β_{max} were user driven parameters.

4.3 Random Movement Along with Attraction Parameter Modification

Yan et al. [53] modified FA to deal with increased problem dimension and feasible region. They modified generalised brightness i.e. $\beta = \beta_0 e^{-\gamma r^\omega}$ by changing r^ω to $r^K\sqrt{n(Range)}$, where k= constant, n= problem dimension and, Range= maximum dimension Range. α was decremented linearly from α_0 to α_{end}. The fireflies were attracted by brightness and winking. Probability $p_w = 0.5 + 0.1\text{count}_i$ decided the winking, where $count_i$ was winking counter of ith firefly. Large counter value (range between 0 to 5) involve greater probability of state shift.

Sulaiman et al. [54] presented a modified FA to address economic dispatch problem. They used minimum variation

distance in-place of Cartesian distance, to increase exploration. They also mutate α.

Wang et al. [55] modified FA's light intensity to tackle premature convergence. For iteration 't', they calculated light intensity difference using $\xi = \frac{\Delta I_{ij}^{(t)}}{\max\{I\}-\min\{I\}}$ formula. Then they modified α, β, γ based on ξ, as $\alpha = \alpha_0(0.02 r_{\max})$ where $\alpha_0 = \begin{cases} \xi, & \xi > \eta_2 \\ \eta_2, & \xi \le \eta_2 \end{cases}$ with new parameter η_2, $\beta_0 = \begin{cases} \xi, & \xi > \eta_1 \\ \eta_1, & \xi \le \eta_1 \end{cases}$ with another new parameter η_1, and $\gamma = \frac{\gamma_0}{r_{\max}^2}$ where $r_{\max} = \max\{d(x_i, x_j) | \forall i, j\}$. This amendment showed that, brighter firefly had small attraction, and less brighter had randomness in step length.

Othman et al. [56] proposed modification in firefly algorithm to optimized size and siting of voltage-controlled distribution generator in power distribution network. They selected optimal location and produced power to minimized power loss for distributed generation. They modified α, γ in each iteration accordance to location and maximum power per location, and considered $\beta = 1$.

Subramanian et al. [57] modified α as $\alpha = \alpha_{\max} - \frac{Itr(\alpha_{\max}-\alpha_{\min})}{MaxGen}$. Dependence of distance function had been shifted from location in feasible region to brightness or functional value calculated using $f(x_b) - f(x_i)$. They considered that, if two fireflies had similar performance, they must be nearby.

α, γ parameters of FA were modified, and used in path planning of autonomous underwater vehicles [58]. α was calculated using $\alpha = \alpha_b + \frac{Itr}{MaxGen}(\alpha_e - \alpha_b)$ where $\alpha_e < \alpha_b$. γ was modified as $\gamma = \gamma_b + \frac{Itr}{MaxGen}(\gamma_e - \gamma_b)$ where $\gamma_e > \gamma_b$. Furthermore, Updation was done using below equation.

$$x_i = x_i + \beta(x_i - x_j) + \alpha r \varepsilon \tag{16}$$

It showed that, α decrease and γ increase, linearly, with increase in iterations. The random movement and attraction decrements as a function of iteration. The Random movement was multiplied with distance, as it was shown in updating Eq. 16.

Fister et al. [59] also modified α, β, γ parameter and proposed a self adaptive firefly algorithm. They used updation equation as $, \psi := \psi + \sigma_\psi N(0,1)$ for $\sigma_\psi := \sigma_\psi e^{\tau' N(0,1) + \tau N(0,1)}$, where $\tau, \tau,$ were learning parameters and $\psi = \{\alpha, \beta, \gamma\}$.

Fu et al. [60] modified γ using $\gamma_{\max} - (\gamma_{\max} - \gamma_{\min})\left(\frac{Itr}{MaxGen}\right)^2$ formula, where limits of γ_{max} was between 2 to 4, and for γ_{min} was between 0.5 to 1. They added a new parameter λ for calculation of α using equation

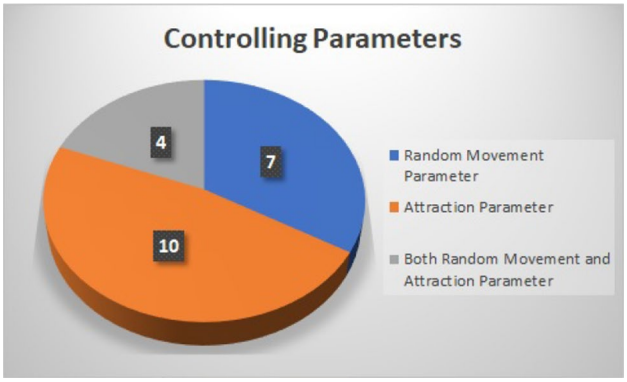


Fig. 10 Studies on controlling parameters

$$\alpha = \alpha_{\max} - (\alpha_{\max} - \alpha_{\min})\left(\frac{Itr - 1}{G_0 - 1}\right)^{\lambda} \tag{17}$$

Where $\alpha = \alpha_{\min}$ at G_0 iteration. As resulted, α reduced faster than linear when λ was in between 0 and 1, reduced linearly at $\lambda = 1$, and it reduced slower than linear at $\lambda > 1$. They used Gaussian distribution to prevent trapping in local optima, that move toward global best solution i.e. $x_b = x_b + x_b N(\mu, \sigma)$. This was used when solution variance was less than η(precision parameter) before M predetermined iterations. They suggested to use cubic mapping for better distribution of initial solution.

Coelho et al. [61] used Gaussian or normal distribution (G_1, G_2) for μ=0, and σ^2=1, and computed α as $\alpha = 0.03|G_2|$, γ as $\gamma = 0.03|G_1|$. They also used it for multi-variable PID controller tuning [62]. Tinkerbell mapping was used for tuning of γ using $\gamma = |G|\bar{x}\frac{itr}{MaxGen}$, where $\bar{x}$ were normalized values ranged between 0 and 1. α was decremented linearly using Eq. 18.

$$\alpha = (\alpha_{final} - \alpha_{initial})\frac{Itr}{MaxGen} + \alpha_{initial} \tag{18}$$

The comparative studies of work done is presented as a graph in Fig. 10 (Table 2)

5 Applications of FA

Many type of classification, Optimization and engineering optimization problems were addressed by firefly algorithm and its various version, as shown graphically in Fig. 11. Applications of FA in different problem domains shows in Fig. 12.

Table 2 Modified variants of Firefly algorithm

References	Modified parameters	Modification done
[39]	Random movement	$\alpha = \alpha_\infty + (\alpha_0 - \alpha_\infty)e^{-Itr}$, added $\alpha\epsilon(x_i - x_b)$
[40–42]	Random Movement	$\alpha^{(Itr)} : \alpha^{(Itr-1)}(1/(2Itr_{Max}))^{1/(Itr_{Max})}$
[43]	Random movement	$\alpha_i^{(Itr+1)} = 1 - \frac{1}{\sqrt{f_b^{(Itr)} - \left(f_i^{(Itr)}\right)^2 + \left(f_i^{(Itr)}\right)^2 + 1}}$ $h_i^{Itr} = \frac{1}{\sqrt{\left(f_{i_{best}}^{(Itr-1)} - f_{i_{best}}^{(Itr-2)}\right)^2 + 1}}$ Where for $f_b^{Itr} = f(x_b)$, $f_i^{Itr} = f\left(x_i\right)$ after Itr iterations
[44]	Random movement	$\alpha = \alpha_0 0.9^{Itr}$
[45]	Random movement	$\alpha_i^{(Itr+1)} = \alpha_i^{(Itr)} - \left(\alpha_i^{(Itr)} - \alpha_{\min}\right)e^{-\frac{Itr\lvert x_{gbest} - x_{ibest}\rvert}{MaxGen}}$
[46]	Attraction	Virtual distance $r' = \frac{r - r_{\min}}{r_{\max} - r_{\min}}$, Where $r_{\min} = 0$ and $r_{\max} = \sqrt{\sum_{i=1}^{d}\left(x_{\max}(i) - x_{\min}(i)\right)^2}$ and $\beta = \beta_0\gamma\left(1 - r'\right)$ and updating equation $x_i = x_i + \beta\left(x_j - x_i\right)\alpha\epsilon$
[15]	Attraction	$\beta_0 = e^{I_{0,j} - I_{0,i}}$
[26]	Attraction	Chaos function to modify β, γ
[47]	Attraction	Chebyshev chaos mapping
[25, 48, 49]	Attraction	Gaussian's mapping
[50]	Attraction	Chaotic mapping, $\alpha = \alpha_{\max} - \left(\alpha_{\max} - \alpha_{\min}\right)\frac{I_{\max} - I_{mean}}{I_{\max} - I_{\min}}$
[51, 52]	Attraction	$\beta = \left(\beta_{\max} - \beta_{\min}\right)e^{-\gamma r^2} + \beta_{\min}$
[53]	Both attraction and random movement	$\beta = \beta_0 e^{-\gamma r^\omega}$ by changing r^ω to $r^K\sqrt{n(Range)}$, and Winking probability $p_w = 0.5 + 0.1\text{count}_i$
[54]	Both attraction and random movement	Minimum variation distance in-place of Cartesian distance
[55]	Both attraction and random movement	$\xi = \frac{\Delta I_{ij}^{(t)}}{\max\{I\} - \min\{I\}}$, $\alpha = \alpha_0\left(0.02r_{\max}\right)$, where $\alpha_0 = \begin{cases} \xi, & \xi > \eta_2 \\ \eta_2, & \xi \le \eta_2 \end{cases}$ with new parameter η_2, $\beta_0 = \begin{cases} \xi, & \xi > \eta_1 \\ \eta_1, & \xi \le \eta_1 \end{cases}$ with another new parameter η_1, and $\gamma = \frac{\gamma_0}{r_{\max}^2}$ where $r_{\max} = \max\left\{d\left(x_i, x_j\right)\lvert\forall i,j\right\}$
[56]	Both attraction and random movement	α, γ in every iteration according to location and maximum power per location, and $\beta = 1$
[57]	Both attraction and random movement	$\alpha = \alpha_{\max} - \frac{Itr(\alpha_{\max} - \alpha_{\min})}{MaxGen}$, Brightness as $f(x_b) - f(x_i)$
[58]	Both attraction and random movement	$\alpha = \alpha_b + \frac{Itr}{MaxGen}\left(\alpha_e - \alpha_b\right)$ where $\alpha_e < \alpha_b$, $\gamma = \gamma_b + \frac{Itr}{MaxGen}\left(\gamma_e - \gamma_b\right)$ where $\gamma_e > \gamma_b$ Updating Equation as $x_i = x_i + \beta(x_i - x_j) + \alpha r\epsilon$
[59]	Both attraction and random movement	Updating equation as, $\psi = \psi + \sigma_\psi N(0,1)$ for $\sigma_\psi := \sigma_\psi e^{\tau' N(0,1) + \tau N(0,1)}$, where $\tau, \tau,$ were learning parameters and $\psi = \{\alpha, \beta, \gamma\}$
[60]	Both attraction and random movement	$\gamma = \gamma_{\max} - \left(\gamma_{\max} - \gamma_{\min}\right)\left(\frac{Itr}{MaxGen}\right)^2$, Where γ_{max} = (2 to 4) and γ_{min} = (0.5 to 1), added $\alpha = \alpha_{\max} - \left(\alpha_{\max} - \alpha_{\min}\right)\left(\frac{Itr-1}{G_0 - 1}\right)^\lambda$, where $\alpha = \alpha_{\min}$ at G_0. Global best solution $x_b = x_b + x_b N(\mu, \sigma)$
[61]	Both attraction and random movement	Gaussian or normal distribution (G_1, G_2) for μ=0, and σ^2=1, Calculated $\alpha = \alpha = 0.03\lvert G_2\rvert$, and $\gamma = \gamma = 0.03\lvert G_1\rvert$
[62]	Both attraction and random movement	Tinkerbell mapping for tuning $\gamma = \lvert G\rvert\bar{x}\frac{itr}{MaxGen}$. Calculated $\alpha = (\alpha_{final} - \alpha_{initial})\frac{Itr}{MaxGen} + \alpha_{initial}$

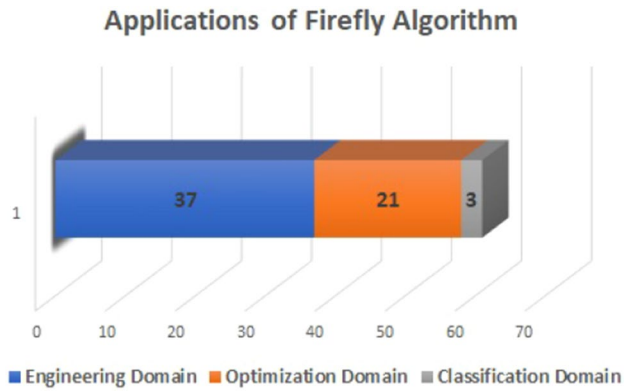


Fig. 11 Application domains of FA

5.1 Optimization Application of Firefly Algorithm

Figure 13 shows graphical representation of different optimization applications of firefly algorithm.

Lukasik et al. [63] was the first author who used FA in constrain optimization . They applied proposed method on 14 benchmark constrained problem and compare the results with PSO. Observed results of FA found is superior than PSO.

FA was used for mixed (both continuous and discrete) Structural optimization problems [64] . They considered their work for pressure vessel designing, stepped cantilever beam design, designing of welded beam, helical compression spring design, car side impact designing and designing of reinforced concrete beam. The results of FA is far better than results collected from PSO, DE and SA. but oscillatory behaviour of FA is observed for searching optimum design. The behaviour is improved by reduction of randomization parameter with optimization progress.

FA is used for shape and size based structural mass optimization problem by Gomes [65]. They took dynamic constrained for consideration. The proposed work shows a extreme non linear behaviour with regards of frequency constraint for shape optimization. Due to eigen values sensitivity for shape modification, they used constraints as penalty functions which affect the fitness.

5.1.1 Scheduling Problems

Quadratic assignment problem was solved by using FA [66]. In this permutation of integer shows the solution. An adaptive discrete function was mapped using attractiveness, movement, and distance.

Sayadi et al. [67] developed a discrete firefly meta-heuristic that minimize makespan of permutation flow shop scheduling problem.

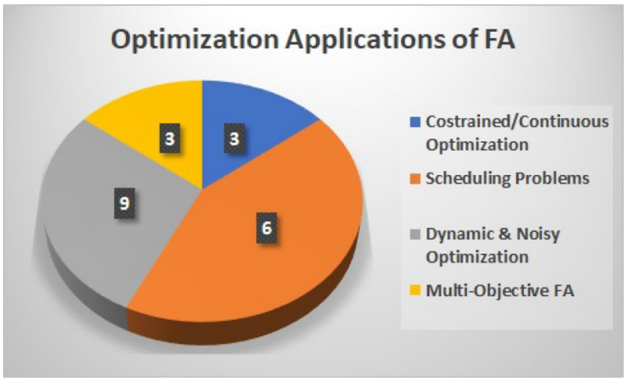


Fig. 13 Different optimization application of FA

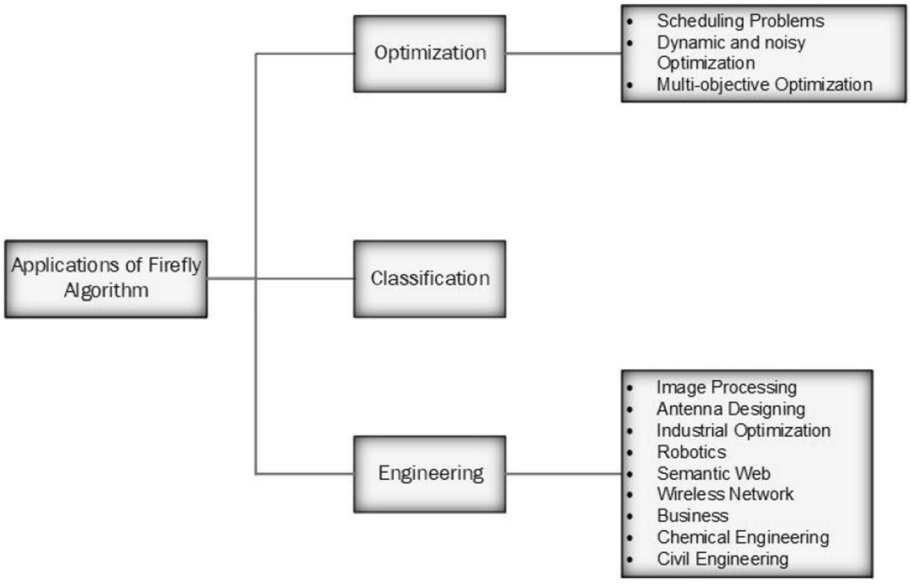


Fig. 12 Various applications of FA

A job scheduling mechanism for cloud computing using FA was proposed by Yousif et al. [68]. They used FA for minimizing the execution time take by different jobs. parameters such as length of job, resource speed, and other identifier considered for this problem.

Honig et al [69] used firefly algorithm 'Task Graph Scheduling'. The proposed algorithm was tested on Task graph having 36,000 task to be schedule. The algorithm takes less computational time than the other meta-heuristics.

Symmetric TSP (Travelling Salesmen Problem) was addressed by FA [70]. The proposed EDFA (Evolutionary discrete firefly algorithm) use permutation representation where city is represented by element array and index denotes tour order. Each inversion mutation of firefly use 'm' generated moves. The results were simulated on some standard TSP library problem where proposed EDFA out perform memetic algorithm.

JSSP (Job Shop scheduling problem) using FA is proposed by Khadwilard et al. [71]. Some JSSP examples were taken from OR library and five popular benchmark functions were used to test the algorithm for finding least makespan. Algorithm mostly got trapped in local optimum but it can provide best solution in some cases.

Auther Liu et al. [72] developed adaptive FA for path planning. Adaptive absorption and random parameter of classical FA provide better solutions and higher convergence rate.

Path planning problem of UCAV (uninhabited combat air vehicle), a modified FA is proposed by Wang [73]. He modified the process of information sharing between top fireflies while updating their light intensity. It increases the global convergence rate while maintaining robustness.

Kwiecien et al. [74] proposed optimize queuing system using firefly algorithm. Queuing theory helps to analyze complex services of communication, transportation, computer system, and manufacturing.

5.1.2 Dynamic and Noisy Optimization

Abshouri et al. [75] used learning automata to tune randomization, absorption coefficient, and attractiveness firefly parameters. In this proposal, splitting of fireflies into different set of information exchanging swarm was done. This sets quickly react to environment changing. Different variants of multi- dynamically moving peaks benchmark functions were used to test its performance.

Chai et al. [76] used FA and Bees meta-heuristics algorithms, to solve various noisy continuous mathematical functions and determine their optimal solutions using two variables.

Farahani [77] proposed a multi-swarm variant of FA for dynamic environment. Provision of splitting firefly population into different groups of communicating fireflies was proposed. Exclusion parameter was used for local interaction and for global interaction.

Stabilizing efforts of firefly movement was also done by Farahani et al. [21]. They presented a mechanishm to direct the movement of firefly toward global best. They considered this behaviour when no nearby solutions were better. Gaussian distribution were used to increase the convergence rate.

Nasiri et al. [78] developed a dynamic environment suitable firefly algorithm. They used

- Preserving of Best Solution
- Adaptive Randomization Parameter.

In power distributing network , Optimal size and location of distributed generation obtained by a variant of FA [79]. The proposed algorithm showed effective working over many different bus distribution test system.

5.1.3 Multi-Objective Firefly Algorithm

For multi-objective optimizing problem, Yang et al. [80] proposed multi objective FA that directly produce Pareto optimal front. A group of multi objective discontinuous Pareto front, convex and non convex function was used to test the proposal by optimizing industrial engineering benchmark problems.

Abedinia et al. [81] modified FA to solve multi objective Environment/ Economic power dispatching problem (EED). This problem has multiple objectives for conflicting scenarios of fuel cost, system loss, and emission. IEEE Standard 30 and 118 bus cases were used to test the developed algorithm.

Another case of optimizing multi objective problem was tackled by Niknam [82]. They obtained Pareto optimal solution (non dominated) using multi-objective FA. Absorption coefficient and randomization parameter were tuned using chaotic maps. This tuning also use self adaptive probabilistic mutation to improve performance.

Table 3 list down different optimization applications.

5.2 Applications in Engineering Domain

Many engineering problems from different engineering domains can be solved by firefly algorithm. Mostly it is applicable for finding solutions of various engineering optimization. Antenna designing, image processing, civil engineering, and business optimization. Graph in Fig. 14 shows different proposals of FA in engineering domain.

5.2.1 Image Processing

Zhang et al. [83] modified firefly algorithm to optimize the image registration. One rotational parameter with two

Table 3 Optimization variants of Firefly algorithm

References	Used for	Category
[63]	Continuous constrained optimization	Constrained optimization
[64]	Mixed variable structural optimization	Constrained optimization
[65]	Structural size and shape optimisation	Constrained optimization
[66]	QAP problem of sage framework	Scheduling problems
[67]	Makespan minimization	Scheduling problems
[68]	Job scheduling	Scheduling problems
[69]	Scheduling task graphs	Scheduling problems
[70]	Travelling salesman problem	Scheduling problems
[71]	Job shop scheduling	Scheduling problems
[72]	Path planning	Scheduling problems
[73]	UCAV path planning	Scheduling problems
[74]	Optimization of queuing systems	Scheduling problems
[75]	Learning automata in dynamic environments	Dynamic optimization
[76]	Noisy non-linear optimisation problems	Dynamic and noisy optimization
[77]	Dynamic environments	Dynamic Optimization
[21]	Gaussian FA	Dynamic and noisy optimization
[78]	Optimization in dynamic environments	Dynamic and noisy optimization
[79]	Optimal allocation and sizing distribution system	Dynamic and noisy optimization
[80]	Continuous optimization	Multi-objective Optimization
[81]	Multi-objective environmental/economic dispatch	Multi-objective Optimization
[82]	Economic emission dispatch	Multi-objective Optimization

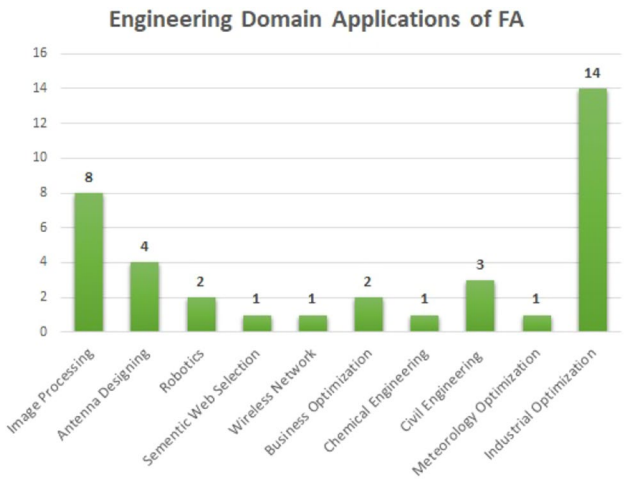


Fig. 14 Various Engineering application of Firefly algorithm

translation parameter was done. They used NCC (Normalized Cross correlation) for evaluating fitness value.

Horng et al. [84] generate vector quantization codebook by using firefly algorithm . Linde-Buzo-Gray (LBG) method was used to initialize FA by which vector quantization algorithm was designed. The proposed LBG-FA is comare with LGB, HBMO-LBG (Honey Bess Mating Optimization-LBG) and PSO-LBG (Particle Swarm Optimization-LBG) and concluded as the fastest one among others [85]. A improved image is reconstructed using LBG-FA which was almost similar to image reconstructed using HBMO-LGB algorithm.

Influencing by firefly algorithm, An Adaptive local enhancement algorithm was proposed by Hassanzadeh et al. [86]. They used FA to find best optimal image enhancing parameter.Entropy with edge number and edge intensity define the evolutionary criteria. The proposed algorithm found better performing as compared to HE(Histogram Equalization), GAIE(image enhancement using GA), PSOIE(image enhancement using PSO) and LCS(Linear contrast starching).

Hassanzadeh [87], accuracy of Segmentation and optimization run time was improved using FA. He used FA to get the optimized solution of otsu's method. The developed FA was know as Maximum variances Intra- cluster FA.

Noor et al. [88] presented a FA to generate multilevel thresholding using otsu's method. Band segmentation from the background of gel electrophoresis(GE) image was done using various image processing technique. They applied some image pre-processing methods to remove noise of the image. The proposed algorithm effectively separates DNA bands from background in a better way.

FA was used to newly design multilevel MET algorithm. Horng et al. [89] proposed maximum entropy based firefly threshold method.

M.Horng [90] proposed a FA based MCET algorithm (minimum cross entry threshold). The generated results were compared with PSO, Quantum PSO, exhaustive search, and HBMO (honey bee mating optimization). It was observed that proposal was outperforming in terms of efficiency to find optimally closed multiple thresholds with five thresholds, which are similar to results obtained by HBMO results. The execution time was comparatively less than other algorithms.

5.2.2 Antenna Designing

Basu et al. [91] used FA and ABC(artificial bee colony) for linear array isotropic antenna design optimization. They considered 3–3 instances of broad side array and steerable linear array having many array pattern direction with many array elements. They evaluated radiation pattern having minimum SLL(Side Lobe Level) for specific FNBW(First Null Beam Width) and HPBW(Half Power Beam Width). The proposed algorithm find out non-uniform excitation of each element. Six different antenna configuration cases were used for testing the effectiveness of proposal.

Basu et al. [92] applied FA meta-heuristics to optimize Side lobe level with FNBW for thinned 2-ring UCCA (Uniform Concentric Circular Array) antenna.Prefix and without prefix thinning was optimized. Firefly algorithm optimized 2 successive concentric ring having thirty five and seventy element of UCCA.

FA and PSO was used to determine the optimum distance between elements of radiating antenna, which were responsible for creation of predefined arbitrary pattern of radiation [93]. FA was compared with PSO for finding global best solution and convergence rate. FA was applied to optimize amplitude distribution of array elements by Chatterjee et al. [94].

5.2.3 Robotics

The concept of coupled biological oscillator having firefly algorithm influence was developed to achieve self organised synchronizing gait pattern for robot walk [95]. The results showed that it was useful to apply nature inspired firefly algorithm to achieve the objective. Severin et al. [96] used natural meta heuristics to automate the optimizing process of blended PTP movement of industrial robots.

5.2.4 Semantic Web Selection

Pop et al developed a hybrid FA for semantic web service composition to determine its optimal selection of solutions [97]. Semantic similarities and QoS attributes were used as selection criterion.

5.2.5 Wireless Network

Breza et al. [98] used nature inspired solution for lightweight WSN decentralization. They used few bio inspired parameter for efficient management of WSN.

5.2.6 Business Optimization

Yang et al. [99] used Non linear support vector machine with A-PSO(accelerated PSO) for solving a business optimizing framework. The proposed work was used for job scheduling, income prediction and optimizing the production.

Giannakouris et al. [100], proposed a nature inspired algorithm for portfolio asset and weight optimization. They combined hybrid intelligent meta-heuristics of FA and ACO for formulating problem of portfolio management.

5.2.7 Chemical Engineering

FA with SCE (Shuffled Complex Evolution) and CMA-ES(Covariant Matrix Adaptation-Evolution Strategy) were evaluated for determination of phase and chemical equilibrium [101]. Performance to determine phase phase stability and phase equilibrium is compared with other methods.

5.2.8 Civil Engineering

Talatahari et al. [102], used FA for optimally designing the towers. They proposed an adaptive FA in which constraints were handled using feasibility methods. The proposed work improve the convergence and inexpensively optimize the large scale structures. They used 3 different towers to evaluate the performance FA.

Azad et al. [103] used FA for optimizing the design of truss structures. They considered size and geometric parameters for this optimization problem. firefly movement were modified to propose this method.

Gholizadeh et al. [104] optimize the shape and size of truss using HS(harmony search), FA and PSO. The converging behaviour of these nature inspired algorithms was focused.

5.2.9 Meteorology Optimization

Santos et al. [105] used FA to optimize the weight for determination of precipitation area . Multi-model approach was used to determine precipitation field where weighted average of rainfall parameter were used for calculation. Finding numerical values for such average was an optimization problem. Firefly algorithm was used for this optimization.

5.2.10 Industrial Optimization

Apostolopoulos et al. [106] used FA to cut off fuel and emission cost from power generating unit . These two are conflicting objectives of economic emission load dispatch problem. he formulated a mathematical and applied proposed meta heuristic firefly algorithm.

Economic dispatch(ED) was addressed using firefly algorithm to find the feasible optimized solution [107]. They tried to optimize the objective having non linear cost and operational constraints. The results showed that proposed FA was suitable option to solve ED problem by determining most economical load.

Mauder et al. [108] proposed an algorithm to achieve high quality products and high rate of production. FA was used to set the controlling parameters and improve the quality of produced steel.

Chatteree et al. obtained dual radiation pattern of concentric ring array based isotropic antenna [109]. Different switch combinations of specific excitation and optimized excitation of array element was determined by firefly algorithm. PSO with same conditions and criteria was used to compare the impact of FA. The considered problem was effectively optimized by FA over PSO. Firefly algorithm was modified to obtain optimal solution for truss designing considering geometry and size of structure [103].

Aungkulanon et al. [110] used FA and PSO for process improvement of simulated manufacturing. They compare convergence rate, result quality and processing time taken for this process and used benchmark methods for analysis.

Rampriya et al. [111] used FA for determining the solution of PBUC (Profit based unit commitment) problem. Lagrangian FA was used to determine on/off schedule of units. they used FA for power generation. It was used for spinning and non-spinning reserve generation. The profit was maximised using the proposed method.

Chandrasekaran tried to solve SCUC(Security Constrained Unit Commitment) problem of Hybrid power systme having reliablility contraints using a variant of firefly algorithm, i.e., BRCFA(Binary Real Coded FA) [112].

Hu et al. [113] solved Energy efficient mathematical model of trains using firefly algorithm. By using FA, they were able to optimize the energy consumption for trains. They also optimize the consumption of material used for power generation. The proposed method were able to optimize the control operation of trains.

Power dispatching optimization was done using FA [114]. They used PSO for comparative analysis . The proposed work was tested on 2 thermal power plant and an IEEE-14. It was found that FA was most efficient to optimize the cost that too in just one second.

Roeva used firefly algorithm for tuning PID controller parameter in smith predictor [115]. Roeva et al. [116] adapt firefly algorithm parameter identification cultivation process of E. coli fed-Batch .

Abedinia et al. [117] proposed a version of FA for Fuzzy PID controller. This proposal address the LFC(Load Frequency Control) problem for deregulated environment. The results were compared with classical PID and fuzzy PID controller using FD indices and ITAE.

Dutta et al [118]. find out that obtaining isospectral system was a error minimizing optimization problem. he used FA to do this optimization based on local search and find it efficient.

The list of various engineering application shown in Table 4.

5.3 Classification Problems

Banati et al. [119] proposed a RST(Rough Set Theory) based Hybrid FA for finding feature set. They simulated the feature selection process.

Horng [120] used FA in the domain of neural network. It was used for training of RBF(Redial Basis Function) type neural network. RBF is an activation function of neural Network. They tested FA on 5 UCI ML Repository data set and check the performance of propose algorithm on the basis of complexity, correctness of classification and mean square error.

Firefly algorithm was used for data object clustering [121]. This makes different groups on the basis of the data attributes. ABC, PSO and other SI algorithms were used for comparing the effectiveness of FA for clustering. CEP(Classification Error Percentage) was used as performance measuring criteria. FA obtained as most efficient algorithm.

Use of Firefly algorithm for classification is showed in Table 5.

6 Conclusions and Future Scope

X.S. Yang Developed firefly algorithm in 2008. after its establishment, a wide expansion has been done in various application domain. all most all practical domain uses FA and is dynamically increasing day by day.

we have seen that firefly algorithm posses properties like :

- Increased convergence rate.
- can be used for multi- optimization problems.
- can work as local search heuristic.
- emerged as global problem solver.
- can be adapted in any problem domain.

much more.

we have categorize application domain of FA as Engineering applications, optimization, and classification.

Table 4 Various Engineering applications of Firefly algorithm

References	Used for	Category
[83]	Rigid image registration	Image processing
[84]	Image vector quantization	Image processing
[85]	Vector quantization for image compression	Image processing
[86]	Non-linear grayscale image enhancement	Image processing
[87]	Image segmentation	Image processing
[88]	Multilevel thresholding	Image processing
[89]	Multilevel image thresholding selection	Image processing
[90]	Cross entropy threshold	Image processing
[91]	Linear array antenna	Antenna designing
[92]	2 ring circular array antenna	Antenna designing
[93]	Linear antenna designing	Antenna designing
[94]	Concentric ring array antenna	Antenna designing
[95]	Self-synchronization of robot	Robotics
[96]	Industrial robots	Robotics
[97]	Optimal semantic web service composition	Semantic Web
[98]	Wireless sensor networks	Wireless network
[99]	Support vector machine	Business Optimization
[100]	Financial portfolio optimization	Business Optimization
[101]	Co-variance matrix adaptation	Chemical engineering
[102]	Tower structures	Civil engineering
[103]	Optimum design of structures	Civil engineering
[104]	Optimum design of trusses	Civil engineering
[105]	Precipitation field	Meteorology optimization
[106]	Economic emissions load dispatch problem	Industry optimization
[107]	Non-convex economic dispatch problems	Industry optimization
[108]	Steel slabs Casting	Industry optimization
[109]	Ring array antenna	Industry optimization
[110]	Simulated manufacturing process	Industry optimization
[111]	Unit commitment	Industry optimization
[112]	Demand response scheduling	Industry optimization
[113]	Train's energy saving	Industry optimization
[114]	Economic power dispatching	Industry optimization
[115]	PID controller	Industry optimization
[116]	*E. coli* cultivation	Industry optimization
[117]	Fuzzy PID	Industry optimization
[118]	Exploring isospectral spring-mass systems	Industry optimization

Table 5 Firefly algorithm for classification problems

References	Used for	Category
[119]	Feature selection approach	Classification
[120]	Data classification and disease diagnosis	Classification
[121]	Clustering	Classification

trending research areas such as image processing, robots, structure designing, antenna designing as well as WSN have used FA to optimize their objectives and results were inspiring towards use of firefly algorithm. various kind of engineering optimization successfully used FA to get global optimum objective.

Combinatorial optimization is an very important domain of optimization where large number of FA variant have been evolved. Continuous and Multi-objective optimization problem can be targeted using modified version of firefly algorithm. Noisy and constrained problem of optimization has been solved in many papers illustrated above.

Not even this, some of NP-Hard problems like TSP(travelling salesmen problem),queuing system optimization, permutation flow shop scheduling, and Graph scheduling were solved using firefly algorithm. which give us an

option to address other NP-Hard problems. this is a very big achievement of FA.

Machine learning is a latest and rapidly growing areas in which we have seen many application to solve classification problem using firefly algorithm. Artificial intelligence and designing of neural networks also have application door open to FA.

All the above studies about firefly algorithm make is very promising algorithm to obtain better, efficient and effective solution of various kind of problems.

in this paper a twofold studies of FA application has been done. the application study shows that FA can be practically used in almost all the domains. the review of above studies Project firefly algorithm as, a simple to use, versatile, and very flexible algorithm. a large number of diverse real world problems can be efficiently solved by FA. many open challenges have been discovered from the above study which can attract researchers to use firefly algorithm in their problem domain in near future.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Blum C, Li X (2008) Swarm intelligence in optimization. Springer, Berlin, pp 43–85
2. Chakraborty A, Kar AK (2017) Swarm intelligence: a review of algorithms. Springer, Cham, pp 475–494
3. Beekman M, Sword GA, Simpson SJ (2008) Biological foundations of swarm intelligence. Springer, Berlin, pp 3–41
4. Yang X-S (2008) Nature-inspired metaheuristic algorithms. Luniver press, London
5. Črepinšek M, Mernik M, Liu S-H (2011) Analysis of exploration and exploitation in evolutionary algorithms by ancestry trees. Int J Innov Comput Appl 3(1):11–19
6. Fister I, Fister I, Yang X-S, Brest J (2013) A comprehensive review of firefly algorithms. Swarm Evol Comput 13:34–46
7. Lee SA (2010) k-phase oscillator synchronization for graph coloring. Math Comput Sci 3(1):61–72
8. Yang X-S (2009) Firefly algorithms for multimodal optimization. In: International symposium on stochastic algorithms. Springer, pp 169–178
9. Yang X-S (2010). Firefly algorithm, stochastic test functions and design optimisation. arXiv:1003.1409
10. Yang X-S (2011) Review of metaheuristics and generalized evolutionary walk algorithm. arXiv:1105.3668
11. Parpinelli RS, Lopes HS (2011) New inspirations in swarm intelligence: a survey. Int J Bio-Inspired Comput 3(1):1–16
12. Zang H, Zhang S, Hapeshi K (2010) A review of nature-inspired algorithms. J Bionic Eng 7(4):S232–S237
13. Eiben AE, Hinterding R, Michalewicz Z (1999) Parameter control in evolutionary algorithms. IEEE Trans Evol Comput 3:124–141
14. Eiben AE, Smith JE et al (2003) Introduction to evolutionary computing, vol 53. Springer, Berlin
15. Tilahun S, Ong HC (2012) Modified firefly algorithm. J Appl Math 467631:11
16. Palit S, Sinha SN, Molla MA, Khanra A, Kule M (2011) A cryptanalytic attack on the knapsack cryptosystem using binary firefly algorithm. In: 2011 2nd international conference on computer and communication technology (ICCCT-2011), pp 428–432
17. Falcon R, Almeida M, Nayak A (2011) Fault identification with binary adaptive fireflies in parallel and distributed systems. In: 2011 IEEE congress of evolutionary computation (CEC), pp 1359–1366
18. Yang H, Elhadef M, Nayak A, Yang X (2008) Network fault diagnosis:aAn artificial immune system approach. In: 2008 14th IEEE international conference on parallel and distributed systems, pp 463–469
19. Falcon R, Almeida M, Nayak A (2010) A binary particle swarm optimization approach to fault diagnosis in parallel and distributed systems. In: IEEE congress on evolutionary computation, pp 1–8
20. Chandrasekaran K, Simon SP (2012) Network and reliability constrained unit commitment problem using binary real coded firefly algorithm. Int J Electr Power Energy Syst 43(1):921–932
21. Farahani SM, Abshouri A, Nasiri B, Meybodi M (2011) A gaussian firefly algorithm. Int J Mach Learn Comput 1(5):448
22. Yang X-S (2011) Metaheuristic optimization: algorithm analysis and open problems. In: Pardalos PM, Rebennack S (eds) Experimental algorithms. Berlin, Heidelberg, pp 21–32
23. Yang X-S (2010) Firefly algorithm, lévy flights and global optimization. In: Bramer M, Ellis R, Petridis M (eds) Research and development in intelligent systems XXVI. Springer, London, pp 209–218
24. Yang X-S (2012) Efficiency analysis of swarm intelligence and randomization techniques. J Comput Theor Nanosci 9(2):189–198
25. dos Santos Coelho L, de Andrade Bernert DL, Mariani VC (2011) A chaotic firefly algorithm applied to reliability-redundancy optimization. In: 2011 IEEE congress of evolutionary computation (CEC), pp 517–521
26. Gandomi A, Yang X-S, Talatahari S, Alavi A (2013) Firefly algorithm with chaos. Commun Nonlinear Sci Numer Simul 18(1):89–98
27. Subutic M, Tuba M, Stanarevic N (2012) Parallelization of the firefly algorithm for unconstrained optimization problems. Latest Adv Inf Sci Appl 22(3):264–269
28. Husselmann AV, Hawick K (2012) Parallel parametric optimisation with firefly algorithms on graphical processing units. In: Proceedings of the international conference on genetic and evolutionary methods (GEM12). Number CSTN-141, Las Vegas, USA, CSREA (16–19 July 2012), pp 77–83
29. Yang X-S, Deb S (2010) Eagle strategy using Lévy Walk and firefly algorithms for stochastic optimization. Springer, Berlin, pp 101–111
30. Luthra J, Pal SK (2011) A hybrid firefly algorithm using genetic operators for the cryptanalysis of a monoalphabetic substitution cipher. In: 2011 world congress on information and communication technologies, pp 202–206
31. Abdullah A, deris SB, Mohamad M, Hashim SM (2012) A new hybrid firefly algorithm for complex and nonlinear problem, 03
32. Fister I, Mernik M, Filipič B (2013) Graph 3-coloring with a hybrid self-adaptive evolutionary algorithm. Comput Optim Appl 54(3):741–770
33. Hassanzadeh T, Faez K, Seyfi G (2012) A speech recognition system based on structure equivalent fuzzy neural network trained

by firefly algorithm. In: 2012 international conference on biomedical engineering (ICoBE), pp 63–67
34. Nandy S, Sarkar PP, Das A (2012) Analysis of a nature inspired firefly algorithm based back-propagation neural network training. arXiv:1206.5360
35. Hassanzadeh T, Meybodi MR (2012) A new hybrid algorithm based on firefly algorithm and cellular learning automata. In: 20th Iranian conference on electrical engineering (ICEE2012), pp 628–633
36. Farahani S, Abshouri A, Nasiri B, Meybodi M (2012) Some hybrid models to improve firefly algorithm performance. Int J Artif Intell 8:97–117
37. Rajini A, David VK (2011) A comparative performance study on hybrid swarm model for microarray data. Int J Comput Appl 975:8887
38. Yang X-S (2010) Engineering optimization: an introduction with metaheuristic applications. Wiley, Hoboken
39. Shafaati M, Mojallali H (2012) Modified firefly optimization for IIR system identification. Control Eng Appl Inform 14:59–69
40. Shakarami MR, Sedaghati R (2014) A new approach for network reconfiguration problem in order to deviation bus voltage minimization with regard to probabilistic load model and DGS. Int J Electr Comput Eng 8(2):430–435
41. Olamaei J, Moradi M, Kaboodi T (2013) A new adaptive modified firefly algorithm to solve optimal capacitor placement problem. In: 18th Electric power distribution conference, pp 1–6
42. Kavousi-Fard A, Samet H, Marzbani F (2014) A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. Expert Syst Appl 41(13):6047–6056
43. Yu S, Yang S, Su S (2013) Self-adaptive step firefly algorithm. J Appl Math 2013
44. Yang X-S (2012) Multiobjective firefly algorithm for continuous optimization. Eng Comput 29:175–184
45. Yu S, Su S, Lu Q, Huang L (2014) A novel wise step strategy for firefly algorithm. Int J Comput Math 91(12):2507–2513
46. Lin X, Zhong Y, Zhang H (2013) An enhanced firefly algorithm for function optimisation problems. Int J Model Ident Control 18(2):166–173
47. Jansi S, Subashini P (2015) A novel fuzzy clustering based modified firefly algorithm with chaotic map for mri brain tissue segmentation. MAGNT Res Rep 3(1):52–58
48. Long NC, Meesad P, Unger H (2015) A highly accurate firefly based algorithm for heart disease prediction. Expert Syst Appl 42(21):8221–8231
49. Abdel-Basset M (2020) Chaotic firefly algorithm for solving definite integral. Int J Inf Technol Comput Sci 6:19–24
50. Khalil A (2014) Improved firefly algorithm for unconstrained optimization problems. Int J Comput Appl Technol Res 4(1):77–81
51. Selvarasu R, Kalavathi MS, Rajan CCA (2013) Svc placement for voltage constrained loss minimization using self-adaptive firefly algorithm. Arch Electr Eng 62(4):649–661
52. Meena S, Chitra K (2015) Modified approach of firefly algorithm for non-minimum phase systems. Indian J Sci Technol 8(23):1
53. Yan X, Zhu Y, Wu J, Chen H (2012) An improved firefly algorithm with adaptive strategies. Adv Sci Lett 16(1):249–254
54. Sulaiman MH, Daniyal H, Mustafa MW (2012) Modified firefly algorithm in solving economic dispatch problems with practical constraints. In: 2012 IEEE international conference on power and energy (PECon), pp 157–161
55. Wang B, Li D-X, Jiang J-P, Liao Y-H (2016) A modified firefly algorithm based on light intensity difference. J Comb Optim 31(3):1045–1060
56. Othman MM, Hegazy YG, Abdelaziz AY (2015) A modified firefly algorithm for optimal sizing and siting of voltage controlled distributed generators in distribution networks. Periodica Polytech Electr Eng Comput Sci 59(3):104–109
57. Subramanian R, Thanushkodi K (2013) An efficient firefly algorithm to solve economic dispatch problems. Int J Soft Comput Eng 2(1):52–55
58. Liu C, Zhao Y, Gao F, Liu L (2015) Three-dimensional path planning method for autonomous underwater vehicle based on modified firefly algorithm. Math Probl Eng
59. Fister I, Yang X-S, Brest J, Fister IJ (2013) 4-memetic self-adaptive firefly algorithm. In: Yang X-S, Cui Z, Xiao R, Gandomi AH, Karamanoglu M (eds) Swarm intelligence and bio-inspired computation. Elsevier, Oxford, pp 73–102
60. Fu Q, Liu Z, Tong N, Wang M, Zhao Y (2015/07) A novel firefly algorithm based on improved learning mechanism. In: International conference on logistics engineering, management and computer science (LEMCS 2015). Atlantis Press
61. dos Santos Coelho L, Mariani VC (2013) Improved firefly algorithm approach applied to chiller loading for energy conservation. Energy Build 59:273–278
62. dos Santos Coelho L, Mariani VC (2012) Firefly algorithm approach based on chaotic tinkerbell map applied to multivariable pid controller tuning. Comput Math Appl 64(8):2371–2382
63. Łukasik S, Żak S (2009) Firefly algorithm for continuous constrained optimization tasks. In: International conference on computational collective intelligence. Springer, pp 97–106
64. Gandomi AH, Yang X-S, Alavi AH (2011) Mixed variable structural optimization using firefly algorithm. Comput Struct 89(23–24):2325–2336
65. Gomes HM (2012) A firefly metaheuristic structural size and shape optimisation with natural frequency constraints. Int J Metaheuristics 2(1):38–55
66. Durkota K (2011) Implementation of a discrete firefly algorithm for the QAP problem within the sage framework. Bachelor thesis, Czech Technical University
67. Sayadi M, Ramezanian R, Ghaffari-Nasab N (2010) A discrete firefly meta-heuristic with local search for makespan minimization in permutation flow shop scheduling problems. Int J Ind Eng Comput 1(1):1–10
68. Yousif A, Abdullah AH, Nor SM, Abdelaziz AA (2011) Scheduling jobs on grid computing using firefly algorithm. J Theor Appl Inf Technol 33(2):155–164
69. Hönig U (2010) A firefly algorithm-based approach for scheduling task graphs in homogeneous systems. Proc Inform 724
70. Jati GK Suyanto (2011) Evolutionary discrete firefly algorithm for travelling salesman problem. I: Bouchachia A (ed) Adaptive and intelligent systems. Springer, Berlin, pp 393–403
71. Khadwilard A, Chansombat S, Thepphakorn T, Chainate W, Pongcharoen P (2012) Application of firefly algorithm and its parameter setting for job shop scheduling. J Ind Technol 8(1):49–58
72. Liu C, Gao Z, Zhao W (2012) A new path planning method based on firefly algorithm. In: *2012 fifth international joint conference on computational sciences and optimization*, pp 775–778
73. Wang G, Guo L, Duan H, Liu L, Wang H et al (2012) A modified firefly algorithm for ucav path planning. Int J Hybrid Inf Technol 5(3):123–144
74. Kwiecień J, Filipowicz B (2012) Firefly algorithm in optimization of queueing systems. Bull Pol Acad Sci Tech Sci 60(2):363–368
75. Abshouri AA, Meybodi MR, Bakhtiary A (2011) New firefly algorithm based on multi swarm & learning automata in dynamic environments. IEEE Proc 13:989–993
76. Chai-Ead N, Aungkulanon P, Luangpaiboon P (2010) Bees and firefly algorithms for noisy non-linear optimisation problems. In: World congress on engineering 2012, July 4–6, 2012. London,

UK., vol. 2189. International Association of Engineers, pp 1449–1454
77. Farahani SM, Nasiri B, Meybodi MR (2011) A multiswarm based firefly algorithm in dynamic environments. In: Third international conference on signal processing systems (ICSPS2011), Citeseer, vol 3, pp 68–72
78. Nasiri B, Meybodi M (2012) Speciation based firefly algorithm for optimization in dynamic environments. Int J Artif Intell 8(S12):118–132
79. Sulaiman MH, Mustafa MW, Azmi A, Aliman O, Rahim S R Abdul (2012) Optimal allocation and sizing of distributed generation in distribution system via firefly algorithm. In: 2012 IEEE international power engineering and optimization conference Melaka, Malaysia, pp 84–89
80. Yang X-S (2013) Multiobjective firefly algorithm for continuous optimization. Eng Comput 29(2):175–184
81. Abedinia O, Amjady N, Naderi MS (2012) Multi-objective environmental/economic dispatch using firefly technique. In: 2012 11th international conference on environment and electrical engineering, pp 461–466
82. Niknam T, Azizipanah-Abarghooee R, Roosta A, Amiri B (2012) A new multi-objective reserve constrained combined heat and power dynamic economic emission dispatch. Energy 42(1):530–545
83. Zhang Y-D, Wu L (2012) A novel method for rigid image registration based on firefly algorithm. In: International journal of research and reviews in soft and intelligent computing (IJRRSIC), vol 2, 01
84. Horng M-H, Jiang T-W (2010) The codebook design of image vector quantization based on the firefly algorithm. In: Pan J-S, Chen S-M, Nguyen N-T (eds) Computational collective intelligence technologies and applications. Springer, Berlin, pp 438–447
85. Horng M-H (2012) Vector quantization using the firefly algorithm for image compression. Expert Syst Appl 39:1078–1091
86. Hassanzadeh T, Vojodi H, Mahmoudi F (2011) Non-linear grayscale image enhancement based on firefly algorithm. In: Proceedings of the second international conference on swarm, evolutionary, and memetic computing—volume part II, SEMCCO'11. Springer, Berlin, pp 174–181
87. Hassanzadeh T, Vojodi H, Moghadam AME (2011) An image segmentation approach based on maximum variance intra-cluster method and firefly algorithm. In: 2011 seventh international conference on natural computation, vol 3, pp 1817–1821
88. Noor MH Mohd, Ahmad AR, Hussain Z, Ahmad KA, Ainihayati AR (2011) Multilevel thresholding of gel electrophoresis images using firefly algorithm. In: 2011 IEEE international conference on control system, computing and engineering, pp 18–21
89. Horng M, Jiang T (2010) Multilevel image thresholding selection based on the firefly algorithm. In: 2010 7th international conference on ubiquitous intelligence computing and 7th international conference on autonomic trusted computing, pp 58–63
90. Horng M-H, Liou R-J (2011) Multilevel minimum cross entropy threshold selection based on the firefly algorithm. Expert Syst Appl 38:14805–14811
91. Basu B, Mahanti G (2011) Fire fly and artificial bees colony algorithm for synthesis of scanned and broadside linear array antenna. Prog Electromagn Res B 32:169–190
92. Basu B, Mahanti G (2012) Thinning of concentric two-ring circular array antenna using fire fly algorithm. Sci Iran 19(6):1802–1809
93. Zaman MA, Matin M Abdul (2012) Nonuniformly spaced linear antenna array design using firefly algorithm. Int J Microw Sci Technol 2012
94. Chatterjee A, Mahanti GK (2013) Minimization in variations of different parameters in different φ planes of a small-size concentric ring array antenna using firefly algorithm. Annal Telecommun 68(5–6):299–306
95. Jakimovski B, Meyer B, Maehle E (2010) Firefly flashing synchronization as inspiration for self-synchronization of walking robot gait patterns using a decentralized robot control architecture. In: Müller-Schloer C, Karl W, Yehia S (eds) Architecture of computing systems (ARCS 2010). Springer, Berlin, pp 61–72
96. Severin S, Rossmann J (2012) A comparison of different metaheuristic algorithms for optimizing blended ptp movements for industrial robots. In: Su C-Y, Rakheja S, Liu H (eds) Intelligent robotics and applications. Springer, Berlin, pp 321–330
97. Pop CB, Chifu VR, Salomie I, Baico RB, Dinsoreanu M, Copil G (2011) A hybrid firefly-inspired approach for optimal semantic web service composition. Scalable Comput Pract Exp 12(3):363–370
98. Breza M, McCann JA (2008) Lessons in implementing bio-inspired algorithms on wireless sensor networks. In: 08 NASA/ESA conference on adaptive hardware and systems, pp 271–276
99. Yang X-S, Deb S, Fong S (2011) Accelerated particle swarm optimization and support vector machine for business optimization and applications. In: Fong S (ed) Networked digital technologies. Springer, Berlin, pp 53–66
100. Giannakouris G, Vassiliadis V, Dounias G (2010) Experimental study on a hybrid nature-inspired algorithm for financial portfolio optimization. In: Konstantopoulos S, Perantonis S, Karkaletsis V, Spyropoulos CD, Vouros G (eds) Artificial intelligence: theories, models and applications. Springer, Berlin, pp 101–111
101. Fateen S-EK, Bonilla-Petriciolet A, Rangaiah GP (2012) Evaluation of covariance matrix adaptation evolution strategy, shuffled complex evolution and firefly algorithms for phase stability, phase equilibrium and chemical equilibrium problems. Chem Eng Res Des 90(12):2051–2071
102. Talatahari S, Gandomi AH, Yun GJ (2014) Optimum design of tower structures using firefly algorithm. Struct Des Tall Spec Build 23(5):350–361
103. Kazemzadeh AS (2011) Optimum design of structures using an improved firefly algorithm. Int J Optim Civ Eng 1(2):327–340
104. Gholizadeh S, Barati H (2012) A comprative study of three metaheuristics for optimum design of trusses
105. dos Santos AF, de Campos Velho HF, Luz EF, Freitas SR, Grell G, Gan MA (2013) Firefly optimization to determine the precipitation field on south america. Inverse Probl Sci Eng 21(3):451–466
106. Apostolopoulos T, Vlachos A (2010) Application of the firefly algorithm for solving the economic emissions load dispatch problem. Int J Comb 2011
107. Yang X-S, Hosseini SSS, Gandomi AH (2012) Firefly algorithm for solving non-convex economic dispatch problems with valve loading effect. Appl Soft Comput 12(3):1180–1186
108. Jeklene O (2011) Optimization of the quality of continuously cast steel slabs using the firefly algorithm. Mater Tehnol 45(4):347–350
109. Chatterjee A, Mahanti GK, Chatterjee A (2012) Design of a fully digital controlled reconfigurable switched beam concentric ring array antenna using firefly and particle swarm optimization algorithm. Prog Electromagn Res 36:113–131
110. Aungkulanon P, Chai-ead N, Luangpaiboon P (2011) Simulated manufacturing process improvement via particle swarm optimisation and firefly algorithms. Lect Not Eng Comput Sci 2:03
111. Rampriya B, Mahadevan K, Kannan SS (2010) Unit commitment in deregulated power system using lagrangian firefly algorithm. In: 2010 international conference on communication control and computing technologies, pp 389–393
112. Chandrasekaran SSK (2011) Demand response scheduling in scuc problem for solar integrated thermal system using firefly algorithm. IET Conf Proc 44–44(1)

113. Hu H (2012) Fa-based optimal strategy of train's energy saving with energy materials. Adv Mater Res 485:93–96
114. Dekhici L, Borne P, Khaled B (2012) Firefly algorithm for economic power dispatching with pollutants emission. Inform Econ 16(2):45–57
115. Roeva O, Slavov T (2012) Firefly algorithm tuning of pid controller for glucose concentration control during e. coli fed-batch cultivation process. In: *FedCSIS*, pp 455–462
116. Roeva O (2012) Optimization of *e. coli* cultivation model parameters using firefly algorithm. Int J Bioautom 16(1):23–32
117. Abedinia O, Amjady N, Kiani K, Shayanfar H (2012) Fuzzy PID based on firefly algorithm: load frequency control in deregulated environment. In: Proceedings on the international conference on artificial intelligence (ICAI), p 1. The Steering Committee of The World Congress in Computer Science, Computer
118. Dutta R, Ganguli R, Mani V (2011) Exploring isospectral spring-mass systems with firefly algorithm. Proc R Soc A: Math Phys Eng Sci 467(2135):3222–3240
119. Banati H, Bajaj M (2011) Fire fly based feature selection approach. Int J Comput Sci Issues (IJCSI) 8(4):473
120. Horng M-H, Lee Y-X, Lee M-C, Liou R-J (2012) Firefly metaheuristic algorithm for training the radial basis function network for data classification and disease diagnosis. Theory New Appl Swarm Intell 4(7):115–132
121. Senthilnath J, Omkar S, Mani V (2011) Clustering using firefly algorithm: performance study. Swarm Evol Comput 1(3):164–171

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Scientific & Industrial Research
Vol. 79, September 2020, pp. 824-828

An Authenticated Enrolment Scheme of Nodes using Blockchain and Prevention of Collaborative Blackhole Attack in WSN

R K Yadav and Rashmi Mishra*

Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

Received 24 August 2019; revised 28 January 2020; accepted 08 May 2020

Security is indispensable concern part of the WSN. The reliable mechanisms and routing schemes are facing different set of encounters in the era of authentic status quo and it is also very knotty. Problem is identified in the recognition of untrusted nodes and routes followed from source to destination along with the restraint of battery status in WSN. No effective technique is there to avert the vindictive node attack. The current research article provides an algorithm of the nodes by employing the blockchain technology as far as solution to the other persisting drawbacks like identification of untrusted nodes and untrusted routes were tackled out by the help of secure AODV using blockchain. The outcomes of this research article by performing the simulation and experimental validation of the existing systems denotes the successful identification and occurrence of malicious nodes, end-to-end delay, packet delivery ratio and through put performance evaluation. The registration of the nodes in blockchain database and behaviour of black hole attack in AODV protocol was also simulated using NS2 blockchain algorithm.

Keywords: AODV, End-to-end delay, Malicious nodes, Performance evaluation, Untrusted routes

Introduction

The most feasible and available technology used for the transfer of information from one node to another node in an environment is refereed as wireless sensor network (WNS). This information is gathered by employing a larger number of sensor nodes in which the transfer of information can occur in between each other in an environment and simultaneously monitored also.[1] During transmission, sensitive data is transmitted from one node to another therefore security is needed for the data and protection of the network functioning is also required.[2] For the security purpose some of the cryptographic algorithm and blockchain had been used.

Aim and Objectives of the Research

The aim of this research is to authenticate the nodes in WSN by using Blockchain. The proposed algorithm reduces the delay time, increase the packet delivery ratio and also increase the throughput by using Secure-AODV BC protocol. The main objectives of this paper are listed below:

- To improve the packet delivery efficiency rate from source to destination.
- To identify the blackhole node and collaborative blackhole nodes in a network.
- To reduce the bandwidth utilization by the nodes.
- To improve the security of the nodes.

Blockchain Network Procedure

WSN required confidentiality, integrity, authentication, time synchronization, secure localization etc. But having so much security characteristics in WSN, there are some type of attacks such as wormhole, selective forwarding, hello flood, sinkhole, false routing attack, blackhole and collaborative blackhole attacks are there.[3,4] In blackhole attack, there is node which drop the packets coming from the authenticated node and drop the packets. A blackhole attack is easy to find but in collaborative blackhole attack, there are multiple blackhole node is exist which change their behaviour during time to time. So, it is harder to detect the collaborative attack. In view of the above security issues Zheng[5] proposed a solution for the third-party trust management in which author use self-organizing ledger system, blockchain is very suitable for the multi hope distributed wireless sensor network.[6] A lot of research was done by many researchers which carried blockchain in routing algorithm.

*Author for Correspondence
E-mail: dtuphd.rashmi@gmail.com

The blockchain is decentralized database which store all the information on every node authorized by process and mainly deal with security problem and enhanced trust in between the nodes. There are some features of blockchain which differentiate it with other schemes. The first properties are distributed ledger which stores all the information or the transaction in the blockchain such as address of the sender, the amount of money he transfers, nonce, timestamp, hash value of the block etc. The ledger is stored on each node in the network and each node participating in the blockchain have to keep this ledger for the information. This property of the distributed ledger provides the monitoring capability to each and every node so that the legality of the transaction is visible to all the nodes. Second feature of the blockchain is asymmetric encryption and authorization technology. The transaction stored in the blockchain is visible to all but the contents of the entity should be encrypted so that no one is able to access the information without the permission of the data owner or temper the stored information. Third feature if consensus algorithm which provide temper proof environment by using Proof of Work, Proof of Stack, Delegated proof of Stack, Proof of Capacity is discussed by.[5–6] Last feature is smart contract, which provides trusted environment, non-tempering and executed the predefined code stored in the blockchain by blockchain miner. Several researchers have proposed several routing schemes based on the blockchain technology.[7–11]

Experimental Details

Proposed Methodology

In our system, we choose the proof of authority (PoA) consensus algorithm which can progression transaction more capably. PoA will pre-authenticate the nodes in a WSN network. Here PoA will work in different entity such as server node, cluster head.

Registration Process using Blockchain: Registration process of any node is completed by using PoA (Proof of Authority). Assumptions considered for implementation are: RREP header is modified with additional field that is Speed of node, Threshold value of node speed is taken as speed threshold = 100 m/s.

1. Source node create the block, containing encrypted MAC address or physical address//time stamp// nonce//hash value of the block//Pus (Public key of Source) using his private key (PRs) and send it to the server node.

   PRs (E, MAC//timestamp//nonce//hash value//PUs)
2. Server node first authenticates the source node then release the blockchain.

If
MAC_Add = ! BC //Address of source (MAC) is not stored in blockchain.
THEN
Server ← MAC_Add //server add the address of source node to the blockchain and generate a token for the node.
Else
Release the blockchain and update the value corresponding to source MAC address to 1.// MAC_Add in BC set to 1.

3. Source node broadcast the generated blockchain to the network.
4. If next node has its MAC address in the block then it passes the block to the next node else add its MAC address to the blockchain, if it's not blackhole node.

Secure AODV Using Blockchain

1. Source S broadcasts RREQ message.
2. IF
{
D replies with RREP
Then
S send the message
}
END IF
3. IF

Node B (intermediate node) replies RREP and packet reaches to the next hope (*) from node A //*preceding node A is in the direction in which RREP is traversing from B towards S.

Then check
{
MAC_address in BC = 1 … (1)
&
(Speed of Node >speed threshold) … (2)
&
(Sequence No> seq_no_threshold) … (3)

// Threshold value is updated every time intermediate node receives a RREQ packet and threshold value of sequence no is calculated as sequence_number_ threshold = sequence number (of RREQ packet) * hop count
If Eqs (1) (2) and (3) are true then
GO TO Step 4.
else
GO TO Step 5.

4. IF (hop count>= 2)

Node X will send a Modified Hello signal with HopCount equal to 2 (in case hopcount = 2)

or

HopCount equal to 3 (in case hopcount > 2) to a Node (*)(*) (say Z) which is few hops (equal to hopcount) away from A.// ** all the values have to be taken from the RREP received from intermediate node B and (*)(*) Z node is in the path through which RREP packet has reached B.

5. IF

A receives acknowledgment from Z successfully then A forwards RREP to S and S will transmit the data.

Else

Node next to B is Blackhole and an alert signal will be transmitted by A to S. Else

Node B is Blackhole node and an alert signal will be transmitted by A to S. Step 5: A forwards RREP to S and S will transmit the data.

Results and Discussion

Implementation and Analysis

Several simulations scenarios on the different approaches are applied. Here represent two different comparison scenarios of the present work. Simulation Parameter used for the implementation are Linux (Ubuntu 12.04, NS-2.35, Number if nodes used for the simulations are 50, 100, 150, 225, 315, Packet size is 512, traffic type is UDP/CBR, Simulation time is 100 sec, antenna type is omni, transmission range is 1000*1000 m and routing protocols used are AODV, Secure-AODV.

The conclusions of the study in this paper are obtainable in three compartments specifically:

Delay Comparison, PDR, Throughput for Single and Cooperative Blackhole Attacks using AODV

The simulation result shows x-axis denotes the simulation node and y-axis shows the performance metrics parameter. Average end-to-end delay is shown in Fig. 1. The average delay of AODV is increased with number of nodes but after a 150-node delay is increased smoothly. The overall performance of average delay for single and cooperative blackhole attack with respect to number of nodes variation are Secure-AODV with hash function verification performance better to AODV protocols. The Performance of packet delivery ratio of single black hole-AODV is increased with 150 nodes. With the variation of number of nodes AODV routing protocol

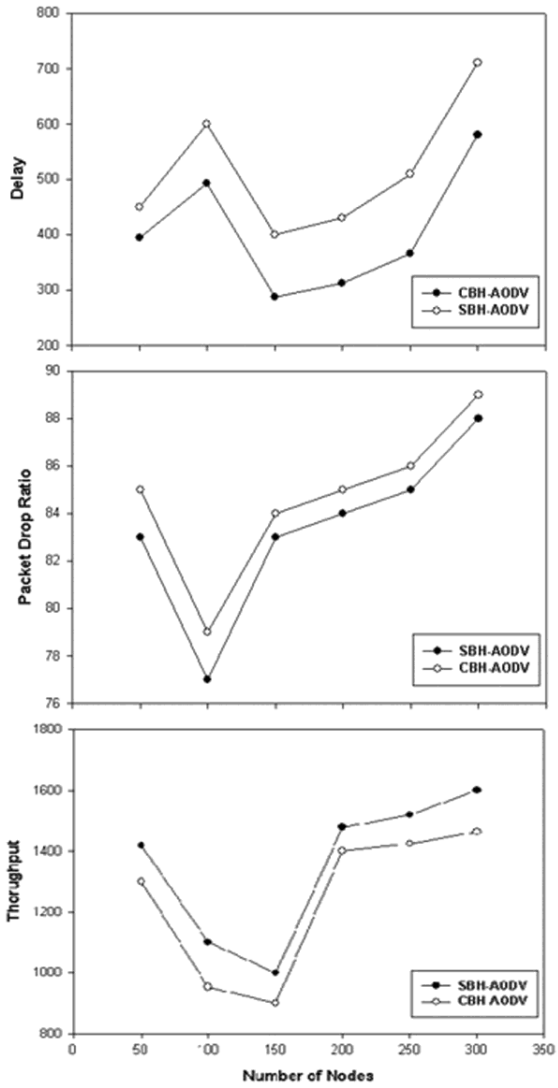


Fig. 1 — End to End Delay, PDR Throughput in AODV

packet delivery ratio is low for the single black hole attack till 150 nodes after 225 nodes it slightly increased as compared to the Cooperative. The performance of throughput for blackhole-AODV for single and cooperative blackhole are almost same for nodes 50, 100, 150 and 225 but throughput after 225 nodes is showing different performance and single blackhole-AODV increase. The maximum delay recorded on comparing both the parameters of CBH and SBH are 35.80 sec and the minimum value recorded so far on the basis of these parameters is 9.30 sec. The maximum PDR recorded on comparing both the parameters of CBH and SBH are 2.53 and the minimum value recorded so far on the basis of these parameters is 1.15. The maximum throughput

recorded on comparing both the parameters of CBH and SBH are 27.27 and the minimum value recorded so far on the basis of these parameters is 6.67.

Single and Cooperative Blackhole Attacks using Secure-AODV using Blockchain

The average end-to-end delay is shown in Fig 2. The average delay of Secure-AODV single blackhole is almost same for node 50, 100 and after 150 nodes its increase and slightly decrease at node 315. The cooperative black holes are showing less delay as compare to single black hole. It's increasing and decreasing with respect to number of node but in AODV-black hole its continuous increasing, so that our proposed Secure-AODV using blockchain is more secure and highly aware to the network. The Performance of packet delivery ratio of single and cooperative black hole in Secure-AODV using blockchain network is continuously increased as compare to AODV- Blackhole; it's increased the packet delivery ratio. The performance of throughput for blackhole-AODV for single and cooperative blackhole almost same for nodes 50, 100, 150 and 225 but as compare to the normal blackhole AODV is better, all results in secure-AODV using blockchain is high. The maximum delay recorded on comparing both the parameters of CBH and SBH is 31.78 sec and the minimum value recorded so far on the basis of these parameters is 5.45 sec. The maximum PDR recorded on comparing both the parameters of CBH and SBH are 3.66 and the minimum value recorded so far on the basis of these parameters are 1.69.The maximum throughput recorded on comparing both the parameters of CBH and SBH are 4.76 and the minimum value recorded so far on the basis of these parameters are zero.

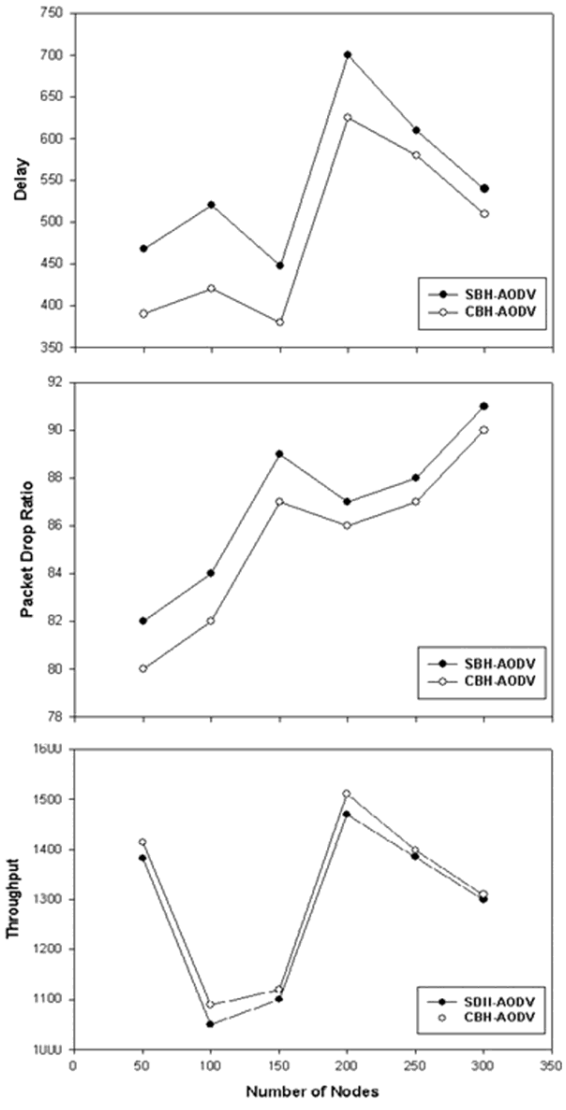


Fig. 2 — End to End Delay, PDR Throughput in secure AODV using blockchain

Conclusions

In this paper, we proposed an authenticated enrolment scheme of nodes using Blockchain and detection and prevention of blackhole and collaborative blackhole attack in WSN. As a decentralized, persistent, distributed ledger, anonymity and auditability in asymmetric encryption and authorization technology blockchain provide a feasible scheme for enrolment process of nodes. So, it is easy to trace the behaviour of each node in a network. By making each node and route perceptible and temper-proof, nodes can acquire dynamic and reliable information on the blockchain network. The detailed AODV and Secure-AODV using blockchain choose the best route and discard the routes having malicious nodes. Finally, experimental and simulation part shows that the result of our system can effectually overwhelm the occurrences of malicious nodes and the end-to-end delay and packet delivery ratio and through performance shows the novelty in research.

References

1 Filippini M, Hunt L C, Energy demand and energy efficiency in the OECD countries, a stochastic demand frontier approach, *Energy J*, **32(2)** (2011) 59–80.

2 Kalkha H, Satori H, & Satori K, A dynamic clustering approach for maximizing scalability in wireless sensor network, *Trans Mach Learn Artif Intel*, **5(4)** August (2017) 637–650.

3 Sarvari S, Sani N F M, Hanapi Z M, Abdullah M T, Wireless local area network, a comprehensive review of attacks and metrics, *J Theor Appl Inf Technol*, 95(13) (2017) 2913–2934.

4 Abraham A, Falcon R, Koeppen M, *Computational Intelligence in Wireless Sensor Networks: Recent Advances and Future Challenges*, Vol. 676, Springer SCI (2017).

5 Angrish A, Craver B, Hasan M, Starly B, A case study for blockchain in manufacturing, a prototype for peer-to-peer network of manufacturing nodes, *Procedia Manuf*, 26 (2018) 1180–1192, arXiv: 1804.01083.

6 Bach L, Mihaljevic B, Zagar M, Comparative analysis of blockchain consensus algorithm, *41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) Opatija*, 2018, pp. 1545–1550, doi: 10.23919/MIPRO.2018.8400278.

7 Lu Z, Sagduyu Y E, Li J H, Securing the backpressure algorithm for wireless networks, *IEEE Trans Mob Computing*, **16** (2017) 1136–1148.

8 Sirisala N, Bindu C S, Recommendations based QoS trusted aggregation and routing in mobile adhoc networks, *Int J Com Network Inf Secur*, **8** (2016) 215.

9 Venkataraman R, Moeller S, Krishnamachari B, Rao T R, Trust-based back pressure routing in wireless sensor networks, *Int J Sens Netw*, **17** (2015) 27–39.

10 Li J, Liang G, Liu T, A novel multi-link integrated factor algorithm considering node trust degree for blockchain-based communication, *KSII Trans Internet Inf Syst*, **11** (2017).

11 Gomez-Aravalli ADLR, Papadimitratos P, Blockchain-based public key infrastructure for inter-domain secure routing, *International Workshop on Open Problems in Network Security (iNetSec)*, May 2017, Rome, Italy. ffhal-01684192f, **5** (2017) 20–38.

12 Devi L, Shantharajah S P, Nirmal Kumar A, Authenticated and security maintance in wireless sensor network by filtering injected false data, *J Sci Ind Res*, **75(12)** (2016) 713–717.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

An overview of conventional and non-conventional hydrogen production methods

Ravindra Kumar [a,b,*], Anil Kumar [b,c,*], Amit Pal [b]

[a] Department of Mechanical and Automation Engineering, ADGITM, New Delhi 110053, India
[b] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India
[c] Centre for Energy and Environment, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Article history:
Received 19 August 2020
Received in revised form 26 August 2020
Accepted 30 August 2020
Available online xxxx

Keywords:
Hydrogen production
Industrial processes
Biological processes

ABSTRACT

H_2 fuel is adopted for the upcoming generation. Leading of H_2 technology is causes of decrement of fossil fuel and change of climate condition. There are multiple methods for the development of producing H_2 from conventional and non-conventional resources. This paper presents the major hydrogen production technologies by the industrial and biological process and a short discussion on the present scenario and economy of H_2 production and hydrogen storage. The collection of industrial techniques are involve in different kind of reforming process such as steam methane reformation (SMR), Partial oxidation (POx), auto thermal reforming (ATR) and steam iron. Biological techniques of H_2 production are adopted by various developing processes such as dark-fermentation, photo fermentation, and bio-photolysis by direct and indirect method. SMR is mostly used in the industrial process and it reaches efficiency 85%. It is available nearly 50% of globally consumption. PO_x and ATR are also another two important processes used for industrial. The biological process is producing low yields but gaining more interest due to easy and large availability of feed materials and low pollutions. Bio-H_2 production methods economy is more expensive than fossil fuel H_2 production. The requirement of more research challenges the yields of bio-H_2 to provide practical solutions. Several researchers have been done more works on bio-hydrogen using water and biomass along with solar energy, but still, it continues to improve the efficiency of development.
© 2020 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Innovations in Clean Energy Technologies.

1. Introduction

The present status of the maximum growth of energy is dependent on traditional resources like fossil fuels [1]. Day by day increasing demand for energy (as a fuel) and limitation of fossil fuels (petroleum, coal and natural gas), there is an essential requirement to make clean fuel and environmental friendly as the environment is very affected by air pollution [2]. The use of renewable energy becomes more important due to its alternative option to produce fuels as a form of energy and it focuses on making different technologies [3]. The climate issue and energy crisis have evoked a deep concern all over the world; speedy depletion of natural energy resources based of 80% is consumed from fossil fuel [4], besides searching for the new sources of energy, which must be clean, pollution free and renewable energy. The Present days, hydrogen gas introduces a potential energy source, attracting worldwide interest and having high content of energy (286 kJ per mol) [5]. Hydrogen is envisaged as alternative energy as a fuel carrier in the coming decades. It can replace the fossil fuels because of it having high potential, non– carbon-based, non-polluted and good symbol for the environment. Hydrogen energy is more gaining attention due to decrement of fossil fuel and increasing demand for energy [6]. It is having energy content/unit weight (142 kJ/g).It is safer and comfortable handling compare to other natural gases. After combustion, it produces energy and water vapors which is environmental friendly. These factors make the ideal condition for replacing fossil fuels. Now, H_2 has been considered worldwide most alternative fossil fuel, safe fuel and pollution less. In the current scenario, hydrogen is developed by natural gases, naphtha and heavy oil, coal and electrolysis etc. Maximum H_2 production (96%) is directly extracted by fossil fuels and nearly

* Corresponding authors at: Department of Mechanical and Automation Engineering, ADGITM, New Delhi 110053, India (R. Kumar), Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India (A. Kumar).
E-mail addresses: ravindra87kumar@gmail.com (R. Kumar), anilkumar76@dtu.ac.in (A. Kumar).

https://doi.org/10.1016/j.matpr.2020.08.793
2214-7853/© 2020 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Innovations in Clean Energy Technologies.

Please cite this article as: R. Kumar, A. Kumar and A. Pal, An overview of conventional and non-conventional hydrogen production methods, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2020.08.793

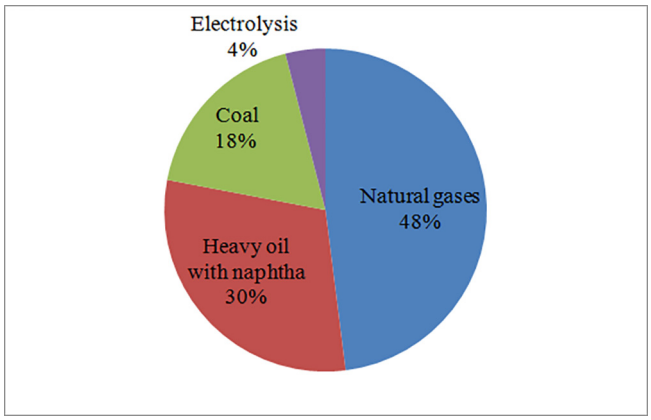


Fig. 1. Fossil fuels used in world H_2 production [9].

4% H_2 production indirectly way by using electricity from fossil fuel as feedstock, depicted in Fig. 1.At present day, biomass is also becoming more favorable H_2 production methods from the biological process due to the utilization of non-conventional recourses, and it uses at ambient temperature and pressure [7–8].

In this review paper, examine the existing industrial-based process and biological methods for H_2 production. The aims of the work to focus on a comparative overviewed of H_2 production methods mainly from hydrocarbon, biomass, and water using solar energy.

2. Present status of hydrogen production

Most of researchers and specialists adopted that hydrogen holds excellent potential as energy in the future. Two factors, such as economic growth and industrialization of developing countries have been grown the world's energy demand.

Balat [10] reported the world H_2 production was nearly 44.5 million tons per annum or 500 Bm^3 by the end of the year 2008 and the estimated demand of energy will touch nearby 600-1000EJ up to 2050 [11]. Hosseini [11] reviewed the H_2 fuel cell is favourable technology in future energy supply. Currently, resources of fossil fuel as coal, natural gas and crude oil provide approximately 80% of planetary energy demand [12]. It is assumed that consume petroleum in each day nearly 18.3 million barrels can replace hydrogen fuel cell-powered vehicles (car and light trucks) by the year 2040. It is also accepting that hydrogen-based powered vehicles are more efficient (2.5 times) of improved vehicles of gasoline and to full fill, the petroleum reduction would meet the requirement of hydrogen nearly 150 million by 2040. The overall net energy can save by the use of petroleum approximate 11 million barrels /day from all H_2 production through petroleum reforming [13]. Rohland et al. [14] focused 36% shares from renewable resources by 2025 increased to world total energy 69% shares by 2050 from the same energy and 11% by 2025 grown to 34% from hydrogen up to 2050.

3. Hydrogen production from industrial process

Many technologies are accessible for hydrogen generation by fossil fuels, but the mains are hydrocarbon reforming and most advanced methods [15].Various kinds of feedstock are available for H_2 production on industrial bases but more ideal feed stock as input material is natural gas due to the availability of huge amount and beneficial price [16].There are also use some light alcohols such as ethanol and methanol, easy handling, and globally distributed [17]. There are some important techniques of hydrogen productions, as refer in Fig. 2

3.1. Steam methane reforming (SMR)

Steam reforming is a kind of technique to produce hydrogen through different feedstock like (higher hydrocarbon, acetone, ethanol, methanol, ethane and methane) with different catalysts. Methane is a more favourable feedstock to others because of better by-product formation [18].This technique can produce hydrogen in multiple stages, as shown in Fig. 3.

Steam methane reforming of natural gas is a well-established technology that is used to generate hydrogen with the presence of steam through the conversion of hydrocarbon and implemented the first industrial application of steam methane in 1930 [19]. The steam methane reforming process involves catalytically reforming or (*syn*-gas) production, gas purification or methanation and water gas shift. It is operated at a higher temperature range 700–900 °C with the endothermic process. The necessary chemical reaction of steam methane reforming as follows [20].

Reformer Equation

$$CH_4 + H_2O \leftrightarrow 3H_2 + CO\ \Delta H = 206.2KJ/mol \text{ at } 298K \quad (1)$$

Water gas shift reaction

$$CO + H_2O \leftrightarrow H_2 + CO_2\ \Delta H = -\frac{41.1KJ}{mol} \text{ at } 298K \quad (2)$$

Methanation reaction

$$CH_4 + 2H_2O \leftrightarrow 4H_2 + CO_2\ \Delta H = -164.9\ \frac{KJ}{mol} \text{ at } 298K \quad (3)$$

The reaction of water gas shift and reverse methanation are involved in a reaction of steam reforming over a catalyst at a higher temperature of more than 700 °C [20].

3.2. Partial oxidation(PO_x)

PO_x is another standard route to H_2 production, which uses oxygen and hydrocarbon or methane fuel. The syngas (H_2 + CO) and a mixture of CO_2, N_2, H_2O, and a small amount of products such as methane are produced through partial oxidation. In a general way, oxidizing fuel is a requirement of energy balance in the presence of air. If sufficient air is not added, the fuel will oxidize incompletely, yielding mainly CO and H_2. However, it provides the sufficient air yielding, mainly CO_2 and H_2O as well. However, if the air present becomes half, at that time the chemical equilibrium will predict the mixture of CO and H_2, along with small traces of CO_2 and H_2O. The general equations are partial oxidation [21].

$$C_nH_m + \frac{n}{2}O_2 \rightarrow nCO + \frac{m}{2}H_2 \quad (4)$$

$$CH_4 + \frac{1}{2}O_2 \leftrightarrow CO + 2H_2\Delta H = -\frac{35KJ}{mol} \text{ at } 298K \quad (5)$$

Two-step reactions consist of partial oxidation of methane. In the first step reaction, total methane combustion in the presence of O_2 to get CO_2 and H_2O. This is now followed by steam reforming of un-reacted methane and CO_2 to receive syngas. The reaction accumulates with the reaction of water gas shift to obtain CO_2 and H_2O.The reactions of reforming are endothermic, however but total oxidation methane is an exothermic process ($\Delta H = -803KJ/mol$ at 298) [22].

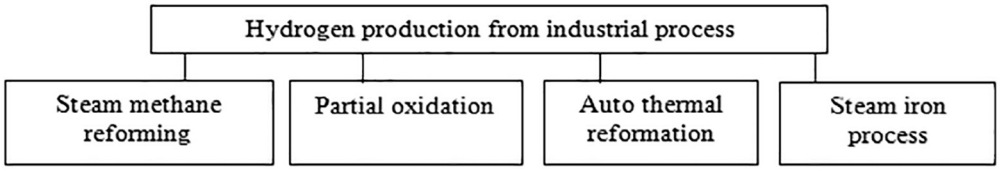


Fig. 2. Methods of H_2 production from industrial process.

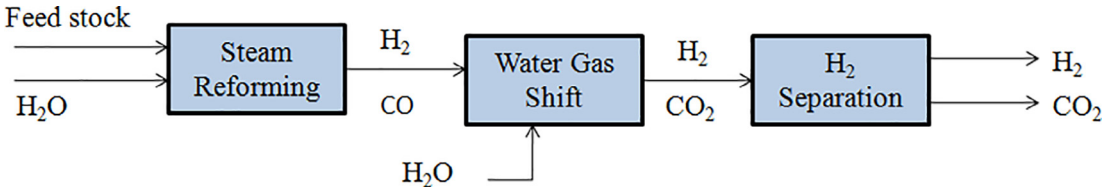


Fig. 3. Flow process of multistage for production of hydrogen [18].

3.3. Auto thermal reforming(ATR)

This approach is combined with steam reforming and PO_x methods in an adiabatic reactor. This technique also provides the neutral reaction of pairing the endothermic as steam reforming and exothermic as partial oxidation. H_2 yield is low than steam reforming, a better alternative option for fuel reforming due to neutral reaction in thermodynamically [23].

General equation:

$$C_nH_m + \frac{n}{2}O_2 + nH_2O \rightarrow \left(n + \frac{m}{2}\right)H_2 + nCO_2; \Delta H^0 \text{ reaction approx. to zero} \quad (6)$$

Natural gas converted into a gas (H_2 + CO) in the furnace of catalytic and then produced CO_2 and H_2 in catalytic shift reaction along with the water when CO reacts. CO_2 can be captured through an absorption process by amines. Aqueous Amine solvent techniques are used for absorption of CO_2 at industrial scale [24].

3.4. Steam iron process

This process is the old technique for H_2 generation. This method is started in the 20th century and mainly for balloons and airships. This is also known as a conventional method in which produces the hydrogen from iron oxide and performed the temperature nearly 540°-900 °C [25]. It is a repeated process for H_2O split by which oil and coal are utilized. This process is done to produce reducing gas by coal gasified and containing H_2 and CO. This type of gas reacts by iron oxides like magnetite (Fe_3O_4), hematite (Fe_2O_3), and wuestite (FeO) to generate an iron oxide with reduced form [25]. The regenerated iron utilize in the reactor of the steam iron, consist of oxidized through steam to generate high abundant gas of H_2 and Fe_3O_4 and touch the temperature nearly (1088 K-1143 K), represented by a chemical reaction.

$$3FeO + H_2O \rightarrow Fe_3O_4 + H_2 \quad (7)$$

Steam iron reactor constitutes the small traces N_2 and CO_2 with 61% steam and 37% H_2 and after condensation steam, improve H_2 concentration nearly 96% with 1.6% CO_2 and 2.5% N_2 [26].

4. Biological techniques of H_2 production

Hydrogen production is prepared by various methods of biological process. The biological process is an alternate technique of generating hydrogen and it is environmental friendly requires less energy-intensive and operated at atmospheric pressure and temperature [27]. Agriculture waste as raw material is the best promising candidate for the production of bio-fuels from anaerobic fermentation process [28]. Currently, the biomass is gaining more attention for hydrogen production and uses as waste material, cellulose, sucrose and starch but earlier used only glucose.

4.1. Dark fermentation

The fermentation process is done from anaerobic bacteria and green algae in dark conditions [29].It is conducted by a biochemical process that can be performed with and without the presence of O_2 and obtained the CO_2, acetone, alcohols and H_2 by the performance of microbial transformation of organic material (feed material) [30]. Carbohydrate is a good example of glucose as organic carbon. As observed from equations no.8 and 9, glucose as biomass produces a 4 mol of hydrogen in each mol of glucose when acetic acid (CH_3COOH) is the fermentation by-product and 2 mol of hydrogen in each mol of glucose can be produced as the butyrate acid (CH_3-CH_2CH_2COOH) is the fermentation by-product [31].

$$C_6H_{12}O_6 + 2H_2O \rightarrow 4H_2 + 2CH_3COOH + 2CO_2 \quad (8)$$

$$C_6H_{12}O_6 \rightarrow 2H_2 + CH_3CH_2CH_2COOH + 2CO_2 \quad (9)$$

Formation of hydrogen from dark fermentation depends on important parameters such as hydraulic retention time (HRT), P_H value and gaseous partial pressure. P_H value must be maintained in between ranges 5 to 6 for the optimal H_2 production [32].

4.2. Photo-Fermentation

This is the process of biochemical that is conducted using organic acid and solar energy under deficient nitrogen. Photosynthetic bacteria can convert organic acids like butyric, lattice and acetic into CO_2 and H_2 with the presence of nitrogenase [33]. For example, photosynthetic non-sulfur (PNS) bacteria have capable of converting volatile fatty acid s (VFA_S) into CO_2 and H_2 with the presence of an oxygenic conditions [34]. For hydrogen production, carbon sources such as sucrose glucose, succinate instead of VFA can also be capable of PNS bacteria [35]. When the acidic acid of fermentation medium, 1 mol of acetic acid generated 4 mol of H_2 as presented in Eq. (10). Ref. [36]

$$CH_3COOH + 2H_2O \overset{Light}{\rightarrow} 4H_2 + 2CO_2 \; \Delta G^O = +104KJ \quad (10)$$

4.3. Hybrid system

It is a group arrangement approach of both photo and dark fermentation from biological wastes. It is adopted as a complete conversion process compares to other processes. This system can produce a maximum yield of H_2 generation and less requirement of light energy. The system is referring in Fig. 4 [37].

Hydrogen can produce from dark conditions using anaerobic bacteria with different carbohydrates and comes out organic acid as a source for bacteria of photosynthetic to obtain additional production of hydrogen [38].The hybrid system is also known as two-stage process and sequential order of dark-fermentation process. These stages are represented by equations as follows [39].

$$\text{Dark Fermentation(StageI)} : C_6H_{12}O_6 + 2H_2O \rightarrow 2CO_2 + 2CH_3COOH + 4H_2 \quad (11)$$

$$\text{Photo fermentation(StageII)} : 2CH_3COOH + 4H_2 \xrightarrow{\text{Light}} 4CO_2 + 8H_2 \quad (12)$$

$$\text{Overallreaction} : C_6H_{12}O_6 + 6H_2O \rightarrow 12H_2 + 6CO_2 \quad (13)$$

4.4. Bio-photolysis

Bio-photolysis is utilized during the photosynthesis process using the ability of plants (mainly algae) to dissociate water. It is also a biological process in which split of water molecules into hydrogen and oxygen presence of sunlight using isolated photosynthetic membrane and other catalysts. This process is mainly decomposition of water using photons with biological catalysts, known as photolysis of water. It is two categories of bio photolysis: Direct and indirect.

4.4.1. Direct bio-photolysis

This process is used microalgae (green algae with cyanobacteria) through the presence of solar energy. It consists of photosynthetic reactions where as water is transformed into hydrogen represented by a general reaction [37].

$$2H_2O + \text{Lightenergy} \rightarrow O_2 + 2H_2 \quad (14)$$

Green algae dissociate water (H_2O) molecules into hydrogen ions along with oxygen ions during photosynthesis. Anabaena cylindrica is an excellent example of cyanobacteria for direct bio-photolysis [40].

4.4.1.1. Indirect bio-photolysis. This process is done in 2 steps; the first step occurs with photosynthesis which generates glucose with O_2 while in the second step, CO_2 and H_2 are generated through glucose along with water and the general reactions as follows [37].

$$12H_2O + 6CO_2 + \text{Solarenergy} \rightarrow 6O_2 + C_6H_{12}O_6 \quad (15)$$

$$C_6H_{12}O_6 + 12H_2O + \text{Solarenergy} \rightarrow 6CO_2 + 12H_2 \quad (16)$$

Hydrogenase and nitrogenase are both responsible for hydrogen formation but more production rate of hydrogenase enzymes by green algae [41]. The process of indirect bio-photolysis is explained as shown in Fig. 5 [37].

5. Overall summary of hydrogen methods

In this section, comparative assessment on the bases of data of various kinds of H_2 production methods has been discussed. The performance evaluation of H_2 production, both technically and economically, are presented in Tables 1 and 2. These methods are highly matured and developed and already exist. They are provided the advantages, disadvantages and efficiency on the bases of hydrocarbon and renewable energy sources (Table 1).

Biomass is renewable, abundant feed material and readily available in the environment. Biological processes generate low yield and need a larger volume reactor.

The total cost of H_2 production methods depend on feedstock, source of energy and capital investment, included in Table 2. Low price of fuel such as coal and natural gas, more efficient as they were using feedstock, conventional techniques are producing H_2 at a low cost nearly 2 US$/Kg, as depicted in Table 1. H_2 production cost/kg considering and without considering CCS ranges 1.34–2.27US$.

Production of hydrogen methods presents with their efficiency and its feedstock, refer table 1. Reforming technologies are most mature and reach efficiency up to 85% for SMR and followed by POx and ATR nearly 75%. The main outlines of the above section are renewable technologies become more competitive, its related cost and need more research to develop a more efficient manner of hydrogen generation.

5.1. Economic aspect

Hay et al. [45] analyzed the production of H_2 from an economical source such as coal and natural gas comprise 0.37–1.82 USD per kg and 2.48–3.15 USD per kg. However, the cost is high as of wind and solar energy. It is assumed that using biomass or waste materials as energy sources can be reduced to H_2 production. In between the year of 2011–2016, globally hydrogen production was evaluated to expand through the compound yearly increase the rate of 5.6% and accepted that generation of H_2 market will

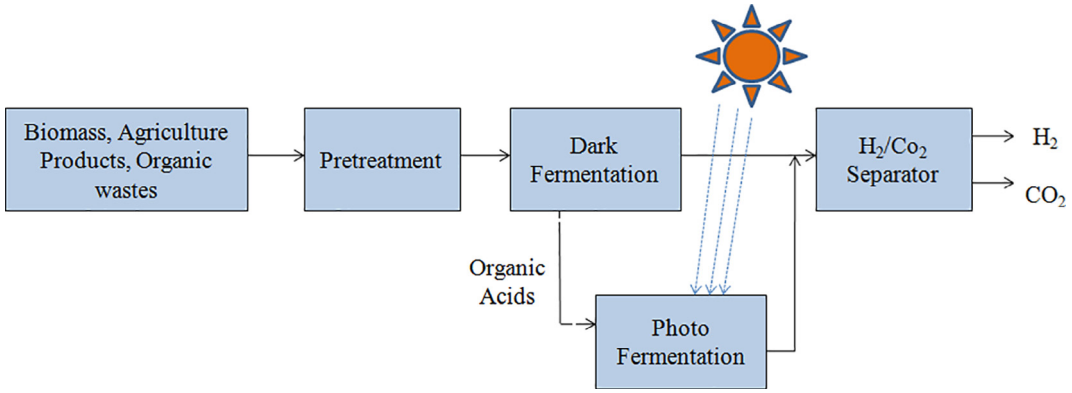


Fig. 4. Flow process of hybrid system [37].

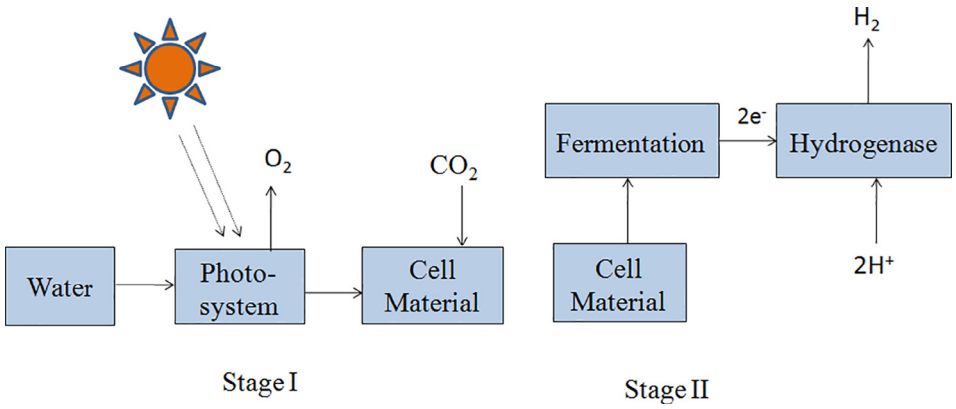


Fig. 5. Flow process of indirect bio-photolysis [37].

Table 1
Different kind of H_2 production methods with their benefits, drawbacks and efficiency [34,42,43,44].

Process	Benefits	Drawbacks	Efficiency (%)
SMR	• Vast industrial experience • No need O_2 • Low temperature required • Higher efficiency • Best H_2/CO ratio for H_2	• Air emissions are high • Complex system	70–85
Pox	• No need catalyst • System is simple • Requirement of desulphurization is less	• Lower H_2/CO ratio • Need O_2 • High operating temperature	60–75
ATR	• Favorable H_2/CO ratio • Lower methane slip • Need less oxygen than Pox • Low operating temperature compare with POx	• Need Air or O_2 • Minimal commercial experience	60–75
Dark Fermentation	• No limitation of O_2 due to anaerobic process • Higher production rate • Use variety of substrates • Simple of reactor technique • H_2 production without light	• Requirement of reactor volume is high • COD removal is low • Conversion efficiency is low • By-product is large amount • Lower H_2 yields	60–80
Photo Fermentation	• Bacteria can use light energy • Use variety of organic waste • Substrate conversion is nearly complete • CO_2 neutral	• Need solar energy • Lower rate of H_2 yields • Conversion efficiency is low • Larger reactor volume	6.6–86
Bio-photolysis by direct	• H_2 production form sun and water	• Need intensity of light	6.04–46
Bio-photolysis by Indirect	• H_2 production from water using Cyanobacteria	• Present 30% O_2 in mixture of gas	6.04–46

Table 2
Summary of various H_2 production processes with their feedstock, capital cost and hydrogen cost [42].

Method	Source of energy	Feedstock	Capital cost (M$)	H_2 cost (US$ per Kg)
Steam methane reforming by CCS (CO_2 capture and storage)	Fossil fuels	CH_4	226.3	2.27
Steam methane reforming with no CCS	Fossil fuels	CH_4	180.5	2.08
Auto thermal reforming of CH_4 by CCS	Fossil fuels	CH_4	183.3	1.48
Dark Fermentation	–	Organic Biomass	–	2.57
Photo-Fermentation	Solar light	Organic Biomass	–	2.83
Bio-photolysis by Direct method	Solar light	Water with algae	50 US$ per m^2	2.13
Bio-photolysis by Indirect method	Solar light	Water with algae	135US$ per m^2	1.42

raise nearly USD 87.3 billion up to 2011 to USD 118 billion up to 2016 using by means of CAGR of 6.21% [45]. Kalamara [46] presented a most favourable and an inexpensive method for production of H_2 is SMR (natural gas). The price of H_2 is nearly 7USD/GJ and this approach is used more than half of globally H_2 production. Hydrogen generation from biomass using pyrolysis and gasification obtain H_2 prices more than 3 times higher compared to H_2 prices produced by SMR process. H_2 price from biomass gasification and pyrolysis estimated ranges from 10 to 14 US$ per GJ and 8.91–5.51 US$ per GJ. Its dependency on feedstock cost, accessibility and equipment [47].

6. Hydrogen storage

A storage device of H_2 is an essential part and a big task of hydrogen storage for the use of a large scale. Hydrogen has high energy density and low energy density by weight and volume respectively, with the comparison to hydrocarbon [48]. The storage parameter such as gravimeter and volumetric should be higher for the devices of hydrogen storage. The material for H_2 storage device is low cost, lightweight, suitable kinetic of desorption, adsorption, and recyclability. The U.S. DOE (Department of Energy) 2015 goals are the capacity of 40 g H_2/litre and 5.5 wt% and 1500 cycles of a

lifetime [49]. 1 kg of H_2 occupies 11 m^3 volume at atmospheric pressure and ambient temperature with a lower density 0.09 kg/m^3 [50].

Hydrogen is store in a gaseous state at ambient temperature but in the liquid state need the lowest temperature (-253 °C or 20 K) using vacuum insulated vessel. The main important methods of H_2 can be stored physically, such as gas under pressure, cryogenic liquid state and frozen (solid) fuel like metal hydrides (combine physical and chemical). Highly pressurized gaseous H_2 storage is a familiar and fully grown technique as a result of its technically simple and high rate of filling, require approximate higher pressure (77 MPa) using mechanical compressor (piston-type) [51].However, the work requirement of compression is higher than 2.21 KWh/kg, consisting of gravimeter (13 wt%) and volumetric density (less than 40 kg/m^3) [52].

Cryogenic tank uses for storage of hydrogen in liquid form by two steps through compression and cooling process in H.E. (Heat Exchanger). Liquid H_2 has a low boiling point, then the work required for compressor is nearly 15.2 kWh/kg; collecting volumetric density (at atm. pressure) is 70.8 kg/m^3 but the parameter gravimeter density depends on size of the tank [51]. H_2 storage in solid state is more effective at moderate pressure, temperature and stored by adsorption (on a solid surface). In the adsorption process, when several atoms of solid surface interacted by gas molecules, it bonded and released in reverse order. A low temperate 77 K and pressure 6 MPa, stored hydrogen by carbon nanotubes and providing volumetric density (10.8 wt%) and gravimeter (41 kg/m^3) [52].

There are various metals used for storing hydrogen. Magnesium is adopted as attractive material because of lightweight, lower cost, and high capacity and it stores H_2 (7.6 wt%) [53]. Metals (lightweight) as Li, Al, B, Mg, Na and Be, a huge variety of compound metal hydrogen [51]. Metal hybrids such as $NaBH_4$, Ca_2H, LiH, MgH_2 and $LIAlH_4$ studied of storage materials for H_2 gas to fuel-cell. Best material such as $NaBH_4$ (Sodium borohydride) is used for high stability, safety purpose and higher storage capacity (10.8 wt%). However, react with water without using any catalyst and not stability in the air to generate hydrogen [54].

Most of the industries apply policy of H_2 need superior package, transportation and storage as of production site to the users. Thus, research requires for hydrogen storage material is safe, commercial and reliable. There is an important criterion for H_2 storage as the reversibility of uptake and release. Therefore, the reversibility criteria excepts the entire covalent and associated compound of hydrogen carbon because hydrogen is just released if it is heated at high temperature more than 800 °C, otherwise the carbon is oxidized [50].

7. Conclusions

In the development of hydrogen production, a tremendous amount of research is going on due to more reasons like its clean, high energy content, safe energy, and produced from different kinds of recourses such as fossil fuel, renewable other sustainability energy. Presently, Hydrocarbon reforming techniques are more used and developed for hydrogen production. Several researchers have been done the more works on bio-hydrogen using water and biomass along with solar energy, but still it continues to improve the efficiency of development.

This review paper analyzed the developments in conventional techniques used by an industrial process, leading enhancement of H_2 production and biological processes described by non-conventional techniques, promising methods for producing H_2 from some advantages such as biomass, plenteous, cleanly and renewable source.

The collection of industrial techniques are involve in different kind of reforming process such as steam methane reformation (SMR), Partial oxidation (POx), auto thermal reforming (ATR) and steam iron. Some other biological techniques of H_2 production are adopted by various developing processes such as dark fermentation, photo fermentation, and bio-photolysis by direct and indirect. The reforming techniques are most efficient and mostly used for H_2 production. Steam methane reforming (SMR) achieved nearly 85% and also reached around 75% by partial oxidation (POx) and auto thermal reforming (ATR) process. The reforming processes are not new methods, but these are the most efficient.

Due to dependency on fossil fuels, developing alternate techniques of H_2 production are necessary. Some techniques of H_2 production are adopted by various developing processes such as the biological process is one of them. The biological procedure is producing low yields but gaining more interest due to easy and large availability of feed materials and low pollutions. Since H_2 can be developed by various kind of renewable sources and efforts in minimal global warming, a lot of research work, introduced the whole world to make methods of H_2 production efficient more and more. Still, H_2 productions of conventional techniques are more attractive and economically, though non-conventional techniques are fields of scientific research.

There are two main factors such as increasing ecological issues as well as the cost of energy. Bio-fuels are generating more interest in the development of H_2 production fields. Bio- H_2 production methods economy is more expensive than fossil fuel H_2 production. The requirement of more research challenges the yields of bio-H_2 to provide practical solutions. The economic analysis point of view reuses of waste areas feed material illustrated the sustainability of commercialization bio-hydrogen, on the other hand, storage of H_2 is a big major issue, presented immediately to process utilization of H_2 fuel, and needs research to the improvement of gravimetric and volumetric density.

CRediT authorship contribution statement

Ravindra Kumar: Writing - review & editing, Writing - original draft. **Anil Kumar:** Supervision, Conceptualization. **Amit Pal:** Visualization, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Authors are highly grateful to Department of Mechanical Engineering and Centre for energy and Environment, Delhi Technological University Delhi (India) for providing basic facilities to compile this work.

References

[1] N. Gaurav, D. Mote, V. Dupont, Hydrogen production from reforming of biogas. Review of technological advances and an Indian perspective, Renew. Sustain. Energy Rev. 76 (2017) 1032–1052.

[2] C. Young N, Y. C. Yang, K. Yoshikawa, Hydrogen generation from biogas reforming using a gliding arc plasma-catalyst reformer Catalysis today 148 2009 283 289

[3] I.P. Jain, Hydrogen the fuel for 21st century, Int. J. Hydrogen Energy 34 (2009) 7368–7378.

[4] Xin Mei Guo, Eric Trably, Eric Latrille, Hélène Carrère, Jean-Philippe Steyer, Hydrogen production from agricultural waste by dark fermentation: A review, Int. J. Hydrogen Energy 35 (19) (2010) 10660–10673.

[5] M. Bao, H. Su, T. Tan, Biohydrogen Production by Dark Fermentation of Starch Using Mixed Bacterial Cultures of Bacillus sp and Brevumdimonas sp, Energy Fuels 26 (2012) 5872–5878.
[6] Maria A. Goula, Nikolaos D. Charisiou, Kiriakos N. Papageridis, Andreas Delimitis, Eleni Pachatouridou, Eleni F. Iliopoulou, Nickel on alumina catalysts for the production of hydrogen rich mixtures via the biogas dry reforming reaction: Influence of the synthesis method, Int. J. Hydrogen Energy 40 (30) (2015) 9183–9200.
[7] P. Sinha, A. Pandey, An evaluative report and challenges for fermentative biohydrogen production, Int. J. Hydrogen Energy 36 (2011) 7460–7478.
[8] D. Das, Advances in biohydrogen production processes: an approach towards commercialization, Int. J. Hydrogen Energy 34 (2009) 7349–7357.
[9] R. Kothari, D. Buddhi, R.L. Sawhney, Comparison of environmental and economic aspects of various hydrogen production methods, Renew. Sustain. Energy Rev. 12 (2) (2008) 553–563.
[10] Mustafa Balat, Potential importance of hydrogen as a future solution to environmental and transportation problems, Int. J. Hydrogen Energy 33 (15) (2008) 4013–4029.
[11] S.E. Hosseini, M.A. Wahid, Hydrogen production from renewable and sustainable energy resources: promising green energy carrier for clean development, Renew. Sustain. Energy Rev. 57 (2016) 850–866.
[12] S.E. Hosseini, M.A. Wahid, Utilization of palm solid residue as a source of renewable and sustainable energy in Malaysia, Renew. Sustain. Energy Rev. 40 (2014) 621–632.
[13] J. N. Armor Catalysis and the hydrogen economy Catal Lett 101 3-4 2005 131 135 http://link.springer.com/10.1007/s10562-005-4877-3
[14] B. Rohland, J. Nitsch, H. Wendt, Hydrogen and fuel cells — the clean energy system, J. Power Sources 37 (1-2) (1992) 271–277.
[15] Ibrahim Dincer, Canan Acar, Review and evaluation of hydrogen production methods for better sustainability, Int. J. Hydrogen Energy 40 (34) (2015) 11094–11111.
[16] Mohammad A. Rakib, John R. Grace, C. Jim Lim, Said S.E.H. Elnashaie, Bahman Ghiasi, Steam reforming of propane in a fluidized bed membrane reactor for hydrogen production, Int. J. Hydrogen Energy 35 (12) (2010) 6276–6290.
[17] Prakash D. Vaidya, Alirio E. Rodrigues, Insight into steam reforming of ethanol to produce hydrogen for fuel cells, Chem. Eng. J. 117 (1) (2006) 39–49.
[18] Trevor L. LeValley, Anthony R. Richard, Maohong Fan, The progress in water gas shift and steam reforming hydrogen production technologies – A review, Int. J. Hydrogen Energy 39 (30) (2014) 16983–17000.
[19] L. Barelli, G. Bidini, F. Gallorini, S. Servili, Hydrogen production through sorption-enhanced steam methane reforming and membrane technology: a review, Energy 33 (2008) 554–570.
[20] M.H. Halabi, M.H.J.M. de Croon, J. van der Schaaf, P.D. Cobden, J.C. Schouten, Low temperature catalytic methane steam reforming over ceria–zirconia supported rhodium, Appl. Catal. A 389 (1-2) (2010) 68–79.
[21] A.E. Lutz, R.W. Bradshaw, L. Bromberg, A. Rabinovich, Thermodynamic analysis of hydrogen production by partial oxidation reforming, Int. J. Hydrogen Energy 29 (2004) 809–816.
[22] A KOH, L CHEN, W KEELEONG, B JOHNSON, T KHIMYAK, J LIN, Hydrogen or synthesis gas production via the partial oxidation of methane over supported nickel–cobalt catalysts, Int. J. Hydrogen Energy 32 (6) (2007) 725–730.
[23] Renika Baruah, Marm Dixit, Pratik Basarkar, Dhrupad Parikh, Atul Bhargav, Advances in ethanol autothermal reforming, Renew. Sustain. Energy Rev. 51 (2015) 1345–1353.
[24] J. Castro, R. Rivera-Tinoco, C. Bouallou, Hydrogen production from natural gas: auto-thermal reforming and CO_2 capture, Chem. Eng. 21 (2010).
[25] Magnus Rydén, Mehdi Arjmand, Continuous hydrogen production via the steam–iron reaction by chemical looping in a circulating fluidized-bed reactor, Int. J. Hydrogen Energy 37 (6) (2012) 4843–4854.
[26] M Steinberg, Modern and prospective technologies for hydrogen production from fossil fuels, Int. J. Hydrogen Energy 14 (11) (1989) 797–820.
[27] S.V. Mohan, V.L. Babu, P.N. Sarma, Anaerobic biohydrogen production from dairy wastewater treatment in sequencing batch reactor (AnSBR): effect of organic loading rate, Enzyme Microb. Technol. 41 (2007) 506–515.
[28] Bin-Bin Hu, Ming-Jun Zhu, Enhanced hydrogen production and biological saccharification from spent mushroom compost by Clostridium thermocellum 27405 supplemented with recombinant β-glucosidases, Int. J. Hydrogen Energy 42 (12) (2017) 7866–7874.
[29] C.Y. Lin, C.H. Jo, Hydrogen production from sucrose using an anaerobic sequencing batch reactor process, J. Chem. Technol. Biotechnol. Int. Res. Process Environ. Clean Technol. 78 (2003) 678–684.
[30] W BALTHASAR, Hydrogen production and technology: today, tomorrow and beyond, Int. J. Hydrogen Energy 9 (8) (1984) 649–668.
[31] F.R. Hawkes, R. Dinsdale, D.L. Hawkes, I. Hussy, Sustainable fermentative hydrogen production: challenges for process optimisation, Int. J. Hydrogen Energy 27 (2002) 1339–1347.
[32] Herbert H.P. Fang, Hong Liu, Effect of pH on hydrogen production from glucose by a mixed culture, Bioresour. Technol. 82 (1) (2002) 87–93.
[33] Meng Ni, Dennis Y.C. Leung, Michael K.H. Leung, K. Sumathy, An overview of hydrogen production from biomass, Fuel Process. Technol. 87 (5) (2006) 461–472.
[34] D. Das, T.N. Veziroglu, Advances in biological hydrogen production processes Int. J. Hydrogen Energy 33(2008) 6046-6057
[35] P Westermann, B Jorgensen, L Lange, B Ahring, C Christensen, Maximizing renewable hydrogen production from biomass in a bio/catalytic refinery, Int. J. Hydrogen Energy 32 (17) (2007) 4135–4141.
[36] Tae-Young Jeong, Gi-Cheol Cha, Sung Ho Yeom, Suk Soon Choi, Comparison of hydrogen production by four representative hydrogen-producing bacteria, J. Ind. Eng. Chem. 14 (3) (2008) 333–337.
[37] Pavlos Nikolaidis, Andreas Poullikkas, A comparative overview of hydrogen production processes, Renew. Sustain. Energy Rev. 67 (2017) 597–611.
[38] S MANISH, R BANERJEE, Comparison of biohydrogen production processes, Int. J. Hydrogen Energy 33 (1) (2008) 279–286.
[39] D Das, Hydrogen production by biological processes: a survey of literature, Int. J. Hydrogen Energy 26 (1) (2001) 13–28.
[40] S. Ghosh, R. Chowdhury, P. Bhattacharya, A review on single stage integrated dark-photo fermentative biohydrogen production: insight into salient strategies and scopes, Int. J. Hydrogen Energy 43 (2018) 2091–2107.
[41] A. A. Tsygankov A. S. Fedorov S. N. Kosourov K. K. Rao Hydrogen production by cyanobacteria in an automated outdoor photobioreactor under aerobic conditions Biotechnol. Bioeng. 80 7 2002 777 783
[42] Abdalla M. Abdalla, Shahzad Hossain, Ozzan B. Nisfindy, Atia T. Azad, Mohamed Dawood, Abul K. Azad, Hydrogen production, storage, transportation and key challenges with applications: A review, Energy Convers. Manage. 165 (2018) 602–627.
[43] D.J. Wilhelm, D.R. Simbeck, A.D. Karp, R.L. Dickenson, Syngas production for gas-to-liquids applications: technologies, issues and outlook, Fuel Process. Technol. 71 (2001) 139–148.
[44] J.D. Holladay, J. Hu, D.L. King, Y. Wang, An overview of hydrogen production technologies, Catal. Today 139 (4) (2009) 244–260.
[45] J.X.W. Hay, T.Y. Wu, J.C. Juan, J.Md. Jahim, Biohydrogen production through photo fermentation or dark fermentation using waste as a substrate: overview, economics, and future prospects of hydrogen usage, Biofuels, Bioprod. Biorefin. 7 (2013) 334–352.
[46] Christos M. Kalamaras, Angelos M. Efstathiou, Hydrogen Production Technologies: Current State and Future Developments, Conference Papers in Energy 2013 (2013) 1–9.
[47] Mustafa Balat, Mehmet Balat, Political, economic and environmental impacts of biomass-based hydrogen, Int. J. Hydrogen Energy 34 (9) (2009) 3589–3603.
[48] A.L. Roes, M.K. Patel, Ex-ante environmental assessments of novel technologies – Improved caprolactam catalysis and hydrogen storage, J. Cleaner Prod. 19 (14) (2011) 1659–1667.
[49] Shengqian Ma, Hong-Cai Zhou, Gas storage in porous metal–organic frameworks for clean energy applications, Chem. Commun. 46 (1) (2010) 44–53.
[50] Andreas Züttel, Materials for hydrogen storage, Mater. Today 6 (9) (2003) 24–33.
[51] Jinyang Zheng, Xianxin Liu, Ping Xu, Pengfei Liu, Yongzhi Zhao, Jian Yang, Development of high pressure gaseous hydrogen storage technologies, Int. J. Hydrogen Energy 37 (1) (2012) 1048–1057.
[52] Li Zhou, Progress and problems in hydrogen storage methods, Renew. Sustain. Energy Rev. 9 (4) (2005) 395–408.
[53] Louis Schlapbach Andreas Züttel Vincent Dusastre Materials for Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group 2010 Co-Published with Macmillan Publishers Ltd, UK 265 270 10.1142/9789814317665_0038 http://www.worldscientific.com/doi/abs/10.1142/9789814317665_0038
[54] Tugce Turhan, Yuksel Guvenilir Avcıbası, Nurettin Sahiner, Versatile p(3-sulfopropyl methacrylate) hydrogel reactor for the preparation of Co, Ni nanoparticles and their use in hydrogen production, J. Ind. Eng. Chem. 19 (4) (2013) 1218–1225.

Analysis of a novel Metamaterial Absorber using equivalent circuit model operating at 3.5 GHz

Priyanka Garg, Priyanka Jain

***Abstract*—** This paper demonstrates a closed meander line shaped metamaterial absorber operation at 3.5 GHz WiMAX band. The proposed metamaterial absorber unit cell has a compact size of $0.11\lambda_o$ X $0.11\lambda_o$ design on an ultrathin FR-4 substrate with thickness $0.018\lambda_o$, where λ_o is the wavelength corresponding to operating frequency. The proposed absorber shows an absorptivity of 98.5% at the intended frequency. An equivalent circuit model is also defined to depict the electrical properties of the structure. The proposed design also shows insensitivity to polarization as well as change in incident angle of the wave over a wide-angle (upto 60˚) for both TE and TM polarization. The proposed structure is a good candidate for radar cross section reduction of an antenna.

***Keywords*—** Metamaterial absorber, meander line, ultra thin.

I. Introduction

ABSORBERS are used to restrict the transmission and reflection of electromagnetic wave for a certain range of frequencies. They are widely used in various wireless applications such as radar cross section reduction of antennas [1], radome, EMI/EMC applications [2], antenna side lobe reduction [3], sensing [4] and so on. Conventional electromagnetic absorbers like Salisbury screen and Jaumann absorbers are thick and bulky.

Use of metamaterial [5] in designing of what we called a 'perfect absorber' has gained popularity in last few decades. Landy et. al. [6] was among the firsts to describe the absorption characteristics of such structures. Since then various metamaterial based absorbers have been proposed in the literature [7], [8].

The proposed work describes in a simplest way to design and analyze an ultrathin metamaterial absorber along with the extraction of its equivalent circuit parameters to study the electrical behaviour of the structure for further integration of lumped components. The design is evolved from a simple square loop to a symmetrical meander line structure whose dimensions are optimized to operate at 3.5 GHz WiMAX band. The structure exhibits 98.5% absorptivity at the intended band along with insensitivity to polarization and oblique incidence.

II. Design and Simulation Approach

The proposed meander line shaped metamaterial absorber is designed on a 1.6 mm thick FR-4 substrate, having dielectric constant (ε_r) of 4.4 and loss-tangent (δ) of 0.02, with unit cell dimensions of 10mm×10mm, as shown in fig. 1(a). The top of the substrate consists of the closed meander line shape and bottom is completely laminated with copper.

Priyanka Garg and Priyanka jain are with the Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi-110042, INDIA (email: priyankagarg_phd2k16@dtu.ac.in, priyajain2000@rediffmail.com)

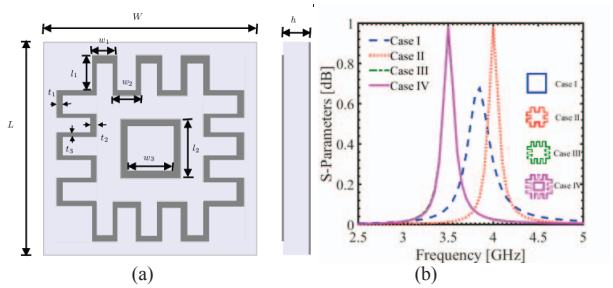


Fig. 1. (a)Top view and side view of the proposed unit cell (b) Stepwise evolution of the proposed unit cell .

The design parameters are given as: W=10mm, L=10mm, w_1=1.25mm, w_2=1.025mm, w_3=2.1mm, l_1=2.1mm, l_2=2.7mm, t_1=0.425mm, t_2=0.35mm, t_3=0.2mm, h=1.6mm. The unit cell is simulated under periodic boundary conditions by applying master-slave boundaries in Ansys HFSS. The unit cell is excited using floquet ports providing electric and magnetic field along x and y axis and propagation vector (k) along z axis.

The absorptivity A(ω) is obtained using Eq. (1)

$$A(\omega) = 1 - |S_{11}(\omega)|^2 - |S_{21}(\omega)|^2 \quad (1)$$

Where, $|S_{11}(\omega)|^2$ and $|S_{21}(\omega)|^2$ represents the reflectivity and transitivity respectively with respect to frequency ω. Since, the bottom of the substrate is completely coated with copper, so further propagation of wave is restricted which indicates zero transmission, thus, $|S_{21}(\omega)|^2$=0. Therefore, Eq. 1 reduces to

$$A(\omega) = 1 - |S_{11}(\omega)|^2 \quad (2)$$

Now, to minimize the reflections, the impedance matching of the free space with the structure is required. Since [9],

$$S_{11}(\omega) = \frac{Z(\omega) - \eta_0}{Z(\omega) + \eta_0} \quad (3)$$

where, Z(ω) is the input impedance of the absorber and η_0 is the free space impedance.

Fig. 1(b) depicts the stepwise evolution of the proposed unit cell. The Case I is a simple rectangular geometry which resulted in a weak absorption at 3.85 GHz. Now, in case II a meander line structure is created with 4 vertical lines on each arm of the rectangle within the same area. This resulted in an increase in the absorptivity however; operating frequency is shifted to 4 GHz. Further increasing the number of vertical lines (Case III), the absorption frequency is shifted to lower side at 3.5 GHz, due to increase in effective length of the path covered by surface current. The absorptivity obtained for this case is 96 %. Finally, on adding a small rectangular loop in the middle of the geometry resulted in increase in absorptivity to 98.5 % due to improved electromagnetic coupling. The final geometry resulted in the miniaturization of structure by 1.2 times.

978-1-7281-3448-2/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:55:55 UTC from IEEE Xplore. Restrictions apply.

III. Results and Discussion

The final meander line structure is simulated using periodic boundary conditions and results are obtained, shown in fig. 1(b). The absorptivity obtained is 98.5 % at 3.5 GHz. Fig. 2(a) shows the normalized impedance of the structure. Since, at 3.5 GHz, the normalized impedance of the structure is close to unity, the reflections from the structure are minimized, as per eqn. 3, giving near perfect absorption.

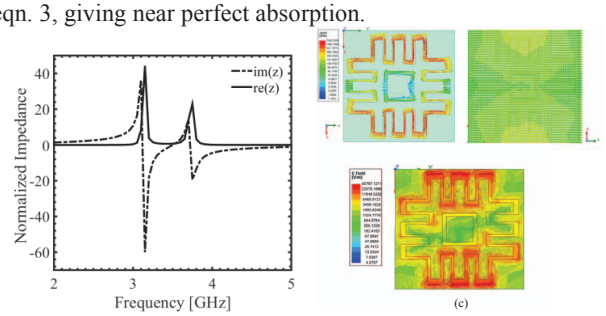


Fig. 2. (a) Normalized impedance of the proposed unit cell (b) Current distribution of top and bottom surface, (c)E-field distribution of the proposed unit cell.

Fig. 2(b) shows the surface current distribution at 3.5 GHz on top and bottom of the structure. It can be observed that the net direction of propagation of current on the upper and lower part of the geometry is from positive y axis to negative y axis and on the bottom, it is from negative y to positive y axis. Thus, the current on top and bottom surfaces are anti-parallel to each other resulting in the formation of magnetic dipole in the direction perpendicular to the applied magnetic field. Now, fig. 2(c) shows the electric field distribution on the structure at 3.5 GHz. It shows that the induced electric field is highly concentrated along the direction of incident electric field (horizontal direction in fig. 2(c)). This gives rise to electric excitation in the design. Thus, we can conclude that, at 3.5 GHz, electromagnetic excitation takes place which results in absorption of wave at that frequency.

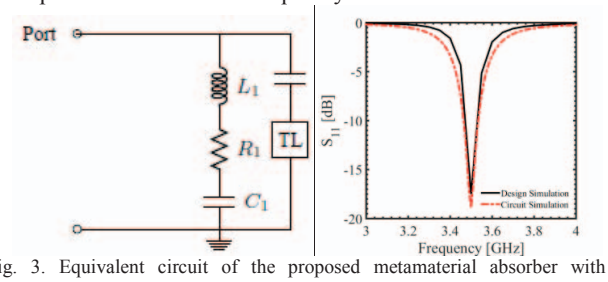


Fig. 3. Equivalent circuit of the proposed metamaterial absorber with reflection coefficient vs frequency response.

To study the electrical behavior of the structure it is necessary to extract the lumped equivalent circuit model of the structure, depicted in fig. III. It can be observed from fig. 2(b) that at 3.5 GHz, the flow of the current is from +y axis to -y axis, which is the direction of applied electric field, on the upper and lower half of the meander line geometry. Therefore, this can be represented as an equivalent inductor (L_1). The capacitor C_1 is the net coupling capacitance between the adjacent unit cells and the two geometries (the meander line and the inner square). The resistance R_1 represents losses in the structure. The *RLC* equivalent of the top geometry is terminated to ground plane through a transmission line equivalent of the dielectric substrate. Finally, the circuit is simulated using Advanced Design System (ADS) and the parameters are optimized using curve fitting tool. The equivalent circuit parameters obtained after optimization are: L_1= 75 *nH*, C_1= 27.8 *fF*, C_c= 250 *fF* and R_1= 57 Ω . Now, the fig. III shows that the circuit is a close approximate model of the proposed metamaterial unit cell.

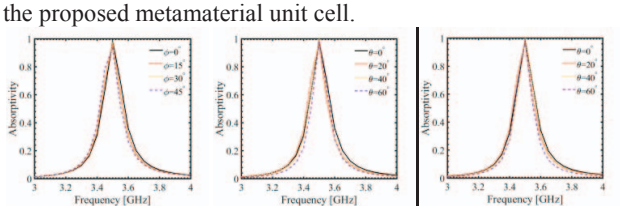


Fig. 4. Simulated absorptivity response of structure under (a) various polarization angles (b) oblique incidence for TE polarization (c) oblique incidence for TM polarization.

Further, the structure is analyzed for different polarization angles(ϕ) (fig. 4(a)) as well as different angle of incidences(θ) of the incident wave for both TE (fig. 4(b)) and TM modes(fig. 4(c)). No significant change in the absorptivity as well as resonant frequency was observed. So the structure is polarization insensitive and offers wide angle stability.

IV. Conclusion

A novel metamaterial absorber based on closed meander line geometry is designed for 3.5 GHz WiMAX band application with an absorptivity of 98.5 %. The structure is analyzed using equivalent circuit model. Also, it was observed that the structure is independent to change in polarization angle and change in angle of incidence of the incident wave. The proposed metamaterial absorber can be used for radar cross section reduction of an antenna.

References

[1] Z.-X. Zhang and J.-C. Zhang, "RCS Reduction for Patch Antenna Based on Metamaterial Absorber," *Prog. Electromagn Res. Symp. (PIERS)*, Shanghai, 2016, pp. 364–368.

[2] P. Munaga, S. Ghosh, S. Bhattacharyya, D. Chaurasiya, and K. V. Srivastava, "An Ultra-thin Dual-Band Polarization-Independent Metamaterial Absorber for EMI/EMC Applications," *9th European Conf. Antenn. Prop. (EuCAP), Lisbon*, 2015, pp. 1–4.

[3] Q. Song and W. Zhang, P. C. Wu, W. Zhu, Z. X. Shen, P. H. J. Chong, Q. X. Liang, Z. C. Yang, Y. L. Hao, H. Cai "Water-Resonator-Based Metasurface: An Ultrabroadband and Near-Unity Absorption," *Adv. Optical Mater.*, 2017.

[4] M. Bakra, M. Karaaslanb, E. Unalb, O. Akgolb, and C. Sabah "Microwave metamaterial absorber for sensing applications," *Opto-Electron. Rev.*, vol. 25, pp. 318–325, 2017.

[5] D. R. Smith, J. B. Pendry, M. C. K. Wiltshire. "Metamaterials and Negative Refractive Index," Science, vol. 305, pp. 788–792, 2004.

[6] N. I. Landy, S. Sajuyigbe, J. J. Mock, D. R Smith, W. J and Padilla. "Perfect metamaterial absorber," *Phys. rev. lett*., vol. 100, 2008.

[7] F. Dincerac, M. Karaaslanbc, E. Unalbc, O. Akgolbc, E. Demireld and C. Sabahe, "Polarization and angle independent perfect metamaterial absorber based on discontinuous cross-wire-strips," *J. Electromagnet. Waves and Applications,* vol. 28, 741–751, 2014.

[8] S. R. Thummaluru, N. Mishra, and R. K. Chaudhary, "Design and Analysis of an Ultrathin X-band polariztion-insensitive metamaterial absorber," *Microw Opt. Techn. Let.*, vol. 58, 2016.

[9] D. R. Smith, D. C. Vier, T. H. Koschny and C. M Soukoulis, "Electromagnetic parameter retrieval from inhomogeneous metamaterials," *Phys.rev. E*, vol. 71, 2005.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:55:55 UTC from IEEE Xplore. Restrictions apply.

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal of Cleaner Production 275 (2020) 124063

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Application of industry 4.0 technologies in SMEs for ethical and sustainable operations: Analysis of challenges

Ravinder Kumar [a], Rajesh Kr. Singh [b, *], Yogesh Kr. Dwivedi [c]

[a] Mechanical Engineering Department, Amity University, Noida, India
[b] Management Development Institute Gurgaon, India
[c] School of Management, Swansea University, Wales, UK

ARTICLE INFO

Article history:
Received 9 March 2020
Received in revised form
21 August 2020
Accepted 31 August 2020
Available online 8 September 2020

Handling editor: Cecilia Maria Villas Bôas de Almeida

Keywords:
Industry 4.0
SMEs
Challenges
DEMATEL
Ethical and sustainable business
Circular economy

ABSTRACT

In the era of Industry 4.0 and circular economy, small and medium enterprises (SMEs) are under huge pressure to make their manufacturing operations ethical and sustainable. Business with ethical and sustainable operations has become the need of the day in the present environment of Industry 4.0 and circular economy. It has been observed that the application of Industry 4.0 technologies may help in achieving the goal of ethical and sustainable operations. Although a lot of research has been done in context to larger enterprises, limited research is available on the application of Industry 4.0 technologies in SMEs for ethical and sustainable operations.

The espousal of Industry 4.0 technologies is a challenging task for SMEs due to various operational and financial constraints. The problem is more acute, specifically in context to developing countries like India. Keeping in mind the role of technologies in ethical business and circular economy, we have identified fifteen challenges, impacting the application of Industry 4.0 technologies in SMEs. A questionnaire was designed for collecting the response from industry and academic experts. On the collected data, the DEMATEL approach has been applied to check the degree of influence and interrelationship among challenges. It has also helped in the categorization of factors as cause and effect. Sensitivity analysis is also performed to validate the results obtained from the DEMATEL approach. Authors have observed that lack of motivation from partners and customers on the application of I4.0 technologies is the leading challenge. Fear of failure of I4.0 technologies is the main effect group challenge. The findings of the study will help SMEs in formulating strategies for implementing Industry 4.0 technologies for ethical and sustainable business processes.

Crown Copyright © 2020 Published by Elsevier Ltd. All rights reserved.

1. Introduction

Under current global scenario, markets all over the world have observed the disruptions in the value chains due to reduction in trade and lack of supplies (Fernandes, 2020). To sustain in such scenario, organizations should restructure their supply chains by sourcing of raw materials and components from different sources including SMEs (Ramelli and Wagner, 2020). SMEs are major contributor to the industrial growth in developing economies all over the globe (Singh and Kumar, 2020; Rauch et al., 2019). Partnership of SMEs with large organizations can enhance opportunities of developing ethical and sustainable operations for them. To become globally competitive for availing emerging opportunities, SMEs need to meet the global standards on quality, technology, sustainability, and pricing (Singh and Kumar, 2020). Many organizations are in transition stage from linear economy to circular economy. Adopting the concept of circular economy (CE) by SMEs may give emerging business opportunities to them (Mura et al., 2020). According to Lieder and Rashid (2016), organizations should adopt the principles of CE for being sustainable in their operations. Technologies may also help organizations in achieving the goals of the circular economy (Prieto-Sandoval et al., 2019; Rosa et al., 2020).

Application of Industry 4.0 technologies may provide a positive direction towards corporate social responsibility and sustainable operations (Kamble et al., 2020; Luthra and Mangla, 2018). According to Battaglia et al. (2018), the experience of the global

* Corresponding author.
E-mail addresses: ravinderkumar.ap@gmail.com (R. Kumar), rajesh.singh@mdi.ac.in (R.Kr. Singh), ykdwivedi@gmail.com (Y.Kr. Dwivedi).

https://doi.org/10.1016/j.jclepro.2020.124063
0959-6526/Crown Copyright © 2020 Published by Elsevier Ltd. All rights reserved.

market helps in acquiring technologies and innovation in processes. SMEs of the manufacturing sector are less reactive to changing business demands of customers in the modern era. They are under pressure due to increasing customization and complexity of products (Dutta et al., 2020). Reduction in time to market, resource optimization, waste minimization, and resource circularity are other few key challenges. The application of technologically advanced sustainable practices imparts a competitive edge to manufacturing organizations of developed economies (Yadav et al., 2020; Mastos et al., 2020). According to Garcia-Muina et al. (2018), Industry 4.0 technologies will help in the transition from linear to CE. Advanced practices of Industry 4.0 can quash costs, meliorate sustainability, and render customizable products to customers (Turner et al., 2019; Machado et al., 2019). Linder (2019) stated that in SMEs, improved communication and information flow can help in achieving efficient processes and cost reduction.

In the cyber-physical environment, machines are made capable of communicating, collecting information, and taking informed decisions by real time data collection through tools like IIoT, AI, big data and clouds (Dutta et al., 2020; Tiwari and Khan, 2020; Lee et al., 2015). Customers' demands and product queries can be collected in real-time by smart technologies. The adoption of intelligent technologies is being predicted as the next industrial revolution (Hofmann and Rusch, 2017). Industry 4.0 technologies will help in effective life cycle management of products in the era of CE (Zhou et al., 2020; Bag et al., 2020c). Therefore, to provide a new innovative environment in the Industry, managers need to adopt or implement the latest technologies such as 3D printing, internet of things (IoT) and cyber-physical system (Almada-Lobo, 2015; Song and Wang, 2018). Rauch et al. (2019) have stated that despite a substantial contribution of SMEs in the economic and employment front, they lack in terms of technology adoption even in developed countries like the European Union and the United States. Effective strategies are required for the application of technologies in small enterprises (Rauch et al., 2019). Many countries like Australia, China, and Thailand are also working on the adoption of Industry 4.0 technologies (Orzes et al., 2020).

1.1. Problem description and research objectives

Ninety-five per cent of Indian manufacturing units with forty per cent value addition fall in small scale category (Singh et al., 2012). SMEs are not able to compete in global markets due to their technological deficiencies and lack of sustainability in operations (Singh and Kumar, 2020). To excel in global competition, manufacturing companies have to upgrade their technologies for ethical and sustainable business operations. The Industry 4.0 techniques are important in the transition from linear to the circular economy (Luthra and Mangla, 2018). Their integration with production systems generate opportunities for sustainable business models aligned with ethical principles of corporate social responsibility (de Sousa Jain et al., 2018 a).

According to Singh and Kumar (2020) global market conditions are very ambitious for Indian SMEs. SMEs should leverage the application of technologies to improve the sustainability of manufacturing operations in the era of the circular economy. In the modern business environment, highly competitive organizations are planning business strategies considering ethical and sustainable business criteria (Longoni and Cagliano, 2015). Mani et al. (2020) have observed that social sustainability affects a firm's performance. Yadav et al. (2020) ascertained that industry 4.0 technologies and circular economy approach could give a competitive edge to supply chains. A focus on the social aspects of sustainability can improve performance and job satisfaction (Sangwan et al., 2019). The efficiency and energy saving of manufacturing processes can be improved by technology adoption (Nascimento et al., 2018). Manufacturing processes can be made efficient and sustainable by effective use of process digitization and quality control tools (Shivajee et al., 2019). Ghobakhloo (2020) has observed that by effectuation of Industry 4.0 technologies, production efficiency, process innovations and sustainability can be improved. To survive and excel in the present business scenario, SMEs need to implement emerging technologies for their sustainable growth (Kumar et al., 2015).

SMEs in developing countries like India are not able to ensure sustainable manufacturing operations due to the high cost of sustainable practices, lack of skills and training, lack of standardized metrics, and lack of adoption of emerging technologies (Kumar, 2020a,b). The incomplete implementation of sustainable and innovative technical processes may impact the performance of SMEs (Shashi et al., 2019). Radziwon et al. (2014) have ascertained that Industry 4.0 technologies can help in improving the sustainability and efficiency of operations. By applying emerging technologies, SMEs can increase productivity, flexibility, responsiveness, and environmental performance (Pedersen et al., 2016). Technologies of Industry 4.0 can be used for resolving sustainability problems (Kumar, 2020a). Kumar et al. (2014) have found that SMEs face challenges in managing their supply chains due to a lack of effective strategies. In recent times, the application of smart technologies has changed the attention of the manufacturing sector drastically (Jain et al., 2017). Effectuation of practices of Industry 4.0 in SMEs faces different issues like security, networking, integration of supply chain, etc. Therefore, to fix such types of problems in SMEs, there is a need to analyze different challenges in the effectuation of emerging technologies (Marques et al., 2017). It is observed that very few studies have holistically examined the challenges coming on the path of new technology adoption by SMEs in developing countries like India (Singh et al., 2019). The majority of the studies are done in context to developed countries and larger enterprises. Therefore, authors are trying to solve the following research questions concerning SMEs of developing countries like India.

RQ1.What are the potential challenges in adopting Industry 4.0 technologies in SMEs for ethical and sustainable operations in context to CE?
RQ2.How these challenges can be prioritized and categorized from a strategy perspective.
RQ3.Which major challenge should be resolved on a priority basis by SMEs for adopting ethical and sustainable business models.

After formulating above research questions and observations from literature, authors feel the need of a study for analyzing challenges in adoption of emerging technologies by SMEs. The remaining part of this study has been organized as follows: Section 2 discourses the literature on the challenges of adopting Industry 4.0 technologies. Section 3 discourses the methodology of the research. Section 4 deals with results and discussion. Section 5 discusses the conclusion with the implications. Section 6 discusses the limitations and future research.

2. Literature review

Review of related literature has been divided into two sections. The first section (2.1) is on the application of Industry 4.0 technologies for ethical and sustainable operations, and the second is on challenges in implementing technologies of Industry 4.0.

2.1. Industry 4.0 technologies applications for ethical and sustainable operations

Industry 4.0 comprises of different technologies like internet of things (IoT), cloud computing, additive manufacturing, cyber security with blockchain, augmented reality with artificial intelligence (AI), big data, system integration, simulation and autonomous robot (Kerin and Pham, 2019; Gurtu and Johny, 2019) (Fig. 1). Techniques of Industry 4.0 have capacities to improve the energy, equipment, and human resource utilization (Lasi et al., 2014). Industry 4.0 is a futuristic construct that nurtures the evolution of autonomous production systems with the application of IoT, CPS, and AI (Pacaux-Lemoine and Trentesaux, 2019). New sensor-based technologies help SMEs in continuously monitoring machine utilization, energy needs, and staff training. By thorough analysis of different Industry 4.0 technologies, data from various IoTs devices can be processed for improving the sustainability of manufacturing operations (Song and Wang, 2017).

For sustainable operations, products need to be manufactured by environment-friendly, socially viable, and economically sound processes. Production systems based on ethical and sustainable manufacturing processes are highly efficient in saving energy and natural resources. Shivajee et al. (2019) have ascertained that manufacturing processes can be made efficient and sustainable by effective use of process digitization and quality control tools. According to Beier et al. (2020), Industry 4.0 is a sociotechnical construct in which technological, social and organizational prospects interacts. Connect of sustainability with Industry 4.0 needs to be studied in depth. For saving energy, reduction of scrap and its impact on the environment, the industrial value chain ought to be oriented towards sustainability (Fatimah et al., 2020). Challenges of ethical and sustainable supply chains can be managed by industry 4.0 and CE concepts (Yadav et al., 2020). According to Garcia-Muina et al. (2018), innovations lead to ethical and sustainable operations when environmental measures are employed across the products life cycle.

Piyathanavong et al. (2019) have observed that knowledge, investment, and training of sustainability concepts are critical requirements for implementing sustainable practices in Thai manufacturing organizations. Efficiency and energy saving of manufacturing processes can be improved by technology adoption (Nascimento et al., 2018). Make to specifications, efficient energy usage, tractability, and closed-loop SCM can be managed by application of Industry 4.0 technologies (Cezarino et al., 2019). Fundamentals of the closed supply chain (circular economy) have introduced a new approach to sustainability. It has become essential for easing reuse and recycle epitome. The closed-loop supply chain adds more sustainability benefits in comparison to the open-loop supply chain (Lieder et al., 2017).

Digital practices can contribute significantly to sustainability by reducing carbon footprints, renewable energy usage, and technology solutions suitable for both individuals and society (Kumar, 2020 b). The evolution of Industry 4.0 helps in the optimal usage of resources in a more transparent manner (Dutta et al., 2020). By effectuation of Industry 4.0 practices, production efficiency and innovation can be improved, which influence the social and environmental sustainability (Ghobakhloo, 2020; Bag et al., 2021). Thakur and Mangla (2019) stated that the professionals of developing economies should emphasize on human, operational, and technological aspects of the sustainable supply chains in the home appliances manufacturing organizations. Government rules and support, awareness of environment protection, and information technologies are the decisive constituents for circular economy implementation (Bhatia et al., 2020). Chauhan et al. (2019) and Cezarino et al. (2019) have observed that with the association of Industry 4.0, manufacturing operations may become sustainable and ethical. De Sousa Jabbour et al. (2018b) also ascertained that ethical and sustainable societal development is possible only by using cleaner production principles. According to Guarnieri and Trojan (2019), in modern times, suppliers are selected by considering the sustainability of operations. There are many benefits contributed individually by different technologies of Industry 4.0. The critical applications of these Industry 4.0 technologies have been summarized in Table 1.

2.2. Challenges in implementing industry 4.0 technologies for ethical and sustainable operations

In the modern business environment of the circular economy, organizations need to use technology for ethical and sustainable value addition. Organizations should not focus only on profit. They need to make a proper balance among different perspectives of performance. To ensure long term growth, adopted business models need to be ethical, sustainable, and transparent without exploiting human values (Machado et al., 2019). Technologies of Industry 4.0 can maintain the whole product life cycle from design to delivery of the product (Hofmann and Rusch, 2017). By IoTs, devices are interconnected via the internet and can share information in the form of commands or data between two or more points (Wollschlaeger et al., 2017). Alqahtani et al. (2019) stated that the IoTs play a critical role in eradicating equivocality about the present and the persisting lives of any product. 3D printing helps in attaining the cheaper and lighter products (DebRoy et al., 2018). System integration is the combination of different software and

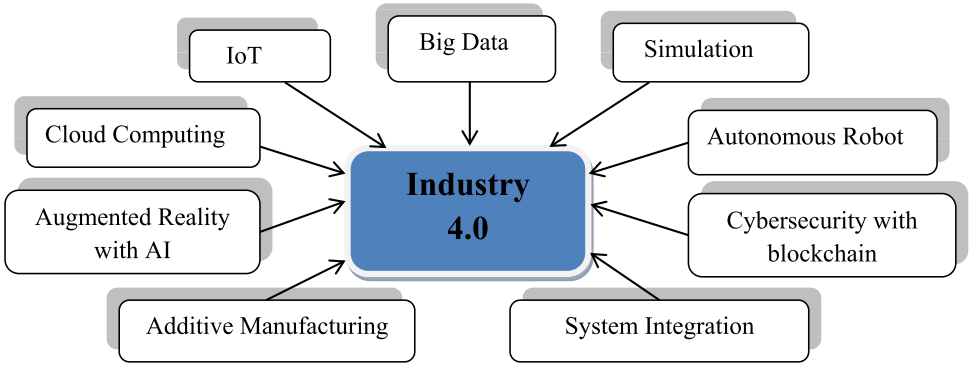


Fig. 1. Technologies of industry 4.0 (source: Kerin and Pham, 2019).

Table 1
Applications of Industry 4.0 technologies for ethical and sustainable operations of SMEs.

Industry 4.0 technologies	Applications of technologies	Reference
IoTs	The convenience of data collection from multiple sources on energy, pollution, efficiency, machine utilization, etc.; cost reduction in manufacturing; technology up-gradation as per ethical and sustainable standards; information on product life cycle can be accessed to promote reuse.	Bhatia et al. (2020); Frank et al. (2019); Kerin and Pham (2019); Thakur and Mangla (2019); Chauhan et al. (2019); Bag and Pretorius (2020)
Cloud computing	All-time anywhere access of available data; transparency and responsiveness of the supply chains; easy sharing of important data; support from supply chain partners in improving ethical and sustainable operations, technology up-gradation, data on product life cycle can be stored & retrieved as per CE philosophy.	Frank et al. (2019); Zhou et al. (2020); Ghobakhloo (2020); Yadav et al. (2020); Cezarino et al. (2019); Pacaux-Lemoine and Trentesaux (2019)
Flexible manufacturing	Reduction in lead time; increased productivity and quality; improved machine utilization; more efficient energy consumption, ethical and sustainable processes.	Yadav et al. (2020)
Additive manufacturing	Negligible scrap generation; environment-friendly process; highly flexible and consistent (intricate designs are easy to produce); testing and prototyping becomes easy, accurate, and affordable for SMEs.	Ford and Despeisse (2016); Bhatia et al. (2020); Thakur and Mangla (2019)
Big Data Analytics	Data gathered from multiple IoTs based devices can be analyzed to get information & trends; data can be used for programming AI devices; machine & human resources utilization can be optimized; traceability of products will improve CE integration & ethical sustainability of operations.	Frank et al. (2019); Horváth and Szabó (2019); Thakur and Mangla (2019); Bag et al. (2020a); Dubey et al. (2020); Pacaux-Lemoine and Trentesaux (2019)
Augmented Reality (AR)	Enables smooth integration of the workforce and resources available within the digital environment; it compounds the real world without substituting it; increased versatility, speed, efficiency, and ethical sustainability of processes; even semi-skilled workforce can work on high-end technologies.	Masood and Egger (2019); Ghobakhloo (2020); Yadav et al. (2020); Guarnieri and Trojan (2019)
Cyber-Physical Systems (CPS)	Machines utilization will improve; data/information on machine tools can be used to improve overall performance; user-friendly human-machine interaction for sustainable operations; autonomous decision-making; perform a task that is difficult for humans	Pater and Gils (2003); Lass and Gronau (2020); Schleinkofer et al. (2019); Alcacer and Cruz-Machlado (2019); Bag et al. (2020d)
Autonomous Robots	Suitability for industrial use in risky surroundings; enhanced versatility, ethical and sustainability of manufacturing operations	Alcacer and Cruz-Machlado (2019); Ghobakhloo (2020); Yadav et al. (2020); Guarnieri and Trojan (2019)

hardware to get easy design modification and maximum value creation (Zhou et al., 2020). By using big data analytics, we can take decisions on future growth and business improvement (Jain et al., 2017). Virtually augmenting (augmented reality) has multiple applications nowadays like gaming, business, and education (Muller et al., 2018a). Cyber security makes a safe and secure system (Muller et al., 2018b). Simulation technology helps in making physical products in the virtual world and save a lot of costs and energy (Pinto Taborga et al., 2018). Robotics and artificial intelligence help in designing and manufacturing products with high accuracy and quality without regular human intervention (Pedersen et al., 2016). IT services of 'Cloud' gives cost-effective data storage (Alsmadi and Prybutok, 2018). Frank et al. (2019) have observed that Industry 4.0 is associated with the espousal of technologies for the entire system.

Few studies support Industry 4.0 implementation and provide a roadmap for its implementation (Zhou et al., 2020). Manufacturing and cyber-physical systems could be integrated by digital learning. Some examples of blended learning could be virtual classrooms, visualization of real data, and adaptive learning (Muller et al., 2018 b). Kumar (2020b) has observed that Indian manufacturing SMEs are more influenced by the technologies of the second industrial revolution and lagging to the fourth industrial revolution. On the other side, Germany's manufacturing sector is more advanced on the technology front (Pfohl et al., 2017). In recent times, the attention on the circular economy has increased, and therefore organizations have started looking for innovative and sustainable technologies (Kumar et al., 2019). As per increasing global competition on the international manufacturing network, there is a need for investigating the challenges of this sector in the Industry 4.0 era (Mishra et al., 2019). Organizations should focus on ethical issues also along with social, economic, and environmental measures while managing their operations for sustainable growth (Guarnieri and Trojan, 2019). Despite many contributions of Industry 4.0 for ethical and sustainable business, many organizations struggle in implementing these technologies in their processes. Fast-changing technological disruptions impose many challenges for SMEs in developing countries (Morrar et al., 2017). SMEs of developing countries face problems due to poor financial condition, lack of technical skills in workers, and the high cost of sustainable practices (Kumar et al., 2014; Moktadir et al., 2018; Bag et al., 2020b; Kumar, 2020a). Apart from it, creating solutions, compatible to environment, culture, and society is also challenging. Lack of awareness about emerging technologies among the workers, the risk of social displacement caused due to unemployment, changes in the market structure due to emerging technologies are other major challenge for sustainable development (Muller et al., 2018b; Satapathy, 2017; Zezulka et al., 2016; Shin et al., 2019; Kumar, 2020b; Cezarino et al., 2019). Lack of global standards and guidelines on the implementation of sustainable technologies are also found challenging for SMEs in developing economies (Shin et al., 2019; Dawson, 2014; Moktadir et al., 2018). The 4th industrial revolution renders a gravid chance to curb these challenges and give a competitive edge in the implementation of sustainable practices (Kumar et al., 2018; Kumar, 2020b).

Machado et al. (2019) have observed that organizations should consider social, environmental, ethical, human rights-related measures while formulating their strategies for sustainable growth. Many developed countries outsource their services to SMEs in developing countries such as India, Bangladesh, Vietnam, etc. to lower the production cost and to avoid restrictive legislation. Transparency of operations is the main challenge for sustainable and ethical services. It becomes more challenging in developing countries to ensure transparency of operations due to lack of IT infrastructure, poor organization culture, lack of legislations and lack of advanced technologies applications (BRICS Business Council,

2017; Luthra and Mangla, 2018; Rauch et al., 2019; Chauhan et al., 2019; Ehrgott et al., 2011). Indian SMEs of manufacturing area lack in awareness about Industry 4.0 contributions to ethical and sustainable production (Dutta et al., 2020).

SMEs are not able to implement technologies efficiently for their sustainable development (Almada-Lobo, 2015). Support of top management is vital for technological changes in SMEs. Senior management should create awareness among employees about the contributions of these technologies for ethical and sustainable operations (Feng et al., 2018; Luthra and Mangla, 2018). According to Dawson (2014), there is a need for knowledge up-gradation on these technologies. SMEs of developing economies faces the problem of lack of funds while implementing Industry 4.0 technologies (Mokhtar et al., 2018; Almada, 2016; Schwab et al., 2019; Dutta et al., 2020; Theorin et al., 2017). Organizations are not aware of government policies on emerging technologies and sustainable business models (BRICS Business Council, 2017; Chauhan et al., 2019). Moktadir et al. (2018) suggested the need for systematic investigation on the application of Industry 4.0 technologies. SMEs' philosophy of short term planning needs to be changed in long term planning for a futuristic approach towards technologies and sustainable operations (Feng et al., 2018; Luthra and Mangla, 2018; Kumar,2020a; Hofmann and Rusch, 2017). SMEs have IT infrastructure related issues (Both software and hardware) in developing economies (Bedekar, 2017; Pfohl et al., 2017; Leitao et al., 2016). SMEs lack in the skilled workforce. There is also need of training about the technologies to upgrade the skills of management and staff (Feng et al., 2018; Sommer, 2015; Luthra and Mangla, 2018; Muller et al., 2018a; Bhatia et al., 2020). From the review of the literature and discussion with Industrial and academic experts, authors have finalized fifteen key challenges affecting the espousal of Industry 4.0 in SMEs for ethical and sustainable value addition (Table 2).

3. Research methodology

Decision Making Trial and Evaluation Laboratory (DEMATEL) technique has been used to develop interrelationship among the challenges and for identifying the most influential challenges. Sensitivity analysis is also performed to check the hardiness of DEMATEL analysis. To showcase the general steps adopted for this study, the authors have designed a research model, as shown in Fig. 2. This model discusses all steps followed by the authors during the study of all concerned issues in this research paper. DEMATEL is a preferred technique over AHP, TISM, ISM, or any MCDM techniques as it divides challenges into cause and effect group and indicates the severity of their effects also (Singh et al., 2019). Policymakers can get observations with a quantifiable and ocular kinship among challenges through matrices or diagraphs (Bai and Satir, 2020). It has a wide range to respond as (0, 1, 2, 3, and 4) to explore the cause-effect relationship among the challenges. The categorization of factors further helps managers in formulating effective strategies to handle them. The DEMATEL is the multiple criteria decision making (MCDM) technique, which helps in developing interrelationship among the challenges or barriers. This tool works as a potent tool for decision making. Rajput and Singh (2019b) have categorized the enablers and challenges of CE and Industry 4.0 using DEMATEL. Yadav and Singh (2020) used Fuzzy-DEMATEL for classifying the blockchain factors in cause and effect groups. Rajput and Singh (2019a) have used DEMATEL for analyzing the enablers of IoTs based system. Kumar and Dixit (2018) applied the DEMATEL approach to develop the framework for analyzing challenges to electronic waste management. Singh et al. (2019) applied the DEMATEL for ICTs use in Indian food SMEs. The detailed procedure for this methodology has been shown in Fig. 2. The overall method of DEMATEL has been divided into four steps. The step by step procedure to apply the DEMATEL approach is as follows:

Step 1**:** Development of Average Direct-relationship Matrix

Experts have given their opinion about the influence of different challenges on each other by analysing the given matrix.

Experts' opinions are collected based on the comparison scale of 0–4. The score is assigned like "0 for no impact, 1 for low impact, 2 for medium impact, 3 for high impact, and 4 for very high impact".

Table 2
Challenges in the espousal of Industry 4.0 technologies.

Code	Challenges	References
Chn1	Lack of awareness about I4.0 contributions to ethical and sustainable production	Almada (2016); Hofmann et al. (2017); Luthra and Mangla (2018); de Sousa Jabbour et al. (2018a), Liao et al. (2018); Dutta et al. (2020); Bhatia et al. (2020); Yadav et al. (2020)
Chn2	Lack of management support for I4.0 technologies	Feng et al. (2018); Luthra and Mangla (2018); Morrar et al. (2017); Turner et al. (2019); Kumar (2020 b)
Chn3	The high initial cost of I4.0 technologies for ethical and sustainable operations	Marques et al. (2017); Dawson (2014); Moktadir et al. (2018); Kumar (2020 b); Bhatia et al. (2020); Yadav et al. (2020)
Chn4	Lack of funds for investment in I4.0 technologies	Moktadir et al. (2018); Chien et al. (2020); Garcia-Muina et al. (2018)
Chn5	Lack of awareness about government policies for I4.0 and sustainability	BRICS Business Council (2017); Luthra and Mangla (2018); Rauch et al. (2019); Chauhan et al. (2019)
Chn6	Lack of dedicated resources for research & development on I4.0 technologies	Mokhtar et al. (2018); Almada (2016), Schwab et al. (2019); Dutta et al. (2020)
Chn7	Lack of long term planning on the adoption of I4.0 technologies for ethical and sustainable operations	Muller et al. (2018a); Feng et al. (2018); Luthra and Mangla (2018); Kumar (2020a)
Chn8	Lack of motivations from customers/OEMs on adopting I4.0 technologies for ethical and sustainable operations	Mokhtar et al. (2018); Marques et al. (2017); Dutta et al. (2020); de Sousa Jabbour et al., (2018b); Shashi et al. (2019)
Chn9	Lack of IT-based infrastructure (Software & Hardware)	Leitao et al.(2016); Bedekar (2017); Pfohl et al. (2017)
Chn10	Lack of trained workforce for sustainable operations & I4.0 technologies	Sommer (2015); Luthra and Mangla (2018); Muller et al. (2018a); Bhatia et al. (2020)
Chn11	Lack of coordination and collaboration among supply chain partners	Wang et al. (2016); Moktadir et al. (2018); Zezulka et al. (2016), Bag et al. (2020b); Kumar (2020a)
Chn12	Fear of unemployment/reduction in workforce	Satapathy (2017); Zezulka et al. (2016); Shin et al. (2019); Kumar (2020b); Cezarino et al. (2019)
Chn13	Fear of failure of I4.0 technologies	Satapathy (2017); Pfohl et al. (2017); Jain et al. (2017); Yadav et al. (2020) and Cezarino et al. (2019)
Chn14	Lack of alternative solutions to the technological breakdown	Dawson (2014); Moktadir et al. (2018); Kumar et al. (2018).
Chn15	Fear of demand uncertainty due to market disruptions	Wang et al. (2016); Theorin et al. (2017); Luthra and Mangla (2018)

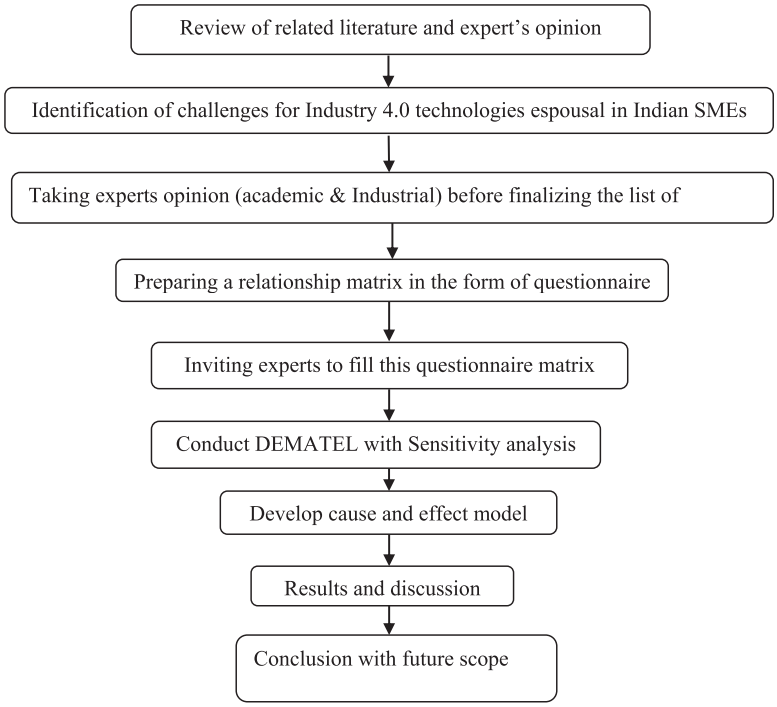


Fig. 2. Research methodology model.

This scale is known as the DEMATEL scale. After collecting the expert's opinion, the average direct-relation matrix is prepared, and it is represented as A_{ij}.

$$Aij = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}$$

Step 2**:** Normalizing the average Direct-Relation Matrix (X)

In this stride, the average Direct-Relation Matrix is further normalized($X = [Xij]n*n$) (Table 4).

Step 3: Development of full direct/Indirect influence Matrix

$$X = z*A \tag{1}$$

Where $z = \min\{1/max1 \le i \le n \sum_{j=1}^{n} a_{ij}; 1/max1 \le i \le n \sum_{i=1}^{n} a_{ij}\}$

$$\lim_{h \to 0} X^h = [0]_{n*n}, \; 0 \le X_{ij} \le 1 \tag{2}$$

Step 4: Obtaining the Total Relation Matrix (T)

$$T = X(I - X)^{-1} \tag{3}$$

Eqn. (3) helps in the formation of the total relation matrix in which I represent the Identity matrix.

Table 3
Average direct − relation matrix.

	Chn1	Chn2	Chn3	Chn4	Chn5	Chn6	Chn7	Chn8	Chn9	Chn10	Chn11	Chn12	Chn13	Chn14	Chn15
Chn1	0	2.67	3	2.34	3	3.67	2.34	3	3	3	1.67	2.67	3	2.67	3
Chn2	2.67	0	2.67	2.67	2.34	3.34	4	2.67	3	3	2	2	2.34	3	1.5
Chn3	3	2.34	0	2.67	2	3.34	3	2	2.34	2.67	3.34	3	2.67	3	1.5
Chn4	2.34	3.34	2	0	2.34	2.67	2	1.67	3.67	3	3	1.67	2	2	3
Chn5	2.34	2	2	2	0	1.67	1.67	1.67	3	2.34	1.34	1.67	2.34	2.34	2
Chn6	2.34	3.67	2.67	3.34	1.67	0	1.67	2	3.34	3	2	1.67	3	3	1.5
Chn7	3.34	2.67	3.67	3.67	2.34	3	0	2.67	2.67	3.34	2.34	2.67	2.67	2.34	1.5
Chn8	3	3.67	2.34	2.34	2.34	3.34	3.34	0	2.34	3	3	3.34	1.67	1.67	2.5
Chn9	4	2.34	2.34	3	2	1.67	2	1.34	0	3.67	2	2	2.67	3	.5
Chn10	3	1.34	3	3	3	2	2.34	1.67	2.34	0	2.34	3	2.67	2.34	1
Chn11	2	1.34	2.67	1.34	1.34	2	2.34	2.34	2.34	2.34	0	1.67	3	2	3.5
Chn12	1.34	2.67	1.67	1.67	1.34	2.34	1.67	1	1.67	2.67	1.34	0	2	1	1.5
Chn13	2	2.34	2.34	2.67	1.67	2.34	2.34	2	2	2.34	1	1.67	0	1.34	1
Chn14	1.67	3.34	2.34	1.67	2	3	2.67	1.67	1.34	2.34	2	1.34	3	0	3
Chn15	1	0	1	1.5	1	1	1.5	3.5	0	1	0	1.5	3	3	0

Step 5. Producing the causal and effect values from Total Relation Matrix

In the Total Relation Matrix, Sum of the all 'i'th row elements represented as *Di* and Sum of the entire '*j*'th row element described as *Rj*. Now *Di* + *Rj* and *Di-Rj* values are obtained, where *Di* + *Rj* indicates the relation of challenges with each other and *Di* - *Rj* suggests the kind of relationship in the form of cause and effect. Where positive values are considered as cause group, and negative values are considered as effect group of challenges.

4. Results obtained and discussion of findings

In this section, authors have discussed the categorization of challenges into cause and effect groups by DEMATEL technique. Sensitivity analysis has been done to validate the consistency of results. Subsections like 4.1 will explore the results of DEMATEL, subsection 4.2 will explore the finding of sensitivity analysis and subsection 4.3 will focus on conclusive discussion on the findings of this study.

4.1. Analysis based on DEMATEL approach

An expert team was formed to take the inputs for analyzing the influence of different factors over each other. Each industrial and academic expert has minimum experience ten years in the relevant field of advanced manufacturing systems and research, respectively. All industrial and academic experts were from organizations located in the National capital region (NCR), Delhi-India. A detailed questionnaire was designed for the survey (Appendix-I). Industrial experts were of manager-level position with professional qualification and had worked in the concerned organization for at least five years. Academic experts were of associate professor level and above with Ph.D. as the minimum qualification. Experts were briefed about the purpose of the survey, standard definition, and title of the study. For collecting the response, the questionnaire was circulated among 80 experts (60 from Industry & 20 from academia). The responders were contacted by different sources of correspondence. A total of 36 experts (21 from Industry & 15 from academia) replied by filling the complete questionnaire. It represents a response rate of 45 per cent.

For the calculation of the DEMATEL approach, first, the average Direct-Relation Matrix A is developed, as shown in Table 3. After this, the average matrix is normalized to express the values in the range of 0–1 (Table 4). Then the total relation matrix is obtained (Table 5). Further calculations are performed as per steps of methodology, and the results are shown in Table 6 and Table 7, respectively. The obtained values of 'Di', 'Rj', 'Di + Rj', and 'Di −Rj' are shown in Table 6. These values will help in analyzing the challenges. The magnitude of 'Di − Rj' values categorizes the challenges into 'Effect' and 'Cause' groups. The positive values of 'Di − Rj' are considered into cause group, and negative values of 'Di − Rj' are considered into effect group.

In this study, it has been observed that lack of motivations from customers/OEMs on adopting I4.0 technologies for ethical and sustainable operations (Chn8) has the highest positive value. Therefore, it is considered as the most critical challenge in the cause category. The exact values for cause and effect challenges are shown in Table 7. Further, it is observed that lack of long term planning on the adoption of I4.0 technologies for ethical and sustainable operations (Chn7), lack of awareness about I4.0 contributions to ethical and sustainable production (Chn1), lack of management support for I4.0 technologies (Chn2) and High initial cost of I4.0 technologies for ethical and sustainable operations (Chn3); are the second, third, fourth and fifth-ranked challenges respectively in the cause category. Further findings of the study also

Table 4
Normalize Direct Relation matrix.

	Chn1	Chn2	Chn3	Chn4	Chn5	Chn6	Chn7	Chn8	Chn9	Chn10	Chn11	Chn12	Chn13	Chn14	Chn15
Chn1	0	0.068409	0.076864	0.059954	0.076864	0.09403	0.059954	0.076864	0.076864	0.076864	0.042788	0.068409	0.076864	0.068409	0.076864
Chn2	0.068409	0	0.068409	0.068409	0.059954	0.085575	0.102485	0.068409	0.076864	0.076864	0.051243	0.051243	0.059954	0.076864	0.038432
Chn3	0.076864	0.059954	0	0.068409	0.051243	0.085575	0.076864	0.051243	0.059954	0.068409	0.085575	0.076864	0.068409	0.076864	0.038432
Chn4	0.059954	0.085575	0.051243	0	0.059954	0.068409	0.051243	0.042788	0.09403	0.076864	0.076864	0.042788	0.051243	0.051243	0.076864
Chn5	0.059954	0.051243	0.051243	0.051243	0	0.042788	0.042788	0.042788	0.076864	0.059954	0.034333	0.042788	0.059954	0.059954	0.051243
Chn6	0.059954	0.09403	0.068409	0.085575	0.042788	0	0.042788	0.051243	0.085575	0.076864	0.051243	0.042788	0.076864	0.076864	0.038432
Chn7	0.085575	0.068409	0.09403	0.09403	0.059954	0.076864	0	0.068409	0.068409	0.085575	0.059954	0.068409	0.068409	0.059954	0.038432
Chn8	0.076864	0.09403	0.059954	0.059954	0.059954	0.085575	0.085575	0	0.059954	0.076864	0.076864	0.085575	0.042788	0.042788	0.064053
Chn9	0.102485	0.059954	0.059954	0.076864	0.051243	0.042788	0.051243	0.034333	0	0.09403	0.051243	0.051243	0.068409	0.076864	0.012811
Chn10	0.076864	0.034333	0.076864	0.076864	0.076864	0.051243	0.059954	0.042788	0.059954	0	0.059954	0.076864	0.068409	0.059954	0.025621
Chn11	0.051243	0.034333	0.068409	0.034333	0.034333	0.051243	0.059954	0.059954	0.059954	0.059954	0	0.042788	0.076864	0.051243	0.089675
Chn12	0.034333	0.068409	0.042788	0.042788	0.034333	0.059954	0.042788	0.025621	0.042788	0.068409	0.034333	0	0.051243	0.025621	0.038432
Chn13	0.051243	0.059954	0.059954	0.068409	0.042788	0.059954	0.059954	0.051243	0.051243	0.059954	0.025621	0.042788	0	0.034333	0.025621
Chn14	0.042788	0.085575	0.059954	0.042788	0.051243	0.076864	0.068409	0.042788	0.034333	0.059954	0.051243	0.034333	0.076864	0	0.076864
Chn15	0.025621	0	0.025621	0.038432	0.025621	0.025621	0.038432	0.089675	0	0.025621	0	0.038432	0.076864	0.076864	0

Table 5
Total relation matrix.

	Chn1	Chn2	Chn3	Chn4	Chn5	Chn6	Chn7	Chn8	Chn9	Chn10	Chn11	Chn12	Chn13	Chn14	Chn15	D
Chn1	0.353	0.415	0.421	0.41	0.369	0.452	0.396	0.371	0.416	0.459	0.328	0.374	0.44	0.402	0.346	5.952
Chn2	0.412	0.345	0.409	0.412	0.349	0.438	0.427	0.357	0.411	0.453	0.332	0.353	0.418	0.402	0.307	5.825
Chn3	0.409	0.393	0.336	0.402	0.334	0.43	0.396	0.335	0.387	0.436	0.355	0.368	0.417	0.394	0.302	5.694
Chn4	0.373	0.39	0.362	0.316	0.323	0.389	0.352	0.31	0.395	0.418	0.328	0.318	0.379	0.352	0.318	5.323
Chn5	0.319	0.308	0.308	0.311	0.221	0.312	0.293	0.263	0.327	0.344	0.246	0.271	0.33	0.307	0.252	4.412
Chn6	0.381	0.408	0.384	0.403	0.315	0.335	0.353	0.322	0.397	0.427	0.313	0.324	0.408	0.381	0.289	5.44
Chn7	0.439	0.421	0.443	0.447	0.361	0.444	0.345	0.368	0.416	0.475	0.351	0.38	0.438	0.399	0.318	6.045
Chn8	0.418	0.43	0.401	0.404	0.349	0.438	0.413	0.295	0.396	0.453	0.354	0.385	0.403	0.373	0.331	5.843
Chn9	0.4	0.36	0.36	0.377	0.308	0.358	0.342	0.291	0.299	0.422	0.298	0.317	0.382	0.362	0.253	5.129
Chn10	0.375	0.336	0.373	0.375	0.328	0.363	0.347	0.298	0.354	0.334	0.305	0.338	0.381	0.345	0.263	5.115
Chn11	0.322	0.303	0.335	0.307	0.263	0.331	0.319	0.291	0.321	0.356	0.222	0.282	0.358	0.311	0.297	4.618
Chn12	0.257	0.285	0.263	0.266	0.222	0.287	0.256	0.215	0.26	0.31	0.215	0.196	0.282	0.239	0.209	3.762
Chn13	0.307	0.313	0.312	0.323	0.259	0.323	0.304	0.266	0.301	0.34	0.236	0.267	0.267	0.279	0.225	4.322
Chn14	0.328	0.363	0.341	0.329	0.291	0.369	0.341	0.287	0.313	0.371	0.282	0.285	0.372	0.276	0.295	4.843
Chn15	0.2	0.18	0.18	0.213	0.174	0.21	0.207	0.234	0.172	0.218	0.146	0.193	0.255	0.238	0.139	2.959
R	5.293	5.25	5.228	5.295	4.466	5.479	5.091	4.503	5.165	5.816	4.311	4.651	5.53	5.06	4.144	

Table 6
The sum of influence given and received on challenges.

Challenges	Di	Rj	Di-Rj	Di + Rj	Overall Ranking
Chn1	5.952	5.293	0.659	11.245	1
Chn2	5.825	5.25	0.575	11.075	3
Chn3	5.694	5.228	0.466	10.922	5
Chn4	5.323	5.295	0.028	10.618	7
Chn5	4.412	4.466	−0.054	8.878	13
Chn6	5.44	5.479	−0.039	10.919	6
Chn7	6.045	5.091	0.954	11.136	2
Chn8	5.843	4.503	1.34	10.346	8
Chn9	5.129	5.165	−0.036	10.294	9
Chn10	5.115	5.816	−0.701	10.931	4
Chn11	4.618	4.311	0.307	8.929	12
Chn12	3.762	4.651	−0.889	8.413	14
Chn13	4.322	5.53	−1.208	9.852	11
Chn14	4.843	5.06	−0.217	9.903	10
Chn15	2.959	4.144	−1.185	7.103	15

indicate that fear of failure of I4.0 technologies (Chn13), fear of demand uncertainty due to market disruptions (Chn15), fear of unemployment/reduction in workforce (Chn12), lack of trained workforce on sustainable operations & I4.0 technologies (Chn10) and lack of alternative solutions to the technological breakdown (Chn14) are top five effect category challenges. The observations from this study will help in developing the framework for analyzing the Industry 4.0 challenges and finding the interrelatedness among these challenges. Authors in current study have also prepared a causal diagram to represent the graphical relationship between the challenges for this study (Fig. 3).

4.2. Sensitivity analysis

Sensitivity analysis can be used to check the hardiness of the optimal solution of any system or model so that it could be applied for effective decision making (Pannell, 1997). We can use this analysis in two ways, firstly by varying the weight assigned to every challenge and secondly by changing the weight assigned to every expert. Xia et al. (2015) applied sensitivity analysis by changing the weight assigned to every expert to check the robustness of the solution for analyzing the cause and effect relationship. This analysis helps to check biasedness for a specific expert, which may affect the outcome of this study. In this study, the authors used this approach by varying the weight of every expert. In Case-1, every expert was assigned with equal weight, and in Case-2, Case-3, and Case-4, higher weight was assigned to one expert, and the other two expert's weights remained the same (Table 8). Then, for different possible scenarios, calculations are performed for sensitivity analysis. After that, in each case, the cause-effect relationship and causal diagrams are generated (Fig. 3-Cases 1–4). The results found from this sensitivity analysis indicate the consistency of findings (Table 9). No considerable change in the results under different conditions confirms the robustness of the model.

4.3. Discussions of findings

Results of study discussed in section 4.1 and section 4.2 indicate that for SMEs, motivation from customers and original equipment manufacturers (Partners) and top management support are very important for adoption of smart technologies. SMEs in developing countries are mostly dependent on their partners for economical and other business advancements. Therefore, there is a need of effective coordination among different members of supply chains. Luthra and Mangla (2018) in their study have found challenges related to finance, government policies, and management supports while implementing Industry 4.0 technologies. Chakraborty et al. (2020) in their study on logistics sector have observed that lack

Table 7
Ranking of challenges subdivided into cause and effect group.

Cause Group		Effect Group	
Challenges	Ranking	Challenges	Rank
Chn8	1	**Chn13**	1
Chn7	2	**Chn15**	2
Chn1	3	**Chn12**	3
Chn2	4	**Chn10**	4
Chn3	5	**Chn14**	5
Chn11	6	**Chn5**	6
Chn4	7	**Chn6**	7
		Chn9	8

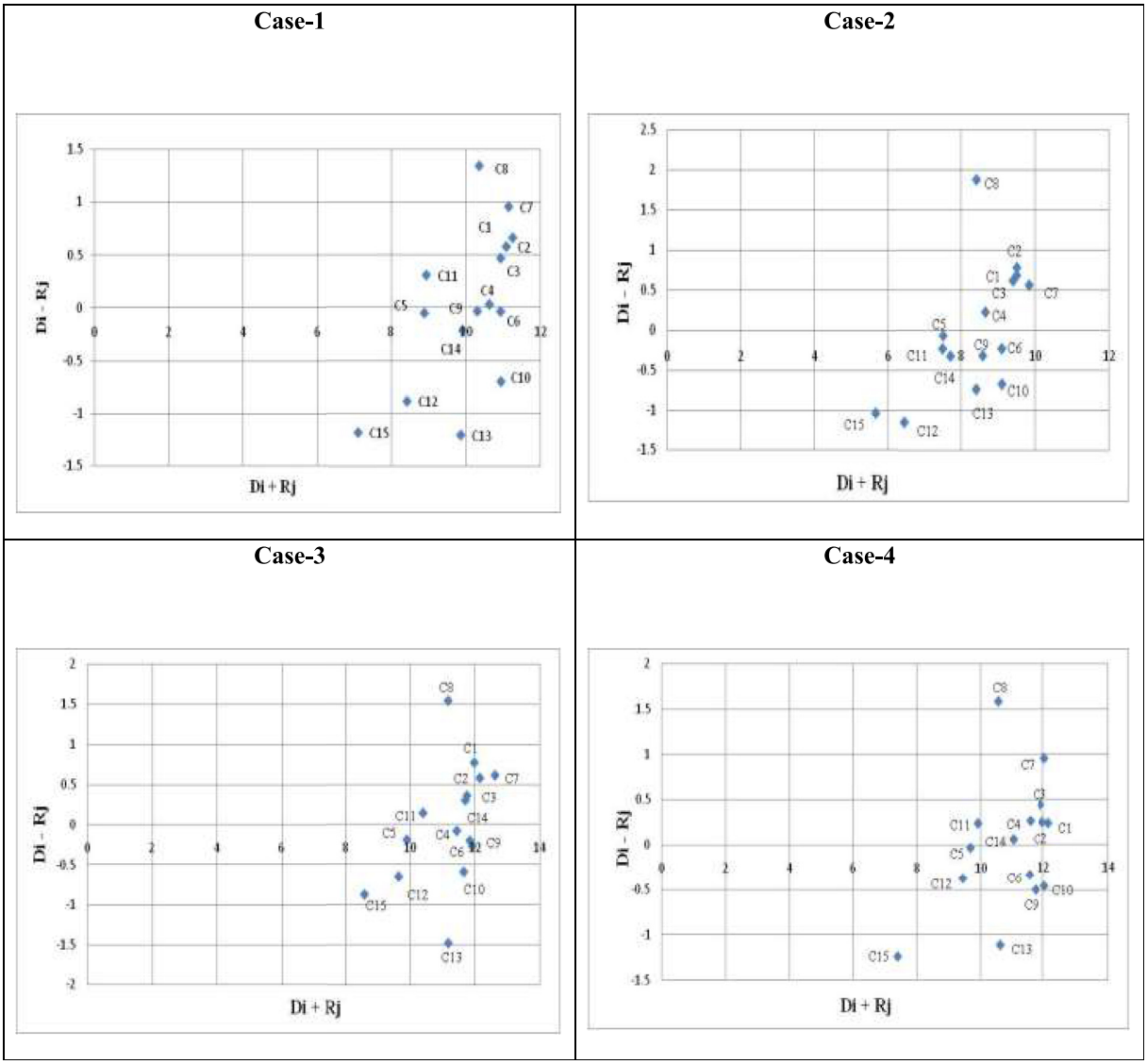


Fig. 3. The Cause-effect diagram (Case 1 to 4).

Table 8
Varying the weight of experts in the sensitivity analysis.

	Case-1	Case-2	Case-3	Case-4
Expert 1	0.33	0.50	0.25	0.25
Expert 2	0.33	0.25	0.50	0.25
Expert 3	0.33	0.25	0.25	0.50

of awareness about technologies is a significant concern for implementing new technologies. Liao et al. (2018) found that organizations of developed countries have more knowledge of advanced technologies in comparison to developing countries. Findings of the current study have highlighted that in developing economy, SMEs face major challenges of awareness, knowledge and funding while adopting Industry 4.0 technologies for ethical and sustainable operations.

Frank et al. (2019) have mentioned about the non-availability of hardware/software while implementation these technologies. Rajput and Singh (2019a) have found out that artificial intelligence (AI), service, and policy framework can enable the implementation of Industry 4.0 technologies for ethical and sustainable operations. The findings of the current study imply that the challenges such as lack of IT based infrastructure, lack of trained workforce, fear of failure of I4.0 technologies, lack of alternative solutions at the time of breakdowns are highly significant for SMEs, and should be given more priority while formulating strategies. Raj et al. (2019) have found lack of strategy for process digitalization and scarcity of resources as major challenges for implementing new technologies. Kamble (2018); Bogoviz et al. (2019) and Chakraborty et al. (2020), have highlighted the importance of regulations, rules, allegiance and willingness for the digitalization of processes in Indian manufacturing organizations. We have also sensed the inadequacy of legislations and commitment on Industry 4.0 technologies in present study on Indian SMEs.

Study of Piyathanavong et al. (2019), in Thai manufacturing organizations have highlighted the issues of investment, training, and knowledge of technologies for implementing sustainability practices. Chakraborty et al. (2020) have discussed in their study about the issues of poor infrastructure, scarcity of finance, poor information sharing, and the inadequacy of expertness of Indian logistics sector in implementing IT tools. For validation of findings,

Table 9
Sensitivity analysis of results.

Challenge Code	Case-1		Case-2		Case-3		Case-4	
	D + R	Ranking	D + R	Ranking	D + R	Ranking	D + R	Ranking
Chn1	11.245	1	9.494	3	11.974	3	12.151	1
Chn2	11.075	3	9.501	2	12.145	2	11.968	3
Chn3	10.922	5	9.395	4	11.746	6	11.901	4
Chn4	10.618	7	8.65	7	11.431	9	11.552	6
Chn5	8.878	13	7.506	12	9.887	13	9.661	12
Chn6	10.919	6	9.092	6	11.915	4	11.524	7
Chn7	11.136	2	9.834	1	12.612	1	12.029	2
Chn8	10.346	8	8.406	9	11.161	11	10.538	10
Chn9	10.294	9	8.576	8	11.836	5	11.716	5
Chn10	10.931	4	9.096	5	11.646	8	12.029	2
Chn11	8.929	12	7.492	13	10.388	12	9.902	11
Chn12	8.413	14	6.456	14	9.635	14	9.423	13
Chn13	9.852	11	8.401	10	11.169	10	10.599	9
Chn14	9.903	10	7.704	11	11.696	7	11.019	8
Chn15	7.103	15	5.649	15	8.579	15	7.386	14

sensitivity analysis needs to be done (Pannell, 1997; Xia et al., 2015). In this study, after doing sensitivity analysis, we have observed that there is no considerable difference in the results under different conditions. It indicates about the robustness of our model. Observations from causal diagram, indicates that lack of motivations from customers/OEMs on adopting I4.0 technologies for ethical and sustainable operations and fear of failure of I4.0 technologies are the key challenges in the espousal of Industry 4.0 technologies and these are the top-ranked challenges in their respective group. Therefore, there is a need for overcoming the fear of failure in the adoption of Industry 4.0 technologies.

5. Conclusion and implications

Across the globe, specifically in developing countries, SMEs are considered the backbone of the economy. In developing countries like India, out of a total of 16% contribution of the manufacturing sector to India's GDP, approximately 8% comes from SMEs (Singh and Kumar, 2020). SMEs are also a significant source of employment in India. To compete in global markets, SMEs need to adopt ethical, efficient, and sustainable business models. Many researchers have observed the importance of Industry 4.0 technologies for ethical and sustainable operations in context to larger enterprises. However, studies in context to SMEs are limited. Therefore, authors wanted to analyze the challenges in the implementation of Industry 4.0 technologies in SMEs for ethical and sustainable operations. Fifteen critical challenges after the review of literature have been identified. For further ranking and categorization of these challenges, the DEMATEL approach is used. Lack of motivation in customers/OEMs on adopting I4.0 technologies for ethical and sustainable operations is found as the most crucial challenge in the cause category. Lack of long term planning on adoption of I4.0 technologies for ethical and sustainable operations, lack of awareness about I4.0 contributions to ethical and sustainable production, lack of management support for I4.0 technologies and the high initial cost of I4.0 technologies for ethical and sustainable operations is other significant challenge under cause category. Under the effect category, fear of failure of I4.0 technologies and fear of demand uncertainty due to market disruptions are significant challenges in the implementation of Industry 4.0 technologies for ethical and sustainable operations in SMEs. Challenges falling in the cause category may influence other challenges, so management should prioritize the strategies accordingly.

5.1. Managerial implications

Findings of this study imply that SMEs need to be motivated to adopt ethical and sustainable business models. SMEs should create awareness about these technologies and their contributions within their organizations and for other stakeholders. Management of SMEs should allocate sufficient funds for such initiatives considering it as long term goals. Usually, SMEs take their decisions based on short term gains. It means overall culture within SMEs needs to be changed. SMEs should integrate these Industry 4.0 technologies with different manufacturing processes for ethical and sustainable operations. As SMEs lack in terms of knowledge and expertise (Kumar, 2020 b), so consultants should be engaged for effective strategy formulation. Integrating digital technologies with manufacturing processes for ethical and sustainable value addition should be part of the strategies. It will not only help them in resource optimization but will also make them responsive to changing market requirements. Telukdarie et al. (2018) have considered the applications of Industry 4.0 technologies for sustainable operations in planning of long term strategies.

5.2. Theoretical implications

Findings of the current study have many theoretical implications too. Government and larger supply chain partners should focus on the need of awareness programs for SMEs in adopting Industry 4.0 technologies. Bogoviz et al. (2019) have stressed on the need for policies and regulations on the implementation of Industry 4.0 in developing economies. To excel on the global front, SMEs need a plan for future and adopt technologies for ethical and sustainable operations. Initiatives taken on time in this direction will make them more resilient and competitive. Researchers and academic experts all over the globe should focus more on case studies and empirical researches for future directions in this context. Multiple researches and observations will help in reducing the challenges of technologies adoption in SMEs of both developing and developed economies.

6. Limitations and future research directions

Despite significant contribution in motivating SMEs for ethical and sustainable business by implementing Industry 4.0 technologies, this study has also got few limitations. The authors have selected fifteen challenges for adopting Industry 4.0 technologies

for ethical and sustainable operations in SMEs from the Indian perspective. SMEs from other sectors and developed countries may face different kind of challenges. Therefore, findings cannot be generalized. DEMATEL methodology has also got its limitations due to the biasedness of experts' opinions while taking inputs. Therefore, the findings of the study cannot be generalized. For further validation, the study can be extended in context to the other countries and sectors by applying other MCDM tools and case studies. It will help in doing the comparative analysis and in generalizing the findings of this study. Sensing the importance of automation and Industry4.0 technologies in current scenario, findings of this study could be very helpful for SMEs all over the globe. Influence of pandemic (COVID-19) has affected the functioning of almost all supply chains. Lockdown, isolation, social distancing and migration of labors will lead the companies to rethink about their location and investment strategies. In the future also for resuming production in new normal, findings of this study will be very helpful for SMEs in developing ethical and sustainable operations.

CRediT authorship contribution statement

Ravinder Kumar: It is certified that all the three authors have contributed in this research work, Introduction and LR have been done by Ravinder Kr, Analysis of results by RAJESH Kr Singh and Discussion with conclusion by Yogesh Dwivedi. **Rajesh Kr. Singh:** It is certified that all the three authors have contributed in this research work, Introduction and LR have been done by Ravinder Kr, Analysis of results by RAJESH Kr Singh and Discussion with conclusion by Yogesh Dwivedi. **Yogesh Kr. Dwivedi:** It is certified that all the three authors have contributed in this research work, Introduction and LR have been done by Ravinder Kr, Analysis of results by RAJESH Kr Singh and Discussion with conclusion by Yogesh Dwivedi.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Authors would like to thank reviewers and Editor of the journal for valuable inputs and suggestions for improvements in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclepro.2020.124063.

References

Alcácer, V., Cruz-Machado, V., 2019. Scanning the industry 4.0: a literature review on technologies for manufacturing systems. Eng. Sci. Technol. Int. J. 22, 899–919. https://doi.org/10.1016/j.jestch.2019.01.006.

Almada-Lobo, F., 2016. The industry 4.0 revolution and the future of manufacturing execution systems (MES). J. Innovat. Manag. 3, 16–21. https://doi.org/10.24840/2183-0606_003.004_0003.

Alqahtani, A.Y., Gupta, S.M., Nakashima, K., 2019. Warranty and maintenance analysis of sensor embedded products using internet of things in industry 4.0. Int. J. Prod. Econ. 208, 483–499. https://doi.org/10.1016/j.ijpe.2018.12.022.

Alsmadi, D., Prybutok, V., 2018. Sharing and storage behavior via cloud computing: security and privacy in research and practice. Comput. Hum. Behav. 85, 218–226. https://doi.org/10.1016/j.chb.2018.04.003.

Bag, S., Gupta, S., Kumar, S., 2021. Industry 4.0 adoption and 10R advance manufacturing capabilities for sustainable development. Int. J. Prod. Econ. 231, 107844. https://doi.org/10.1016/j.ijpe.2020.107844.

Bag, S., Gupta, S., Luo, Z., 2020d. Examining the role of logistics 4.0 enabled dynamic capabilities on firm performance. Int. J. Logist. Manag. https://doi.org/10.1108/IJLM-11-2019-0311 ahead-of-p.

Bag, S., Pretorius, J.H.C., 2020. Relationships between industry 4.0, sustainable manufacturing and circular economy: proposal of a research framework. Int. J. Organ. Anal. https://doi.org/10.1108/IJOA-04-2020-2120 ahead-of-p.

Bag, S., Wood, L.C., Mangla, S.K., Luthra, S., 2020b. Procurement 4.0 and its implications on business process performance in a circular economy. Resour. Conserv. Recycl. 152, 104502. https://doi.org/10.1016/j.resconrec.2019.104502.

Bag, S., Wood, L.C., Xu, L., Dhamija, P., Kayikci, Y., 2020a. Big data analytics as an operational excellence approach to enhance sustainable supply chain performance. Resour. Conserv. Recycl. 153, 104559. https://doi.org/10.1016/j.resconrec.2019.104559.

Bag, S., Yadav, G., Wood, L.C., Dhamija, P., Joshi, S., 2020c. Industry 4.0 and the circular economy: resource melioration in logistics. Resour. Pol. 68, 101776. https://doi.org/10.1016/j.resourpol.2020.101776.

Bai, C., Satir, A., 2020. Barriers for green supplier development programs in manufacturing industry. Resour. Conserv. Recycl. 158, 104756. https://doi.org/10.1016/j.resconrec.2020.104756.

Battaglia, D., Neirotti, P., Paolucci, E., 2018. The role of R&D investments and export on SMEs' growth: a domain ambidexterity perspective. Manag. Decis. 56, 1883–1903. https://doi.org/10.1108/MD-02-2017-0136.

Bedekar, A., 2017. Opportunities & challenges for IoT in India. Online available at: http://www.startupcity.com/leaders-talk/-opportunities-challenges-for-IoT-in-india-nid-3444.html.

Beier, G., Ullrich, A., Niehoff, S., Reißig, M., Habich, M., 2020. Industry 4.0: how it is defined from a sociotechnical perspective and how much sustainability it includes – a literature review. J. Clean. Prod. 259, 120856. https://doi.org/10.1016/j.jclepro.2020.120856.

Bhatia, M.S., Jakhar, S.K., Mangla, S.K., Gangwani, K.K., 2020. Critical factors to environment management in a closed loop supply chain. J. Clean. Prod. 255, 120239. https://doi.org/10.1016/j.jclepro.2020.120239.

Bogoviz, A.V., Osipov, V.S., Chistyakova, M.K., Borisov, M.Y., 2019. Comparative Analysis of Formation of Industry 4.0 in Developed and Developing Countries, pp. 155–164. https://doi.org/10.1007/978-3-319-94310-7_15.

BRIC Business Council, 2017. Skill development for industry 4.0. A white paper by BRICS skill development working group, BRICS business Council, India group. Online available at: http://www.globalskillsummit.com/Whitepaper-Summary.pdf.

Cezarino, L.O., Liboni, L.B., Oliveira Stefanelli, N., Oliveira, B.G., Stocco, L.C., 2019. Diving into emerging economies bottleneck: industry 4.0 and implications for circular economy. Manag. Decis. https://doi.org/10.1108/MD-10-2018-1084. MD-10-2018-1084.

Chakraborty, S., Sharma, A., Vaidya, O.S., 2020. Achieving sustainable operational excellence through IT implementation in Indian logistics sector: an analysis of barriers. Resour. Conserv. Recycl. 152, 104506. https://doi.org/10.1016/j.resconrec.2019.104506.

Chauhan, C., Sharma, A., Singh, A., 2019. A SAP-LAP linkages framework for integrating Industry 4.0 and circular economy. Benchmark Int. J. https://doi.org/10.1108/BIJ-10-2018-0310 ahead-of-p.

Chien, C.F., Tseng, M.L., Tan, R.R., Tan, K., Velek, O., 2020. Industry 3.5 for sustainable transition and total resource management, 2020. Resour. Conserv. Recycl. 152, 104482. https://doi.org/10.1016/j.resconrec.2019.104482.

Dawson, T., 2014. Industry 4.0 – Opportunities and Challenges for Smart Manufacturing. http://blog.ihs.com/q13-industry-40-opportunities-and-challenges-for-smart-manufacturing.

de Sousa Jabbour, A.B., Jabbour, C.J.C., Godinho Filho, M., Roubaud, D., 2018b. Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations. Ann. Oper. Res. 270, 273–286. https://doi.org/10.1007/s10479-018-2772-8.

de Sousa Jabbour, A.B.L., Jabbour, C.J.C., Foropon, C., Godinho Filho, M., 2018a. When titans meet – can industry 4.0 revolutionise the environmentally-sustainable manufacturing wave? The role of critical success factors. Technol. Forecast. Soc. Change 132, 18–25. https://doi.org/10.1016/j.techfore.2018.01.017.

DebRoy, T., Wei, H.L., Zuback, J.S., Mukherjee, T., Elmer, J.W., Milewski, J.O., Beese, A.M., Wilson-Heid, A., De, A., Zhang, W., 2018. Additive manufacturing of metallic components – Process, structure and properties. Prog. Mater. Sci. 92, 112–224. https://doi.org/10.1016/j.pmatsci.2017.10.001.

Dubey, R., Gunasekaran, A., Childe, S.J., Papadopoulos, T., Luo, Z., Wamba, S.F., Roubaud, D., 2019. Can big data and predictive analytics improve social and environmental sustainability? Technol. Forecast. Soc. Change 144, 534–545. https://doi.org/10.1016/j.techfore.2017.06.020.

Dutta, G., Kumar, R., Sindhwani, R., Singh, R.K., 2020. Digital transformation priorities of India's discrete manufacturing SMEs – a conceptual study in perspective of Industry 4.0. Compet. Rev. An Int. Bus. J. 30, 289–314. https://doi.org/10.1108/CR-03-2019-0031.

Ehrgott, M., Reimann, F., Kaufmann, L., Carter, C.R., 2011. Social sustainability in selecting emerging economy suppliers. J. Bus. Ethics 98 (1), 99–119. https://doi.org/10.1007/sl0551-010-0537-7.

Fatimah, Y.A., Govindan, K., Murniningsih, R., Setiawan, A., 2020. Industry 4.0 based sustainable circular economy approach for smart waste management system to achieve sustainable development goals: a case study of Indonesia. J. Clean. Prod. 269, 122263. https://doi.org/10.1016/j.jclepro.2020.122263.

Feng, L., Zhang, X., Zhou, K., 2018. Current problems in China's manufacturing and countermeasures for industry 4.0. EURASIP J. Wirel. Commun. Netw. 90 https://

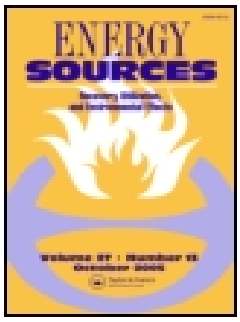
doi.org/10.1186/s13638-018-1113-6, 2018.
Fernandes, N., 2020. Economic effects of coronavirus outbreak (COVID-19) on the world economy. SSRN Working paper, available at: https://ssrn.com/abstract=3557504.
Ford, S., Despeisse, M., 2016. Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. J. Clean. Prod. 137, 1573–1587. https://doi.org/10.1016/j.jclepro.2016.04.150.
Frank, A.G., Dalenogare, L.S., Ayala, N.F., 2019. Industry 4.0 technologies: implementation patterns in manufacturing companies. Int. J. Prod. Econ. 210, 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004.
Garcia-Muiña, F., González-Sánchez, R., Ferrari, A., Settembre-Blundo, D., 2018. The paradigms of industry 4.0 and circular economy as enabling drivers for the competitiveness of businesses and territories: the case of an Italian ceramic tiles manufacturing company. Soc. Sci. 7, 255. https://doi.org/10.3390/socsci7120255.
Ghobakhloo, M., 2020. Industry 4.0, digitization, and opportunities for sustainability. J. Clean. Prod. 252, 119869. https://doi.org/10.1016/j.jclepro.2019.119869.
Guarnieri, P., Trojan, F., 2019. Decision making on supplier selection based on social, ethical, and environmental criteria: a study in the textile industry. Resour. Conserv. Recycl. 141, 347–361. https://doi.org/10.1016/j.resconrec.2018.10.023.
Gurtu, A., Johny, J., 2019. Potential of blockchain technology in supply chain management: a literature review. Int. J. Phys. Distrib. Logist. Manag. 49, 881–900. https://doi.org/10.1108/IJPDLM-11-2018-0371.
Hofmann, E., Rüsch, M., 2017. Industry 4.0 and the current status as well as future prospects on logistics. Comput. Ind. 89, 23–34. https://doi.org/10.1016/j.compind.2017.04.002.
Horváth, D., Szabó, R.Z., 2019. Driving forces and barriers of Industry 4.0: do multinational and small and medium-sized companies have equal opportunities? Technol. Forecast. Soc. Change 146, 119–132. https://doi.org/10.1016/j.techfore.2019.05.021.
Jain, S., Shao, G., Shin, S.-J., 2017. Manufacturing data analytics using a virtual factory representation. Int. J. Prod. Res. 55, 5450–5464. https://doi.org/10.1080/00207543.2017.1321799.
Kamble, S., Gunasekaran, A., Dhone, N.C., 2020. Industry 4.0 and lean manufacturing practices for sustainable organisational performance in Indian manufacturing companies. Int. J. Prod. Res. 58, 1319–1337. https://doi.org/10.1080/00207543.2019.1630772.
Kamble, S.S., Gunasekaran, A., Sharma, R., 2018. Analysis of the driving and dependence power of barriers to adopt industry 4.0 in Indian manufacturing industry. Comput. For. Ind. 101, 107–119. https://doi.org/10.1016/j.compind.2018.06.004.
Kerin, M., Pham, D.T., 2019. A review of emerging industry 4.0 technologies in remanufacturing. J. Clean. Prod. 237, 117805. https://doi.org/10.1016/j.jclepro.2019.117805.
Kumar, A., Dixit, G., 2018. An analysis of barriers affecting the implementation of e-waste management practices in India: a novel ISM-DEMATEL approach. Sustain. Prod. Consum. 14, 36–52. https://doi.org/10.1016/j.spc.2018.01.002.
Kumar, R., 2020a. Sustainable Supply Chain Management in the Era of Digitialization, pp. 446–460. https://doi.org/10.4018/978-1-5225-8933-4.ch021.
Kumar, R., 2020b. Espousal of Industry 4.0 in Indian Manufacturing Organizations, pp. 245–252. https://doi.org/10.4018/978-1-7998-2772-6.ch013.
R. Kumar, R., Singh, K., Shankar, R., 2014. Strategy development by Indian SMEs for improving coordination in supply chain Compet. Rev. 24, 414–432. https://doi.org/10.1108/CR-06-2012-0016.
Kumar, R., Singh, R.K., Shankar, R., 2015. Critical success factors for implementation of supply chain management in Indian small and medium enterprises and their impact on performance. IIMB Manag. Rev. 27, 92–104. https://doi.org/10.1016/j.iimb.2015.03.001.
Kumar, R., Singh, S.P., Lamba, K., 2018. Sustainable robust layout using Big Data approach: a key towards industry 4.0. J. Clean. Prod. 204, 643–659. https://doi.org/10.1016/j.jclepro.2018.08.327.
Kumar, V., Sezersan, I., Garza-Reyes, J.A., Gonzalez, E.D.R.S., AL-Shboul, M.A., 2019. Circular economy in the manufacturing sector: benefits, opportunities and barriers. Manag. Decis. 57, 1067–1086. https://doi.org/10.1108/MD-09-2018-1070.
Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M., 2014. Industry 4.0. Bus. Inf. Syst. Eng. 6, 239–242. https://doi.org/10.1007/s12599-014-0334-4.
Lass, S., Gronau, N., 2020. A factory operating system for extending existing factories to Industry 4.0. Comput. Ind. 115, 103128. https://doi.org/10.1016/j.compind.2019.103128.
Lee, J., Bagheri, B., Kao, H.-A., 2015. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. Manuf. Lett. 3, 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001.
Liao, Y., Loures, E.R., Deschamps, F., Brezinski, G., Venâncio, A., 2018. The impact of the fourth industrial revolution: a cross-country/region comparison. Production 28. https://doi.org/10.1590/0103-6513.20180061.
Lieder, M., Asif, F.M.A., Rashid, A., Mihelič, A., Kotnik, S., 2017. Towards circular economy implementation in manufacturing systems using a multi-method simulation approach to link design and business strategy. Int. J. Adv. Manuf. Technol. 93, 1953–1970. https://doi.org/10.1007/s00170-017-0610-9.
Lieder, M., Rashid, A., 2016. Towards circular economy implementation: a comprehensive review in context of manufacturing industry. J. Clean. Prod. 115, 36–51. https://doi.org/10.1016/j.jclepro.2015.12.042.
Linder, C., 2019. Customer orientation and operations: the role of manufacturing capabilities in small- and medium-sized enterprises. Int. J. Prod. Econ. 216, 105–117. https://doi.org/10.1016/j.ijpe.2019.04.030.
Longoni, A., Cagliano, R., 2015. Environmental and social sustainability priorities. Int. J. Oper. Prod. Manag. 35, 216–245. https://doi.org/10.1108/IJOPM-04-2013-0182.
Luthra, S., Mangla, S.K., 2018. Evaluating challenges to Industry 4.0 initiatives for supply chain sustainability in emerging economies. Process Saf. Environ. Protect. 117, 168–179. https://doi.org/10.1016/j.psep.2018.04.018.
Machado, C.G., Winroth, M.P., Ribeiro da Silva, E.H.D., 2020. Sustainable manufacturing in Industry 4.0: an emerging research agenda. Int. J. Prod. Res. 58, 1462–1484. https://doi.org/10.1080/00207543.2019.1652777.
Mani, V., Jabbour, C.J.C., Mani, K.T.N., 2020. Supply chain social sustainability in small and medium manufacturing enterprises and firms' performance: empirical evidence from an emerging Asian economy. Int. J. Prod. Econ. 227, 107656. https://doi.org/10.1016/j.ijpe.2020.107656.
Marques, M., Agostinho, C., Zacharewicz, G., Jardim-Gonçalves, R., 2017. Decentralized decision support for intelligent manufacturing in Industry 4.0. J. Ambient Intell. Smart Environ. 9, 299–313. https://doi.org/10.3233/AIS-170436.
Masood, T., Egger, J., 2019. Augmented reality in support of Industry 4.0—implementation challenges and success factors. Robot. Comput. Integrated Manuf. 58, 181–195. https://doi.org/10.1016/j.rcim.2019.02.003.
Mastos, T.D., Nizamis, A., Vafeiadis, T., Alexopoulos, N., Ntinas, C., Gkortzis, D., Papadopoulos, A., Ioannidis, D., Tzovaras, D., 2020. Industry 4.0 sustainable supply chains: an application of an IoT enabled scrap metal management solution. J. Clean. Prod. 269, 122377. https://doi.org/10.1016/j.jclepro.2020.122377.
Mishra, S., Singh, S.P., Johansen, J., Cheng, Y., Farooq, S., 2019. Evaluating indicators for international manufacturing network under circular economy. Manag. Decis. 57, 811–839. https://doi.org/10.1108/MD-05-2018-0565.
Moktadir, M.A., Ali, S.M., Kusi-Sarpong, S., Shaikh, M.A.A., 2018. Assessing challenges for implementing Industry 4.0: implications for process safety and environmental protection. Process Saf. Environ. Protect. 117, 730–741. https://doi.org/10.1016/j.psep.2018.04.020.
Morrar, R., Arman, H., 2017. The fourth industrial revolution (industry 4.0): a social innovation perspective. Technol. Innov. Manag. Rev. 7, 12–20. https://doi.org/10.22215/timreview/1117.
Müller, J.M., Buliga, O., Voigt, K.-I., 2018b. Fortune favors the prepared: how SMEs approach business model innovations in Industry 4.0. Technol. Forecast. Soc. Change 132, 2–17. https://doi.org/10.1016/j.techfore.2017.12.019.
Müller, J.M., Kiel, D., Voigt, K.-I., 2018a. What drives the implementation of industry 4.0? The role of opportunities and challenges in the context of sustainability. Sustainability 10, 247. https://doi.org/10.3390/su10010247.
Mura, M., Longo, M., Zanni, S., 2020. Circular economy in Italian SMEs: a multi-method study. J. Clean. Prod. 245, 118821. https://doi.org/10.1016/j.jclepro.2019.118821.
Nascimento, D.L.M., Alencastro, V., Quelhas, O.L.G., Caiado, R.G.G., Garza-Reyes, J.A., Rocha-Lona, L., Tortorella, G., 2019. Exploring Industry 4.0 technologies to enable circular economy practices in a manufacturing context. J. Manuf. Technol. Manag. 30, 607–627. https://doi.org/10.1108/JMTM-03-2018-0071.
Orzes, G., Poklemba, R., Towner, W.T., 2020. Implementing industry 4.0 in SMEs: a focus group study on organizational requirements. In: Industry 4.0 for SMEs. Springer International Publishing, Cham, pp. 251–277. https://doi.org/10.1007/978-3-030-25425-4_9.
Pacaux-Lemoine, M.-P., Trentesaux, D., 2019. Ethical risks OF human-machine symbiosis IN industry 4.0: insights from the human-machine cooperation approach. IFAC-PapersOnLine 52, 19–24. https://doi.org/10.1016/j.ifacol.2019.12.077.
Pannell, D.J., 1997. Sensitivity analysis: strategies, methods, concepts, examples. Agric. Econ. 16, 139–152.
Pater, A., Van Gils, A., 2003. Stimulating ethical decision-making in a business context. Eur. Manag. J. 21, 762–772. https://doi.org/10.1016/j.emj.2003.09.016.
Pedersen, M.R., Nalpantidis, L., Andersen, R.S., Schou, C., Bøgh, S., Krüger, V., Madsen, O., 2016. Robot skills for manufacturing: from concept to industrial deployment. Robot. Comput. Integrated Manuf. 37, 282–291. https://doi.org/10.1016/j.rcim.2015.04.002.
Pfohl, H.-C., Yahsi, B., Kurnaz, T., 2017. Concept and Diffusion-Factors of Industry 4.0 in the Supply Chain, pp. 381–390. https://doi.org/10.1007/978-3-319-45117-6_33.
Pinto Taborga, C., Lusa, A., Coves, A.M., 2018. A proposal for a green supply chain strategy. J. Ind. Eng. Manag. 11, 445. https://doi.org/10.3926/jiem.2518.
Piyathanavong, V., Garza-Reyes, J.A., Kumar, V., Maldonado-Guzmán, G., Mangla, S.K., 2019. The adoption of operational environmental sustainability approaches in the Thai manufacturing sector. J. Clean. Prod. 220, 507–528. https://doi.org/10.1016/j.jclepro.2019.02.093.
Prieto-Sandoval, V., Jaca, C., Santos, J., Baumgartner, R.J., Ormazabal, M., 2019. Key strategies, resources, and capabilities for implementing circular economy in industrial small and medium enterprises. Corp. Soc. Responsib. Environ. Manag. csr 1761. https://doi.org/10.1002/csr.1761.
Radziwon, A., Bilberg, A., Bogers, M., Madsen, E.S., 2014. The smart factory: exploring adaptive and flexible manufacturing solutions. Procedia Eng. 69, 1184–1190. https://doi.org/10.1016/j.proeng.2014.03.108.
Raj, A., Dwivedi, G., Sharma, A., Lopes de Sousa Jabbour, A.B., Rajak, S., 2020. Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: an inter-country comparative perspective. Int. J. Prod. Econ. 224, 107546. https://doi.org/10.1016/j.ijpe.2019.107546.
Rajput, S., Singh, S.P., 2019a. Identifying Industry 4.0 IoT enablers by integrated PCA-

ISM-DEMATEL approach. Manag. Decis. 57, 1784–1817. https://doi.org/10.1108/MD-04-2018-0378.
Rajput, S., Singh, S.P., 2019b. Connecting circular economy and industry 4.0. Int. J. Inf. Manag. 49, 98–113. https://doi.org/10.1016/j.ijinfomgt.2019.03.002.
Ramelli, S., Wagner, A., 2020. What the stock market tells us about the consequences of COVID-19. In: Mitigating the COVID Economic Crisis: Act Fast and Do Whatever, vol. 63.
Rauch, E., Dallasega, P., Unterhofer, M., 2019. Requirements and barriers for introducing smart manufacturing in small and medium-sized enterprises. IEEE Eng. Manag. Rev. 47, 87–94. https://doi.org/10.1109/EMR.2019.2931564.
Rosa, P., Sassanelli, C., Urbinati, A., Chiaroni, D., Terzi, S., 2020. Assessing relations between Circular Economy and Industry 4.0: a systematic literature review. Int. J. Prod. Res. 58, 1662–1687. https://doi.org/10.1080/00207543.2019.1680896.
Sangwan, K.S., Bhakar, V., Digalwar, A.K., 2019. A sustainability assessment framework for cement industry – a case study. Benchmarking An Int. J. 26, 470–497. https://doi.org/10.1108/BIJ-01-2018-0021.
Satapathy, S., 2017. An analysis of barriers for plastic recycling in the Indian plastic industry. Benchmark Int. J. 24, 415–430. https://doi.org/10.1108/BIJ-11-2014-0103.
Schleinkofer, U., Klöpfer, K., Schneider, M., Bauernhansl, T., 2019. Cyber-physical systems as part of frugal manufacturing systems. Procedia CIRP 81, 264–269. https://doi.org/10.1016/j.procir.2019.03.046.
Schwab, L., Gold, S., Reiner, G., 2019. Exploring financial sustainability of SMEs during periods of production growth: a simulation study. Int. J. Prod. Econ. 212, 8–18. https://doi.org/10.1016/j.ijpe.2018.12.023.
Shashi, Centobelli, P., Cerchione, R., Singh, R., 2019. The impact of leanness and innovativeness on environmental and financial performance: insights from Indian SMEs. Int. J. Prod. Econ. 212, 111–124. https://doi.org/10.1016/j.ijpe.2019.02.011.
Shin, H., Hwang, J., Kim, H., 2019. Appropriate technology for grassroots innovation in developing countries for sustainable development: the case of Laos. J. Clean. Prod. 232, 1167–1175. https://doi.org/10.1016/j.jclepro.2019.05.336.
Shivajee, V., Singh, R.K., Rastogi, S., 2019. Manufacturing conversion cost reduction using quality control tools and digitization of real-time data. J. Clean. Prod. 237, 117678. https://doi.org/10.1016/j.jclepro.2019.117678.
Singh, R.K., Kumar, R., 2020. Strategic issues in supply chain management of Indian SMEs due to globalization: an empirical study. Benchmark Int. J. 27, 913–932. https://doi.org/10.1108/BIJ-09-2019-0429.
Singh, R.K., Kumar, R., Shankar, R., 2012. Supply chain management in SMEs: a case study. Int. J. Manuf. Res. 7, 165. https://doi.org/10.1504/IJMR.2012.046801.
Singh, R.K., Luthra, S., Mangla, S.K., Uniyal, S., 2019. Applications of information and communication technology for sustainable growth of SMEs in India food industry. Resour. Conserv. Recycl. 147, 10–18. https://doi.org/10.1016/j.resconrec.2019.04.014.
Sommer, L., 2015. Industrial revolution - industry 4.0: are German manufacturing SMEs the first victims of this revolution? J. Ind. Eng. Manag. 8 https://doi.org/10.3926/jiem.1470.
Song, M., Wang, S., 2017. Participation in global value chain and green technology progress: evidence from big data of Chinese enterprises. Environ. Sci. Pollut. Res. 24, 1648–1661. https://doi.org/10.1007/s11356-016-7925-1.
Song, M., Wang, S., 2018. Market competition, green technology progress and comparative advantages in China. Manag. Decis. 56, 188–203. https://doi.org/10.1108/MD-04-2017-0375.
Telukdarie, A., Buhulaiga, E., Bag, S., Gupta, S., Luo, Z., 2018. Industry 4.0 implementation for multinationals. Process Saf. Environ. Protect. 118, 316–329. https://doi.org/10.1016/j.psep.2018.06.030.
Thakur, V., Mangla, S.K., 2019. Change management for sustainability: evaluating the role of human, operational and technological factors in leading Indian firms in home appliances sector. J. Clean. Prod. 213, 847–862. https://doi.org/10.1016/j.jclepro.2018.12.201.
Theorin, A., Bengtsson, K., Provost, J., Lieder, M., Johnsson, C., Lundholm, T., Lennartson, B., 2017. An event-driven manufacturing information system architecture for Industry 4.0. Int. J. Prod. Res. 55, 1297–1311. https://doi.org/10.1080/00207543.2016.1201604.
Tiwari, K., Khan, M.S., 2020. Sustainability accounting and reporting in the industry 4.0. J. Clean. Prod. 258, 120783. https://doi.org/10.1016/j.jclepro.2020.120783.
Turner, C., Moreno, M., Mondini, L., Salonitis, K., Charnley, F., Tiwari, A., Hutabarat, W., 2019. Sustainable production in a circular economy: a business model for Re-distributed manufacturing. Sustainability 11, 4291. https://doi.org/10.3390/su11164291.
Wang, S., Wan, J., Li, D., Zhang, C., 2016. Implementing smart factory of industrie 4.0: an outlook. Int. J. Distributed Sens. Netw. 12, 3159805. https://doi.org/10.1155/2016/3159805.
Wollschlaeger, M., Sauter, T., Jasperneite, J., 2017. The future of industrial communication: automation networks in the era of the internet of things and industry 4.0. IEEE Ind. Electron. Mag. 11, 17–27. https://doi.org/10.1109/MIE.2017.2649104.
Xia, X., Govindan, K., Zhu, Q., 2015. Analyzing internal barriers for automotive parts remanufacturers in China using grey-DEMATEL approach. J. Clean. Prod. 87, 811–825. https://doi.org/10.1016/j.jclepro.2014.09.044.
Yadav, G., Luthra, S., Jakhar, S.K., Mangla, S.K., Rai, D.P., 2020. A framework to overcome sustainable supply chain challenges through solution measures of industry 4.0 and circular economy: an automotive case. J. Clean. Prod. 254, 120112. https://doi.org/10.1016/j.jclepro.2020.120112.
Yadav, S., Singh, S.P., 2020. Blockchain critical success factors for sustainable supply chain. Resour. Conserv. Recycl. 152, 104505. https://doi.org/10.1016/j.resconrec.2019.104505.
Zezulka, F., Marcon, P., Vesely, I., Sajdl, O., 2016. Industry 4.0 – an introduction in the phenomenon. IFAC-PapersOnLine 49, 8–12. https://doi.org/10.1016/j.ifacol.2016.12.002.
Zhou, X., Song, M., Cui, L., 2020. Driving force for China's economic development under Industry 4.0 and circular economy: technological innovation or structural change? J. Clean. Prod. 271, 122680. https://doi.org/10.1016/j.jclepro.2020.122680.

Dr. Ravinder Kumar is an Associate Professor at Amity University, Noida, India. He holds a bachelor's degree in Mechanical Engineering with Honors. He did his Master's and Ph.D. from Delhi College of Engineering, University of Delhi, India. He has published about 45 research papers in reputed international journals and Conferences. His areas of interest include supply chain management and quality management in SMEs. He has published papers in journals such as Competitiveness Review: An International Journal, Global Journal of Flexible Systems and Management, IIMB Management Review, Benchmarking: an International Journal, International Journal of Pharmaceutical and Healthcare Marketing, International Journal of Business Information Systems, International Journal of Manufacturing Research, etc.

Dr. Rajesh Kr. Singh is a Professor at the Management Development Institute, Gurgaon, India. His areas of interest include supply chain management, Industry 4.0, quality management and circular economy. He has about 170 research papers published in reputed international/national journals and conferences. He has published papers in journals such as Resources, Conservation and Recycling, Journal of Cleaner Production, Industrial Management and Data Systems, Production planning and Control, Singapore Management Review, International Journal of Productivity and Performance Management, Benchmarking: An International Journal, Journal of Modelling in Management, Competitiveness Review: An International Business Journal, Global Journal of Flexible Systems and Management, International Journal of Productivity and Quality Management, International Journal of Services and Operations Management, International Journal of Automotive Industry and Management, International Journal of Logistic Systems Management, IIMB Management Review and Productivity Promotion.

Yogesh K. Dwivedi is a Professor of Digital Marketing and Innovation, Founding Director of the Emerging Markets Research Centre (EMaRC) and Co-Director of Research at the School of Management, Swansea University, Wales, UK. Professor Dwivedi is also currently leading the *International Journal of Information Management* as its Editor-in-Chief. His research interests are at the interface of Information Systems (IS) and Marketing, focusing on issues related to consumer adoption and diffusion of emerging digital innovations, digital government, and digital marketing particularly in the context of emerging markets. Professor Dwivedi has published more than 300 articles in a range of leading academic journals and conferences that are widely cited (more than 15 thousand times as per Google Scholar). Professor Dwivedi is an Associate Editor of the *European Journal of Marketing, Government Information Quarterly* and *International Journal of Electronic Government Research*, and Senior Editor of the *Journal of Electronic Commerce Research*. More information about Professor Dwivedi can be found at: http://www.swansea.ac.uk/staff/som/academic-staff/y.k.dwivedi/.



Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Artificial neural network models for global solar energy and photovoltaic power forecasting over India

Gulnar Perveen , M. Rizwan , Nidhi Goel & Priyanka Anand

To cite this article: Gulnar Perveen , M. Rizwan , Nidhi Goel & Priyanka Anand (2020): Artificial neural network models for global solar energy and photovoltaic power forecasting over India, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, DOI: 10.1080/15567036.2020.1826017

To link to this article: https://doi.org/10.1080/15567036.2020.1826017

Published online: 12 Oct 2020.

Submit your article to this journal

Article views: 9

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ueso20

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
https://doi.org/10.1080/15567036.2020.1826017

Taylor & Francis
Taylor & Francis Group

Check for updates

Artificial neural network models for global solar energy and photovoltaic power forecasting over India

Gulnar Perveen [a,b,c,d,e], M. Rizwan [a,b,c,d,e], Nidhi Goel [a,b,c,d,e], and Priyanka Anand [a,b,c,d,e]

[a]Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India; [b]Department of Electrical Engineering, College of Engineering, Qassim University, Buraydah,-Qassim, Saudi Arabia; [c]Department of Electrical Engineering, Delhi Technological University, Delhi, India; [d]Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India; [e]Department of Electrical Engineering, Bhagat Phool Singh Mahila Vishwavidyala, Khanpur Kalan, Sonipat, India

ABSTRACT

For design and development of solar-based photovoltaic (PV) system, accuracy in obtaining data of solar energy has been a prime concern. Forecasting of global solar energy in clear or sunny weather conditions are easily calculated; however, the challenge lies in estimating under the impact of foggy and cloudy conditions. Therefore, this research investigates and presents the Artificial Neural Network (ANN) models for global solar energy forecasting for distinct sky-conditions such as sunny, hazy, partly, and fully cloudy conditions. This research further applied for distinct climatic conditions namely composite, warm & humid, moderate, hot & dry, and cold & cloudy climate zone using meteorological parameters. Further, PV power forecasting has been essential for planning the operating capability of energy systems and its reliability. The present work explores the implementation of proposed model for short-term photovoltaic power forecast in solar energy systems. Lastly, comparative analysis of the developed model with other neural network models are carried out using statistical evaluation indices. It has been observed from the overall analysis that by using the Radial-Basis-Function-Neural-Network (RBFNN) model, a significant reduction in error has been noticed as compared to other methods. Further, it is noted that the sunny model provides improved performance for hot & dry climate zone; hazy model gives favorable output for composite climate zone; partly cloudy model is favorable for warm & humid climate zone; and lastly, fully cloudy model provides favorable results for cold & cloudy climate zone. The model has been implemented for broader application series and it has been observed that for short-term photovoltaic power forecasting under composite climatic zone, hazy model emerges out to provide better performance with Mean Absolute Percentage Error (MAPE) of 0.0019% followed by sunny, partly, and fully cloudy model with MAPE 0.024%, 0.054%, and 0.109% respectively. The models presented in this research will be useful for scholars to design and engineer the solar energy systems.

ARTICLE HISTORY
Received 2 July 2020
Revised 4 September 2020
Accepted 13 September 2020

Keywords
Artificial neural network; meteorological parameters; sky-condition; solar energy forecasting; solar photovoltaic

Introduction

The Introduction section has been classified into four sub-section, i.e. motivation and significance, literature review, research gap, and novelty of the present research work. The details of each sub-section are as follows.

CONTACT Gulnar Perveen gulegulnar@gmail.com Department of Electronics and Communication Engineering, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042,India

© 2020 Taylor & Francis Group, LLC

Motivation and significance

Global warming, energy crisis, and energy security is becoming the major challenge soon and therefore it becomes necessary to use and develop alternative, sustainable, and clean energy resources for power generation. Thus, the use of Renewable Energy Sources (RES) is invigorated; since solar energy is inexhaustible in nature, therefore, considered to be the most promising renewable resources for large scale power generation. Further, the power generation from solar PV system has introduced significant economic and environmental interests and may be able to reduce the carbon-dioxide (CO_2) emissions significantly and becoming a source of creating employment as well. Therefore, an accurate analysis for predicting RES can be very helpful in this perspective. For monitoring these RES, measuring equipment are installed at the projected sites; however, such meteorological sites are insufficient for providing radiation data because of costly equipment's, its maintenance, proper calibration of the instruments, and sufficient record period. Due to unavailability of data, a need has arisen for introducing intelligent approaches for global solar energy forecasting. Further, concept of PV power generation becomes imperative for investors to market and make a decisive policy for proper energy planning. Since the solar energy changes throughout the day from dawn to dusk with changes in seasonal variation, time, and geographical locations; therefore, if investors could understand the concept of short-term photovoltaic power generation, it would help them to strategize the market policy for investment purpose.

Literature review

Many researches were carried out for developing solar forecasting models. The first set of models were empirical models to forecast solar energy. A literature survey conveys that most of the developed regression models have been based on sunny sky; however, such models were found unsuitable for foggy/cloudy sky. Gouda et al. proposed about 220 empirical models to estimate solar energy over China using different meteorological parameters. It is concluded from this research that intelligent modeling techniques exhibit improved performance for solar energy estimation as compared to empirical models (Gouda et al. 2020). Bayrakci et al. presented regression models to estimate solar energy over Turkey with aid of meteorological parameters. It is concluded from the results that Veeran and Chegaar models are observed within ±10% acceptable error limits (Bayrakci, Demircan, and Kecebas 2018). Khalil et al. presented regression models to estimate solar energy at Saudi Arabia region. This research concludes that the solar insolation gives maximum value of solar energy during summer months, and lowest value can be observed between autumn and winter season (Khalil and Aly 2018). Perveen et al. presented sunshine-based & empirical model for solar energy forecasting correlating parameters for diversified climate zone over India. It is concluded from the results that considered parameters have strong influence to forecast solar energy (Perveen, Rizwan, and Goel 2018a).

Some uncertainty in the environment could be because of the presence of dust, moisture, and cloudy weather conditions. For India, wherein around more than 80 days are cloudy, it becomes very difficult for obtaining accuracy to forecast solar energy based on multiple regression analysis approaches. Since the regression models developed were satisfactory but applied only to clear sky-condition (Angstrom 1924). So, for meeting the uncertainty of variation in weather conditions, approaches based on intelligent modeling techniques are introduced. The second category comprises fuzzy logic models. The fuzzy set concept was presented by Zadeh using sunshine hours as a meteorological parameter for estimating global solar energy (Sen 1998). Since then various models are proposed which implies fuzzy logic models for solar energy forecasting. In recent times, a fuzzy logic models based on sky-conditions were introduced by Perveen et al. to forecast solar energy. It is concluded from the obtained results that using dew-point as an additional parameter, the model accuracy has significantly improved (Perveen, Rizwan, and Goel 2018b). Dhimish et al. proposed algorithm for a photovoltaic system based on fuzzy logic modeling and artificial neural network. It is

concluded from this research that artificial neural network model outperforms fuzzy logic modeling (Dhimish et al. 2018). For systems that employ complex computation of a large number of data sets, maintaining accuracy using fuzzy logic modeling is a complicated task and henceforth, the third category of models were introduced, i.e. artificial neural network models. Chang et al. presented RBFNN model for short-term photovoltaic power and wind speed forecast. This research concludes that the developed model gives better accuracy to forecast wind speed and photovoltaic power maintaining efficiency (Chang et al. 2017). These models used modeling techniques in diverse engineering branch, like neurology, mathematics, and meteorology. These models find its wider aspects of applications including optimum forecasting and optimization of the weather forecast. Zamen et al. proposed an optimization method to carry out various artificial intelligent approach. The photovoltaic-thermal (PV/T) efficiency has been forecasted using amb. temp., recirculation flow rate, and solar irradiation as input parameters. This research concludes that ANN model provides better performance than other networks (Zamen et al. 2019). Ahmadi et al., proposed an artificial neural network and Group Method of Data Handling (GMDH) approach in evaluating carbon-dioxide emissions. This research concludes that ANN model provides more accurate results to forecast carbon-dioxide emissions with absolute relative error of 2.3% which otherwise demonstrate the precision of forecasting model (Ahmadi et al. 2019). Notton et al. presented smart persistence and artificial neural network model for estimating solar energy over France. This research concludes that the improvement made by using artificial neural network model has increased the forecasting horizons as compared to small persistence model (Notton et al. 2019).

A recent study reveals various artificial neural network approaches to forecast solar energy using clustering & classification approach, and hybrid method. Yan et al. proposed intelligent models which includes Adaptive-Neuro-Fuzzy Inference System (ANFIS) and Generalized Regression Neural Network (GRNN) models for predicting methane adsorption using statistical evaluation indexes. This research concludes that ANFIS model gives more accuracy as compared to other models (Yan et al. 2018). Ustun et al. proposed reliable solar predicting model know as Simple Membership Function and Fuzzy Rule Generating Technique (SMGRT). This research concludes that artificial driven neural network based techniques were found better as compared to others (Ustun et al. 2020). A recent research by Kuhe et al. proposed ANN's ensemble to predict solar energy in Nigeria with aid of statistical analysis. It is concluded from this research that the ANN ensemble gives improved performance for evaluating solar energy for places where no measurements are done (Kuhe, Achirgbenda, and Agada 2019). Ghimire et al. proposed and evaluated artificial neural network model and compared with other data-driven models for forecasting global solar energy using evaluation indexes. It is concluded from this research that the performance of developed model was improved in comparison to other models namely Genetic Programming (GP), Support Vector Regression (SVR), and Gaussian Process Machine Learning (GPML) models resulting in lower values of statistical indices (Ghimire et al. 2019). Recently, it has been observed that for modeling complex problems, soft computing approaches were used. Many types of researches were done using Linear Neural Network (LNN) and GRNN which is trained for modeling static and dynamic systems. Many researches were done based on soft computing methods. Rohani et al., proposed a soft computing method, i.e. Gaussian Process Regression (GPR) model to estimate solar energy over Iran. This research concludes that GPR model gives better performance that other models with low value of evaluation indexes (Rohani, Taki, and Abdollahpour 2018). Feng et al. developed models to estimate evapotranspiration with use of amb. temperature as an input parameter for China. It is revealed in this research that artificial neural network models gives better performance than other models (Feng et al. 2017). Kasaeian et al. proposed artificial intelligence analysis for estimating the temperature changes, and also finding data quality and solar chimney temperature using statistical indices. It is concluded from this research that the experimental data of solar chimney were found competent with no noise (Kasaeian et al. 2017). Loni et al. presented artificial neural network modeling for evaluating concentrator collector performance using cavity receiver. This research concludes that developed model provides better performance in comparison to other methods for each prediction (Loni et al. 2017).

Perveen et al. proposed comparison of intelligent models to forecast solar energy. The models were developed using the meteorological parameters. It is concluded from this research that the artificial neural network model perform better than the fuzzy logic model (Perveen, Rizwan, and Goel 2019).

Further, most of the researchers have worked toward the accuracy in forecasting PV power in solar energy systems. Many factors influence the PV power generation namely solar irradiance, variation in sky-conditions, topographical locations, etc. Due to inconsistency of solar irradiation and factors affecting environment such as sky-conditions, PV power generation has become a stochastic process that apart from affecting the system stability and robustness affects the capital and maintenance costs also. Kumar et al. proposed artificial neural network model to forecast power generation. It is concluded from results that the developed model gives improved accuracy (Kumar and Kalavathi 2018). Nespoli et al. proposed methods based on ANN which are trained on the same dataset. It is concluded from this research that during sunny sky-condition, the forecasting method which feeds on the available dataset provides improved performance as compared to the hybrid method whereas for cloudy days, hybrid method gives improved performance than the forecasting method (Nespoli et al. 2019). Ahmadi et al. proposed ANN, and neuro-fuzzy model for estimating solar PV-thermal collector performance. It is concluded in this research that the neuro-fuzzy model outperformed the ANNs models. So, for advancing the system stability, and accurate PV power forecast is must taking into consideration the influence of sky-conditions (Ahmadi et al. 2020). Mahmood et al. presented Multi-layer Perceptron (MLP) network for developing biogas predictor model. It is concluded from overall analysis that the proposed model gives high accuracy (Mahmood et al. 2020). Mehran et al. proposed artificial intelligence approach for forecasting temperature changes due to radiation variations for solar chimney with the aid of statistical indicators. It is concluded from the overall analysis that the experimental data of solar chimney were qualified with no noise (Mehran et al. 2017). Olusola et al. presented artificial neural network models to forecast solar radiation and photovoltaic parameters over Nigeria. It is concluded from the results that the developed models predict solar irradiance and solar PV parameters with accuracy (Olusola et al. 2020). Yildiz et al. presented artificial neural network for forecasting very short-term power in solar photovoltaic system. The Kernel Extreme Learning Machine (KELM) has been proposed for developing PV power forecast. It is concluded from the results that the KELM gives more reliable performance (Yildiz and Acikgoz 2020). Perveen et al. proposed comparison of empirical and intelligent models to forecast solar energy with aid of meteorological parameters. It is concluded from overall analysis that the performance of the model using intelligent modeling techniques is better than the empirical models using multiple regression analysis (Perveen, Rizwan, and Goel 2019).

Research gap

A detailed analysis reveals that many intelligent models found in the literature survey were based on clear sky-condition for solar energy forecast; however, a very little research has elucidated the concept of using intelligent modeling approach for distinct sky-condition namely sunny, hazy, partly, and fully cloudy conditions for distinct climate zone across India. Further, most of the previous researches exploited the use of commonly available meteorological parameters namely sunshine duration, amb. temp., etc.; however, very few researchers have explicated the concept of using dew-point; the inclusion of this parameter has considerably enhanced the model accuracy.

Novelty

The novelty of this research lies in the development of ANN models for distinct sky-conditions namely sunny, hazy, partly, and fully cloudy conditions to forecast global solar energy which further has been applied for distinct climate zone namely warm & humid, moderate, composite, hot & dry, and cold & cloudy climate. Further, the obtained results are realized for forecasting short-term photovoltaic power in the solar energy systems. Simulations are being carried out using MATLAB platform.

Lastly, a comparison has been made with other neural network models namely cascade forward, elman, generalized regression, layered recurrent, linear network, fuzzy logic, and regression models using statistical evaluation indices for checking the model validity.

Methodology

Climate zone

As per the study carried out by Bansal and Minke in the year 1988, five climate zones are defined by evaluation of the average monthly data that are considered from 233 meteorological stations across India. The zones are defined as warm & humid, moderate, composite, hot & dry, and cold & cloudy climatic zone as presented in Table 1. One such criterion for assigning the site to these zones lies in weather conditions existing for more than six months. The solar radiation resource stations across India are presented in Table 1 (Ajit 2009; Bansal and Minke 1988).

Collection of meteorological data

Data collection for five meteorological stations across India

In the present work, hourly data for 15 years are considered which is obtained from Solar Energy Center (SEC), now transformed into an independent and autonomous body named as National Institute of Solar Energy (NISE), Gurugram, and Center for Wind Energy Technology (CWET), Chennai in collaboration with Indian Meteorological Department (IMD), India.

Meteorological parameters namely sunshine hours, ambient temp., wind speed, rel. humidity, and dew-point are used as input parameters whereas global solar energy is the output parameter to forecast solar energy. Inclusion of these parameters has made the model accuracy more significant. These data are collected for 24 hours as per Indian Standard Time (IST). Sunshine hour are measured using Campbell – Stokes during the availability of sunshine hours as per Local Apparent Time (LAT). Bimetallic thermograph measures air temperature and dew-point; hygrometer is used for measuring relative humidity. For measuring wind speed, electrical anemograph is used whereas global solar energy is measured using pyranometer and shown in Tables 26 respectively.

Data collection for 210 W_p PV module power plant

Data pertaining to 210 W_p polycrystalline module photovoltaic power plant are considered which is installed at National Institute of Solar Energy, Gurugram. For the PV power plant, meteorological parameters such as temperature of cell, solar irradiance, open circuit voltage, and short circuit current are considered as input parameter whereas photovoltaic power is the output.

Data normalization

Data such as ambient temperature, sunshine duration, global solar energy, dew-point, relative humidity, and wind speed are standardize to evade convergence problems and defined with 0.1–0.9 interval as shown in Eq. (1).

$$L_s = (((\frac{X_{\max} - X_{mix}}{L_{\max} - L_{mix}}) * (L - L_{min})) + X_{min}) \tag{1}$$

where L represents the measured data set,

L_s denotes standardize data set,

L_{max} denotes max. data range,

L_{min} denotes min. data range,

X_{max} denotes max. limit,

X_{min} denotes min. limit.

Classification of sky-condition

Models covered under different sky – conditions can be categorized as (Chel and Tiwari 2011):

Sunny: If H_d is equal or less than 25% of H_g and sunshine duration is either equal to or greater than 9 hours.

Hazy: If H_d is either less than 50% or more than 25% of H_g and sunshine duration is between 79 hours.

Partly cloudy: If H_d is either less than 75% or more than 50% of H_g and sunshine duration is between 57 hours.

Fully cloudy: If H_d is more than 75% of H_g and sunshine duration is less than 5 hours.

Design and development of artificial neural network models

ANN-based models are intended in a way such that output parameter is calculated from input parameters by arranging basic connections. In this, models based on radial-basis-function and feed-forward neural network are simulated using input parameters namely amb. temp., dew-point, wind speed, duration of sunshine hours, and rel. humidity for forecasting output parameter i.e. global solar energy. Training algorithm such as Levenberg-Marquardt is employed, the details of which are explored in MATLAB. For the development of ANN model, the following are considered: (i) standardize value of measured and target data; (ii) formation of the neural model; (iii) training and simulating the neural model; (iv) generating output data set; (v) de-standardize the output data; (vi) comparison of measured and target data. The model validation is carried out using statistical evaluation indices.

The neural network comprises input level, a hidden level with 'tansig' function, and an output level. MATLAB is used for algorithm implementation. 'TRAINLM' are used for training the data set and is modeled by using Eq. (2) as:

$$y = \sum_{j=1}^{n} W_{ij} x_{ij} + \theta_i \tag{2}$$

where x_{ij} is the incoming signal, represents i neuron bias,

w_{ij} represents connected weightage focused from j to i neuron (Puneet, Nidhi, and Ansari 2017).

Determination of statistical evaluation indices

For model validation, statistical evaluation indices are used and defined as (Dolara et al. 2018):

Mean Absolute Percentage Error (MAPE): It represents the ratio of summing up the absolute error in each period and the measured data which is averaged to fixed percentages and given by Eq. (3) as:

$MAPE_{\%}$ =(3)

Normalized Mean Absolute Error (NMAE): Since data changes considerably i.e. from dawn to dusk, so NMAE is chosen which is described by Eq. (4) as:

$NMAE_{\%}$ =(4)

Normalized Root Mean Square Error (nRMSE): This definition of error is normalized over the maximum hourly measured data and given by Eq. (5) as:

$nRMSE_{\%}$ = (5)

where A = (m_i _ f_i) represents the absolute error, x denotes number of observations, m_i and f_i represents i_{th} measured and forecasted data.

Implementing the artificial neural network model in the solar energy system

Since the power generated in the solar photovoltaic system gets affected by factors namely dust, clouds, moisture and temperature differences in the upper atmospheric region; therefore, it becomes

a necessity for accurately forecasting the power since it improves the system robustness and reliability which otherwise help manufacturer's to expedient some methodologies for achieving better energy management. The developed is model further implemented for short-term photovoltaic power forecasting in the solar energy system which uses 210 W_p polycrystalline PV module operated at Maximum Power Point (MPP) conditions. Its specification detail is listed in Table 7 as:

Based on Standard Test Condition (STC), and influenced by variables like cell temperature and solar irradiance, the power generation can be defined using Eq. (6) and Eq. (7) as follows (Panapakidis and Dagoumas 2016):

P_{PV} =* (6)

and – 20)(7)

where $P_{PV,\ STC}$ represents solar energy system output at Standard Test Condition (STC) (W); P_{PV} represents photovoltaic panel output (W); G_T represents solar irradiance at STC (W/m^2); N_{PVS} represents the photovoltaic panel in series (number); is a temperature parameter; N_{PVP} represents the photovoltaic panel in parallel (number); T_{amb} is amb. temp. (^{o}C); T_j is the photovoltaic panel temp. (^{o}C); N_{OCT} is constant.

Simulation results and discussion

The artificial neural network models are established using radial-basis-function and feed-forward neural network using meteorological parameters based on weather-conditions for different climatic conditions. In this work, the real-time data are obtained from NISE and CWET in collaboration with IMD and are recorded on hourly basis for global solar energy forecasting. Dataset comprises geographical and meteorological parameters for the period between 2009 and 2018. Data are segregated into two sections wherein 60% of the data are used for training; 20% of the data are used for testing; and remaining 20% are used for validation, and shown in Tables A1A5 of Appendix A. After training the data set, the models are tested with data set not found in training. The training details are simulated in MATLAB platform. For model validation, statistical evaluation indices are used for validating the proposed model for distinct sky-conditions and are presented in Tables 8 and 9 respectively. The overall analysis based on Tables 8 and 9 reveals that the radial-basis-function network model meets accuracy and precision while comparing with feed-forward model. Some points are elucidated for models based on different sky-conditions which is further applied to five meteorological locations as:-

Sunny Sky: Analysis carried out for sunny weather condition as shown in Table 8 reveals that by using radial-basis-function network model for simulating measured and forecasted data, mean absolute percent error achieved is 0.0110×10^{-6}%, normalized mean absolute error is 0.0101×10^{-6}%, and normalized root mean square error is 0.01363%; whereas for the feed-forward neural model, the mean absolute percentage error observed is 0.447%, normalized mean absolute error is 0.4079%, and normalized root mean square error is 0.3088% as presented in Table 9. Further, it is noticed that in each of the developed model, the error obtained is minimum for Jodhpur station which is categorized under hot and dry climatic conditions. This is because of the reason that Jodhpur climate zone is hot and dry wherein bright/sunny sky-conditions prevails throughout year, with varying temperature prevails. In the summer period, the ambient temp. lies between 30 and 35°C; whereas, in winter, varies between 15 and 25°C; therefore, a variation of the order of 10°C is noticed as presented in Table 4. Jodhpur climate is extreme, with high solar insolation during day time because of which the surrounding region got heated up very quickly and at night time also there is a clear sky because of which the heat absorbed by atmosphere gets dissipated to higher atmospheric surface. The Jodhpur climate is hot and dry with sandy ground conditions as the relative humidity varies from 24–62% which is very low due to water bodies.

Hazy Sky: For hazy sky-condition, it is observed from Table 8 that by using radial-basis-function model for simulating the measured and forecasted data, mean absolute percentage error obtained is 0.0126×10^{-6}%, normalized mean absolute error is 0.0105×10^{-6}%, and normalized root mean square

Table 4. Standardize data for hot and dry climatic conditions.

	Sunshine Hours (hrs)		H_g (MJ/m^2)		Rel. Humidity (%)		Amb. Temp. (^{o}C)		Wind Speed (m/s)		Dew Point (^{o}C)	
Months	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize
Jan	10.12	0.53	17.01	0.61	60.89	0.50	20.79	0.47	2.50	0.51	10.59	0.50
Feb	10.22	0.74	19.65	0.59	49.82	0.44	22.28	0.44	6.34	0.21	13.01	0.50
Mar	10.89	0.70	21.67	0.61	42.28	0.46	25.28	0.50	4.24	0.35	5.09	0.31
Apr	9.87	0.65	23.35	0.64	45.47	0.59	30.12	0.60	5.07	0.49	10.74	0.56
May	9.83	0.73	24.40	0.62	56.22	0.61	30.15	0.41	10.76	0.51	18.40	0.53
Jun	5.76	0.49	19.94	0.61	77.83	0.41	26.02	0.65	9.48	0.47	22.39	0.49
Jul	5.27	0.28	14.12	0.51	87.12	0.52	24.05	0.55	8.94	0.31	20.79	0.55
Aug	5.32	0.44	17.45	0.44	86.25	0.52	24.13	0.56	8.74	0.53	22.89	0.53
Sep	6.57	0.54	19.41	0.54	85.25	0.56	23.78	0.39	4.68	0.52	21.60	0.58
Oct	8.67	0.60	19.79	0.54	75.97	0.62	23.59	0.56	2.28	0.31	20.07	0.59
Nov	8.46	0.61	17.53	0.47	71.97	0.41	22.41	0.49	2.01	0.46	18.13	0.42
Dec	8.74	0.74	17.11	0.73	63.03	0.41	19.27	0.51	2.44	0.51	16.34	0.66

Table 6. Standardize data for cold and cloudy climatic conditions.

	Sunshine Hours (hrs)		H_g (MJ/m^2)		Rel. Humidity (%)		Amb. Temp. (^oC)		Wind Speed (m/s)		Dew Point (^oC)	
Months	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize
Jan	7.06	0.73	14.44	0.64	76.60	0.43	9.95	0.64	3.62	0.37	7.45	0.48
Feb	6.36	0.61	16.58	0.52	70.84	0.53	10.25	0.56	3.81	0.41	7.37	0.50
Mar	7.22	0.64	21.01	0.59	60.02	0.46	15.55	0.46	5.65	0.39	10.39	0.53
Apr	3.80	0.36	19.97	0.55	62.58	0.54	18.26	0.52	7.64	0.48	16.25	0.65
May	4.85	0.49	18.43	0.43	49.13	0.54	20.70	0.58	4.30	0.37	19.20	0.66
Jun	3.49	0.46	16.42	0.60	86.25	0.57	21.14	0.61	3.64	0.36	20.81	0.54
Jul	2.71	0.36	16.14	0.62	86.78	0.54	21.25	0.62	3.15	0.37	20.58	0.46
Aug	2.16	0.42	14.24	0.53	90.12	0.57	20.74	0.63	1.23	0.19	21.24	0.50
Sep	2.75	0.36	13.99	0.44	86.45	0.49	20.01	0.48	0.87	0.34	19.16	0.65
Oct	5.88	0.56	15.01	0.47	79.48	0.66	18.40	0.61	2.30	0.47	17.80	0.45
Nov	7.06	0.62	15.63	0.55	76.25	0.53	15.23	0.40	2.64	0.34	12.64	0.46
Dec	7.60	0.64	15.64	0.65	75.16	0.60	11.89	0.35	0.84	0.27	10.15	0.53

Table 1. Topographical features of meteorological stations with distinct climatic conditions.

Climatic zone	Lat. (°N)	Long. (°E)	Meteorological Stations	T_{amb} (°C)	Rel. Hum. (%)	N Number of sunny days
Hot & dry	26.28	73.02	IIT Jodhpur, Rajasthan	>30	< 55	>20
Warm & humid	13.08	80.27	CWET Chennai, Tamil Nadu	>30	>55	< 20
Composite	28.61	77.2	Gurgaon, Haryana	Such condition exists when six months or more do not arise within any of the below-mentioned classifications.		
Moderate	18.52	73.85	Karad, Pune	25–30	< 75	< 20
Cold & cloudy	25.56	91.88	Meghalaya	<25	>55	< 20

error is 0.014784%; whereas for feed-forward network model, mean absolute percentage error is 0.5061%, normalized mean absolute error is 0.5101%, and normalized root mean square error is 0.099% as illustrated in Table 9. It has been noted for hazy sky model, that analysis based on statistical evaluation indices provides minimum error for Delhi meteorological station which is categorized under the composite climate zone. The contributing factor for this sky-condition could be the varying relative humidity from 30–40% in the dry period to 60–80% in the wet period as presented in Table 2. The presence of the high amount of relative humidity has been one of the main reasons that categorize Delhi under composite climate zone and not under hot and dry climatic conditions. Further, it has been noticed that solar intensity is high in summer and low in monsoon with predominant diffuse solar radiation. The sky is generally overcast and dull during monsoon; it is clear in winter and often hazy during summer.

Partly Cloudy Sky: Analysis carried out for partly cloudy sky as illustrated in Table 8 reveals that by using the radial-basis-function neural model, mean absolute percentage error obtained is 0.0130×10^{-6}%, normalized mean absolute error is 0.0109×10^{-6}%, and normalized root mean square error is 0.02140% by performing simulation between measured and forecasted data; whereas in the feed-forward neural network model, mean absolute percentage error is 0.6099%, normalized mean absolute error is 0.2094%, and normalized root mean square error is 0.1984% as shown in Table 9. For partly cloudy sky, the statistical analysis reveals that the obtained error is minimum for Chennai station which is categorized under the warm and humid climatic conditions.

The reason behind is the fact that the sky is partly cloudy as diffuse energy is quite intense in sunny days and the heat dissipation is marginal in nighttime due to cloud presence. During summer, the temperature reaches 30–35°C; whereas, during winter, ambient temperature remains between 25 and 30°C during day time as presented in Table 3. Hence, the variation of ambient temperature is comparatively lower. Most important features of this sky-condition is the high amount of humidity present in the atmosphere varying from 65 to 85% throughout the year.

Fully Cloudy Sky: It is learnt from Table 8 that by using radial-basis-function model, mean absolute percentage error obtained is 0.0121×10^{-6}%, normalized mean absolute error is 0.0111×10^{-6}%, and normalized root mean square error is 0.015369% by performing simulation between measured and forecasted data; whereas for the feed-forward neural network model, the mean absolute percentage error is 0.01562%, normalized mean absolute error is 0.1207%, and normalized root mean square error is 0.0205% as illustrated in Table 9. For fully cloudy sky, the statistical analysis reveals that obtained error is minimum for Shillong station, categorized under cold & cloudy climate zone.

One of the possible reason could be the availability of solar intensity which is comparatively lower in winter due to diffuse solar energy, the presence of which has made winters extremely cold. In summer, the variation in ambient temperature lies between 20 and 25°C and reaches low from 10 to 15°C during winter as presented in Table 6. Further, it has been noticed that relative humidity is high which lies between 70 and 90%. The weather for this climatic conditions is usually overcast and cloudy excluding small summers.

It has been noticed that for the radial-basis-function-neural network model, error obtained is comparatively lower comparing with feed-forward network model for distinct climate zone. Further, the graph representing the comparative analysis of measured and forecasted data are plotted based on

Table 2. Standardize data for composite climatic conditions.

	Sunshine Hours (hrs)		H_g (MJ/m^2)		Rel. Humidity (%)		Ambient Temp. (^oC)		Wind Speed (m/s)		Dew Point (^oC)	
Months	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize
Jan	7.72	0.66	14.99	0.43	66.49	0.39	15.12	0.48	6.16	0.43	4.59	0.45
Feb	7.94	0.60	17.79	0.45	60.65	0.50	19.57	0.56	8.85	0.29	4.89	0.45
Mar	7.4	0.73	20.12	0.67	54.30	0.41	23.74	0.36	8.35	0.43	5.63	0.48
Apr	9.2	0.72	26.22	0.61	37.19	0.46	31.03	0.54	9.42	0.36	5.71	0.44
May	8.85	0.65	25.23	0.57	35.32	0.36	35.14	0.66	9.52	0.41	8.65	0.54
Jun	7.13	0.60	21.90	0.64	53.57	0.37	34.40	0.53	10.59	0.46	16.15	0.50
Jul	4.59	0.44	20.39	0.42	71.64	0.60	31.49	0.42	10.40	0.51	24.59	0.63
Aug	5.54	0.54	18.81	0.54	80.36	0.43	30.15	0.60	9.58	0.38	26.07	0.61
Sep	6.67	0.56	14.86	0.54	70.28	0.38	30.73	0.52	9.43	0.56	24.50	0.64
Oct	9.33	0.72	19.334	0.53	65.52	0.54	27.18	0.50	7.34	0.37	12.44	0.63
Nov	7.20	0.55	15.57	0.35	50.81	0.44	21.93	0.63	7.54	0.49	7.07	0.55
Dec	5.27	0.60	13.13	0.58	66.69	0.49	16.98	0.47	6.94	0.45	3.43	0.37

Table 3. Standardize data for warm and humid climatic conditions.

	Sunshine Hours (hrs)		H_g (MJ/m^2)		Rel. Humidity (%)		Amb. Temp. (^{o}C)		Wind Speed (m/s)		Dew Point (^{o}C)	
Months	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize
Jan	8.94	0.701	17.64	0.524	71.32	0.515	25.46	0.223	8.22	0.43	20.33	0.47
Feb	9.75	0.866	21.07	0.899	76.29	0.549	26.61	0.788	9.06	0.45	21.13	0.52
Mar	9.05	0.882	23.53	0.594	73.75	0.318	27.84	0.799	6.98	0.50	20.77	0.54
Apr	9.37	0.61	23.86	0.472	71.26	0.412	30.53	0.475	8.78	0.42	23.07	0.47
May	8.83	0.626	22.87	0.753	67.14	0.452	31.71	0.506	7.49	0.58	24.49	0.50
Jun	7.35	0.64	21.51	0.683	64.04	0.532	30.73	0.539	8.48	0.56	24.65	0.44
Jul	6.18	0.534	18.87	0.661	62.25	0.465	30.53	0.649	10.04	0.45	22.74	0.50
Aug	4.78	0.511	19.04	0.583	71.30	0.504	29.07	0.554	8.85	0.49	24.52	0.49
Sep	6.16	0.582	19.66	0.596	79.09	0.542	29.10	0.216	8.41	0.44	24.63	0.44
Oct	6.52	0.622	17.28	0.633	80.06	0.461	27.75	0.595	6.25	0.41	24.20	0.70
Nov	5.81	0.547	15.13	0.566	83.73	0.467	25.64	0.519	11.58	0.46	22.71	0.58
Dec	7.07	0.638	13.89	0.527	78.13	0.527	26.09	0.272	9.49	0.50	21.16	0.49

Table 5. Standardize data for moderate climatic conditions.

	Sunshine Hours (hrs)		H_g (MJ/m^2)		Rel. Humidity (%)		Amb. Temp. (^{o}C)		Wind Speed (m/s)		Dew Point (^{o}C)	
Months	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize	Measured	Standardize
Jan	9.23	0.69	16.21	0.60	45.33	0.40	18.10	0.47	6.38	0.51	2.24	0.53
Feb	9.72	0.70	19.30	0.49	41.86	0.41	19.87	0.54	6.12	0.38	4.38	0.49
Mar	9.15	0.64	22.06	0.50	30.92	0.44	26.12	0.47	7.74	0.41	4.69	0.49
Apr	9.87	0.64	23.56	0.66	23.98	0.38	32.92	0.62	5.70	0.35	6.69	0.44
May	10.22	0.76	26.07	0.65	35.81	0.43	34.88	0.48	8.66	0.47	9.65	0.52
Jun	8.94	0.69	23.36	0.59	45.92	0.57	33.53	0.56	14.04	0.56	19.39	0.68
Jul	8.04	0.65	19.04	0.62	60.79	0.43	31.52	0.54	14.10	0.55	25.08	0.51
Aug	8.10	0.66	20.16	0.72	62.34	0.43	31.45	0.48	5.99	0.47	25.40	0.50
Sep	9.73	0.67	23.09	0.55	59.60	0.48	29.71	0.52	6.75	0.48	23.35	0.51
Oct	9.80	0.78	19.97	0.53	42.03	0.51	28.46	0.50	4.38	0.42	15.30	0.63
Nov	9.33	0.78	17.35	0.48	42.01	0.48	22.08	0.48	3.29	0.36	9.57	0.45
Dec	8.49	0.81	14.68	0.32	48.59	0.56	18.64	0.58	3.22	0.37	7.83	0.54

Table 7. Specification details of solar PV module.

S. No.	Specification Details	
1.	Short-circuit current (I_s)	5.57A
2.	Open-circuit voltage (V_{oc})	50.9 V
3.	Module efficiency	16.7%
4.	Cell efficiency	18.9%;
5.	Amb. Temp. (T_{amb})	− 4°F to 115°F
6.	N_{OCT} (Normal operating cell temperature)	114.8°F

Table 8. Radial-basis function neural network model based on sky-conditions for different climatic conditions.

		H_g (MJ/m^2)		MAPE	NMAE	nRMSE
Climate Zone	Sky-condition	Measured	RBFNN	(x10^{-6}%)	(x10^{-6}%)	(%)
CWET Chennai (Warm & humid)	Sunny	23.62	23.62	0.0941	0.0537	0.01304136
	Hazy	21.04	21.04	0.0757	0.0622	0.06649930
	Partly cloudy	18.40	18.40	0.0130	0.0109	0.02140537
	Fully cloudy	11.88	11.88	0.0620	0.0240	0.02187008
IIT Jodhpur (Hot & dry)	Sunny	21.76	21.76	0.0110	0.0101	0.0136354
	Hazy	19.79	19.80	0.0655	0.0745	0.0153404
	Partly cloudy	18.16	18.16	0.0381	0.0436	0.0164766
	Fully cloudy	15.75	15.78	0.0772	0.0439	0.0444673
SEC Gurgaon (Composite)	Sunny	20.58	20.58	0.0620	0.0764	0.0231622
	Hazy	18.12	18.10	0.0126	0.0105	0.0147846
	Partly cloudy	12.51	12.51	0.0438	0.0378	0.0153404
	Fully cloudy	11.37	11.30	0.0414	0.0582	0.0594645
Karad, Pune (Moderate)	Sunny	21.49	21.52	0.0781	0.0871	0.0127561
	Hazy	20.41	20.42	0.0699	0.0861	0.0248540
	Partly cloudy	18.71	18.71	0.0524	0.0323	0.0144679
	Fully cloudy	13.73	13.73	0.0681	0.0620	0.0706488
Meghalaya (Cold & cloudy)	Sunny	22.85	22.85	0.0576	0.0988	0.0889760
	Hazy	19.25	19.25	0.0838	0.0949	0.0785640
	Partly cloudy	15.93	15.93	0.0759	0.0616	0.0194640
	Fully cloudy	11.22	11.22	0.0121	0.0111	0.0153698

Table 9. Feed-forward neural network model based on sky-conditions for different climatic conditions.

		H_g (MJ/m^2)		MAPE	NMAE	nRMSE
Climate Zone	Sky-condition	Measured	RBFNN	(%)	(%)	(%)
CWET Chennai (Warm & humid)	Sunny	23.62	23.62	0.8088	0.7558	0.7601
	Hazy	21.04	21.04	0.7512	0.7271	0.3079
	Partly cloudy	18.40	18.40	0.6099	0.2094	0.1984
	Fully cloudy	11.88	11.88	0.8557	0.5777	0.6175
IIT Jodhpur (Hot & dry)	Sunny	21.76	21.76	0.4471	0.4079	0.3089
	Hazy	19.79	19.80	0.8232	0.6524	0.2857
	Partly cloudy	18.16	18.16	0.8565	0.7390	0.5030
	Fully cloudy	15.75	15.78	0.6163	0.6107	0.1107
SEC Gurgaon (Composite)	Sunny	20.58	20.58	0.9098	0.8237	0.6491
	Hazy	18.12	18.10	0.5062	0.5101	0.0993
	Partly cloudy	12.51	12.51	0.7915	0.7410	0.5725
	Fully cloudy	11.37	11.30	0.4830	0.9092	0.5297
Karad, Pune (Moderate)	Sunny	21.49	21.52	0.6986	0.6432	0.7195
	Hazy	20.41	20.42	0.8042	0.6319	0.1171
	Partly cloudy	18.71	18.71	0.8603	0.4329	0.4152
	Fully cloudy	13.73	13.73	0.6225	0.7257	0.3801
Meghalaya (Cold & cloudy)	Sunny	22.85	22.85	0.5298	0.5954	0.9759
	Hazy	19.25	19.25	0.8151	0.6580	0.2527
	Partly cloudy	15.93	15.93	0.9176	0.5167	0.6443
	Fully cloudy	11.22	11.22	0.1562	0.1207	0.0206

weather conditions and simulated for different climatic conditions namely warm & humid climate [CWET Chennai, Tamil Nadu (Latitude: 13.08°N, Longitude: 80.27°E)], hot & dry climate [IIT

Jodhpur, Rajasthan (Latitude: 26.23°N, Longitude: 73.02°E)], composite climate [SEC, Gurgaon (Latitude: 28.70°N, Longitude:77.10°E)], moderate climate [Karad, Pune (Latitude: 18.52°N, Longitude: 73.85°E)], and cold & cloudy climate zone [Meghalaya (Latitude: 25.57°N, Longitude: 91.89°E)] and presented in Figures 1 5 respectively.

It has been observed from Figures 15, that the radial basis function neural network model which is represented by red dashed line exactly follows the black solid line representing the measured data; whereas for the feed-forward neural network model which is represented by the blue line shows some deviation w.r.t. the measured data. Therefore, it is concluded that the radial basis function neural network model outperforms the feed-forward neural network model.

Further, it is noticed that among the four models investigated, clear/sunny model gives better performance for hot & dry climatic conditions as the forecasted data matches with measured data as presented in Figure 2A; similarly, the hazy model provides improved performance for composite climatic conditions as shown in Figure 3B; partly cloudy model performs better for warm & humid climatic conditions as presented in Figure 1C; and lastly, fully cloudy model gives better performance for cold & cloudy zone as shown in Figure 5D.

As we know that the power generating for the solar energy system is fluctuating in nature and becomes problematic. The output power varies because of various meteorological factors attributed to it which includes cloud cover, dust on panels, geographical locations and seasonal variations. Such a problem may be overcome by forecasting power using intelligent modeling techniques. Therefore, the proposed model is further implemented for simulating with input parameters such as solar irradiation, and cell temperature for forecasting short-term PV power and are presented in Table 10 for composite climatic conditions.

It has been observed from Table 10, that ANN models are established with input parameters namely solar irradiance, cell temp., open circuit voltage, and short-circuit current to forecast short-term PV power. The value of mean absolute percentage error is 0.007%, normalized mean absolute error is 0.092% and normalized root mean square error obtained is 0.109% which is achieved by

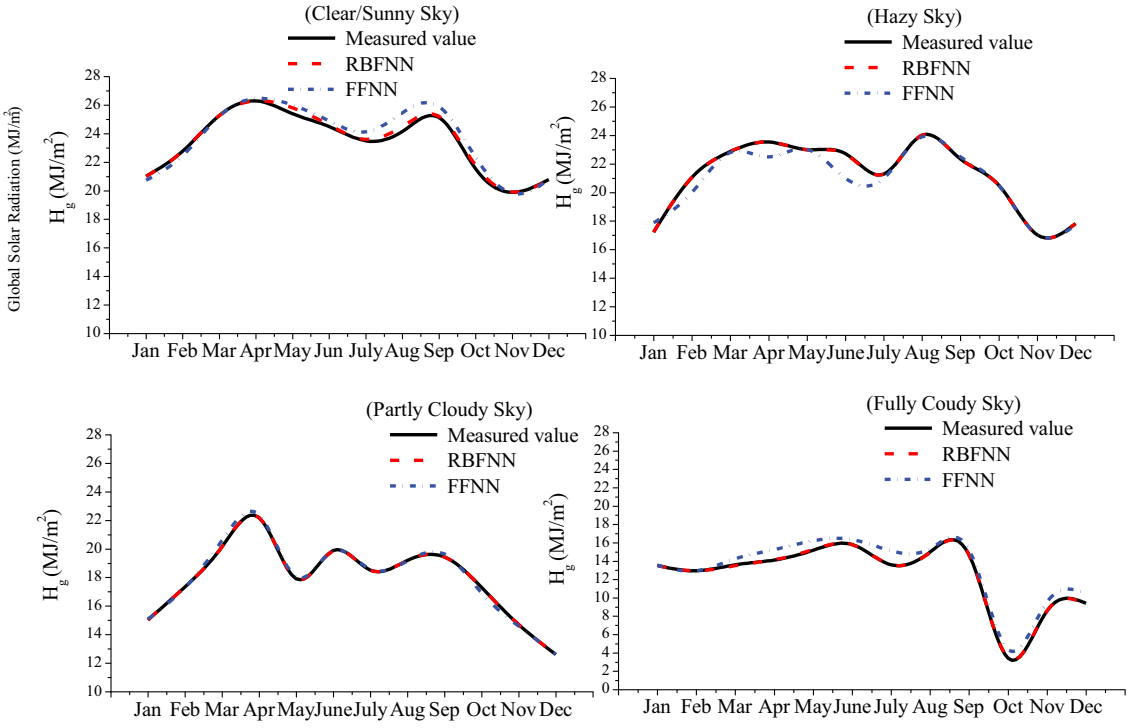


Figure 1. ANN model showing measured and forecasted H_g for warm and humid climatic conditions (a) sunny; (b) hazy; (c) partly cloudy; (d) fully cloudy sky.

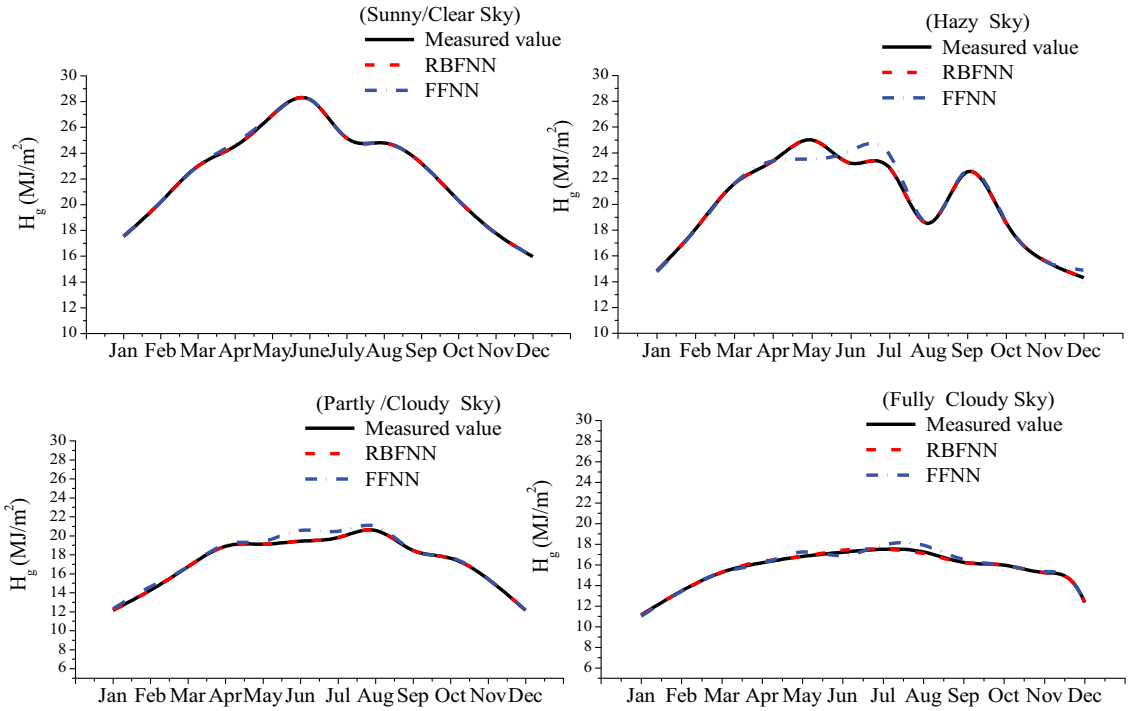


Figure 2. ANN model showing measured and forecasted H_g for hot & dry climatic conditions (a) sunny; (b) hazy; (c) partly cloudy; (d) fully cloudy sky.

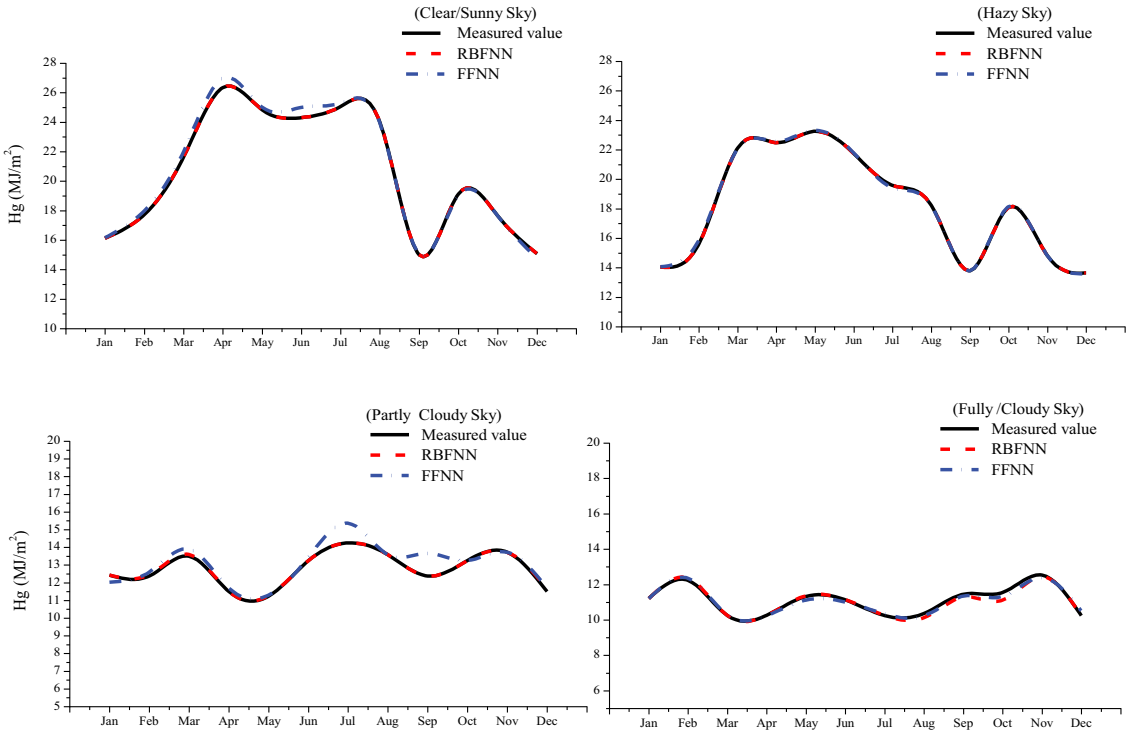


Figure 3. ANN model showing measured and forecasted H_g for composite climatic conditions (a) sunny; (b) hazy; (c) partly cloudy; (d) fully cloudy sky.

comparing the measured and forecasted data in a solar PV system employing 210 W_p polycrystalline PV module installed at NISE.

Further, the power generated in a solar energy system depends on many factors affecting the output of solar system but one of the most influential factors is varying weather condition which otherwise

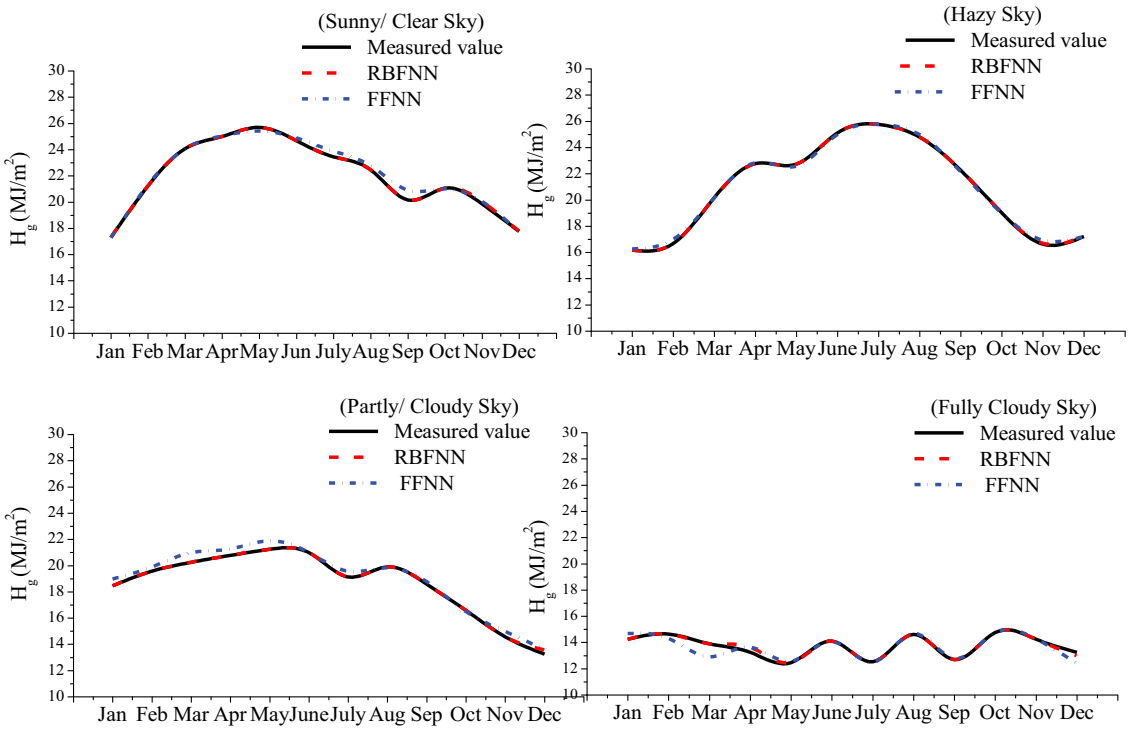


Figure 4. ANN model showing measured and forecasted H_g for moderate climatic conditions (a) sunny; (b) hazy; (c) partly cloudy; (d) fully cloudy sky.

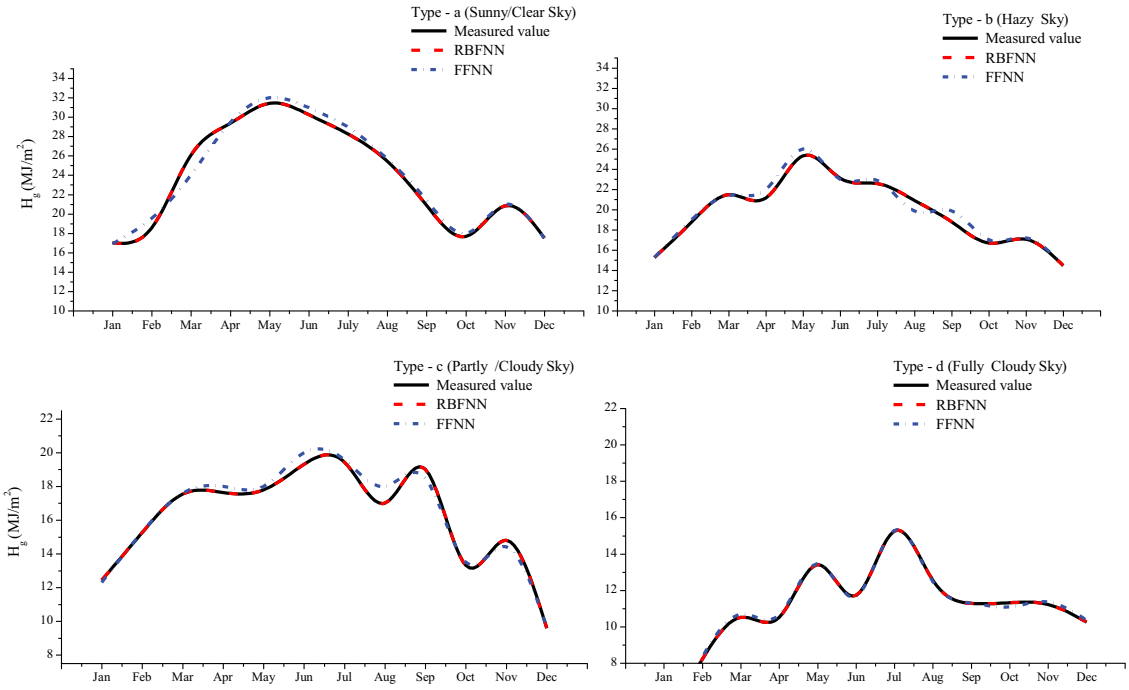


Figure 5. ANN model showing measured and forecasted H_g for cold & cloudy climatic conditions (a) sunny; (b) hazy; (c) partly cloudy; (d) fully cloudy sky.

makes it hard for examining the system performance by using a single model. Therefore, different sky-conditions are considered i.e. sunny, hazy, partly, and fully cloudy weather-conditions for short-term PV power forecast under composite climate and evaluation indexes are used for model validation and are presented in Figures 69 respectively.

Table 10. Short-term PV power forecast using ANN models for composite climatic conditions.

Month	Solar Irradiance (W/m^2)	V_{oc} (V)	I_{sc} (A)	Cell Temp. (°C)	Power (W)		MAPE (%)	NMAE (%)	nRMSE (%)
					Measured	ANN			
Jan	361.15	81.47	0.42	28.14	20.65	20.61	0.007	0.120	0.140
Feb	461.52	82.03	0.61	35.69	29.89	29.87	0.002	0.040	0.049
March	548.24	83.64	0.62	40.12	31.56	30.61	0.006	0.030	0.025
Apr	575.12	79.62	0.60	42.53	30.23	30.26	0.005	0.018	0.031
May	559.67	77.35	0.59	46.22	30.53	30.53	0.002	0.022	0.028
Jun	537.17	76.92	0.55	45.24	26.14	26.14	0.008	0.193	0.217
Jul	537.81	76.66	0.05	48.28	25.68	25.71	0.020	0.167	0.199
Aug	428.80	76.90	0.47	53.27	22.13	22.13	0.001	0.027	0.031
Sep	437.51	77.78	0.50	51.62	23.68	23.68	0.016	0.236	0.288
Oct	466.57	78.74	0.63	53.21	29.91	29.90	0.001	0.016	0.019
Nov	369.82	78.48	0.44	45.08	21.20	21.23	0.014	0.205	0.257
Dec	370.84	80.66	0.47	43.36	25.60	25.61	0.004	0.023	0.029
Avg.	471.18	79.19	0.50	44.40	26.43	26.36	0.007	0.092	0.109

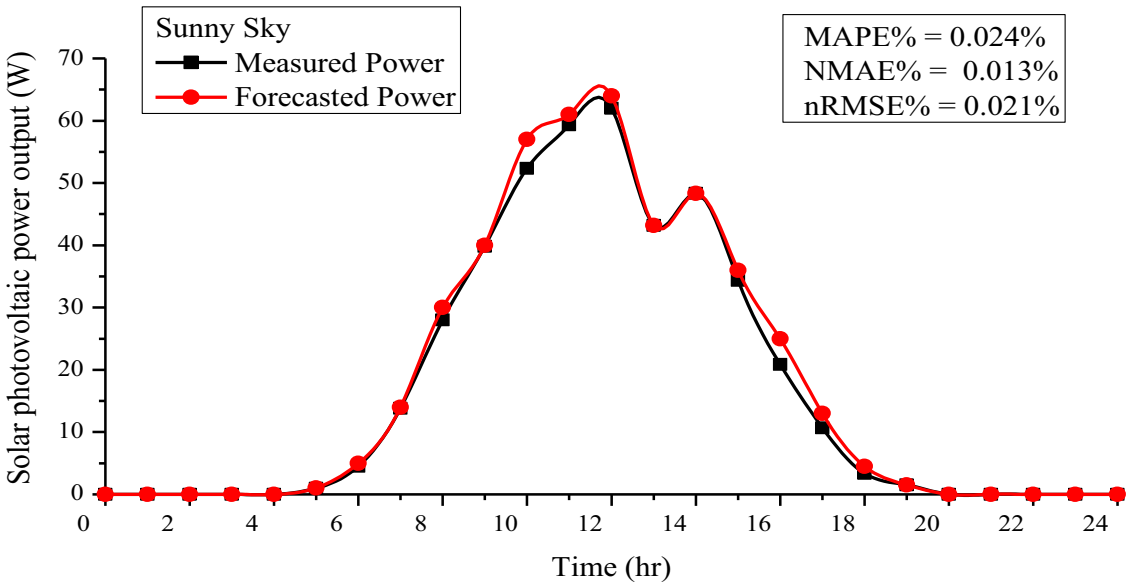


Figure 6. ANN-based model for short-term PV power forecast in a clear/sunny day.

Sunny Sky: 1st June was observed as a sunny day after evaluating the average of 5 years of data as the value of as diffuse solar energy is lower than 25% of global solar energy and the value of sunshine duration equals or greater than 9 hours. It is noticed that the sunny model gives the maximum power output of 16.92 W_p with average mean absolute percentage error of 0.024%, normalized mean absolute error of 0.013%, and normalized root mean square error of 0.021% as shown in Figure 6. In this sky-condition, time factor majorly affects the solar radiation than the factor of temperature.

Hazy Sky: Similarly, 26th December was observed as a hazy day after evaluating the average of 5 years of data as diffuse solar energy is lower than 50% of global solar energy and sunshine duration lies between 7 and 9 hours. For hazy sky model, the forecasted power is comparatively smaller than the sunny sky model with a maximum power output of 13.44 W_p, average mean absolute percentage error of 0.0019%, normalized mean absolute error of 0.361%, and normalized root mean square error of 0.7488% as presented in Figure 7. In this, the solar radiation gets clogged and temperature factor majorly affects the solar insolation in comparison to the factor of time, unlike sunny sky-condition.

Partly Cloudy Sky: Similarly, 3rd August was observed as a partly cloudy sky after evaluating the average of 5 years of data as the value of diffuse solar energy is lower than 75% of global solar energy and the sunshine hour lies between 5 and 7 hours. For partly cloudy sky, the power curve is biased during noon hours with a maximum power output of only 6.42 W_p, average mean absolute percentage

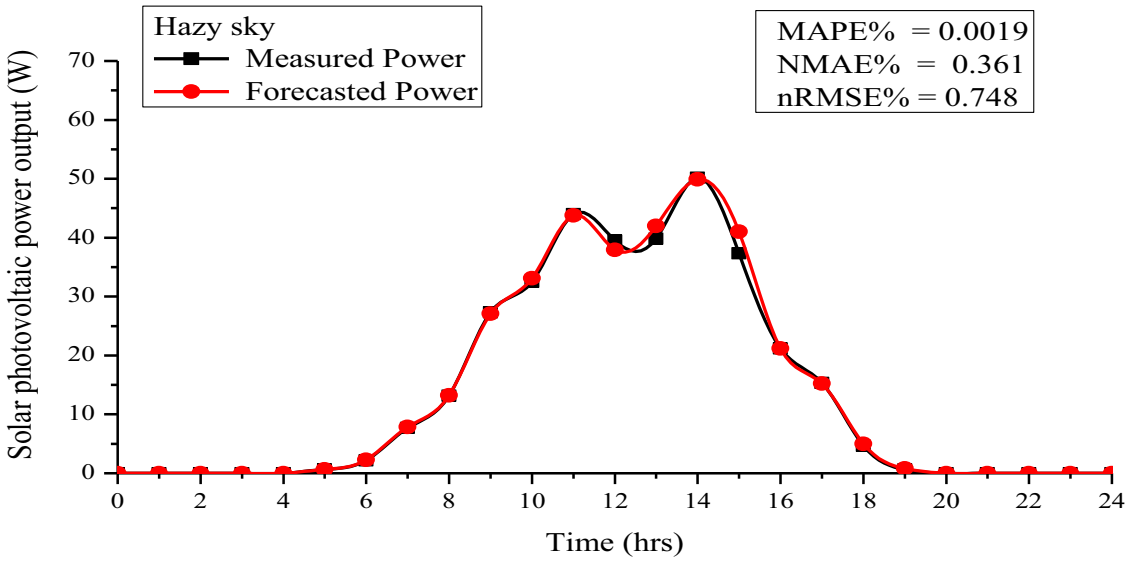


Figure 7. ANN-based model for short-term PV power forecast in a hazy day.

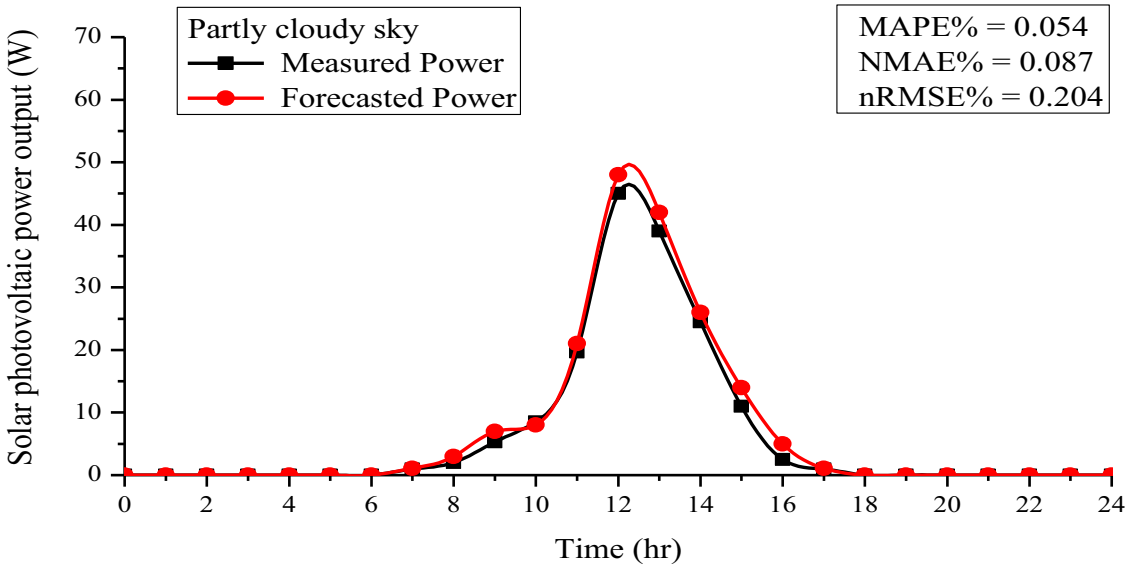


Figure 8. ANN-based model for short-term PV power forecast in a partly cloudy day.

error of 0.054%, normalized mean absolute error of 0.087%, and normalized root mean square error of 0.204% as shown in Figure 8. In this model, the sun rays are partly absorbed by the solar photovoltaic system and partly by cloud cover and both factors of time and temperature affect solar energy.

Fully Cloudy Sky: Similarly, 3rd January was observed as a fully cloudy sky after evaluating the average of 5 years of data as the value of diffuse solar energy is greater than 75% of global solar energy and sunshine duration is less than 5 hours. For fully cloudy sky, the power curve is biased during noon hours with a maximum power output of 2.97 W_p, mean absolute percentage error of 0.109%, normalized mean absolute error of 0.716%, and normalized root mean square error of 1.58% as presented in Figure 9. For this model, solar rays become fully obstructed by the presence of cloud; time and temperature factor affects the solar energy.

It has been observed from Figures 69 that the forecasted bell-shaped power curve on hourly basis shown by the red line matches the measured power represented by the black line using artificial neural network model. The duration considered is 24 hours. It is noticed that among the four models investigated, hazy model gives better performance with mean absolute percent error 0.0019% followed

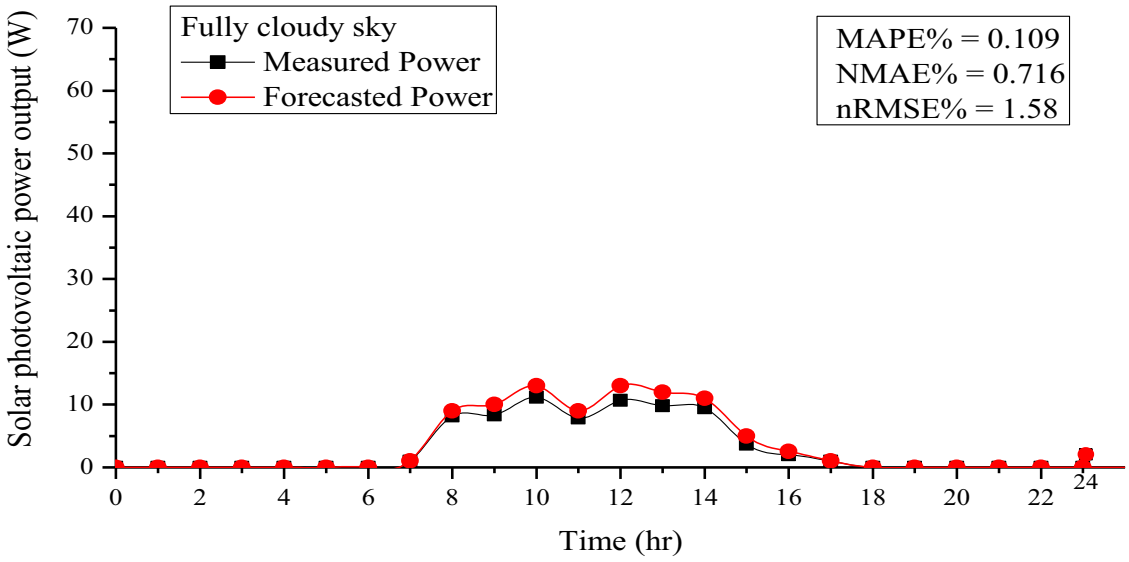


Figure 9. ANN-based model for short-term PV power forecast in a fully cloudy day.

by sunny, partly, and fully cloudy model with mean absolute percent error of 0.024%, 0.054%, and 0.109% respectively under composite climate zone.

Comparative analysis of the proposed model with other models

A comparison of the proposed model is done with other developed neural network models namely cascade-forward, elman, generalized regression, layered recurrent, linear layer neural network, fuzzy logic, and regression models for global solar energy forecasting and are presented in Table 11. The neural network can be summarized as:

Cascade-Forward Neural Network (CFNN): These models are the same as that of the feed-forward networks but such network creates a connection of its previous layer to the subsequent layers. Similar to the feed-forward neural network models, more than two layers of the cascade network have finite relation between input and output variables.

Elman Neural Network (ENN): These are feed-forward neural networks with the layered recurrent connection additions with tap delays. Such networks with more than one hidden layers can have a dynamic relation between input and output variables with sufficient number of neurons in the hidden layers. However, these networks make use of simple calculations at the cost of less reliable learning which results in a trade-off of reduced training calculations, but the risk of poorer accuracy.

Generalized Regression Neural Network (GRNN): These networks are used for function approximation. To closely fit the data, a spread smaller than the typical distance between input vectors are used and smoothly fitting the data, a larger spread has been used.

Layered Recurrent Neural Network (LRNN): It is same as that of the distributed delay neural networks with finite input responses. These networks with two and more layers can estimate dynamic output from past inputs with enough hidden neurons and recurrent layer delays.

Linear Neural Network (LNN): These are trained for modeling static and dynamic systems, with low learning rate to be stable.

Fuzzy Logic-based Models: Fuzzy logic-based models are established for comparison of the obtained results with the proposed model. Here, five membership functions are described with fuzzy-terms namely VVVL, VVL, VL, L, and LL lying in 0.1–0.9 range for sunshine duration meteorological parameter and certain rules are defined in Fuzzy Inference System (FIS) to forecast global solar energy as shown in Figure 10.

Table 11. Comparison of the proposed model with other developed neural network models.

Month	Data Actual H_g (MJ/m^2)	Forecasting	Proposed Model (RBFNN)	CFNN	ENN	GRNN	LRNN	LNN	Fuzzy	Regression
January	13.98	H_g (MJ/m^2)	13.95	13.98	13.98	13.91	13.93	13.95	13.85	14.02
		MAPE (%)	0.01	0.03	0.02	0.26	0.04	0.02	0.79	0.85
February	16.79	H_g (MJ/m^2)	16.79	16.76	16.8	16.76	16.82	16.79	16.62	16.79
		MAPE (%)	0.00	0.04	0.02	0.31	0.02	0.00	0.82	0.85
March	21.12	H_g (MJ/m^2)	21.12	21.08	21.16	21.30	21.1	21.12	21.12	21.11
		MAPE (%)	0.00	0.04	0.02	0.27	0.02	0.04	0.19	0.56
April	25.12	H_g (MJ/m^2)	25.21	25.31	25.15	25.26	25.21	25.22	25.16	25.25
		MAPE (%)	0.00	0.03	0.02	0.20	0.01	0.01	0.13	0.45
May	24.23	H_g (MJ/m^2)	24.23	24.21	24.24	24.28	24.25	24.23	23.58	24.22
		MAPE (%)	0.00	0.02	0.01	0.18	0.03	0.01	0.25	0.56
June	20.91	H_g (MJ/m^2)	20.91	20.88	20.91	21.17	20.86	20.91	21.23	20.93
		MAPE (%)	0.00	0.08	0.04	0.36	0.02	0.01	0.58	0.66
July	19.38	H_g (MJ/m^2)	19.38	19.48	19.36	19.31	19.32	19.38	19.53	19.38
		MAPE (%)	0.00	0.04	0.04	0.47	0.04	0.04	0.65	0.74
August	18.8	H_g (MJ/m^2)	18.80	18.94	18.75	18.95	18.78	18.8	18.98	18.85
		MAPE (%)	0.00	0.12	0.04	0.54	0.03	0.01	0.69	0.85
September	13.85	H_g (MJ/m^2)	13.85	13.87	13.86	13.83	13.85	13.85	13.89	13.9
		MAPE (%)	0.00	0.03	0.01	0.24	0.01	0	0.74	0.98
October	18.33	H_g (MJ/m^2)	18.33	18.4	18.31	18.36	18.33	18.33	18.72	18.36
		MAPE (%)	0.00	0.03	0.01	0.21	0.01	0.01	1.12	0.39
November	14.56	H_g (MJ/m^2)	14.56	14.59	14.56	14.59	14.57	14.56	14.48	14.58
		MAPE (%)	0.00	0.03	0.01	0.21	0.01	0.02	1.01	0.76
December	12.12	H_g (MJ/m^2)	12.12	12.26	12.2	12.21	12.13	12.13	12.3	12.12
		MAPE (%)	0.00	0.06	0.03	0.37	0.03	0.03	1.56	3.27
Avg.	18.27	H_g (MJ/m^2)	18.27	18.31	18.27	18.33	18.26	18.27	18.29	18.29
		MAPE (%)	0.001	0.050	0.020	0.300	0.020	0.020	0.711	0.910

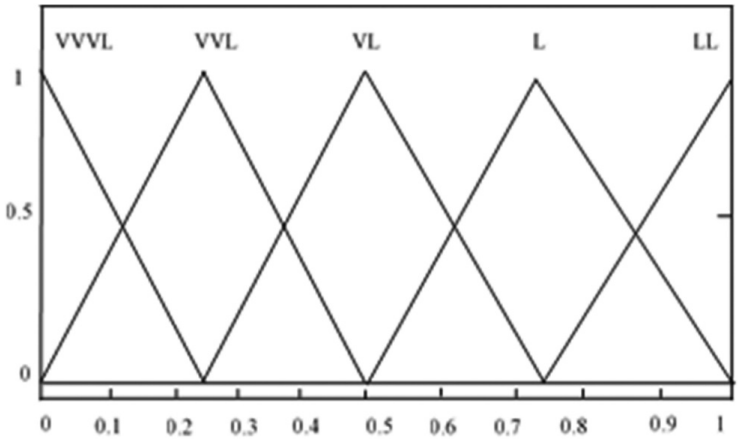


Figure 10. Membership functions defined for sunshine duration.

Fuzzy rules are then applied using MATLAB fuzzy logic toolbox for model development with parameters namely rel. humidity, sunshine duration, wind speed, dew-point, and ambient temperature for forecasting global solar energy as output and presented in Figure 11.

Empirical models: Empirical models are developed with input variables namely sunshine hours, amb. temp., wind speed, dew point, and rel. humidity whereas clearness index is the output (H_g/H_o). The forecasted H_g is achieved on multiplication of forecasted (H_g/H_o) with H_o, where H_g represents global solar energy and H_o represents extra-terrestrial solar energy which is obtained by standard geometric procedures. The proposed model is processed through Principal Component Analysis (PCA) so to obtain correlation with a high value of correlation coefficients. Some methodologies are used to define empirical relationships for solar energy forecasting. The value of H_o is expressed by using Eq. (8) as:-

), (8)

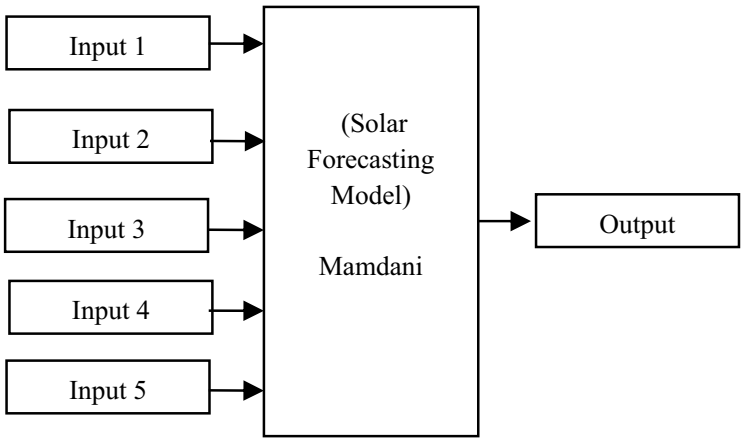


Figure 11. Development of Fuzzy logic models to forecast global solar energy.

where G_{sc} represents solar constant and equals 1367 Wm^{-2}, represents site latitude; represents mean sunrise angle; represents solar declination angle; denotes number of days in the year, n_{day} equals 1 for 1st Jan and 365 for 31st Dec.

It has been learnt from Table 11 that the proposed model using radial-basis-function model outperforms other developed neural network model. The overall analysis reveals that the error obtained in case of the proposed model is less with mean absolute percentage error of 0.001% whereas the mean absolute percentage error obtained for other developed neural network models such as CFNN is (0.050%), ENN (0.020%), GRNN (0.300%), LRNN (0.020%), LNN (0.020%), fuzzy (0.711%), and regression models (0.910%) and gives comparatively better accuracy as compared to other neural network models.

Conclusion

In the present work, artificial neural network models are established using the radial-basis-function and feed-forward neural network models for global solar energy forecasting. Simulations are achieved for distinct weather conditions namely sunny, hazy, partly, and fully cloudy sky-conditions which is applied to five meteorological stations across India. Model validation has been carried out w.r.t unidentified data set, and the overall analysis reveals that by employing radial-basis-function neural network model, there is a significant reduction in error for each of the climatic zones across India.

Further, it has been noticed that among the four models investigated, sunny sky model provides better performance for hot & dry climatic conditions with mean absolute percent error obtained of 0.0110×10^{-6}%, normalized mean absolute error of 0.0101×10^{-6}%, and normalized root mean square error of 0.01363%.

Similarly, hazy model gives favorable output for composite climate zone with mean absolute percentage error of 0.0126×10^{-6}%, normalized mean absolute error of 0.0105×10^{-6}%, and normalized root mean square error of 0.014784%; partly cloudy model is favorable for warm & humid climate zone with mean absolute percentage error of 0.0130×10^{-6}%, normalized mean absolute error of 0.0109×10^{-6}%, and normalized root mean square error of 0.02140%; and lastly, fully cloudy model provides favorable results for cold & cloudy climate zone with mean absolute percentage error of 0.0121×10^{-6}%, normalized mean absolute error of $0.0111\text{x}10^{-6}$%, and normalized root mean square error of 0.015369%.

The observed accuracy is found within the acceptable limits which validates the proposed model. One of its advantages is the better convergence rate, therefore, gives more accuracy in comparison to other developed neural network models. Several sites that have been used along with geographical range concludes that the developed model is valid to forecast global solar energy for distinct climate zone.

Further, the proposed model has been implemented for wider application series and the obtained results are further exploited for short-term photovoltaic power forecasting based on sky-conditions employing 210 W_p polycrystalline photovoltaic module under composite climatic conditions. The overall analysis reveals that among the four models investigated, hazy model gives better performance with mean absolute percentage error of 0.0019% followed by sunny, partly, and fully cloudy model with mean absolute percentage error of 0.024%, 0.054%, and 0.109% respectively under the composite climate zone. The model has been implemented for a wide range of application series.

Further, the developed model are compared with other developed neural network models and overall analysis reveals that the error obtained in case of the proposed model is lesser with mean absolute percentage error of 0.001% whereas the mean absolute percentage error obtained for other developed neural network models such as CFNN is (0.050%), ENN (0.020%), GRNN (0.300%), LRNN (0.020%), LNN (0.020%), fuzzy (0.711%), and regression models (0.910%). The radial-basis-function model provides better accuracy comparative to other developed neural network models.

Abbreviation

$f_{i,}$ i^{th} forecasted value (dimensionless); $G_{sc,}$ Solar constant (W/m^2); $G_{T,}$ Solar irradiance at STC (W/m^2); $H_{d,}$ Diffuse solar radiation (MJ/m^2); $H_{g,}$ Global solar radiation (MJ/m^2); $H_{o,}$ Extra-terrestrial solar radiation on horizontal surface (MJ/m^2); $I_{s,}$ Short-circuit current (A); $m_{i,}$ i_{th} measured value (dimensionless); $P_{PV,STC,}$ SPV system output (W); $P_{PV,}$ PV panel output (W); $N_{PVP,}$ PV panel in parallel (number); $N_{PVS,}$ PV panel in series (number); $T_{amb,}$ Ambient temp. (^oC); $T_{j,}$ PV panel temperature (^oC); x, Number of observations (dimensionless); x_{ij}, Incoming signal; $X_{max,}$ Maximum limit; $X_{min,}$ Minimum limit; $V_{oc,}$ Open circuit voltage (V); $w_{ij,}$ Connected weightage focused from j to i neuron.

ORCID

Gulnar Perveen http://orcid.org/0000-0001-6789-7455
M. Rizwan http://orcid.org/0000-0003-2839-7241

Nomenclature

ANN	Artificial Neural Network
CWET	Centre for Wind Energy Technology
IMD	Indian Meteorological Department
MAPE	Mean Absolute Percentage Error
NISE	National Institute of Solar Energy
NMAE	Normalized Mean Absolute Error
nRMSE	Normalized Root Mean Square Error
Greek Symbols	
	latitude of the site (^{o})
ω_s	mean sunrise hour angle (^{o})
n_{day}	number of days of the year starting from 1st January (dimensionless)
η_o	optical efficiency (%)
δ	solar declination angle (^{o})
γ	temperature parameter (dimensionless)
θ_i	i neuron bias (dimensionless)

References

Ahmadi, M. H., A. Baghban, M. Sadeghzadeh, M. Zamen, A. Mosavi, S. Shamshirband, R. Kumar, and M. Mohammad. 2020. Evaluation of electrical efficiency of photovoltaic thermal solar collector. *Engineering Applications of Computational Fluid Mechanics* 14 (1):545–65. doi:10.1080/19942060.2020.1734094.

Ahmadi, M. H., H. Jashnani, K. Chau, R. Kumar, and M. A. Rosen. 2019. Carbon dioxide emissions prediction of five Middle Eastern countries using artificial neural networks. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects.* doi:10.1080/15567036.2019.1679914.

Ajit, P. T. 2009. Handbook of solar radiant energy over India. *ndia Meteorological Department, Ministry of Earth Sciences, India* 270–2292.

Angstrom, A. 1924. Solar and terrestrial radiation. *Quarterly Journal of Royal Meteorological Society* 50:121–25. doi:10.1002/qj.49705021008.

Bansal, N. K., and G. Minke. 1988. Climatic zones and rural housing in India. In *Scientific Series of the International Bureau, Kernforschungsanlage Julich GmbH*, ed. N. K. BansaI and G. Minke, Vol. 35, 1–292. Julich, Germany: Zentralbibliothek.

Bayrakci, H. C., C. Demircan, and A. Kecebas. 2018. The development of empirical models for estimating global solar radiation on a horizontal surface: A case study. *Renewable and Sustainable Energy Reviews* 81 (P2):2771–82. doi:10.1016/j.rser.2017.06.082.

Chang, G. W., H. J. Lu, Y. R. Chang, and Y. D. Lee. 2017. An improved neural network-based approach for short-term wind speed and power forecast. *Renewable Energy* 105:301–11. doi:10.1016/j.renene.2016.12.071.

Chel, A., and G. N. Tiwari. 2011. A case study of a typical 2.32 kW stand-alone photovoltaic (SAPV) in a composite climate of New Delhi (India). *Applied Energy* 88:1415–26. doi:10.1016/j.apenergy.2010.10.027.

Dhimish, M., V. Holmes, B. Mehrdadi, and M. Dales. 2018. Comparing mamdani sugeno fuzzy logic and RBF ANN network for PV fault detection. *Renewable Energy* 117:257–74. doi:10.1016/j.renene.2017.10.066.

Dolara, A., F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari. 2018. Comparison of training approaches for photovoltaic forecasts by means of machine learning. *Applied Sciences 8(2), Article ID: 228*. 1–16. doi:10.3390/app8020228.

Feng, Y., Y. Peng, N. Cui, D. Gong, and K. Zhang. 2017. Modelling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. *Computers and Electronics in Agriculture* 136:71–78. doi:10.1016/j.compag.2017.01.027.

Ghimire, S., R. C. Deo, N. J. Downs, and N. Raj. 2019. Global solar radiation prediction by ANN integrated with European centre for medium range weather forecast fields in solar rich cities of Queensland Australia. *Journal of Cleaner Production* 216:288–310. doi:10.1016/j.jclepro.2019.01.158.

Gouda, G. S., Z. Hussein, S. Luo, and Q. Yuan. 2020. Review of empirical solar radiation models for estimating global solar radiation of various climate zones of China. *Progress in Physical Geography: Earth and Environment* 44 (2):168–88. doi:doi.10.1177/0309133319867213.

Kasaeian, A., G. Mehran, M. H. Ahmadi, and G. Mehrdad. 2017. GMDH algorithm for modeling the outlet temperatures of a solar chimney based on the ambient temperature. *Mechanics & Industry 18 (2), Article ID: 216* 1–10. doi:doi-10.1051/meca/2016034.

Khalil, S. A., and H. A. S. Aly. 2018. Comparative and evaluate of empirical models for estimation of global solar radiation in Al-Baha, KSA. *Journal of Earth Science and Climate Change* 9 (9):1–10.

Kuhe, A., V. T. Achirgbenda, and M. Agada. 2019. Global solar radiation prediction for Makurdi, Nigeria using neural networks ensemble. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects.* doi:10.1080/15567036.2019.1637481.

Kumar, K. R., and M. S. Kalavathi. 2018. Artificial intelligence based forecast models for predicting solar power generation. *Proceedings on Materials Today* 5 (1):796–802. doi:10.1016/j.matpr.2017.11.149.

Loni, R., A. Kasaeian, K. Shahverdi, E. A. A. Ardeh, B. Ghobadian, and M. H. Ahmadi. 2017. ANN model to predict the performance of parabolic dish collector with tubular cavity receiver. *Mechanics & Industry 18 (4), Article ID: 408* 1–8. doi:doi-10.1051/meca/2017016.

Mahmood, F., M. Behnam, A. Ahmad, M. H. Ahmadi, and M. A. Sheremet. 2020. Precise prediction of biogas thermodynamic properties by using ANN algorithm. *Renewable Energy* 147 (1):179–91. doi:10.1016/j.renene.2019.08.112.

Mehdizadeh, S., J. Behmanesh, and K. Khalili. 2017. Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data. *Environmental Earth Sciences 76, Article ID* - 325:1–16. doi:doi.10.1007/s12665-017-6607-8..

Mehran, G., A. Kasaeian, M. H. Ahmadi, and G. Mehrdad. 2017. Evolving ICA and HGAPSO algorithms for prediction of outlet temperatures of constructed solar chimney. *International Journal of Low-Carbon Technologies* 12 (2):84–95.

Nespoli, A., E. Ogliari, S. Leva, A. M. Pavan, A. Mellit, V. Lughi, and A. Dolara. 2019. Day-Ahead photovoltaic forecasting: A comparison of the most effective techniques. *Energies* 12:1–15. doi:10.3390/en12091621.

Notton, B., A. Fouilloy, J. L. Duchaud, C. Voyant, and R. Dizene. 2019. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. *Renewable Energy* 132:871–84. doi:10.1016/j.renene.2018.08.044.

Olusola, B., O. Ariyo, O. Sandra, O. Emmanuel, A. Gaylord, and H. Qi. 2020. Application of deep learning for solar irradiance and solar photovoltaic multi-parameter forecast. *Energy Sources Part A Recovery Utilization and Environmental Effects*. doi:10.1080/15567036.2020.1801903.

Perveen, G., M. Rizwan, and N. Goel., 2018a. Development of empirical models for forecasting global solar energy. 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES-2018), Delhi, India, 317–24.

Perveen, G., M. Rizwan, and N. Goel. 2018b. Intelligent model for solar energy forecasting and its implementation for solar photovoltaic applications. *Journal of Renewable and Sustainable Energy* 10 (6):1–23, Article ID - 063702.

Panapakidis, I. P., and A. S. Dagoumas. 2016. Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy* 172:132–51. doi:10.1016/j.apenergy.2016.03.089.
Perveen, G., M. Rizwan, and N. Goel. 2019. Comparison of intelligent modelling techniques for solar energy forecasting and its application in solar PV systems. *IET Energy Systems Integration* 1 (1):34–51. doi:10.1049/iet-esi.2018.0011.
Puneet, K., S. Nidhi, and M. A. Ansari. 2017. Solar radiation forecasting using artificial neural network with different meteorological variables. *Communication and Computing System* 487–91. doi:10.1201/9781315364094-88.
Rohani, A., M. Taki, and M. Abdollahpour. 2018. A novel soft computing model (Gaussian Process Regression with k-fold cross-validation) for daily and monthly solar radiation forecasting. *Renewable Energy* 115:411–22. doi:10.1016/j.renene.2017.08.061.
Sen, Z. 1998. Fuzzy algorithm for estimation of solar irradiation from sunshine duration. *Solar Energy* 63 (1):39–49. doi:10.1016/S0038-092X(98)00043-7.
Ustun, I., F. Unes, I. Mert, and C. Karakus. 2020. A comparative study of estimating solar radiation using machine learning approaches: DL, SMGRT, and ANFIS. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 1–24. doi:10.1080/15567036.2020.1781301.
Yan, H., Y. Mou, X. Xu, J. Du, R. Wang, and P. Liu. 2018. A comparative assessment of predicting CH_4 adsorption on different activated carbons using Generalized Regression Neural Network (GRNN), and Adaptive Network-Based Fuzzy Inference System (ANFIS). *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 41 (16):-1983–92. doi:10.1080/15567036.2018.1548527.
Yildiz, C., and H. Acikgoz. 2020. A kernel extreme learning machine-based neural network to forecast very short-term power output of an on-grid photovoltaic power plant. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 1–18. doi:10.1080/15567036.2020.1801899.
Zamen, M., A. Baghban, S. M. Pourkiaei, and M. H. Ahmadi. 2019. Optimization methods using artificial intelligence algorithms to estimate thermal efficiency of PV/T system. *Energy Science & Engineering* 7 (3):821–34. doi: Doi:10.1002/ese3.312.

Appendix A

Table A1. Training and testing data set for CWET Chennai (warm & humid climatic conditions).

Data Set	Parameter	Maximum	Minimum	Mean	St. Deviation
Training	H_g (MJ/m^2)	21.83	12.78	17.56	2.45
	Amb. Temp. (^{o}C)	36.79	37.05	25.35	2.35
	Sunshine Hours (hrs)	10.89	3.10	8.83	1.97
	Wind Velocity (m/s)	12.18	5.56	8.13	1.49
	Rel. Hum. (%)	81.43	60.48	71.21	5.45
	Dew-Point (^{o}C)	21.85	18.94	20.42	0.69
Testing	H_g (MJ/m^2)	20.69	6.06	13.79	4.24
	Amb. Temp. (^{o}C)	34.89	23.63	26.16	2.36
	Sunshine Hours (hrs)	10.46	3.78	7.18	2.85
	Wind Velocity (m/s)	15.97	3.45	9.63	3.70
	Rel. Hum. (%)	88.36	66.26	78.25	5.16
	Dew-Point (^{o}C)	23.24	19.28	21.28	1.18

Table A2. Training and testing data set for IIT Jodhpur (hot & dry climatic conditions).

Data Set	Parameter	Maximum	Minimum	Mean	St. Deviation
Training	H_g (MJ/m^2)	19.24	11.57	16.35	1.56
	Amb. Temp. (^{o}C)	22.23	14.66	18.12	1.58
	Sunshine Hours (hrs)	10.65	4.45	9.27	1.41
	Wind Velocity (m/s)	12.73	3.49	6.48	4.36
	Rel. Hum. (%)	79.73	25.19	45.41	9.42
	Dew-Point (^{o}C)	6.15	−2.44	2.34	2.52
Testing	H_g (MJ/m^2)	16.90	13.75	14.71	0.75
	Amb. Temp. (^{o}C)	21.28	14.75	18.58	1.75
	Sunshine Hours (hrs)	9.300	2.22	8.63	1.31
	Wind Velocity (m/s)	7.750	2.89	3.39	1.56
	Rel. Hum. (%)	56.79	37.96	48.42	5.08
	Dew-Point (^{o}C)	11.82	−13.72	0.35	5.57

Table A3. Training and testing data set for SEC, Gurgaon (composite climatic conditions).

Data Set	Parameter	Maximum	Minimum	Mean	St. Deviation
Training	H_g (MJ/m^2)	17.78	11.15	13.53	1.25
	Amb. Temp. (^{o}C)	17.86	10.63	14.36	1.85
	Sunshine Hours (hrs)	10.20	2.55	7.85	1.78
	Wind Velocity (m/s)	11.86	2.96	5.28	3.28
	Rel. Hum. (%)	81.12	56.68	65.69	6.26
	Dew-Point (^{o}C)	8.89	−15.25	−4.57	5.66
Testing	H_g (MJ/m^2)	15.45	7.36	12.35	2.23
	Amb. Temp. (^{o}C)	19.28	13.38	15.86	1.74
	Sunshine Hours (hrs)	8.43	3.78	5.39	2.55
	Wind Velocity (m/s)	11.47	2.57	5.03	2.96
	Rel. Hum. (%)	87.69	45.52	65.79	8.78
	Dew-Point (^{o}C)	9.52	−5.12	0.23	3.71

Table A4. Training and testing data set for Karad, Pune (moderate climatic conditions).

Data Set	Parameter	Maximum	Minimum	Mean	St. Deviation
Training	H_g (MJ/m^2)	18.85	14.25	16.38	1.28
	Ambient Temperature (^{o}C)	21.32	17.14	19.89	0.47
	Sunshine Hours (hrs)	10.02	8.78	9.56	0.39
	Wind Velocity (m/s)	2.25	1.53	1.50	0.78
	Rel. Hum. (%)	68.24	52.22	59.93	4.63
	Dew-Point (^{o}C)	17.46	5.96	11.63	3.86
Testing	H_g (MJ/m^2)	18.85	10.55	17.25	1.82
	Amb. Temp. (^{o}C)	22.15	16.36	19.45	1.63
	Sunshine Hours (hrs)	9.78	4.22	8.87	1.79
	Wind Velocity (m/s)	3.99	1.13	2.59	0.98
	Rel. Hum. (%)	75.35	55.48	63.12	4.96
	Dew-Point (^{o}C)	19.46	10.15	16.45	2.15

Table A5. Training and testing data set for Meghalaya (cold & cloudy climatic conditions).

Data Set	Parameter	Maximum	Minimum	Mean	St. Deviation
Training	H_g (MJ/m^2)	19.84	3.79	14.54	3.25
	Ambient Temperature (^{o}C)	12.75	5.70	9.63	1.24
	Sunshine Hours (hrs)	9.02	5.15	7.25	2.36
	Rel. Hum. (%)	97.63	60.25	75.75	8.63
	Wind Velocity (m/s)	11.05	2.12	3.83	2.78
	Dew-Point (^{o}C)	12.41	1.19	6.52	2.59
Testing	H_g (MJ/m^2)	18.88	9.14	15.71	2.25
	Ambient Temperature (^{o}C)	15.45	10.48	11.91	1.22
	Sunshine Hours (hrs)	9.28	4.63	7.79	1.59
	Rel. Hum. (%)	89.15	51.25	74.35	9.28
	Wind Velocity (m/s)	4.78	1.36	2.25	1.24
	Dew-Point (^{o}C)	14.36	5.26	10.49	2.36

ATT: Attention-based Timbre Transfer

Deepak Kumar Jain*, Akshi Kumar†, Linqin Cai§, Siddharth Singhal†‡ and Vaibhav Kumar†‡
*Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing,
College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China.
§Key Laboratory of Industrial Internet of Things and Networked Control, Ministry of Education,
Chongqing University of Posts and Telecommunications, Chongqing 400065, China.
†Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India.

***Abstract*—In this paper, we tackle the issue of timbre transfer on a given monophonic music sample. The objective is to change the timbre of source audio from one instrument to another while preserving features such as loudness, pitch, and rhythm. Existing approaches use image-to-image translation techniques on the entire region of time-frequency representations of the raw audio wave, which may lead to the addition of unwanted elements in the final audio waveform. We propose Attention-based Timbre Transfer (ATT), an attention-based pipeline for transferring timbre. To the best of our knowledge, ATT is the first approach which leverages attention for achieving timbre transfer. Further, ATT uses MelGAN for spectrogram inversion, which provides a fast and parallel alternative to other autoregressive music generation approaches, without compromising on the quality. ATT shows promising results, thus efficaciously transferring timbre with minimal offset to other physical characteristics.**

***Index Terms*—Generative Adversarial Networks, Timbre transfer, Style transfer, mel-frequency spectrogram.**

I. Introduction

Music is composed of different elements such as rhythm, dynamics, texture, timbre and melody together creating both short and long-range structure. One of the distinctive features of a musical sound is timbre. The American Standards Association (1960) define timbre as "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar". They further add, "Timbre depends primarily upon the frequency spectrum, although it also depends upon the sound pressure and the temporal characteristics of the sound", thereby making timbre a multidimensional entity and modelling it a challenging task.

In this paper, we propose a novel attention-based approach for timbre transfer, i.e. given a monophonic music sample, we aim to transfer the timbre of source audio from one instrument to another while preserving features such as loudness, pitch and rhythm. For instance, given a music sample composed using a violin, we aim to obtain an audio waveform for the same composition but with the instrument being a piano. To the best of our knowledge, this is the first application of attention-guided image-to-image translation for transferring timbre.

Further, to convert the translated time-domain representation back to audio, multiple methods can be used. One such method is Griffin-Lim (GL) (Griffin Lim, 1984) which, however, is known to introduce disturbances in the output wave [1]. Recently, neural network-based models such as WaveNet [2] and SampleRNN [3] have been highly successful in generating high-quality audio. However, the inference time of these models is high as the audio needs to be generated sequentially, making them inefficient for real-time usage. Therefore, we have used MelGAN [4], a Generative Adversarial Network based approach for raw audio generation which uses Mel-frequency spectrogram. MelGAN is faster than other Mel-frequency spectrogram inversion methods and can be used for generating audio in parallel. Therefore, making it ideal for real-time applications.

Building upon the above components, we propose attention-based timbre transfer (ATT). For training, ATT uses non-parallel monophonic sound samples. For time-frequency representation, we have used Mel-frequency spectrogram; this is because Mel-frequency distribution is done in a perceptually motivated manner. Empirically, ATT shows promising results, successfully transferring timbre on a given set of monophonic sounds, while preserving other characteristics such as pitch and loudness. Additionally, it achieves high score on a test aimed at classifying musical instrument based upon time-frequency representation.

II. Related works and motivations

A. *Generative Adversarial Network (GAN)*

Generative adversarial network (GAN) [5] is a technique for implicitly learning the distribution of a dataset. The task is formed as a zero-sum game between a generator network(G) and a discriminator network(D). During the training process, G and D are trained sequentially. First, the discriminator learns to minimize the classification loss by assigning a high score to natural data samples and a low score to synthetic data samples. Second, the generator learns to synthesize synthetic data samples such that the score assigned by the discriminator to them increases. The objective function is defined as follows:

$$G^*, D^* = \arg\min_G \max_D -\frac{1}{2}E_{x\sim p_{\text{data}}(x)}[\log D(x)] - \frac{1}{2}E_{z\sim p_z(z)}[\log(1 - D(g(z)))]$$

Where, $p_{\text{data}}(x)$ denotes the real data distribution, $p_z(z)$ denotes a prior distribution such as Gaussian distribution. GANs

‡Authors have contributed equally.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

have been successfully applied to a range of tasks such as image super-resolution [6], image-to-image translation [7], and text-to-image synthesis [8].

B. Unpaired image-to-image translation

Unsupervised image-to-image translation techniques aim to create a mapping between the source and the target domains without the presence of paired data samples. In their seminal work, [9] proposed CycleGAN, a generative adversarial network based approach which uses cyclic consistency for learning the desired mapping between two given domains of data. For the source domain X and the target domain Y. The model learns the mappings $F_{Y\to X} : Y \to X$ and $F_{X\to Y} : X \to Y$ along with adversarial discriminators D_X and D_Y, such that D_X distinguishes between the elements of X and translated samples created by $F_{Y\to X}$ and D_Y distinguishes between the elements of Y and artificial samples created by $F_{X\to Y}$. The objective function is given as :

$$\begin{aligned}\mathcal{L}\left(F_{X\to Y}, F_{Y\to X}, D_X, D_Y\right) = \mathcal{L}^x_{\mathrm{GAN}}\left(F_{X\to Y}, D_Y\right) \\ + \mathcal{L}^y_{\mathrm{GAN}}\left(F_{Y\to X}, D_X\right) \\ + \lambda\mathcal{L}_{\mathrm{cyc}}(F_{X\to Y}, F_{Y\to X})\end{aligned}$$

where,

$$\begin{aligned}\mathcal{L}^x_{\mathrm{GAN}}\left(F_{X\to Y}, D_Y\right) = \mathbb{E}_{y\sim p_{\mathrm{data}}(y)}\left[\log D_Y(y)\right] \\ + \mathbb{E}_{x\sim p_{\mathrm{data}}(x)}\left[\log\left(1 - D_Y(F_{X\to Y}(x)\right]\end{aligned}$$

$$\begin{aligned}\mathcal{L}_{\mathrm{cyc}}(F_{X\to Y}, F_{Y\to X}) = \mathbb{E}_{x\sim p_{\mathrm{data}}\,(x)}\left[\|F_{Y\to X}(F_{X\to Y}(x)) - x\|_1\right] \\ + \mathbb{E}_{y\sim p_{\mathrm{data}}\,(y)}\left[\|F_{X\to Y}(F_{Y\to X}(y)) - y\|_1\right]\end{aligned}$$

Building upon CycleGAN, [10] propose an attention-guided image-to-image translation technique. The main advantage of using attention-guided architecture is that mapping mechanism translates only those parts of the source image sample which are the most relevant for the translation operation, minimizing changes to any other elements in the background. They introduce attention networks in generators of CycleGAN architecture. The attention networks are trained in tandem with the generator networks. These attention networks are then used to create attention maps for regions which the discriminator networks 'perceives' to be the most discriminative between the given source and target domains. We delineate upon the architectural details in the next section.

C. Audio style transfer

Similar to unpaired image-to-image translation, GAN-based methods have also been applied for unpaired audio-to-audio timbre transfer. [11] proposed Timbretron, a timbre transfer pipeline which uses log amplitude Constant-Q-Transform (CQT) for the time-frequency representation of the audio-waveform. A translation between the source and the target CQT is then performed using CycleGAN [9]. Finally, they use WaveNet [2] for generating the target audio waveform from the CQT representation. Since, Timbretron uses CycleGAN for image-to-image translation, unwanted artifacts might be introduced in the target representation. Further, as WaveNet generates audio waveform in a non-parallel manner, it is unable to scale to real-time environments.

GAN based methods have been further applied to voice transfer applications and singing voice separation [12]. Additionally, [13] propose an encoder-decoder approach for successfully translating music across different instruments and styles.

D. Audio generation from time-frequency representation

This process involves the creation of audio waveform from its corresponding time-frequency representation such as the Mel-frequency spectrogram. Various techniques have been proposed for the inversion of spectrogram to audio; we briefly highlight them here. One of the most popular approaches for inverting a short-time Fourier transform is Griffin-Lim (Griffin Lim, 1984) method. Griffin-Lim method iteratively estimates the unknown phases by continuously converting back and forth between the time and frequency domain by using short-time Fourier transform and its inverse, replacing the magnitude of each frequency component by that of the predicted magnitude in each iteration. While simplistic, the approach is known to introduce powerful artificial features in the final output as observed by [1].

Wavenet [2] is a fully convolutional, autoregressive model that is capable of producing life-like speech samples. Further, it can also generate high-quality music samples. However, audio samples are required to be generated in a sequential manner which makes the speed of these methods unsuitable for real-time applications.

Recently, GAN based methods have also been used for audio generation. One such approach is MelGAN [4]. MelGAN is a lightweight and parallel GAN architecture for conditional audio synthesis. MelGAN is capable of producing high-quality audio-waveforms parallelly, therefore, making it suitable for real-time applications. Hence, We have employed MelGAN in our approach for constructing the audio waveform from the translated Mel-frequency spectrogram representation.

III. Proposed method

The aim of audio timbre transfer is to construct a mapping from a source audio domain X to a destination audio domain Y such that properties such as pitch and loudness are preserved, while transforming the timbre of the instrument in X to that of target instrument in Y. To this end, we propose **ATT**: **A**ttention-based **T**imbre **T**ransfer pipeline. In this work, we have only considered monophonic sound samples. We intend to work on polyphonic sound samples in the future.

As shown in figure 1, the working of ATT can be divided into three stages:

1) Convert the raw input audio waveform into its equivalent Mel-frequency spectrogram representation.
2) Considering the Mel-frequency spectrogram representation as an image, we model the timbre transfer as an image-to-image translation problem and perform unsupervised attention-guided operation proposed by [10].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

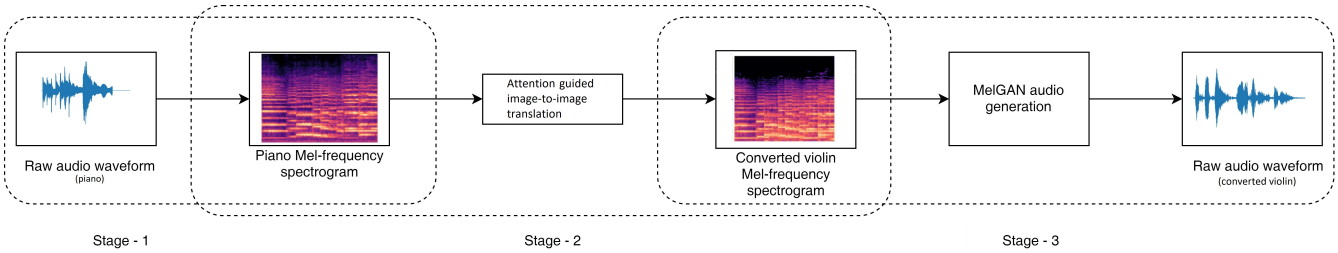


Fig. 1. ATT: Attention-based timbre transfer

3) Finally, the image resulting from the image-to-image translation operation is used as the source in the audio synthesis process using MelGAN [4]. In the subsequent sections, we describe each of these stages in detail.

A. Music representation using Mel-frequency spectrogram

An audio waveform is a variation of pressure with respect to time, carrying the information about changes in amplitude with time. However, utilizing crude audio wave for modelling sound is practically infeasible. This is because crude audio waves are usually sampled at a very high frequency (usually 16Khz or above) for an enhanced temporal resolution, thus, creating a deluge of data, making it unsuitable for modelling purposes. Therefore, modelling methodologies operate upon a simplified representation of the original audio waveform. This representation is lower in resolution than the initial raw wave, but it is easier to work with. In our approach, we have used the Mel-frequency spectrogram for representing the sound. Mel-frequency scale is obtained by the application of non-linear transformation on the linear frequency axis of a spectrogram obtained from short-term Fourier transform (STFT). The transformation results in a frequency scale in which frequencies are represented in a manner which is perceived by the human auditory system, i.e. higher resolution is given to low frequencies and lower resolution is given to high frequencies. Properties such as these have made Mel-frequency spectrogram an attractive choice for training neural networks on large scale audio corpus [14].

B. Timbre transfer on mel-frequency spectrogram Representation

We have employed the attention-guided architecture proposed by [10] for image-to-image translation on the Mel-frequency spectrogram. The advantage of this approach over other methods such as CycleGAN [9] and DualGAN [15] is its ability to discern the most discriminative parts of the image, thereby making changes only to the most relevant parts and adding minimal or no changes to the background. It achieves its objective by creating foreground and background attention maps.

By using attention maps on Mel-frequency spectrograms, we can focus on only those areas of the spectrogram which are most important for discerning the characteristics of the given audio wave. Therefore, When subjected to a image-translation procedure, changes are brought about in only these regions. This reduces the addition of unnecessary features which may deteriorate the quality of audio wave which is obtained from the resulting spectrogram. We now discuss our approach based upon the methodology proposed by [10] in detail.

Having obtained the Mel-frequency spectrogram for the source audio domain X and the target audio domain Y. The attention networks are referred to as A_x and A_y respectively. Each attention network creates attention maps which are denoted as follows: $A_X : X \rightarrow X_a$, $A_Y : Y \rightarrow Y_a$ where X_a, Y_a refer to the attention maps induced by X and Y respectively and consist of [0,1] values per-pixel.

For an input Mel-frequency spectrogram $x \in X$ the transfered representation using the mapping function $F_{X \rightarrow Y}$ and the attention network A_X is defined as :

$$x' = x_a \odot F_{X \rightarrow Y}(x) + (1 - x_a) \odot x \tag{1}$$

where, x_a denotes the attention map $A_X(x)$ and $\odot$ denotes an element-wise product. In the expression above, the former part of the sum denotes the foreground of the image (which is focused upon by the attention map) and the latter part of the sum refers to the background. Figure 2 shows the procedure on actual Mel-Frequency spectrograms. The adversarial loss is then defined as:

$$\mathcal{L}^x_{GAN}\left(F_{X \rightarrow Y}, A_X, D_Y\right) = \mathbb{E}_{y \sim \mathrm{P}_{data}(y)}\left[\log\left(D_Y(y)\right)\right] + \mathbb{E}_{x \sim \mathrm{P}_{data}(x)}\left[\log\left(1 - D_Y\left(x'\right)\right)\right]$$

Further, similar to [9], [10] introduce a cyclic consistency loss which is defined as follows:

$$\mathcal{L}^x_{cyc}\left(x, x''\right) = \left\|x - x''\right\|_1$$

where, x'' is obtained from x' via $F_{Y \rightarrow X}$ and A_Y in a same manner as equation 1.

The final objective function is then defined as:

$$\mathcal{L}\left(F_{X \rightarrow Y}, F_{Y \rightarrow X}, A_X, A_Y, D_X, D_Y\right) = \mathcal{L}^x_{GAN} + \mathcal{L}^y_{GAN} + \lambda_{cyc}\left(\mathcal{L}^x_{cyc} + \mathcal{L}^y_{cyc}\right) \tag{2}$$

During our training procedure we set the value λ_{cyc} as 10. Further, the input to the discriminator network is also modified using the attention maps so that it consider only attention mapped regions in the synthetic and the real image distribution. Lastly, [10] threshold the learned attention maps in order to prevent attention maps from interfering with the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

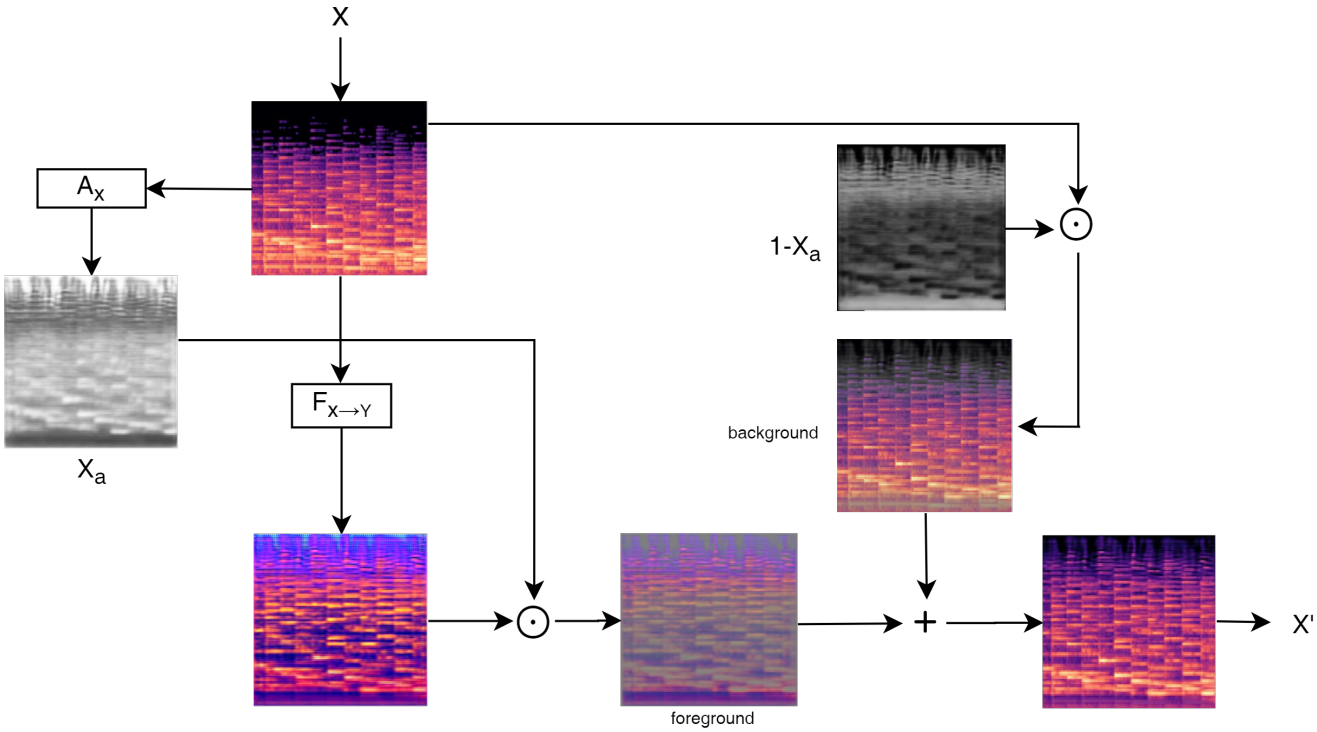


Fig. 2. The generation of output using the mapping function $F_{X\to Y}$ and the attention network A_X. The network works in a symmetric manner for domain Y, mapping function $F_{Y\to X}$, and attention network A_Y.

representation of actual images. We use the threshold value as 0.1 as proposed in the original setup. For further details regarding the architecture, the readers may refer [10].

C. Raw audio waveform generation from mel-frequency spectrogram

After processing the input Mel-frequency spectrogram under the methodology described in the previous section, we obtain the Mel-frequency spectrogram of the audio waveform of the timbre transferred audio. However, the absence of phase information makes it non-trivial to convert the resultant spectrogram into an audio waveform. Therefore, we have used MelGAN for obtaining the audio waveform. MelGAN comprises of generative adversarial networks to produce high-quality audio waveforms given a Mel-frequency spectrogram.

The objective minimized during the training process, as proposed by [4], are defined as follows:

$$\min_{D_k} \mathbb{E}_x \left[\min\left(0, 1 - D_k(x)\right)\right] + \mathbb{E}_{s,z}\left[\min\left(0, 1 + D_k(G(s,z))\right)\right], \forall k = 1,2,3 \quad (3)$$

$$\min_{G}\left(\mathbb{E}_{s,z}\left[\sum_{k=1,2,3} -D_k(G(s,z))\right] + \lambda \sum_{k=1}^{3} \mathcal{L}_{\mathrm{FM}}\left(G, D_k\right)\right) \quad (4)$$

where, G denotes the generator network, D denotes the discriminator network, x is the raw audio waveform, z denotes Gaussian noise vector and s denotes the conditional information (the input Mel-frequency spectrogram) and k denotes the index of the respective discriminator network. $\mathcal{L}_{\mathrm{FM}}$ denotes the feature matching objective, used for reducing the L1 distance of discriminator feature maps of synthetic and real audio. we use $\lambda = 10$ during our training procedure. For further details, the readers may refer [4].

IV. Experiments and Results

A. Datasets

We have used two datasets, namely, MIDI-BACH and real-world dataset for the training and evaluation of our proposed approach.

MIDI-BACH: It is a catalog of J.S Bach's compositions using different Instruments. We use the audio data generated from piano and violin for our analysis.

Real World Dataset: We collected two hours of real world recording from Youtube videos for each instrument [16] [17] [18] [19]. We subsequently converted these video files to audio files. Real World dataset helps us in creating a robust model, effective against noise inherent to real-world audio samples, unlike the MIDI-BACH

These recordings from both the datasets were segmented into six seconds of audio samples. Mel-frequency spectrograms generated for these audio samples were used for training and testing our model. For generating spectrograms, we used 512

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

samples between successive frames and FFT window of length 2048. The resolution of each spectrogram was [432*432*3]. We generated 2000 Mel-frequency spectrograms for each instrument. We further split this dataset into training, validation, and test set using a 60:20:20 split.

We used an initial learning rate of 0.00005 for the complete network with Adam [20] optimizer for training our network.

B. Evaluation of results

Figures 3 and 6 depict the Mel-frequency spectrograms generated for the audios recorded using piano. These spectrograms were fed as input to our model. Figures 5 and 8 represent the spectrograms for violin generated after attention guided image-to-image translation. [10], using figures 3 and 6 respectively as input. Figures 4 and 7 are the attention maps generated for our first and second input respectively.

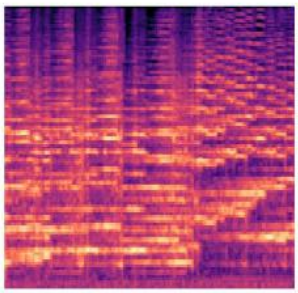


Fig. 3. Input spectrogram of piano: MS_{P1}

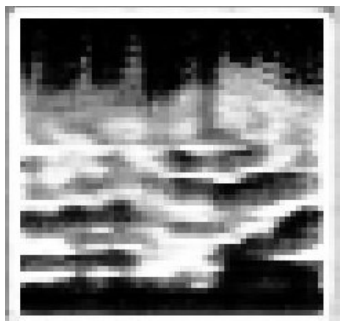


Fig. 4. Generated attention map for MS_{P1}

*1) **Convolutional Neural Network based evaluation**:* We trained a Convolutional Neural Network for identifying the instrument used for generating the input Mel-frequency spectrogram. The classifier was trained on Mel-frequency spectrograms obtained using 1400 audio samples for each instrument category, i.e., piano and violin.

Table I provides a brief overview of the architecture of CNN used during the quality evaluation phase. RGB images [128*128*3] of Mel-frequency Spectrograms were fed as input to the model, and the output was the label for each image.

Our model achieves an accuracy of 99.2% during the testing phase. Figure 9 depicts the Receiver Operating Characteristic

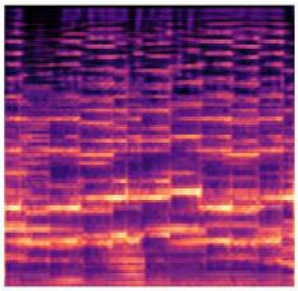


Fig. 5. Generated spectrogram of violin: MS_{P1V}

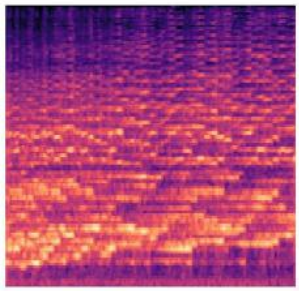


Fig. 6. Input spectrogram of piano: MS_{P2}

curve plot obtained for our Convolutional Network during the testing phase. This represents the high accuracy of our Convolutional Neural Network in identifying the source instrument for the input Mel-frequency Spectrogram.

This CNN was subsequently used for predicting the class labels for our generated images. We obtained an accuracy of 98.89% while predicting the class of our generated spectrograms for a given target label, indicating the good quality and superiority of our proposed architecture in performing timbre transfer.

*2) **MelGAN based evaluation**:* We used MelGAN for reconstructing audio waveform from our generated Mel-frequency spectrograms. These generated audio waveforms were subjected to human listening tests. While the audio-

Conv2d Layers with BatchNorm2d

Filter	Stride	Output	Activation	Max Pooling kernel size
3	1	8	LReLu	2
5	1	32	LReLu	2

Fully Connected Linear Layer

Input Features	Output features
32768	8192
8192	4000
4000	2000
2000	500
500	50
50	2

TABLE I
CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

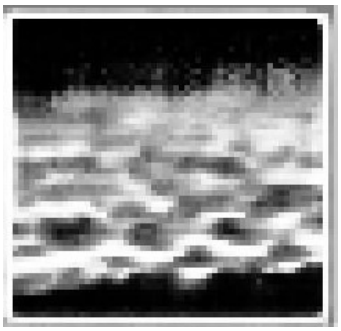


Fig. 7. Generated attention map for MS_{P2}

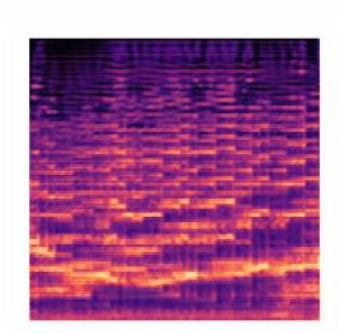


Fig. 8. Generated spectrogram of violin: MS_{P2V}

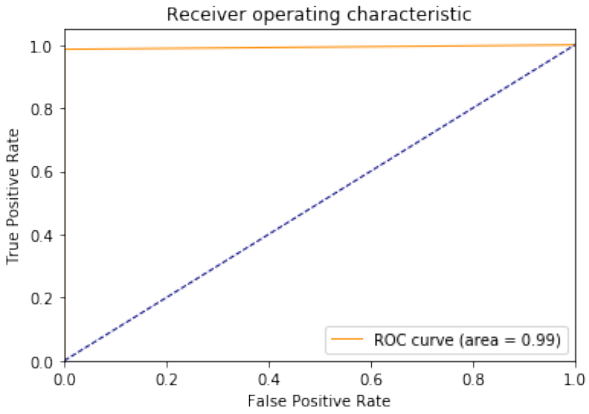


Fig. 9. ROC curve plot for CNN

waveform generated sounded close to its target instrument domain, improvements can be done with regard to reduction of unwanted noises.

V. CONCLUSION

In this paper, we propose **A**ttention-based **T**imbre **T**ransfer (**ATT**), an attention-based pipeline for musical timbre transfer. ATT is capable of manipulating the timbre of a given monophonic audio sample to that of target instrument while adding a minimal offset to other physical characteristics of sound such as loudness and pitch. By using attention, ATT is capable of manipulating only those regions of the Mel-frequency spectrogram which carry the highest importance for determining the timbre, thereby reducing the addition of unwanted characteristics in the translated image representation. Further, we also use MelGAN [4], which increases the speed of spectrogram inversion process without degrading the quality. In future, we would like to modify the training process for noise reduction as well as extend ATT towards polyphonic sound samples.

REFERENCES

[1] Y. Wang, R. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio *et al.*, "Tacotron: Towards end-to-end speech synthesis," *arXiv preprint arXiv:1703.10135*, 2017.

[2] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499*, 2016.

[3] S. Mehri, K. Kumar, I. Gulrajani, R. Kumar, S. Jain, J. Sotelo, A. Courville, and Y. Bengio, "Samplernn: An unconditional end-to-end neural audio generation model," *arXiv preprint arXiv:1612.07837*, 2016.

[4] K. Kumar, R. Kumar, T. de Boissiere, L. Gestin, W. Z. Teoh, J. Sotelo, A. de Brébisson, Y. Bengio, and A. C. Courville, "Melgan: Generative adversarial networks for conditional waveform synthesis," in *Advances in Neural Information Processing Systems*, 2019, pp. 14 881–14 892.

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[6] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[8] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," *arXiv preprint arXiv:1605.05396*, 2016.

[9] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[10] Y. A. Mejjati, C. Richardt, J. Tompkin, D. Cosker, and K. I. Kim, "Unsupervised attention-guided image-to-image translation," in *Advances in Neural Information Processing Systems*, 2018, pp. 3693–3703.

[11] S. Huang, Q. Li, C. Anil, X. Bao, S. Oore, and R. B. Grosse, "Timbretron: A wavenet (cyclegan (cqt (audio))) pipeline for musical timbre transfer," *arXiv preprint arXiv:1811.09620*, 2018.

[12] H. Choi, J.-h. Lee, and K. Lee, ""singing voice separation using generative adversarial networks,"," in *Proc. 31st Conf. Neural Information Processing Systems (NIPS 2017)*, vol. 5.

[13] N. Mor, L. Wolf, A. Polyak, and Y. Taigman, "A universal music translation network," *arXiv preprint arXiv:1805.07848*, 2018.

[14] K. Choi, G. Fazekas, M. Sandler, and K. Cho, "Convolutional recurrent neural networks for music classification," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 2392–2396.

[15] Z. Yi, H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2849–2857.

[16] "https://www.youtube.com/watch?v=xkzvya69wco."

[17] "https://www.youtube.com/watch?v=corkefuzsj0."

[18] "https://www.youtube.com/watch?v=0sdlezkik-wt=629s."

[19] "https://www.youtube.com/watch?v=wtbit8alneat=21s."

[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:30:14 UTC from IEEE Xplore. Restrictions apply.

arXiv:2010.01304v1 [math.CV] 3 Oct 2020

BOHR RADIUS FOR SOME CLASSES OF HARMONIC MAPPINGS

S. SIVAPRASAD KUMAR AND KAMALJEET GANGANIA

Abstract. We introduce a general class of sense-preserving harmonic mappings defined as follows:

$$\mathcal{S}^0_{h+\bar{g}}(M) := \{f = h + \bar{g} : \sum_{m=2}^{\infty} (\gamma_m |a_m| + \delta_m |b_m|) \leq M,\ M > 0\},$$

where $h(z) = z + \sum_{m=2}^{\infty} a_m z^m$, $g(z) = \sum_{m=2}^{\infty} b_m z^m$ are analytic functions in $\mathbb{D} := \{z \in \mathbb{C} : |z| \leq 1\}$ and

$$\gamma_m,\ \delta_m \geq \alpha_2 := \min\{\gamma_2, \delta_2\} > 0,$$

for all $m \geq 2$. We obtain Growth Theorem, Covering Theorem and derive the Bohr radius for the class $\mathcal{S}^0_{h+\bar{g}}(M)$. As an application of our results, we obtain the Bohr radius for many classes of harmonic univalent functions and some classes of univalent functions.

2010 AMS Subject Classification. 30C45, 30C50, 30C80.
Keywords and Phrases. Bohr radius, Harmonic mappings, Univalent analytic functions, Growth theorem.

1. Introduction

Let $\mathcal{H}$ be the class of complex valued harmonic functions f (which satisfy the Laplacian equation $\Delta f = 4f_{z\bar{z}} = 0$) defined on the unit disk $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$, then we can write $f = h + \bar{g}$, where h and g are analytic and satifies $f(0) = g(0)$. We say that f is sense-preserving in $\mathbb{D}$ if the Jacobian $J_f := |h'|^2 - |g'|^2 > 0$. Let $\mathcal{H}_0$ be the class of functions f with $f_{\bar{z}}(0) = 0$ and $f = h + \bar{g}$, where

$$h(z) = z + \sum_{m=2}^{\infty} a_m z^m \quad \text{and} \quad g(z) = \sum_{m=2}^{\infty} b_m z^m$$

are analytic functions in $\mathbb{D}$. Note that $\mathcal{H}_0$ contains the class $\mathcal{A}$ of normalized functions $f(0) = 0 = f_z(0) - 1$ for the case $g \equiv 0$. We denote the class of harmonic and univalent functions by $\mathcal{S}^0_{\mathcal{H}}$, which clearly contains the well known class of normalized univalent functions $\mathcal{S}$. Now we recall the definition of subordination. We say f is subordinate to ϕ, written as $f \prec \phi$, if $f(z) = \phi(w(z))$, where w is a Schwarz function. Further if ϕ is univalent, then $f \prec \phi$ if and only if $f(\mathbb{D}) \subseteq \phi(\mathbb{D})$ and $f(0) = \phi(0)$.

The work of the second author is supported by University Grant Commission, New-Delhi, India under UGC-Ref. No.:1051/(CSIR-UGC NET JUNE 2017).

Recently, Cho and Dziok [8] considered the subclass of $\mathcal{S}^0_{\mathcal{H}}$, motivated by Sakaguchi [24] class of starlike functions with respect to the symmetric points using subordination, given by

$$\mathcal{S}^{**}_{\mathcal{H}}(C,D) := \left\{ f \in \mathcal{H}_0 : \frac{2\mathcal{D}_{\mathcal{H}} f(z)}{f(z)-f(-z)} \prec \frac{1+Cz}{1+Dz}, -D \leq C < D \leq 1 \right\},$$

where $\mathcal{D}_{\mathcal{H}} f(z) := zh'(z) - \overline{zg'(z)}$. Further motivated by Silverman [25], they defined the class $\mathcal{S}^{**}_{\tau}(C,D) := \tau^0 \cap \mathcal{S}^{**}_{\mathcal{H}}(C,D)$, where $\tau^a (a \in \{0,1\})$ is the class of functions in $\mathcal{H}_0$, where $a_m = -|a_m|$ and $b_m = (-1)^a |b_m|$. Dziok [10] also considered the following classes involving Janowski functions:

$$\mathcal{S}^{*}_{\mathcal{H}}(C,D) := \left\{ f \in \mathcal{H}_0 : \frac{2\mathcal{D}_{\mathcal{H}} f(z)}{f(z)} \prec \frac{1+Cz}{1+Dz}, -D \leq C < D \leq 1 \right\},$$

and $\mathcal{S}^{c}_{\mathcal{H}}(C,D) := \{ f \in \mathcal{S}^0_{\mathcal{H}} : D_{\mathcal{H}} f(z) \in \mathcal{S}^*_{\mathcal{T}}(C,D) \}$, and further defined the classes $\mathcal{S}^{*}_{\tau}(C,D) := \tau^0 \cap \mathcal{S}^{*}_{\mathcal{H}}(C,D)$ and $\mathcal{S}^{c}_{\tau}(C,D) := \tau^1 \cap \mathcal{S}^{c}_{\mathcal{H}}(C,D)$. Singh [26] studied the subclass $\mathcal{F}(\lambda) := \{ f \in \mathcal{A} : |f'(z)-1| < \lambda, \lambda \in (0,1] \}$ of close-to-convex functions. Further Nagpal and Ravichandran [19] considered its harmonic extension defined as:

$$\mathcal{F}^0_{\mathcal{H}}(\lambda) := \{ f = h + \bar{g} : |f_z(z) - 1| < \lambda - |f_{\bar{z}}(z)|, \lambda \in (0,1] \}.$$

Gao and Zohu [12] studied the subclass of close to convex functions, $W(\mu,\rho) := \{ f \in \mathcal{A} : \Re(h'(z) + \mu z h''(z)) > \rho,\ \mu \geq 0, 0 \leq \rho < 1 \}$. Rajbala and Prajapat [23] also studied a subclass of close-to-convex harmonic mappings defined as:

$$W^0_{\mathcal{H}}(\mu,\rho) := \{ f = h + \bar{g} \in \mathcal{H}_0 : \Re(h'(z) + \mu z h''(z) - \rho) > |g'(z) + \mu z g''(z)| \},$$

where $\mu \geq 0$ and $0 \leq \rho < 1$, which is the harmonic extension of $W(\mu,\rho)$ and generalizes the classes studied by Gosh and Vasudevarao [13], and Nagpal and Ravichandran [20]. In a similar way, Dixit and Porwal [9] considered the class $R_{\mathcal{H}}(\beta) = \{ f = h + \bar{g} : \Re\{h'(z) + g'(z)\} \leq \beta, 2 \geq \beta > 1 \}$, where $h(z) = z + \sum_{m=2}^{\infty} |a_m| z^m$ and $g(z) = \sum_{m=1}^{\infty} |b_m| \bar{z}^m$, where $|b_1| < 1$. Now if we consider $b_1 = 0$, then we have the class

$$R^0_{\mathcal{H}}(\beta) := \{ f = h + \bar{g} \in \mathcal{H}_0 : \Re\{h'(z) + g'(z)\} \leq \beta, \beta > 1 \},$$

consists of the functions with positive coefficients, which reduces to the class $R(\beta)$ studied by Uralegaddi for the case $g \equiv 0$. Altinkaya et al. [4] studied the class $k - \widetilde{ST}^{-}_{q}(\alpha)$ of functions in $\mathcal{A}$ with negative coefficients associated with the conic domains defined by $\Re(p(z)) > k|p(z)-1| + \alpha$, where $0 \leq k < \infty$, $0 \leq \alpha < 1$, $0 < q < 1$, $p(z) = z(\widetilde{D}_q f)(z)/f(z)$ and

$$(\widetilde{D}_q f)(z) = \begin{cases} \frac{f(qz)-f(q^{-1}z)}{(q-q^{-1})z}, & z \neq 0 \\ f'(0), & z = 0 \end{cases} \quad \text{and} \quad [\widetilde{m}]_q = \frac{q^m - q^{-m}}{q - q^{-1}}.$$

In literature, sufficient conditions for many classes of harmonic and analytic mappings are obtained in terms of coefficients. See [4, 8, 9, 10, 19, 23]. In a similar way, we introduce a new subclass of $\mathcal{H}_0$ as follows:

$$\mathcal{S}^0_{h+\bar{g}}(M) := \{ f = h + \bar{g} : \sum_{m=2}^{\infty} (\gamma_m |a_m| + \delta_m |b_m|) \leq M,\ M > 0 \},$$

where

$$\gamma_m,\ \delta_m \geq \alpha_2 := \min\{\gamma_2, \delta_2\} > 0, \tag{1.1}$$

for all $m \geq 2$. Thus we observe that the classes $S^{**}_{\tau}(C, D)$, $S^{*}_{\tau}(C, D)$, $S^{c}_{\tau}(C, D)$ and $k-\widetilde{ST}^{-}_{q}(\alpha)$ are all become subclasses of $\mathcal{S}^{0}_{h+\bar{g}}(M)$ for some suitable choices of γ_m, δ_m and M.

Now let us recall the Bohr phenomenon for the harmonic mappings:

Definition 1.1. Let $f(z) = h(z)+\overline{g(z)} = z+\sum_{m=2}^{\infty} a_m z^m+\overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{H}_0$. Then the Bohr-phenomenon is to find the constant $r^* \in (0, 1]$ such that the inequality $r + \sum_{m=2}^{\infty}(|a_m| + |b_m|)r^m \leq d(f(0), \partial\Omega)$ holds for all $|z| = r \leq r^*$, where $d(f(0), \partial\Omega)$ denotes the Euclidean distance between $f(0)$ and the boundary of $\Omega := f(\mathbb{D})$. The largest such r^* is called the Bohr radius.

The work on Bohr radius originated with the work of H. Bohr [7] on the inequality $\sum_{m=0}^{\infty} |a_m|r^m \leq 1$ for an analytic function having the power series $\sum_{m=0}^{\infty} a_m z^m$, which holds for $r \leq 1/3$, known as Bohr Theorem. In recent times, finding the Bohr radius for such inequalities with different prospects is an active area of research. For some recent work on Bohr phenomenon related to the Ma-Minda [16] classes of starlike functions, bounded harmonic mappings and families of analytic functions defined by subordination and many more, see [1, 2, 3, 6, 7, 14, 15, 17, 18] and references therein.

In this article, we find the sharp growth theorem, covering theorem and finally derive the Bohr radius for the class $\mathcal{S}^{0}_{h+\bar{g}}(M)$. As an application of our results, we obtain the Bohr radius for the classes $\mathcal{S}^{**}_{\tau}(C, D)$, $S^{*}_{\tau}(C, D)$, $S^{c}_{\tau}(C, D)$, $k - \widetilde{ST}^{-}_{q}(\alpha)$ and many more. Further, Bohr radius for the classes $W^{0}_{\mathcal{H}}(\mu, \rho)$ and $\mathcal{F}^{0}_{\mathcal{H}}(\lambda)$ is also derived.

2. The class $\mathcal{S}^{0}_{h+\bar{g}}(M)$ and its applications

From the condition (1.1), we see that

$$\sum_{m=2}^{\infty}(\gamma_m|a_m| + \delta_m|b_m|) \leq M \tag{2.1}$$

implies

$$\sum_{m=2}^{\infty}(|a_m| + |b_m|) \leq \frac{M}{\alpha_2}, \tag{2.2}$$

where γ_m, δ_m and α_2 are as defined in (1.1) and the equality in (2.2) holds for the function $f(z) = z + (M/\alpha_2)z^2$. Now using the inequality $|a| - |b| \leq |a \pm b| \leq |a| + |b|$ and then (2.2), we have for $|z| = r$

$$|z| - \left|\sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m}\right| \leq |f(z)| \leq |z| + \left|\sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m}\right|,$$

which immediately yields the following growth theorem:

Theorem 2.1 (Growth Theorem)**.** *Let* $f(z) = z+\sum_{m=2}^{\infty} a_m z^m+\overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{S}^0_{h+\bar{g}}(M)$*. Then*

$$r - \frac{M}{\alpha_2} r^2 \leq |f(z)| \leq r + \frac{M}{\alpha_2} r^2 \quad (|z| = r).$$

The inequalities are sharp for the functions $f(z) = z \pm \frac{M}{\alpha_2} z^2$ *and* $f(z) = z \pm \frac{M}{\alpha_2} \bar{z}^2$ *with the suitable choice of* α_2*.*

From the above result, we also see that functions in the class $\mathcal{S}^0_{h+\bar{g}}(M)$ are bounded. Now taking r tending to 1^-, we get the covering theorem:

Corollary 2.2 (Covering Theorem)**.** *Let* $f(z) = z+\sum_{m=2}^{\infty} a_m z^m+\overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{S}^0_{h+\bar{g}}(M)$*. Then*

$$\left\{ w \in \mathbb{C} : |w| \leq 1 - \frac{M}{\alpha_2} \right\} \subset f(\mathbb{D}).$$

Now we are ready to obtain the Bohr-radius for the class $\mathcal{S}^0_{h+\bar{g}}(M)$.

Theorem 2.3. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{S}^0_{h+\bar{g}}(M)$*. Then*

$$|z| + \sum_{m=2}^{\infty} (|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$*, where*

$$r^* = \frac{-1 + \sqrt{1 + 4\left(\frac{M}{\alpha_2}\right)\left(1 - \frac{M}{\alpha_2}\right)}}{2\left(\frac{M}{\alpha_2}\right)}.$$

Bohr radius r^* *is achieved by the function* $f(z) = z - \frac{M}{\alpha_2} z^2$*.*

Proof. From the Growth Theorem 2.1 and Covering Theorem 2.2, we see that the distance between origin and the boundary of $f(\mathbb{D})$ satisfies

$$d(f(0), \partial f(\mathbb{D})) \geq 1 - \frac{M}{\alpha_2}. \tag{2.3}$$

Let us consider the continuous function defined in $(0, 1)$ as

$$H(r) := r + \frac{M}{\alpha_2} r^2 - \left(1 - \frac{M}{\alpha_2}\right),$$

such that $H'(r) > 0$ for $r \in (0, 1)$ with $H(0) < 0$ and $H(1) > 0$. Therefore, by Intermediate Value Theorem for continuous functions, we let r^* be the unique positive root in $(0, 1)$ as mentioned in the Theorem statement, and thus for $r = r^*$, we have

$$r^* + \frac{M}{\alpha_2} (r^*)^2 = 1 - \frac{M}{\alpha_2}.$$

Now from (2.2), (2.3) and the above equality, it follows that

$$|z| + \sum_{m=2}^{\infty} (|a_m| + |b_m|)|z|^m \leq r + \frac{M}{\alpha_2} r^2 \leq r^* + \frac{M}{\alpha_2} (r^*)^2 \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| = r \leq r^*$. Let us consider the analytic function $f : \mathbb{D} \to \mathbb{C}$

$$f(z) = z - \frac{M}{\alpha_2} z^2,$$

which by suitably choosing α_2 and using (2.1) belongs to $\mathcal{S}^0_{h+\bar{g}}(M)$. Further for $|z| = r^*$, we have

$$|z| + \sum_{m=2}^{\infty} (|a_m| + |b_m|)|z|^m = d(f(0), \partial f(\mathbb{D})).$$

Hence the result is sharp. □

Remark 2.1. Note that we can easily extend our results by considering the analytic functions of the following form:

$$h(z) = z + \sum_{m=k}^{\infty} a_m z^m \quad \text{and} \quad g(z) = \sum_{m=k}^{\infty} b_m z^m, \quad (k \geq 2)$$

and the class

$$\mathcal{S}^0_{h+\bar{g}}(k, M) := \{f = h + \bar{g} : \sum_{m=k}^{\infty} (\gamma_m |a_m| + \delta_m |b_m|) \leq M,\ M > 0\},$$

where

$$\gamma_m,\ \delta_m \geq \alpha_k := \min\{\gamma_k, \delta_k\} > 0,$$

for all $m \geq k$. Thus we see that $\mathcal{S}^0_{h+\bar{g}}(2, M) \equiv \mathcal{S}^0_{h+\bar{g}}(M)$. Precisely, for the class $\mathcal{S}^0_{h+\bar{g}}(k, M)$, we have

$$r - \frac{M}{\alpha_k} r^k \leq |f(z)| \leq r + \frac{M}{\alpha_k} r^k, \quad (|z| = r)$$

and the Bohr radius r^* is the unique positive root in $(0, 1)$ of the equation

$$r + \frac{M}{\alpha_k}(r^k) - \left(1 - \frac{M}{\alpha_k}\right) = 0.$$

Now we can also obtain the Bohr radius for the classes (see [11, Theorem 8, 9]) studied by Dziok [11] .

2.1. **Applications.** Cho and Dziok [8] showed that $f \in \mathcal{S}^{**}_{\tau}(C, D)$ if and only if

$$\sum_{m=2}^{\infty} (|\alpha_m||a_m| + |\beta_m||b_m|) \leq D - C, \tag{2.4}$$

where

$$\alpha_m = m(1 + D) - \frac{(1 + C)(1 - (-1)^m)}{2}$$

and

$$\beta_m = m(1 + D) + \frac{(1 + C)(1 - (-1)^m)}{2}.$$

Note that for all $m \geq 2$, we have $\alpha_m < \beta_m$, which shows that $0 < \alpha_2 \leq \alpha_m < \beta_m$. Therefore from (2.4) we obtain that

$$\sum_{m=2}^{\infty}(|a_m| + |b_m|) \leq \frac{D-C}{\alpha_2}$$

and also the condition in (1.1) holds by choosing $\gamma_m = \alpha_m$, $\delta_m = \beta_m$ and $M = D - C$. Thus using Theorem 2.2 and Theorem 2.3, we have:

Theorem 2.4. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{S}_\tau^{**}(C, D)$. *Then*

$$r - \frac{D-C}{2(1+D)}r^2 \leq |f(z)| \leq r + \frac{D-C}{2(1+D)}r^2 \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{-1 + \sqrt{1 + 4\left(\frac{D-C}{\alpha_2}\right)\left(1 - \frac{D-C}{\alpha_2}\right)}}{2\left(\frac{D-C}{\alpha_2}\right)},$$

where α_2 *is as defined in* (2.4). *Bohr radius* r^* *is achieved by the function* $f(z) = z - \frac{D-C}{2(1+D)}z^2$.

Dziok [10] proved that $f \in S_\tau^*(C, D)$ if and only if

$$\sum_{m=2}^{\infty}(\alpha_m|a_m| + \beta_m|b_m|) \leq D - C \tag{2.5}$$

and $f \in S_\tau^c(C, D)$ if and only if

$$\sum_{m=2}^{\infty}(m\alpha_m|a_m| + m\beta_m|b_m|) \leq D - C, \tag{2.6}$$

where $\alpha_m = m(1 + D) - (1 + C)$ and $\beta_m = m(1 + D) + (1 + C)$. We note that $\beta_m > \alpha_m \geq \alpha_2 > 0$ for all $m \geq 2$. Therefore from (2.5), we obtain that

$$\sum_{m=2}^{\infty}(|a_m| + |b_m|) \leq \frac{D-C}{\alpha_2}.$$

Now choosing $\gamma_m = \alpha_m$, $\delta_m = \beta_m$ and $M = D - C$, the condition (1.1) holds and thus using Theorem 2.2 and Theorem 2.3 we get:

Theorem 2.5. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in S_\tau^*(C, D)$. *Then*

$$r - \frac{D-C}{1+2D-C}r^2 \leq |f(z)| \leq r + \frac{D-C}{1+2D-C}r^2, \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{-1 + \sqrt{1 + 4\left(\frac{D-C}{1+2D-C}\right)\left(1 - \frac{D-C}{1+2D-C}\right)}}{2\left(\frac{D-C}{1+2D-C}\right)}.$$

Bohr radius r^* *is achieved by the function* $f(z) = z - \frac{D-C}{1+2D-C}z^2$.

Again using $\beta_m > \alpha_m \geq \alpha_2$ for all $m \geq 2$ in (2.6), we obtain that if $f \in S_\tau^c(C, D)$, then the following inequality holds:

$$\sum_{m=2}^{\infty}(|a_m| + |b_m|) \leq \frac{D-C}{2\alpha_2}.$$

Now choosing $\gamma_m = m\alpha_m$, $\delta_m = m\beta_m$ and $M = D - C$, the condition (1.1) holds and thus using Theorem 2.2 and 2.3, we obtain the following result:

Theorem 2.6. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in S_\tau^c(C, D)$. *Then*

$$r - \frac{D-C}{2(1+2D-C)}r^2 \leq |f(z)| \leq r + \frac{D-C}{2(1+2D-C)}r^2, \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where* α_2 *is defined in* (2.6) *and*

$$r^* = \frac{-1 + \sqrt{1 + 2\left(\frac{D-C}{\alpha_2}\right)\left(1 - \frac{D-C}{\alpha_2}\right)}}{\left(\frac{D-C}{\alpha_2}\right)}.$$

Bohr radius r^* *is achieved by the function* $f(z) = z - \frac{D-C}{2(1+2D-C)}z^2$.

Now we obtain the result for the class $\mathcal{F}_{\mathcal{H}}^0(\lambda)$. Using the condition

$$\sum_{m=2}^{\infty} m(|a_m| + |b_m|) \leq \lambda, \tag{2.7}$$

convolution properties, radius of starlikeness and certain inclusion relationships were studied in [19] for the class $\mathcal{F}_{\mathcal{H}}^0(\lambda)$. Now with the same condition, we arrive at the following result using Theorem 2.2 and 2.3:

Theorem 2.7. *If* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in \mathcal{F}_{\mathcal{H}}^0(\lambda)$ *and also satisfy the condition* $\sum_{m=2}^{\infty} m(|a_m| + |b_m|) \leq \lambda$. *Then*

$$r - \frac{\lambda}{2}r^2 \leq |f(z)| \leq r + \frac{\lambda}{2}r^2, \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{-1 + \sqrt{1 + 2\lambda\left(1 - \frac{\lambda}{2}\right)}}{\lambda}$$

and the Bohr radius r^* *for* $\mathcal{F}^0_{\mathcal{H}}(\lambda)$ *is obtained when* $f(z) = z - \frac{\lambda}{2}z^2$.

If $g \equiv 0$, then $\mathcal{F}^0_{\mathcal{H}}(\lambda)$ reduces to the Singh [26] class $\mathcal{F}(\lambda)$, which is contained in the MacGregor subclass $\mathcal{F} := \{f \in \mathcal{A} : |f'(z) - 1| < 1\}$ of close-to-convex functions.

Corollary 2.8. *Bohr radius for the classes* $\mathcal{F}(\lambda)$ *and* $\mathcal{F}^0_{\mathcal{H}}(\lambda)$ *is same, whenever the condition* (2.7) *holds.*

The following two theorems are for the classes $k - \widetilde{ST}_q^-(\alpha)$ and $R^0_{\mathcal{H}}(\beta)$ respectively:

Lemma 2.1. [4] *Let* $0 \leq k < \infty$, $0 < q < 1$ *and* $0 \leq \alpha < 1$. *Then* $f \in k - \widetilde{ST}_q^-(\alpha)$ *if and only if*

$$\sum_{m=2}^{\infty}([\widetilde{m}]_q(k+1) - (k+\alpha))a_m \leq 1 - \alpha.$$

From Lemma 2.1, we see that choosing $\gamma_m = [\widetilde{m}]_q(k+1) - (k+\alpha)$, $\delta_m = 0$ and $M = 1 - \alpha$, condition in (1.1) holds, and therefore applying Theorem 2.2 and 2.3, we get the following result:

Theorem 2.9. *Let* $f(z) = z - \sum_{m=2}^{\infty} a_m z^m \in k - \widetilde{ST}_q^-(\alpha)$. *Then*

$$r - \frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}r^2 \leq |f(z)| \leq r + \frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}r^2,$$

where $|z| = r$ *and*

$$|z| + \sum_{m=2}^{\infty} a_m |z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{-1 + \sqrt{1 + 4\left(\frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}\right)\left(1 - \frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}\right)}}{2\left(\frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}\right)}$$

and the Bohr radius r^* *is obtained when* $f(z) = z - \frac{q(1-\alpha)}{(q^2+1)(k+1) - q(k+\alpha)}z^2$.

Lemma 2.2. [9] *Let* $f \in R^0_{\mathcal{H}}(\beta)$. *Then the following inequality*

$$\sum_{m=2}^{\infty}(m(|a_m| + |b_m|)) \leq \beta - 1$$

is necessary and sufficient for the functions to be in $R^0_{\mathcal{H}}(\beta)$.

From Lemma 2.2, we see that choosing $\gamma_m = \delta_m = m$ and $M = \beta - 1$, condition in (1.1) holds, and therefore applying Theorem 2.2 and 2.3, we get the following result:

Theorem 2.10. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in R^0_{\mathcal{H}}(\beta)$. *Then*

$$r - \frac{\beta - 1}{2} r^2 \leq |f(z)| \leq r + \frac{\beta - 1}{2} r^2, \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty} (|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{-1 + \sqrt{1 + 2(\beta - 1)\left(1 - \frac{\beta - 1}{2}\right)}}{\beta - 1}$$

and the Bohr radius r^* *for the class* $R^0_{\mathcal{H}}(\beta)$ *is obtained when* $f(z) = z - \frac{\beta - 1}{2} z^2$.

Corollary 2.11. *Bohr radius for the classes* $R(\beta)$ *and* $R^0_{\mathcal{H}}(\beta)$ *is same.*

Silverman considered the class with negative coefficients as follows:

$$\mathcal{T} := \left\{ f \in \mathcal{S} : f(z) = z - \sum_{m=2}^{\infty} a_m z^m, a_m \geq 0 \right\}.$$

Using this recently, Ali et al. [3] considered the following general class defined as:

$$\mathcal{T}(\alpha) := \left\{ f \in \mathcal{T} : \sum_{m=2}^{\infty} g_m a_m \leq 1 - \alpha \right\},$$

where $g_m \geq g_2 > 0$ and $\alpha < 1$. Note that if we choose in (1.1), $\gamma_m = g_m$, $\delta_m = 0$ and $M = 1 - \alpha$, then the class $\mathcal{S}^0_{h+\bar{g}}(M)$ contains $\mathcal{T}(\alpha)$, which satisfies the required conditions. Thus, using Theorem 2.1 and 2.3, we have the following result obtained in [3]:

Theorem 2.12. [3] *Let* $f(z) = z - \sum_{m=2}^{\infty} a_m z^m \in \mathcal{T}(\alpha)$. *Then*

$$r - \frac{1 - \alpha}{\gamma_2} r^2 \leq |f(z)| \leq r + \frac{1 - \alpha}{\gamma_2} r^2, \quad (|z| = r)$$

and

$$|z| + \sum_{m=2}^{\infty} |a_m||z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where*

$$r^* = \frac{2(\gamma_2 + \alpha - 1)}{\gamma_2 + \sqrt{\gamma_2^2 + 4\gamma_2(1 - \alpha) - 4(1 - \alpha)^2}}$$

and the Bohr radius r^* *is obtained by the function* $f(z) = z - \frac{1 - \alpha}{\gamma_2} z^2$.

Choosing $\gamma_m = m - \alpha$ and $\gamma_m = m(m-\alpha)$ in (1.1), the class $\mathcal{S}^0_{h+\bar{g}}(M)$ contains the Silverman classes $\mathcal{TST}(\alpha) := \mathcal{T} \cap \mathcal{ST}(\alpha)$ and $\mathcal{TCV}(\alpha) := \mathcal{T} \cap \mathcal{CV}(\alpha)$ of starlike and convex functions with negative coefficients respectively and Theorem 2.3 provides the Bohr radius as obtained in [3, Theorem 2.3] and [3, Theorem 2.4].

Ali et al. [3] considered the class $\mathcal{TF}_\alpha := \mathcal{T} \cap \mathcal{F}_\alpha$, $0 \leq \alpha \leq 1$, where $\mathcal{F}_\alpha$ is the class of close to convex functions and showed that if $f \in \mathcal{TF}_\alpha$, then $\sum_{m=2}^{\infty} m(2n+\alpha)a_m \leq 2+\alpha$. Thus choosing $\gamma_m = m(2n+\alpha)$ and $\delta_m = 0$ and $M = 2+\alpha$ in (1.1), the class $\mathcal{TF}_\alpha$ satisfies all conditions of $\mathcal{S}^0_{h+\bar{g}}(M)$, and Theorem 2.3 (or Theorem 2.12) reduces to [3, Corollary 2.6].

For $\alpha > 1$, Owa and Nishiwaki [21] considered the classes of analytic functions $\mathcal{M}(\alpha) := \{f \in \mathcal{A} : \Re(zf'(z)/f(z)) < \alpha\}$ and $\mathcal{N}(\alpha) := \{f \in \mathcal{A} : 1 + \Re(zf''(z)/f'(z)) < \alpha\}$. They showed that the following conditions:

$$\sum_{n=2}^{\infty}((m-\mu) + |m+\mu-2\alpha|)|a_m| \leq 2(\alpha-1)$$

and

$$\sum_{n=2}^{\infty} m(m-\mu+1+|m+\mu-2\alpha|)|a_m| \leq 2(\alpha-1),$$

where $0 \leq \mu \leq 1$ are sufficient for the function $f(z) = z + \sum_{m=2}^{\infty} a_m z^m$ to be in $\mathcal{M}(\alpha)$ and $\mathcal{N}(\alpha)$, respectively. It is easy to see that the above two conditions also become necessary for the classes $\mathcal{TM}(\alpha) := \mathcal{T} \cap \mathcal{M}(\alpha)$ and $\mathcal{TN}(\alpha) := \mathcal{T} \cap \mathcal{N}(\alpha)$, respectively. Thus choosing $\gamma_m = (m-\mu)+|m+\mu-2\alpha|$, $\gamma_m = m(m-\mu+1+|m+\mu-2\alpha|)$ with $\delta_m = 0$ and $M = 2(\alpha-1)$, from Theorem 2.2 and 2.3, we obtain the following result respectively:

Theorem 2.13. *Let $f(z) = z - \sum_{m=2}^{\infty} a_m z^m \in \mathcal{TM}(\alpha)$ (or $\mathcal{TN}(\alpha)$). Then*

$$r - \frac{2(\alpha-1)}{\gamma_2} r^2 \leq |f(z)| \leq r + \frac{2(\alpha-1)}{\gamma_2} r^2, \quad (|z| = r),$$

where $\gamma_2 := (2-\mu) + |\mu - 2(1-\alpha)|$ (or $2(3-\mu+|\mu-2(1-\alpha)|)$) and

$$|z| + \sum_{m=2}^{\infty} |a_m||z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, where

$$r^* = \frac{-1 + \sqrt{1 + 4\left(\frac{2(\alpha-1)}{\gamma_2}\right)\left(1 - \frac{2(\alpha-1)}{\gamma_2}\right)}}{2\left(\frac{2(\alpha-1)}{\gamma_2}\right)}.$$

The radius r^*, achieved for the function $f(z) = z - \frac{2(\alpha-1)}{\gamma_2} z^2$ is the Bohr radius for the class $\mathcal{TM}(\alpha)$ (or $\mathcal{TN}(\alpha)$).

3. Bohr-radius for the class $W^0_{\mathcal{H}}(\mu,\rho)$

Let us recall the subclass of close-to-convex harmonic mappings introduced by Rajbala and Prajapat [23]:

$$W^0_{\mathcal{H}}(\mu,\rho) := \{f = h + \bar{g} \in \mathcal{H}_0 : \Re(h'(z) + \mu z h''(z) - \rho) > |g'(z) + \mu z g''(z)|\},$$

where $\mu \geq 0$ and $0 \leq \rho < 1$.

Lemma 3.1. [23] *let $f = h + \bar{g} \in W^0_{\mathcal{H}}(\mu,\rho)$. Then for $m \geq 2$ the following sharp inequality holds:*

$$|a_m| + |b_m| \leq \frac{2(1-\rho)}{m(1+\mu(m-1))}.$$

Lemma 3.2. [23] *let $f = h + \bar{g} \in W^0_{\mathcal{H}}(\mu,\rho)$. Then for $|z| = r$, we have the sharp inequality*

$$|f(z)| \geq |z| - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))}|z|^m.$$

Theorem 3.1. *Let $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in W^0_{\mathcal{H}}(\mu,\rho)$. Then*

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, where r^* is the unique positive root in $(0,1)$ of

$$r + \sum_{m=2}^{\infty} \frac{2(1-\rho)}{m(1+\mu(m-1))} r^m = 1 - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))}.$$

The radius r^* is the Bohr radius for the class $W^0_{\mathcal{H}}(\mu,\rho)$.

Proof. From Lemma 3.2, it follows that the distance between origin and the boundary of $f(\mathbb{D})$ satisfies

$$d(f(0), \partial f(\mathbb{D})) \geq \left(1 - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))}\right). \tag{3.1}$$

Let us consider the continuous function

$$H(r) := r + \sum_{m=2}^{\infty} \frac{2(1-\rho)}{m(1+\mu(m-1))} r^m - \left(1 - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))}\right).$$

Now

$$H'(r) = 1 + \sum_{m=2}^{\infty} \frac{2m(1-\rho)}{m(1+\mu(m-1))} r^{m-1} > 0$$

for all $r \in (0,1)$, which implies that H is a strictly increasing continuous function. Note that $H(0) < 0$ and

$$H(1) = \sum_{m=2}^{\infty} \frac{2(1-\rho)}{m(1+\mu(m-1))} + \sum_{m=2}^{\infty} \frac{2(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))} > 0.$$

Thus by Intermediate Value Theorem for continuous function, we let r^* be the unique root of $H(r) = 0$ in $(0, 1)$. Now using Lemma 3.1 and the inequality (3.1), we have

$$\begin{aligned}|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m &\leq r + \sum_{m=2}^{\infty} \frac{2(1-\rho)}{m(1+\mu(m-1))} r^m \\ &\leq r^* + \sum_{m=2}^{\infty} \frac{2(1-\rho)}{m(1+\mu(m-1))} (r^*)^m \\ &= \left(1 - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))}\right) \\ &\leq d(f(0), \partial f(\mathbb{D})),\end{aligned}$$

which hold for $r \leq r^*$. Now consider the analytic function

$$f(z) = z + \sum_{m=2}^{\infty} \frac{2(-1)^{m-1}(1-\rho)}{m(1+\mu(m-1))} z^m.$$

Then clearly $f \in W^0_{\mathcal{H}}(\mu, \rho)$ and at $|z| = r^*$, we get $|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m = d(f(0), \partial f(\mathbb{D}))$. Hence the radius r^* is the Bohr radius for the class $W^0_{\mathcal{H}}(\mu, \rho)$. □

Now using Theorem 3.1, we can obtain the Bohr radius for the classes $W^0_{\mathcal{H}}(\mu, 0) = W^0_{\mathcal{H}}(\mu)$, $W^0_{\mathcal{H}}(0, \rho) = P^0_{\mathcal{H}}(\rho)$, $W^0_{\mathcal{H}}(1, 0) = W^0_{\mathcal{H}}$ and $W^0_{\mathcal{H}}(0, 0) = P^0_{\mathcal{H}}$. Here we mention the following:

Corollary 3.2. *Let* $f(z) = z + \sum_{m=2}^{\infty} a_m z^m + \overline{\sum_{m=2}^{\infty} b_m z^m} \in P^0_{\mathcal{H}}$. *Then*

$$|z| + \sum_{m=2}^{\infty}(|a_m| + |b_m|)|z|^m \leq d(f(0), \partial f(\mathbb{D}))$$

for $|z| \leq r^*$, *where the Bohr radius* r^* *is the unique positive root in* $(0, 1)$ *of*

$$r + \sum_{m=2}^{\infty} \frac{2}{m} r^m = 1 - 2\sum_{m=2}^{\infty} \frac{(-1)^{m-1}}{m}. \tag{3.2}$$

For the case $g \equiv 0$, the class $W^0_{\mathcal{H}}(\mu, \rho)$ reduces to $W(\mu, \rho)$, the class introduced by Gao and Zohu [12], and we have:

Corollary 3.3. *The Bohr radius of the classes* $W(\mu, \rho)$ *and* $W^0_{\mathcal{H}}(\mu, \rho)$ *is same.*

4. Conclusion and future scope

We studied the Bohr phenomenon for the class $\mathcal{S}^0_{h+\bar{g}}(M)$ and pointed out its several applications in context of various known classes. Further the Bohr radius for the classes of q-starlike and q-convex functions studied in [5] can be obtained by mere an application of our result. Similarly many other classes can also be dealt for Bohr radius. Since $\mathcal{S}^{**}_{\tau}(C, D) \subset \mathcal{S}^{**}_{\mathcal{H}}(C, D)$, $S^*_{\tau}(C, D) \subset \mathcal{S}^*_{\mathcal{H}}(C, D)$ and $S^c_{\tau}(C, D) \subset \mathcal{S}^c_{\mathcal{H}}(C, D)$. If we let r_* be the Bohr radius for a well defined class $\mathcal{F}$, then we conclude that $r_* \leq r^*$ whenever $\mathcal{F}$ is $\mathcal{S}^{**}_{\mathcal{H}}(C, D)$ or $\mathcal{S}^*_{\mathcal{H}}(C, D)$ or $\mathcal{S}^c_{\mathcal{H}}(C, D)$. However finding r_* is still open.

Conflict of interest

The authors declare that they have no conflict of interest.

References

[1] Ahuja, O.P, Anand, S. and Jain, N. K.: Bohr Radius Problems for Some Classes of Analytic Functions Using Quantum Calculus Approach, Mathematics. 8 (2020), 623.

[2] Ali, R.M., Barnard, R.W. and Solynin, A.Yu.: A note on the Bohr's phenomenon for power series. J. Math. Anal. Appl. 449 (2017), 154–167

[3] Ali, R.M., Jain, N.K. and Ravichandran, V.: Bohr radius for classes of analytic functions. Results Math. 74 (2019), Art. 179, 13 pp.

[4] Altinkaya, S., Kanas, S. and Yalcin, S.: Subclass of -Uniformly Starlike Functions Defined by symmetric -Derivative Operator. Ukrains'kyi Matematychnyi Zhurnal, 70(11) (2018), pp. 1499-1510.

[5] Altintas, O. and Mustafa, N.: Coefficient bounds and distortion theorems for the certain analytic functions, Turkish J. Math. 43 (2019), 985–997.

[6] Boas, H.P. and Khavinson, D.: Bohr's power series theorem in several variables. Proc. Amer. Math. Soc. 125 (1997), 2975–2979

[7] Bohr, H.: A theorem concerning power series. Proc. London Math. Soc. 13 (1914), 1–5

[8] Cho, N.E. and Dziok, J.: Harmonic starlike functions with respect to symmetric points. Axioms 2020, 9, 3.

[9] Dixit, K.K. and Porwal, S.: A subclass of harmonic univalent functions with positive coefficients, Tamkang J. Math. 41 (2010), no. 3, 261–269.

[10] Dziok, J.: On Janowski harmonic functions, J. Appl. Anal. 21(2) (2015), 99–107.

[11] Dziok, J.: Classes of harmonic functions associated with Ruscheweyh derivatives. RACSAM 113 (2019), 1315–1329.

[12] Gao, C.Y. and Zohu, S.Q.: Certain Subclass of starlike functions, Appl. Math. Comput. 187 (2007), 176–182.

[13] Ghosh, N. and Vasudevarao, A.: On a subclass of harmonic close-to-convex mappings, Monatsh Math. 188(2) (2019), 247–267.

[14] Kumar, S.S. and Gangania, K.: On certain Generalizations of $\mathcal{S}^*(\psi)$, arXiv:2007.06069v1 (2020)

[15] Liu, M., Ponnusamy, S. and Wang, J.: Bohr's phenomenon for the classes of Quasi-subordination and K-quasiregular harmonic mappings, RACSAM **144**. 115 (2020)

[16] Ma, W.C. and Minda, D.: A unified treatment of some special classes of univalent functions. *Proceedings of the Conference on Complex Analysis, Tianjin*, Conf Proc Lecture Notes Anal., I Int Press, Cambridge, MA. 157–169 (1992)

[17] Muhanna, Y.A.: Bohr's phenomenon in subordination and bounded harmonic classes. Complex Var. Elliptic Equ. 55 (2010), 1071–1078

[18] Muhanna, Y.A., Ali, R.M., Ng, Z.C. and Hansi, S.F.M.: Bohr radius for subordinating families of analytic functions and bounded harmonic mappings. J. Math. Anal. Appl. 420 (2014), 124–136

[19] Nagpal, S. and Ravichandran, V.: A subclass of close-to-convex harmonic mappings, Complex Var. Elliptic Equ. 59(2) (2014), 204–216.

[20] Nagpal, S. and Ravichandran, V.: Construction of subclasses of univalent harmonic mappings, J. Korean Math. Soc. 53 (2014), 567–592.

[21] Owa, S. and Nishiwaki, J.: Coefficient estimates for certain classes of analytic functions, J. Inequal. Pure Appl. Math. 3(5) (2002), Article 72, 5 pp.

[22] Paulsen, V.I., Popescu, G. and Singh, D.: On Bohr's inequality. Proc. London Math. Soc. 85 (2002), 493–512

[23] Rajbala and Prajapat, J.K.: Certain geometric properties of close-to-convex harmonic mappings, https://doi.org/10.1142/S1793557121501023, (2020)

[24] Sakaguchi, K.: On a certain univalent mapping. J. Math. Soc. Jpn., 2 (1959), 72–75.

[25] Silverman, H.: Harmonic univalent functions with negative coefficients, J. Math. Anal. Appl. 220 (1998), 283–289.

[26] Singh, V.: Univalent functions with bounded derivative in the unit disc, Ind. J. Pure Appl. Math. 8 (1977), 1370-1377.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

Email address: spkumar@dce.ac.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India

Email address: gangania.m1991@gmail.com

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344379230

Comparative Study of different Missile shapes using Computational Fluid dynamics

Article *in* International Journal Of Advance Research And Innovative Ideas In Education · September 2020

CITATIONS
0

READS
11

4 authors, including:

Gaurav Kumar
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Comparative Study of different Missile shapes using Computational Fluid dynamics View project

All content following this page was uploaded by Gaurav Kumar on 25 September 2020.

The user has requested enhancement of the downloaded file.

Volume 8 Issue 3 (2020) 264-271 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Comparative Study of different Missile shapes using Computational Fluid dynamics

Gaurav Kumar, Rohit Kumar, Sagar Dagar, Raj Kumar Singh
Department of Mechanical Engineering, Delhi Technological University, Delhi, India

Article Info
Article history:
Received 11 January 2020
Received in revised form 29 May 2020
Accepted 30 July 2020
Available online 15 September 2020

Keywords: Missiles, Fluid Dynamics, Flow Separation, Drag Force, Drag Coefficient, Comparative Study, Aerodynamics CFD.

Abstract: Aerodynamics study using computational fluid dynamics is very famous among the engineers and scientists, because it not only reduces the cost of the entire project but also saves a lot more time. The results of the CFD simulations needed to validate through experiments. So, we can say that CFD simulation studies reduce the no. Of experiments taken, if it cannot eliminate. In this research paper, we made four different aerodynamics missiles shapes CAD models in solid works by using underlying principles, mathematical equations of different curves and engineering judgement, one of them is a missile of standard basic shape. We have analysed and compared them with basic shape of missile. Here, in this study, Volume is taken a constant parameter. Drag Coefficient is the main parameter which is evaluated and studied at different Mach no's and at a constant angle of attack. Reasons behind the magnificent drop in drag coefficient explained in discussion section.

1. Introduction

For development or modification of expensive machine or system there is a need arise to make a model before making it in real size so that we can clarify most of the problems in less perimeter which results low cost, but for making some specific parts like jet engine missiles etc. It is not suitable at all. So there is a need of computer science by which we can easily get the final result approximately equal to the actual result and also very complexity is being solved by using the computer and this process of imitation is called the simulation process

CFD simulation is one of the complex simulations by which many fluid dynamics processes and systems are being simulated over the decades CFD simulation uses many complex relations and numerical algorithms. As the human needs increases there is a continuous improvement in the CFD simulation, and very expensive systems such as wind tunnel and various types of experimental testing systems are present in this software. CFD simulation is being proved a very useful and inexpensive tool over the last few decades, but for analysis of the high complexity there is a requirement of great performance of the processor, graphic tools etc. Mostly it is used for the engineers for finding the flaws in the model and all the possibilities of technical failure and for study and analysis of the dynamics of the fluid over it. Advancing in the turbo machines and reduced noise are some common difference and considerable changes can be made by the CFD simulation. In India most of the missile designers are heavily dependent on the CFD simulation for getting results very close to the actual in propulsive and aerodynamic parameters.

Missiles are the propulsive explosive which can destroy a large area and its construction in real life is too dangerous so it is being necessity to design and analyse missiles on the CFD tool. CFD tool contains the Euler, navier stokes and other compressible fluid flow equation for the analysis of high speed aircrafts and missiles. In the CFD simulation tool we are able to explode it at any position such as in air or land on a specific co-ordinate, also we can measures the total drag experienced by the missile during the total flight. Also we can easily determine the speed and time and exact position if the wind speed varies from different positions. There may be some other flaws like development of oblique shock waves and heat transfer during flight air friction by which energy is being lost. So, at what rate of propulsion is required for getting a specific speed with highest precision. High precision is being very necessary for blasting an aircraft so, CFD simulation can also be helpful for determining the possibility of the successive collision of missiles.

Corresponding Author,
E-mail address: gauravdtu@hotmail.com**,** rk868616@gmail.com, sagardtu@yahoo.com, rajkumarsingh@dce.ac.in
All rights reserved: http://www.ijari.org

2. Computational fluid dynamics (CFD) Modelling

The development of computer plays an important role in the modern science and technology. The computer brings about computational fluid dynamics (CFD), computational structure mechanics, computational electromagnetism, etc., which produce great impact on modern aeronautics and astronautics. CFD has now been an essential tool for the aerodynamic design of aeronautic and astronautic vehicles, Next, CFD simulations were carried out to validate the results obtained from the VLM approach and to analyze the flow around the wing at high angles of attack.[1]

Whether the missile can be launched safely or not from a helicopter is a very important problem which should be taken into consideration in the design of a helicopter. The complete set[3] of three-dimensional (3-D) time-dependent Navier–Stokes equations is solved in a time-accurate manner for simulations of unsteady flow fields associated with projectiles during flight. domain of aerospace and aviation, the existence of flow separation leads to the generation of noise, the reduction of lift and the instability of flying vehicles. Researchers have been searching for the flow separation control method for a long time.[4] Generally, quasi-steady flow computation methods are adopted to calculate various release trajectories of external stores under aircraft with fixed wings. However, unsteady flow computation methods should be used to simulate the flow around a helicopter because of its rotating rotor. When it occurs, the increment of lift coefficient with angle of attack began to decrease until the maximum lift coefficient value is reached. Afterwards, when angle of attack is still increased, the lift decrease appears [5].

This paper challenges the CFD community to take on the full flight simulation problem. This challenge is referred to as Digital Flight and is defined as the ability to simulate in a computer a flight manoeuvre satisfying the governing flow equations, the aircraft Aeroelasticity characteristics, the 6-DOF equations, the flight control system, and the propulsion system. In short, we want to fly an aircraft in the computer. The missile model reliability was proved comparing numerical and experimental results in regards to pressure coefficients and aerodynamic forces (i.e. drag, lift and moment coefficients)[2]. Separation on smooth curved surfaces is more challenging as the point or line of separation is not fixed in space and is very sensitive to external flow properties. [6]

Since separation is associated with significant performance losses, its mitigation becomes important [7]. A Blasius boundary layer is known to transition to turbulence through several mechanisms, known as the K (Klebanoff) [8] Transitional flow is a complicated phenomenon. The exact point of laminar-turbulent flow transition is challenging to define and is often based on empirical data with large uncertainties obtained in the experiments mostly using air or water [9].

However, there are some situations that the transition regime cannot be avoided, such as upgrading the system as it working originally in the laminar flow or in the accidental scenarios. Until now, limited

work in the transition region has been reported due to the complexity of the influence factors [10]. The transition process is initiated through the growth and development of disturbances originating on the body or in the freestream. Environmental disturbances include atmospheric turbulence, entropy spottiness, particulates, and electrostatic discharges [11].

3. Missiles models used in this paper for comparative study

3.1 Case 1

Missile 1, as shown in figure 1, looks like a normal cylindrical shape, which is looking like a half oval as a cap resting on that cylinder, but at the front the oval is keeping sharp for minimize the coefficient of drag.

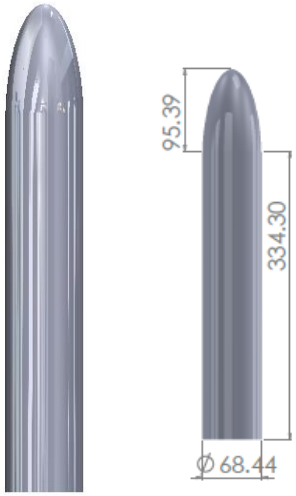


Fig1: Missile 1 (Dimensions are in mm)

Fig.2: Missile 2 (Dimensions are in mm)

Fig.3: Missile 3 (Dimensions are in mm)

As shown in figure 2, the main body of this missile is also remained cylindrical shape but in the front portion, width of the missile is larger than diameter of the cylindrical part, front part is looks like a Case 2 with a sharp nose.

From figure 3, the main body of this missile is also remaining in cylindrical shape but as we moving at the top its diameter is increases and at the point where the diameter is becomes almost 3/2 of cylinder diameter and then it becomes a half spherical shape with a sharp nose.

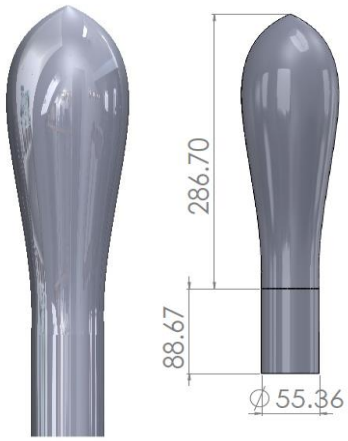


Fig.4: Missile 4 (Dimensions are in mm)

4.Results and discussion

4.1 Case 1

4.1.1 At Mach no. – 0.2

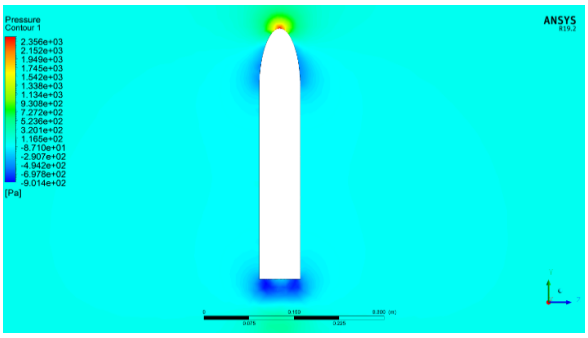


Fig. 5: Pressure Effects

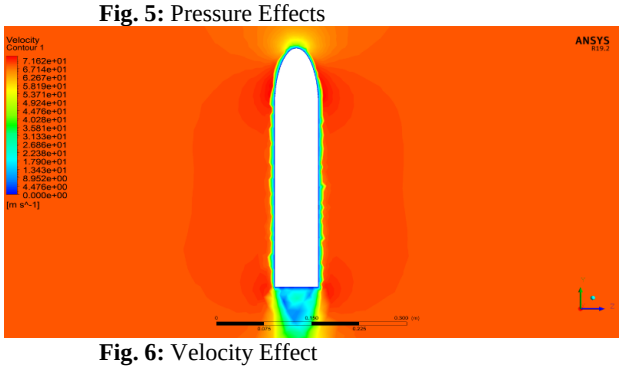


Fig. 6: Velocity Effect

4.1.2 At Mach no. – 0.5

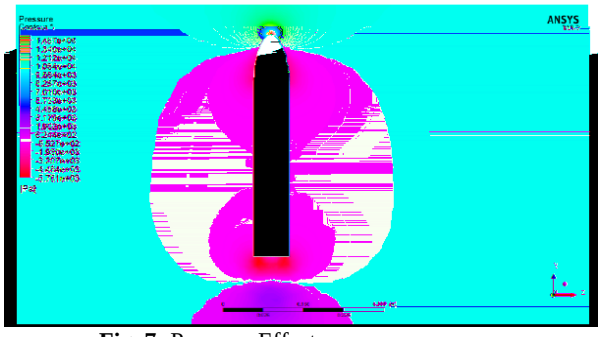


Fig. 7: Pressure Effects

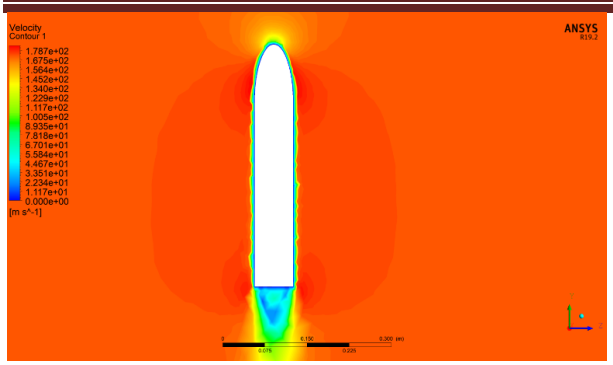


Fig. 8: Velocity Effects

4.1.3 At Mach no. – 0.75

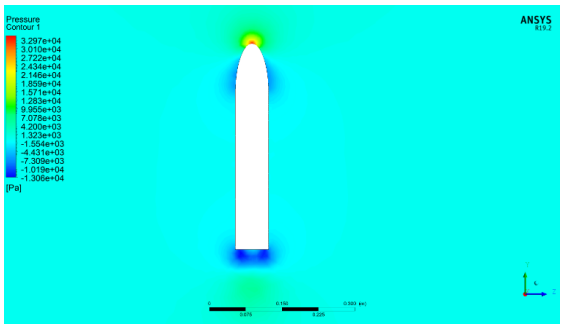


Fig. 9: Pressure Effects

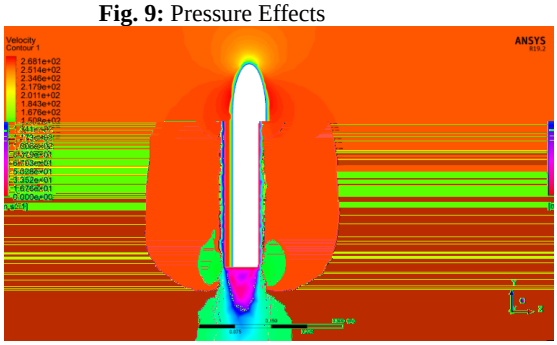


Fig. 10: Velocity Effects

4.1.4 At Mach no. – 1.0

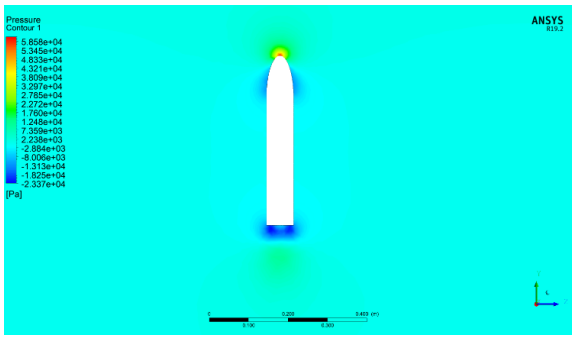


Fig. 11: Pressure Effects

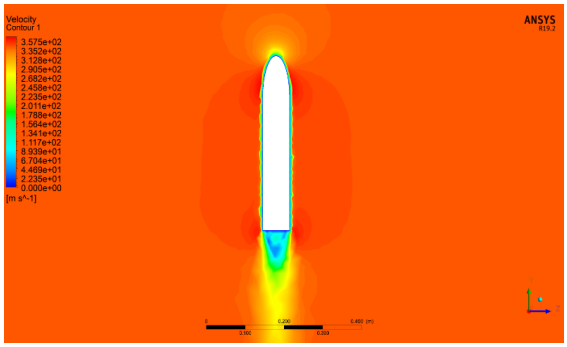


Fig. 12: Velocity Effects

4.1.5 At Mach no. – 2.0

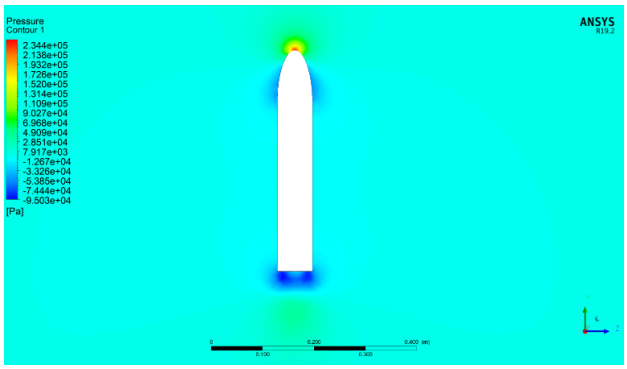


Fig. 13: Pressure Effects

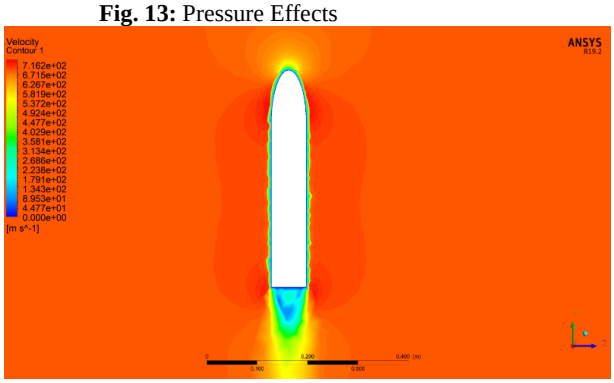


Fig. 14: Velocity Effects

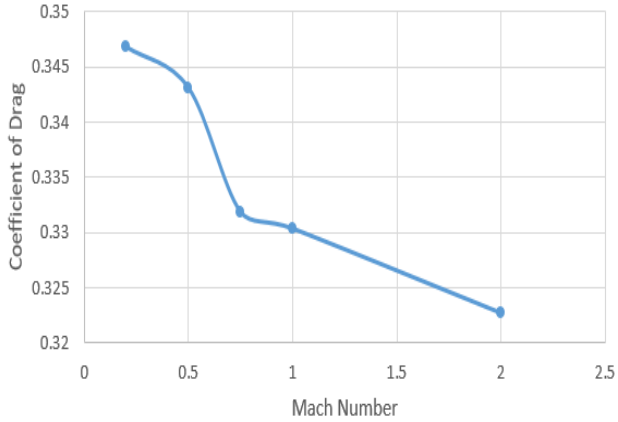


Fig.15: Coefficient of drag vs Mach number for Case 1

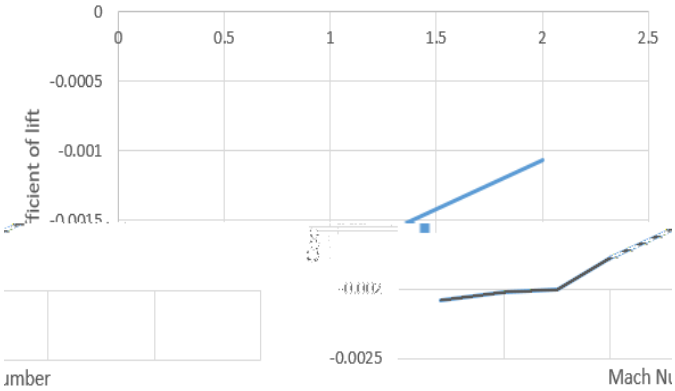


Fig.16: Coefficient of lift vs Mach number for Case 1

In Case 1, there is very high pressure at the tip of the missile and there is a vacuum pressure after the tip and a very high vacuum pressure is being created on the tail of the missile ,for boundary layer at the tip of the missile there is a thick boundary layer and then a very thin boundary layer is formed and again at the tail of the missile a very thick velocity boundary layer is formed.

For Case 1, coefficient of drag is decreasing nonlinearly with increase in the Mach number. After reaching 0.5 Mach number it decreases rapidly till 0.75, again after 0.75 Mach number it decreases in small amount. After 1 Mach number it decreases approximately linearly with respect to Mach number.
Coefficient of lift for this missile having very little increment up to 0.75 Mach number, after reaching 0.75 Mach number the curve is approximately linear but very little change is being occurred in coefficient of lift.

4.2 Case 2

4.2.1 At Mach no. – 0.2

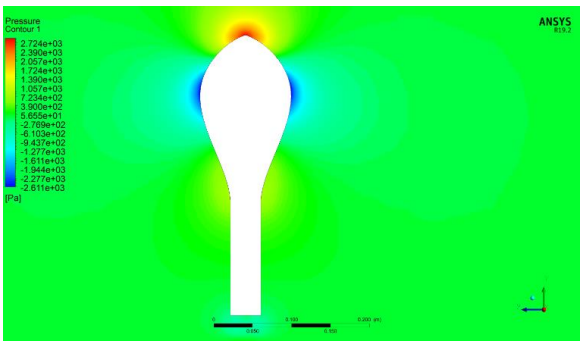


Fig. 17: Pressure Effects

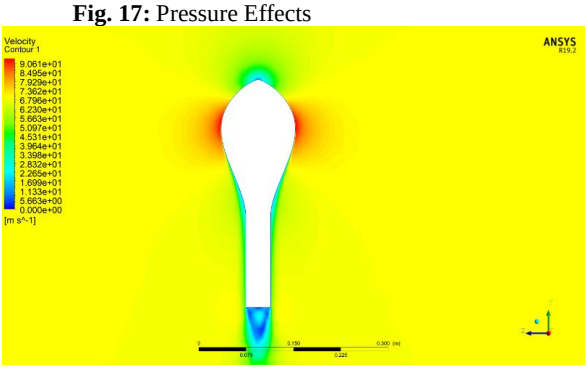


Fig. 18: Velocity Effects

4.2.2 At Mach no. – 0.5

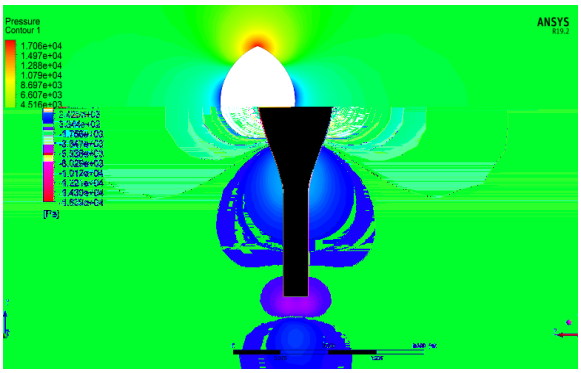


Fig. 19: Pressure Effects

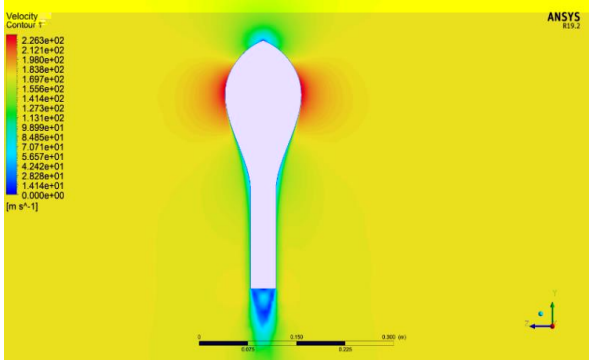


Fig.20: Velocity Effects

4.2.3 At Mach no. – 0.75

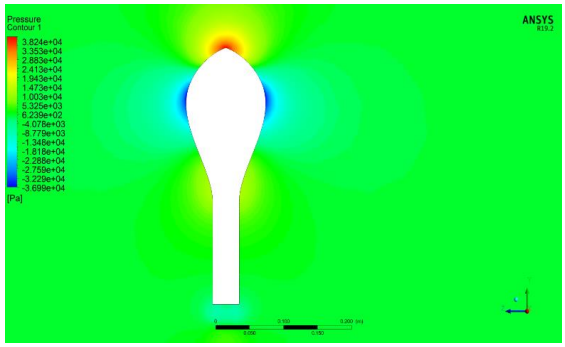


Fig. 21: Pressure Effects

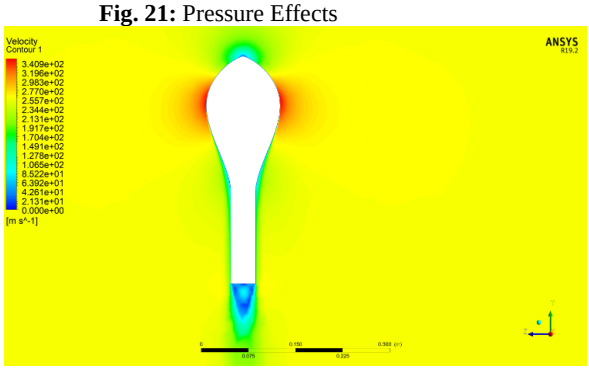


Fig. 22: Velocity Effects

4.2.4 At Mach no. – 1.0

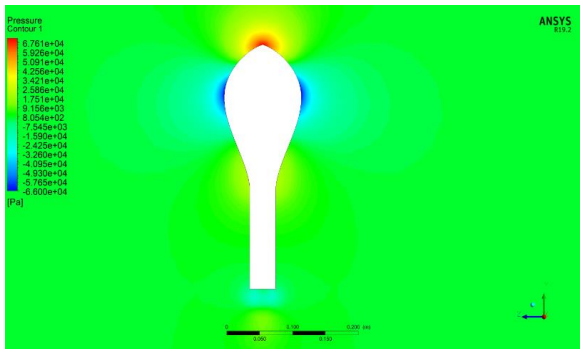


Fig. 23: Pressure Effects

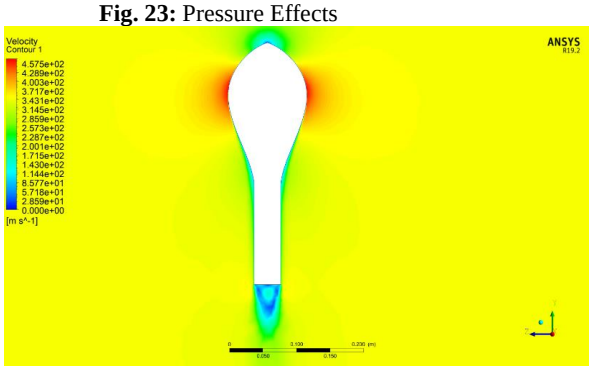


Fig. 24: Velocity Effects

4.2.5 At Mach no. – 2.0

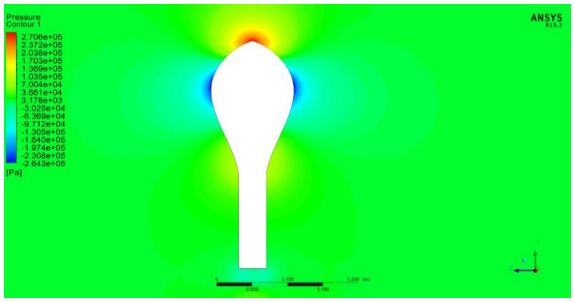


Fig. 25: Pressure Effects

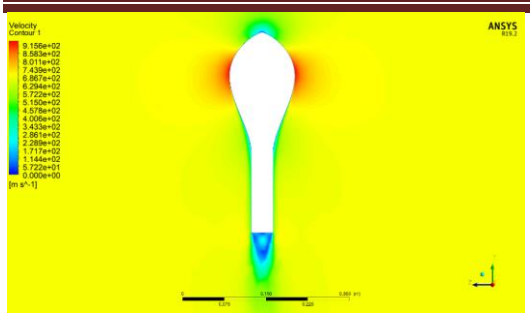


Fig. 26: Velocity Effects

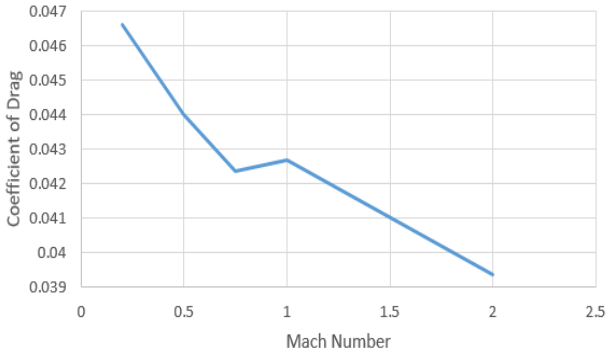


Fig. 27: Coefficient of drag vs Mach number for Case 2 missile

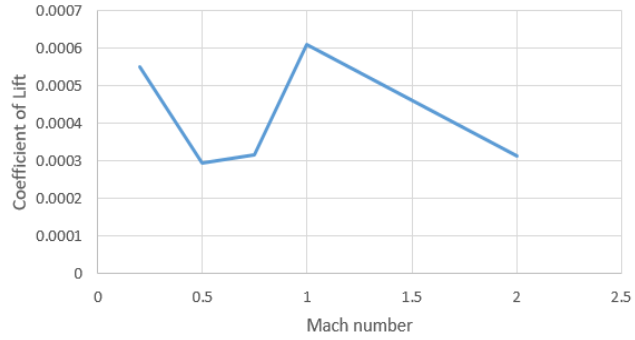


Fig. 28: Coefficient of lift vs Mach number for Case 2 missile

Missile in case 2 there is very high pressure at the tip of the missile and then after the tip of the missile there is very high vacuum pressure is developed at very small region, for boundary layer, at the tip of the missile there is sudden decrease velocity and a thick boundary layer is formed and its thickness decreases at the curve of the Case 2 where velocity is even more than free stream velocity and at the tail of the missile the boundary layer separation may be occur and a very thick boundary layer is formed.

The coefficient of drag is free meaning law as compared to Case 1, interest system coefficient of drag is decreases linearly up to 0.75 Mach number, and then it increases negligible till 1 Mach number, and again linear decrement is followed till 2 Mach number.

Coefficient of lift for this missile is varying w.r.t Mach number in irrespective nature, as from 0.2 to 0.5 it decreases, and from 0.5 to 0.75 it remains constant after 0.75 Mach number it increases and after one Mach number it decreases linearly.

4.3 Case 3

4.3.1 At Mach no. – 0.2

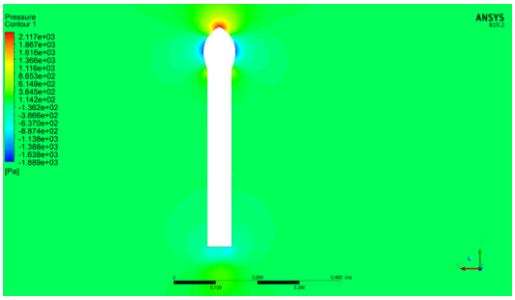


Fig. 29: Pressure Effects

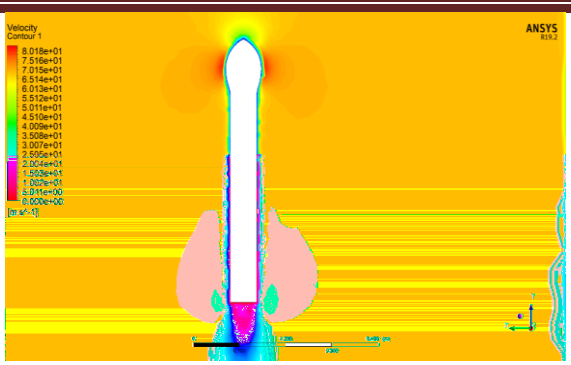


Fig. 30: Velocity Effects

4.3.2 At Mach no. – 0.5

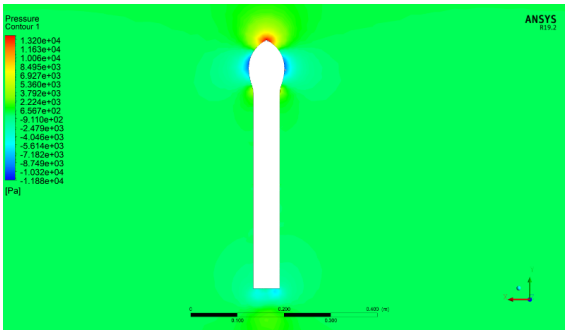


Fig. 31: Pressure Effects

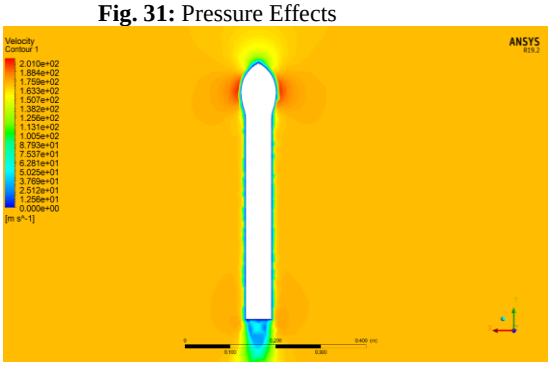


Fig. 32: Velocity Effects

4.3.3 At Mach no. – 0.75

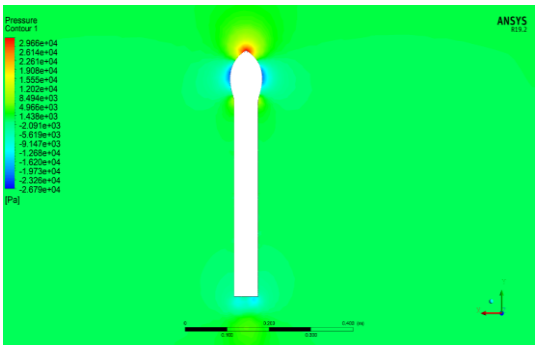


Fig. 33: Pressure Effects

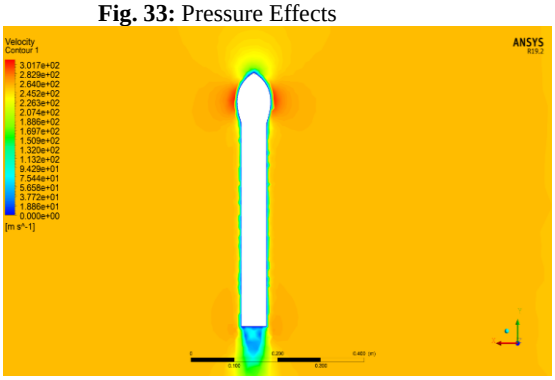


Fig. 34: Velocity Effects

4.3.4 At Mach no. – 1.0

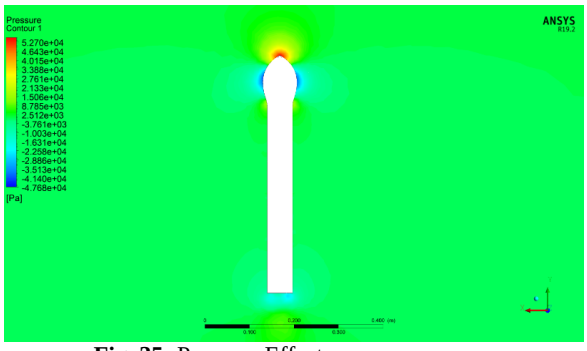


Fig. 35: Pressure Effects

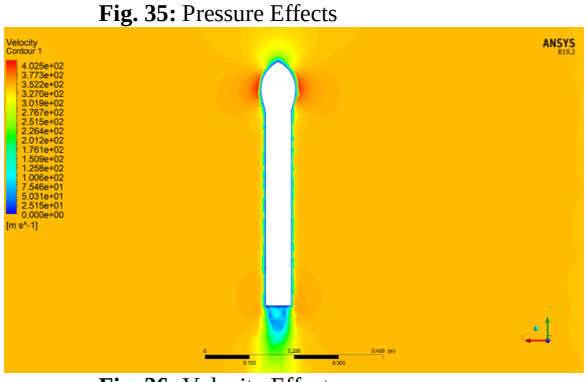


Fig. 36: Velocity Effects

4.3.5 At Mach no. – 2.0

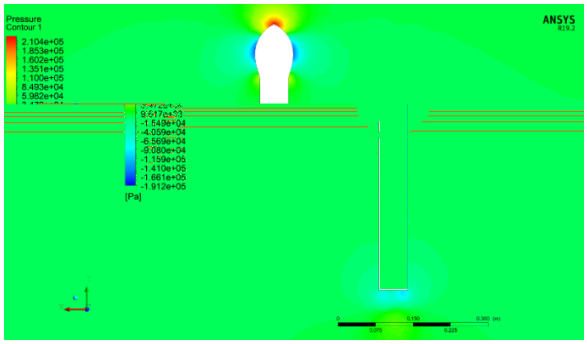


Fig. 37: Pressure Effects

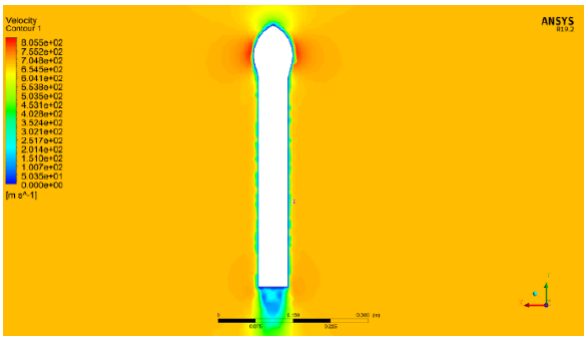


Fig. 38: Velocity Effects

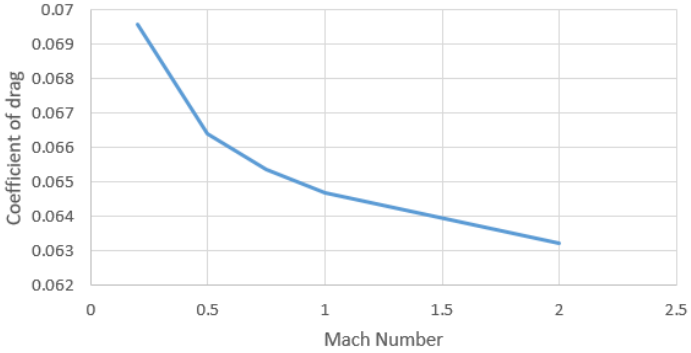


Fig. 39: Coefficient of drag vs Mach number for Case 3

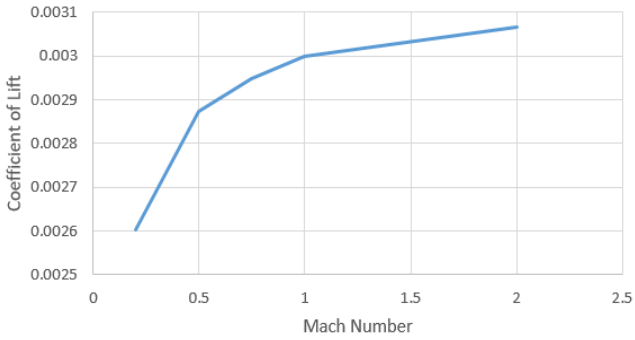


Fig. 40: Coefficient of lift vs Mach number for Case 3.

Here, there is very high pressure at the tip of the missile and at few distance from the tip where vacuum pressure is being found in very small area and at the tail low vacuum pressure is created, for boundary layer , at the tip of the missile there is sudden decrease in the velocity and after a few distance high velocity profile is being created even more than free stream velocity and at the tail of the missile a thick boundary layer is formed also boundary layer separation may be occur.For Case 3 missile system coefficient of drag is low as compared to Case 1, coefficient of drag is decreases inversely with respect to Mach number, as from 0.2 to 0.5 Mach number it decreases rapidly and then its rate of decrement of C_d is also decreases continuously.

4.4 Case 4

4.4.1 At Mach no. – 0.2

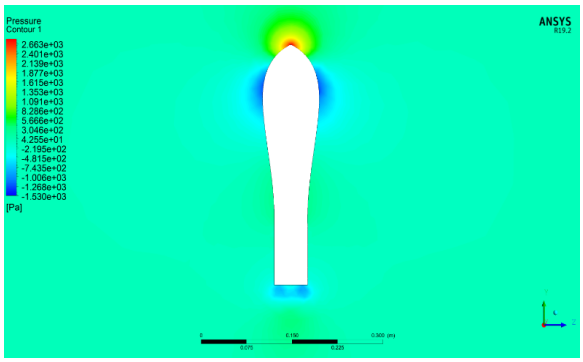


Fig. 41: Pressure Effects

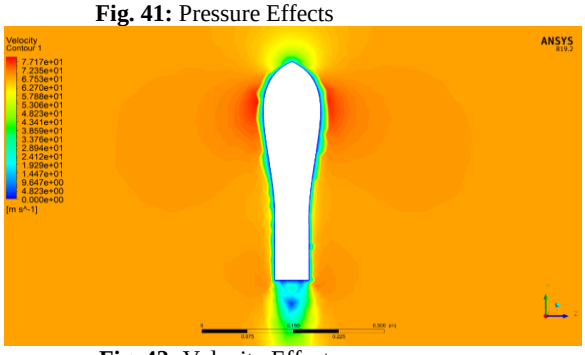


Fig. 42: Velocity Effects

4.4.2 At Mach no. – 0.5

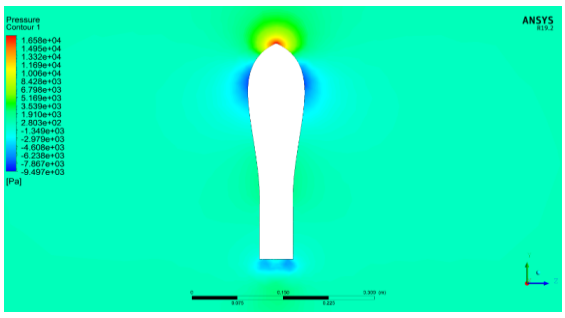


Fig. 43: Pressure Effects

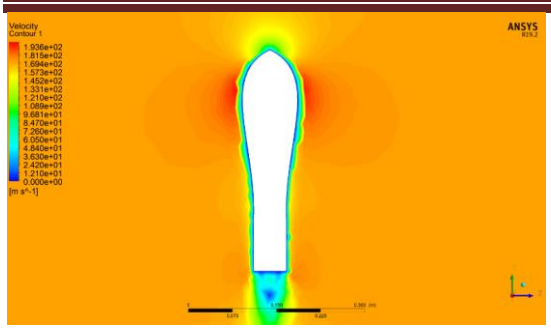


Fig. 44: Velocity Effects

4.4.3 At Mach no. – 0.75

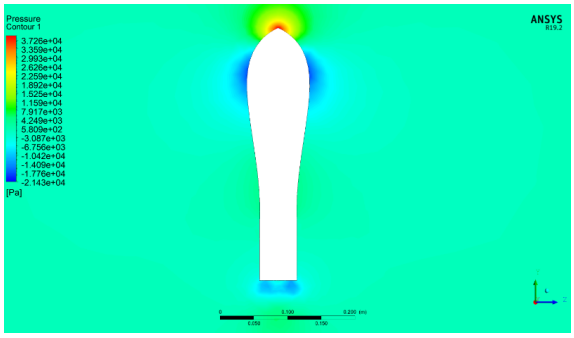


Fig. 45: Pressure Effects

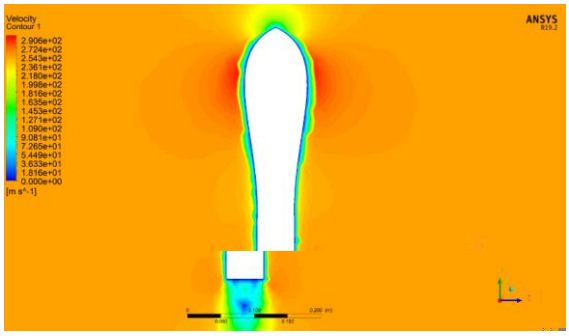


Fig. 46: Velocity Effects

4.4.4 At Mach no. – 1.0

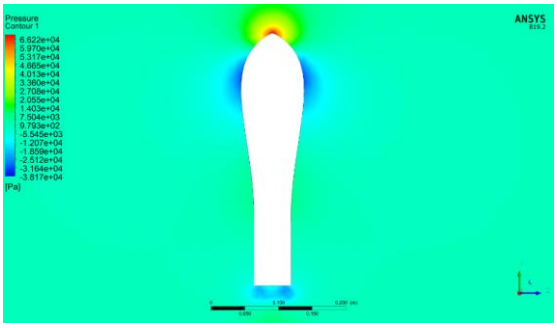


Fig. 47: Pressure Effects

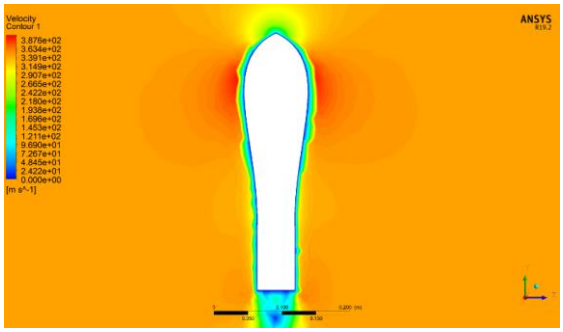


Fig. 48: Velocity Effects

4.4.5 At Mach no. – 2.0

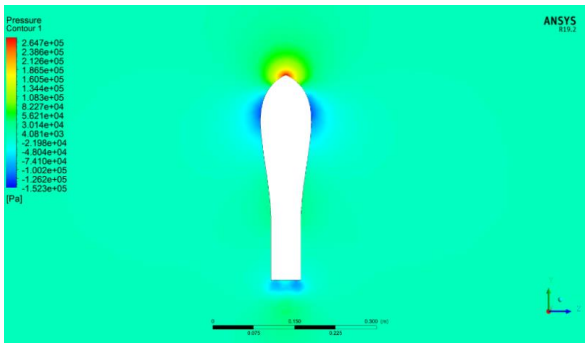


Fig. 49: Pressure Effects

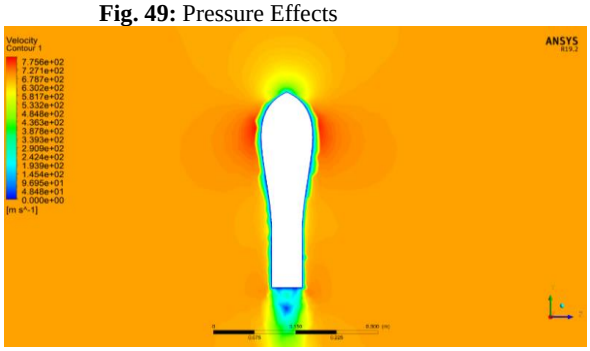


Fig. 50: Velocity Effects

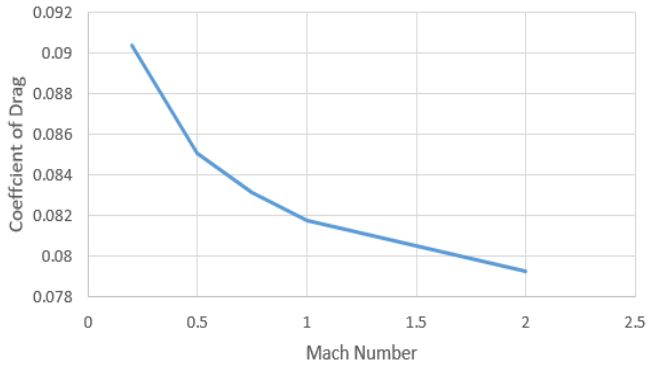


Fig. 51: Coefficient of drag vs Mach number for Case 4

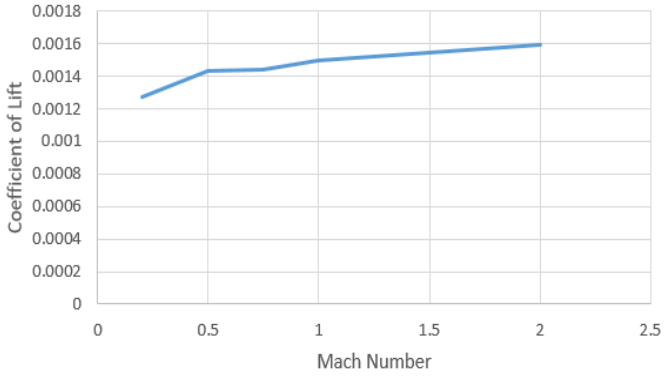


Fig. 52: Coefficient of lift vs Mach number for Case 4

In case 4 missile system there is high pressure at the tip of the missile which is followed by a vacuum pressure is created at the curve which is at the short distance away from the tip and again vacuum pressure is being created at the tail of the missile, for boundary layer, there is a sudden decrease in the velocity at the tip which is followed by high velocity even more than free stream velocity for the total curved area, and at the tail again very low velocity profile is being developed or thick boundary layer is formed.

For Case 4 missile system coefficient of drag is inversely proportional to the Mach number in the early stage from 0.2 to 0.5

Mach number there is a sudden decrease in coefficient of drag and its decreasing rate is reduced continuously.

Coefficient of lift for Case 4 missile is increasing with reducing rate as for Mach number 1 to 2 there is very little change is observed.

Table 1: Coefficient of Drag for all Mach numbers

	Mach Number				
Case	**0.2**	**0.5**	**0.75**	**1**	**2**
-1	0.3468	0.3431008	0.3318876	0.3303848	0.3227552
-2	0.046602	0.044008	0.042367	0.042706	0.03938
-3	0.069586	0.066393	0.065363	0.064702	0.063243
-4	0.090396	0.085087	0.083129	0.081788	0.079249

Table 2: Coefficient of Lift for all Mach Numbers

	Mach Number				
Cases	**0.2**	**0.5**	**0.75**	**1**	**2**
1	-0.0020808	-0.00202011	-0.0020002846	-0.0017713388	-0.0010720455
2	0.0005499	0.000296	0.000317	0.000611	0.0003138
3	0.0026028	0.0028746	0.0029493	0.0029983	0.0030666
4	0.0012723	0.0014327	0.0014435	0.0015005	0.0015948

5.Conclusions

As the numerical study has demonstrated we are able to know about the drag coefficient acting on the different missile systems. Drag coefficient is depend on many factors such as roughness, variation in velocity and pressure of the fluid and flow separation. The effect of flow separation in the missile or any vehicle is play a great importance to the coefficient of drag, it increases the coefficient of drag to great extent and in this project we are simulating four types of missile using CFD simulation by ANSYS, for this simulation we should take equal free stream velocity, equal amount of volume and weight, identical material so roughness is being equal and identical atmospheric pressure (approximately).

Case 1 is very simple type of missile and in this missile very low vacuum pressure is developed as compared to the other types of missile systems whereas the flow separation is not being affect very much and finally due to low vacuum pressure there is very high drag coefficient. For Case 2 missile we create the front part of the missile to look like a Case 2 so that drag coefficient to be minimum for the foremost part but by using this theory the flow separation is affected the coefficient of drag, flow separation is occurred at rearward part of the Case 2 so, drag coefficient is increased but it is lesser than the Case 1 . For Case 3nd type of missile system the shape of the foremost part is may be considered as the mixture of both Case 2 and Case 1 , this type of missile is often used in the real life application in this type the coefficient of drag is minimum in the all four types of missiles, in this type there is very high vacuum pressure is being created at the most emerged part of the foremost portion in very small region and no flow separation is being observed or flow separation is being occurred insignificantly so, it is the most optimized design of the missile system it may be possible to decrease the coefficient of drag by making it Case 4 . For Case 4 type of missile system the figure of the foremost part is a little bit smaller than the Case 2 type missile and it is modified by designing it as a Case 4 so that no possibility of flow separation, in this system the high vacuum pressure is also generated at the most emerged part of the foremost portion, in this system coefficient of drag is very close to the drag coefficient of the Case 3nd missile but due to lower vacuum pressure than Case 3nd , the coefficient of drag of the Case 4 missile is being more (approximately) than the drag coefficient of Case 3nd system. But this system shows more continuous curve of drag coefficient vs Mach number. It is also used in small missile applications.

Now we concluded that Case 3 missile system having lowest drag coefficient. Case 4 system having little bit more drag coefficient than the Case 3 but in this system no flow separation is occurred so, it is being the best design to develop and most optimized design. Case 2 missile having coefficient of drag lesser than Case 1 system. Case 4 missile system is become the most optimized design in perspective of the flow separation and other factors of coefficient of drag, but we have not study the variation of the stresses developed in under the missile and some other factors. So, if you want to further study about these topics you can do it in future.

Reference

[1] S Piedra, F Martinez, CA Escalante-Velazquez, SMA Jimenez. Computational aerodynamics analysis of a light sport aircraft: compliance study for stall speed and longitudinal stability certification requirements. Aerospace Science and Technology. 2018, doi:10.1016/j.ast.2018.09.016

[2] F Rizzo, V D'Alessandro, S Montelpare, L Giammichele . Computational study of a bluff body aerodynamics: Impact of the laminar-to-turbulent transition modelling. International Journal of Mechanical Sciences, 178, 2020, 105-620. doi:10.1016/j.ijmecsci.2020.105620

[3] J Sahu, KR Heavey. Parallel CFD computations of projectile aerodynamics with a flow control mechanism. Computers & Fluids, 88, 678–687. doi:10.1016/j.compfluid.3 (15), 2013.

[4] Y Zhou, L Hou, D Huang. The effects of Mach number on the flow separation control of airfoil with a small plate near the leading edge. Computers & Fluids, 156, 2017, 274–282.

[5] R Placek, P Ruchała. The flow separation development analysis in subsonic and transonic flow regime of the laminar airfoil. Transportation Research Procedia, 29, 2018, 323–329. doi:10.1016/j.trpro.2018.02.029

[6] P Razi, P Tazraei, S Girimaji. Partially-averaged Navier–Stokes (PANS) simulations of flow separation over smooth curved surfaces. International Journal of Heat and Fluid Flow, 66, 2017, 157–171. doi:10.1016/j.ijheatfluidflow.2017.05.005

[7] G Jones, M Santer, M Debiasi, G Papadakis. Control of flow separation around an airfoil at low Reynolds numbers using periodic surface morphing. Journal of Fluids and Structures, 76, 2018, 536–557. doi:10.1016/j.jfluidstructs.2017.11.008

[8] KL Kang, KS Yeo. Combined effects of amplitude, frequency and bandwidth on wavepackets in laminar turbulent transition. Computers & Fluids, 197, 2020, 104-358. doi:10.1016/j.compfluid.2019.104358

[9] C Mills. High frequency pressure loss measurements of laminar-turbulent transitional flow. Flow Measurement and Instrumentation, 74,2020, 101-770. doi:10.1016/j.flowmeasinst.2020.101770

[10] W Chang, G Pu-zhen, W Zhan-wei, T Si-chao. Experimental study of transition from laminar to turbulent flow in vertical narrow channel. Annals of Nuclear Energy, 47, 2012, 85–90. doi:10.1016/j.anucene.2012.04.018

[11] SP Schneider. Hypersonic laminar–turbulent transition on circular cones and scramjet forebodies. Progress in Aerospace Sciences, 40(1-2), 2018, 1–50. doi:10.1016/j.paerosci.2003.11.001

View publication stats

Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Crowd Emotion Analysis Using 2D ConvNets

Gaurav Tripathi
Department of ECE,
Delhi Technological University,
Delhi, 110042, India
gauravtripathy@gmail.com

Kuldeep Singh
Department of Electronics &
Communication Engineering, MNIT,
Jaipur-302017
kuldeep.ece@mnit.ac.in

Dinesh Kumar Vishwakarma
Department of IT, Delhi
Technological University, Delhi,
110042, India
dinesh@dtu.ac.in

***Abstract*—Timely detection of the crowd emotions can lead to an effective crowd management measure. Crowd can exhibit multiple behaviour having separate emotions underneath it. Accurate emotion analysis of the crowd from a crowded image plays a critical role in the prudent examination of the crowded area. A popular non-intrusive solution for crowd behavior analysis is based on videos obtained through ambient camera, and the corresponding methods usually require a large dataset to train a classifier and are inclined to be influenced by the image quality. Detection of emotion of crowds becomes more pertinent in case of religious and political rallies. Peaceful conduct of these events is important for saving of human lives. The paper presents a novel method of deciphering crowd emotions using 2D convolutional neural network (ConvNets). This framework is then used to predict the crowd emotions, which leads to the subtle hints about the overall behaviour of crowd. The paper presents a suitable classifier for crowd emotion using the self-curated datasets for proving the concept that emotions can be extracted using ConvNets. Experiments have verified the proposed scheme on crowd behavior benchmark with fair accuracy.**

Keywords—Crowd emotions, crowd behaviour, convolutional neural network, deep learning, XceptionNet

1. INTRODUCTION

Crowd management has become one of the most important field in the modern cities. The advent of surveillance medium and usage of CCTV has increased the capability of surveillance. Effective strategies to manage crowd using surveillance is being developed. Crowd behaviour can contain many activities such as protesting, shouting, looting and so on. These behaviour has one underlying principle which is common across all its behaviour. This is known as crowd emotions. Emotions are an important part of any behaviour analysis whether it concerns individual or group of people However deciphering mass emotions becomes tricky as it is simply not the summation of individual emotional. To cater to the need of effective management of crowd, underlying emotions needs to be read and applied to the overall behaviour deciphering scheme for a proper crowd surveillance.

Cognitive aspects of human mind play a conscious role in the deciphering human emotions. Individual emotions fuse themselves and evolve in to more profuse crowd emotions. As per [1], cognitive aspects of human mind emotions, has an impact in determining crowd behaviour. Collective emotions form a basis of many mass activities which may be normal or abnormal events. The underlying basis of collective emotion lies in the individual emotions manifesting in the separate individuals [2]. As per the convergence theory of mass psychology [3]states that characteristics shown by a crowd is directly proportional to the emotion of likeminded entities. It has been categorically proved that emotion forms a strong basis of human behaviour
[4]. Simple emotions have shown unique set of behavioral and psychological patterns [5]. Thus, individual emotions on a common theme that can result in a common collective behaviour manifestation in the crowd behaviour. Based on above arguments this paper presents an approach to predict crowd emotions.

The computer vision has presented many challenges. Crowd behaviour analysis is one such challenge which evolved from the nodal concepts of object tracking, object detection, and their respective route analysis. A crowd is gathering of large number of people displaying some sort of coherent movement [6]. Another viewpoint of labelling crowd comes from mathematical model-based psychologists which describes the crowd in terms of their path followed and allow mobilization of people only in some emergency [7]. Inherent behaviour analysis of crowd is needed for a successful implementation of law and order in the big cities. In a recent survey [8] for crowd behaviour it was described that overall crowd behaviour analysis consist of the following categories:

- **People detection and tracking: 34%**
- **Abnormality detection 30%.**
- **People counting: 17%**
- **Crowd behavior analysis: 19%**

Effective crowd management and monitoring strategies is needed in public places such as shopping malls, temples, stadium etc. The major theme of crowd analysis is to ensure public safety. It is concluded that 19 percent of the crowd displays the emotional behaviour. Emotions may not be visible while tracking the person but in behaviour analysis

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

emotions play a vital role. Hence emotions become important for crowd behaviour study. The underlying emotions of crowd needs to be studied and deciphered for a more successful crowd management strategy. Some sample images of angry crowd emotions are as follows:

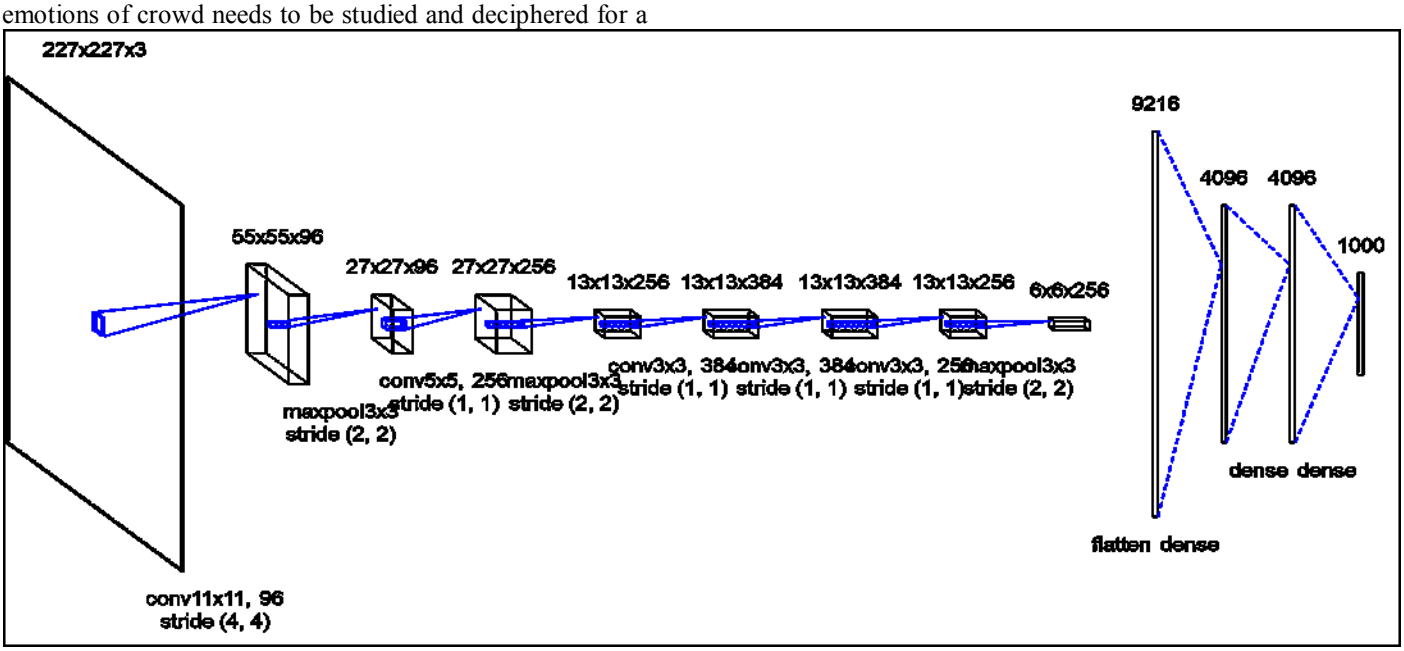


Fig.1. A standard 2D ConvNets

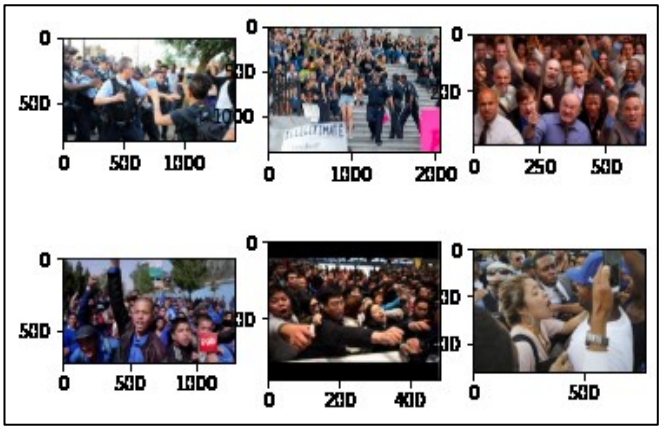


Fig.2 Sample Angry Crowd Images

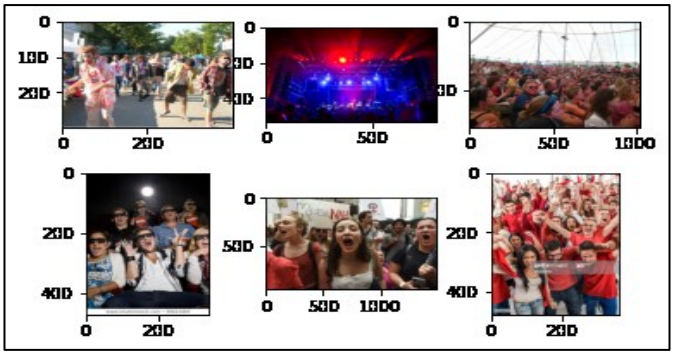


Fig. 3. Sample of Excited and Happy Images

The above images give the idea about crowd emotions. Since there is a large gathering of crowds for a unifying cause, it is important to detect crowd emotions using surveillance so that governing authorities must be aware of the inherent crowd emotions development. There have been previous attempts to study crowd emotions. Computer vision is the new normal which is breeding fast for various vision related problems. Getting cues from the surveillance cameras, smart phone and images from social media is an apt way to get the input to the ConvNets. These ConvNets process the images to classify the images as containing protests or non-protesting images. Our contributions in this paper is as follows:

- Curating image datasets for crowd emotions. Following are the emotions that have been catered in the paper:

Angry	Excited	Sad	Happy	Scared	Neutral

- A deep convnets based emotion classification model using a model the variation of Xception network.
- Training a 2D Convolutional Neural Network (ConvNets) model from scratch.

The paper is organized in following major sections. Section 2 presents a brief introduction of ConvNets and Transfer Learning. Section 3 presents the existing work in the area of emotion classification. Section 4 presents examine the overall framework of the emotion classification architecture. Section 5 presents experiments and results of the emotion classification. Section 6 ends the paper with discussion and key takeaways of the results.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

2. CONVNETS DESCRIPTION

ConvNets has become a significant tool for challenging computer vision applications [9] [10] [11]. Convolutional neural networks (CNNs) facilitates the feature extraction and classification process for vision related problems [12] .An apt discussion can be found on ConvNets in [13]. ConvNets are suited for feature extraction process as well classification problems especially in computer vision problems. Since ConvNets have progressed, two popular options exist while training the images. One option is to train the domain specific problem statement from the scratch and the other is to use pre trained model, which is often known as transfer learning. This paper uses training from scratch using random initialization of weights. Automatic extraction of discriminant features is the highlight of any competent. Following diagram aptly represents the inputs and series of convolution, pooling and classification operations that is followed in any ConvNets operation. The paper utilizes the 2D Convolution Neural Network learning approach for classifying crowd emotions on the curated dataset.

3. RELATED WORK

Emotions form an important basis of human actions. Every single human being shows traits of emotions in daily activities. Similarly, crowd also exhibits emotions. Individual emotions may be different but crowd comes for a common cause probably to show dissent or celebrate o protest and so on. Each of the crowd activities is imbibed with an inherent emotion. Normal human emotions have been shown in table 1. There have been previous attempts to link personality traits and emotions. Different personalities exhibit different emotions. Psychological factors such as personality traits and emotions are also explored in detecting crowd behaviors by [14] . One of the given experiments [15] discusses about the crowd behaviour such as the aggressive behaviour, impulsive behaviour, tense behaviour, assertive behaviour and active behaviour. Such simulation has shown a direct link between emotions and crowd actions. A barrage of positive and negative emotions has been studied within the context of crowd evacuation activities in [16]. It utilized the artificial bee colony for deciphering optimal situations and strategies. [17] experiments with psychological elements of emotions when a group of people changes from audience to mob. It utilizes OCC emotional theory with PAD factors and OCEAN personality model for identification of crowd behaviour in a controlled simulated environment. Similar experiments involving OCEAN and OCC theory has been experiment in [18]. It also involves psychological analysis in to crowd behaviors. Deep learning has been applied to predict crowd behavior on the basis of emotions [19]. The point is that deciphering emotions in the crowds can be a key factor in crowd related activities and that is what explored in the present work. A virtual reality system has also proven the significance of emotions in a therapeutic cause post-accident. Significant improvement was seen in the therapy and mental simulations as explained in [20]. A recent survey also discusses about neural networks with respect to the human performances. This discusses about the neural architectures utilized in various applications for automatic classification and explores the significance of automated feature extraction [21].

4. PROPOSED METHODOLOGY

The paper focusses on the variation of Xception network [22]. This network consists of stacking of depth wise separable convolutional with residual connections. This depth wise steps reduces the computational cost and memory requirements. Also, the introduction of residual connection [23] facilitates to remove issue of vanishing gradients and creates a shortcut in the sequential network. This operation helps the architecture to access earlier outputs as an input to the later layer after utilizing summation operation. The input layer to the model consists of $300 \times 300 \times 3$ dimensions of images. Then the model deploys the 2D Convolutions along with rectifier linear unit activation function. Followed by max pooling function. The last dense layer has output filter set to 6 for classifying the images as angry, sad, happy, scared and neutral images. The following diagram represents the architectural diagram used in protest classification in the current framework.

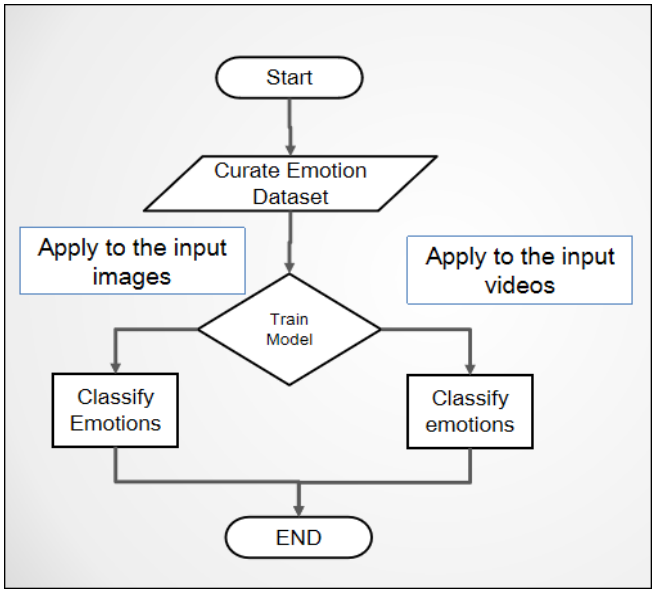


Fig.4. Flow Chart of the proposed scheme

In the above flowchart, it is clearly viewed that first the sic types of specific emotions are curated and stored in a dataset. Then the model is trained for suitable number of iterations. This training generates an emotion to classify.h5 files. Then the model is saved and whenever there is a need the model is deployed. Then this model is deployed in flask server. The input to this server is given by an input request and the response is recorded in json format. Various emotions that is seen in the images is stored in the separated folder so as to access the system. If the images are geotagged, the system

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

knows that the emotional classification is linked to the geo-location of the images. This helps the system to identify the developing emotions in a particular geo-location here crowds have gathered and displaying their respective activities.

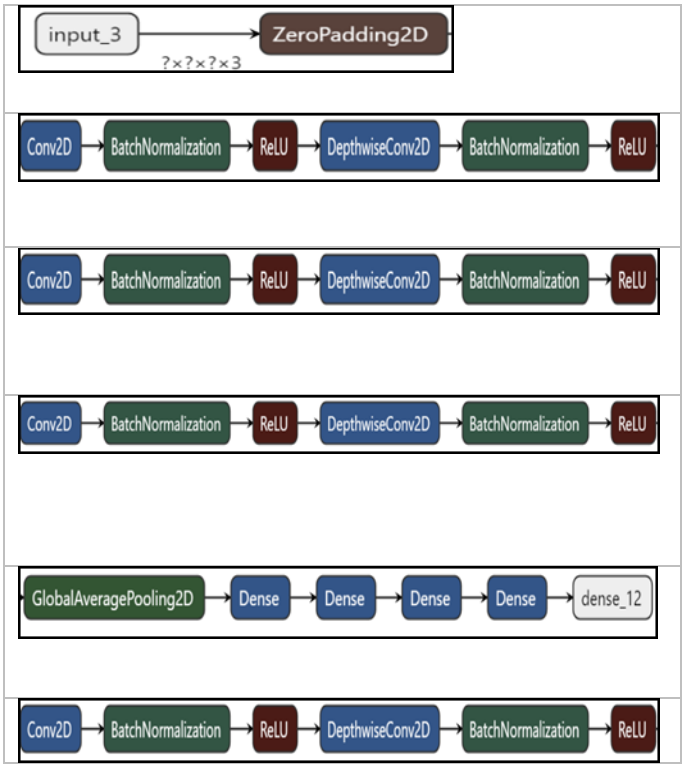


Fig.5. 2D ConvNets Architecture

It is clearly inferred that the 2D ConvNets uses series of convolutional network with max pooling function. It is important to note that the last dense layer is used for classification among two groups of datasets, namely protest and celebration. A proper protest dataset was curated using Bing image downloader [24] .Our dataset contains 6000 images which have been collected using Bing image downloader from popular search engine. The 2D ConvNets based classification is used only for images.

5. Experiment and Results

The concept is proven by utilizing the dataset that is curated using popular search engines and Bing image downloader. The primary aim is to detect protests in any given image. To justify the protests activity detection, the celebration is chosen as the other dataset. Between these two activities, the model is initiated for training on this dataset. For the implementation of proposed protest detection in crowded places, the standard Keras application is used with Python on Ubuntu operating system and experiments were conducted on a standard machine (Intel Core i5,8 GB RAM) equipped with NVIDIA TITAN X GPU. The training was performed to generate a model file. Since the dataset is small the number of epochs is set to 30. The model is described in figure 3. This is a variation of Xception network which has given positive results in the given dataset. The model uses Keras [25] version 2.3.1 and the underlying environment is that of Tensorflow. Some of the sample attributes of the Xception network model is described below:

Table 1: Node Properties of Pretrained Model network

	Zero Padding	Conv2D	Depthwise Conv2D	Dense
Activation Function		linear	linear	softmax
Bias Initializer		Zeros ()	Zeros ()	Zeros ()
Data Format	float32	float32	float32	float32
Dilation rate		1, 1	1, 1	-
Filter		32	32	-
Kernel Initializer		distribution: "uniform"	distribution: "uniform"	distribution: "uniform"
Kernel Size		3, 3	3, 3	-
Padding	[0, 1], [0, 1]	valid	valid	-
Strides		2,2	2,2	-
Trainable	false	true	true	true
Use Bias		false	false	true

The above table clearly brings out the attributes of the nodes that is part of the mini Xception network. The combination of Conv2d and max pooling is repeated six times in the network for training purpose. After training a classification model is generated and is used for classification of protesting images. Another method for pre trained network is used. Moreover this research work has also used the metric accuracy for image classification evaluation. Accuracy is defined as the proportion of tweets that has been correctly classified among all image content. Accuracy is a very intuitive metric. It is computed using below equation, where TP is True Positive, TN is True Negative, FP is False Positive, FN is False Negative;

$$\textbf{Accuracy} = \frac{\textbf{TP} + \textbf{TN}}{\textbf{TP} + \textbf{TN} + \textbf{FP} + \textbf{FN}} \times \textbf{100\%} \qquad (\textbf{1})$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

Our model achieved **93.243**% accuracy on the validation set:

Table 2: Class wise emotions accuracy

Angry	Excited	Sad	Happy	Scared	Neutral
92.41	**93.11**	**93.21**	**93.42**	**94.1**	**90.27**

.

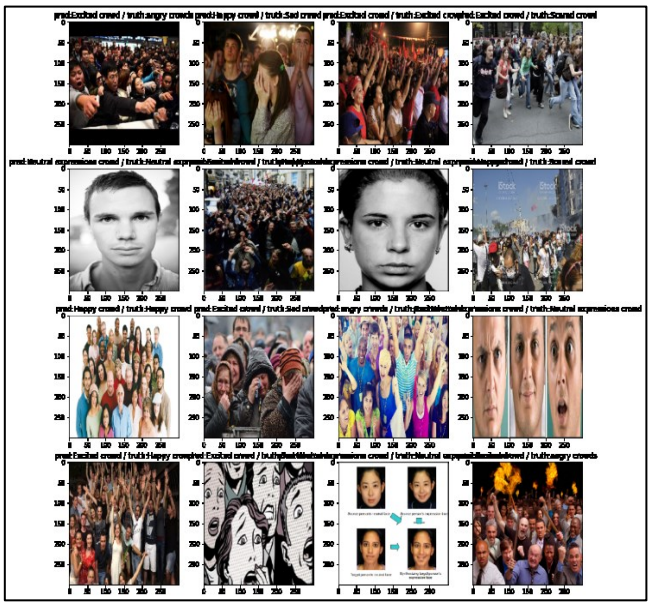


Fig.6. Emotion classification results

The above figure clearly brings out the results of our approach for successful modelling of crowd emotion classification. Angry, sad, happy, sacred and neutral emotions have been displayed in the above figure. The model accuracy is displayed in the figure below:

Fig 7. Accuracy result of the 2D ConvNets emotion classification

Fig 8. Model loss result of the 2D ConvNets emotion classification

Emotion is seen everywhere. Showing dissent and disagreement is part of human psychology. All governments around the globe allows protests as part of governing process. To make the process peaceful and non-violent, protest identification and classification absolutely necessary.

6. CONCLUSION

From this research work it is concluded that to make the process more peaceful and non-violent, the emotion identification and classification becomes absolutely necessary. If the process is automated, crowd emotion classification serves the purpose of alerting the authorities. With the advent of smart cities and huge internet infrastructure, crowd emotion classification becomes an advantageous tool in the surveillance process of cities and key public areas. Automatic emotion classification using deep learning has made the job easier for humans. Deep learning frameworks have unique capability of learning discriminant features. These features are preserved for training and generating an efficient model for classification purposes. The contribution of this paper is unique as it classifies crowd emotion in six categories. The interest in crowd emotion classification has become more significant as gathering of public, mob or crowd can turn violent and can harm human lives and public properties. The model can be extended for single image, bulk images and for video also. The approach and modelling of deep learning framework for emotion classification can be advantageous for surveillance across all the cities. The limitations of this proof of concept is that the number of images is less as compared to very large datasets. The emotion displayed in the images can be interpreted differently by human operators. It is difficult to equate individual emotions to crowd emotions.

The model can be deployed on to local server for in-house organization process, or on a web server where authorities can see the real time results of developing crowd emotion classification. In the future this model can be integrated with any Android technology for more comprehensive and localized solution. Also dataset is small and there is good chance of overfitting. Hence this architecture can be repeated over some large curated datasets on a wide variety of crowd emotions.

References

[1] W. Challenger, W. Clegg and A. Robinson, "Understanding crowd behaviours: Guidance and lessons identified," *UK Cabinet Office,* pp. 11--13, 2009.

[2] O. J. Urizar, E. I. Barakova, L. Marcenaro, C. S. Regazzoni and M. Rauterberg, "Emotion estimation in crowds: a survey," 2017.

[3] R. H. Turner, L. M. Killian and others, Collective behavior, Prentice-Hall Englewood Cliffs, NJ, 1957.

[4] R. F. Baumeister, K. D. Vohs, D. C. Nathan and L. Zhang, "HHow emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation," *Personality and social psychology review,* vol. 11, no. 2, pp. 167--203, 2007.

[5] Y. Baveye, C. Chamaret, E. Dellandr{\'e}a and L. Chen, "Affective video content analysis: A multidisciplinary insight," *IEEE Transactions on Affective Computing,* vol. 9, no. 4, pp. 396--409, 2017.

[6] R. Mehran, A. Oyama and M. Shah, "Abnormal crowd behavior detection using social force model," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2009, pp. 935--942.

[7] J. Wang and Z. Xu, "Spatio-temporal texture modelling for real-time crowd anomaly detection," *Computer Vision and Image Understanding,* pp. 177-187, 2016.

[8] G. Tripathi, K. Singh and D. K. Vishwakarma, "Convolutional neural

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

networks for crowd behaviour analysis: a survey," *The Visual Computer,* vol. 35, no. 5, pp. 753--776, 2019.

[9] A. Krizhevsky, I. Sutskever and E. H. Geoffrey, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 2012.

[10] K. Simonyan and A. Zisserman, "Two-stream convolutional networks for action recognition in videos," in *Advances Neural Information Processing Systems*, 2014.

[11] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *CVPR*, 2014.

[12] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE international symposium on circuits and systems*, 2010.

[13] M. Perez, A. C. Kot and A. Rocha, "Detection of Real-world Fights in Surveillance Videos," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

[14] S. J. Guy, S. Kim, M. C. Lin and D. Manocha, "Simulating heterogeneous crowd behaviors using personality trait theory," in *Proceedings of the 2011 ACM SIGGRAPH/Eurographics symposium on computer animation*, 2011, pp. 43--52.

[15] A. Bera, T. Randhavane and D. Manocha, "Aggressive, Tense or Shy? Identifying Personality Traits from Crowd Videos," in *IJCAI*, 2017, pp. 112--118.

[16] G. Zhang, D. Lu and H. Liu, "Strategies to utilize the positive emotional contagion optimally in crowd evacuation," *IEEE Transactions on Affective Computing,* 2018.

[17] F. Durup{\i}nar, U. G{\"u}d{\"u}kbay, A. Aman and N. I. Badler, "Psychological parameters for crowd simulation: From audiences to mobs," *IEEE transactions on visualization and computer graphics,* vol. 22, no. 9, pp. 2145--2159, 2015.

[18] L. Saifi, A. Boubetra and F. Nouioua, "An approach for emotions and behavior modeling in a crowd in the presence of rare events," *Adaptive Behavior,* vol. 24, no. 6, pp. 428--445, 2016.

[19] E. B. Varghese and S. M. Thampi, "A deep learning approach to predict crowd behavior based on emotion," in *International Conference on Smart Multimedia*, 2018.

[20] S. Smys, "Virtual reality gaming technology for mental stimulation and therapy," *Journal of Information Technology,* vol. 1, no. 1, pp. 19--26, 2019.

[21] A. Bashar, "Survey on evolving deep learning neural network architectures," *Journal of Artificial Intelligence,* vol. 1, no. 02, pp. 73--82, 2019.

[22] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251--1258.

[23] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770--778.

[24] "https://github.com/ostrolucky/Bulk-Bing-Image-downloader," [Online]. [Accessed 06 July 2020].

[25] F. Chollet and o. , "Keras," GitHub, 2015.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:10:18 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Data Aggregation Algorithms in IoT: An Organized Evaluation of The Literature

Rajesh Kumar Yadav
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rkyadav@dtu.ac.in

Monica Gupta
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
monicagupta809@gmail.com

***Abstract*— IoT is an innovative idea in technology and communication, and by a good data communication network the data receiving and sending is possible. There are few challenges faced by IoT as there are restricted hardware and communicational functioning ability as well as small scopes. As a result, essential limitations in the productivity of accumulation, broadcast, communication between the nodes, and, very importantly, degrading its routing performance are all caused by these inherent challenges. Seeing these situations, data aggregation (DA) and broadcast methods in these networks are planned more logically to handle the problem of reduced time and unsteady communications between the nodes. DA has a significant part in the IoT; thus, its important mechanism has been analyzed by a systematic and comprehensive study. So, this survey evaluates the DA approached methodically. Three main groups are categorized by the DA mechanism, which are clustering, central, and tree. And, each class contains a detailed comparison of essential techniques which brings a reference for more studies.**

Keywords*— *Internet of Things, Wireless Sensor Network, Data Aggregation, Cluster Head

I. INTRODUCTION

Internet of Things: In the IT industry, IoT became one of the top and progressive technology [1]. In the coming time, most of the study and findings will be in the direction towards IoT only [2]. For accessing essential information from the environment, it includes devices in good quantity that are positioned all over the target zone, and then the actuator acts. These devices generate a majority of data nonstop, which requires the conversion of combined data. Such a smart environment has a vital role in data routing [3], [4]. Energy usage, traffic volume, increase the lifetime, and accuracy can be reduced by a suitable DA technique.

Summary of Wireless Sensor Network: In today's time, WSN is trending because of features like communication, calculation, and detecting abilities [5]. Many devices together make this network. Every device consists of a transceiver, transducer, energy source, and microprocessor. Data or output is transmitted to the main database by a transmitter. Power Unit provides the energy to all stations. Due to this restrictive factor, protocols for WSNs should deliberate low energy usage.

Basically, the sensor field is comprised of the sensor network, in which the nodes are positioned throughout the area. Fig. 1 represents that the data is transmitted to the end-user.

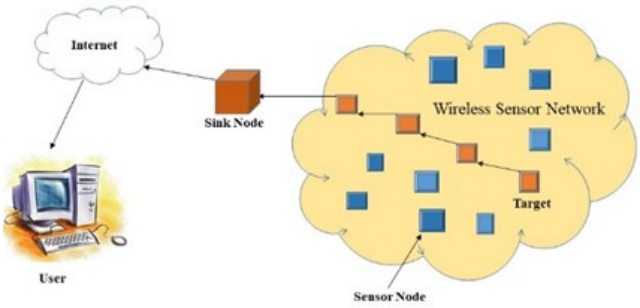


Fig. 1. WSN

II. DATA AGGREGATION

A. Data Aggregation: An Overview

IoT uses sensor nodes that provide data by seeing the real-life environment, along with using multi-hop communication. It then transmits them to the sink node. Distributed sensor nodes gather the information that is calculated. The detected data is transferred to the sink node via relay nodes, which have additional processing ability than the end devices.

Because of redundancy, to lower energy consumption, sink nodes receives the relevant data transmitted by the intermediate nodes ([6], [7]). Transmitting data this way improves resource utilization, and this process is called data aggregation. DA method is the procedure where certain functions like MIN, AVG, MAX, SUM, COUNT, etc. are used on the gathered information ([8], [9]).

DA Function:

- $\text{Min} = f(Q1, Q2\ldots Qn) = \min\{ Q\alpha \mid \alpha = 1,2\ldots n\}$
- $\text{Avg} = f(Q1, Q2\ldots Qn) = \Sigma\, Q\alpha / n.\ (\alpha = 1,2\ldots n)$
- $\text{Max} = f(Q1, Q2\ldots Qn) = \max\{ Q\alpha \mid \alpha = 1,2\ldots n\}$
- $\text{Sum} = f(Q1, Q2\ldots Qn) = \Sigma\, Q\alpha\ .\ (\alpha = 1,2\ldots n)$
- $\text{Count} = f(Q1, Q2\ldots Qn) = |\{ Q\alpha \mid \alpha = 1,2\ldots n\}|$
- $\text{Median} = f(Q1,Q2\ldots Qn) = \Sigma\, Qr\ .\ (r = (n+1)/2)$

In the above DA functions, $Q\alpha$ represents the data from node α, where α varies from 1 to n. Fig. 2. shows two representations- one conventional transmission and another with aggregation model. Devices 1, 2, 3, 5, 6, and 7 at the bottom sense data from the nearby area then transmit to the above level, i.e., nodes 4 and 8, which are called relay as well as aggregator nodes. These nodes act by aggregating the information received from a lower level. In first representation, no aggregation is performed, and all the data

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:15:56 UTC from IEEE Xplore. Restrictions apply.

is transmitted further by node 4, but in other representations, node 8 sends a cumulative packet to the final station.

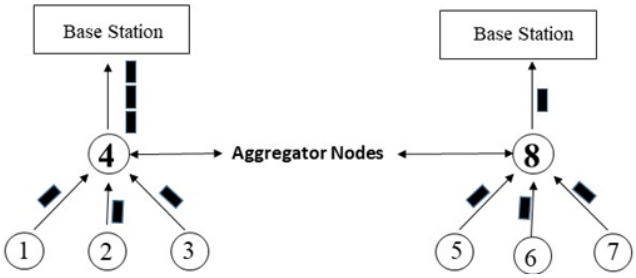


Fig. 2. Data representation with and without aggregation technique.

Fig. 3 shows the structure of the DA algorithm. The method receives the information then actions are taken based on DA protocols. Then the information is transmitted to the final station by the further mechanism of the selected DA protocol.

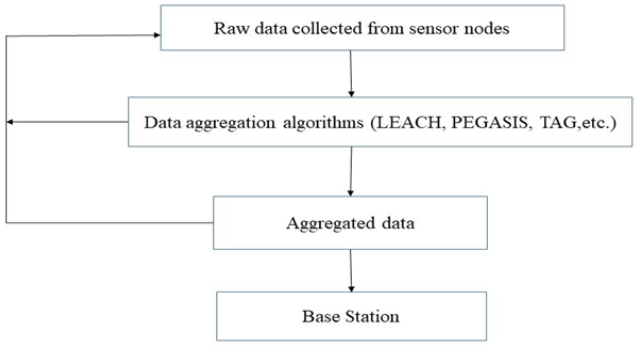


Fig. 3. The architecture of DA.

Below, Table I shows the five main goals of DA, that are considered for protocol designing.

TABLE I. DATA AGGREGATION

Goals	Description
Energy efficiency	• Energy usage must be minimized by reducing the data transmission. • The data aggregation allows the nodes to switch off the transceivers, thus helps in energy saving.
Reducing data propagation latency	• Applications based on real time data analysis, need to be on time. • It is the delay associated with routing, data transmission, and data aggregation in the IoT.
Data accuracy	• Aggregation protocol reduces the amount of transmitted data, which may result in loss of information. • Data accuracy must be increased by number of nodes participating.
Aggregation freshness	• If the data collected during the same time period are aggregated together, then it is definitely ensured by aggregation freshness. • Aggregation freshness is considered as an important goal in monitoring applications due to the requirement for time-stamping data.
Avoiding collisions	• Avoiding collisions can be helped by the data aggregation protocol by minimizing the amount of data to be carried by the network. • Also, it can decrease the number of collisions by a better selection of the data path and the adequate data aggregation

B. DA Techniques

DA methods can be divided according to their topology, QoS, the basis of network and so on. Here, the categorization is done based on network topology. In Table II, two categories are defined that is a flat and hierarchical network.

- *DA in Flat Networks:* Here, every sensor node has the same capabilities and battery power. The data-centric path is followed, where the final station sends a query to all the stationed nodes, and they respond with the sensed data [10]. Due to lots of communication with the sink node, the battery is consumed rapidly, failing the sink node.
- *DA in Hierarchical Networks:* In this, because of the failure of sink node, techniques are designed to improve efficiency and scalability [11]. The amount of data forwarded to the upper level is reduced by some special nodes. [10].

TABLE II. CHARACTERISTICS OF HIERARCHICAL VS FLAT NETWORKS

Hierarchical Networks	Flat Networks
Data aggregation performed by cluster heads or a leader note.	Data aggregation is performed by different nodes along the multi-hop path
Overhead involved in cluster or chain formation throughout the network.	Data aggregation routes are formed only in regions that have data for transmission.
Even if ne cluster head fails, the network may still be operational.	The failure of sink node may result in the breakdown of entire network.
Lower latency is involved since sensor nodes perform short-range transmissions to the cluster head.	Higher latency is involved in data transmission to the sink via a multi-hop path.
Routing structure is simple but not necessarily optimal.	Optimal routing can be guaranteed with additional overhead.
Node heterogeneity can be exploited by assigning high energy nodes as cluster heads.	Does not utilize head heterogeneity for improving energy efficiency.

III. DA IN HIERARCHICAL NETWORKS

A. Network-based on Cluster Topology

In this approach, information is transmitted to the base station via the cluster head of all clusters formed. This node gathers the sensed data from other associate nodes of the group, performs the aggregation, and sends it to the sink, as shown in Fig. 4. LEACH protocol is one of the examples of such an approach.

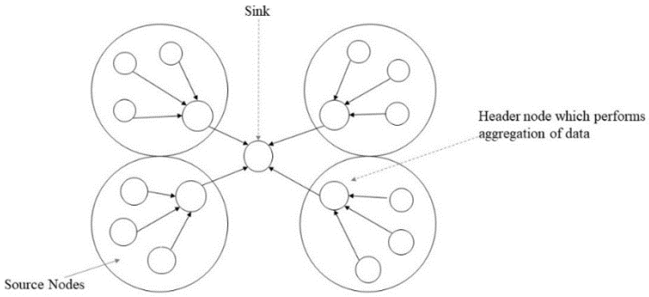


Fig. 4. Cluster-based architecture.

B. Network-based on Centralized Topology

Here, every node transmits the data considering a node at the center and powerful amongst all, known as a header node. To forward the data, nodes select the shortest path, as shown in Fig. 5.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:15:56 UTC from IEEE Xplore. Restrictions apply.

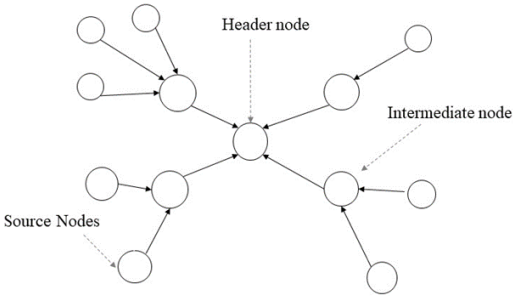


Fig. 5. Centralized architecture.

C. Network-based on Tree Topology

In this approach, the network is based on the tree structure. Information forwarding id did via intermediate nodes, and those nodes aggregate the data at that level and pass it to the next level, as shown in Fig. 6. TREEPSI protocol is one of the examples of such an approach.

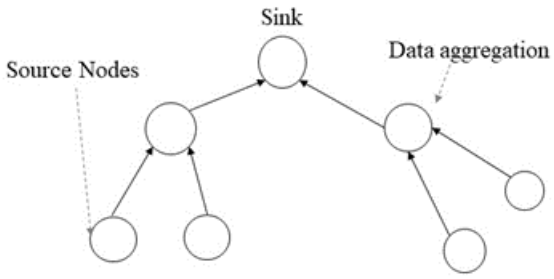


Fig. 6. Tree-based architecture.

Below is the diagrammatic arrangement of primary DA methods in Fig. 7.

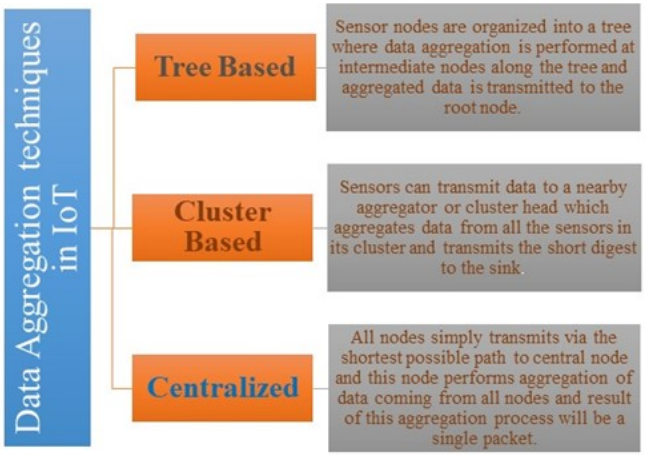


Fig. 7. Classification of DA methods.

Table III shows significant benefits and weaknesses of DA techniques based on their energy consumption, accuracy, security, scalability, and many more factors.

TABLE III. ADVANTAGES AND DISADVANTAGES OF THESE APPROACHES

Category	Advantages	Weakness
Tree-based	Low energy consumption	Low accuracy
	High network lifetime	Low security
	High Scalability	
	Low traffic load	
Cluster-based	Low Traffic load	High Latency
	High Scalability	Low network lifetime
	High fault tolerant	
Centralized	High security	Low fault tolerant

IV. CLUSTER-BASED DA TECHNIQUE

A. LEACH (Low-Energy Adaptive Clustering Hierarchy)

In WSN, LEACH is considered as primary clustering protocol, which is dynamic. CH for each round is selected if the value calculated by that node lies below the threshold,

$$T(n) = p/(1-p*(t \bmod (1/p))), \text{ if } n \in H$$

$$T(n) = 0, \text{ otherwise} \quad [12]$$

Here, t is the current round number, p is the desired percentage of clusters, and H is the set of nodes that have not been selected as CH in the last 1/ p rounds.

To balance the energy load, CH is randomly rotated. The transmission of information from nodes to the base station involves data fusion to reduce the number of packets [12]. It consists of many cycles, where each cycle is divided into 2 parts: arrangement and steady phase. The cluster setup phase means the cluster is formed along with a selection of CH. CH then broadcast to all the devices in the field, declaring they are the CH. Based on the signal strength, nodes decide the cluster for themselves. During the steady phase, the schedule for every node is decided, that when can they transmit. Periodically, information is collected by CH, then they perform the accumulation and forward it towards the station.

B. CAG (Clustered Aggregation)

CAG algorithm forms the cluster based on the values sensed if they are in the provided threshold. Till the time, all the nodes reach under the given threshold; no change is done to the cluster. In this, for every cluster, only one data packet is forwarded, whereas, in TAG (TinyAGgregation), data is transmitted by all the nodes. CAG performs well in comparison to TAG according to the result of [13]. CH values were not considered to calculate aggregate in LEACH, whereas in CAG, similar values are sensed.

C. Analysis of LEACH and CAG

TABLE IV. ANALYSIS OF LEACH AND CAG

Protocol	Advantage	Disadvantage
LEACH	Cluster head is selected randomly	• Cluster formation is expensive • Terminates in a constant number of iterations
CAG	Offers energy efficient aggregation results with small and negligible error	• Saves energy only when few nodes change clusters. • Resilient to the packet loss.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:15:56 UTC from IEEE Xplore. Restrictions apply.

V. CHAIN BASED DA TECHNIQUE

A. PEGASIS (Power-Efficient Gathering in Sensor Information System)

PEGASIS was planned with some enhancement in LEACH and is based on chain structure. In this scheme, communication holds between each of the nodes and its neighbor node, then transmits the information to the base station in their scheduled time, which helps in consuming less power. This approach allocates the load evenly throughout the network. Through the assistance of some greedy algorithms, chains of nodes are formed. According to [14], the structure of a network is known to all the nodes. One can easily construct a chain by using a greedy approach, and this is done before the starting of communication. In the building of the chain, [14] declares that the external sensors start the process so that they can have close neighbors. Every round sensor sends the data to its neighbor, where it merges the information with its personal sensed information and give it to another device. To understand it properly, [14] represented it with the help of an example, as shown in Fig. 8.

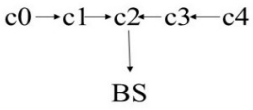


Fig. 8. Passing of token in PEGASIS.

A token is passed from C2, which is a leader, towards both the nodes at the end of the chain. Both C0 and C4 will forward the sensed data to their neighbors that are C1 and C3, respectively. As all the intermediate nodes have extra capabilities of data fusion, C1 and C3 will merge their data with the received data and will pass the packet to the C2 node.

B. COSEN (Chain Oriented Sensor Network for Efficient Data Collection)

COSEN protocol is comprised of 2 phases: chain creation stage and data collection state.

Chain Formation Phase: COSEN protocol requires all the nodes in the area to form many fixed length chains at the lower level. The creation of chain commences from the device that is positioned furthest from BS can be done with the help of a greedy algorithm. Then this node adds the nearest node to its chain if it has not become the part of other chains. Until its length exceeds the fixed length, this chain process works continuously. The formation of a new chain begins only when it exceeds the length.

Data Collection and Transmission Phase: The Data collection process starts after the chains are formed, and the leaders are selected. The process of token passing is similar to PEGASIS. From Fig. 9, a token is passed from node n3 to both the ends n1 and n5, where the forwards the sensed data to the next device, and it merges the information with its data and sends it into another device. In this way, the leader node receives the data from all the nodes, merges it, and forwards the info to the nodes at a higher level. At last, the leader, in the end, directs the aggregated info to the station.

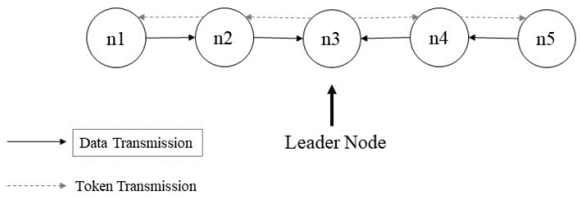


Fig. 9. Shows the token passing approach in COSEN

C. Analysis of PEGASIS and COSEN

TABLE V. ANALYSIS OF PEGASIS AND COSEN

Protocol	Advantage	Disadvantage
PEGASIS	Network energy dissipation is balanced	• Chain leader is chosen by taking turns • Improver data delay time
COSEN	Improve the consumption of energy and transmission delay	Lots of unneeded transmission paths

VI. TREE-BASED TECHNIQUE

A. TREEPSI (Tree-based Efficient Protocol for Sensor Information)

Different approaches are used by researchers for the association between the nodes in TREEPSI. Here, the devices are stationed haphazardly in the required zone. The system is formed based on a tree structure, considering a node as a root node. There are two approaches to passing for building the network. In the first method, the base station informs all the devices about the structure of the path, by itself calculating it. Another way involves where all the devices build a similar network utilizing a conventional procedure. Subsequently, the formation of a tree, one of the following methods can be used for data collection [16]:

- Here root initiates the information collection process by using standard tree traversal techniques. It transmits the control packets to the lower levels.

- The process is initiated by the end nodes or leaf nodes. They convey the data to the upper-level devices. All the intermediate devices have the capability of fusing the gathered info with their personal sensed info and forward it to another upper level. This method keeps on going until the final data is reached to the root note.

Both have different methods, but at last, it's the root node that gathers the information and directs the final info packet to the sink. If any node expires because of battery issues, the tree structure is calculated again. Then, the round of data communication goes on until another node dies.

B. TCDGP (Tree-Clustered Data Gathering Protocol)

Due to certain limitations in the previous protocols, TCDGP is designed by the combo of a tree and cluster-based approach. By the work done in [17], it was concluded that the tree-based approach is suitable according to the other two. When the data make diversions for nodes, then only a few sensor nodes will send information to the sink. Increased power dissipation was observed during the process of data collection. For the bigger target area, this issue occurs in the process of making a binary tree path. Therefore, [17] came up with the model of a new approach which reduces the power dissipation. In Fig. 10, the flow chart represents the working of this algorithm. It is a combo of a tree and cluster-based approach, divided into setup phase and transmission

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:15:56 UTC from IEEE Xplore. Restrictions apply.

phase. Setup phase consists of cluster establishment and formation of a tree structure in each cluster. Then, in the transmission phase, info is gathered by every node and is passed to upper level till it reaches the root node.

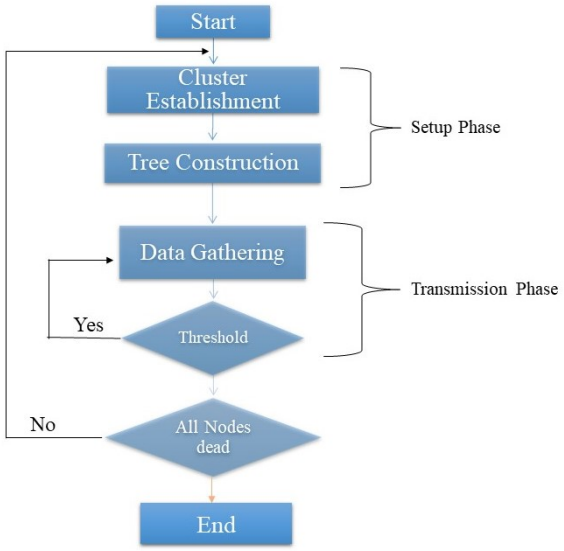


Fig. 10. The flowchart of TCDGP.

C. *Analysis of TREEPSI and TCDGP*

TABLE VI. ANALYSIS OF TREEPSI AND TCDGP

Protocol	Advantage	Disadvantage
TREEPSI	Consumes less power during data transmission	Path has made an alternative route in the topology
TCDGP	Reduces consumption of energy	Require methods for recovery procedures

VII. CONCLUTION

Sensor nodes in good quantity handover the info, over the wireless network, are included under IoT. The lifetime of the network is restricted because the devices have many constraints. Thus, many methods to improve the lifetime of the systems have been proposed by various researchers. Here, few vital DA algorithms are described that enhances the network lifetime. So, a theoretical survey is directed in this research to get a clear idea of all these. For each classification, two basic protocols were discussed that were LEACH, CAG, PEGASIS, COSEN, TREEPSI and TCDGP. Further, based on network architecture, a complete study of different protocols is presented.it shows pros, cons in Table VI, V and VI respectively along with some relevant features of the DA techniques. Based on the discussed techniques and Table III, it is concluded that the network based on tree topology has good features compared to the rest.

REFERENCES

[1] H. Rahman, N. Ahmed, and I. Hussain, Comparison of DA Techniques in Internet of Things (IoT)

[2] M.H, Homaei, Ely Salwana, Shahab Shamshirband, An Enhanced distributed DA method in the internet of things

[3] S. Cirani, G. Ferrari, M. Picone, and L. Veltri, Internet of Things: Architectures, Protocols and Standards. John Wiley & Sons, 2018.

[4] F. Montori, L. Bedogni, and L. Bononi, "A collaborative internet of things architecture for smart cities and environmental monitoring," IEEE Internet Things J., vol. 5, no. 2, pp. 592–605, 2018.

[5] Yu-CheeTseng et al. (2004) "Location Tracking in a Wireless Sensor Network by Mobile Agents and Its Data Fusion Strategies" The Computer Journal, Vo. 47, No. 4.

[6] M. Azharuddin, P. Kuila, and P. K. Jana, "Energy efficient fault tolerant clustering and routing algorithms for wireless sensor networks," Computers and Electrical Engineering, vol. 41, pp. 177 – 190, 2015.

[7] N. Pantazis, S. Nikolidakis, and D. Vergados, "Energy-Efficient Routing Protocols in Wireless Sensor Networks: A Survey," Communications Surveys Tutorials, IEEE, vol. 15, pp. 551–591, Feb 2013.

[8] V. Pandey, "A Review on DA Techniques in Wireless Sensor Network," Journal of Electronic and Electrical Engineering, vol. 1, no. 2, 2010.

[9] S. Sirsikar and S. Anavatti, "Issues of DA Methods in Wireless Sensor Network: A Survey ," Procedia Computer Science, vol. 49, pp. 194 – 201, 2015. Proceedings of 4th International Conference on Advances in Computing, Communication and Control (ICAC).

[10] VaibhavPandey, AmarjeetKaur and Narottam Chand (2010) A Review on DA Techniques in Wireless Sensor Network. Journal of Electronic and Electrical Engineering, Vol.1, Issue 2.

[11] Thangaraj, M. and PunithaPonmalar, P. (2011) A Survey on DA Techniques in Wireless Sensor Networks. International Journal of Research and Reviews in Wireless Sensor Networks, Vol. 1, No. 3.

[12] Heinzelman, W., Chandrakasan, A. and Balakrishnan, H. (2000) An Energy-Efficient Communication Protocol for Wireless Microsensor 85 Networks. Proceedings of the 33rd Hawaii International Conference on System Sciences. IEEE

[13] SunHee Yoon and Cyrus Shahabi (2007) The Clustered AGgregation (CAG) Technique Leveraging Spatial and Temporal Correlations in Wireless Sensor Networks. ACM Transactions on Sensor Networks (TOSN), Vol. 3, Issue 1, No. 3.

[14] Stephanie Lindsey and Cauligi S. Raghavendra (2002) PEGASIS Powerefficient Gathering in Sensor Information Systems. In Proceeding of IEEE Aerospace Conference. Montana IEEE Aerospace and Electronic Systems Society.

[15] SunHee Yoon and Cyrus Shahabi (2007) The Clustered AGgregation (CAG) Technique Leveraging Spatial and Temporal Correlations in Wireless Sensor Networks. ACM Transactions on Sensor Networks (TOSN), Vol. 3, Issue 1, No. 3.

[16] Satapathy S. S. and Sarma, N. (2006) TREEPSI: tree based energy efficient protocol for sensor information. IFP International Conference.

[17] Gurpree Sigh Chhabra and Dipesh Sharma (2011) Cluster-Tree based Data Gathering in Wireless Sensor Network. International Journal of Soft Computing and Engineering, Vol. 1, Issue 1.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:15:56 UTC from IEEE Xplore. Restrictions apply.

Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease

Surendrabikram Thapa[1], Priyanka Singh[2], Deepak Kumar Jain[3], Neha Bharill[4], Akshansh Gupta[5], Mukesh Prasad[2]

[1]Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India
[2]School of Computer Science, FEIT, University of Technology Sydney, Sydney, Australia
[3]Department of Computer Science, Chongqing University of Posts and Telecommunications, China
[4]Department of Computer Science and Engineering, Mahindra Ecole Centrale, Hyderabad, India
[5]Central Electronics Engineering Research Institute, Pilani, Rajasthan, India

Abstract: Alzheimer's disease (AD) is a neurodegenerative disorder resulting in memory loss and cognitive decline caused due to the death of brain cells. It is the most common form of dementia and accounts for 60-80% of all dementia cases. There is no single test for diagnosis of AD, the doctors rely on medical history, neuropsychological assessments, computed tomography (CT) or magnetic resonance imaging (MRI) scan of the brain, etc. to confirm a diagnosis. In terms of the treatment, currently, there is neither a cure nor any way to slow the progression of AD. However, for people with mild or moderate stages of this disease, there are some medications available to temporarily reduce symptoms and help to improve quality of life. Hence, early diagnosis of AD is extremely crucial for overall better management of the disease. The researches have shown some relation between neuropsychological scores and atrophies of the brain. This can be leveraged for the early diagnosis of AD. This paper makes use of feature selection techniques to extract the most important features in the diagnosis of AD. This paper demonstrates the need to combine neuropsychological scores like mini-mental state examination (MMSE) with MRI features to provide better decisional space for early diagnosis of AD. Through the experiments, including MMSE along with other features are found to improve the classification of AD, significantly.

Keywords: Alzheimer's disease, neuropsychological scores, feature selection, mini-mental state examination

I. INTRODUCTION

Alzheimer's disease (AD) was first identified and discussed by a German physicist and neuro-pathologist Alois Alzheimer [1]. According to the World Alzheimer Report [2], there were around 50 million people in the world living with dementia in 2018 and about two-third of them have AD. The number of patients suffering from this disease is likely to rise to around 152 million in 2050 and the cost of this disease is forecasted to be 2 trillion US dollars in 2030 [2]. Currently, the treatment of AD predominantly involves the use of this disease-modifying or delaying drugs rather than drugs that reverse or permanently stop its progression. Thus, there is a necessity for the prediction of AD in the early stage so that it becomes possible to better manage this disease by delaying the occurrence of cognitive decline.

AD is characterized by loss of neurons and synapses in the cerebral cortex and certain subcortical regions which results in gross atrophy of the affected regions, including degeneration in the temporal lobe and parietal lobe, and parts of the frontal cortex and cingulate gyrus [3]. The previous studies have shown that various impairments in AD are correlated with atrophy in various regions like the hippocampus [4], temporal lobe [5], and amygdala [6]. Moreover, the determination of the key atrophied regions across the entire brain could be used as parameters for the delineation of AD patients from cognitively normal (CN) subjects [7].

Earlier, the clinical measures like clinical dementia rating (CDR) were used in the diagnosis of AD [8]. With the rapid advancement in computation and development of imaging techniques, machine learning techniques in early diagnosis of AD have been widely used using neuroimaging images like MRI and positron emission tomography (PET) [9-11]. Among numerous machine learning techniques that have been explored, the widely used technique for AD prediction is Support Vector Machines (SVM) [12]. The various studies have been carried out with single modality or multiple modalities. Using a single modality makes the job easier with less cost and time but is accompanied by more erroneous predictions. The experiments that make use of imaging modalities without neuropsychological scores attain lesser accuracy than the experiments which used imaging modalities with neuropsychological scores. Hinrichs et al. [11] were able to improve the accuracy of classification between AD and Control Normal (CN) by around 5% after incorporating the Mini-Mental State Examination (MMSE) in their experiment. Ewers et al. [13] also combined the primary MRI and Cerebrospinal fluid (CSF) biomarkers with a neuropsychological test for making the classification of AD and CN better. They also used similar combinations to predict the conversion from Mild Cognitive Impairment (MCI) to AD. Similarly, Zhou et al. [7] had an accuracy score of 83.1% using subcortical volume alone whereas the accuracy of classification between AD and CN using subcortical volume with MMSE was 92.3% with a

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

significant increment of more than 9% in accuracy. Some of the neuropsychological tests can be a good determinant of atrophy in various regions of the brain. For example, naming impairment in AD is associated with left anterior temporal lobe atrophy [5]. Due to this reason, the use of neuropsychological test scores can increase accuracy in the classification.

This study discusses the classification of CN, MCI, and AD individuals by combining the primary MRI with neuropsychological test scores. This study uses the MMSE as a neuropsychological test, which concentrates on cognitive aspects of mental function [14]. It is a 30-point questionnaire that is extensively used to examine the individual's orientation to time and place, repeating abilities, attention and calculation, recall, language, and response to simple commands. The proposed study uses MRI images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The structural measures are extracted from raw T1 structural MRI images using Free Surfer. The machine learning algorithms are trained on different CN, MCI, and AD subjects at the screening phase with some feature selection techniques to yield the best possible results. The results obtained in this experiment are expected to portray the role of MMSE along with the volume of important features in the early detection of AD.

II. METHODOLOGY

Figure 1 shows the flow diagram of the proposed approach for the classification of various stages of AD. The process starts with dataset collection and goes through data preprocessing, feature selection, use of learning algorithms, and finally the result obtained from the proposed approach. The framework for the proposed approach is simple and can be replicated with any machine learning algorithms. Each component is described below in detail.

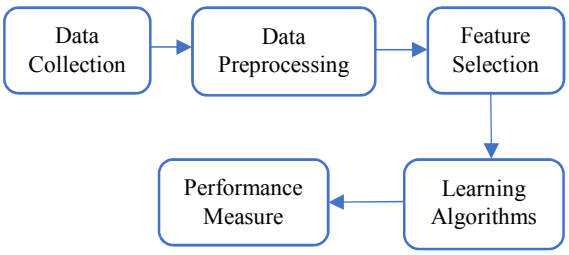


Fig. 1: Flow diagram of the proposed approach

A. Data Collection

The proposed approach uses the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [15]. The ADNI dataset was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI dataset was to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

The ADNI dataset is queried for basic, clinical, and MRI data from CN, MCI, and AD subjects screening. The query results are processed to obtain useful information and to drop the irrelevant information for this paper and only three subsets are considered. The first, second, and third subset contains 119 CN, 155 MCI, and 46 AD subjects, respectively. This paper uses T1 weighted MRI images of 1.5 T. The basic demographic data of the ADNI subjects included in this study are mentioned in Table 1. Data are given as mean ± standard deviation (σ) except for the gender for which frequencies are mentioned and represented by *M* and *F* for males and females, respectively.

B. Data Preprocessing

The attributes of the ADNI dataset are in different ranges such as the range for attribute age and MMSE. The classifiers like SVM work comparatively better with data that is scaled rather than the unscaled data. The main advantage of scaling is to avoid the features in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the calculation. This is because kernel values depend usually on the inner products of the feature vectors. For the linear kernel and the polynomial kernel, attribute values with a greater numeric range might cause numerical problems. So, for better computation and accuracy, the data are scaled using the *z*-score as shown in Eq. (1).

$$z = \frac{x - \bar{x}}{\sigma} \tag{1}$$

where x is the original feature vector, $\bar{x}$ = is the mean of that feature vector, and σ is the standard deviation.

The number of observations for CN, MCI, and AD is 119, 155, and 46, respectively. Imbalanced data often have poor performance in terms of sensitivity and specificity. So, to nullify this effect, some over-sampling techniques are required. The Synthetic Minority Over-sampling Technique (SMOTE) is used to solve the class imbalance problem by synthesizing new instances from existing minority cases in this paper [16].

C. Feature Selection

The irrelevant features can simply overfit the model, hence there is a need for a feature selection technique to handle the overfitting issue. The feature selection techniques improve the reliability of the model out of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

sample if the irrelevant features are dropped out. This paper uses a filter-based feature selection method, which is independent of any classifier [17]. The feature selection method consists of a forward selection (Weka's *Bestfirst*) to look for a combination of attributes with the high individual predictive value of the diagnostic class and low inter-correlation (Wekas's *CfsSubsetEval*) [17]. It evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. The significant features obtained for different classification problems are mentioned in Table 2.

Table 1: Demographic data of ADNI subjects included in this paper

Attributes	CN (N= 119)	MCI (N= 155)	AD (N= 46)
Gender	60 F/ 59 M	49 F/ 106 M	24 F/ 22 M
Age	76.12 ± 5.06	75.33 ± 7.45	76.08 ± 5.92
Years of Education	15.79 ± 3.01	15.74 ± 2.88	24.19 ± 1.49
MMSE	29.11 ± 0.97	27.3 ± 1.75	24.19 ± 1.49

Table 2: Significant features for CN vs AD, MCI vs AD and CN vs MCI

Classification	Rank	Variables
CN *vs* AD	1	MMSE
	2	Hippocampus (Left) Volume
	3	Hippocampus (Right) Volume
MCI *vs* AD	1	MMSE
	2	Hippocampus (Right) Volume
	3	Hippocampus (Left) Volume
CN *vs* MCI	1	MMSE
	2	Hippocampus (Left) Volume
	3	Hippocampus (Right) Volume

D. Machine Learning Algorithms

The best learning algorithm for the given classification problem is very difficult to select by instinct. So, different machine learning algorithms are compared with the proposed approach. Some of the machine learning algorithms used in this paper can be studied under two broad headings viz. Classification and Ensemble Learning algorithms. The ensemble learners are chosen because sometimes ensemble learners outperform the classification algorithms by combining the predictions of multiple base estimators in order to improve the robustness of a given single classifier. Five classification algorithms are used in this paper. Wherever applicable, ten-fold cross-validation is performed to optimize the hyper-parameters of models using a grid-search technique. Ensemble learning algorithms are expected to give better performances than the base learning algorithms. Multiple hypotheses are generated by running the same base learner and are combined to form a better hypothesis. Among some of the popular ensemble learning algorithms, AdaBoost, XGBoost, and Random Forest are used.

Logistic Regression: Logistic regression is a statistical method to model a binary classification problem [18], which fits with the binary classification problems in this paper such as CN vs AD, CN vs MCI, and MCI vs AD. There are certain parameters to be tuned in logistic regression like *C* (regularization) and *penalty*. A grid search method is performed to tune these parameters. The values of parameters *C* and *penalty* for CN vs AD without MMSE are 0.1 and '*L2*', and with MMSE are 1 and '*L2*'. Similarly, the values of parameters *C* and *penalty* for CN vs MCI without MMSE are 0.1 and '*L1*', and with MMSE are 0.01 and '*L2*'. The values of parameters *C* and *penalty* for MCI vs AD without MMSE are 0.1 and '*L2*', and with MMSE are 0.1 and '*L1*'.

Decision Tree: Decision tree is one of the simplest and yet most useful machine learning algorithms. It can be used to solve both classification and regression problems [19]. There are some parameters like *maximum depth* which is the longest path from the tree root to a leaf and *minimum samples* that are required to be at a leaf node. A Grid search method is used to tune

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

these parameters. The values of parameters *maximum depth* and *minimum samples* for CN vs AD without MMSE are 7 and 1, and with MMSE are 1 and 1. Similarly, the values of parameters *maximum depth* and *minimum samples* for CN vs MCI without MMSE are 1 and 100, and with MMSE are 5 and 5. The values of parameters *maximum depth* and *minimum samples* for MCI vs AD without MMSE are 7 and 1, and with MMSE are and 12 and 1.

Naïve Bayes: Naïve Bayes classifier is a probabilistic machine learning model that is used to solve classification problems. A naïve Bayes classifier assumes that all attributes are independent given the class variable [20].

Factorization Machines: The Factorization machine is a supervised learning algorithm that can be used in solving a wide variety of prediction tasks. It was introduced by Steffen Rendle in 2010 and can be used to solve various prediction tasks like regression, binary classification, and ranking [21]. The values for tuned parameters of rank, order, and learning rates are 10, 3, and 0.001, respectively.

Support Vector Machines: Support Vector Machines are based on the concept of decision planes that define decision boundaries. Support Vector Machine constructs a hyperplane in high-dimensional space which can be used to classify data into two groups [12]. It can also be used for tasks like regression and outlier detection. The performance of SVM is affected by parameters *C* (regularization) and γ (gamma). A Grid search method is performed to tune these parameters. The values of parameters *C* and γ for CN vs AD without MMSE are 32 and 2, and with MMSE are 0.125 and 0.5. Similarly, the values of parameters *C* and γ for CN vs MCI without MMSE are 2048 and 0.0001, and with MMSE are 8 and 2. The values of parameters *C* and γ for MCI vs AD without MMSE are 8 and 8, and with MMSE are 8 and 8.

AdaBoost: AdaBoost, a short form for Adaptive Boosting is a meta-estimator that was developed by Freund and Schapire which has its applications in numerous fields. It can be used in conjunction with many other learning algorithms to improve the accuracy of especially weak learners. This is because AdaBoost combines weak hypotheses by summing their probabilistic predictions [22]. The values of parameters *learning rate*, *loss function* and *maximum number of estimators* for AD vs CN without MMSE are 1, linear and 50, and with MMSE are 0.01, linear, and 100. Similarly, the values of parameters *learning rate, loss function,* and *maximum number of estimators* for CN vs MCI without MMSE are 0.1, exponential and 100, and with MMSE are 0.01, square and 50. The values of parameters *learning rate*, *loss function* and *maximum number of estimators* for MCI vs AD without MMSE are 0.3, linear and 100, and with MMSE are 0.1, square and 100.

XGBoost: Extreme Gradient Boosting or popularly known as XGBoost is a tree-based ensemble learning technique that uses gradient boosting [23]. XGBoost has recently been widely used and has shown great performance in various machine learning competitions. Hyperparameters like *learning rate*, *maximum depth*, *minimum samples* that are required to be at a leaf node, and *number of estimators* are tuned using grid search. The values of parameters *learning rate*, *maximum depth*, *minimum samples,* and *number of estimators* for CN vs AD without MMSE are 0.01, 2, 3 and 1000, and with MMSE are 0.01, 2, 3, 1000. Similarly, the values of parameters *learning rate*, *maximum depth*, *minimum samples,* and *number of estimators* for CN vs MCI without MMSE are 0.01, 2, 3 and 1000, and with MMSE are 0.01, 2, 3, 1000. The value of parameters *learning rate*, *maximum depth*, *minimum samples,* and *number of estimators* for MCI vs AD without MMSE are 0.01, 2, 3 and 1000, and with MMSE are 0.01, 2, 3, 1000.

Random Forest: Random forest is a meta-estimator used for solving both regression and classification problems. The random forest as the name suggests is a forest of many decision trees that outputs the class that is calculated from aggregating the classification outputs of each decision tree of that forest [24]. There are certain parameters that affect the classification performance of the Random Forest classifier. Grid search is performed to find optimal values of the function to measure the quality of a split, and the number of estimators. The values of parameters *criterion* and *number of estimators* for CN vs AD, CN vs MCI and MCI vs AD are *gini* and 200, respectively for both without and with MMSE.

E. Performance Measures

The performance of the proposed approach is measured by different evaluation criteria such as accuracy, precision, recall, specificity, sensitivity, F1-score, the area under the curve (AUC) as shown in Eqs. (2)-(7).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

$$F1-Score = 2.\frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

$$Specificity = \frac{TN}{TN+FP} \quad (6)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (7)$$

where *TP*, *TN*, *FP,* and *FN* indicate True Positive, True Negative, False Positive and False Negative, respectively.

III. EXPERIMENTAL RESULTS

The feature selection techniques give significant attributes as shown in Table 2. This shows that atrophy in the hippocampal region is more significant in the distinction of different stages of AD when used with neuropsychological scores (MMSE). The classification accuracy with and without MMSE for different classifiers are shown in figs. 2-4. It is observed that SVM with RBF kernel outperforms the other classification techniques.

The classification performance in terms of accuracy after feature selection for SVM with RBF kernel is shown in Table 3. From Table 3, it can be observed that with MMSE, the performance is seen better in all three classification types viz. CN vs AD, CN vs MCI, and MCI vs AD. Table 4 compares AUC values with MMSE and without MMSE for various classifiers under the studied experiments. The best performance is shown by SVM with RBF kernel, Extreme Gradient Boosting, and Random Forest. While taking other measures of performance, it is observed that SVM with RBF kernel performs better than other classifiers.

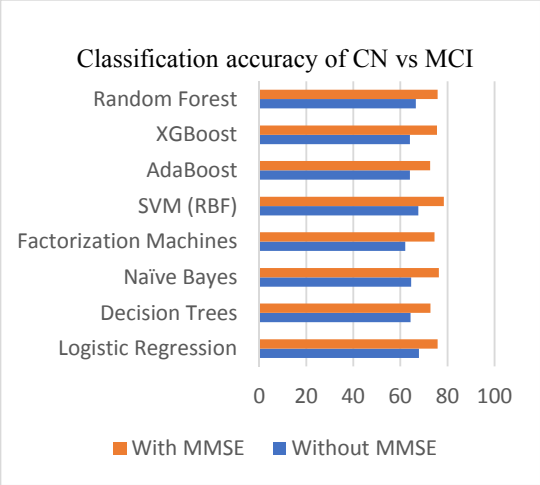


Fig. 2: Classification accuracy for CN vs MCI with and without MMSE

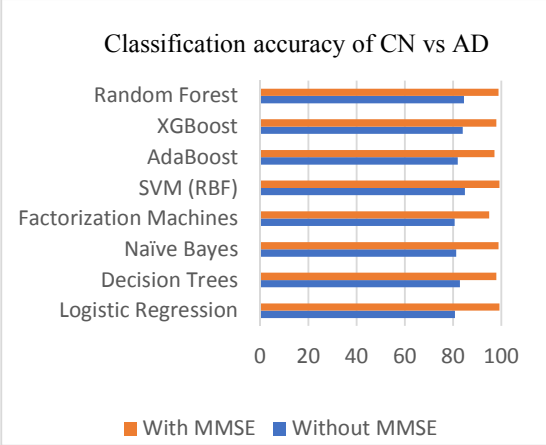


Fig. 3: Classification accuracy for CN vs AD with and without MMSE

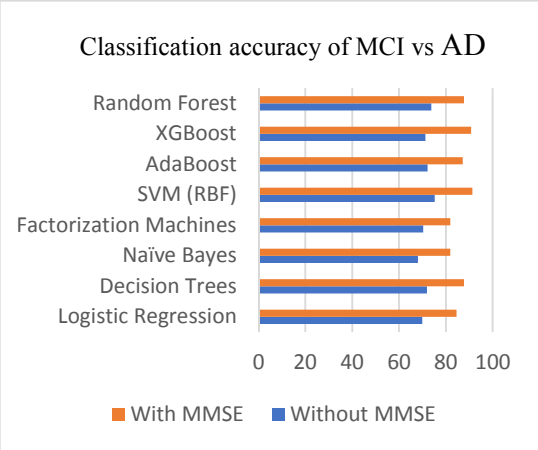


Fig. 4: Classification accuracy for MCI vs AD with and without MMSE

Table 5 shows that SVM with RBF performs better with MMSE in terms of all performance measures. It can be also observed that the performance of CN vs AD is better than MCI vs AD and CN vs MCI. Table 6 compares the performance of the proposed approach with other approaches for different stages of Alzheimer's disease. It can be observed that with the addition of neuropsychological tests to imaging modalities, the accuracy increases significantly. The performance of the proposed approach is better in terms of accuracy, sensitivity, and specificity with various MRI modalities and multimodal techniques as well. The addition of neuropsychological tests with imaging modalities provides better decisional space for the classification of different stages of Alzheimer's disease.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

IV. CONCLUSION

The research for finding new biomarkers for Alzheimer's disease is undoubtedly of great importance but a combination of neuropsychological test scores like MMSE with MRI measures can also give more accurate results. This paper proposes a feature selection based model which enables to learn good patterns for the given data and enhances the performance of the model. This paper substantiates that the role of neuropsychological test scores can be vital in improving the performance of the classification models as it carries some important information about the brain. Also, the multimodal methods in early diagnosis of AD are proved to be more accurate such as image modality with neuropsychological scores. In the future, the proposed study can be used to find how atrophy in different regions of interest is correlated with test scores. Also, different neuropsychological tests other than MMSE can be used for proving more robust and accurate results.

Table 3: Performance in terms of accuracy by SVM based on feature selection

Classification Type	With MMSE	Without MMSE
CN vs AD	99.2	84.9
CN vs MCI	78.5	67.6
MCI vs AD	91.3	75.2

Table 4: Comparison of AUC values of different classifiers with and without MMSE

Model	CN vs AD		CN vs MCI		MCI vs AD	
	Without MMSE	With MMSE	Without MMSE	With MMSE	Without MMSE	With MMSE
Logistic Regression	0.807	0.992	0.677	0.752	0.70	0.845
Decision Trees	0.823	0.979	0.642	0.723	0.719	0.877
Naïve Bayes	0.811	0.987	0.645	0.67	0.681	0.819
Factorization Machines	0.807	0.949	0.621	0.745	0.704	0.821
Support Vector Machines	0.849	0.992	0.677	0.784	0.752	0.913
AdaBoost	0.819	0.971	0.639	0.726	0.723	0.819
XGBoost	0.840	0.979	0.639	0.755	0.713	0.906
Random Forest	0.845	0.987	0.665	0.758	0.739	0.877

Table 5: Performance of svm based on feature selection with and without MMSE as shown in table 2

Performance Measure	CN vs AD		CN vs MCI		MCI vs AD	
	Without MMSE	With MMSE	Without MMSE	With MMSE	Without MMSE	With MMSE
Accuracy	84.89	99.2	67.56	78.5	75.15	91.3
Precision	0.85	0.99	0.68	0.79	0.75	0.91
Recall	0.85	0.99	0.67	0.78	0.75	0.91
F1-Score	0.85	0.99	0.67	0.78	0.75	0.91
AUC	0.849	0.992	0.674	0.784	0.752	0.913
Sensitivity	0.84	1.00	0.59	0.72	0.81	0.92
Specificity	0.86	0.98	0.76	0.85	0.70	0.90

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

Table 6: Performance comparison of the proposed approach with other methods

Method	Target	Modality	Data	Performance		
				Acc (%)	Sens.	Spec.
Zhou et al., 2014 [25]	CN vs AD	MRI	(127 AD, 59 CN) Private	78.2	0.685	0.755
	aMCI vs CN			66.7	0.55	0.725
	naMCI vs CN			62.1	0.515	0.703
	CN vs AD	MRI + MMSE	(127 AD, 59 CN) Private	92.4	0.84	0.961
	aMCI vs CN			74.9	0.611	0.834
	naMCI vs CN			74.1	0.552	0.823
Hinrichs et al., 2011 [11]	CN vs AD	MRI + PET	(129 AD, 60 CN) ADNI	87.6	0.789	0.938
	CN vs AD	MRI+PET+ CSF+ APOE+ Cognitive Scores	(129 AD, 60 CN) ADNI	92.4	0.867	0.966
Zhou et al., 2014 [7]	CN vs AD	MRI	(129 AD, 60 CN) Private	83.1	0.779	0.856
	CN vs AD	MRI + MMSE	(129 AD, 60 CN) Private	92.3	0.882	0.942
Kim et al., 2017 [26]	CN vs AD	MRI	(160 AD, 208 CN) ADNI	92.84	0.885	0.961
	MCI vs CN			78.28	0.863	0.651
Cuingnet et al., 2011[#] [10]	CN vs AD	MRI	(162 CN, 76 MCIc, 137 AD) ADNI	-	0.81	0.95
	CN vs MCIc			-	0.73	0.85
Proposed approach	CN vs AD	MRI	(46 AD, 155 MCI, 119 CN) ADNI	84.89	0.84	0.86
	CN vs MCI			67.56	0.59	0.76
	MCI vs AD			75.15	0.81	0.70
	CN vs AD	MRI + MMSE	(46 AD, 155 MCI, 119 CN) ADNI	99.2	1.00	0.98
	CN vs MCI			78.5	0.72	0.85
	MCI vs AD			91.3	0.92	0.90

[#] Best performance out of ten different methods

REFERENCES

[1] R. A. Stelzmann, H. Norman Schnitzlein, and F. Reed Murtagh, "An English translation of Alzheimer's 1907 paper, "Über eine eigenartige Erkankung der Hirnrinde"," Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists, vol. 8, no. 6, pp. 429-431, 1995.

[2] C. Patterson, "The state of the art of dementia research: New frontiers," World Alzheimer Report 2018, 2018.

[3] L. J. Herrera, I. Rojas, H. Pomares, A. Guillén, O. Valenzuela, and O. Baños, "Classification of MRI images for Alzheimer's disease detection," in 2013 International Conference on Social Computing, 2013, pp. 846-851: IEEE.

[4] D. H. Adler et al., "Characterizing the human hippocampus in aging and Alzheimer's disease using a computational atlas derived from ex vivo MRI and histology," Proceedings of the National Academy of Sciences, vol. 115, no. 16, pp. 4252-4257, 2018.

[5] K. Domoto-Reilly, D. Sapolsky, M. Brickhouse, B. C. Dickerson, and A. s. D. N. Initiative, "Naming impairment in Alzheimer's disease is associated with left anterior temporal lobe atrophy," Neuroimage, vol. 63, no. 1, pp. 348-355, 2012.

[6] S. P. Poulin, R. Dautoff, J. C. Morris, L. F. Barrett, B. C. Dickerson, and A. s. D. N. Initiative, "Amygdala

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

atrophy is prominent in early Alzheimer's disease and relates to symptom severity," Psychiatry Research: Neuroimaging, vol. 194, no. 1, pp. 7-13, 2011.
[7] Q. Zhou, M. Goryawala, M. Cabrerizo, W. Barker, R. Duara, and M. Adjouadi, "Significance of normalization on anatomical MRI measures in predicting Alzheimer's disease," The Scientific World Journal, vol. 2014, pp. 1-12, 2014.
[8] M. M. Williams, M. Storandt, C. M. Roe, and J. C. Morris, "Progression of Alzheimer's disease as measured by Clinical Dementia Rating Sum of Boxes scores," Alzheimer's & Dementia, vol. 9, no. 1, pp. S39-S44, 2013.
[9] J. Zhang, B. Yan, X. Huang, P. Yang, and C. Huang, "The diagnosis of Alzheimer's disease based on voxel-based morphometry and support vector machine," in 2008 Fourth International Conference on Natural Computation, 2008, vol. 2, pp. 197-201: IEEE.
[10] R. Cuingnet et al., "Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database," neuroimage, vol. 56, no. 2, pp. 766-781, 2011.
[11] C. Hinrichs, V. Singh, G. Xu, S. C. Johnson, and A. D. N. Initiative, "Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population," Neuroimage, vol. 55, no. 2, pp. 574-589, 2011.
[12] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, no. 3, pp. 273-297, 1995.
[13] M. Ewers et al., "Prediction of conversion from mild cognitive impairment to Alzheimer's disease dementia based upon biomarkers and neuropsychological test performance," Neurobiology of aging, vol. 33, no. 7, pp. 1203-1214. e2, 2012.
[14] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician," Journal of psychiatric research, vol. 12, no. 3, pp. 189-198, 1975.
[15] (2019, Jul 15). Data Sharing and Publication Policy. Available: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_DSP_Policy.pdf
[16] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of artificial intelligence research, vol. 16, pp. 321-357, 2002.
[17] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.
[18] Z. Y. Algamal and M. H. Lee, "Penalized logistic regression with the adaptive LASSO for gene selection in high-dimensional cancer classification," Expert Systems with Applications, vol. 42, no. 23, pp. 9326-9332, 2015.
[19] J. R. Quinlan, "Induction of decision trees," Machine learning, vol. 1, no. 1, pp. 81-106, 1986.
[20] I. Rish, "An empirical study of the naive Bayes classifier," in IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001, vol. 3, no. 22, pp. 41-46.
[21] S. Rendle, "Factorization machines," in 2010 IEEE International Conference on Data Mining, 2010, pp. 995-1000: IEEE.
[22] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation," IEEE transactions on medical imaging, vol. 29, no. 1, pp. 30-43, 2009.
[23] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 785-794: ACM.
[24] M. Pal, "Random forest classifier for remote sensing classification," International Journal of Remote Sensing, vol. 26, no. 1, pp. 217-222, 2005.
[25] Q. Zhou et al., "An optimal decisional space for the classification of Alzheimer's disease and mild cognitive impairment," IEEE Transactions on Biomedical Engineering, vol. 61, no. 8, pp. 2245-2253, 2014.
[26] J. Kim and B. Lee, "Automated discrimination of dementia spectrum disorders using extreme learning machine and structural T1 MRI features," in 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2017, pp. 1990-1993: IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 05,2020 at 05:28:08 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Deep NeuralNet For Violence Detection Using Motion Features From Dynamic Images

Aayush Jain, Dinesh Kumar Vishwakarma
Department of Information Technology
Delhi Technological University
Delhi-110042, India
aayush89j@gmail.com, dvishwakarma@gmail.com

***Abstract*—Violent Action Recognition has been a challenging topic in monitoring human activities, especially in public premises. Deep neural nets and Transfer learning have proven highly successful in the detection of violent activities. In this paper, a novel deep NeuralNet system is proposed for the task of Detecting Violence by extraction of motion features from RGB Dynamic Images (DI). Motion feature extraction and prediction of violent content using a stream of RGB DI is done effectively by fine-tuning the pre-trained Inception-Resnet-V2 model. Improved and highly effective methods for intelligent analysis are highly demanded. For performance validation of the proposed novel model, tests are performed on three popular and publically available benchmarks – Hockey Fight dataset, Real Life Violence Dataset, and movie dataset.**

***Keywords*—Violence Detection; Deep NeuralNet; Transfer Learning; Dynamic Images; Inception-Resnet-V2; CNN;**

I. INTRODUCTION

With the spike in violent activities like the recent Delhi riots [1] or the recent mass protests in the US resulting in violent events [2], it is becoming more and more necessary to develop efficient and highly accurate automated intelligent systems which can analyze the public premise for detecting and avoiding suspicious events and violent activities for ensuring better public safety and enabling agencies in deploying better security measures. Recent state-of-the-art has been quite successful and effective in creating surveillance systems and detection methods which can enable enforcement agencies to avoid such incidents [3]–[6]. CNN's have gained remarkable success in recent years in the field of image and video classification with incomparable accuracy and modeling capacity [7]–[9]. Prior studies in the area of violence detection [10]–[16] have ascertained that temporal, spatial, and motion information act as an important trait to classify and detect violent activities in surveillance footage and real-life situations in an effective manner. Researchers have also done extensive work in detecting violent activities using a single-stream architecture extracting spatial information or temporal information or motion information and also by multi-stream architectures to extract spatial-temporal or spatial-temporal-motion data that has improved CNN based systems accuracy by examining RBG and depth based appearance and motion details.

Yet most of the approaches like [17] do not necessarily model the dynamics of the video. RGB images, extracted as a stack of frames from the video, are used to learn discriminative attributes that allow the model to obtain the highest accuracy in tasks of recognition and classification since the primary aim is to represent the action class rather than the motion. Whereas motion features contain a decent amount of evidence for distinguishing the movement and human action. In this regard, optical flow [14], [18]–[20], and dense trajectories [21] are used extensively for illustrating the movement of the entity in videos. Various methods proposed descriptors, based on optical flow orientation and magnitude changes for example ViF [22], [23], OViF [24], HOF and HOG [25] , HOMO [10], etc. However, these approaches regardless of being beneficial also have some disadvantages like high computational time limiting the possibility of real-time implementation of such systems. Similarly, dense trajectories can be extremely sensitive to camera viewpoint.
Considering all such approaches and their shortcomings [26] proposed a new approach which is an extremely powerful, simple, and yet efficient depiction of the video called Dynamic Images, summarizing the motion dynamics present in the video into a single image. Dynamic Images (DIs) can encrypt information in a conventional content-agonistic manner which results in highly effective long-term, stable representation of motion for classification and recognition tasks and also for other tasks like human pose estimation [27], [28], as well. Key benefits of using Dynamic Images which are as follows:

1. DIs can be attributed to various forms of sequences.
2. DIs are very efficient, fast, and simple.
3. Extraction of DIs reduces computational time, also reduces video analysis to single RBG image analysis.
4. Compact Representation of video.
5. DIs can be useful for large scale indexing
6. DIs can be processed by any CNNs

The sole objective of this paper is to emphasize the distinct

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

motion features embedded and encoded as RGB-DIs and how a deep learning neural net can be used to capture such motion characteristics and improve the task of violence detection. The fine-tuned pre-trained Inception-Resnet-V2 model is used to evaluate how well the proposed framework performs and establish the novelty of the work on the three publically available datasets or benchmarks – hockey fight, real-life violence, and movie dataset.

The paper is structured as follows: Description of the relevant work is discussed in section II, emphasizing the contribution in the field of violence detection. Section III deals in detail with the proposed framework and the results of conducted experiments are present in section IV. Lastly, concluded in Section V.

II. Related Work

Violence Detection systems dependent on deep learning like [29]–[31] are burgeoning with ever-increasing accuracies and advantages over the other. The evolutions of single-stream architectures into multi-stream state-of-the-art have also shown a significant increase in performance since these architectures incorporate the amalgamation of different information cues – motion, spatial, and temporal. Previous works based on extracting spatiotemporal features from video or pyramid of RGB image frames extracted from videos have shown tremendous potential in detecting violence with accuracies nearing almost 100% mark. [32] extracted frame-level characteristics from video using AlexNet as CNN followed by LSTM that uses convolutional gates, aggregating the frame-level characteristics and [33] used pre-trained light-weight MobileNet to reduce bulk processing of useless frames followed by 3D CNN for feature extraction. [6] used spatiotemporal encoder and bidirectional ConvLSTM architecture for better performance on complex and heterogeneous datasets like Violent Flow datasets, [34] also used CNN and LSTM architecture for extracting spatial features and learning temporal relation respectively resulting in outperforming speed as compared to previous models. These advancements are majorly based on spatiotemporal feature extraction in reaching extremely high accuracy. But with increasing complexities in such multi-stream architectures, it is also necessary to improve the quality of information that each stream captures.

Motion information is a rich tool for gathering knowledge dependent on human action. Techniques such as optical flow, optical flow based ViF, OViF are used to capture motion information. Optical flow-based methodologies record the optical-flow within consecutive image frames encapsulating the motion using key components. [22] used ViF descriptors, object detection method, for representing stats collected for short frame sequences, and bags of feature method for feature extraction. The fundamental aim of the strategy is to identify the transition of violent to non-violent action having the shortest delay since the change occurred. [24] used OViF for feature extraction and captures the essence of information about the motion magnitude change. However, only the local dynamics are provided by these methodologies, and local motion analysis is performed using simple summarization techniques. CNN can extract meaningful information from the input provided hence it is highly important to decide how the video information is provided to CNN. Recent advancements in capturing the motion information using Dynamic Images for the task of action recognition [35] have shown significant improvement in the action recognition task using motion information. Dynamic Images can summarize the whole video content into a single image (fig.1).

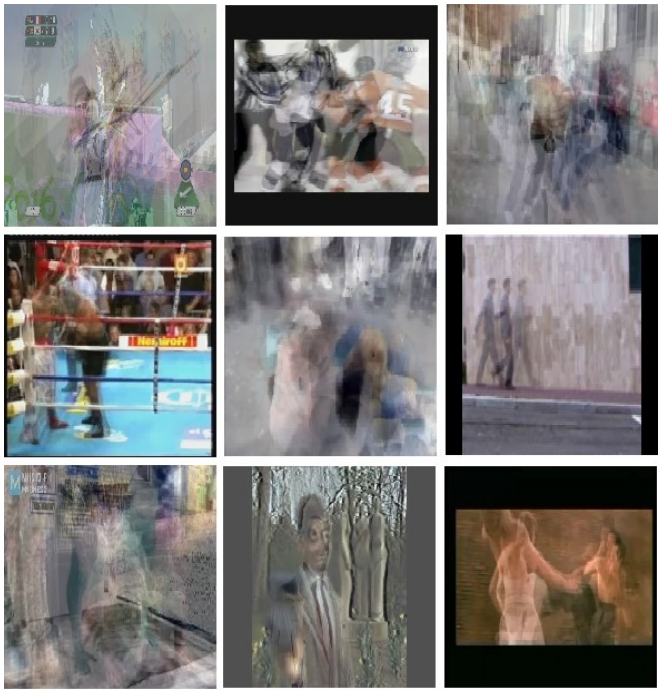


Fig. 1: Dynamic Images captured from 3 datasets - Hockey Fight dataset, Real-Life Violence Dataset and Movie Dataset summarizing violent and non-violent behavior

III. Proposed Methodology

The proposed deep neural net architecture, fine-tuned pre-trained Inception Resnet V2 based CNN model for Violence Detection is shown in Fig.2. The architecture is designed to learn motion features and use motion information for the task of Violence Detection. The motion content or the motion information present in the video is captured by transforming RGB video into Dynamic Image (DI). The transfer learning approach is used for recognizing violent and non-violent acts with the assistance of fine-tuned pre-trained Inception-Resnet-V2. DIs typically focuses on the salient object's movement by combining and averaging the background pixels as well as the movement patterns while retaining the long-term kinetics. The shape of the obtained DI is identical to the original frame. The constructed DI is transferred through the implemented architecture to compute the motion characteristics.

The architecture is a combination of Inception ResnetV2, pre-trained on the ImageNet Dataset, followed by a set of Dense and Dropout Layers. ReLU is used as the activation function on the Dense Layers expects the end Dense Layer. Last Dense Layer used the Softmax Activation. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

architecture of our proposed model is as follows :

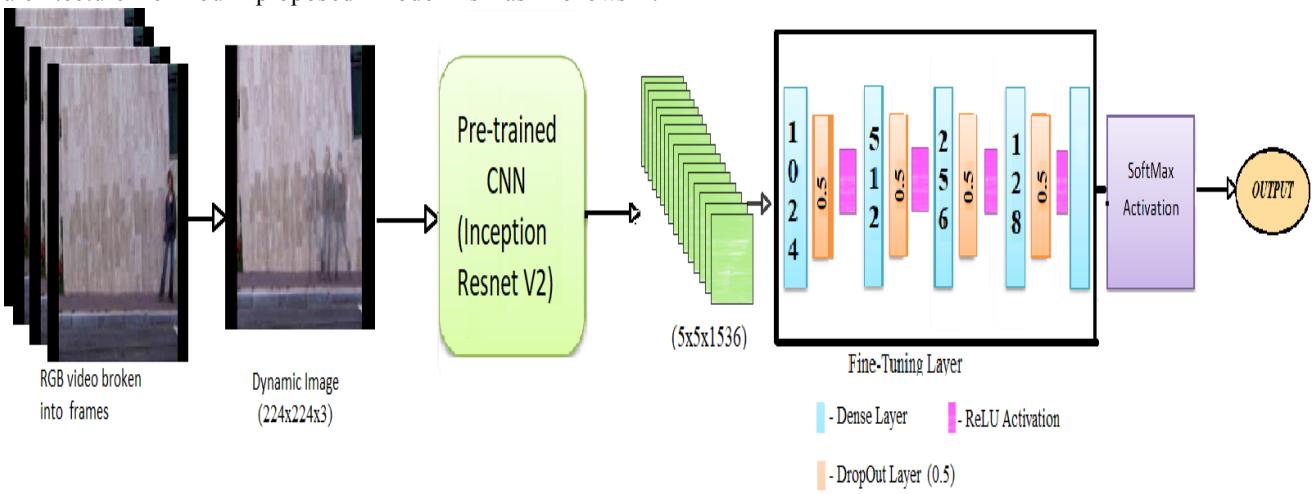


Fig. 2: Schematic Block Diagram of Proposed Framework

Input DI (224x224x3) – Inception-Resnet-V2 () – Intermediate Vector Shape (5x5x1536) – Dense Layer (1024) – DropOut Layer (0.5) – ReLU – Dense Layer (512) – DropOut Layer (0.5) – ReLU – Dense Layer (256) – DropOut Layer (0.5) – ReLU – Dense Layer (128) – DropOut Layer (0.5) – ReLU – Dense Layer (2) – Softmax ()

The pre-trained Inception-Resnet-V2 layers are not trained and kept frozen during the whole training duration. The layers for fine-tuning the model are trained end-to-end for updating weights according to the training sample. Since the data size is small, overfitting will be an issue. To counter it, the Dropout Layer helps to handle the overfitting phenomena as well as the training data size was also doubled by capturing the 2 dynamic images from a single video, created using distinct mutually exclusive frames meaning the frame used to create first DI was not used for creating the second DI. Early stopping was used so that model doesn't over-train. Without early stopping the validation loss increases and accuracy degrades. Model Checkpoint is used to obtain the best-trained model weights during training having the lowest validation loss. For testing purposes, the same saved model checkpoint having the best weights is used to achieve the desired results.

IV. Experimental Results

For substantiating the achievement of our proposed violence detection framework, three publically available benchmarks – Hockey Fight Dataset, Real Life Violence Dataset, and Movie Datasets are used.

A. Hockey Fight Dataset

The Hockey Fight Dataset was introduced by [36] containing 1000 violent and non-violent action clips each, recorded during a hockey game of the NHL. Each video has a frame rate of 25fps consisting of 50frames of 720x576 pixels. The dataset was labeled into a fight and non-fight category. All the clips have the same background with only Ice Hockey players entering the frames.

B. Real-Life Violence Dataset

A new and quite challenging benchmark is introduced [37], The Real-Life Violence Dataset, as compared to Hockey Fight and movie dataset with 2000 videos divided into 1000 violent and non-violent genre videos, having a wide range of gender, age, and race collected from diverse backgrounds and environment. Violent videos are captured in an environment like a prison, street, and schools whereas the non-violent videos are captured in environments like sports fields and arena featuring events such as swimming, doing archery, playing basketball. The average duration of the clip is 5seconds. With maximum duration being 7seconds and a minimum duration of 3 seconds with a frame rate of 25fps.

C. Movie Dataset

The Movie Dataset [36] is a considerably small dataset containing only 200 movie clips from action scenes bifurcated equally into violent and non-violent video genres. The movie dataset has also varying backgrounds. The clips have an average duration of 1 second and some clips have a duration of 2seconds, having a frame rate of either 25fps or 30fps.

In the experiment, the end-to-end training of fine-tuned layers of the proposed model takes place. Before the training phase begins, the dataset is subdivided into testing and training samples using the 30-70 splitting strategy, and training sample videos are processed to get dynamic images. For each training video, 2 dynamic images are computed by adopting the methodology proposed by [35]. RGB video was broken into a stack of frames and 2 sets of frames were produced, one starting with 1st frame having stride of frameRate/2 and the other set starting with 1st frame = frameRate/4 and having stride of frameRate/2. Each frame in the 2 sets was taken and split into R, G, and B channels separately. A coefficient was computed for each frame and a list of frames split by channels was multiplied by the coefficients. The weighted aggregate of all 3 channels - R, G, and B in each frame separately is collected to obtain the R, G, B channel of the DI. Once the R, G, B channels of DI is achieved, they were then merged to form a single DI

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

normalized into pixels having a value between 0-255 as shown in Fig. 3.

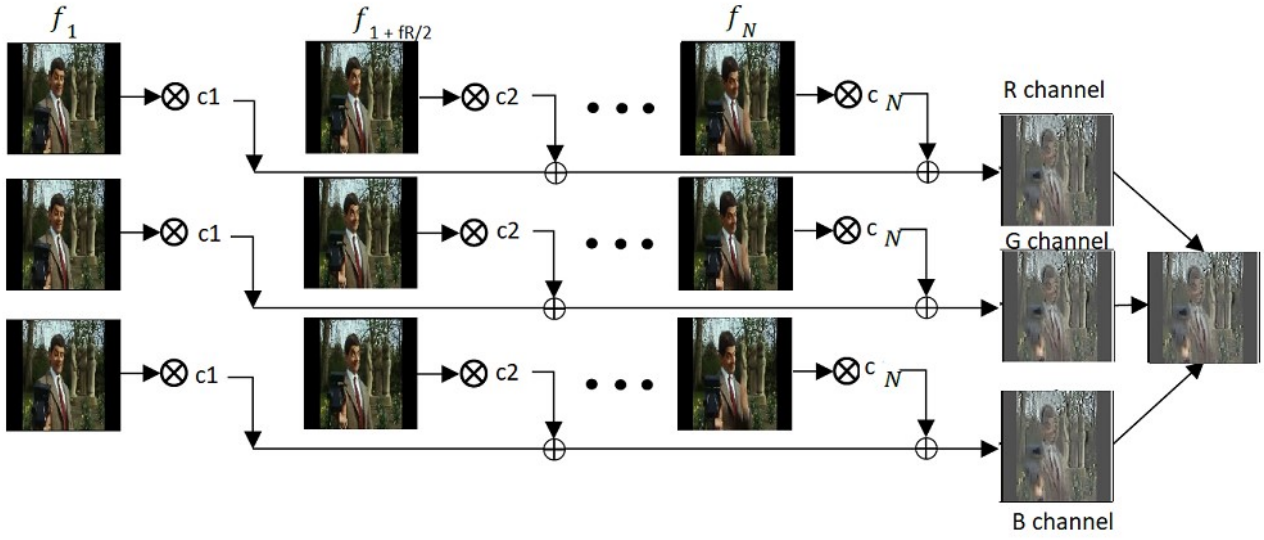


Fig. 3: Dynamic Image Formation

The size of the generated DI was the same as the original frame. Once DI were obtained for the training sample, motion features were extracted from DI using pre-trained Inception-Resnet-V2 on ImageNet dataset followed by fine-tuning layers using Adam optimizer with epoch = 50, and batch size = 64. Callbacks were used to save the best model weights based on minimum validation loss and early stopping with patience = 20 was used to stop the model in case model weights were not improving, hence accelerating the training process. Training input was further divided into actual training set and validation set using 80-20 splitting strategy. In Testing, the model checkpoint saved during training is loaded and DI is computed for each testing video simultaneously. The performance of our proposed architecture on all 3 benchmarks is compiled in Table I highlighting the highest training and validation accuracy achieved and the loss incurred during the training phase as well as the highest testing accuracy achieved during the testing phase.

Table I. Training, Validation, and Testing Accuracy along-with Training, Validation Loss achieved on the 3 publically available benchmarks for our proposed model

Dataset Name	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Testing Accuracy
Hockey Fight Dataset	100 %	0.0212 %	94.99 %	17.18 %	93.33 %
Real-Life Violence Dataset	100 %	0.0252 %	92.8571 %	21.58 %	86.7892 %
Movie Dataset	100 %	7.30e-03 %	100 %	5.75e-05 %	100 %

Table II. Violence Detection Accuracy (%) Comparison on Hockey Fight Dataset

Method	Accuracy
ViF [22]	81.6%
OViF [24]	84.2%
ViF + OViF [24]	86.03%
HOMO	89.3%
Our Proposed Approach	93.33%

Table II also shows the comparison between the accuracy achieved by our proposed framework and similar works based on motion feature extraction on the Hockey Fight dataset and it is evident that our proposed model has outperformed the rest.

Table III. Violence Detection Accuracy (%) Comparison on Movie Dataset

Method/Classifier		Feature Extracted	Accuracy
[32]		Spatiotemporal	100%
[33]		Spatiotemporal	99.9%
ViF [38]	SVM	Motion	96.7%
	Adaboost		92.8%
	Random Forest		88.9%

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

Our Proposed Approach	Motion	100%

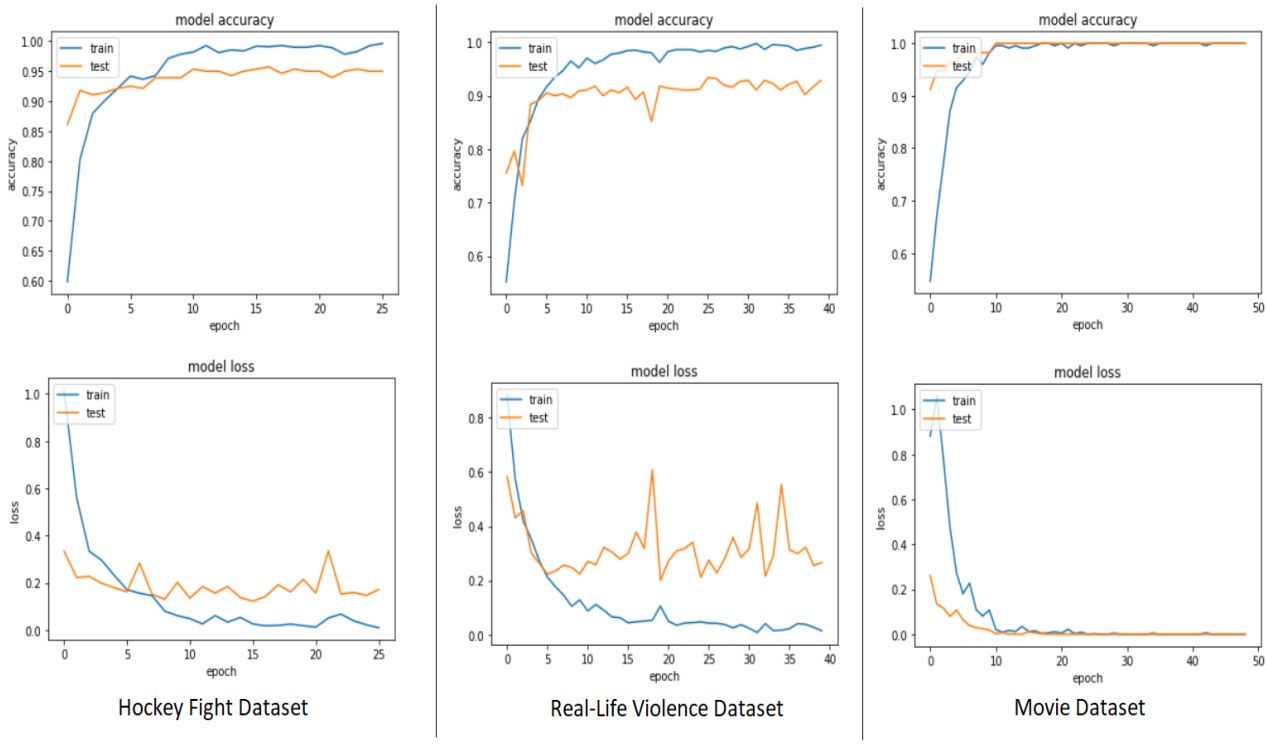


Fig. 4: Plot of Training-Validation Accuracy and Training-Validation Loss during each iteration in the training phase for all 3 dataset

It is also observed that our proposed framework is at par in terms of accuracy achieved with methods based on spatiotemporal feature extraction for detecting violence. Table III shows the accuracy comparison between our framework based on the extraction of motion features and other methodologies based on either extraction of motion or spatiotemporal feature on the movie dataset. The plot of training and validation accuracy, as well as loss incurred on training and validation sample after each iteration in the training phase for all 3 datasets, are shown in Fig 4. On average, it took nearly 1 second or less to convert single RGB video into RGB DI and predict the label as violent or non-violent.

It is also worth noting that even though training accuracy reaches 100% mark, training loss is yet a non-zero value. The reason behind this behavior is that accuracy and loss often appear to be inversely proportional but there is no concrete mathematical relationship between these two metrics [39]. Accuracy can be seen as a count of correct predictions whereas loss can be seen as a distance between predicted probability and true value. The Real-Life Violence Dataset being quite new to the domain isn't used extensively yet, the best testing accuracy achieved by [37] was 88.2% which is higher than achieved by our proposed model. Inception-Resnet-V2 was chosen over other common pre-trained models due to its better top-1 and top-5 accuracy [40].

V. CONCLUSION

With the rising population and increasing need for surveillance has posed growing demands of systems that are capable to detect violent acts automatically. Violence detection is already an interesting research field and CNN's have made an exceptional breakthrough in the field of detecting violence. Although, a great deal of success has been achieved by CNN's yet there a lot of issues that require further investigation. In this work, a novel deep architecture focused on single-stream RGB-DI is introduced that capitalizes on motion features obtained from violent or non-violent videos. DIs are simple yet extremely powerful representations of videos summarizing videos into a single image. DIs can encrypt the summary of the video in a very compact environment allowing for excellent performance by using motion features. Any existing or future CNN architectures can also utilize the DI as input to achieve better results. Furthermore, incorporating DIs in motion stream used in multi-stream state-of-the-arts can enable in learning meaningful results as well. Applying DIs with very recent very deep CNN – pre-trained InceptionResnetV2 with fine-tuning enabled in achieving more speed and better accuracy as compared to existing violence detection methods based on motion feature extraction. Experiments on publically available violence detection dataset benchmarks are conducted to validate the performance and experiments demonstrate the effectiveness of DIs in achieving impressive performance despite their simplicity. In the future, spatiotemporal features of video is aimed to extract along with motion features in a multi-stream architecture to make violence detection more

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

robust enhancing the performance of the system.

REFERENCES

[1] "New Delhi Streets Turn Into Battleground, Hindus vs. Muslims - The New York Times." [Online]. Available: https://www.nytimes.com/2020/02/25/world/asia/new-delhi-hindu-muslim-violence.html. [Accessed: 21-Jun-2020].

[2] "Fiery Clashes Erupt Between Police and Protesters Over George Floyd Death - The New York Times." [Online]. Available: https://www.nytimes.com/2020/05/30/us/minneapolis-floyd-protests.html. [Accessed: 21-Jun-2020].

[3] B. Jiang, F. Xu, W. Tu, and C. Yang, "Channel-wise Attention in 3D Convolutional Networks for Violence Detection," *Proc. - 2019 Int. Conf. Intell. Comput. Its Emerg. Appl. ICEA 2019*, pp. 59–64, 2019.

[4] E. Ditsanthia, L. Pipanmaekaporn, and S. Kamonsantiroj, "Video Representation Learning for CCTV-Based Violence Detection," *TIMES-iCON 2018 - 3rd Technol. Innov. Manag. Eng. Sci. Int. Conf.*, pp. 1–5, 2019.

[5] M. Baba, V. Gui, C. Cernazanu, and D. Pescaru, "A sensor network approach for violence detection in smart cities using deep learning," *Sensors (Switzerland)*, vol. 19, no. 7, pp. 1–17, 2019.

[6] A. Hanson, K. Pnvr, S. Krishnagopal, and L. Davis, "Bidirectional convolutional LSTM for the detection of violence in videos," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11130 LNCS, pp. 280–295, 2019.

[7] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[8] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[9] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[10] J. Mahmoodi and A. Salajeghe, "A classification method based on optical flow for violence detection," *Expert Syst. Appl.*, vol. 127, pp. 121–127, 2019.

[11] K. Lloyd, P. L. Rosin, D. Marshall, and S. C. Moore, "Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM)-based texture measures," *Mach. Vis. Appl.*, vol. 28, no. 3–4, pp. 361–371, 2017.

[12] A. A. Einstein, "DETECTION OF REAL-WORLD FIGHTS IN SURVEILLANCE VIDEOS Mauricio Perez , Alex C . Kot School of Electrical and Electronic Engineering University of Campinas Institute of Computing," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 2662–2666, 2019.

[13] A. Singh, D. Patil, and S. N. Omkar, "Eye in the sky: Real-time drone surveillance system (DSS) for violent individuals identification using scatternet hybrid deep learning network," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2018-June, pp. 1710–1718, 2018.

[14] Z. Guo, F. Wu, H. Chen, J. Yuan, and C. Cai, "Pedestrian violence detection based on optical flow energy characteristics," *2017 4th Int. Conf. Syst. Informatics, ICSAI 2017*, vol. 2018-Janua, no. Icsai, pp. 1261–1265, 2017.

[15] A. M. R. Abdali and R. F. Al-Tuma, "Robust Real-Time Violence Detection in Video Using CNN and LSTM," *SCCS 2019 - 2019 2nd Sci. Conf. Comput. Sci.*, pp. 104–108, 2019.

[16] X. Yang, X. Yang, M. Y. Liu, F. Xiao, L. S. Davis, and J. Kautz, "Step: Spatio-temporal progressive learning for video action detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019-June, pp. 264–272, 2019.

[17] C. Feichtenhofer, A. Pinz, and A. Zisserman, "Convolutional Two-Stream Network Fusion for Video Action Recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, no. i, pp. 1933–1941, Apr. 2016.

[18] T. Z. Ehsan and M. Nahvi, "Violence detection in indoor surveillance cameras using motion trajectory and differential histogram of optical flow," *2018 8th Int. Conf. Comput. Knowl. Eng. ICCKE 2018*, no. Iccke, pp. 153–158, 2018.

[19] L. Lazaridis, A. Dimou, and P. Daras, "Abnormal behavior detection in crowded scenes using density heatmaps and optical flow," *Eur. Signal Process. Conf.*, vol. 2018-Septe, pp. 2060–2064, 2018.

[20] P. D. Garje, M. S. Nagmode, and K. C. Davakhar, "Optical Flow Based Violence Detection in Video Surveillance," *2018 Int. Conf. Adv. Commun. Comput. Technol. ICACCT 2018*, pp. 208–212, 2018.

[21] H. Wang, A. Kläser, C. Schmid, and C. L. Liu, "Action recognition by dense trajectories," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3169–3176, 2011.

[22] T. Hassner, Y. Itcher, and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 1–6, 2012.

[23] P. Vashistha, C. Bhatnagar, and M. A. Khan, "An architecture to identify violence in video surveillance system using ViF and LBP," *Proc. 4th IEEE Int. Conf. Recent Adv. Inf. Technol. RAIT 2018*, pp. 1–6, 2018.

[24] Y. Gao, H. Liu, X. Sun, C. Wang, and Y. Liu, "Violence detection using Oriented VIolent Flows," *Image Vis. Comput.*, vol. 48–49, pp. 37–41, 2016.

[25] A. A. Mishra and G. Srinivasa, "Automated detection of fighting styles using localized action features," *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 1385–1389, 2018.

[26] H. Bilen, B. Fernando, E. Gavves, and A. Vedaldi, "Action Recognition with Dynamic Image Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2799–2813, 2018.

[27] C. Dhiman, M. Ieee, D. K. Vishwakarma, and S. Member, "View-invariant Deep Architecture for Human Action Recognition using late fusion."

[28] C. Dhiman and D. K. Vishwakarma, "View-Invariant Deep Architecture for Human Action Recognition Using Two-Stream Motion and Shape Temporal Dynamics," *IEEE Trans. Image Process.*, vol. 29, no. DI, pp. 3835–3844, 2020.

[29] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 6479–6488, 2018.

[30] S. R. Dinesh Jackson *et al.*, "Real time violence detection framework for football stadium comprising of big data analysis and deep learning through bidirectional LSTM," *Comput. Networks*, vol. 151, pp. 191–200, Mar. 2019.

[31] M. Marsden, K. McGuinness, S. Little, and N. E. O'Connor, "ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification," *2017 14th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2017*, 2017.

[32] S. Sudhakaran and O. Lanz, "Learning to detect violent videos using convolutional long short-term memory," *2017 14th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2017*, 2017.

[33] F. U. M. Ullah, A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, "Violence detection using spatiotemporal features with 3D convolutional neural network," *Sensors (Switzerland)*, vol. 19, no. 11, pp. 1–15, 2019.

[34] B. Peixoto, B. Lavi, J. P. Pereira Martin, S. Avila, Z. Dias, and A. Rocha, "Toward Subjective Violence Detection in Videos," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2019-May, pp. 8276–8280, 2019.

[35] H. Bilen, B. Fernando, E. Gavves, A. Vedaldi, and S. Gould, "Dynamic Image Networks for Action Recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 3034–3042, 2016.

[36] E. Bermejo Nievas, O. Deniz Suarez, G. Bueno García, and R. Sukthankar, "Violence detection in video using computer vision techniques," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6855 LNCS, no. PART 2, pp. 332–339, 2011.

[37] M. M. Soliman, M. H. Kamal, M. A. El-Massih Nashed, Y. M. Mostafa, B. S. Chawky, and D. Khattab, "Violence Recognition from Videos using Deep Learning Techniques," *Proc. - 2019 IEEE 9th Int. Conf. Intell. Comput. Inf. Syst. ICICIS 2019*, pp. 80–85, 2019.

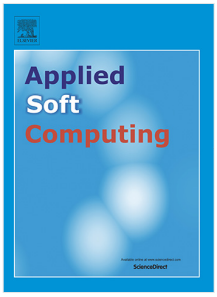
[38] I. S. Gracia, O. D. Suarez, G. B. Garcia, and T. K. Kim, "Fast fight detection," *PLoS One*, vol. 10, no. 4, pp. 1–19, 2015.

[39] "Accuracy and Loss - AI Wiki." [Online]. Available: https://docs.paperspace.com/machine-learning/wiki/accuracy-and-loss. [Accessed: 03-Jul-2020].

[40] Keras Documentation, "Keras Applications," 2017. [Online]. Available: https://keras.io/api/applications/. [Accessed: 03-Jul-2020].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:24:21 UTC from IEEE Xplore. Restrictions apply.

Journal Pre-proof



Deep transfer with minority data augmentation for imbalanced breast cancer dataset

Manisha Saini, Seba Susan

PII: S1568-4946(20)30697-9
DOI: https://doi.org/10.1016/j.asoc.2020.106759
Reference: ASOC 106759

To appear in: *Applied Soft Computing Journal*

Received date : 3 April 2020
Revised date : 23 August 2020
Accepted date : 24 September 2020

Please cite this article as: M. Saini and S. Susan, Deep transfer with minority data augmentation for imbalanced breast cancer dataset, *Applied Soft Computing Journal* (2020), doi: https://doi.org/10.1016/j.asoc.2020.106759.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier B.V.

Journal Pre-proof

Highlights

- A novel deep transfer approach is proposed for the imbalanced Breast cancer dataset.
- The assembled deep transfer network comprises of pre-trained VGG16 layers at the lower level.
- Higher level contains 2D convolution, pooling, dropout, dense layers with Batch Normalization.
- Deep Convolutional Generative Adversarial Network is used for minority data augmentation.
- Proposed approach outperforms state of the art networks for different Magnification factors.

Journal Pre-proof

*Manuscript
Click here to view linked References

Journal Pre-proof

Deep Transfer with Minority Data Augmentation for Imbalanced Breast Cancer Dataset

Manisha Saini* and Seba Susan

Delhi Technological University, Delhi, India

Email: manisha.saini44@gmail.com

Abstract

Clinical diagnosis of breast cancer is a challenging problem in the biomedical domain. The BreakHis breast cancer histopathological image dataset consists of two classes: Benign (Minority class) and Malignant (Majority class). The imbalanced class distribution results in the degradation of performance of the classifier model due to biased classification towards the majority class. To tackle this problem, a novel learning strategy that involves a deep transfer network has been proposed in this paper, in collaboration with Deep Convolution Generative Adversarial network (DCGAN). DCGAN is used in the initial phase for data augmentation of the minority class only. The dataset, with the class distribution now balanced, is applied as input to the deep transfer network. The proposed deep transfer architecture has at its core, the initial pre-trained layers (until block 4 pool layer) of the VGG16 deep network architecture pre-trained on the ImageNet object classification dataset. The higher end of our transfer network comprises of batch normalization, 2D convolutional (CONV2D) layer, Global Average Pooling 2D, dropout and dense layers that are added to enhance the network's performance. Experiments on the benchmark BreakHis dataset for different magnification factors: 40X, 100X, 200X and 400X validate the efficiency of the proposed deep transfer learning approach due to the high scores achieved as compared to the state-of-the-art deep networks.

Keywords: Data augmentation; Imbalanced dataset; DCGAN; VGG16; transfer learning; batch normalization.

Journal Pre-proof

Journal Pre-proof

1. Introduction

Cancer is of serious concern worldwide since it has started affecting a higher rate of population [1]. Hence there is need to have correct and timely diagnosis of cancer which will ultimately speedup the process of taking up right preventive measures at an early stage to save the life of patients worldwide. Amongst cancer types, breast cancer is increasing at an alarming rate nowadays [2]. For detection of breast cancer cells, histopathological images are analyzed with the help of tissue biopsy, followed by which histopathological microscopic images are generated at different magnification factors such as: 40X, 100X, 200X and 400X. These histopathological images further help to diagnose breast cancer cells in order to distinguish the cancerous cells from the normal tissue cells [3]. A manual process is usually undertaken to analyze the histopathological images, which is a time-consuming and difficult task [4]. It is notable to mention that there is high probability of inaccurate diagnosis due to the lack of resources and human errors. The advancement in technologies: computer vision, machine learning, and deep learning [5] help to automate the process for analyzing the histopathological images to segregate the Benign samples from the Malignant samples containing the cancerous cells. With the high availability of computational power (in the form of Graphical Processing Units (GPUs), large compute clusters), deep Convolutional Neural Networks (CNN) are now conveniently used for computer vision applications [6]. Broadly categorizing, there are two variations of CNN: (a) trained from scratch [7] and (b) pre-trained models [8]. Researchers have come up with effective techniques to detect cancer using pre-trained networks. Fabio Spanhol et al. (2016) conducted an experiment using CNN based on deep learning for BREAKHIS cancer dataset, and they had further proposed a new approach using image patches with the sliding window mechanism [9]. Bardou et al. (2018) had compared various approaches such as

handcrafted features extracted by using locality linear coding, the Bag of visual words learned by SVM classifier and CNN [10]. It was found from the study that the CNN outperformed the handcrafted features. Bayramoglu et al. (2016) proposed a single task CNN as well as multi-task convolutional neural networks (MTCNN) [11]. Litjens et al. (2016), presented a tool based on deep learning and CNN that can effectively perform cancer detection using a network trained on histopathological images of prostate and breast tissue by applying patch extraction [12]. Another variant of CNN, pre-trained networks are already trained on a larger standard dataset and knowledge is transformed from one domain to another domain which makes it quite useful in healthcare sector. Many researchers have worked on pre-trained networks in the biomedical field, such as Xie et al. (2019) who conducted supervised and unsupervised experiments using a transfer learning approach [13]. From the supervised classification results, it was found that the Inception ResNet V2 pre-trained network performs well. Further, the authors had also developed a novel set of deep features using Inception ResNet v2 pre-trained with the auto-encoder and K-means algorithm. Deniz et al. (2018) proposed concatenating the features extracted by combining the layers from AlexNet and VGG16 pre-trained model. The new features were learnt by support vector machines (SVM) for the breast cancer classification task [14]. Garud et al. (2017) proposed a pre-trained GoogleNet based approach for separating Benign samples from Malignant using microscopic high magnification multi-view samples [15]. In [17], the authors had compared deep neural network, Bag of visual words and data augmentation for classification of a small imbalanced dataset and found from the analysis that deep neural network works well in comparison to other techniques.

From the literature review, it is noted that many researchers have indeed applied CNN and pre-trained models in order to improve the performance of classification models for diagnosis of

Journal Pre-proof

breast cancer. But there are still various issues that need to be carefully dealt with, such as the imbalanced nature of the biomedical datasets, which occurs because of the uneven distribution of samples related to cancerous and non-cancerous cells. To tackle the class-imbalance problem in data mining [18] [19] [20], various methods have been proposed such as oversampling, undersampling, and hybrid sampling approaches. Whereas in computer vision, the general method of data augmentation has been applied in several works to improve the class-distribution and the overall performance of the model. Data augmentation is achieved by applying affine transformation operations to images, including rotation, scaling, translating, etc. Antoniou et al. (2017), proposed the use of generative adversarial network for augmentation of data samples. It was observed that the data augmentation technique was able to adapt to the training distribution and improved the performance of classifiers on various datasets vastly. Their approach was able to improve the classification results on VGG-Face and EMNIST datasets [16]. Also Lyu et al. (2020) proposed a novel de-noising approach based on generative adversarial network variant that made use of VGG architecture [21].

Frid-Adar et al. (2018) presented a novel model in [22] using Generative Adversarial Networks (GAN) using data augmentation for liver lesion image dataset. In [23], the authors synthetically increased the number of images in the minority class by applying various data augmentation operations on the minority class only in order to balance the class distribution. Extending the idea further, in our current paper, Deep Convolution Generative Adversarial network (DCGAN) [24] for data augmentation has been applied only on the minority class in order to increase the number of image samples of the minority class, which in our experiments is the Benign (non-cancerous) class, by generating fake images, and to also balance the class-distribution. In this work, we have proposed a novel transfer learning-based architecture to

Journal Pre-proof

Journal Pre-proof

transfer the knowledge from object detection domain (ImageNet dataset) to the biomedical cancer domain.

The highlights of this paper are summarized as follows: (a) Successfully created a novel deep transfer model using DCGAN as data augmentation technique for cancer-related biomedical imbalanced datasets for different magnification factors: 40X, 100X, 200X and 400X. (b) Mitigated the effect of covariant shift by using the Batch Normalization layer in the proposed deep network. (c) Proposed a novel transfer learning approach by transferring the knowledge from the source ImageNet object dataset to the target Breast cancer dataset effectively. (d) Analyzed the proposed transfer learning approach in comparison to other state-of-the-art networks for distinguishing the Benign samples from Malignant in case of imbalanced dataset. The rest of the paper is organized as follows. In section 2, the proposed deep transfer network with DCGAN for data augmentation is discussed in an elaborative manner. In section 3, discussion regarding the BreakHis dataset and the implementation details required for performing the experimental task related to the proposed work and other state-of-the-art networks are outlined. The final results are compiled and discussed in section 4. Finally, conclusion of the paper is presented in section 5.

2. Proposed Deep Transfer Network With DCGAN For Data Augmentation

In this paper, we have presented a novel deep transfer learning approach for the classification of BreakHis dataset [25], specifically for biomedical datasets. Further, DCGAN has been used for data augmentation of the minority class to overcome the class-imbalance problem and overall boost the performance of the classification model as shown in Figure 1.

Journal Pre-proof

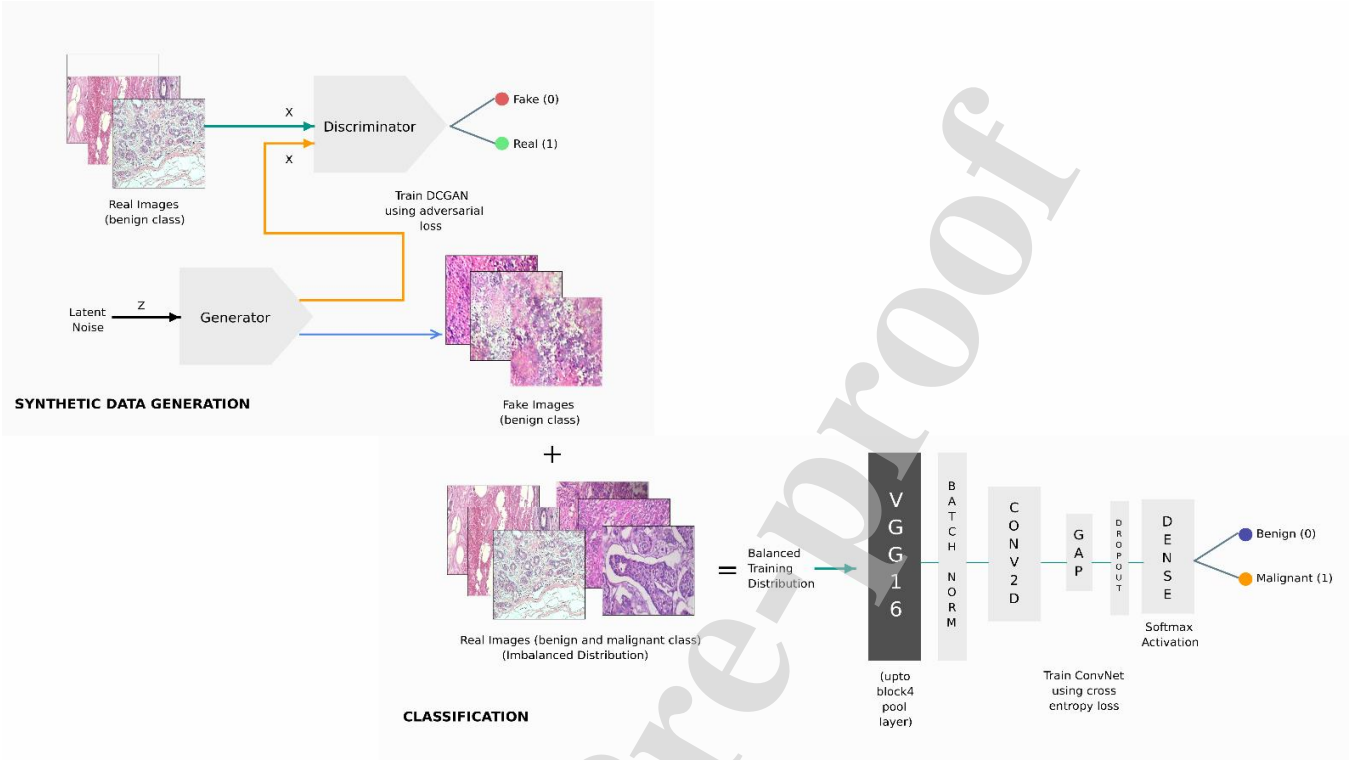


Figure 1: Proposed novel deep transfer network in collaboration with Deep Convolutional Generative Adversarial network (DCGAN).

In the proposed approach, Visual Geometry Group (VGG16) network [26] is modified. We have used VGG16 network in the proposed approach since VGG16 works well for images of different scales as observed from the work given in original paper of VGG16. Since our work involves images of different magnification factors from the BreakHis dataset (40X, 100X, 200X and 400X), Therefore VGG16 is a suitable choice of network for dealing with microscopic images of varying magnification factors. Another distinguishing factor of the VGG16 network would be that it uses smaller receptive fields in comparison to AlexNet and previous Convnets. In the proposed work, we have used configuration D as given in the original paper of VGG16, where 3x3 Convolution and 16 layers are taken into consideration. Moreover, configuration D is

Journal Pre-proof

Journal Pre-proof

computationally less expensive than VGG19 network [26] and yet is able to achieve significant accuracy on large-scale image classification tasks. The different convolutional layers of VGG16 make use of different receptive field sizes allowing it to learn multi-scale (even, high-level and low-level) information easily which will ultimately be useful for edge detection [45]. This makes VGG16 network more suitable to work with images of different scales and varying aspect ratios. Hence, when dealing with histopathological images of different zoom factors VGG16 network is robust and more adept in comparison to other networks.

The previous works [40] [46] [47] [48] have inspired us to modify the VGG16 network for our proposed methodology. The Original VGG16 network architecture consists of large number of parameters which results memory constraint problems and large number of floating point operations. In this paper, we intend to reduce the shortcomings in the original VGG16 network architecture. As inspired from the work done by Baheti et al. (2018) in [48], we have reduced the number of parameters by replacing the flatten and three dense layers present in the original network architecture with the global average pooling (GAP), dropout and one dense layer, as shown in the model overview in Figure 1. This replacement of layers have ultimately reduced the memory requirements and improves the time taken to train the network. Addition of one convolutional layer

prior to the GAP layer is inspired from the concept given by Yosinski et al. (2014) [29], that the top layers should learn high-level features specific to the target biomedical dataset. Whereas on the other hand the lower layers comprising of the pre-trained layers of VGG16 upto block 4 pool layer, are intended for learning the more generic features. Further, the DCGAN approach has been used at data level for generating fake images in order to balance the dataset so as to bring the minority class population at par with the majority class population, overcoming the

imbalance class problem.

To elaborate the proposed network, we have embedded the VGG16 pre-trained layers (from initial layer up to block 4 pool layer), from the lower level of the hierarchy and fused it with the newly learnt layers trained on the target biomedical dataset. The original VGG16 pre-trained end-to-end architecture consists of convolutional layers along with pooling layers and fully connected layers, from which only the layers till block 4 have been taken, instead of considering all the layers of the pre-trained network so as to extract the useful bottleneck features. After the block 4 pool layer, the rest of the layers of the original pre-trained VGG16 network are replaced by new layers

as shown in the model overview in Figure 1, which were selected to improve the overall performance of the classification model. Batch normalization [27] layer has been added along with 2D convolution (CONV2D), Global Average Pooling 2D [28], dropout and dense layers as illustrated in the detailed network architecture shown in Figure 2 and Figure 3. Features extracted from VGG16 network till block4 pool layer result in features of dimensions 14 x 14, which will be more informative for edge localization than the 7 x 7 features resulting from block 5 pool layer [45]. Since, we introduce an additional convolutional layer in our architecture to learn the high-level features specific to the targeted biomedical classification problem, we require to choose a robust bottleneck endpoint from the pre-trained model. The use of block 4 instead of block 3 or even block 2 ensures that the number of trainable parameters in the added convolutional layer remain minimum helping us reduce overall training time. Hence, bottleneck features extracted till block 4 pool layer works well as verified by conducting experiments as elaborated in the results and discussion section.

Journal Pre-proof

Journal Pre-proof

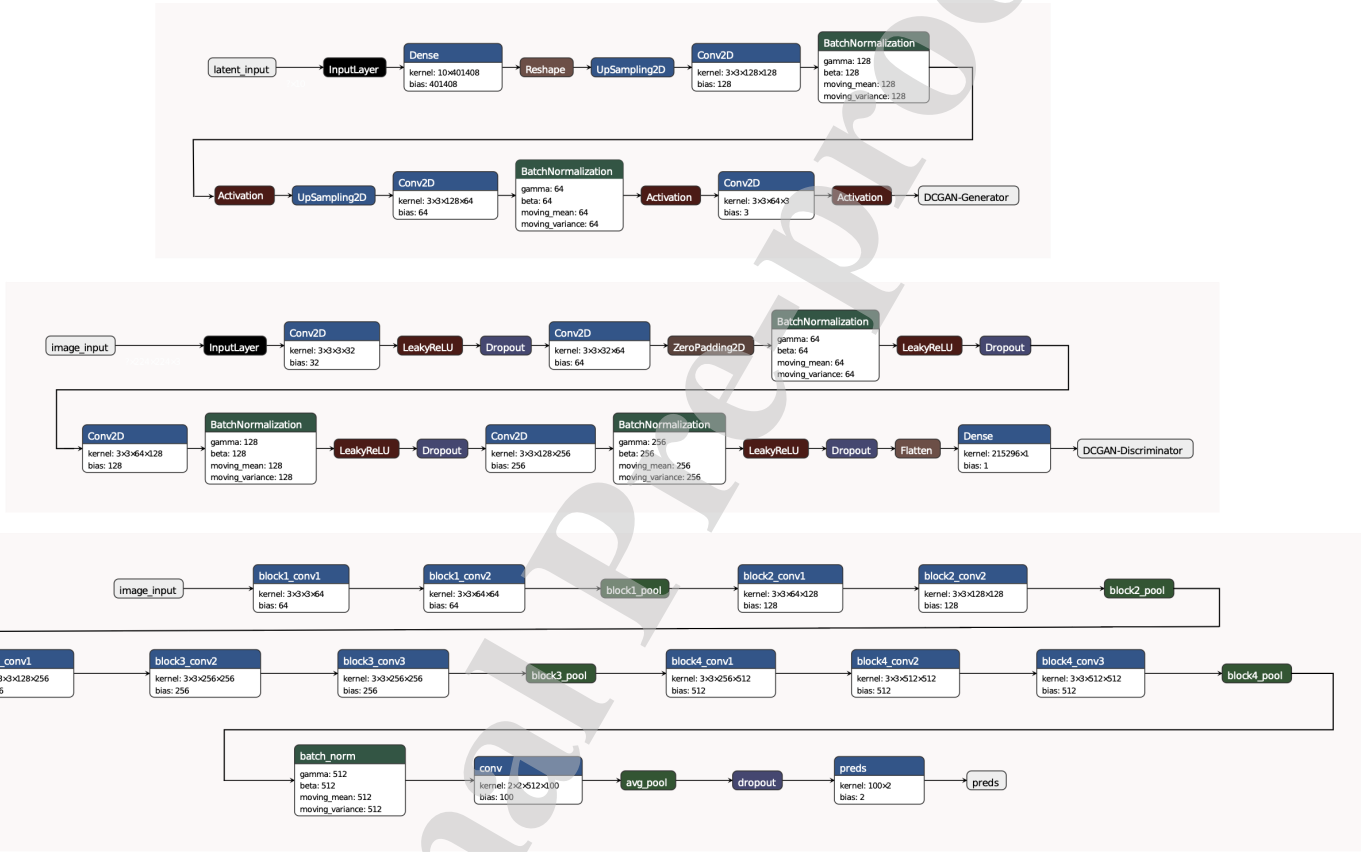


Figure 2. (i) Generator architecture of DCGAN (top row). (ii) Discriminator architecture of DCGAN (middle row). (iii) Proposed network architecture with VGG16 pre trained layers till block4_pool (bottom row).

Journal Pre-proof

Journal Pre-proof

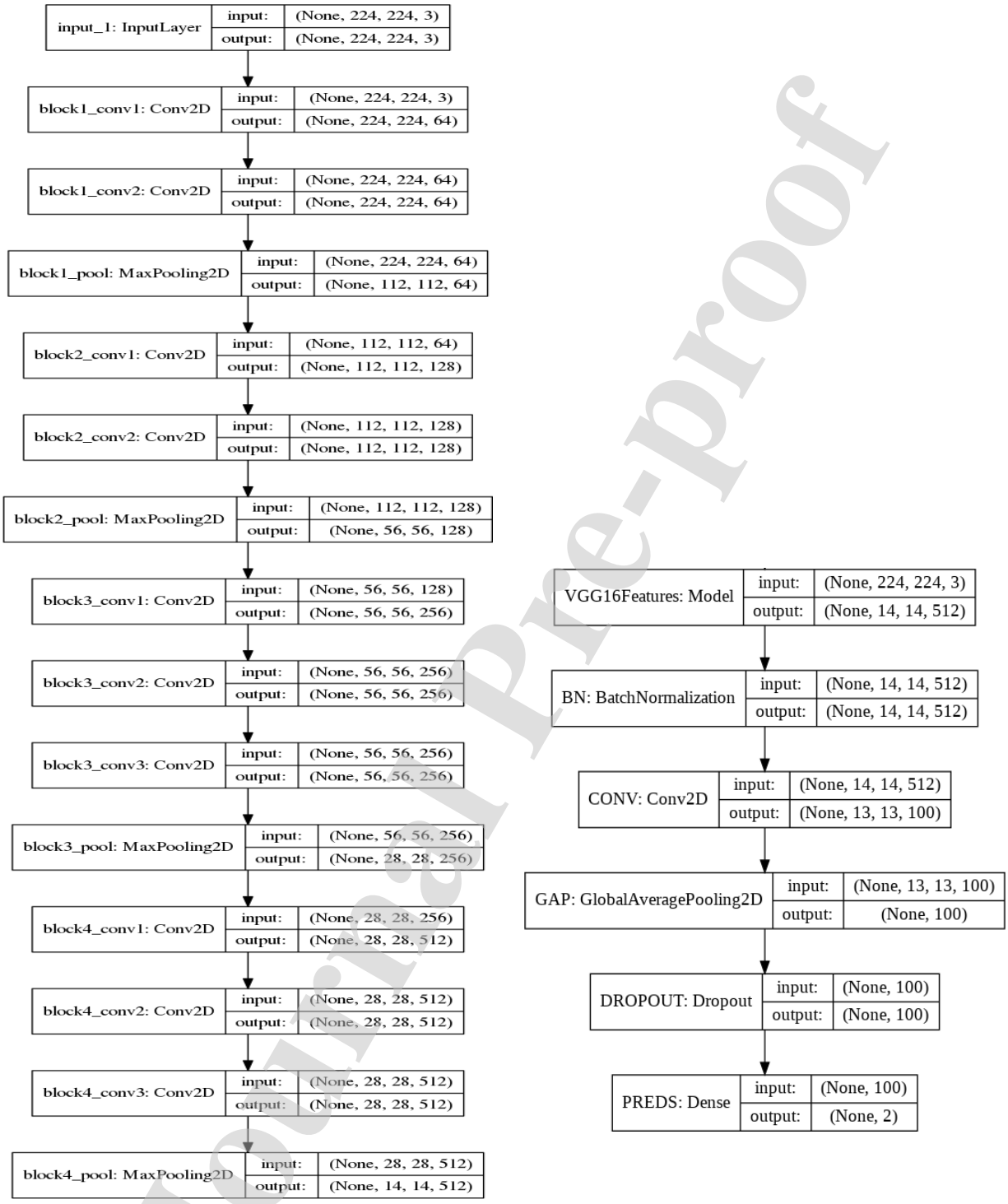


Figure 3. (i) VGG16 architecture upto block4_pool layer (left column)

(ii) Layers added after block4_pool layer for proposed network architecture (right column)

The batch normalization layer normalizes the data values to mitigate the effective covariant shift, and not applying the batch normalization will lead to biased results as observed in the results section by performing suitable experiments. In addition to that, convolutional (CONV2D) layer is applied, based on the concept proposed by Yosinski et al. (2014) [29] that if both the source as well as target datasets are dissimilar, then features needs to be extracted from the target-specific domain at higher layers to improve the overall performance of the model. Here we have considered the source domain as ImageNet dataset and considered the target domain as breast cancer dataset. So based on the same concept, low-level features are extracted from lower layers specific to ImageNet dataset and the high-level features specific to cancer domain are extracted from the higher layers that we have inserted in our new architecture. Visualization is presented in Figure 4 to represent the various feature maps obtained from the convolutional (CONV2D) layer.

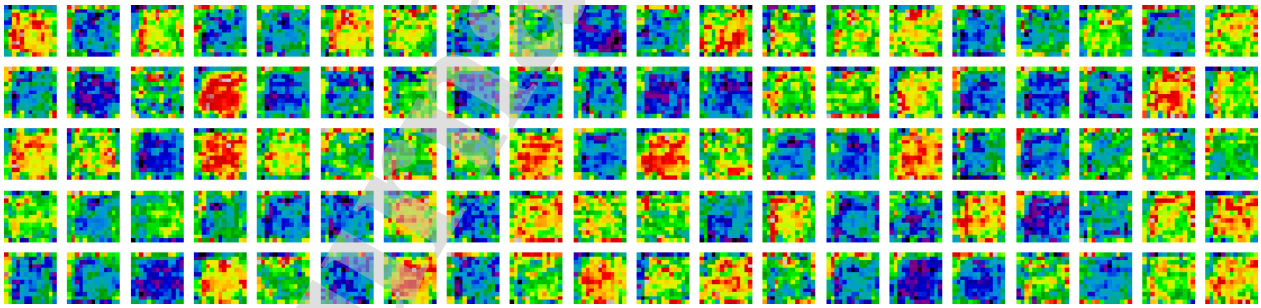


Figure 4. Illustration of feature maps obtained by applying filters at the convolutional (CONV2D) layer of the proposed model.

After CONV2D, Global Average Pooling 2D layers is used, instead of max pooling,

Journal Pre-proof

Journal Pre-proof

because this makes the model more robust to spatial translations in the data [27] and receptive to higher intensities. Further, dropout layer is added to improve the regularization and reduce the overfitting problem in the proposed network, and finally the dense layer is used at the end of the network to make the network specific to solving the binary classification problem by using the softmax function in the last layer of network.

Lastly, we describe in detail, a most important component, the Deep Convolutional Generative Adversarial Network (DCGAN) [24] [30] used for minority data augmentation in the initial phase of our experiments. DCGAN is used to generate synthetic high quality fake images from the available distribution of the minority data. As DCGAN uses a stable architecture, it is able to learn good representation of hierarchical features, more precisely capturing higher details from images quickly. DCGAN network with the help of this architecture, trains quite well and generates better quality of fake samples than Generative Adversarial Network (GAN) which is only based upon fully connected neurons. Fake images generated using GAN is generally of low quality and possesses higher noise ratio. It is evident from our analysis (comparison of FID as discussed later), that addition of convolutional layers to generative networks improve their quality of synthetic sample generation as seen in the case of DCGAN, over basic GAN. Even addition of Batch Normalization layer in the DCGAN network has significantly helped in training the network by normalizing the intermediate input values. DCGAN is applied only on the minority class in order to balance out the distribution samples of both the classes, to make the sample distribution of the minority class (Benign) equivalent to that of the majority class (Malignant), which will ultimately lead to an overall improvement in the performance of the proposed deep transfer network, as shown in Figure 5.

Journal Pre-proof

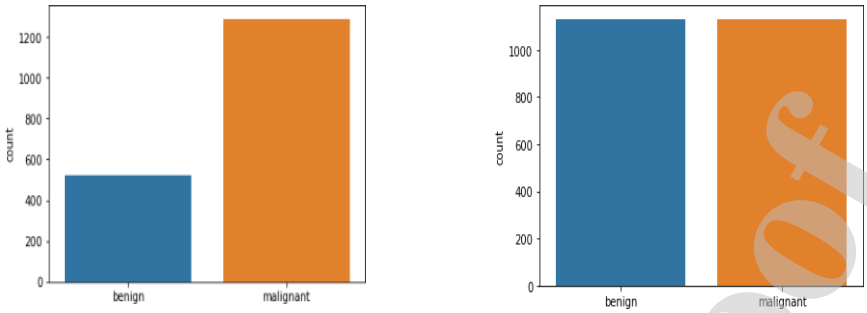


Figure 5. Sample distribution of the classes before and after applying DCGAN for data augmentation on the BreakHis imbalanced dataset.

Goodfellow et al. (2014) proposed the Generative adversarial networks (GAN), using two important components: generator (G) and discriminator (D) modules [31]. While the generator generates fake images, the role of the discriminator is to detect whether the fake images generated are actually fake images or real. In GAN, the generator and discriminator are rapidly competing with each other by following the min-max strategy as represented below.

$$min_G max_D\, V(D,G) = \mathrm{E}_{x\sim p_{data}(x)}[\log D(x)] + \mathrm{E}_{z\sim p_{z\,(z)}}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (1)$$

In equation (1), *D* is denoted as discriminator and *G* is represented as a generator; whereas *p* denotes the sample and *z* represents the noise components. Radford et al. **(**2015), further proposed a variant of GAN known as DCGAN based upon the same concept of generator and discriminator by using deep convolutional neural networks for generator and discriminator [24]. The generator and discriminator architectures used are illustrated in Figure 6 (a) and 6 (b) respectively.

The generator used in the experiment, uses the input vector of size 100 sampled from a normal distribution with zero mean and unit variance, and the output of the generator is of size

224 x 224 x 3. The discriminator will take input of size 224 x 224 x 3 and the output of the discriminator will be a single value of either 0 or 1 (based upon whether the image is detected fake or real).

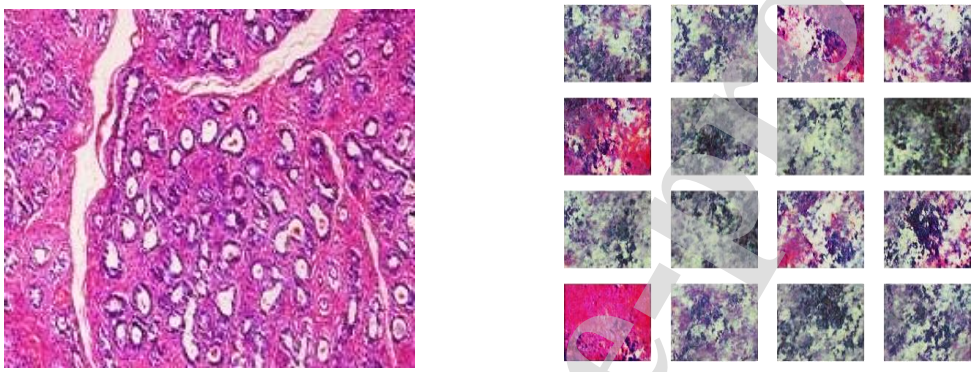
The network architecture of the generator and the discriminator used in DCGAN includes batch normalization layers that normalize the values during gradient propagation as well as forward pass. Further, this has been observed to be effective in tackling the class imbalance problem and to boost the performance of the classification model. However the same is lacking in the GAN [24] approach. So we have used DCGAN technique instead of GAN for data augmentation to synthetically generate the fake images in the proposed approach. Experiments have been conducted as illustrated in result section (Table 2) using Fréchet Inception Distance (FID) evaluation criterion for comparing GAN and DCGAN. The evaluation criterion FID is defined as:

$$\text{FID} = \| \mu_r - \mu_g \|^2 + \text{Tr}\left(\sum_r + \sum_g - 2\left(\sum_r \sum_g\right)^{1/2}\right) \quad (2)$$

Where μ_r is the mean of feature vector calculated from the real images, μ_g is mean of the feature vector calculated from the fake images, $\sum_r$ is the covariance of feature vector calculated from the real images, $\sum_g$ is the covariance of feature vector calculated from the fake images. FID is to calculate the quality of images generated and lower scores indicate that the generated image correlates well with the high quality real images. Mathematically, Fréchet distance is used to calculate the measure of similarity between curves on the basis of location and ordering of points. While calculating FID, we make use of the feature vector of both real and fake images.

The original imbalanced distribution from where the images have been sampled to train the various classifiers have imbalance factor of 2.30, 2.45, 2.46, 2.32 respectively in 40X, 100X, 200X and 400X Magnification factors. After augmenting the minority images in the training set

through DCGAN, the imbalance ratio obtained is 1:1 for the Benign and Malignant class populations, for all Magnification factors. Figure 7 shows some samples of the fake images generated, of size 224 x 224, corresponding to the original images of Benign class.



(a) (b)

Figure 7. (a) Original image sample from Benign class (b) Fake images samples generated for Benign class.

Since the pre-trained VGG16 network takes the input RGB image of size 224 x 224, so all the images including original and fake are presented to the network in the specific size of 224 x 224. Figure 8 displays the visually prominent localization features identifying the cancer cell from Benign and Malignant images.

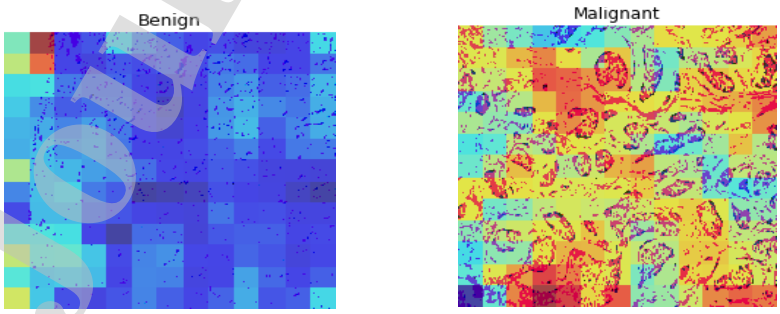


Figure 8. Activation map corresponding to the Benign and Malignant Images to illustrate the

Journal Pre-proof

Journal Pre-proof

prominent features to detect cancer cells.

3. Experimental Setup

3.1. Dataset

The BreakHis dataset consists of 7,909 histopathological images [25]. This dataset includes all the images from various categories such as: Adenosis (A), Fibroadenoma (F), Tubular Adenoma (TA), Phyllodes tumors (PT), Ductal Carcinoma (DC), Lobular carcinoma (LC), Mucinous carcinoma (MC) and papillary carcinoma (PC) respectively. All these categories of images are placed into two subcategories: Benign and Malignant. Benign consists of images from Adenosis (A), Fibroadenoma (F), Tubular Adenoma (TA), Phyllodes tumors (PT) categories. Malignant consists of images from Ductal Carcinoma (DC), Lobular carcinoma (LC), Mucinous carcinoma (MC) and papillary carcinoma (PC) categories. All the images of Benign and Malignant are of different magnification factors 40X, 100X, 200X and 400X and are resized to 224 x 224 x 3 using bilinear interpolation method to make the images suitable for the VGG16 pre-trained model. Table 1, indicates the disproportionate class distribution of samples in Benign and Malignant. Benign class consists of less number of samples so it is considered as the Minority class and Malignant consists of more number of samples in comparison to Benign so it is termed as the Majority class. Further, the complete dataset is split into train and test sets. The test set contains 100 samples each from the Malignant and Benign categories each Magnification factor. Lower number of samples of Benign (Minority class) indicates the class imbalance problem. Therefore for conducting the experimental task, equal number of test samples are chosen from both Benign and Malignant classes so as to avoid a bias in the testing phase [32].

Journal Pre-proof

Journal Pre-proof

Table 1: Class distribution of Benign and Malignant sample images with respect to the different Magnification factors

	Magnification Factor			
Class	40X	100X	200X	400X
Benign	625	644	623	588
Malignant	1,370	1,437	1,390	1,232
Images: 7,909	1,995	2,081	2,013	1,820

3.2. Implementation details

We have conducted all our experiments on GCP (Google Cloud Platform) virtual machine with 8 GB RAM, dual-core Intel Xeon processor and a single NVIDIA Tesla K80 GPU that has 12 GB DDR5 virtual memory. For all deep learning related model training, we use the Keras v2.2.3 deep learning framework with TensorFlow v1.15.3 backend. Additionally, we have used support vector classifier (SVM) from Scikit-learn. For training the deep network, we make use of the standard cross entropy loss (log likelihood) without any form of L2 regularization. In case of DCGAN, ReLU activation function is applied along with the Tanh activation function for generator and LeakyReLU activation function is used for all the layers of the discriminator, along with that, sigmoid activation function is used in the last layer of discriminator.

While training all the deep learning networks, Stochastic Gradient Descent (SGD) with momentum optimizer has been used to bring stability to the training. In all our experiments, the Scholastic gradient descent with momentum (using SGDR learning rate schedule) optimizer has been implemented, which also has overall improved the consistency of the learning curve for the last few set of epochs. According to Loshchilov et al. (2017) [38], in case of SGDR optimizer, it is recommended to start with a smaller cycle weight and increase the cycle steps linearly when scaling up. As SGDR performs warm restarts after cyclic intervals, it ensures that higher learning

Journal Pre-proof

rate would always be able to push the gradient in the direction away from the local optima. SGD approach uses cosine decay annealing to decay the learning rate after interval epochs and performs warm restart after every iteration cycle. Hence, we executed all our deep learning experiments starting with a default value of T_0= 1, T_{mul}=2, for the cosine decay learning rate schedule. As per the experiments performed in SGDR paper, a learning rate of 0.05 adapts well to the CIFAR-10 dataset. Initially for all of the deep learning experiments, we have used the initial learning in the experiments by setting the value to 0.05 which is tuned later to further improve the performance of our approach using grid search strategy. All our models are trained with the net budget of 100 epoch with the exception of ResNet-50, Inception v3 and VGG16 model which was trained till 50 epochs as they couldn't converge significantly even after 50 epochs. The choice of SGDR optimizer for our experiments helps us to determine a recommendation for various hyperparameters without the need of separate validation dataset, as emphasized in the original SGDR work [38].

We next state the experimental settings for various state-of-the-art networks and learning models used for comparison. The conventional Bag of visual words (BOVW) [33] used for the comparison uses the SIFT (scale-invariant feature transform) features along with K-Means clustering (number of clusters set to 10) and SVM classifier (kernel used in classifier is 'linear', C is set as 1.0 and gamma is set as '1/number of features'). For the conventional CNN used as the comparison method, a five layers architecture is used which consists of repeated blocks of Convolution with Batch normalization and Relu activation and pooling layers along with two fully connected layers. In addition, dropout with the drop rate of 0.2 have been used applied after the first fully connected layer. For all of our experiments we use the ImageDataGenerator from Keras framework with the batch size set to 128. In case of pre-trained convolutional neural

Journal Pre-proof

networks, Inception V3 [34], VGG16 [26] and ResNet [35], we use fixed pre-trained features and the terminal dense layer (classifier) was trained. The experiments conducted related to deep features extracted from pre-trained networks were done in with SVM classifiers. In the case of SVM classifier [36], both linear and RBF kernels were used while conducting the experiments.

3.3. Evaluation Measures

For the performance evaluation of the various classification models, two important criteria are considered while dealing with the imbalanced dataset: accuracy and F1 score; defined by Eq (2) to Eq (6); both the measures are calculated with the help of confusion matrix [37] which will help to measure the overall performance of the model for four different magnification factors 40x, 100x, 200x and 400x respectively. Accuracy helps us to determine how well model is able to distinguish the Benign samples from Malignant. F1 score is formed from the measure of precision and recall. However it is advisable that accuracy is not only considered as the evaluation measure while dealing with an class imbalance. F1 score must be considered in order to truly analyze the performance of the model while dealing with the imbalanced dataset. Along with accuracy and F1 score other evaluation criteria were considered to measure the performance of imbalanced dataset such as Receiver Operating Characteristics (ROC) analysis, Area Under the curve (AUC), Mathews correlation, and Cohen's kappa coefficient. Mathews correlation is used to find the correlation that occurs between actual and target variables for binary classification. Moreover, the Cohen Kappa is used to measure the agreement that occurs between raters. Both the measure are used specifically while dealing with imbalance problem, and are defined in Eq (7) and Eq (8) respectively.

$$Accuracy = \frac{true\ po.\ + true\ neg.}{true\ pos.+\ true\ neg.+\ false\ pos.+\ false\ neg.} \quad (3)$$

$$Precison\ = \frac{true\ pos.}{true\ pos.+\ false\ pos.} \quad (4)$$

$$Recall\ = \frac{true\ pos.}{true\ pos.+\ false\ neg.} \quad (5)$$

$$F1\ score = \frac{2*(Precison*Precison)}{(Precison+Precison)} \quad (6)$$

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)\ (TN+FP)(TN+FN)}} \quad (7)$$

$$Cohen\ kappa = \frac{p_0-p_c}{1-p_0} \quad (8)$$

4. Results And Discussionss

From experimental results, it was found that the proposed deep transfer learning-based network using DCGAN as data augmentation approach works significantly well in comparison to other conventional approaches in case of 40X, 100X, 200X and 400X Magnification factors, as illustrated in the results compiled in Table 3. We have compared the proposed approach with the other state-of-the-art models: (i) Bag of visual words (BOVW) (ii) CNN network trained from scratch (iii) pre-trained models (ResNet-50, Inception V3 and VGG16) (iv) deep feature extracted from pre-trained models, and trained via machine learning models. Various deep features and classifier combinations tested are: Inception V3 + SVM (with RBF and linear kernel); VGG16 + SVM (with RBF and linear kernel) and ResNet-50 +SVM (with RBF and linear kernel).

The evaluation was done on the basis of various evaluation criteria: Accuracy, F1 score, Mathews correlation coefficient, Cohen kappa coefficient and ROC with AUC. The pre-trained

Journal Pre-proof

networks (Inception V3, ResNet, and VGG16) are not working effectively in comparison to other conventional approaches, since the results are biased towards either the majority or minority class. Features extracted from pre-trained networks: Inception V3, VGG16 and ResNet-50, when given to SVM classifier (having kernel value set to 'rbf' and 'linear') has shown improvement in the performance of the result in comparison to the pre-trained networks. The BOVW technique, is also able to detect the minority class and majority class separately but still they are not able to show significant performance. CNN trained from scratch is also working well in comparison to the pre-trained networks but they have not shown remarkable performance in comparison to the proposed network in conjunction with SVM classifier, in case of 40X and 200X magnification factors. But in the case of 100X and 400X magnification factors, CNN has shown improvement over few deep features and classifier combinations as shown in Table 3 (i) and (ii).

For our experiments, we make use of the mean normalization technique which helps to rescale intensity of each pixel and standardize them to a zero mean and unit variance normal distribution before it is input to the network. This preprocessing step was particularly useful when the CNN is trained from scratch and for it we have used RGB normalization of the BreakHis dataset. We have also found that VGG16, ResNet50 networks make use of the same preprocessing technique with an exception that it uses mean and standard deviation from the ImageNet dataset as we use these network with pre-trained weights. The Inception-style preprocessing aims to rescale image pixels to a range [-1, 1] without performing any standardization of the RGB values. As an additional step, we also tried to use the Inception-style preprocessing technique for CNN trained from scratch but with it results were seen to reduce significantly following which we decided to move forward with the mean normalization step for

Journal Pre-proof

all forms of pre-processing. Experiments reflect that for this imbalanced breast cancer dataset the mean normalization step is better suited than rescaling (Inception-style [-1, +1] feature scaling) due to improved generalization with RGB mean subtraction. Our approach is based on pre-trained VGG16 network which also uses RGB normalization based preprocessing.

The proposed deep transfer network developed by considering DCGAN and Batch normalization with VGG16 layers till block 4 pool layer along with the Convolutional layer, Batch normalization, Global Average Pooling 2D, dropout, and dense layers, is working well in comparison to other approaches as shown in Table 3. Before selecting the appropriate combination, we have performed experiments with the different combinations to analyze the effect of DCGAN synthetic sample generation and Batch normalization on the proposed transfer network. Initially, the combination of layers were tried (i) without batch normalization (iii) secondly without batch normalization and with the DCGAN samples, it was found that the results were biased towards the minority or majority class. But application of batch normalization with the DCGAN samples leads to overall improvement in the performance of the network as shown in Table 3 (i) and (ii). The effect of the DCGAN, and the effect of covariance shift by use of batch normalization layer on the proposed deep transfer network has been analyzed experimentally. Another set of experiments are performed to find the appropriate deep features for the proposed approach. Experimental study is performed to extract the appropriate features from VGG16 by analyzing and comparing the deep features from different blocks of VGG16 network as shown in Table 4 (i) and (ii) for different magnification factors. From the analysis, it was found that extraction of layers till block 4 layer of VGG16 architecture is the appropriate choice to be used in the proposed network in comparison to other layers. So we can conclude that the proposed deep transfer network incorporating initial layers till block 4 pool

Journal Pre-proof

layer of VGG16 is deriving the relevant features. Further the tuning of hyperparameters experiments are conducted on the proposed approach using a grid search based strategy in order to find the best hyperparameter combination by changing the initial learning rate, T_{mul} and T_0 value of the SGDR optimizer as shown in case of 40X, 100X, 200x and 400X magnification factors in Table 5. In Table 6, performance evaluation of proposed approach for different combinations of SGDR hyperparameters are shown for 40X, 100X, 200X and 400X magnification factors, to find the best combination of hyperparameters.

Figure 9 shows the significance of SGDR optimizer in different networks. From the analysis it was found that proposed approach is more stable in comparison to other approaches using SGDR optimizer, which further leads to the improved convergence. Performance evaluation of the proposed approach with the help of ROC (Receiver operating curve) with AUC as been shown in Figure 10 that also depicts the comparison of the proposed approach with the other approaches. This proves that the proposed approach is the best choice in comparison to the other conventional approaches.

The performance improvement of our approach is primarily based on the use of DCGAN for synthetic Benign sample generation. We conducted experiments with GAN and DCGAN in an initial set of experiments. Based upon Fréchet Inception Distance (FID) evaluation criteria as illustrated in Table 2, the observation drawn from the analysis is that DCGAN produces higher quality samples in comparison to GAN. Fréchet Inception Distance (FID) score is the metric that has been used for the performances analysis of generative modelling. FID is used to calculate the difference between Inception features extracted from real images and fake images generated using Generative Adversarial Networks and its variants [44]. It is a useful metric indicator for quality is to noise analysis of synthetic samples where a lower FID usually indicates better

Journal Pre-proof

Journal Pre-proof

generation of fake samples.

Table 2: Performance analysis of GAN and DCGAN based upon FID evaluation criteria

Magnification factor	**No of Real Images**	**No of fake Images**	**FID of GAN**	**FID of DCGAN**
40X	522	718	386482.139	401.592
100X	544	793	165530.598	480.946
200X	523	767	122297.189	465.874
400X	488	644	143042.212	517.281

Results obtained from experiments infer that use of transfer learned weights on architectures of previously trained available pre-trained networks (which are already trained on large scale ImageNet dataset) helps improve the classification task. The use of transfer learning is also appropriate our case as we have comparatively much less number of samples present in the available biomedical image dataset. Training of CNN from scratch would require a large set of samples and it has to be ensured that the large number of parameters of the CNN be trained effectively with much generalization in order to yield good results. It is also experimentally observed that when the batch normalization layer is not included in the proposed network, the results get completely biased towards the majority class. The batch normalization layer scales intermediate input values of the VGG16 block 4 pool features resulting in effective training of the classifier. Without it, the classifier suffers from the exploding gradient problem and we incur NaN loss during training due to quite high log-likelihood value with the imbalanced class setting.

Application of SGDR optimizer has made the proposed network and other comparison approaches more stable as shown experimentally. Also, it helps to generalize well allowing faster

convergence. Experimental task has been performed by using the BreakHis dataset because it consists of images of different magnification factors and also consists of adequate image samples and is one of the few available datasets related to breast cancer. Further, extensive hyperparameter tuning have been conducted. From the analysis it was found that choice of the exact hyperparameters plays crucial role in overall learning and performance of the model. Using our proposed network, the best determined hyperparameters leads to the better generalization. The hyperparameter tuning of the proposed network have been performed using a simple grid search strategy and we choose test accuracy as the maximization objective of this search step. Due to use of the SGDR learning rate schedule, as per recommendation of Loschilov et al. (2017) [38] an extra validation would not be required.

In this study, we have also compared and analyzed the features extracted from the different VGG16 blocks for our proposed network in order to find relevant features. Experimental results indicate that the features extracted till block 4 pool proves to be quite suitable. So we have used features till block 4 pool layers only in the proposed work instead of till block 5 as there in the original architecture of VGG16. Features extracted till block 4 pool layers are considered bottleneck endpoint are hence considered optimal for our experiments. So, 14 x 14 x 512 features extracted from the VGG16 network is passed onto the higher-end layers, in the proposed deep transfer network.

Journal Pre-proof

Journal Pre-proof

Table 3 (i): Performance evaluation on 40X and 100X Magnification factors

Magnification Factor	40X					100X				
Approach	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa
		Benign	Malignant				Benign	Malignant		
BOVW [33]	0.65	0.49	0.73	0.3904	0.3	0.615	0.4	0.72	0.3266	0.2299
CNN [39]	0.67	0.51	0.75	0.4525	0.3399	0.88	0.87	0.89	0.7655	0.76
VGG16 [26]	0.5	0.0	0.67	0.0	0.0	0.5	0.67	0.00	0.0	0.0
InceptionV3 [34]	0.56	0.21	0.69	0.2526	0.12	0.50	0.06	0.66	0.0	0.0
ResNet50 [35]	0.53	0.11	0.68	0.1758	0.06	0.51	0.06	0.67	0.02	0.0714
VGG16 + Linear SVM [42]	0.855	0.84	0.87	0.7295	0.71	0.87	0.86	0.88	0.7473	0.74
InceptionV3 + Linear SVM [41]	0.86	0.85	0.87	0.7319	0.72	0.815	0.80	0.83	0.6353	0.63
ResNet50 + Linear SVM [43]	0.91	0.91	0.91	0.8226	0.82	0.91	0.91	0.91	0.8241	0.82
VGG16 + RBF SVM [42]	0.85	0.83	0.87	0.7249	0.7	0.855	0.84	0.87	0.7231	0.71
InceptionV3 + RBF SVM [41]	0.82	0.79	0.84	0.6627	0.64	0.795	0.75	0.83	0.6298	0.59
ResNet50 + RBF SVM [43]	0.89	0.88	0.90	0.7901	0.78	0.86	0.85	0.87	0.7319	0.72
Proposed Network (w/o Batch Normalization)	0.5	0.67	0.0	0.0	0.0	0.5	0.67	0.0	0.0	0.0
Proposed Network (w/o Batch Normalization and w/ DCGAN samples)	0.5	0.67	0.0	0.0	0.0	0.5	0.67	0.0	0.0	0.0
Proposed Network (w/ BatchNormalization and w/ DCGAN samples)	**0.935**	**0.93**	**0.94**	**0.8774**	**0.87**	**0.925**	**0.92**	**0.93**	**0.8551**	**0.85**
Proposed Network (w/ BatchNormalization, w/ DCGAN samples and w/ hyper parameter tuning)	**0.965**	**0.96**	**0.97**	**0.9304**	**0.9299**	**0.94**	**0.94**	**0.94**	**0.8815**	**0.88**

Journal Pre-proof

Journal Pre-proof

Table 3 (ii): Performance evaluation on 200X and 400X Magnification factors

Magnification Factor	200X					400X				
Approach	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa
		Benign	Malignant				Benign	Malignant		
BOVW[33]	0.56	0.24	0.69	0.2211	0.12	0.575	0.35	0.68	0.2072	0.15
CNN [39]	0.775	0.72	0.81	0.5972	0.55	0.885	0.88	0.89	0.7731	0.77
VGG16 [26]	0.5	0.67	0.0	0.0	0.0	0.5	0.0	0.67	0.0	0.0
InceptionV3 [33]	0.51	0.08	0.67	0.0586	0.02	0.5	0.0	0.67	0.0	0.0
ResNet50 [35]	0.5	0.0	0.67	0.0	0.0	0.5	0.0	0.67	0.0	0.0
VGG16 + Linear SVM [42]	0.835	0.82	0.85	0.6852	0.6699	0.835	0.81	0.85	0.6958	0.6699
InceptionV3 + Linear SVM [41]	0.795	0.77	0.81	0.6034	0.59	0.82	0.80	0.84	0.656	0.64
ResNet50 + Linear SVM [43]	0.93	0.93	0.93	0.86	0.8627	0.915	0.91	0.92	0.835	0.83
VGG16 + RBF SVM [42]	0.84	0.81	0.86	0.7128	0.6799	0.82	0.78	0.85	0.6805	0.64
InceptionV3 + RBF SVM [41]	0.76	0.7	0.8	0.5729	0.52	0.795	0.75	0.83	0.6298	0.59
ResNet50 + RBF SVM [43]	0.86	0.84	0.88	0.75	0.72	0.885	0.87	0.89	0.7813	0.77
Proposed Network (w/o Batch Normalization)	0.5	0.67	0.0	0.0	0.0	0.5	0.67	0.0	0.0	0.0
Proposed Network (w/o Batch Normalization and w/ DCGAN samples)	0.5	0.67	0.0	0.0	0.0	0.5	0.67	0.00	0.0	0.0
Proposed Network (w/ BatchNormalization and w/ DCGAN samples)	**0.955**	**0.95**	**0.96**	**0.91**	**0.91**	**0.915**	**0.91**	**0.92**	**0.832**	**0.83**
Proposed Network (w/ BatchNormalization, w/ DCGAN samples and w/ hyper parameter tuning)	**0.955**	**0.95**	**0.96**	**0.9111**	**0.91**	**0.93**	**0.93**	**0.93**	**0.8627**	**0.86**

Journal Pre-proof

Journal Pre-proof

Table 4 (i): Performance analysis of different layers of proposed network on 40X, 100X Magnification factors

Magnification Factor	40X					100X				
Approach	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa
		Benign	Malignant				Benign	Malignant		
Proposed Network (using block3_pool layer)	0.895	0.89	0.90	0.799	0.79	0.91	0.90	0.92	0.8307	0.82
Proposed Network (using block4_pool layer)	**0.935**	**0.93**	**0.94**	**0.8774**	**0.87**	**0.925**	**0.92**	**0.93**	**0.8551**	**0.85**
Proposed Network (using block5_pool layer)	0.855	0.84	0.87	0.7231	0.71	0.87	0.86	0.88	0.7552	0.74

Table 4 (ii): Performance analysis of different layers of proposed network on 200X, 400X Magnification factors

Magnification Factor	200X					400X				
Approach	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa	Accuracy	F1 score		Matthews Correlation Coefficient	Cohen's Kappa
		Benign	Malignant				Benign	Malignant		
Proposed Network (using block3_pool layer)	0.93	0.93	0.93	0.8643	0.86	**0.935**	**0.93**	**0.94**	**0.8753**	**0.87**
Proposed Network (using block4_pool layer)	**0.955**	**0.95**	**0.96**	**0.91**	**0.91**	0.915	0.91	0.92	0.832	0.83
Proposed Network (using block5_pool layer)	0.865	0.85	0.88	0.7435	0.73	0.855	0.84	0.86	0.716	0.71

Table 5: Performance analysis of best values of hyperparameters obtained in 40X, 100X, 200X and 400X magnification factors

Magnification Factors	Accuracy	Initial Learning Rate	SGDR T_0	SGDR T_{mul}
40X	**0.965**	0.01	1	2
100X	**0.94**	0.1	10	2
200X	**0.955**	0.005	10	2
400X	**0.93**	0.01	10	2

Journal Pre-proof

Journal Pre-proof

Table 6: Performance evaluation of proposed approach on various SGDR hyperparameters
(i) 40X (ii) 100X (iii) 200X (iv)400X (top to bottom)

40X

Hyperparameters	Evaluation Metric	SGDR (T_0=1, T_{mul}=2)	SGDR (T_0=10, T_{mul}=2)	SGDR (T_0=50, T_{mul}=1)
LR=0.001	Accuracy	0.91	0.905	0.91
	Matthews Correlation Coefficient	0.8281	0.8192	0.8281
	Cohen's Kappa	0.82	0.81	0.82
LR=0.005	Accuracy	0.945	0.935	0.94
	Matthews Correlation Coefficient	0.8921	0.8735	0.8828
	Cohen's Kappa	0.89	0.87	0.88
LR=0.01	Accuracy	**0.965**	0.95	0.955
	Matthews Correlation Coefficient	**0.9304**	0.9007	0.9111
	Cohen's Kappa	**0.9299**	0.9	0.91
LR=0.05	Accuracy	0.935	0.955	0.96
	Matthews Correlation Coefficient	0.8774	0.91	0.9201
	Cohen's Kappa	0.87	0.91	0.92
LR=0.1	Accuracy	0.96	0.95	0.96
	Matthews Correlation Coefficient	0.92	0.9028	0.9201
	Cohen's Kappa	0.92	0.9	0.92

100X

Hyperparameters	Evaluation Metric	SGDR (T_0=1, T_{mul}=2)	SGDR (T_0=10, T_{mul}=2)	SGDR (T_0=50, T_{mul}=1)
LR=0.001	Accuracy	0.935	0.935	0.93
	Matthews Correlation Coefficient	0.8753	0.8735	0.8643
	Cohen's Kappa	0.87	0.87	0.86
LR=0.005	Accuracy	0.935	0.935	0.93
	Matthews Correlation Coefficient	0.8721	0.8735	0.8627
	Cohen's Kappa	0.87	0.87	0.86
LR=0.01	Accuracy	0.925	0.93	0.93
	Matthews Correlation Coefficient	0.8551	0.8643	0.8627
	Cohen's Kappa	0.85	0.86	0.86
LR=0.05	Accuracy	0.93	0.94	0.925
	Matthews Correlation Coefficient	0.8627	0.8815	0.8534
	Cohen's Kappa	0.86	0.88	0.85
LR=0.1	Accuracy	0.93	**0.94**	0.925
	Matthews Correlation Coefficient	0.8627	**0.8815**	0.8534
	Cohen's Kappa	0.86	**0.88**	0.85

Journal Pre-proof

Journal Pre-proof

200X

Hyperparameters	Evaluation Metric	SGDR (T_0=1, T_{mul}=2)	SGDR (T_0=10, T_{mul}=2)	SGDR (T_0=50, T_{mul}=1)
LR=0.001	Accuracy	0.92	0.93	0.92
	Matthews Correlation Coefficient	0.8442	0.8662	0.8461
	Cohen's Kappa	0.84	0.86	0.84
LR=0.005	Accuracy	0.955	**0.955**	0.945
	Matthews Correlation Coefficient	0.9104	**0.9111**	0.8921
	Cohen's Kappa	0.91	**0.91**	0.89
LR=0.01	Accuracy	0.945	0.94	0.955
	Matthews Correlation Coefficient	0.8936	0.8815	0.9111
	Cohen's Kappa	0.89	0.88	0.91
LR=0.05	Accuracy	0.955	0.935	0.955
	Matthews Correlation Coefficient	0.91	0.8735	0.9111
	Cohen's Kappa	0.91	0.87	0.91
LR=0.1	Accuracy	0.895	0.95	0.95
	Matthews Correlation Coefficient	0.8046	0.9016	0.9007
	Cohen's Kappa	0.79	0.9	0.9

400X

Hyperparameters	Evaluation Metric	SGDR (T_0=1, T_{mul}=2)	SGDR (T_0=10, T_{mul}=2)	SGDR (T_0=50, T_{mul}=1)
LR=0.001	Accuracy	0.895	0.895	0.905
	Matthews Correlation Coefficient	0.7967	0.7967	0.8169
	Cohen's Kappa	0.79	0.79	0.81
LR=0.005	Accuracy	0.925	0.925	0.92
	Matthews Correlation Coefficient	0.8520	0.8534	0.8442
	Cohen's Kappa	0.85	**0.85**	0.84
LR=0.01	Accuracy	0.925	**0.93**	0.925
	Matthews Correlation Coefficient	0.8534	**0.8627**	0.8520
	Cohen's Kappa	0.85	**0.86**	0.85
LR=0.05	Accuracy	0.915	0.925	0.92
	Matthews Correlation Coefficient	0.832	0.8534	0.8427
	Cohen's Kappa	0.83	0.85	0.84
LR=0.1	Accuracy	0.895	0.92	0.925
	Matthews Correlation Coefficient	0.8046	0.8415	0.852
	Cohen's Kappa	0.79	0.84	0.85

Journal Pre-proof

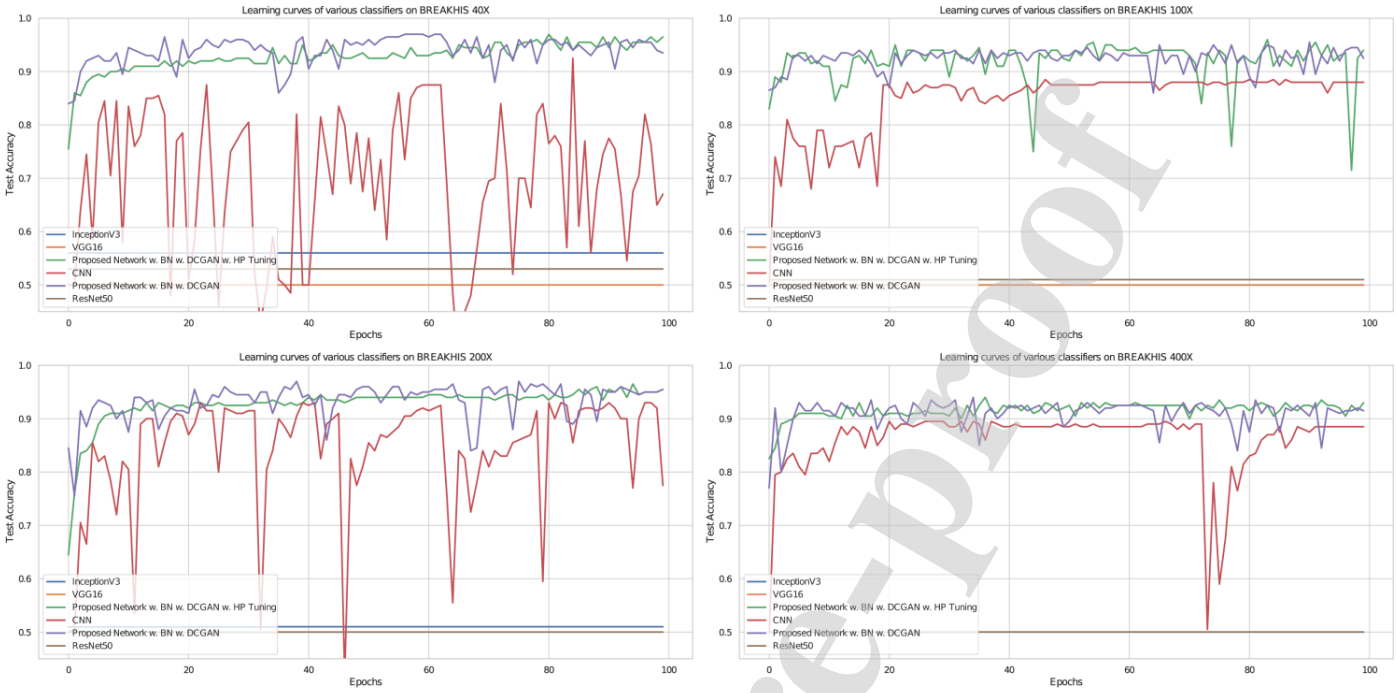
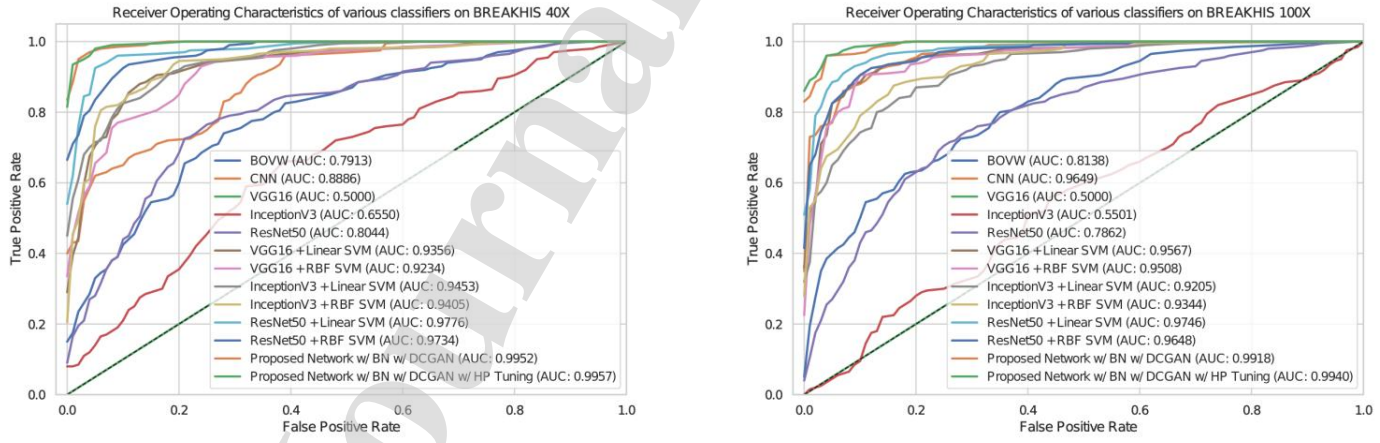


Figure 9. Learning curve depicting test accuracy across epochs and demonstrating better anytime performance with proposed approach in comparison to other approaches



Journal Pre-proof

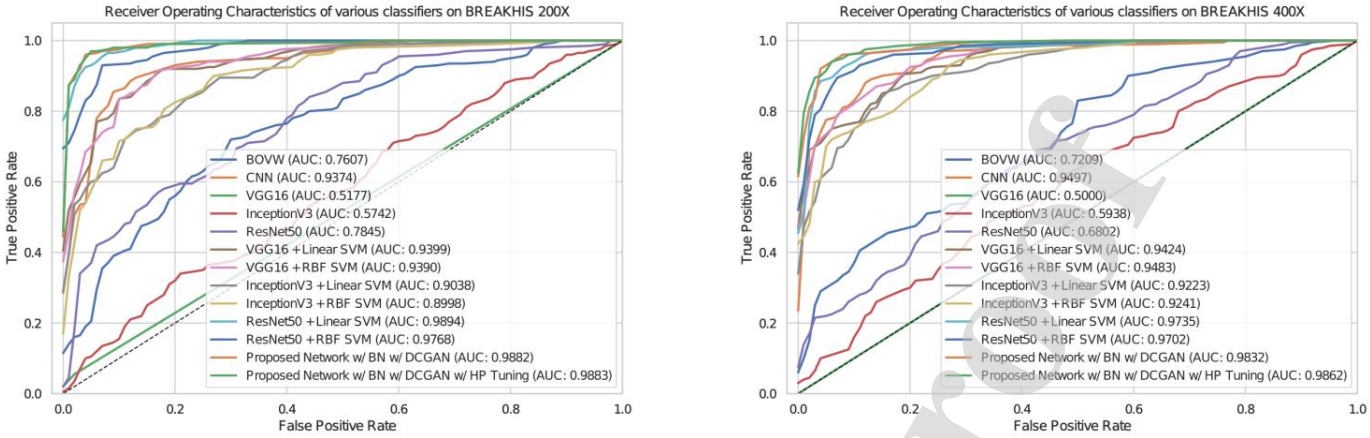


Figure 10. Comparison of proposed approach with other approaches using ROC (Receiver operating Curve)

5. Conclusion

We have proposed a novel deep transfer learning based approach for the classification of an Imbalanced breast cancer dataset. In this paper, we have also explored the effect of DCGAN and the effect of Batch normalization on the performance of the proposed transfer network architecture. The proposed network contains the VGG16 pre-trained model layers till block 4 pool layers along with the Batch Normalization, convolutional (CONV2D), Global Average Pooling 2D layer, dropout, and dense layers. It would help in the accurate detection of cancer cells at an early stage.

The limitation of the proposed work is that applying DCGAN at data level is directly dependent on number of samples of minority class as we employ deep generative modelling for enhancing the performance of the classifier. For scenarios with much less number of minority samples, the DCGAN training distribution would not be able to generalize well and would fail to generate high quality samples resulting in sub-optimal performance. Despite of the limitations, the

Journal Pre-proof

proposed methodology is able to tackle the class-imbalance problem that adversely affects many state-of-the-art deep learning networks as well as another traditional machine learning techniques. The proposed architecture is able to learn fine grained features on top of ImageNet pre-trained deep features, specific to the biomedical datasets. Total number of network parameters is reduced by introducing the global average pooling layer instead of flatten in the original VGG16 architecture. This also helps to reduce the number of FLOPS. The proposed approach is able to work well even when the microscopic images are very different at data level as in the BreakHis dataset that contains images of four different magnification factors. Our focus was on the combination of pre-trained network with CNN. We have included VGG16 pre-trained network since it is a popular pre-trained network with its ability to work well with images of different scales and aspect ratios. In future, we shall be exploring few advanced pre-trained network architectures for incorporating into our work. Also, we will explore other variants of GAN models as a data augmentation approach to improve the overall performance of our deep transfer network and also test it on other biomedical datasets.

References

1. Parkin, D. Max, Freddie Bray, J. Ferlay, and Paola Pisani. "Global cancer statistics, 2002." CA: a cancer journal for clinicians 55, no. 2 (2005): 74- 108.
2. Can't Veer, Laura J., Hongyue Dai, Marc J. Van De Vijver, Yudong D. He, Augustinus AM Hart, Mao Mao, Hans L. Peterse et al. "Gene expression profiling predicts clinical outcome of breast cancer." nature 415, no. 6871 (2002): 530.
3. Xu, Jun, Lei Xiang, Qingshan Liu, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi. "Stacked sparse autoencoder (SSAE) for nuclei detection on breast

Journal Pre-proof

cancer histopathology images." IEEE transactions on medical imaging 35, no. 1 (2015): 119-130.

4. Belsare, A. D., and M. M. Mushrif. "Histopathological image analysis using image processing techniques: An overview." Signal & Image Processing 3, no. 4 (2012): 23.
5. Burt, Jeremy R., Neslisah Torosdagli, Naji Khosravan, Harish RaviPrakash, Aliasghar Mortazi, Fiona Tissavirasingham, Sarfaraz Hussein, and Ulas Bagci. "Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks." The British journal of radiology 91, no. 1089 (2018): 20170545.
6. Lemley, Joe, Shabab Bazrafkan, and Peter Corcoran. "Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision." IEEE Consumer Electronics Magazine 6, no. 2 (2017): 48-56.
7. Tai, Cheng, Tong Xiao, Yi Zhang, and Xiaogang Wang. "Convolutional neural networks with low-rank regularization." arXiv preprint arXiv:1511.06067 (2015).
8. Shin, Hoo-Chang, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers. "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning." IEEE transactions on medical imaging 35, no. 5 (2016): 1285-1298.
9. Spanhol, Fabio A., Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. "A dataset for breast cancer histopathological image classification." IEEE Transactions on Biomedical Engineering 63, no. 7 (2015): 1455-1462.
10. Bardou, Dalal, Kun Zhang, and Sayed Mohammad Ahmad. "Classification of breast cancer based on histology images using convolutional neural networks." IEEE Access 6 (2018): 24680-24693.

Journal Pre-proof

Journal Pre-proof

11. Bayramoglu, Neslihan, Juho Kannala, and Janne Heikkilä. "Deep learning for magnification independent breast cancer histopathology image classification." In 2016 23rd International conference on pattern recognition (ICPR), pp. 2440-2445. IEEE, 2016.
12. Litjens, Geert, Clara I. Sánchez, Nadya Timofeeva, Meyke Hermsen, Iris Nagtegaal, Iringo Kovacs, Christina Hulsbergen-Van De Kaa, Peter Bult, Bram Van Ginneken, and Jeroen Van Der Laak. "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis." Scientific reports 6 (2016): 26286.
13. Xie, Juanying, Ran Liu, I. V. Joseph Luttrell, and Chaoyang Zhang. "Deep Learning Based Analysis of Histopathological Images of Breast Cancer." Frontiers in genetics 10 (2019).
14. Deniz, Erkan, Abdulkadir Şengür, Zehra Kadiroğlu, Yanhui Guo, Varun Bajaj, and Ümit Budak. "Transfer learning based histopathologic image classification for breast cancer detection." Health information science and systems 6, no. 1 (2018): 18.
15. Garud, Hrushikesh, Sri Phani Krishna Karri, Debdoot Sheet, Jyotirmoy Chatterjee, Manjunatha Mahadevappa, Ajoy K. Ray, Arindam Ghosh, and Ashok K. Maity. "High-magnification multi-views based classification of breast fine needle aspiration cytology cell samples using fusion of decisions from deep convolutional networks." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 76-81. 2017.
16. Antoniou, Antreas, Amos Storkey, and Harrison Edwards. "Data augmentation generative adversarial networks." arXiv preprint arXiv:1711.04340 (2017).
17. Saini, Manisha, and Seba Susan. "Comparison of Deep Learning, Data Augmentation and Bag of-Visual-Words for Classification of Imbalanced Image Datasets." In International Conference on Recent Trends in Image Processing and Pattern Recognition, pp. 561-571.

Journal Pre-proof

Journal Pre-proof

Springer, Singapore, 2018.

18. Susan, Seba, and Amitesh Kumar. "SSOMaj-SMOTE-SSOMin: Three- step intelligent pruning of majority and minority samples for learning from imbalanced datasets." Applied Soft Computing 78 (2019): 141-149.
19. Susan, Seba, and Amitesh Kumar. "Hybrid of Intelligent Minority Oversampling and PSOBased Intelligent Majority Undersampling for Learning from Imbalanced Datasets." In International Conference on Intelligent Systems Design and Applications, pp. 760-769. Springer, Cham, 2018.
20. Susan, Seba, and Amitesh Kumar. "Learning Data Space Transformation Matrix from Pruned Imbalanced Datasets for Nearest Neighbor Classification." In 2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), pp. 2831-2838. IEEE, 2019.
21. Lyu, Qiongshuai, Min Guo, and Zhao Pei. "DeGAN: Mixed noise removal via generative adversarial networks." Applied Soft Computing (2020): 106478.
22. Frid-Adar, Maayan, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification." Neurocomputing321 (2018): 321-331.
23. Saini, Manisha, and Seba Susan. "Data Augmentation of Minority Class with Transfer Learning for Classification of Imbalanced Breast Cancer Dataset Using Inception-V3." In Iberian Conference on Pattern Recognition and Image Analysis, pp. 409-420. Springer, Cham, 2019.
24. Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning

Journal Pre-proof

with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434 (2015).

25. Databases – Laboratório Visão Robótica e Imagem. https://web.inf.ufpr.br/vri/databases/. Accessed: 28-11-2019.
26. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for largescale image recognition." arXiv preprint arXiv:1409.1556 (2014).
27. Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167 (2015).
28. Cui, Yin, Feng Zhou, Jiang Wang, Xiao Liu, Yuanqing Lin, and Serge Belongie. "Kernel pooling for convolutional neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921-2930. 2017.
29. Yosinski, Jason, Jeff Clune, Yoshua Bengio, and Hod Lipson. "How transferable are features in deep neural networks?." In Advances in neural information processing systems, pp. 3320-3328. 2014
30. Keras-GAN/dcgan at master · eriklindernoren/Keras-GAN. https://github.com/eriklindernoren/Keras-GAN/tree/master/dcgan. Accessed: 28-11-2019.
31. Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial nets." In Advances in neural information processing systems, pp. 2672-2680. 2014.
32. Brodersen, Kay Henning, Cheng Soon Ong, Klaas Enno Stephan, and Joachim M. Buhmann. "The balanced accuracy and its posterior distribution." In 2010 20th International Conference on Pattern Recognition, pp. 3121-3124. IEEE, 2010.
33. Suh, Hyun K., Jan Willem Hofstee, Joris IJsselmuiden, and Eldert J. van Henten. "Sugar

Journal Pre-proof

beet and volunteer potato classification using Bag-of- Visual-Words model, Scale-Invariant Feature Transform, or Speeded Up Robust Feature descriptors and crop row information." Biosystems engineering 166 (2018): 210-226

34. Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. 2016.
35. Wang, Fei, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang. "Residual attention network for image classification." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3156-3164. 2017.
36. Vapnik, Vladimir, Isabel Guyon, and Trevor Hastie. "Support vector machines." Mach. Learn 20, no. 3 (1995): 273-297. Camps-Valls, Gustavo, Luis Gomez-Chova, Jordi Muñoz-Marí, Joan Vila-Francés, and Javier Calpe-Maravilla. "Composite kernels for hyperspectral image classification." IEEE geoscience and remote sensing letters 3, no. 1 (2006): 93-97.
37. Alom, Md Zahangir, Mahmudul Hasan, Chris Yakopcic, and Tarek M. Taha. "Inception recurrent convolutional neural network for object recognition." arXiv preprint arXiv:1704.07709 (2017).
38. Loshchilov, Ilya, and Frank Hutter. "Sgdr: Stochastic gradient descent with warm restarts." arXiv preprint arXiv:1608.03983 (2016).
39. Araújo, Teresa, Guilherme Aresta, Eduardo Castro, José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, and Aurélio Campilho. "Classification of breast cancer histology images using convolutional neural networks." PloS one 12, no. 6 (2017): e0177544.
40. Perdana, Anugrah Bintang, and Adhi Prahara. "Face Recognition Using Light-Convolutional

Journal Pre-proof

Neural Networks Based On Modified Vgg16 Model." In 2019 International Conference of Computer Science and Information Technology (ICoSNIKOM), pp. 1-4. IEEE, 2019.

41. Li, Chen, Hao Chen, Le Zhang, Ning Xu, Dan Xue, Zhijie Hu, He Ma, and Hongzan Sun. "Cervical Histopathology Image Classification Using Multilayer Hidden Conditional Random Fields and Weakly Supervised Learning." IEEE Access 7 (2019): 90378-90397.
42. Vijayalakshmi, A. "Deep learning approach to detect malaria from microscopic images." Multimedia Tools and Applications(2019): 1-21.
43. Wu, Chenchen, Jun Ruan, Guanglu Ye, Jingfan Zhou, Simin He, Jianlian Wang, Zhikui Zhu, Junqiu Yue, and Yanggeling Zhang. "Identifying Tumor in Whole-Slide Images of Breast Cancer Using Transfer Learning and Adaptive Sampling." In 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI), pp. 167-172. IEEE, 2019.
44. Shmelkov, Konstantin, Cordelia Schmid, and Karteek Alahari. "How good is my GAN?." In Proceedings of the European Conference on Computer Vision (ECCV), pp. 213-229. 2018.
45. Liu, Yun, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, and Xiang Bai. "Richer convolutional features for edge detection." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3000-3009. 2017.
46. Kumar, Abhinav, Sanjay Kumar Singh, Sonal Saxena, K. Lakshmanan, Arun Kumar Sangaiah, Himanshu Chauhan, Sameer Shrivastava, and Raj Kumar Singh. "Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer." Information Sciences 508 (2020): 405-421.
47. Aravind, Krishnaswamy R., Purushothaman Raja, Rajendran Ashiwin, and Konnaiyar V. Mukesh. "Disease classification in Solanum melongena using deep learning." Spanish

Journal Pre-proof

Journal of Agricultural Research 17, no. 3 (2019): 0204.

48. Baheti, Bhakti, Suhas Gajre, and Sanjay Talbar. "Detection of distracted driver using convolutional neural network." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1032-1038. 2018.

Manisha Saini is currently pursuing Ph.D. in Computer science and Engineering Department from Delhi Technological University, Delhi, India and also working as Assistant Professor in the Department of Computer Science and Engineering, GD Goenka University, Gurgaon, Haryana, India. Her area of interests includes Computer Vision, Neural Networks, Machine Learning and Deep Learning.

Dr. Seba Susan is an Associate Professor in Information Technology Department, Delhi Technological University, Delhi, India. She completed her Ph.D. from Electrical Engineering department of the Indian Institute of Technology (IIT) Delhi, in 2014. Her current research interests includes Statistical inferencing and soft computing tools for Computer Vision, Natural Language Processing, Data Mining and Machine Learning.

Journal Pre-proof

***Credit Author Statement**

Journal Pre-proof

Manisha Saini: Software, Conceptualization, Methodology, Validation, Investigation, Resources, Writing- Original draft preparation.

Seba Susan.: Conceptualization, Methodology, Investigation, Writing - Review & Editing, Supervision.

Journal Pre-proof

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 8, 2020, accepted September 30, 2020, date of publication October 12, 2020, date of current version October 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3030020

Diode Based Reconfigurable Microwave Filters for Cognitive Radio Applications: A Review

HASHINUR ISLAM[1], (Graduate Student Member, IEEE), SAUMYA DAS[1], TANUSHREE BOSE[1], AND TANWEER ALI[2], (Senior Member, IEEE)

[1]Department of Electronics and Communication Engineering, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Gangtok 737136, India
[2]Department of Electronics and Communication, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India

Corresponding author: Tanweer Ali (tanweer.ali@manipal.edu)

This work was supported by the T. M. A. Pai Research Grant, Sikkim Manipal University, India, under Grant 118/SMU/REG/UOO/104/2019.

ABSTRACT The cognitive radio paradigm for developing next-century wireless communication systems is rapidly entering the mainstream, and various aspects of it are currently being applied in 5G technology, aeronautical engineering, military communications, emergency, and public safety applications, satellite communication, and healthcare. Cognitive radio focuses on the existence of software defined radio architectures that allow dynamic reconfiguration. Many researchers have taken initiatives in the last decade to achieve the reconfiguration ability in cognitive radio systems to support the concept of dynamic spectrum access. As cognitive radio adapts dynamic spectrum allocation for its users, the physical implementation requires reconfigurable filters that can alter the carrier frequencies and bandwidth. Although there are many ways to reconfigure filter operation, diode based reconfiguration has received utmost attention among researchers because of its shorter response delay and easy implementation. In the last decade, researchers have reported several diode-based reconfigurable filters, including their characteristics such as filter function, filter combination, tuning range, variation in bandwidth, isolation, and resonance. However, to examine the potential of these filters in the application of cognitive radio, a comprehensive review needs to be pursued. In this review article, the descriptions of several diode based reconfigurable filters are illustrated with their exhibiting characteristics. The detailed information provided in this article has disclosed that primarily three different types diode based reconfigurable filters have been reported by researchers: Tunable, Switchable, and Hybrid (Both Tunable and Switchable). It is also found that each type of reconfiguration can further be segregated in terms of filter function, centre frequency variation, and bandwidth variation. The detailed categorization of the reconfiguration presented in this paper provides a systematic approach to select the correct reconfigurable filter for the desired frequency reconfiguration in cognitive radio.

INDEX TERMS Cognitive radio, reconfigurable filter, tunable filter, switchable filter, hybrid reconfigurable filter.

I. INTRODUCTION

Cognitive Radio (CR) is considered to have tremendous potential to alleviate the spectrum scarcity problem in wireless communication [1]–[3]. It exploits underused spectrums to find a new way of spectrum access. CR systems are capable of reconfiguring themselves depending on spectrum availability, traffic congestion, channel interference, etc. The intelligence of CR systems is widely being explored by 5G technology [4], [5], military communication [6], healthcare [7], aeronautical applications [8], RADAR [9], emergency and public safety applications [10], and satellite communication [11], as depicted in Fig. 1. Further, the use of cognitive radio is expected to grow exponentially through the different vertical sectors in the coming years. By 2022, the scale of the cognitive radio industry is projected to exceed USD 7.44 billion, with a compound annual growth rate (CAGR) of 16.6 percent [12].

In CR system, two types of spectrum users are present: 1) Primary user who owns the spectrum and has the highest priority for using it. 2) Secondary user who can access the spectrum when the primary user is not in operation. Spectrum utilization measurement has shown that many primary user bands are left unused and underused over different space and

The associate editor coordinating the review of this manuscript and approving it for publication was Shaohua Wan.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

FIGURE 1. Cognitive Radio Applications.

time and can be adaptively utilized by the secondary user with the implementation of CR technology. CR can sense the spectrum holes that are not occupied by primary users and allows secondary users to capture that spectrum to start communication. At any frequency, if the primary communication gets affected by secondary, then secondary has to stop its operation over that band and needs to find another inactive primary user band to continue its communication. Hence the secondary user has to be adaptive as spectrum holes are dynamic, and it must be reconfigurable to maintain its communication all the time.

From the hardware aspect, reconfigurability of secondary user system can be achieved with the use of reconfigurable filters. Reconfigurable filters can provide the secondary user the ability to adapt to the dynamic nature of spectrum holes and avoid the interference caused by congested communication applications. So reconfigurable filter is an inevitable hardware for the CR applications.

Reconfigurable microwave filters are the filters which can adapt multiple bands of operation with a single filter structure [13]. It will be highly desirable in future wireless communication with the increase in number of multifunctional devices. Multiple bands of a bandpass filter and bandstop filter can enable the secondary user of CR to select the desired spectrum and reject the interference in an opportunistic approach. Thus, a cognitive approach would pave the way for multiple applications to be introduced without spectrum constraints.

Two different types of filter reconfiguration are considered by researchers. Continuous variation or tuning and discrete variation or switching. The continuous variation or tuning technique exhibits continuous change of centre frequency and/or bandwidth of filter response and is achieved by using varactor diodes [14], Micro-Electrical Mechanical Systems (MEMS) capacitors [15], phase shifters [16], ferroelectric materials [17], photonic microwave filters [18], optical delay line, and ferromagnetic materials [19], [20] etc. In contrast, discrete variation or switching can be generally obtained by PIN diodes [21], MEMS switches [22], optical frequency comb [23] etc. In addition, there are research work where both continuous and discrete variation have been reported.

Although there are many ways to reconfigure filter operation, diode based reconfiguration has gained the utmost attention from researchers for its shorter response delay and easy implementation. Therefore, a study on diode based reconfigurable filter has been considered in this article for exploring spectrum reconfiguration in CR implementation.

A varactor diode is advantageous in continuous variation or tuning over other methods because of its small dimension, low cost, and high tuning speed. PIN diode is also inexpensive and small in dimension. In addition, it exhibits very low loss, good isolation, low parasitic reactance, and high switching speed. PIN diodes can also control large radio frequency signals using low power levels. PIN diodes need a simple DC bias network for acting as a switching circuit. These qualities of PIN diodes make them superior to other methods for achieving discrete variation or switching in reconfigurability.

In this article, three categories of diode based reconfigurable filters have been discussed, mentioning their specifications and specialties. The first category includes varactor based continuous variation or tunable reconfigurable filters. Second is PIN diode based discrete variation or switchable reconfigurable filters. Third is a hybrid approach where both varactor and PIN diodes have been included in a single structure to explore more working options. All three categories are precisely able to adapt reconfigurability for CR systems and thereby have the potential to become the enabling hardware technology for next generation wireless communication. To assist the reader in choosing the suitable method for designing appropriate reconfigurable filters as per the requirement, a well-organized summary diagram of diode based filter reconfiguration and a probable research direction on diode based reconfigurable filter designing for the implementation of CR is also presented.

II. VARACTOR DIODE BASED RECONFIGURABLE FILTERS

Varactor diodes are semiconductor diodes whose capacitance varies with varying bias voltages. Varactor diode is essentially a p-n junction diode that exhibits the characteristics of variable capacitance under the influence of reverse bias voltage. The junction capacitance decreases with the increase in reverse bias voltage, as depicted in Fig. 2. Thus, if a varactor diode is integrated with the filter structure, tuning of resonating frequency is possible with the varying reverse

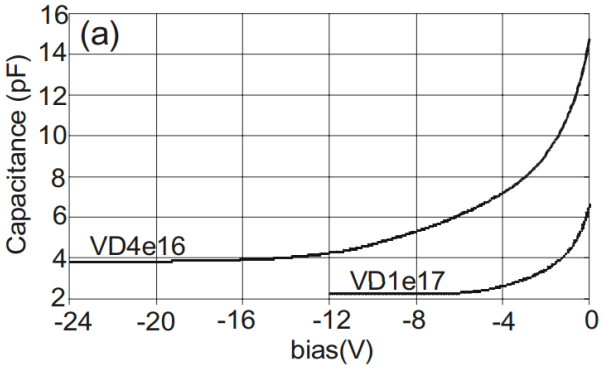


FIGURE 2. Variation in capacitance with bias voltage in varactor diode [24].

bias voltage. So reconfigurability can be attained with the use of varactor diode in the filter circuit.

Table 1 includes underlying information on the number of tunable elements, filter dimensions, tuning range, bandwidth variation, and the value of insertion loss within the stated band from few recent research works on varactor diode based reconfigurable filter. This table finds that tunability of varactor diode can explore mainly three possible combinations of reconfigurable filter characteristics: 'Variable centre frequency constant bandwidth', 'Constant centre frequency variable bandwidth', and 'Variable centre frequency variable bandwidth'. The three categories of varactor diode based reconfigurable filters are explicitly illustrated in Fig. 3 referring their related research papers from Table 1. Discussions on each of the three categories depicting examples are included in the following subsections.

A. VARIABLE CENTRE FREQUENCY CONSTANT BANDWIDTH

In this case, tuning operation in diode changes the central resonating frequency without affecting bandwidth of the filter. This can be seen in a research work [30] where four varactor diodes are engaged to achieve a dual band reconfigurable bandpass filter. Fig. 4 shows the geometric configuration and frequency response to the filter. Reflection coefficient and transmission coefficient curve of this dual band BPF illustrates the centre frequency tuning from 0.98-1.22 GHz in lower band and 1.63-1.95 GHz in upper band with constant absolute bandwidth of 76 MHz and 100 MHz respectively under the influence of two bias voltages.

B. CONSTANT CENTRE FREQUENCY VARIABLE BANDWIDTH

In this case of reconfiguration, tuning affects the bandwidth of the filter but keeps the centre frequency constant. Fig. 5 shows a geometric configuration and frequency response characteristics diagram for a research work [41] where six varactor diodes are used to vary the bandwidth from 440 to 680 MHz keeping the centre frequency constant at 2 GHz.

C. VARIABLE CENTRE FREQUENCY VARIABLE BANDWIDTH

This causes both centre frequency and bandwidth of filter variable under the influence of variable bias voltage. An example of this type [43] where centre frequency tuning from 2.17-2.72 GHz and 3 dB BW tuning range 140-208 MHz around 2.17 GHz, 231-355 MHz around 2.45 GHz, and 314-435 MHz around 2.72 GHz is achieved with the use of varactor diodes. Fig. 6 presents the necessary graph depicting the scattering parameters for variable centre frequency variable bandwidth tunable filter.

It is noteworthy that the capacitance variation in varactor diode does not cause any transformation in filter characterization. So BPF remains BPF, and BSF remains BSF under varactor diode based filter reconfiguration.

III. PIN DIODE BASED RECONFIGURABLE FILTERS

The PIN diode acts as a variable resistance diode under the control of bias voltage. Under reverse bias conditions, a small series junction capacitance causes high impedance to restrict the current from flowing through junction. A forward biasing in PIN diode removes that junction capacitance to exhibit very low impedance and thereby a sharp increase in flowing current through the junction. This behaviour of PIN diode makes it appropriate for switching function. Fig. 7 depicts the characteristics graph for PIN diode where current flow increases with the increase in bias voltage.

PIN diode explores a wide range of reconfiguration in filter characteristics. A large change of centre frequency of reconfigurable filter can be achieved with the use of PIN diodes. Additionally, this reconfiguration technique can transform the filter characteristics exploring the ON and OFF state of PIN diode. So a BPF can be transformed into BSF or vice versa by changing the biasing voltage of PIN diodes. Similarly, BPF can be reconfigured as LPF or HPF or vice versa with the use of PIN diode reconfiguration.

Table 2 includes information on the number of switchable elements, filter dimensions, possible combinations, centre frequency of each combination, bandwidth of each combination, and the value of insertion loss within the stated band from few recent research articles on PIN diode based reconfigurable filters. This table reveals that the biasing voltage of PIN diodes is varied to meet one of the four possible combinations: 'Variable centre frequency constant bandwidth', 'Constant centre frequency variable bandwidth', 'Variable centre frequency variable bandwidth' and 'Filter mode transformation.' These four categories are explicitly illustrated in Fig. 8 referring their related research papers from Table 2.

The first three cases seem to be same as cases of varactor diode based reconfigurable filter. However, in the PIN diode reconfiguration, all the cases are the outcome of switching action instead of tuning action. Because of the switching action, there is a long hop in centre frequency and/or in bandwidth of the filter.

TABLE 1. Varactor diode based tunable filters of different reconfigurable categories.

Ref.	Dimensions	Number of Varactor diodes	Performance/Outcome			
			Filter Type	Tuning range (GHz)	Bandwidth variation	IL (dB)
[25]	$0.09\lambda_g \times 0.1\lambda_g$	7	BPF	0.97-1.53	5.5% (1 dB FBW)	4.2-2
[26]	$0.12\lambda_g \times 0.16\lambda_g$	4	BSF	0.66-0.99	18% (3 dB FBW)	30.0-38.0
[27]	$0.25\lambda_g \times 0.18\lambda_g$	2	BPF	0.542-0.926	105 MHz (3 dB ABW)	1.5-2.56
[28]	$0.005\lambda_g^2$	2	LPF	1.15-2.15	Constant BW	0.2-0.4
[29]	$0.19\lambda_g \times 0.19\lambda_g$	3	BPF	0.6-1.015	15.5% (3 dB FBW)	1.1-2.8
[30]	$0.21\lambda_o \times 0.17\lambda_o$	4	Dual BPF	0.98-1.22 (Lower), 1.63-1.95 (Upper)	76 MHz (Lower), 100 MHz (Upper) (3 dB ABW)	2.1-3.0 (Lower), 2.8-4.0 (Upper)
[31]	$0.29\lambda_o \times 0.19\lambda_o$	2	Dual BSF	1.47-1.84 (Lower), 2.76 (Upper)	370 MHz (Lower), 730 MHz (Upper) 3 dB ABW	More than 33.0 (Lower), More than 10.0 (Upper)
[32]	60 mm × 42 mm	4	BPF	1.60-2.27	137 MHz (3 dB ABW)	4.17-1.99
[33]	65 mm × 20 mm	4	BPF	1.5-1.9	85 MHz (1 dB ABW)	6.2-6.7
[34]	$0.12\lambda_o \times 0.08\lambda_o$	6	BPF	0.5-0.825	56 MHz (3 dB ABW)	2.2-3.3
[35]	$0.065\lambda_g \times 0.065\lambda_g$	3	BPF	0.9-1.5	290 MHz (3 dB ABW)	2.3-1.2
[36]	$0.125\lambda_g \times 0.125\lambda_g$	2	Notch filter	0.71-3.04	450 MHz	More than 20.0
[37]	$0.0189\lambda_g^2$	4	Dual BPF	0.80-1.02 (Lower), 2.02-2.48 (Upper)	170 MHz (Lower), 240 MHz (Upper) 3 dB ABW	1.12-2.93 (Lower), 1.45-4.89 (Upper)
[38]	$0.395\lambda_g \times 0.197\lambda_g$	4	BSF	1.73-2.2	50 MHz (20 dB ABW)	More than 10.0
[39]	$0.029\lambda_g^2$	6	Dual BPF	2.2-2.7 (Lower), 3.45-4.2 (Upper)	500 MHz (Lower), 750 MHz (Upper)	Less than 6.0
[40]	120 mm × 40.2 mm	11	BPF	1.0	46-482 MHz (1 dB BW)	17.7-3.12
[41]	34 mm × 21 mm	6	BPF	2.0	440-680 MHz (3 dB BW)	0.3-0.8
[42]	$0.0088\lambda_g^2$	2	Dual BPF	0.69-0.88 (Lower), 2.67-3.78 (Upper)	23.5%-40.27% (Lower), 30%-14% (Upper) 3 dB FBW	1.83-0.78 (Lower), 1.76-2.02 (Upper)
[43]	$0.24\lambda_g \times 0.189\lambda_g$	5	BPF	2.17-2.72	140-435 MHz (3 dB BW)	1.7-4.8
[44]	28 mm × 65 mm	4	BPF BPF BPF BPF	1.5 1.29-1.71 1.20-1.74 1.37-1.72	87-165 MHz (3 dB BW) 132-137 MHz (3 dB BW) 7.9%-8.0% (3 dB FBW) 161-142 MHz (3 dB BW)	2.4-4.8 2.7-3.2 2.9-4.4 2.5-3.1
[45]	17.6 mm × 10.9 mm	2	BPF	0.61-2.39	150-260 MHz (Lower), 290-520 MHz (Upper) 3 dB ABW	2.1-4.4
[46]	$0.19\lambda_g \times 0.11\lambda_g$	3	BPF	1.52-2.76	240-350 MHz (3 dB BW)	2.4-4.8
[47]	$0.1\lambda_g \times 0.03\lambda_g$	6	BPF	0.56-1.15	65-180 MHz (1 dB BW)	1.4-4.5
[48]	$1.87\lambda_g \times 0.43\lambda_g$	3	BPF	3.75-4.0	140-280 MHz (3 dB BW)	Less than 6.0
[49]	$0.11\lambda_g \times 0.08\lambda_g$	4	BPF	0.8-1.52	63-140 MHz (3 dB BW)	2.4-5.8
[50]	$0.65\lambda_g \times 0.164\lambda_g$	6	BPF	1.11-1.5	46-130 MHz (3 dB BW)	2.62-7.0
[51]	$0.35\lambda_g \times 0.62\lambda_g$	6	BPF	2.45-3.02	630-1500 MHz	0.75-1.1
[52]	$0.09\lambda_g \times 0.20\lambda_g$	4	BPF	1.3-1.6	120-420 MHz (1 dB BW)	Less than 0.9
[53]	$0.12\lambda_g \times 0.08\lambda_g$	10	BPF	0.58-1.22	65-180 MHz (3 dB BW)	1.8-4.6
[54]	$0.138\lambda_g \times 0.131\lambda_g$	8	BPF	0.84-1.15	9.6-12% (3 dB BW)	2.7-1.6
[55]	$0.12\lambda_g \times 0.07\lambda_g$	6	BPF	1.7-2.7	50-110 MHz (1 dB BW)	4.9-3.8
[56]	$0.41\lambda_g \times 0.12\lambda_g$	7	BPF	0.58-0.91	115-315 MHz (1 dB BW)	1.53-1.19
[57]	$0.24\lambda_g \times 0.18\lambda_g$	6	BPF	0.669-1.215	140-644 MHz (3 dB BW)	1.0-2.0
[58]	37 mm × 26 mm	6	Dual BPF	0.77-1.0 (Lower), 1.57-2.0 (Upper)	20.3%-24.7 (Lower 3 dB FBW), 120 MHz (Upper 3 dB ABW)	0.7-1.4 (Lower), 2.74-3.93 (Upper)
[59]	40 mm × 30 mm	4	BPF	0.43-0.6	16-55 MHz (1 dB BW)	1.4-4.63
[60]	11.9 mm × 17.2 mm	10	BPF	1.55-2.1	40-120 MHz (1 dB BW)	6.0-4.5
[61]	$0.19\lambda_o \times 0.04\lambda_o$	3	Dual BPF	0.79-0.86 (Lower), 1.20-1.33 (Upper)	4%-4.5% (Lower), 9%-10.5% (Upper) 3 dB FBW	2.6-3.1 (Lower), 1.8-2.7 (Upper)
[62]	29 mm × 16 mm	5	BPF	2.9-3.5	4%-12% (3 dB FBW)	1.0-3.0
[63]	38 mm × 26 mm	6	Dual BPF	1.48-1.8 (Lower), 2.40-2.88 (Upper)	5.76%-8.55% (Lower), 8.28%-12.42% (Upper) 3 dB FBW	1.99-4.4 (Lower), 1.60-4.2 (Upper)
[64]	10.4 mm × 14.8 mm	5	BPF	1.75-2.25	70-100 MHz (1 dB BW)	7.2-3.2
[65]	$0.18\lambda_g \times 0.21\lambda_g$	9	BPF	1.5-2.2	50-170 MHz (1 dB BW)	5.1-3.2
[66]	48 mm × 36 mm	6	BPF	0.8-1.43 (Lower), 3.0-5.6 (Upper)	25 MHz (Lower), 45-33 MHz (Upper) 1 dB BW	Less than 3.1 (Lower), Less than 2.8 (Upper)

IL: Insertion loss, BPF: Bandpass Filter, BSF: Bandstop Filter, FBW: Fractional bandwidth, LPF: Lowpass Filter, BW: Bandwidth, ABW: Absolute bandwidth

The PIN diode action for each of the four cases of reconfiguration is depicted in the following subsections with appropriate examples.

A. VARIABLE CENTRE FREQUENCY CONSTANT BANDWIDTH

This case of reconfigurability can be explained with a recent research work [69] where in first combination, two PIN diodes are used to get the dual band BPF with centre frequency at 1.2 GHz and 3.5 GHz keeping fractional bandwidth almost constant as 3.5% and 4% respectively. Further, the biasing voltages of two PIN diodes are controlled to stop 1.2 GHz or 3.5 GHz in other combinations. Fig. 9 provides necessary reflection coefficient and transmission coefficient graph of this PIN diode based reconfigurable filter for different combinations.

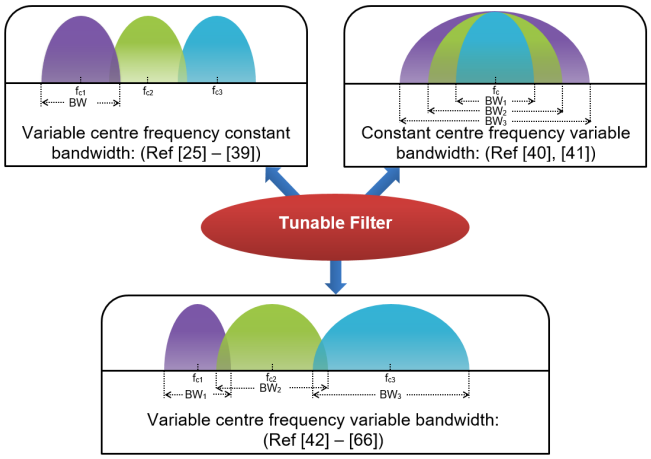


FIGURE 3. Varactor diode based tunable filter of different reconfigurable categories.

B. CONSTANT CENTRE FREQUENCY VARIABLE BANDWIDTH

This type of reconfiguration displays change in bandwidth keeping centre frequency constant. Fig. 10 illustrates the frequency response of an example [70] where fractional bandwidth is varied from 58.5% to 75% maintaining centre frequency constant at 2.4 GHz by controlling the biasing voltages of 4 PIN diodes.

C. VARIABLE CENTRE FREQUENCY VARIABLE BANDWIDTH

In this case, Both centre frequency and bandwidth of the filter vary under the control of PIN diodes bias voltages. This can be viewed in research work in Fig. 11 where two PIN diodes are used to reach three centre frequencies 2.4/3.5/5.2 GHz with their fractional bandwidth percentage 7.9/6/9.1, respectively [74]. Fig. 12 depicts four different combinations of filter output depending on the ON and OFF states of PIN diodes.

D. FILTER MODE TRANSFORMATION

This can only be achieved with the use of PIN diode in diode based reconfiguration circuit. This can change the filter characteristics from BPF to BSF, BSF to BPF, BPF to LPF, and BSF to LPF, etc. Fig. 13 shows an example of filter transformation [86] where different bandpass and bandstop options are explored with biasing of two PIN diodes in three different modes. In first mode, bandpass filter (Passband: 2.8-5.4 GHz) with a notch at 3.5 GHz is obtained, keeping PIN1 at ON state and PIN2 at OFF state. Second mode provides a bandstop filter (Stopband: 3.2-7.0 GHz) keeping PIN1 at OFF state and PIN2 at ON state while the third mode brings a full bandpass filter (Passband: 2.8-5.4 GHz) holding PIN1 at OFF state and PIN2 also at OFF state. The necessary reflection coefficient and transmission coefficient diagram for three different modes are shown in Fig. 13. Further, an almost constant group delay in bandpass region (Fig. 13 (d)) satisfies the linear phase requirement of the filter.

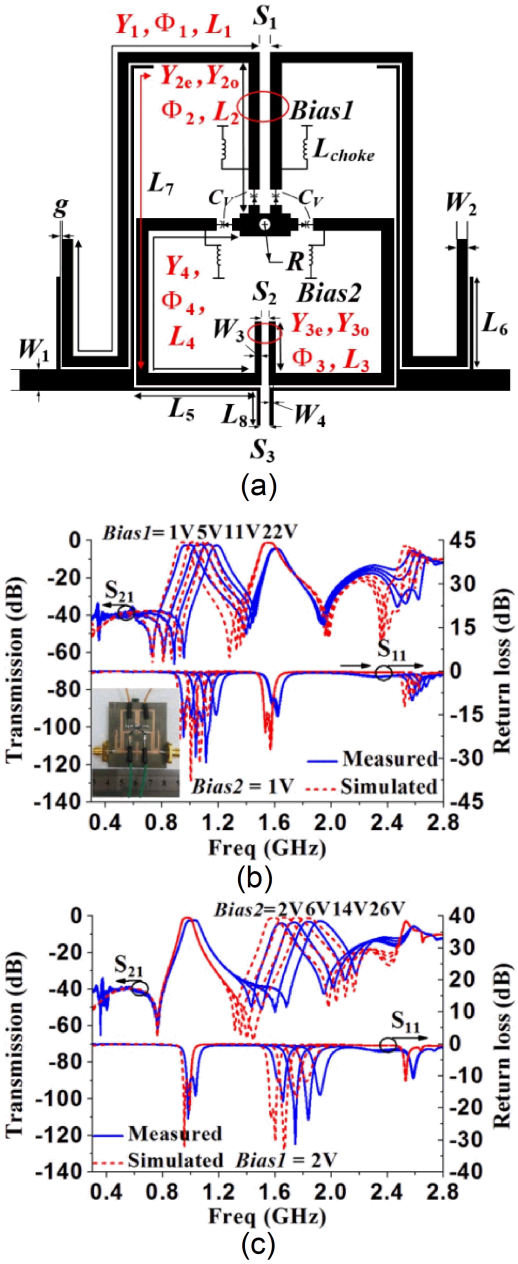


FIGURE 4. The varactor diode based dual band tunable BPF for 'Variable centre frequency constant bandwidth' [30] (a) geometric configuration (b) frequency responses of lower band (c) frequency response of upper band.

Because of its four possible reconfigurable options, PIN diodes are widely used in many applications of reconfigurable filters.

IV. HYBRID DIODE BASED RECONFIGURABLE FILTER

In many recent research works, hybrid structure of reconfigurable filters is proposed in order to bring advantages of both tuning and switching action. These filters include both varactor and PIN diodes in its configuration. The varactor diode causes the filter response to varying gradually while the PIN diode can transform the filter function, e.g., BPF to BSF or BPF to LPF, as well as can change the centre frequency

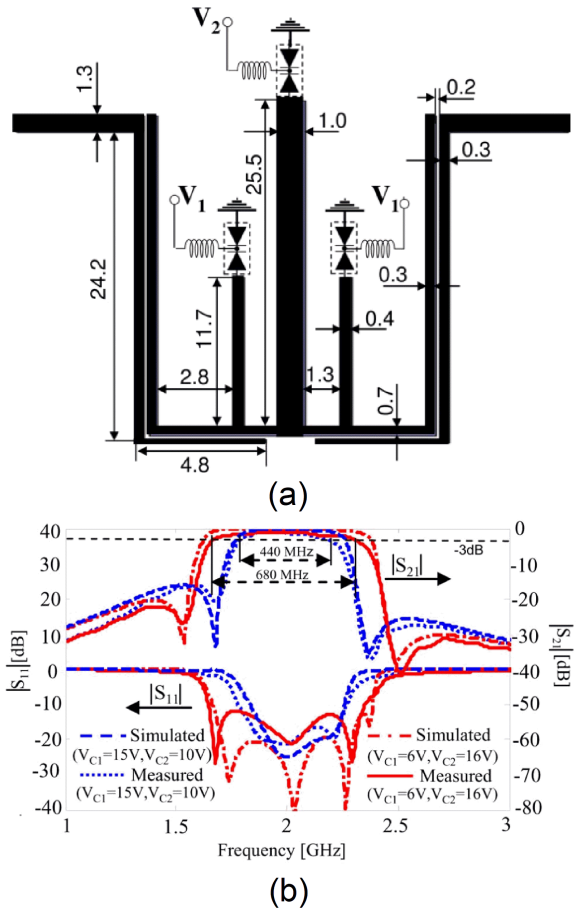


FIGURE 5. The (a) geometric configuration, and (b) frequency responses of varactor based tunable filter for 'Constant centre frequency variable bandwidth' [41].

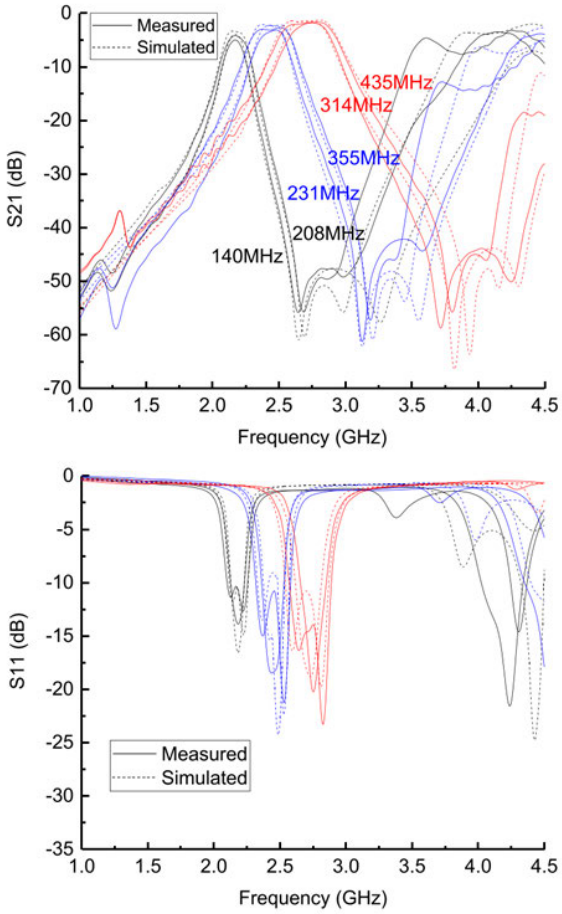


FIGURE 6. The frequency responses of varactor based tunable filter for 'Variable centre frequency variable bandwidth' [43].

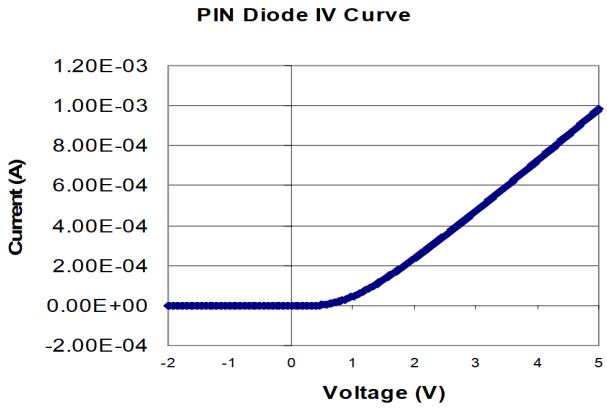


FIGURE 7. Current-Voltage relation for PIN diode [67].

and bandwidth significantly. Table 3 includes information on the number of different diode elements, effect of PIN and varactor diodes on filter characteristics, tuning range, bandwidth variation, and the value of insertion loss within the stated band from few recent research works on hybrid reconfigurable filter. Presence of both varactor and PIN diodes results in three different categories of filter characteristics: 'Variable centre frequency constant bandwidth,' 'Variable centre frequency variable bandwidth,' and 'Multimode filter transformation'. These three categories are explicitly illustrated in Fig. 14 referring their related research papers from Table 3. The action of varactor diode and PIN diode for each of the three cases of hybrid reconfiguration are depicted in the following subsections with appropriate examples.

A. VARIABLE CENTRE FREQUENCY CONSTANT BANDWIDTH

In this case of hybrid reconfigurable filter, gradual shifting of centre frequency under the influence of varactor diode biasing has been implemented at lower and upper range of frequencies keeping bandwidth constant. The function of PIN diode is to switch between lower and upper range of frequencies. A recent research [93] can be considered for its illustration. Five varactor diodes and two PIN diodes are used to bring a tuning range from 255 MHz to 455 MHz and from 455 MHz to 725 MHz at lower and higher band respectively. Bandwidth is kept constant as 75 MHz for both the bands. Lower band and higher band are achieved keeping PIN diodes are OFF and ON, respectively. Tuning in both bands is obtained by controlling three biasing voltages: V_L, V_{od}, and V_{in}. Fig. 15 illustrates the geometric configuration and frequency response of the reconfigurable bandpass filter.

TABLE 2. PIN diode based switchable filters of different reconfigurable categories.

Ref.	Dimensions	Number of PIN diodes	Performance/Outcome			
			Combination with Characteristics	Center Frequency (GHz)	Percentage Bandwidth (%)	IL (dB)
[68]	50 mm × 13 mm	5	i. Five BSFs ii. BSF with no rejection iii. Four BSFs iv. Three BSFs v. Three BSFs	2.412,2.422, 2.437,2.457, 2.484 ... 2.412,2.422, 2.437,2.484 2.422,2.457, 2.484 2.412,2.422, 2.484	4% (3 dB FBW)	More than 10.0
[69]	$0.21\lambda_g \times 0.064\lambda_g$	2	i. Dual band BPF ii. BPF iii. BPF	1.2/3.5 3.5 1.2	3.5%/4% (3 dB FBW) 3.5% (3 dB FBW) 4% (3 dB FBW)	2.79/2.96 2.96 2.79
[70]	$0.514\lambda_g \times 0.514\lambda_g$	4	i. BPF ii. BPF	2.4 2.4	58.5% (3 dB FBW) 75% (3 dB FBW)	Less than 1.1
[71]	29.4 mm × 29.4 mm	4	i. BPF ii. BPF iii. WB-BPF	5.69 5.68 5.66	34.8% (FBW) 48.4% (FBW) 56.5% (FBW)	Less than 1.4
[72]	68 mm × 68 mm	2	Dual BPF	2.4, 2.4	4.8%, 1.5% (3 dB FBW)	1.2, 1.5
[73]	$0.54\lambda_g \times 0.15\lambda_g$	1	i. WB-BPF ii. NB-BPF	1.5 1.5	50% (3 dB FBW) 29% (3dB FBW)	1.1 0.9
[74]	$0.20\lambda_g \times 0.15\lambda_g$	2	i. Tri Band BPF ii. Dual Band BPF iii. Dual Band BPF	2.4/3.5/5.2 3.5/5.2 2.4/3.5	7.9%/6%/9.1% (3 dB FBW) 6%/9.1% (3 dB FBW) 7.9%/6% (3 dB FBW)	Less than 1.9
[75]	30×18 mm^2	3	Multiband BPF	1.2, 1.5, 1.8, 1.9, 2.4, 2.45, 3.2	4.5% for GPS, 5.4% for GSM, 9% for ISM, 27.5% for WiMAX	Less than 3.0
[76]	29.8 mm × 30.8 mm	4	i. Dual band BPF ii. Tri band BPF iii. Quad band BPF	1.71, 1.94 1.71, 1.94, 2.75 1.71, 1.94, 2.75, 2.85	12%, 13% (3 dB FBW) 12%, 13%, 25% (3 dB FBW) 12%, 13%, 25%, 11% (3 dB FBW)	Less than 1.0 Less than 2.5 Less than 3.0
[77]	$0.26\lambda_{g_o} \times 0.13\lambda_{g_o}$	2	i. Tri band BPF ii. Dual band BPF	1.87/2.62/3.95 2.06/3.95	19.6%/8%/24% (3 dB FBW) 30%/24% (3 dB FBW)	Less than 1.5 Less than 3.0
[78]	20 mm × 20 mm	3	i. BPF ii. BPF iii. BPF	2.4 3.5 1.65-1.89	20% (3 dB FBW) for WLAN 11% (3 dB FBW) for WiMAX 14% (3 dB FBW)	#
[79]	$0.75\lambda \times 0.50\lambda$	4	Dual BPF	1.5, 2.0	11.4%, 10.8% (3 dB FBW)	3.05, 3.1
[80]	24 mm × 30 mm	2	Dual band BPF BPF	2.45, 5.2 5.2	6.2%, 5.4% (3 dB FBW) 7% (3 dB FBW)	2.45, 2.1 2.1
[81]	30×30 mm^2	2	Dual BPF BPF	1.8, 3.5 3.5	9%, 7% (3 dB FBW) 7% (3 dB FBW)	Nearly 3.0
[82]	$0.50\lambda \times 0.25\lambda$	6	i. UWB-BPF ii. NB-BPF iii. UWB-BSF	2.57 2.9 2.54	95.7% (3 dB FBW) 25.3% (3 dB FBW) 106% (3 dB FBW)	0.94 0.8 More than 10.0
[83]	$0.41\lambda_g \times 0.76\lambda_g$	2	Bandpass to bandstop	2.4 with 5.0 stopband 5.0 with 2.4 stopband	100.76% (10 dB BW) 45.78% (10 dB BW)	0.47 1.28
[84]	22.5 mm × 24.3 mm	4	i. SB-BSF ii. ASF iii. SB-BPF iv. DB-BSF	2.0 dc to 4.3 1.92 1.61/2.36	50.2% (10 dB FBW) 4.3 GHz 9.7% (3 dB FBW) 21.1%/20% (10 dB FBW)	More than 10.0 12.0 1.95 More than 25.0
[85]	$0.41\lambda_g \times 0.24\lambda_g$	5	i. BPF ii. BSF iii. UWB BPF iv. UWB with notched band v. Narrow band BPF	1.83 1.95, 1.88, 1.83, 1.77 1.84-3.43 1.80-1.85 & 3.43-3.50 1.77	14.8% (FBW) 9.7% for UMTS, 9.8% for PCS, 11% for GSM, 12.5% for DCS 63% (FBW) 2% (FBW) 18%(FBW)	1.01 25.5 25.0 40.0 21.0 0.25 27 0.15
[86]	$0.15\lambda_g \times 0.13\lambda_g$	2	i. BPF with one notch band ii. BSF iii. BPF	4.15 with notch at 3.56 4.73 4.15	2.8-3.52 GHz and 3.63-5.4 GHz with 3.2% (5 dB FBW) for notch band 3.2-7.0 GHz (80% 3 dB FBW) 2.8-5.4 GHz (63% 3 dB FBW)	0.9 for BPF 16.0 for notch band More than 20.0 0.9
[87]	$0.75\lambda \times 0.50\lambda$	4	i. Dual band lowpass bandpass filter ii. LPF iii. BPF iv. All reject	1.5, 2.4 1.5 2.4 ...	1.5 GHz BW, 9%(3 dB BW) 1.5 GHz BW 9% (3 dB BW) ...	0.37, 3.3 0.37 3.3 33.0, 26.0
[88]	5.0×4.4 mm^2	2	Passband Stopband	5.5 ...	5% (FBW) ...	2.35 More than 20.0

B. VARIABLE CENTRE FREQUENCY VARIABLE BANDWIDTH

This case includes a change in both centre frequency and bandwidth. Involvement of both tuning and switching action enables the filter reconfigurability to be implemented in a wide frequency range. Fig. 16 illustrates an example of this case [99], where four varactor diodes are used to tune the centre frequency from 1.25 GHz to 1.45 GHz, and one PIN

TABLE 2. *(Continued.)* **PIN diode based switchable filters of different reconfigurable categories.**

Ref.	Dimensions	Number of PIN diodes	Performance/Outcome			
			Combination with Characteristics	Center Frequency (GHz)	Percentage Bandwidth (%)	IL (dB)
[89]	21.8 mm × 22 mm	1	i. BPF ii. Band notched	1.7 2.4	118% (3 dB FBW) 15% (3 dB FBW)	0.8 More than 20.0
[90]	$0.25\lambda_g \times 0.23\lambda_g$	1	i. Band notched ii. BPF	3.52 3.1-5.0	3.2% (5 dB rejection FBW) 60% (3 dB FBW)	16.0 0.7
[91]	8.3 mm × 7 mm	1	i. Notch band ii. UWB BPF	5.67 3.0-10.9	2.3% (10 dB BW) 138% (3 dB BW)	More than 10.0 0.2
[92]	39.2 mm × 6 mm	2	i. UWB BPF ii. Notch at 5GHz iii. Notch at 2.4GHz iv. Dual notches	1.8-7.8 5.15 2.4 5.15, 2.4	160% (3 dB FBW) 4% (3 dB FBW) 21% (3 dB FBW) 4%, 21% (3 dB FBW)	0.5-7.0 27.0 38.0 27.0, 38.0

IL: Insertion loss, BSF: Bandstop Filter, FBW: Fractional bandwidth, BPF: Bandpass Filter, WB-BPF: Wide-band bandpass filter, NB-BPF: Narrow-band bandpass filter, #: Not mentioned, UWB: Ultra wideband, SB-BSF: Single-band bandstop filter, ASF: All-stop filter, SB-BPF: Single-band bandpass filter, DB-BSF: Dual-band bandstop filter, LPF: Lowpass Filter

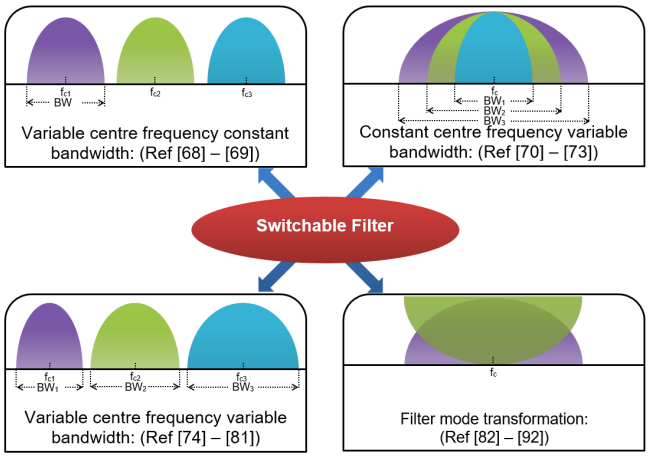


FIGURE 8. **PIN diode based switchable filter of different reconfigurable categories.**

diode is used to switch the FBW in two discrete states of 8% and 12%. The varactor biasing voltages are varied from 0 to 20 V.

C. MULTIMODE FILTER TRANSFORMATION

Multimode transformation in reconfigurable mechanism enables gradual tuning as well as filter characteristics transformation. An illustration of this case can be viewed in a recent research work [103]. A geometric configuration of this reconfigurable filter is shown in Fig. 17. In this work, six varactor diodes and one PIN diode have been utilized to explore two different modes: dual bandpass mode and dual bandstop mode. In the dual bandpass mode, the lower passband tuning range is from 1.7 to 2.2 GHz, and the higher passband tuning range is from 2.2 to 2.7 GHz. For the dual bandstop mode, the lower stopband tuning range is from 1.7 to 2.3 GHz, and the higher stopband tuning range is from 2.3 to 2.9 GHz. A picture of reflection coefficient and transmission coefficient under the control of PIN diode and varactor diode biasing is shown in Fig. 18 for depicting the filter responses at different modes.

V. DIODE BASED RECONFIGURABLE FILTER FOR CR APPLICATIONS

So far, three different categories of reconfigurability in the design of diode based filters have been discussed: Tunable filters, Switchable filters, and Hybrid reconfigurable filters. Based on reflection coefficient, and transmission coefficient value in the frequency band, each category of reconfigurability again explored multiple options of their possible applications. Fig. 19 summarizes all the possible combinations explored by the researchers in the domain of diode based reconfigurable filter designing.

The main aim of the CR framework is to bring various heterogeneous devices and systems together to provide smart services with minimal device resonance requirements, i.e., power, hardware complexity, cost etc. [106]. Under all circumstances, to attain this objective, the communication between devices should always be continued. This demands cyclic execution of four routine phases of cognitive radio: sensing, decision-making, sharing, and mobility (Fig. 20). The cognition process starts with the step of spectrum sharing (SS), in which various SS methods are used to identify the available spectrum resources across the specified spectrum band. Depending on the results of the detection, the decision is made for sharing the band concurrently, or to cease transmission within that band.

Once a CR takes a decision to use the band, a proper Medium Access Control (MAC) protocol is utilized, and power allocation should be reasonably considered to protect the primary user. Finally, the mobility process is performed by switching applications from one band to another.

The authors in [107] proposed the expansion of four traditional cognition phases by introducing certain additional CR modules, as shown in Fig. 21. The diagram also offers a briefing on the role of each module.

Diode based reconfigurable filters can contribute to the physical implementation of reconfigurable modules or the mobility phase of cognition cycle for CR. Out of several reconfigurable parameters, as shown in Fig. 21, frequency, and bandwidth can be cognitively controlled by the reconfigurable filters presented in this article.

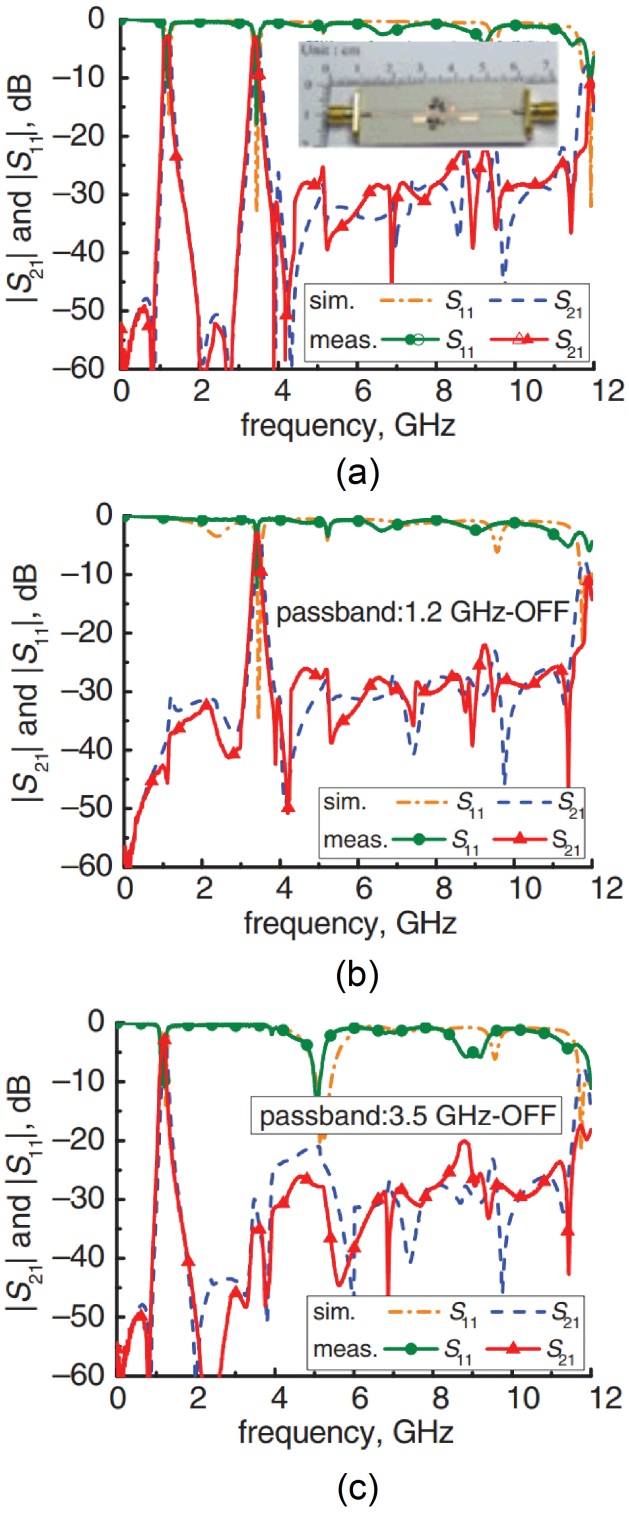


FIGURE 9. The four different frequency responses of PIN diode based switchable filter for 'Variable centre frequency constant bandwidth' [69] (a) dual band BPF (1.2/3.5 GHz) (b) BPF (3.5 GHz) (c) BPF (1.2 GHz).

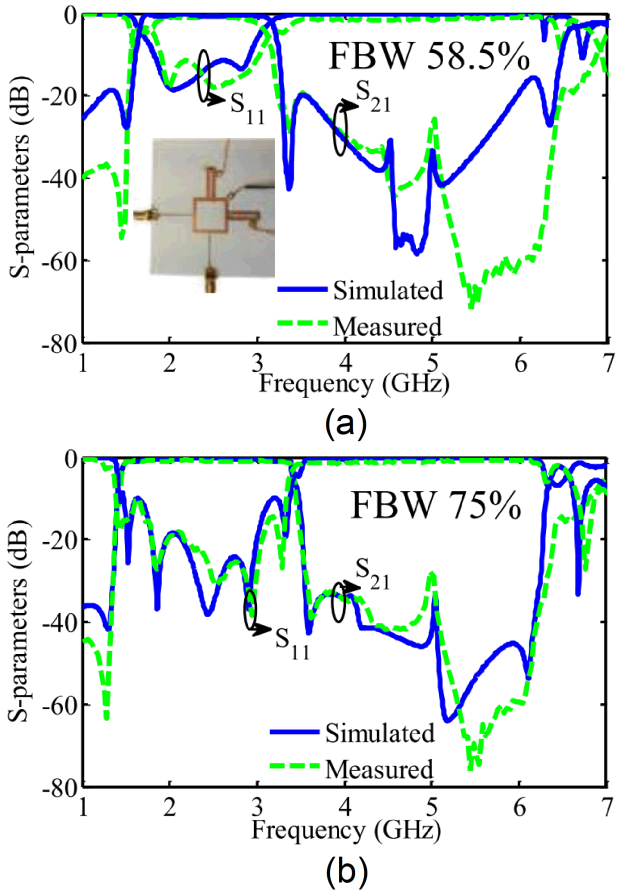


FIGURE 10. The frequency responses of PIN diode based switchable filter for 'Constant centre frequency variable bandwidth' [70] (a) BPF (2.4 GHz) with BW 58.5% when PIN diodes are at OFF state (b) BPF (2.4 GHz) with BW 78% when PIN diodes are at ON state.

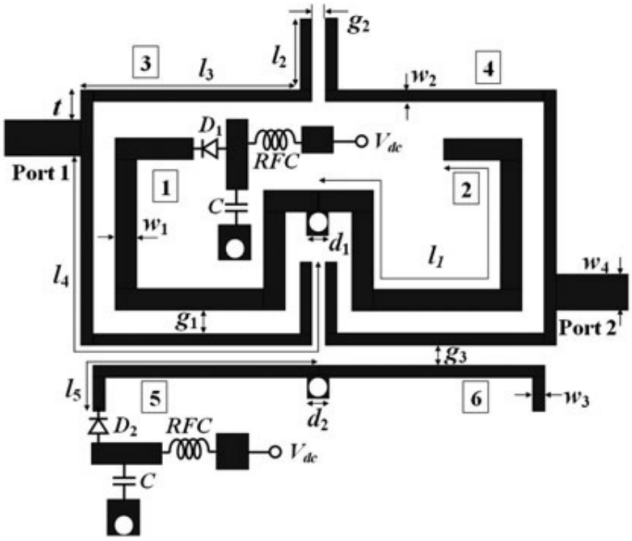


FIGURE 11. Geometric configuration of PIN diode based reconfigurable filter for 'Variable center frequency variable bandwidth' [74].

VI. FUTURE PERSPECTIVE

In the upcoming domain of wireless communication, certain application possibilities of reconfigurable filters could create opportunities in the implementation of cognitive systems. A brief discussion on future perspective of different categories of reconfigurable filters are presented in this section.

TABLE 3. Hybrid Diode Based Reconfigurable Filters.

Ref.	Dimensions	Number of diodes	Performance/Outcome				
			Combination with Characteristics	PIN effect	Varactor effect (Centre frequency tuning, GHz)	Bandwidth	IL (dB)
[93]	$0.2\lambda_g \times 0.14\lambda_g$	5 Varactor, 2 PIN	i. BPF ii. BPF	i. Lower BPF when PIN OFF ii. Higher BPF when PIN ON	0.255-0.455 0.455-0.725	75 MHz (3 dB ABW) 75 MHz (3 dB ABW)	1.4-1.8 2.5-2.9
[94]	60 mm × 46 mm	4 Varactor, 2 PIN	BPF	i. Frequency tunable passband with constant ABW when PIN ON ii. In band isolation is optimized when PIN OFF	0.785-0.566	51 MHz (3 dB ABW)	2.52-4.08
[95]	$0.06\lambda_g \times 0.27\lambda_g$	7 Varactor, 2 PIN	i. BPF ii. BPF	i. Lower BPF when PIN ON ii. Higher BPF when PIN OFF	1.1-1.5 1.5-2.1	40 MHz (3 dB ABW)	4.4-6.1
[96]	25 mm × 55 mm	1 Varactor, 1 PIN	i. BPF ii. BPF	i. BPF BW 36 MHz when PIN ON ii. BPF BW 45 MHz when PIN OFF	0.72-0.82 0.72-0.82	36 MHz (3 dB ABW) 45 MHz (3 dB ABW)	2.8 2.8
[97]	35 mm × 47 mm (2-pole) 36 mm × 58 mm (3-pole)	4 Varactor, 4 PIN (2-pole) 6 Varactor, 4 PIN (3-pole)	i. BPF (2-ploe) ii. BPF (3-pole)	i. low band when Sw 2 ON and Sw 1 OFF ii. high band when Sw 2 OFF and Sw 1 ON iii. low band when Sw 2 ON and Sw 1 OFF iv. high band when Sw 2 OFF and Sw 1 ON	0.55-0.99 0.99-1.9 0.54-1.006 0.96-1.8	92 MHz ($3dB$ BW) 91 MHz (3 dB BW) 89 MHz (3 dB BW) 91 MHz (3 dB BW)	3.2-4.4 3.2-4.0 4.2-5.4 4.0-5.4
[98]	6.12 mm × 4.12 mm (unit cells)	16 Varactor or 16 PIN	i. Notch ii. Notch	i. When OFF BW is 5.7% ii. When ON BW is 6.75%	5.09-5.43	450 MHz 340 MHz	More than 10.0
[99]	40 mm × 40 mm	4 Varactor, 1 PIN	i. BPF ii. BPF	i. BPF with narrow BW when PIN OFF ii. BPF with wide BW when PIN ON	1.25-1.45 1.25-1.45	8% FBW 12% FBW	Nearly 3.0 Nearly 3.0
[100]	$0.16\lambda_g \times 0.11\lambda_g$	4 Varactor, 1 PIN	i. BPF ii. BSF	i. BPF when PIN OFF ii. BSF when PIN ON	0.79-1.59 0.7-1.47	22% (3 dB FBW) 7% (15 dB FBW)	1.3-2.1 More than 20.0
[101]	$0.52\lambda_g \times 0.42\lambda_g$	4 Varactor, 4 PIN	i. BPF ii. BSF	i. BPF when PIN ON ii. BSF when PIN OFF	1.8-2.5 1.0-1.5	127-29 MHz ...	2.5-3.2 28.0
[102]	$0.14\lambda_g \times 0.12\lambda_g$	3 Varactor, 2 PIN	i. BSF ii. BPF	i. BSF when PIN ON ii. BPF when PIN OFF	0.95-1.35 0.95-1.35	3%-7% (3 dB FBW) 15% (3 dB FBW)	More than 13.0 4.7-5.6
[103]	$0.333\lambda_g \times 0.346\lambda_g$	6 Varactor, 1 PIN	i. Dual BPF ii. Dual BSF	i. BPF when PIN ON ii. BSF when PIN OFF	1.7-2.2 (lower), 2.2-2.7 (upper) 1.7-2.3 (lower), 2.3-2.9 (upper)	200 MHz (lower), 400 MHz (upper) 200 MHz (lower), 500 MHz (upper)	4.6-5.0 (lower), 4.6-5.6 (upper) More than 20.0
[104]	$0.084\lambda_g \times 0.084\lambda_g$	5 Varactor, 3 PIN	i. LPF ii. BPF I iii. BPF II	i. LPF when PIN OFF ii. BPF when PIN ON iii. BPF when PIN ON	0.59-1.48 0.85-1.25 1.04-1.6	NA 550-800MHz 140-215MHz	0.82-0.43 1.65-1.16 2.84-1.47
[105]	58 mm × 54 mm	4 Varactor, 3 PIN	i. BPF ii. BSF	i. Bandstop mode when D1 ON and D2-D3 are OFF ii. Bandpass mode when D1 is OFF and D2-D3 are ON	0.805-1.032 0.760-1.228	92-166 MHz (3 dB) 52-85 MHz (3 dB)	3.1-5.9 More than 40.0

IL: Insertion loss, BPF: Bandpass Filter, BSF: Bandstop Filter, ABW: Absolute bandwidth, FBW: Fractional bandwidth

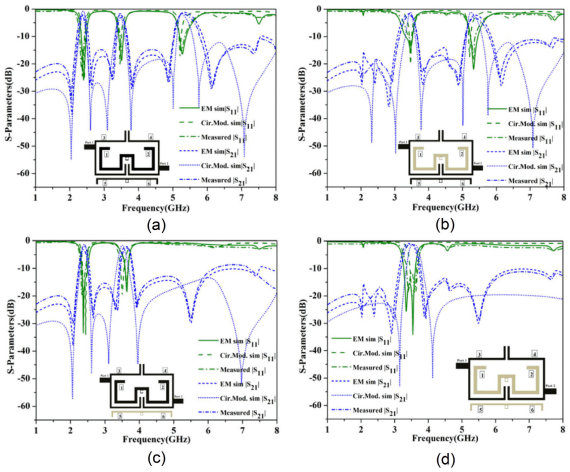


FIGURE 12. The frequency responses of PIN diode based switchable filter for 'Variable centre frequency variable bandwidth' [74] (a) Tri Band BPF (2.4/3.5/5.2 GHz) with FBW 7.9%/6%/9.1% when diode D1 and D2 are OFF (b) Dual Band BPF (3.5/5.2 GHz) with FBW 6%/9.1% when diode D1 is ON, and D2 is OFF (c) Dual Band BPF (2.4/3.5 GHz) with FBW 7.9%/6% when diode D1 is OFF, and D2 is ON (d) BPF (3.5 GHz) with FBW 6% when diode D1 and D2 are ON.

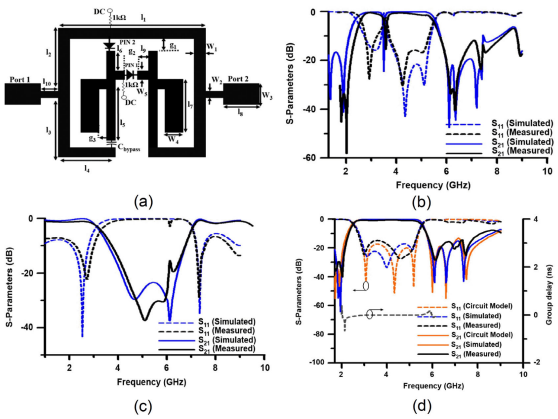


FIGURE 13. The geometric configuration and frequency response of PIN diode based switchable filter for 'Filter mode transformation' [86] (a) filter structure (b) BPF with notch band at 3.56 GHz response when PIN1 is ON, and PIN2 is OFF (c) BSF when PIN1 is OFF, and PIN2 is ON (d) BPF when PIN1 and PIN2 are OFF.

The three reconfiguration abilities of tunable filters, 'Variable centre frequency constant bandwidth', 'Constant centre frequency variable bandwidth', and 'Variable centre frequency variable bandwidth' can be involved in different forms of cognitive radio implementation. 'Variable centre frequency constant bandwidth' can facilitate the secondary

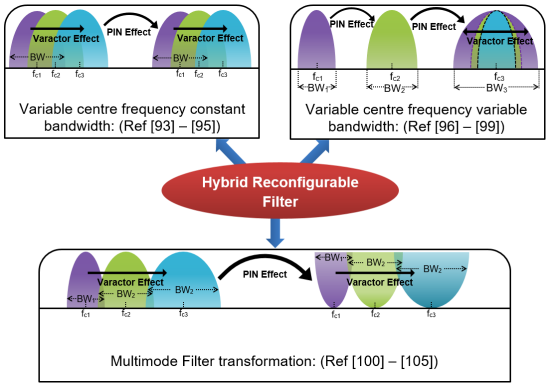


FIGURE 14. Hybrid reconfigurable filter categories.

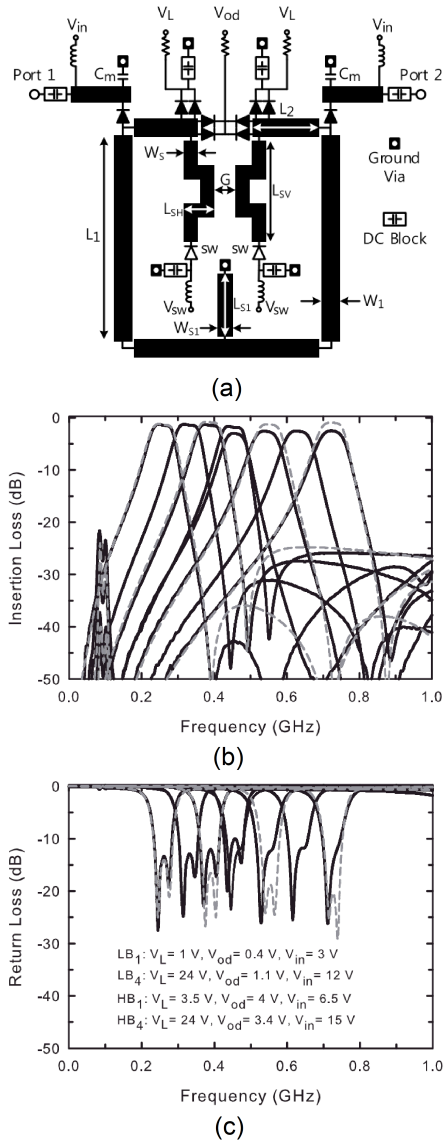


FIGURE 15. The hybrid reconfigurable filter for 'Variable centre frequency constant bandwidth' [93] (a) geometric configuration (b) insertion loss of both lower band (Both PIN diodes are OFF) and higher band (Both PIN diodes are ON) under the control of biasing voltages of varactor diodes (c) return loss of both lower band (Both PIN diodes are OFF) and higher band (Both PIN diodes are ON) under the control of biasing voltages of varactor diodes.

user to continue the communication by altering the carrier frequency as per the spectrum availability without changing

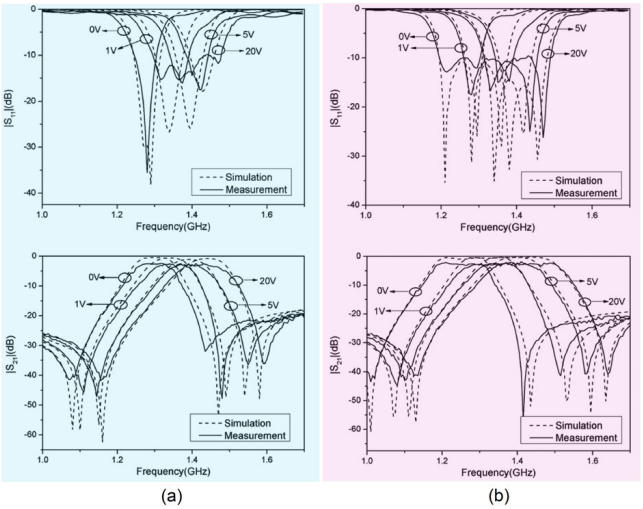


FIGURE 16. The frequency responses of hybrid reconfigurable filter for 'Variable centre frequency variable bandwidth' [99] (a) return loss and insertion loss BPF (Centre frequency tuning 1.25-1.45 GHz) with 8% FBW when PIN diode is OFF (b) return loss and insertion loss BPF (Centre frequency tuning 1.25-1.45 GHz) with 12% FBW when PIN diode is ON.

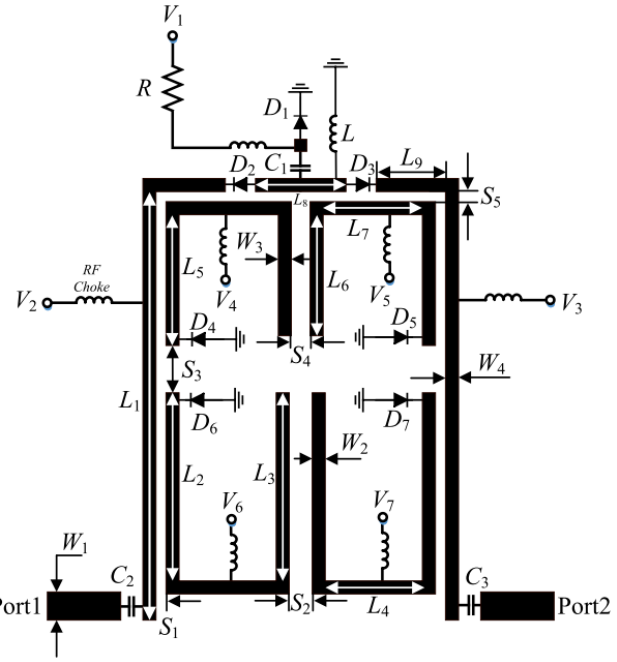


FIGURE 17. Geometric configuration of hybrid reconfigurable filter for Multimode transformation [103].

the required bandwidth while the 'Constant centre frequency variable bandwidth' can enable the CR to adapt the rate of data transfer as per the availability of bandwidth following Shannon's theorem on channel capacity. This is extremely useful in the implementation of cognitive radio of variable bandwidth. 'Variable centre frequency variable bandwidth' can explore the opportunity of implementing secondary user communication with cognitive centre frequency and cognitive bandwidth approach.

The switchable filter uncovers 'Filter mode transformation' in addition to possibilities explored by tunable filters. CR may be in need to restrict the secondary user to

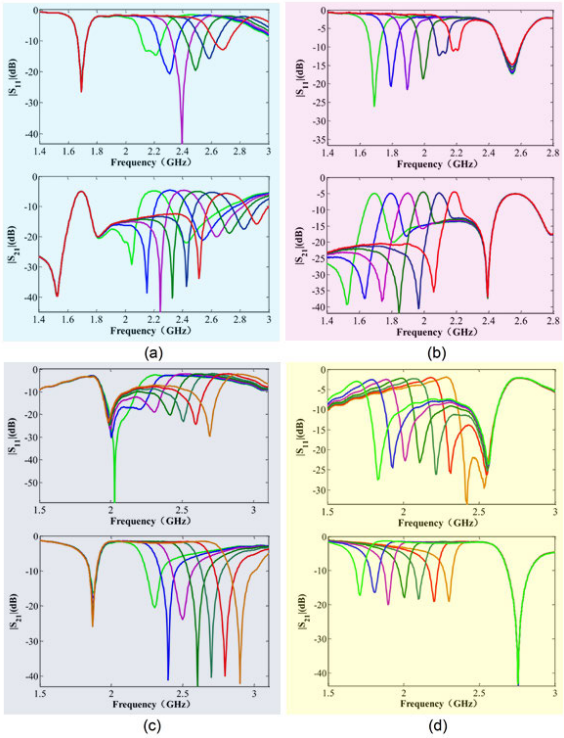


FIGURE 18. The frequency responses of hybrid reconfigurable filter for multimode transformation [103] (a) Dual Bandpass with fixed lower band but variable higher band when PIN diode is ON and V2=V3=30V, V4=V5=0∼6V, V6=V7=0V (b) Dual Bandpass with variable lower band but fixed higher band when PIN diode is ON and V2=V3=30V, V4=V5=3V, V6=V7=0∼30V (c) Dual Band stop with fixed lower band and variable higher band when PIN diode is OFF and V2=V3=0V, V4=0∼18V, V5=0.4∼30V, V6=1.9V, V7=1.6V (d) Dual Bandstop with variable lower band and fixed higher band when PIN diode is OFF and V2=V3=0V, V4=7.6V, V5=12V, V6=0.1∼30V, V7=0∼26V.

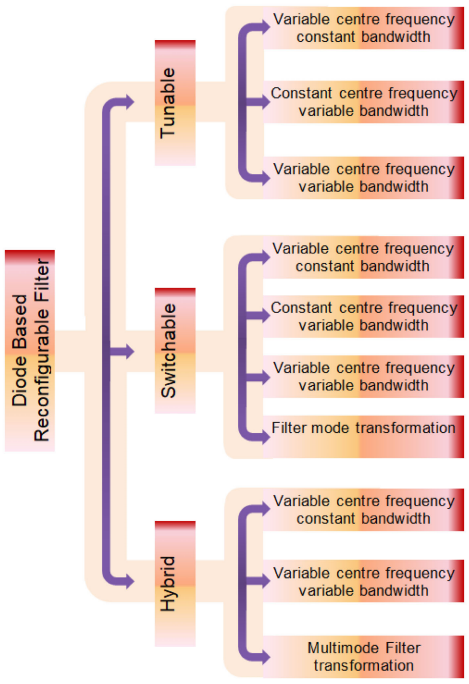


FIGURE 19. Summary for diode based reconfigurable filter categories.

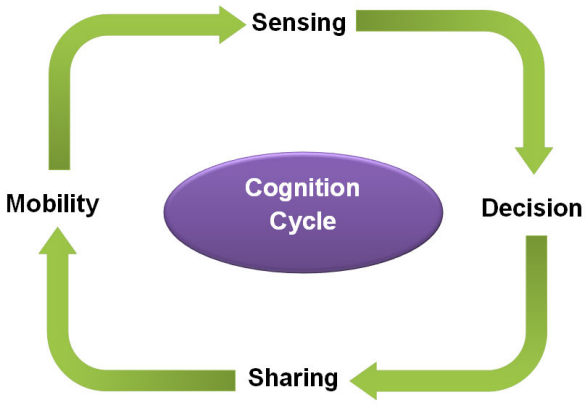


FIGURE 20. Common cognition cycle.

Spectrum sensing module	•Employs energy detection, matched filter detection, or feature detection to monitor spectrum and identify spectrum frequencies used by PU/SU. •Uses cooperative sensing to avoid misdetections. •Controls the sensing parameters.
Spectrum analysis module	•Analyzes the sensing results to identify capacity, condition, interference, occupancy, PU presence, and other link-layer related parameters of spectrum frequencies. •Provides output to the spectrum decision module and to the history and prediction module.
History and prediction module	•Keeps a history of the channel occupancy in each band and identifies frequencies in use by utilizing prediction models. •Transfers this information to the sensing module in order to avoid sensing of those frequencies.
Spectrum decision module	•Selects the spectrum to be used and its operating parameters, i.e., central frequency, bandwidth, modulation, power, etc. •Considers the remaining energy in the battery and application-specific traffic QoS requirements from the upper layers.
Spectrum-aware routing module	•Combines the routing algorithms with a spectrum-aware metric. •Also considers the remaining battery to avoid overutilizing of machines with low battery.
Spectrum- mobility module	•This is a cross-layer module to perform spectrum handovers in case of a PU activity or if the current spectrum cannot meet application-specific requirements.
Security and privacy module	•Ensures the secure and privacy preserving operations such as secure object configuration and management, information security, privacy enhancing technologies, trust modeling, and avoiding PU emulation attacks.
Reconfiguration module	•Reconfigures parameters (frequency, bandwidth, modulation, channel coding, output power, and other operational parameters, i.e., receiver sensitivity, noise threshold, BER, etc.) of the SDR-based RF front-end according to the selected communication technology/protocol.

FIGURE 21. An enhanced version of cognitive radio modules [107].

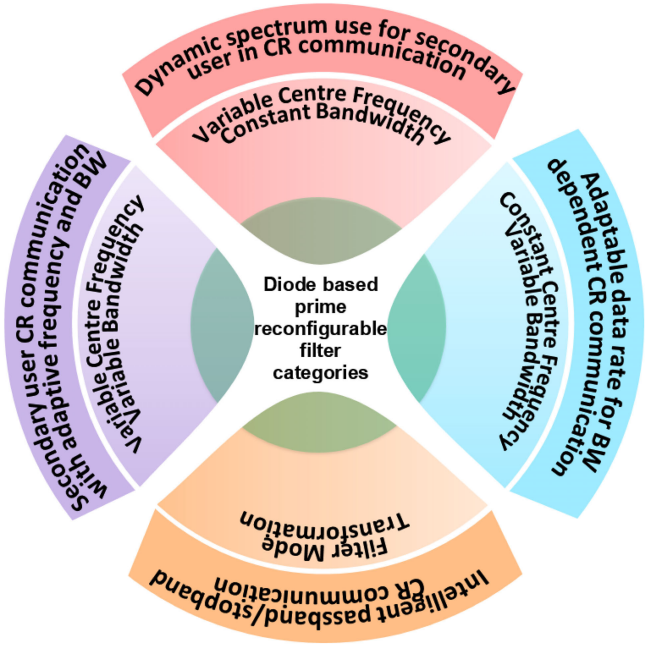


FIGURE 22. Prospective CR applications of diode based prime reconfigurable filter categories.

allow the primary in some particular frequency band. 'Filter mode transformation' can convert the filter from bandpass to bandstop to limit the secondary from participating in any licensed band communication.

The hybrid approach brings both tuning and switching action in filter characteristics with the use of both varactor diode and PIN diode. One new possibility has been explored with this complex approach as 'Multimode filter transformation.' This can enable the CR to adapt multiband acceptance or multiband rejection for fulfilling the need of a collaborative CR network.

Fig. 22 illustrates some prospective applications of prime reconfigurable filter categories for CR implementation. All of these possible outcomes could pave a new research path for the implementation of CR in the use of different customized wireless communication requirements.

VII. CONCLUSION

This article has addressed a comprehensive survey on diode based reconfigurable filters for cognitive radio applications. Three different types of reconfigurable filters based on diode were primarily reviewed: Tunable filter, Switchable filter, and Hybrid reconfigurable filter. Each type is further divided into subgroups. Filter response and controlling action of each subgroup have been illustrated with examples. Several research papers of each type, including the recently published ones, have been enlisted for exploring possibilities for CR implementation. A summary has been drawn to show the potential use of three types of reconfigurable filters for realizing CR systems. The work presented here may be of assistance to the researchers working on reconfigurable filter development for CR devices. Some application possibilities of reconfigurable filters may open up new opportunities in the implementation of cognitive systems in the future arena of wireless communication.

REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] J. Marinho and E. Monteiro, "Cognitive radio: Survey on communication protocols, spectrum decision issues, and future research directions," *Wireless Netw.*, vol. 18, no. 2, pp. 147–164, Feb. 2012.

[3] M. Song, C. Xin, Y. Zhao, and X. Cheng, "Dynamic spectrum access: From cognitive radio to network radio," *IEEE Wireless Commun.*, vol. 19, no. 1, pp. 23–29, Feb. 2012.

[4] X. Hong, J. Wang, C.-X. Wang, and J. Shi, "Cognitive radio in 5G: A perspective on energy-spectral efficiency trade-off," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 46–53, Jul. 2014.

[5] D. Wang, B. Song, D. Chen, and X. Du, "Intelligent cognitive radio in 5G: AI-based hierarchical cognitive cellular networks," *IEEE Wireless Commun.*, vol. 26, no. 3, pp. 54–61, Jun. 2019.

[6] L.-W. Zhao, Z.-J. Zhou, Z.-M. Miao, and Y. Hui, "An introduction to application of the cognitive radio in military communication," *Wireless Commun. Technol.*, vol. 4, pp. 1–32, Sep. 2007, doi: 10.1007/978-981-10-1389-8_59-1.

[7] S. Sodagari, B. Bozorgchami, and H. Aghvami, "Technologies and challenges for cognitive radio enabled medical wireless body area networks," *IEEE Access*, vol. 6, pp. 29567–29586, 2018.

[8] P. Jacob, R. P. Sirigina, A. S. Madhukumar, and V. A. Prasad, "Cognitive radio for aeronautical communications: A survey," *IEEE Access*, vol. 4, pp. 3417–3443, 2016.

[9] S. Haykin, "Cognitive dynamic systems: Radar, control, and radio [Point of View]," *Proc. IEEE*, vol. 100, no. 7, pp. 2095–2103, Jul. 2012.

[10] R. Ferrus, O. Sallent, G. Baldini, and L. Goratti, "Public safety communications: Enhancement through cognitive radio and spectrum sharing principles," *IEEE Veh. Technol. Mag.*, vol. 7, no. 2, pp. 54–61, Jun. 2012.

[11] P. Victor Rodrigues Ferreira, R. Paffenroth, A. M. Wyglinski, T. M. Hackett, S. G. Bilén, R. C. Reinhart, and D. J. Mortensen, "Multiobjective reinforcement learning for cognitive satellite communications using deep neural network ensembles," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 5, pp. 1030–1041, May 2018.

[12] *Cognitive Radio Market Size, Share and Global Market Forecast to 2022 | MarketsandMarkets*. Accessed: Sep. 3, 2020. [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/cognitive-radio-market-219954724.html

[13] J.-S. Hong, "Reconfigurable planar filters," *IEEE Microw. Mag.*, vol. 10, no. 6, pp. 73–83, Oct. 2009.

[14] M. Sanchez-Renedo, R. Gomez-Garcia, J. I. Alonso, and C. Briso-Rodriguez, "Tunable combline filter with continuous control of center frequency and bandwidth," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 1, pp. 191–199, Jan. 2005.

[15] A. Abbaspour-Tamijani, L. Dussopt, and G. M. Rebeiz, "Miniature and tunable filters using mems capacitors," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 7, pp. 1878–1885, Jul. 2003.

[16] E. Ojefors, S. Cheng, K. From, I. Skarin, P. Hallbjorner, and A. Rydberg, "Electrically steerable single-layer microstrip traveling wave antenna with varactor diode based phase shifters," *IEEE Trans. Antennas Propag.*, vol. 55, no. 9, pp. 2451–2460, Sep. 2007.

[17] S. Courreges, Y. Li, Z. Zhao, K. Choi, A. Hunt, and J. Papapolymerou, "A low loss X-Band quasi-elliptic ferroelectric tunable filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 19, no. 4, pp. 203–205, Apr. 2009.

[18] W. Zhang and R. A. Minasian, "Widely tunable single-passband microwave photonic filter based on stimulated Brillouin scattering," *IEEE Photon. Technol. Lett.*, vol. 23, no. 23, pp. 1775–1777, Dec. 1, 2011.

[19] M. S. Rasras, C. K. Madsen, M. A. Cappuzzo, E. Chen, L. T. Gomez, E. J. Laskowski, A. Griffin, A. Wong-Foy, A. Gasparyan, A. Kasper, J. Le Grange, and S. S. Patel, "Integrated resonance-enhanced variable optical delay lines," *IEEE Photon. Technol. Lett.*, vol. 17, no. 4, pp. 834–836, Apr. 2005.

[20] M. Liu, B. M. Howe, L. Grazulis, K. Mahalingam, T. Nan, N. X. Sun, and G. J. Brown, "Voltage-Impulse-Induced non-volatile ferroelastic switching of ferromagnetic resonance for reconfigurable magnetoelectric microwave devices," *Adv. Mater.*, vol. 25, no. 35, pp. 4886–4892, Sep. 2013.

[21] C. Lugo and J. Papapolymerou, "Six-state reconfigurable filter structure for antenna based systems," *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 479–483, Feb. 2006.

[22] K. Y. Chan, S. Fouladi, R. Ramer, and R. R. Mansour, "RF MEMS switchable interdigital bandpass filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 22, no. 1, pp. 44–46, Jan. 2012.

[23] P. Shen, N. J. Gomes, P. A. Davies, P. G. Huggard, and B. N. Ellison, "Analysis and demonstration of a fast tunable Fiber-Ring-Based optical frequency comb generator," *J. Lightw. Technol.*, vol. 25, no. 11, pp. 3257–3264, Nov. 2007.

[24] K. Buisman, L. K. Nanver, T. L. M. Scholtes, H. Schellevis, and L. C. N. de Vreede, "High-performance varactor diodes integrated in a silicon-on-glass technology," in *Proc. 35th Eur. Solid-State Device Res. Conf. (ESSDERC)*, Sep. 2005, pp. 117–120.

[25] L. Gao and G. M. Rebeiz, "A 0.97–1.53-GHz tunable four-pole bandpass filter with four transmission zeroes," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 3, pp. 195–197, Mar. 2019.

[26] A. Ebrahimi, T. Baum, J. Scott, and K. Ghorbani, "Continuously tunable dual-mode bandstop filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 5, pp. 419–421, May 2018.

[27] Y.-J. Zhang, J. Cai, and J.-X. Chen, "Wideband varactor-tuned balanced bandpass filter with stable passband selectivity," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 28, no. 9, Nov. 2018, Art. no. e21429.

[28] T. Zhang, Z. Cai, Y. Yang, J. Bao, and Y. Wang, "Compact tunable lowpass filter with sharp roll-off and low insertion loss," *Microw. Opt. Technol. Lett.*, vol. 59, no. 10, pp. 2619–2623, Oct. 2017.

[29] W. Qin, J. Cai, Y.-L. Li, and J.-X. Chen, "Wideband tunable bandpass filter using optimized varactor-loaded SIRs," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 9, pp. 812–814, Sep. 2017.

[30] Z.-H. Chen and Q.-X. Chu, "Dual-band reconfigurable bandpass filter with independently controlled passbands and constant absolute bandwidths," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 2, pp. 92–94, Feb. 2016.

[31] Z.-H. Chen, S.-X. Zhang, and Q.-X. Chu, "Dual-band bandstop filter with tunable lower stopband based on double-layer structure," *Microw. Opt. Technol. Lett.*, vol. 58, no. 9, pp. 2273–2276, Sep. 2016.

[32] W.-J. Zhou and J.-X. Chen, "High-selectivity tunable balanced bandpass filter with constant absolute bandwidth," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 64, no. 8, pp. 917–921, Aug. 2017.
[33] P.-L. Chi and T. Yang, "Novel 1.5-1.9 GHz tunable Single-to-Balanced bandpass filter with constant bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 12, pp. 972–974, Dec. 2016.
[34] Z.-H. Chen and Q.-X. Chu, "Reconfigurable bandpass filter with electric coupling and constant absolute bandwidth," *Microw. Opt. Technol. Lett.*, vol. 58, no. 6, pp. 1401–1404, Jun. 2016.
[35] N. Kumar and Y. K. Singh, "Compact constant bandwidth tunable wideband BPF with second harmonic suppression," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 11, pp. 870–872, Nov. 2016.
[36] C.-W. Tang and W.-C. Chen, "A compact tunable notch filter with wide constant absolute bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 25, no. 3, pp. 151–153, Mar. 2015.
[37] B. You, L. Chen, Y. Liang, and X. Wen, "A high-selectivity tunable dual-band bandpass filter using stub-loaded stepped-impedance resonators," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 11, pp. 736–738, Nov. 2014.
[38] X. Y. Zhang, C. H. Chan, Q. Xue, and B.-J. Hu, "RF tunable bandstop filters with constant bandwidth based on a doublet configuration," *IEEE Trans. Ind. Electron.*, vol. 59, no. 2, pp. 1257–1265, Feb. 2012.
[39] E. E. Djoumessi, M. Chaker, and K. Wu, "Varactor-tuned quarter-wavelength dual-bandpass filter," *IET Microw., Antennas Propag.*, vol. 3, no. 1, pp. 117–124, Feb. 2009.
[40] R. Gomez-Garcia, J.-M. Munoz-Ferreras, J. Jimenez-Campillo, F. Branca-Roncati, and P. Martin-Iglesias, "High-order planar bandpass filters with electronically-reconfigurable passband width and flatness based on adaptive multi-resonator cascades," *IEEE Access*, vol. 7, pp. 11010–11019, 2019.
[41] X. Huang, Q. Feng, and Q. Xiang, "Bandpass filter with tunable bandwidth using quadruple-mode stub-loaded resonator," *IEEE Microw. Wireless Compon. Lett.*, vol. 22, no. 4, pp. 176–178, Apr. 2012.
[42] B. Pal, M. K. Mandal, and S. Dwari, "Varactor tuned dual-band bandpass filter with independently tunable band positions," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 4, pp. 255–257, Apr. 2019.
[43] J. Guo, B. You, and G. Q. Luo, "A miniaturized eighth-mode substrate-integrated waveguide filter with both tunable center frequency and bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 7, pp. 450–452, Jul. 2019.
[44] D. Tian, Q. Feng, and Q. Xiang, "A fully tunable two-pole bandpass filter using the resonator with asymmetrical capacitances," *Microw. Opt. Technol. Lett.*, vol. 61, no. 12, pp. 2843–2846, Dec. 2019.
[45] B.-Z. Lan, H.-X. Li, C.-J. Guo, and J. Ding, "A 0.61-2.39 GHz fully tunable bandpass filter with minimum number of loaded varactors," *Microw. Opt. Technol. Lett.*, vol. 61, no. 2, pp. 469–473, 2019.
[46] B. Lan, C. Guo, and J. Ding, "A fully tunable two-pole bandpass filter with wide tuning range based on half mode substrate integrated waveguide," *Microw. Opt. Technol. Lett.*, vol. 60, no. 4, pp. 865–870, Apr. 2018.
[47] G. Zhang, Y. Xu, and X. Wang, "Compact tunable bandpass filter with wide tuning range of centre frequency and bandwidth using short coupled lines," *IEEE Access*, vol. 6, pp. 2962–2969, 2018.
[48] A. Kumar and N. P. Pathak, "Varactor-incorporated bandpass filter with reconfigurable frequency and bandwidth," *Microw. Opt. Technol. Lett.*, vol. 59, no. 8, pp. 2083–2089, Aug. 2017.
[49] S.-X. Zhang, Z.-H. Chen, Q.-X. Chu, and S. Member, "Compact tunable balanced bandpass filter with novel multi-mode resonator," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 1, pp. 43–45, Jan. 2017.
[50] B. You, S. Lu, L. Chen, and Q. J. Gu, "A half-mode substrate-integrated filter with tunable center frequency and reconfigurable bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 3, pp. 189–191, Mar. 2016.
[51] Z.-H. Chen and Q.-X. Chu, "Wideband fully tunable bandpass filter based on flexibly multi-mode tuning," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 10, pp. 789–791, Oct. 2016.
[52] X.-M. Zhao, Q.-Y. Wu, X.-J. Liu, and X.-W. Shi, "Compact varactor-tuned dual-mode filter with center frequency and bandwidth control," *Microw. Opt. Technol. Lett.*, vol. 58, no. 4, pp. 803–806, Apr. 2016.
[53] H. Zhu and A. M. Abbosh, "Tunable balanced bandpass filter with wide tuning range of center frequency and bandwidth using compact coupled-line resonator," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 1, pp. 7–9, Jan. 2016.
[54] X.-L. Zhao, L. Gao, and J.-X. Xu, "Tunable balanced bandpass filter with high common-mode suppression," *Electron. Lett.*, vol. 51, no. 24, pp. 2021–2023, Nov. 2015.
[55] P.-L. Chi, T. Yang, and T.-Y. Tsai, "A fully tunable two-pole bandpass filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 25, no. 5, pp. 292–294, May 2015.
[56] X. Luo, S. Sun, and R. B. Staszewski, "Tunable bandpass filter with two adjustable transmission poles and compensable coupling," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 9, pp. 2003–2013, Sep. 2014.
[57] J.-R. Mao, W.-W. Choi, K.-W. Tam, W. Q. Che, and Q. Xue, "Tunable bandpass filter design based on external quality factor tuning and multiple mode resonators for wideband applications," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 7, pp. 2574–2584, Jul. 2013.
[58] X. Huang, L. Zhu, Q. Feng, Q. Xiang, and D. Jia, "Tunable bandpass filter with independently controllable dual passbands," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 9, pp. 3200–3208, Sep. 2013.
[59] H.-J. Tsai, N.-W. Chen, and S.-K. Jeng, "Center frequency and bandwidth controllable microstrip bandpass filter design using loop-shaped dual-mode resonator," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 10, pp. 3590–3600, Oct. 2013.
[60] Y.-C. Chiou and G. M. Rebeiz, "Tunable 1.55–2.1 GHz 4-pole elliptic bandpass filter with bandwidth control and >50 dB rejection for wireless systems," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 117–124, Jan. 2013.
[61] C.-F. Chen, "A compact reconfigurable microstrip dual-band filter using varactor-tuned stub-loaded stepped-impedance resonators," *IEEE Microw. Wireless Compon. Lett.*, vol. 23, no. 1, pp. 16–18, Jan. 2013.
[62] A. L. C. Serrano, F. S. Correra, T.-P. Vuong, and P. Ferrari, "Synthesis methodology applied to a tunable patch filter with independent frequency and bandwidth control," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 484–493, Mar. 2012.
[63] G. Chaudhary, Y. Jeong, and J. Lim, "Dual-band bandpass filter with independently tunable center frequencies and bandwidths," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 107–116, Jan. 2013.
[64] Y.-C. Chiou and G. M. Rebeiz, "A quasi elliptic function 1.75–2.25 GHz 3-Pole bandpass filter with bandwidth control," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 2, pp. 244–249, Feb. 2012.
[65] Y.-C. Chiou and G. M. Rebeiz, "A tunable three-pole 1.5–2.2-GHz bandpass filter with bandwidth and transmission zero control," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 11, pp. 2872–2878, 2011.
[66] H. Joshi, H. H. Sigmarsson, S. Moon, D. Peroulis, and W. J. Chappell, "High-Q fully reconfigurable tunable bandpass filters," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 12, pp. 3525–3533, Dec. 2009.
[67] T. Hodson, B. Miao, C. Chen, A. Sharkawy, and D. Prather, "Silicon based photonic crystal electro-optic modulator utilizing the plasma dispersion effect," in *Proc. Conf. Lasers Electro-Optics (CLEO)*, May 2007, pp. 1–2.
[68] M. Donelli, M. Manekiya, and S. K. Menon, "Compact microstrip reconfigurable filter based on spiral resonators," *Microw. Opt. Technol. Lett.*, vol. 61, no. 2, pp. 417–424, Feb. 2019.
[69] S.-C. Weng, W.-H. Tu, and K.-W. Hsu, "Compact and switchable dual-band bandpass filter with high selectivity and wide stopband," *Electron. Lett.*, vol. 49, no. 20, pp. 1275–1277, Sep. 2013.
[70] S. Arain, P. Vryonides, M. A. B. Abbasi, A. Quddious, M. A. Antoniades, and S. Nikolaou, "Reconfigurable bandwidth bandpass filter with enhanced out-of-band rejection using π-section-loaded ring resonator," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 1, pp. 28–30, Jan. 2018.
[71] T. Cheng and K.-W. Tam, "A wideband bandpass filter with reconfigurable bandwidth based on cross-shaped resonator," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 10, pp. 909–911, Oct. 2017.
[72] P. Vryonides, S. Nikolaou, S. Kim, and M. M. Tentzeris, "Reconfigurable dual-mode band-pass filter with switchable bandwidth using PIN diodes," *Int. J. Microw. Wireless Technol.*, vol. 7, no. 6, pp. 655–660, Dec. 2015.
[73] W.-H. Tu, "Compact low-loss reconfigurable bandpass filter with switchable bandwidth," *IEEE Microw. Wireless Compon. Lett.*, vol. 20, no. 4, pp. 208–210, Apr. 2010.
[74] A. Bandyopadhyay, P. Sarkar, T. Mondal, and R. Ghatak, "A high selective tri-band bandpass filter with switchable passband states," *Int. J. Microw. Wireless Technol.*, vol. 12, no. 2, pp. 103–108, Mar. 2020.
[75] S. Kingsly, M. Kanagasabai, G. N. A. Mohammed, S. Subbaraj, Y. Panneer Selvam, and R. Natarajan, "Multi-band reconfigurable microwave filter using dual concentric resonators," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 28, no. 6, Aug. 2018, Art. no. e21290.
[76] F. Qiu, J. Huang, D. Lei, Z. Tang, and M. Yao, "Dual-band to tri-band to quad-band passband switchable bandpass filter," *Microw. Opt. Technol. Lett.*, vol. 59, no. 9, pp. 2307–2311, Sep. 2017.

[77] N. Kumar and Y. K. Singh, "Compact tri to dual passband switchable bandpass filter using stub-loaded split-ring resonator with improved bandwidth," *Electron. Lett.*, vol. 51, no. 19, pp. 1510–1512, Sep. 2015.
[78] J. Mazloum, A. Jalali, and M. Ojaroudi, "Miniaturized reconfigurable band-pass filter with electronically controllable for WiMAX/WLAN applications," *Microw. Opt. Technol. Lett.*, vol. 56, no. 2, pp. 509–512, Feb. 2014.
[79] P.-H. Deng and J.-H. Jheng, "A switched reconfigurable high-isolation dual-band bandpass filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 2, pp. 71–73, Feb. 2011.
[80] X.-J. Liao, H.-C. Yang, and N. Han, "A switchable bandpass filter using pin diodes on/off characteristics for WLAN application," *J. Electromagn. Waves Appl.*, vol. 25, no. 10, pp. 1402–1411, Jan. 2011.
[81] B. Lui, F. Wei, and X. Shi, "Switchable bandpass filter with two-state frequency responses," *Electron. Lett.*, vol. 47, no. 1, pp. 40–41, Jan. 2011.
[82] W. Feng, Y. Shang, W. Che, R. Gomez-Garcia, and Q. Xue, "Multifunctional reconfigurable filter using transversal signal-interaction concepts," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 11, pp. 980–982, Nov. 2017.
[83] H. Nachouane, A. Najid, A. Tribak, and F. Riouch, "A switchable bandstop-to-bandpass reconfigurable filter for cognitive radio applications," *Int. J. Microw. Wireless Technol.*, vol. 9, no. 4, pp. 765–772, May 2017.
[84] J. Xu, "A microstrip switchable filter with four operating modes," *IEEE Microw. Wireless Compon. Lett.*, vol. 26, no. 2, pp. 101–103, Feb. 2016.
[85] H. A. Mohamed, H. B. El-Shaarawy, E. A. F. Abdallah, and H. M. El-Hennawy, "Frequency-reconfigurable microstrip filter with dual-mode resonators using RF PIN diodes and DGS," *Int. J. Microw. Wireless Technol.*, vol. 7, no. 6, pp. 661–669, Dec. 2015.
[86] K. Rabbi and D. Budimir, "Highly selective reconfigurable filter for UWB systems," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 3, pp. 146–148, Mar. 2014.
[87] P.-H. Deng, J.-T. Tsai, and R.-C. Liu, "Design of a switchable microstrip dual-band lowpass-bandpass filter," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 9, pp. 599–601, Sep. 2014.
[88] H. Tang, J.-X. Chen, Y. Ge, Q.-Y. Lu, J. Shi, and Z.-H. Bao, "Novel LTCC switchable bandpass filter using vertically short-ended sir," *Microw. Opt. Technol. Lett.*, vol. 56, no. 4, pp. 999–1002, Apr. 2014.
[89] K. Rabbi, L. Athukorala, C. Panagamuwa, J. C. Vardaxoglou, and D. Budimir, "Highly linear microstrip wideband bandpass filter with switchable notched band for wireless applications," *Microw. Opt. Technol. Lett.*, vol. 55, no. 6, pp. 1331–1335, Jun. 2013.
[90] K. Rabbi, J. C. Vardaxoglou, C. Panagamuwa, D. Budimir, and L. Athukorala, "Compact UWB bandpass filter with reconfigurable notched band," *Electron. Lett.*, vol. 49, no. 11, pp. 709–711, May 2013.
[91] K. Song, S. Hu, Y. Fan, and Q. Xue, "Compact ultra-wideband (UWB) slotline bandpass filter with a switchable notch band," *J. Electromagn. Waves Appl.*, vol. 27, no. 6, pp. 691–699, Apr. 2013.
[92] K. Thirumalaivasan and R. Nakkeeran, "Development of UWB bandpass filter with reconfigurable notches to suppress IEEE 802.11 a/b services," *Microw. Opt. Technol. Lett.*, vol. 54, no. 10, pp. 2423–2426, Oct. 2012.
[93] M. Jung and B.-W. Min, "A widely tunable compact bandpass filter based on a switched varactor-tuned resonator," *IEEE Access*, vol. 7, pp. 95178–95185, 2019.
[94] Y.-J. Zhang, J. Cai, and J.-X. Chen, "Design of novel reconfigurable filter with simultaneously tunable and switchable passband," *IEEE Access*, vol. 7, pp. 59708–59715, 2019.
[95] C.-F. Chen, G.-Y. Wang, and J.-J. Li, "Microstrip switchable and fully tunable bandpass filter with continuous frequency tuning range," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 6, pp. 500–502, Jun. 2018.
[96] J. Shu, C. Cheng, C. Ge, and Y. Zhang, "A switchable absolute bandwidth tunable filter with etched slot," *J. Electromagn. Waves Appl.*, vol. 32, no. 14, pp. 1733–1747, Sep. 2018.
[97] F. Lin and M. Rais-Zadeh, "Continuously tunable 0.55–1.9-GHz bandpass filter with a constant bandwidth using switchable varactor-tuned resonators," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 3, pp. 792–803, Mar. 2017.
[98] L. Kurra, M. P. Abegaonkar, A. Basu, and S. K. Koul, "Switchable and tunable notch in ultra-wideband filter using electromagnetic bandgap structure," *IEEE Microw. Wireless Compon. Lett.*, vol. 24, no. 12, pp. 839–841, Dec. 2014.
[99] B. Liu, F. Wei, H. Zhang, X. Shi, and H. Lin, "A tunable bandpass filter with switchable bandwidth," *J. Electromagn. Waves Appl.*, vol. 25, nos. 2–3, pp. 223–232, Jan. 2011.
[100] M. Fan, K. Song, Y. Zhu, and Y. Fan, "Compact Bandpass-to-Bandstop reconfigurable filter with wide tuning range," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 3, pp. 198–200, Mar. 2019.
[101] S. S. Maragheh, M. Dousti, M. Dolatshahi, and B. Ghalamkari, "A dual-mode tunable bandpass filter for GSM, UMTS, WiFi, and WiMAX standards applications," *Int. J. Circuit Theory Appl.*, vol. 47, no. 4, pp. 561–571, Apr. 2019.
[102] K. Song, W. Chen, S. R. Patience, Y. Chen, A. M. Iman, and Y. Fan, "Compact wide-frequency tunable filter with switchable bandpass and bandstop frequency response," *IEEE Access*, vol. 7, pp. 47503–47508, 2019.
[103] F.-C. Chen, R.-S. Li, and J.-P. Chen, "A tunable dual-band Bandpass-to-Bandstop filter using p-i-n diodes and varactors," *IEEE Access*, vol. 6, pp. 46058–46065, 2018.
[104] N. Kumar and Y. K. Singh, "Compact tunable low-pass to CFBW bandpass switchable filter using concentric resonators," *IET Microw., Antennas Propag.*, vol. 12, no. 14, pp. 2225–2233, Nov. 2018.
[105] R.-S. Li and F.-C. Chen, "A tunable bandpass-to-bandstop filter with controllable bandwidth and high rejection level," *Microw. Opt. Technol. Lett.*, vol. 59, no. 1, pp. 110–113, Jan. 2017.
[106] F. A. Awin, Y. M. Alginahi, E. Abdel-Raheem, and K. Tepe, "Technical issues on cognitive radio-based Internet of things systems: A survey," *IEEE Access*, vol. 7, pp. 97887–97908, 2019.
[107] A. A. Khan, M. H. Rehmani, and A. Rachedi, "Cognitive-radio-based Internet of Things: Applications, architectures, spectrum related functionalities, and future research directions," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 17–25, Jun. 2017.

HASHINUR ISLAM (Graduate Student Member, IEEE) received the B.Tech. degree in electronics and telecommunication engineering from the Institute of Electronics and Telecommunication Engineers (IETE), India, in 2014, and the M.Tech. degree in digital electronics and communication engineering from the Sikkim Manipal Institute of Technology, Sikkim Manipal University, in 2018. He is currently pursuing the Ph.D. degree with Sikkim Manipal University.

He has published over 12 refereed papers in journals and conference proceedings. His research interests include multiband antennas, fractal antenna, circular polarization on antennas, feeding mechanism for antenna, PIFA antenna, wearable antenna, antenna for embedded systems, microstrip filters, and so on. He is an Associate Member of IETE, India.

SAUMYA DAS received the B.Tech. degree in electronics and telecommunication engineering from the Institute of Electronics and Telecommunication Engineers (IETE), India, and the M.E. degree in electronics and communication engineering from the Delhi College of Engineering, University of Delhi.

He has more than 15 years of teaching experience with expertise in various subjects, such as electromagnetic field theory, antenna theory, microwave devices and circuits, signals and system, digital signal processing, and adaptive signal processing. He is currently an Assistant Professor with the Department of Electronics and Communication Engineering, Sikkim Manipal Institute of Technology, India. He has published several articles in reputed peer-reviewed international journals and conferences. His research interests include different feeding techniques for microstrip and dielectric resonator antennas, flexible and wearable antenna for tracking and medical application, computational mathematics for antenna designing, RF exposure testing on human-bodies, electromagnetic compatibility, and ultra wideband antenna designing. He is an Associate Member of IETE, India. He serves as a Reviewer for *Microwave and Optical Technology Letters (MOTL)* (Wiley), and *IETE Journal of Research* (Taylor & Francis).

TANUSHREE BOSE received the Ph.D. degree in design of printed antennas using neural network from the Department of Electronics and Communication Engineering, BIT, Mesra, India, in 2011.

She worked in RF and microwave antennas with soft computation. She is currently an Associate Professor with the Sikkim Manipal Institute of Technology, India. She has published over 30 refereed papers in journals and conference proceedings. She has completed one Government (AICTE) Sponsored Research Project on Patch Antenna Designing Using Artificial Neural Network and successfully guided her research scholars. Her current research interests include multiband antennas for handsets and networks, miniature antennas, reconfigurable antennas, EBG application on conformal antennas, flexible and wearable antennas, smart antennas, and so on. She served on many national and international technical committees.

TANWEER ALI (Senior Member, IEEE) is currently an Assistant Professor-Senior Scale with the Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal. He is also an Active Researcher in microstrip antennas, wireless communication, and microwave imaging. He has published more than 90 papers in reputed peer-reviewed international journal and conferences. He is a member of the IEEE Antenna and Propagation Society and an Associate Member of IETE, India. He serves on the board of reviewers of journals, such as the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, the IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, IEEE ACCESS, the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES, the IEEE SENSORS, *IET Microwaves, Antennas and Propagation*, *Electronics Letters* (IET), *Wireless Personal Communication (WPC)* (Springer), *AEU-International Journal of Electronics and Communications*, *Microwave and Optical Technology Letters (MOTL)* (Wiley), the *International Journal of Antennas and Propagation* (Hindawi), *Advanced Electromagnetics*, *Progress in Electromagnetic Research (PIER)*, *KSII Transactions on Engineering Science*, Korea, the *International Journal of Microwave and Wireless Technologies*, *Frequenz*, *Radioengineering*, and so on.

End to End Binarized Neural Networks for Text Classification

Harshil Jain[1]*, **Akshat Agarwal**[2]*, **Kumar Shridhar**[3]*, **Denis Kleyko**[4,5]
[1]Computer Science and Engineering, IIT Gandhinagar, Gujarat, India
[2]Electrical Engineering, Delhi Technological University, Delhi, India
[3]NeuralSpace, London
[4]Redwood Center for Theoretical Neuroscience, University of California, Berkeley
[5]Intelligent Systems Lab, Research Institutes of Sweden
jain.harshil@iitgn.ac.in, akshat.agarwal0311@gmail.com
kumar@neuralspace.ai, denis.kleyko@ri.se

arXiv:2010.05223v1 [cs.LG] 11 Oct 2020

Abstract

Deep neural networks have demonstrated their superior performance in almost every Natural Language Processing task, however, their increasing complexity raises concerns. In particular, these networks require high expenses on computational hardware and training budget is a concern for many. Even for a trained network, the inference phase can be too demanding for resource-constrained devices, thus limiting its applicability. The state-of-the-art transformer models are a vivid example. Simplifying the computations performed by a network is one way of relaxing the complexity requirements. In this paper, we propose an end to end binarized neural network architecture for the intent classification task. In order to fully utilize the potential of end to end binarization, both input representations (vector embeddings of tokens statistics) and the classifier are binarized. We demonstrate the efficiency of such architecture on the intent classification of short texts over three datasets and for text classification with a larger dataset. The proposed architecture achieves comparable to the state-of-the-art results on standard intent classification datasets while utilizing $\sim$ 20-40% lesser memory and training time. Furthermore, the individual components of the architecture, such as binarized vector embeddings of documents or binarized classifiers, can be used separately with not necessarily fully binary architectures.

1 Introduction

In recent years, deep neural networks have achieved great success in a variety of domains including but not limited to autonomous driving (Grigorescu et al., 2020), image recognition (Krizhevsky et al., 2012), and text classification (Kim, 2014). However, deep networks are becoming more and more computationally expensive due to the ever-growing size of the models being trained. This tendency has been noticed in (Strubell et al., 2019) and it has been recommended that academia and industry researchers should draw their attention towards more computationally efficient methods. At the same time, many important application areas such as chatbots, IoT devices, mobile devices and other types of power-constrained and resource-constrained platforms require solutions that would be highly computation and memory efficient. Such use-cases limit the potential use of the state-of-the-art deep networks. One viable solution is the transformation of these high-performance neural networks to a more computationally efficient architecture. For example, continuous values of the network's parameters (e.g., float 32-64 bits) can be quantized into discrete integer values. Recently, Binarized Convolutional Neural Networks (BNN) (Hubara et al., 2016) have been developed where both weights and activations are restricted to $\{+1, -1\}$. This results in highly computationally efficient networks with a much lower memory footprint. However, to the best of our knowledge, in the area of text classification in general and intent classification, in particular, no end to end trainable binarized architecture has been demonstrated yet.

In this paper, we introduce an architecture for the task of intent classification that fully utilizes the power of binary representations. The input representations are tokenized and embedded in binary high-dimensional (HD) vectors forming distributed representations using the approach known as hyperdimensional computing (Kanerva, 2009), (Rachkovskij, 2001). The binary input representations are used

* The authors contributed equally to this research and work done at NeuralSpace

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: http://creativecommons.org/licenses/by/4.0/

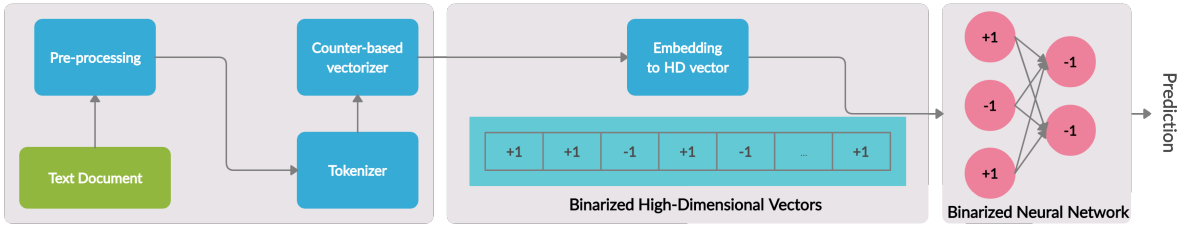


Figure 1: A schematic diagram of the end to end binarized classification architecture for text classification

for training an end to end BNN classifier for intent classification. Classification performance-wise, the binarized architecture achieves the results comparable to the state-of-the-art on three standard intent classification datasets (Braun et al., 2017) but performs less accurately on a large text classification 20NewsGroups (Lang, 1995) dataset. The efficiency of the proposed architecture is shown in terms of its time and memory complexity relative to non-binarized architectures. For example, it reduces the memory consumption by about 40% when compared to a non-binarized neural network using GloVe embeddings (Pennington et al., 2014) as the input representation.
Thus, the main contributions of this paper are as follows:

- We introduce a fully end to end binarized architecture for the task of intent classification that uses binary representations followed by the binarized classifier.
- We show that the proposed architecture preserves classification performance while being more time and memory-efficient compared to the non-binarized architecture.
- Finally, we show that the components of the architecture can be used individually as parts of solutions for other NLP tasks.

2 Related Work

2.1 Tokenization

Classical methods such as word-based tokenizers were one of the most popular tokenization methods to form a dictionary. They, however, suffer from out-of-dictionary problems for unseen words (Bojanowski et al., 2017). n-gram based tokenization estimates the out-of-dictionary words better. They can also be used to represent words (Joulin et al., 2016). As an alternative to word/token representation (Shridhar et al., 2019) introduced the idea of Subword Semantic Hashing that marks the start and end of a word with a specialized token and defines a hashing function to form n-gram-based subword tokens. Byte-Pair-Encoding (BPE) (Sennrich et al., 2015) is another popular subword segmentation method widely used in many neural machine translation systems. The main idea of the BPE is to split the sentence into individual tokens and to merge the most frequent tokens to get the final representation. The area was further extended by (Kudo, 2018) using a unigram language model. Most recently, a highly successful Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al., 2018) has proven to be a good tokenization choice where the tokenization is based on BERT's masking objective.

2.2 Representation of tokens

Conventional n-gram statistics uses sparse localist representations. In the special case, when $n = 1$, it can be seen as count-based vectors and compatible with Term Frequency Inverse Document Frequency (TF-IDF). Note that these representations are proportional to the length of the tokens' dictionary. Moreover, the length of representations grows exponentially with n, making them highly inefficient in terms of memory. Hyperdimensional computing-based embedding of n-gram statistics (Joshi et al., 2016), (Alonso et al., 2020) is used to represent the tokens extracted from the input text into a vector of fixed size, which does not depend on n. Previous works have shown the usefulness of these embeddings for the tasks of language identification (Rahimi et al., 2016), (Kleyko et al., 2019) and news articles

classification (Najafabadi et al., 2016). This work differs from the ones mentioned above, as in these studies, the embeddings were not combined with a trainable architecture. In contrast, here, the binary embeddings are used as a part of the architecture representing the input text, which enters a BNN.

As an alternative to n-gram based representations, dense trainable word embeddings (e.g., word2vec (Mikolov et al., 2013) or GloVe (Pennington et al., 2014)) use a fixed size vector representation. These embeddings are based on the distributional hypothesis and are expected to represent words with similar meanings nearby in the embedding space. However, the drawbacks of such embeddings are that they require pre-training on large text corpora and that storage of the trained embeddings demands a non-negligible memory footprint.

2.3 Classification models

In recent years, convolutional neural networks have been shown to be useful for NLP classification tasks at the level of sentences (Kim, 2014) when trained on top of the pre-trained word embeddings. Moreover, token-level convolutional networks (Zhang et al., 2015) have achieved state-of-the-art or competitive results on several large scale classification datasets. BNNs (Hubara et al., 2016), on the other hand, have been extensively used for computer vision tasks and their application to NLP tasks has been limited. One such example is (Zheng and Tang, 2016), where the binarized recurrent neural network has been used for the language modeling task. In this work, we show that BNNs are applicable for intent classification. We further show that the binarization capabilities are enhanced when using a fully binarized architecture, including embeddings and classifiers.

3 Methods

This section introduces a variety of methods used in the experiments. Figure 1 presents a schematic overview of the architecture. Given an input text document D, we first pre-process the document (see Section 4.2 for more details). The pre-processed document is then tokenized into the corresponding tokens $< T_1, T_2, ..., T_n >$ using one of the methods below. The tokens are used as an input to a count-based vectorizer. The representation of vectorizers, which is sparse and localist, is embedded into an HD vector (distributed representation) using hyperdimensional computing (Rahimi et al., 2017). Hyperdimensional computing is tightly related to randomly connected neural networks (Kleyko et al., 2020), (Kleyko et al., 2017) but it allows constructing more complex representations such as data structures (Osipov et al., 2017). Moreover, HD vector representing the counter's content can be binarized. It is used as an input to a classifier. The main classifier studied in this work is BNN but other classifiers are also considered for benchmarking.

3.1 Tokenizers

3.1.1 Word tokenizer

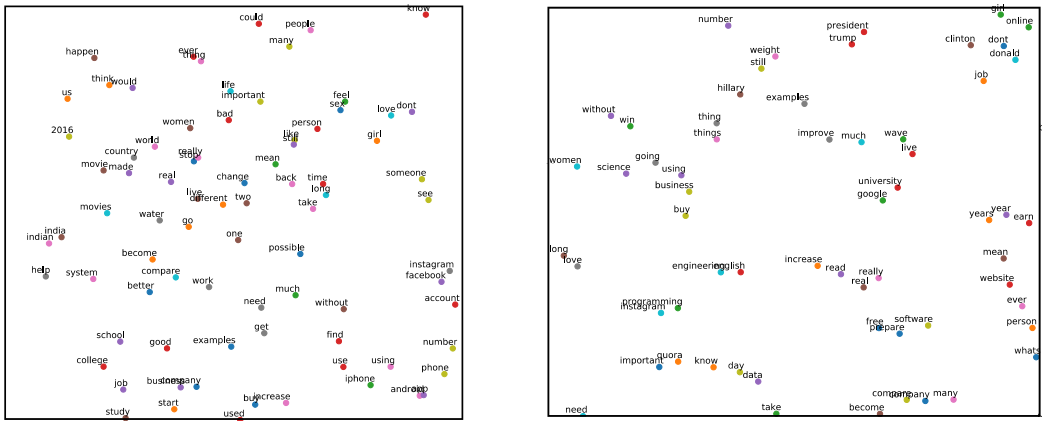
Given a pre-processed text D, a word tokenizer splits the input text D into tokes $< T_1, T_2, ..., T_n >$ corresponding to words. Words are split based on spaces while throwing away certain tokens, such as punctuation marks, unidentified tokens, hyphen symbols and others. For a sample text sentence D, [*"hello! how are you?"*], its word tokens would be [*"hello", "how", "are", "you"*].

3.1.2 Semantic Hashing tokenizer

First, Subword Semantic Hashing (SemHash) splits the input text D into word tokens $< T_1, T_2, ..., T_n >$. Next, for each word token T_i, SemHash creates subword tokens using a hashing method. For a given pre-processed input sample text D, e.g., [*"hello how are you"*], the output of the word split-step would be: [*"hello", "how", "are", "you"*]. A hashing function $\mathcal{H}(T_i)$ adds $\#$ at the beginning and end of each word token. Subwords are extracted for every token as n-grams for a given value of n, for example, when $n = 3$, $\mathcal{H}(hello)$=[*"#he", "hel", "ell", "llo", "lo#"*].

3.1.3 Byte Pair Encoding tokenizer

Given as input a pre-processed text D, BPE (Sennrich et al., 2015) splits the sentences into individual words with special end of word token $\dot{g}$ and merges the most frequent tokens to get the final tokens.



(a) t-SNE projection with GloVe embeddings

(b) t-SNE projection with HD embeddings

Figure 2: (a) shows the t-SNE projection with GloVe embeddings, while (b) shows the t-SNE projection with HD vectors; GloVe is pre-trained on the Quora question pairs dataset[2]while HD vectors are directly obtained from n-gram statistics of the words present in the plot. Colours of points are not associated to any meaning.

This is highly efficient for representing the unknown words into known frequent token sets. For a given input sample text D, e.g., [*"hello how are you"*], the output of the BPE tokenizer would be: [*"hello"*, *"how"*, *"are"*, *"you"*].

3.1.4 Char level Byte Pair Encoding tokenizer

For better subword sampling, (Kudo, 2018) proposed a new subword segmentation method based on a unigram language model. The unigram language model is based on an independent subword occurrence assumption with the probability of a subword sequence defined as the product of the subword occurrence probabilities. For a given input sample text D, e.g., [*"hello how are you"*], the output of the Char level BPE tokenizer would be: [*"h"*, *"e"*, *"l"*, *"l"*, *"o"*, *"h"*, *"o"*, *"w"*, *"a"*, *"r"*, *"e"*, *"y"*, *"o"*, *"u"*] where $\dot{g}$ is the special token to mark the end of the word.

3.1.5 BERT-based tokenizer

The BERT-based tokenizer is based on the BERT architecture, as proposed by (Devlin et al., 2018). It learns the token splits based on the masking objective on an unsupervised dataset. For a given input sample text D, e.g., [*"hello how are you"*], the output of the BERT-based tokenizer would be: [*"[cls]"*, *"hello"*, *"how"*, *"are"*, *"you"*, *"[sep]"*] where "[cls]" is the special token used to mark the start of the sentence and "[sep]" is the sentence separator.

3.2 High-Dimensional embedding of vectorized representations

The vectorized representations containing the statistics of tokens extracted from the pre-processed documents are usually sparse and have high dimensionality. In order to reduce the dimensionality of representations, we use hyperdimensional computing (Frady et al., 2020). First, each unique token T_i is assigned with a random d-dimensional bipolar HD vector, where d would be a hyperparameter of the method. HD vectors are stored in the item memory, which is a matrix $\mathbf{H} \in [d \times n]$. Thus, for a token T_i there is an HD vector $\mathbf{H}_{T_i} \in \{-1, +1\}^{[d \times 1]}$. To construct composite representations from the basic HD vectors stored in $\mathbf{H}$, hyperdimensional computing defines three key operations: permutation (ρ), binding ($\odot$, implemented via position-wise multiplication), and bundling ($+$, implemented via position-wise addition). For detailed properties and definitions of hyperdimensional computing operations, please see (Kanerva, 2009). The bundling operation allows storing information in HD vectors (Frady et al., 2018). Three operations above allow embedding vectorized representations based on n-gram statistics into an HD

[2] https://www.kaggle.com/c/quora-question-pairs

vector (Joshi et al., 2016). Note that a vectorizer that simply counts the number of times each token is present in the document is the special case of n-gram statistics for $n = 1$.

We first generate $\mathbf{H}$, which has an HD vector for each token. The permutation operation ρ is applied to $\mathbf{H}_{T_j}$ j times ($\rho^j(\mathbf{H}_{T_j})$) to represent a relative position of token T_j in an n-gram. A single HD vector corresponding to an n-gram (denoted as $\mathbf{m}$) is formed using the consecutive binding of permuted HD vectors $\rho^j(\mathbf{H}_{T_j})$ representing tokens in each position j of the n-gram. For example, the trigram '#he' will be embedded to an HD vector as follows: $\rho^1(\mathbf{H}_{\#}) \odot \rho^2(\mathbf{H}_{\mathrm{h}}) \odot \rho^3(\mathbf{H}_{\mathrm{e}})$. In general, the process of forming HD vector of an n-gram is formalized as follows:

$$\mathbf{m} = \prod_{j=1}^{n} \rho^j(H_{T_j}),$$

where T_j is token in jth position of the n-gram; the consecutive binding operations applied to n HD vectors are denoted by $\prod$. Once it is known how to form an HD vector for an individual n-gram, embedding the n-gram statistics into an HD vector $\mathbf{h}$ is achieved by bundling together all n-grams observed in the document:

$$\mathbf{h} = [\sum_{i=1}^{k} f_i \mathbf{m}_i = \sum_{i=1}^{k} f_i \prod_{j=1}^{n} \rho^j(H_{T_j})],$$

where k is the total number of n-grams; f_i is the frequency of ith n-gram and $\mathbf{m}_i$ is the HD vector of ith n-gram; $\sum$ denotes the bundling operation when applied to several HD vectors; $[*]$ denotes the binarization operation, which is implemented via the sign function. The usage of $[*]$ is optional, so we can either obtain binarized or non-binarized $\mathbf{h}$. If $\mathbf{h}$ is non-binarized, its components will be integers in the range $[-k, k]$ but these extreme values are highly unlikely since HD vectors for different n-grams are quasi-orthogonal, which means that in the simplest (but not practical) case when all n-grams have the same probability the expected value of a component in $\mathbf{h}$ is 0. Due to the use of $\sum$ for representing n-gram statistics, two HD vectors embedding two different n-gram statistics might have very different amplitudes if the frequencies in these statistics are very different. When HD vectors $\mathbf{h}$ are binarized, this issue is addressed. In the case of non-binarized HD vectors, we address it by using the cosine similarity, which is imposed by normalizing each $\mathbf{h}$ by its ℓ_2 norm, thus, all $\mathbf{h}$ have the same norm and their dot product is equivalent to their cosine similarity. Figure 2 provides an example of t-SNE projections of words represented either as GloVe embeddings or as embedding of n-gram statistics to HD vectors.

3.3 Binarized Neural Networks

Based on the work of (Hubara et al., 2016), we construct BNNs capable of working with representations of texts. To take the full advantage of binarized HD vectors, we constraint the weights and activations of the network layers to be $\{+1, -1\}$. This constraint is highly efficient in terms of hardware and memory, as bit-wise operations are used instead of multiply-accumulate operations. For example, a multiplication on binary values can be performed using an XNOR logical operation.

The vectorized representations of tokens embedded into HD vectors are binarized with all values $\{+1, -1\}$. In the case of HD vectors, we binarize the result of the bundling operation using the sign function. Similarly, the sign function is used in the BNN for every weight or activation to restrict them into $\{+1, -1\}$ as follows:

$$b(x) = [x] = \mathrm{sign}(x) = \begin{cases} +1 & \text{if } x \geq 0, \\ -1 & \text{otherwise} \end{cases} \tag{1}$$

where, x can be any weight or activation value.

We further define a convolutional 1D layer that creates a convolution kernel that is convolved with the input HD vector over a single spatial dimension to produce a tensor of outputs. Since gradient descent methods make small changes to the value of the weights, which cannot be done with binary values, we

	Chatbot		AskUbuntu		WebApplication		20NewsGroups	
Tokenizers	*Text-LeNet*	BNN	*Text-LeNet*	BNN	*Text-LeNet*	BNN	*Text-LeNet*	BNN
Word	**0.80**	0.73	0.51	**0.79**	0.56	**0.78**	0.54	**0.56**
SemHash	**0.94**	0.90	**0.87**	0.84	0.79	**0.83**	**0.78**	0.69
BPE	**0.80**	0.58	0.54	**0.67**	0.52	**0.75**	0.38	**0.42**
Char BPE	**0.92**	0.81	**0.76**	**0.76**	**0.55**	0.53	**0.55**	0.48
SentencePiece	0.80	**0.99**	0.70	**0.72**	0.50	**0.70**	0.41	**0.43**
BERT	**0.89**	0.88	**0.72**	0.71	0.70	**0.77**	**0.60**	**0.60**

Table 1: F_1 performance comparison of binarized *Text-LeNet* (BNN) architecture with non-binarized *Text-LeNet* for the task of intent classification on various datasets.

Datasets	Binarized GloVe	Binarized SemHash	Binarized HD vectors
Chatbot	0.74	0.91	**0.99**
AskUbuntu	0.86	**0.87**	0.84
WebApplication	0.66	0.80	**0.83**
20NewsGroups	0.62	0.64	**0.69**

Table 2: F_1 performance comparison of Binarized GloVe vectors, Binarized SemHash vectors and Binarized HD vectors. All vectorizers use the same binarized *Text-LeNet* architecture as classifier.

use the straight-through estimator (STE) idea, as mentioned in (Yin et al., 2019). We also define a value over which we clip the gradients in the backward pass:

$$\frac{\delta b(x)}{\delta x} = \begin{cases} +1 & \text{if } |x| < \text{clip value}, \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

This ensures that the entire architecture is end to end trainable using the gradient descent optimization.

4 Empirical Analysis

4.1 Datasets

All the experiments in this paper are performed on four datasets, namely: the *Chatbot Corpus* (Chatbot), the *Ask Ubuntu Corpus* (AskUbuntu), the *Web Applications Corpus* (WebApplication), and the *20 News Groups Corpus* (20NewsGroups). The Chatbot, AskUbuntu and WebApplication datasets are used for short text intent classification and were taken from (Braun et al., 2017). The datasets are available on GitHub[3]; 20NewsGroups dataset is available within the sklearn (Buitinck et al., 2013) library.

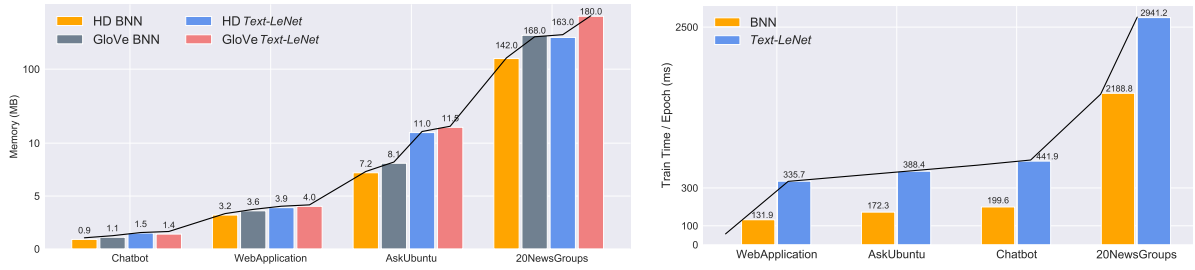
Chatbot Corpus: The Chatbot Corpus was collected from questions posted using a Telegram Chatbot. The classification task was performed on 2 intents: *(Departure Time and Find Connection)* with 206 questions. The train set consists of 100 samples, while 106 samples were used as the test set.

AskUbuntu Corpus: The AskUbuntu Corpus has been extracted from the AskUbuntu platform. It consists of 5 intents *(Make Update, Setup Printer, Shutdown Computer, Software Recommendation, and None)* with 53 training samples and 109 test samples.

WebApplication Corpus: The WebApplication Corpus consists of 88 samples and 8 intents *(Change Password, Delete Account, Download Video, Export Data, Filter Spam, Find Alternative, Sync Accounts, and None)* with 30 training samples and 58 test samples.

20NewsGroups Corpus: The 20NewsGroups consists of news post categorized into 20 categories. Each category has exactly 18846 text samples with 11314 samples for training and 7532 samples for testing. A detailed data distribution for each dataset is provided in the Appendix.

[3]Under the Creative Commons CC BY-SA 3.0 license: `https://github.com/sebischair/NLU-Evaluation-Corpora`



(a) Memory Comparison

(b) Train Time per epoch Comparison

Figure 3: (a) shows the memory comparisons for all 4 datasets using HD *Text-LeNet*, HD BNN, GloVe *Text-LeNet*, GloVe BNN and (b) shows the training time per epoch comparison for all 4 datasets using BNN and *Text-LeNet*

4.2 Experimental Details

On all datasets, six tokenization methods were used, namely: Word-based, SemHash, BPE, Char level BPE, Sentence Piece and BERT-based. For Word-based tokenizer, the datasets were pre-processed using the Spacy library to remove stop words from data. Spacy pre-trained model "en_core_web_lg" was used to parse the datasets. All text samples were also pre-processed by removing control tokens, in particular, the ones in the set *[Cc]*, which includes Unicode tokens from *U+0000* to *U+009F*. The small datasets (Chatbot, WebApplication and AskUbuntu) data samples within each class were made equal to the largest class sample size by augmenting the data. For SemHash, n value was kept at 3, and all trigrams were considered. For BPE, a dictionary size of 1000 was used for Chatbot, WebApplication and AskUbuntu, and 2000 for 20NewsCorpus dataset because of its larger size. Similar values were used for Char-based BPE and Sentence Piece tokenizers. The difference between Sentence Piece and BPE is that Sentence Piece uses stop words removal before tokenization while BPE and Char-based BPE do not. BERT-based tokenizer was pre-trained on a large corpus "bert-large-cased-vocab.txt"[4] that contains 29213 unique words and has been provided by the Huggingface[5] library.

For the purpose of benchmarking, nine sklearn classifiers were applied to the intent classification datasets: MLP, Random Forest, Linear SVC, Passive Aggressive, SGD Classifier[6], Ridge Classifier, Nearest Centroid, Bernoulli NB, KNN Classifier with an HD vector size of $d = \{512, 1024, 4098, 8192, 16384\}$. The detailed results for these classifiers are provided in the Appendix. CNN-based architecture used $d = \{512, 1024, 4098, 8192\}$ as embedding sizes. 20NewsGroups dataset was run with MLP, Random Forest, Linear SVC and SGD classifiers for embedding sizes $d = \{512, 1024, 4096, 16384\}$ and with CNN for sizes $d = \{512, 1024, 4096\}$.

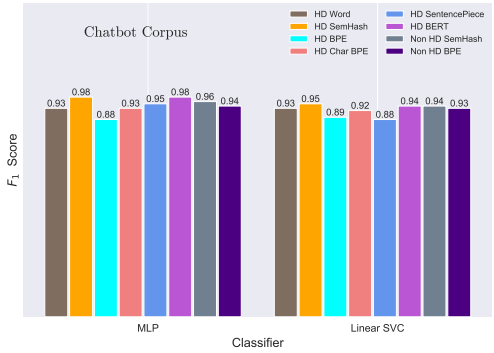
For CNN-based architecture, 5 hidden layers were used: 3 convolutional 1D layers followed by 2 dense layers. Due to its resemblance to the original LeNet architecture (LeCun et al., 1998), we refer to this architecture as to *Text-LeNet*. The detailed configuration is provided in the Appendix. We noted in the experiments that Batch Normalization (Ioffe and Szegedy, 2015) is very helpful for faster convergence of the network and, therefore, it was added after every convolutional and dense layer. The clip value for the gradients was set to 1 during the backward pass. RMSProp (Hinton et al., 2012) was used as the optimizer with a learning rate of 1e-3 and categorical cross-entropy as the loss function. All the checkpoints were saved in the h5 format. Larq Compute Engine (Geiger and Team, 2020), a highly optimised engine for quantization of networks, was used to convert the h5 files to tflite format for BNNs.

For classifiers, that were a part of the grid-based search, the paper reports the results for the best hyperparameters. For all other classifiers the default hyperparameter settings provided by the sklearn library were used. A 5-fold cross-validation was used in the experiments. A total of 5 simulations were performed and the average results are reported in the paper. All sklearn-based classifier experiments

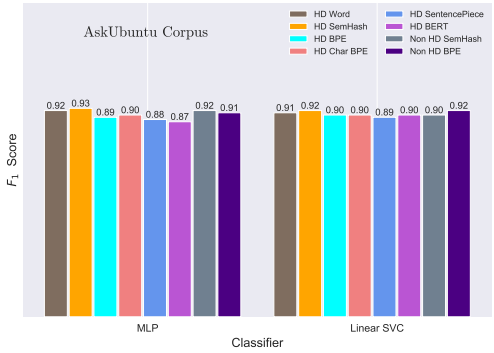
[4] `https://s3.amazonaws.com/models.huggingface.co/bert/bert-large-cased-vocab.txt`

[5] `https://github.com/huggingface/`

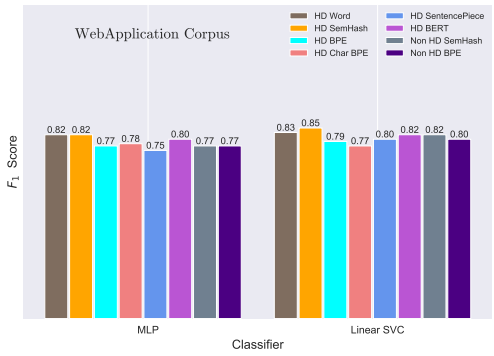
[6] The SGD classifier here refers to the SVM classifier trained using SGD optimization as per the sklearn library and this notation is used henceforth.



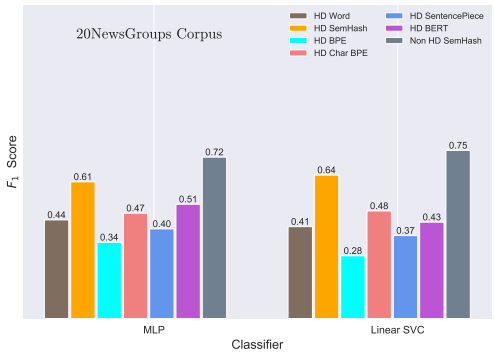
(a) Chatbot Corpus



(b) AskUbuntu Corpus



(c) WebApplication Corpus



(d) 20NewsGroups Corpus

Figure 4: (a), (b), (c) and (d) shows the F_1 score comparison of MLP and Linear SVC classifier with HD and non-HD based tokenizers on Chatbot, AskUbuntu, WebApplication and 20NewsGroups corpus respectively.

were performed on the CPU and CNN-based experiments were performed on NVIDIA Tesla GPUs.

4.3 Results and Discussions

To demonstrate the efficiency of binarized architecture for the intent classification task, we compare the results of binarized HD vectors with the binarized *Text-LeNet* (BNN) architecture as classifier against non-binarized HD vectors with non-binarized *Text-LeNet*. The F_1 scores are compared in Table 1 where BNN performed equally well to a *Text-LeNet* architecture while being 20% to 40% more memory efficient, as shown in Figure 3(a). Due to the specifics of implementation, BNNs use 32 bit float values as *Text-LeNet*. The memory efficiency of BNNs can be further improved by 4x when 8-bit representations are used. On performance side, BNNs outperforms the *Text-LeNet* for AskUbuntu and WebApplication datasets on 4 out of 6 tokenizers. The results reported in Table 1 used 512 dimensional HD vectors for Chatbot, AskUbuntu and WebApplication corpus, while 1024 dimensional HD vectors were used for 20NewsGroups dataset. The choice of dimensionality was based on the size of the datasets and a larger dimensionality was used for larger datasets.

One thing to note here is that *Text-LeNet* also used HD vectors with the mentioned tokenizers, but the HD vectors were non-binarized. HD vectors in itself are already faster and much more efficient than counter-based representations, as shown in (Alonso et al., 2020). When experimenting with other embedding methods like GloVe, the training was significantly slower; therefore, HD vectors were used for all the experiments. In addition to that, using binarized classifier (BNN) further improves the training time upto 50% compared to non-binarized classifier on all four datasets as shown in Figure 3 (b). Furthermore, when compared to GloVe embeddings with *Text-LeNet*, HD BNN used around 20 - 40%

Platform	Chatbot	AskUbuntu	WebApplication	Average
Botfuel	0.98	0.90	0.80	0.89
Luis	0.98	0.90	0.81	0.90
Dialogflow	0.93	0.85	0.80	0.86
Watson	0.97	0.92	0.83	0.91
Rasa	0.98	0.86	0.74	0.86
Snips	0.96	0.83	0.78	0.86
Recast	**0.99**	0.86	0.75	0.87
TildeCNN	**0.99**	0.92	0.81	0.91
FastText	0.97	0.91	0.76	0.88
SemHash	0.96	0.92	**0.87**	**0.92**
BPE	0.95	**0.93**	0.85	0.91
HD vectors (Alonso et al., 2020)	0.97	0.92	0.82	0.90
Binarized HD vectors with the best classifier	0.98	**0.93**	0.84	**0.92**
HD *Text-LeNet*	0.94	0.87	0.79	0.88
HD BNN	**0.99**	0.84	0.83	0.91

Table 3: F_1 score comparison of various platforms on intent classification datasets of short texts with methods used in the paper. Some results are taken from (Alonso et al., 2020)

lesser memory for all the intent classification datasets.

To justify that the binarization of HD vectors is effective and needed in this work, we benchmarked the binarized HD vectors with binarized 300-dimensional GloVe vectors and the binarized version of counter-based representation for SemHash tokenizer for all the datasets. Table 2 summarizes the results of the comparison. All the binarized representations were trained with the same BNN classifier. Binarized HD vectors performed significantly better than other binarized methods outperforming binarized GloVe by 4 - 25% and binarized SemHash by 2 - 8% on 2 out of 3 smaller intent classification datasets and achieves comparable results for AskUbuntu dataset. The trend continued for 20NewsGroups with binarized HD achieving 5 - 7% better F_1 scores. Note that for the SemHash counter-based vectorizer, we put a sign function $sign(x) = +1$ for $x > 0$ and -1 otherwise.

In Figure 4, MLP and Linear SVC with all the tokenizers with HD vectors as representation are compared with MLP and Linear SVC classifiers with SemHash tokenizers and counter-based vectorizer as representation from (Alonso et al., 2020). The F_1 score is comparable to the state-of-the-art for both MLP and SVC. For all small intent classification datasets, binarized HD vectors have achieved better results than non-HD vectors. The proposed architecture beats the non-HD baselines by $+2\%$ for AskUbuntu and Chatbot Corpus, and $+5\%$ for WebApplication Corpus. However, for 20NewsGroups the results of binarized HD Vectors are lower than non-HD Vectors. This is mainly due to the large size of the dataset and simple classifiers like LinearSVC failed to perform with just binarized values. The results for all the other classifiers are provided in the Appendix.

Table 3 compares the F_1 scores of various platforms on the intent classification datasets. We report the results of binarized HD vectors with the best classifiers from one of the nine classifiers mentioned (Binarized HD vectors with the best classifier), non-binarized HD vectors with *Text-LeNet* (HD *Text-LeNet*) and binarized HD vectors with binarized *Text-LeNet* (HD BNN). Our end to end binarized architecture (HD BNN) achieved the state-of-the-art results for Chatbot dataset. The approach where only HD vectors were binarized (binarized HD vectors with the best classifier) achieved the state-of-the-art results for AskUbuntu dataset. The results on WebApplication dataset are comparable to the state-of-the-art (0.87 with SemHash): 0.84 for binarized HD vectors with the best classifier and 0.83 for HD BNN. The average performance of both binarized HD vectors with the best classifier (0.92) and HD BNN (0.91) was also comparable to the best non-binarized approach (0.92 with Semhash).

5 Conclusion

In this work, we present an end to end binarized architecture for the intent classification task. We demonstrate that it is possible to achieve comparable to the state-of-the-art results while using the binarized representations of all the components of the NLP classification architecture. This allows exploring the effectiveness of binary representations both for reducing the memory footprint of the architecture and for increasing the energy-efficiency of the inference phase due to effectiveness of binary operations. This work takes a step towards enabling NLP functionality on resource-constrained devices.

Acknowledgement

The authors would like to thank P. Alonso, E. Osipov, and M. Liwicki with whom the previous research, which is the foundation of this work, has been done. The work of DK was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Individual Fellowship Grant Agreement 839179 and in part by the DARPA's VIP (Super-HD Project) and AIE (HyDDENN Project) programs.

References

P. Alonso, K. Shridhar, D. Kleyko, E. Osipov, and M. Liwicki. 2020. HyperEmbed: Tradeoffs between Resources and Performance in NLP Tasks with Hyperdimensional Computing Enabled Embedding of n-gram Statistics. *arXiv:2003.01821*.

P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

D. Braun, A. Hernandez-Mendez, F. Matthes, and M. Langen. 2017. Evaluating Natural Language Understanding Services for Conversational Question Answering Systems. In *Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 174–185.

L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux. 2013. API Design for Machine Learning Software: Experiences from the Scikit-learn Project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805*.

E. P. Frady, D. Kleyko, and F. T. Sommer. 2018. A Theory of Sequence Indexing and Working Memory in Recurrent Neural Networks. *Neural Computation*, 30:1449–1513.

E. P. Frady, D. Kleyko, and F. T. Sommer. 2020. Variable Binding for Sparse Distributed Representations: Theory and Applications. *arXiv:2009.06734*, pages 1–16.

L. Geiger and P. Team. 2020. Larq: An Open-Source Library for Training Binarized Neural Networks. *Journal of Open Source Software*, 5(45):1746, January.

S. Grigorescu, B. Trasnea, T. Cocias, and G. Macesanu. 2020. A Survey of Deep Learning Techniques for Autonomous Driving. *Journal of Field Robotics*, 37(3):362–386.

G. Hinton, N. Srivastava, and K. Swersky. 2012. Neural Networks for Machine Learning Lecture 6a Overview of Mini-batch Gradient Descent.

I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio. 2016. Binarized Neural Networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1–9.

S. Ioffe and C. Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167.

A. Joshi, J. T. Halseth, and P. Kanerva. 2016. Language Geometry Using Random Indexing. In *Quantum Interaction (QI)*, pages 265–274.

A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov. 2016. Bag of Tricks for Efficient Text Classification. *arXiv:1607.01759*.

P. Kanerva. 2009. Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognitive Computation*, 1(2):139–159.

Y. Kim. 2014. Convolutional Neural Networks for Sentence Classification. *arXiv:1408.5882*.

D. Kleyko, E. P. Frady, and E. Osipov. 2017. Integer Echo State Networks: Hyperdimensional Reservoir Computing. *arXiv:1706.00280*, pages 1–10.

D. Kleyko, E. Osipov, D. De Silva, U. Wiklund, V. Vyatkin, and D. Alahakoon. 2019. Distributed Representation of n-gram Statistics for Boosting Self-Organizing Maps with Hyperdimensional Computing. In *International Andrei Ershov Memorial Conference on Perspectives of System Informatics (PSI)*, pages 64–79.

D. Kleyko, M. Kheffache, E. P. Frady, U. Wiklund, and E. Osipov. 2020. Density Encoding Enables Resource-Efficient Randomly Connected Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, PP(99):1–7.

A. Krizhevsky, I. Sutskever, and G. E. Hinton. 2012. Imagenet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1097–1105.

T. Kudo. 2018. Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates. *arXiv:1804.10959*.

K. Lang. 1995. Newsweeder: Learning to filter Netnews. In *Twelfth International Conference on Machine Learning*, pages 331–339.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

T. Mikolov, K. Chen, G. Corrado, and J. Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781*.

F. R. Najafabadi, A. Rahimi, P. Kanerva, and J. M. Rabaey. 2016. Hyperdimensional Computing for Text Classification. In *Design, Automation and Test in Europe Conference (DATE)*, pages 1–1.

E. Osipov, D. Kleyko, and A. Legalov. 2017. Associative Synthesis of Finite State Automata Model of a Controlled Object with Hyperdimensional Computing. In *Annual Conference of the IEEE Industrial Electronics Society (IECON)*, pages 3276–3281.

J. Pennington, R. Socher, and C. D. Manning. 2014. GloVe: Global Vectors for Word Representation. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, volume 14, pages 1532–1543.

D. A. Rachkovskij. 2001. Representation and Processing of Structures with Binary Sparse Distributed Codes. *IEEE Transactions on Knowledge and Data Engineering*, 3(2):261–276.

A. Rahimi, P. Kanerva, and J. M. Rabaey. 2016. A Robust and Energy Efficient Classifier Using Brain-Inspired Hyperdimensional Computing. In *IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*, pages 64–69.

A. Rahimi, S. Datta, D. Kleyko, E. P. Frady, B. Olshausen, P. Kanerva, and J. M. Rabaey. 2017. High-dimensional Computing as a Nanoscalable Paradigm. *IEEE Transactions on Circuits and Systems I*, 64(9):2508–2521.

R. Sennrich, B. Haddow, and A. Birch. 2015. Neural Machine Translation of Rare Words with Subword Units. *arXiv:1508.07909*.

K. Shridhar, A. Dash, A. Sahu, G. Grund Pihlgren, P. Alonso, V. Pondenkandath, G. Kovacs, F. Simistira, and M. Liwicki. 2019. Subword Semantic Hashing for Intent Classification on Small Datasets. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–6.

E. Strubell, A. Ganesh, and A. McCallum. 2019. Energy and Policy Considerations for Deep Learning in NLP. In *57th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 3645–3650.

P. Yin, J. Lyu, S. Zhang, S. Osher, Y. Qi, and J. Xin. 2019. Understanding Straight-through Estimator in Training Activation Quantized Neural Nets. *arXiv:1903.05662*.

X. Zhang, J. Zhao, and Y. LeCun. 2015. Character-level Convolutional Networks for Text Classification. In *Advances in Neural Information Processing Systems (NIPS)*, pages 649–657. Curran Associates, Inc.

W. Zheng and Y. Tang. 2016. Binarized Neural Networks for Language Modeling. *Technical Report cs224d, Stanford University*.

A Appendix

A.1 Datasets

The experiments in the paper were performed on four datasets: the *Chatbot Corpus* (Chatbot), the *Ask Ubuntu Corpus* (AskUbuntu), the *Web Applications Corpus* (WebApplication), and the *20 News Groups Corpus* (20NewsGroups).

A chatbot was created on Telegram, where questions of the public transport of Munich were posted. The chatbot replied to the questions and thus data was collected for the Chatbot corpus. A detailed test and train split are provided in Table 1.

Intent	Train original	Train Augmented	Test
Departure Time	43	57	35
Find Connection	57	57	71

Table 1: Data sample distribution for the Chatbot dataset

The AskUbuntu and WebApplication datasets are questions and answers from the StackExchange platform. A detailed breakdown is provided in Table 2 and 3 respectively.

Intent	Train original	Train Augmented	Test
Make Update	10	17	37
Setup Printer	10	17	13
Shutdown Computer	13	17	14
Software Recommendation	17	17	40
None	3	17	5

Table 2: Data sample distribution for the AskUbuntu dataset

Intent	Train original	Train Augmented	Test
Change Password	2	7	6
Delete Account	7	7	10
Download Video	1	7	0
Export Data	2	7	3
Filter Spam	6	7	14
Find Alternative	7	7	16
Sync Accounts	3	7	6
None	2	7	4

Table 3: Data sample distribution for the WebApplication dataset

The 20NewsGroups dataset comprises news posts labelled into several categories and a detailed breakdown is provided in Table 4.

A.2 Experimentation Details

The *Text-LeNet* architecture used in the experiments is defined as follows:

Text-LeNet : $[128, M, BN, 256, M, BN, 512, M, F, 128_D, DO, C]$

where, numbers 128, 256 and 512 represents the filters of *Convolution layer* which is followed by an activation function. *M* represents the *Max Pooling layer* and *BN* represents the *Batch Normalization layer*. *F* refers to a *Flatten layer*. 128_D represents the *Dense layer* of size 128 followed by a *Dropout layer* denoted by *DO*. Finally, *C* represents the *Linear classification layer* of dimension (128, number of classes).

Categories	Train	Test
alt.atheism	11314	7532
comp.graphics	11314	7532
comp.os.ms-windows.misc	11314	7532
comp.sys.ibm.pc.hardware	11314	7532
comp.sys.mac.hardware	11314	7532
comp.windows.x	11314	7532
misc.forsale	11314	7532
rec.autos	11314	7532
rec.motorcycles	11314	7532
rec.sport.baseball	11314	7532
rec.sport.hockey	11314	7532
sci.crypt	11314	7532
sci.electronics	11314	7532
sci.electronics	11314	7532
sci.space	11314	7532
soc.religion.christian	11314	7532
talk.politics.guns	11314	7532
talk.politics.mideast	11314	7532
talk.politics.misc	11314	7532
talk.religion.misc	11314	7532

Table 4: Data sample distribution for the 20NewsGroups dataset

Other hyper-parameters settings include:

Hyper-parameter	**Value**
Convolution Kernel Size	3
Convolution layer Padding	valid
Max-Pooling Kernel Size	3,2,3 for the three M layers respectively
Optimizer	RMSprop
Loss	Categorical cross entropy
Activation Function	Rectified Linear Unit (ReLU)
Batch Size	4 (small datasets) and 64 (20NewsGroups)
Learning Rate	0.001
Number of Epochs	15-40 (Depending upon the dataset)
Initializer	Xavier initialization

Table 5: Hyper-parameters for the experiments

To find the best hyperparameters, a grid-based search was applied to all the datasets for the MLP classifier. Three different hidden layer configurations were considered: $[(100, 50), (300, 100), (300, 200, 100)]$. The best parameters were chosen based on `GridSearchCV`. The maximal number of iterations for MLP was set to 200.

In the case of Random Forest, two hyperparameters include optimized number of estimators ($[50, 60, 70]$) and minimum samples leaf ($[1, 11]$); we used 50 estimators and 1 leaf. In the case of KNN, the number of neighbors between 3 and 7 was considered; 3 neighbors were used in the experiments. For all the other classifiers the default hyperparameter settings provided by the sklearn library were used.

A.3 Experiments

Table 6-8 compares the F_1 scores of nine classifiers: MLP, Random Forest, Linear SVC, Passive Aggressive, SGD Classifier, Ridge Classifier, Nearest Centroid, Bernoulli NB, KNN Classifier on all three small datasets with HD versions of six tokenization methods and Non HD versions of SemHash and BPE for small datasets and Non HD SemHash for 20NewsGroups Corpus. SP is an acronym for Sentence Piece. Table 9 compares the results of all the tokenizers using four classifiers: MLP, SGD Classifier, Linear SVC and Random Forest for the 20NewsGroups dataset. SemHash tokenizer, in general, achieved better results compared to other tokenizers followed by BERT tokenizer on all four datasets.

Classifier	HD Word	HD SemHash	HD BPE	HD Char BPE	HD SP	HD BERT	Non HD SemHash	Non HD BPE
MLP	0.92	**0.93**	0.89	0.90	0.88	0.87	0.92	0.91
Passive Aggr.	0.92	**0.93**	0.90	0.88	0.91	0.90	0.92	**0.93**
SGD Classifier	0.86	**0.89**	0.85	0.88	0.86	**0.89**	**0.89**	**0.89**
Ridge Classifier	**0.92**	**0.92**	0.88	0.87	0.91	**0.92**	0.90	0.91
KNN Classifier	0.81	**0.84**	0.78	0.60	0.80	0.79	0.79	0.72
Nearest Centroid	**0.91**	**0.91**	0.88	0.86	0.86	0.85	0.90	0.89
Linear SVC	0.91	**0.92**	0.90	0.90	0.89	0.90	0.90	**0.92**
Random Forest	0.90	**0.92**	0.82	0.83	0.83	0.82	0.88	0.90
Bernoulli NB	0.76	0.87	0.74	0.84	0.65	0.79	0.91	**0.92**

Table 6: F_1 scores of all sklearn classifiers for AskUbuntu dataset.

Classifier	HD Word	HD SemHash	HD BPE	HD Char BPE	HD SP	HD BERT	Non HD SemHash	Non HD BPE
MLP	0.93	**0.98**	0.88	0.93	0.95	**0.98**	0.96	0.94
Passive Aggr.	**0.96**	**0.96**	0.92	0.94	0.89	0.94	0.95	0.91
SGD Classifier	0.95	**0.97**	0.93	0.89	0.89	0.92	0.93	0.93
Ridge Classifier	0.94	**0.95**	0.77	0.91	0.84	0.88	0.94	0.94
KNN Classifier	0.86	0.86	0.80	0.83	0.80	**0.91**	0.75	0.71
Nearest Centroid	0.79	0.89	0.90	0.86	0.88	0.88	0.89	**0.94**
Linear SVC	0.93	**0.95**	0.89	0.92	0.88	0.94	0.94	0.93
Random Forest	0.88	0.91	0.77	0.83	0.85	0.92	**0.95**	**0.95**
Bernoulli NB	0.81	0.88	0.81	0.89	0.84	0.91	**0.93**	**0.93**

Table 7: F_1 scores of all sklearn classifiers for Chatbot dataset.

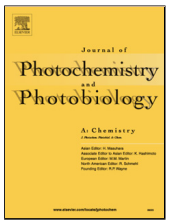
Classifier	HD Word	HD SemHash	HD BPE	HD Char BPE	HD SP	HD BERT	Non HD SemHash	Non HD BPE
MLP	**0.82**	**0.82**	0.77	0.78	0.75	0.80	0.77	0.77
Passive Aggr.	**0.83**	**0.83**	0.76	0.80	0.80	0.80	0.82	0.80
SGD Classifier	0.76	0.79	0.66	0.64	0.64	**0.84**	0.75	0.74
Ridge Classifier	0.79	**0.84**	0.79	0.78	0.76	0.81	0.79	0.80
KNN Classifier	0.80	0.81	0.78	0.67	0.80	**0.84**	0.72	0.75
Nearest Centroid	0.77	**0.80**	0.72	0.76	0.75	0.75	0.74	0.73
Linear SVC	0.83	**0.85**	0.79	0.77	0.80	0.82	0.82	0.80
Random Forest	0.77	0.80	0.61	0.68	0.67	0.78	**0.87**	0.85
Bernoulli NB	0.61	0.70	0.57	0.62	0.57	0.61	0.74	**0.75**

Table 8: F_1 scores of all sklearn classifiers for WebApplication dataset.

Classifier	HD Word	HD SemHash	HD BPE	HD Char BPE	HD SP	HD BERT	Non HD SemHash
MLP	0.44	0.61	0.34	0.47	0.40	0.51	**0.72**
SGD Classifier	0.43	0.59	0.41	0.26	0.45	0.49	**0.70**
Linear SVC	0.41	0.64	0.28	0.48	0.37	0.43	**0.75**
Random Forest	0.17	0.33	0.13	0.20	0.13	0.21	**0.58**

Table 9: F_1 scores of sklearn classifiers for 20NewsGroups dataset.

Journal Pre-proof



Experimental and Theoretical Interpretations of Spectral Behavior of 6-Methoxyflavone

Nisha Fatma, Mohan Singh Mehata, Nupur Pandey, Sanjay Pant

PII: S1010-6030(20)30742-5

DOI: https://doi.org/10.1016/j.jphotochem.2020.112945

Reference: JPC 112945

To appear in: *Journal of Photochemistry & Photobiology, A: Chemistry*

Received Date: 3 August 2020

Revised Date: 15 September 2020

Accepted Date: 22 September 2020

Please cite this article as: Fatma N, Mehata MS, Pandey N, Pant S, Experimental and Theoretical Interpretations of Spectral Behavior of 6-Methoxyflavone, *Journal of Photochemistry and amp; Photobiology, A: Chemistry* (2020), doi: https://doi.org/10.1016/j.jphotochem.2020.112945

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

Journal Pre-proof

Experimental and Theoretical Interpretations of Spectral Behavior of 6-Methoxyflavone

Nisha Fatma[1], Mohan Singh Mehata[2*], Nupur Pandey[1], Sanjay Pant[1]

[1]Photophysics Laboratory, Department of Physics, Centre of Advance Study, DSB Campus, Kumaun University, Nainital-263002, India.

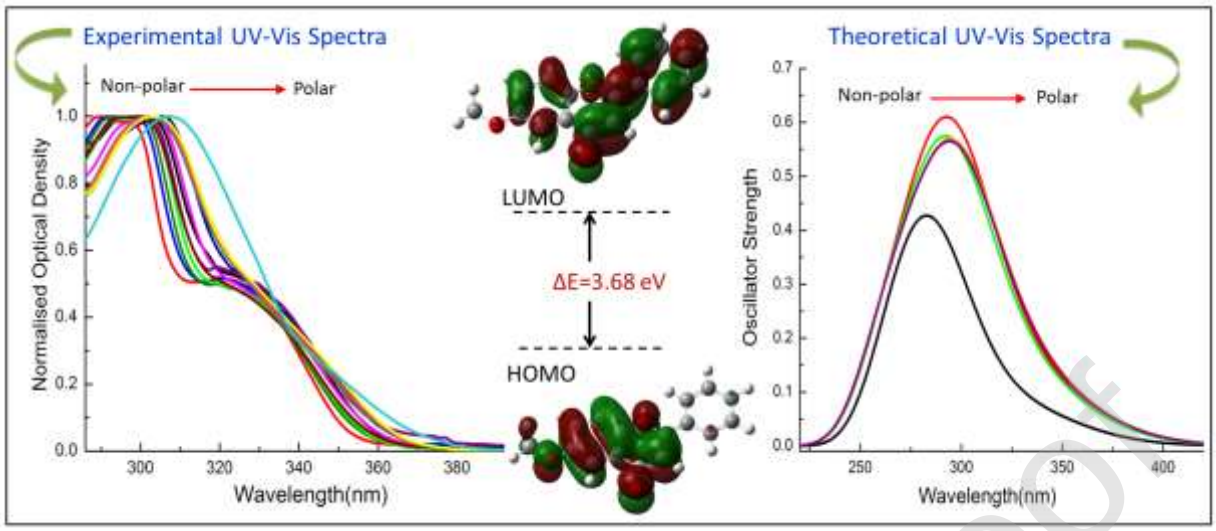
[2]Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India.

*Corresponding author: msmehata@gmail.com (MSM)

Journal Pre-proof

Journal Pre-proof

GRAPHICAL ABSTRACT



Highlights

- The ground and the excited state dipole moments of the 6-methoxyflavone were estimated from several methods.
- Charge transfer mechanism is found to be the cause of larger excited state dipole moment.
- Solvent dependent spectral shifts in absorption and fluorescence spectra were analyzed by multiple linear regression analysis.
- This work is characterized by detailed quantum chemical studies on the nature and extent of solvent–solute interactions.

Abstract:

The present work emphasizes on the experimental and computational analysis of the solvatochromic behavior of 6-methoxyflavone (6MF). The steady-state absorption and fluorescence spectra were observed in certain polar protic, polar aprotic and non-polar/weakly polar solvents to examine their solvatochromism and determine the electrical dipole moment.

Journal Pre-proof

Bathochromic shift in fluorescence spectra with solvent polarity was ascribed to intramolecular charge transfer (ICT) interaction. Calculations based on various solvent correlation techniques denote that the excited singlet-state dipole moment value is higher than the ground-state counterpart, hence indicates substantial redistribution of π-electron density in the excited state. Quantum chemical investigations were performed using density functional theory (DFT) and time-dependent density functional theory (TD-DFT), respectively, in vacuum and solvents. The vertical transitions, along with the bright singlet states and HOMO-LUMO, were obtained from the DFT/TD-DFT calculations. The experimentally calculated ground and excited state dipole moments and different photophysical parameters adequately reflect the values obtained from computational methods. Multiple linear regression (MLR) method based on Kamlet-Taft and Catalan solvent parameters demonstrate that the photophysical behavior of 6MF is attributed to both specific and non-specific solute-solvent interactions. Thus, the study helps to demonstrate the sensing ability of 6MF towards organic solvents and the charge transfer/hydrogen donor /acceptor properties of complex systems.

Keywords: Methoxyflavone, dipole moment, solvent effect, charge transfer, ground excited electronic states.

1. Introduction

Flavones are a widespread class of natural dyestuff abundantly present in the plant kingdom. Flavones are ubiquitously found in the human diet, mostly in fruits and vegetables. Flavones are biosynthesized under enzymatic control, including the amino acid phenylalanine, cinnamic acids, and malonic acid [1-3]. Flavones and their derivatives are captivating groups of organic molecules in the sphere of natural medicines, biosciences and pharmaceutical activities. The inherent properties of flavones such as antioxidant, antiviral, antibacterial, anti-inflammatory, anticancer, angioprotective and low toxicity are of paramount importance for many scientific

Journal Pre-proof

centers [4-6]. The flavones framework consists of a 2-phenyl-1-benzopyran-4one moiety. Most natural flavones carry either hydroxy or methoxy groups or are associated with a sugar residue ("glycones") [7,8]. Among the various substituted flavones derivatives, photophysical characteristics of hydroxy flavones have been extensively reviewed over the past few years. Hydroxyflavone (HF) is a captivating array of organic fluorophores that plays a prominent role in analytical chemistry and designing of probes [9-12]. Photophysics of hydroxyl derivatives have been examined thoroughly in organic solvents [7,11], supramolecular assemblies [13], micelles [14,15], microemulsions [16,17], liposomes [18], cyclodextrins [19,20] and different metal ions [21-26], etc. The phenomenon of excited-state intramolecular proton transfer (ESIPT) in hydroxyflavone has been explored in-depth and used to illustrate the excited state properties of flavanol and related compounds [27-29].

Moreover, methoxyflavones are influential due to their accessibility in natural and synthetic forms. These molecules are utilized as building blocks in the organic synthesis of potential novel anticancer drugs [30]. Owing to their importance, photophysical parameters (absorption, fluorescence, fluorescence lifetime and pK_a values) of some methoxy derivatives have been proclaimed in recent times, and most of the information is restricted to methanol and sulphuric acid as a solvent [7,8,31,32]. The photophysical properties of 3, 5 and 7 substituted methoxyflavones have been elaborated in the literature. Studies have explored the quantitative formation of the triplet states in 7MF, and illustrated its electron donor and acceptor properties. Conversely, 3MF exhibits reduced triplet formation, which behaves as an electron donor [33]. Smith *et al*. [3] explained the fluorescence spectra and decreased quantum yields of 5-, 7- and 8-mono-methoxyflavones in 3-methyl pentane solvent, based on the thermally induced vibronic coupling of the zero-order excited states. Because of the limited number of literature reports, various photophysical aspects of some methoxy-substituted flavones are still unresolved.

On account of the previous reports, it is apparent that the position of the $-OCH_3$ groups in flavonoids may significantly affect the photochemical responses and spectroscopic properties

Journal Pre-proof

as well. One of the positional substituents is 6-methoxyflavone. 6MF possesses good protein binding ability that influences the dissemination, metabolism, excretion and interaction of some medicinal molecules with tissues [34]. 6MF was quantitatively detected in human plasma by liquid chromatography-tandem mass spectrometry [35]. Some studies have illustrated that 6MF is a novel antinociceptive drug that might be a therapeutic option in the management of neuropathic pain linked with limited effectiveness of chemotherapy-induced neuropathy treatments [36].

The spectral behaviour and biological activities of an organic molecule are primarily analogous to its structure in both the ground and excited states. The solvent-solute interaction analysis is important in view of the fact that the solvents can significantly influence the physical and chemical properties of the molecule [37,38]. Estimation of the ground state and excited-state dipole moments of a molecule plays a crucial role in research, as it brings light on the distribution of charge near the fluorophore, electron density, electronic and geometrical configuration of a probe molecule in the solvent. The insight into the effect of solvents on electronic transitions and dipole moment of an excited molecule (μ_e) are of considerable interest as they explain the nature of the excited state and progression of its photochemical transformations. It can be implemented in the parameterisation of semi-empirical quantum chemical methods employed for the system. Several solvatochromic techniques, including those developed by Bilot Kawski [39-41], Lippert [42-44], Bakhshiev [45], Kawski-Chamma-Viallet [46-49], Suppan [28,44] and Ravi *et al.* [50] were accustomed to evaluate the dipole moments in the ground and excited states. These approaches have been used from the last decades till date to estimate the dipole moment of organic molecules. Imperative information on the dipole moments of molecules in both the ground state (GS) and excited state (ES) conveyed by quantum-chemical calculations enables the estimation of absolute values of μ_g and μ_e of molecules in the gas phase [51-60].

The focus of the current analysis is to estimate μ_g and μ_e using solvatochromic shift and computational methods. As far as we know (based on reported literature), there are no current

Journal Pre-proof

reports available regarding 6-methoxyflavone (6MF) as a solvatochromic probe. Considering 6MF to be biological and photochemical active, it is necessary to accomplish information on the geometry of the molecule. Therefore, this research work emphasizes on understanding the effect of general solvents along with alcohols on its electronic transitions. This study utilizes the Bilot-Kawski, Lippert-Mataga, Bakhshiev, Kawski-Chamma-Viallet, and Reichardt's method to describe the ground state (μ_g) and excited state (μ_e) dipole moments and experimental results of 6MF. Additionally, density functional theory (DFT) and time-dependent density functional theory (TD-DFT) were applied to determine the value of μ_g and μ_e in the gas phase along with the vertical transitions.

2. Materials and Methods

2.1. Materials

6-Methoxyflavone of 99% purity was procured from Sigma–Aldrich, USA and used without further purification. Spectroscopic grade solvents used in the study were purchased from Sisco Research Laboratories (SRL) Pvt. Ltd and were of the highest available purity (99%, HPLC grade).

2.2. Instrumentation

The electronic absorption spectra of 6MF were recorded in different solvents at room temperature in the wavelength region of 250-550 nm using dual-beam UV-VIS-NIR (Lambda-750 Perkin Elmer) spectrophotometer. The fluorescence spectra were measured using Horiba Fluorolog-3 spectrofluorometer equipped with a double grating for both excitation and emission paths and a 450W xenon lamp for excitation. Transparent quartz cuvettes were used for the absorption and emission measurements to avoid the inner filter effect. Data analysis was carried out using Origin software supplied by Horiba. Fluorescence spectra of 6MF in different solvents were measured on exciting by absorption maxima of the respective solvent. To calculate quantum yield, excitation wavelength (λ_{ex}~300 nm) was kept fixed for fluorescence measurements in all solvents.

Journal Pre-proof

2.3. Quantum chemical computational details

Quantum-chemical calculations were performed to calculate the vertical transitions, μ_g, and μ_e of 6MF. The computations were carried out using GAUSSIAN 09W [61], and software package GAUSSVIEW 5.0 [62] was used to draw the molecular structure. Density functional theory (DFT) [63,64] and Hartree-Fock (HF) [65] methods with the Hybrid Becke3-Lee-Yang-Parr (B3LYP) functional levels, adopting the standard basis set 6-31G(d,p), was opted for the energy optimization and calculation of dipole moment in the ground state. The theoretical study of ES was performed using a time-dependent DFT (TD-DFT) level of theory [66,67]. Additionally, various molecular properties such as the energy of the Highest Occupied Molecular Orbital (HOMO), energy of the Lowest Unoccupied Molecular orbital (LUMO), electronegativity and ionization potential, etc., have been described in the present work. Absorption and emission spectra were estimated with the TD-DFT approach.

3. Results and Discussion

3.1. Steady-state measurements: Effect of solvents on absorption and emission spectra

UV-Vis absorption and fluorescence spectra of 6MF (for 10^{-5} M concentration) were examined at room temperature in various non-polar, polar aprotic and polar protic solvents (isoctane (ISCTN), toluene (TOL), diethylether (DEE), dioxane (Dioxane), ethylacetate (EtOAc), dimethylformamide (DMF), acetonitrile (ACN), tetrahydrofuran (THF), dimethylsulphoxide (DMSO), methanol (MeOH), ethanol (EtOH), propanol (PrOH), butanol (BuOH), heptanol (HepOH), octanol (OctOH), trifluoroethanol (TFE)). The obtained values of physicochemical parameters are provided in Table 1. The absorption and fluorescence spectra of 6MF are represented in Figs.1 and 2, respectively. It is noteworthy that an absorption and fluorescence spectrum of 6MF in MeOH is in complete agreement with the reported values [7]. The absorption maxima of 6MF were found to be in 290–306 nm wavelength range, whereas the fluorescence maxima were obtained in the wavelength range of 351-446 nm. It is worth mentioning here that the fluorescence emission of 6MF was not observable in some non-polar and aprotic solvents

Journal Pre-proof

viz cyclohexane and acetone. From Figs. 1 and 2 it can be noticed that, as the polarity of the solvent increases, an apparent red shift was observed in the fluorescence spectra of 6MF ($\Delta\lambda_f \sim$ 95 nm) when compared to its absorption spectra ($\Delta\lambda_a \sim$ 16 nm). The significant fluorescence spectral shift suggested that the titled molecule is more polarized in the singlet ES over the GS, and therefore, the expected value of μ_g would be less than μ_e. The positive solvatochromism and higher value of dipole moment in the excited state indicate towards delocalization of π electrons [52-60]. Excitation spectra of 6MF were also recorded in polar solvents and it quietly resembles the absorption spectra (Fig. S1 of Supplementary Information).

The huge difference in Stokes shift value (5322 cm^{-1} to 10258 cm^{-1}) was observed in polar solvents in comparison with non-polar solvents. This infers that redshift and intramolecular charge transfer (ICT) arise due to π-π^* transition and affirms that the molecule was more stabilized in polar solvents than in solvents of low polarity (Table 1). The bathochromic shift has emerged as a result of the strong interaction of the singlet ES solute molecule with polar solvents, which may lead to a change in charge distribution between the GS and singlet ES of the solute molecule. In polar solvents, the redshift of the fluorescence spectra seems to occur through solute-solvent hydrogen-bonding interplay in the singlet ES [56] (see Table 1). Plausibly the hydrogen bonding in the ES of 6MF is reinforced, as the charge density at carbonyl oxygen is enhanced in the ICT state. Variation in Stokes shift value with the polarity of solvents assists the significant change in molecule's geometry in ES upon electronic transition. Perhaps it would influence solvent parameters like hydrogen bond acceptor (HBA), hydrogen bond donor (HBD) strength and solvent polarity. The solvatochromic nature of 6MF is found to be similar to some benzo[a]xanthene [52], methoxycoumarin [55], and chalcone derivatives [56]. In 6MF, the maximum redshift in fluorescence is observed while going from ISCTN to TFE, whereas the fluorescence is blue-shifted in higher alcohols (TFE, MeOH, EOH, PrOH, BuOH, OctOH, HepOH). Alcohols possess hydrogen bond donating (HBD) parameter (α), and TFE has the largest 'α =1.51' value comparative to other solvents as well as alcohols. With increasing, α

Journal Pre-proof

value of the solvent red shift is observed in alcohol. The redshift in the alcoholic solvents is caused by solute-solvent hydrogen bonding interactions in the singlet ES. Hydrogen bonding is strengthened in the ES as the charge density at the carbonyl oxygen is enhanced in the ICT excited state.

Fluorescence quantum yield (QY) calculations using Eq. (S25) represents that 6MF shows the highest quantum yield in trifluoroethanol (TFE) and the lowest one in toluene (TOL) solvent. A higher quantum yield in polar solvents may be considered due to the ICT mechanism. ICT is a plausible predominant ES process in polar solvents, in which it hinders more efficient quenching by a non-radiative process, thereby leading to an increase in fluorescence QY. QY of certain aromatic aldehydes and ketones, for example, 1-formylpyrene and 1-acetylpyrene, often enhance in protic solvents. In such aromatic hydrocarbons, the lowest-lying transition may be the n→π* transition. The increase in the conjugation of the π electrons causes the shifting of absorption and fluorescence spectra towards the longer wavelength regions and also increased fluorescence quantum yields. In those systems, it is believed that protic solvents serve to destabilize the n→π* excited state through hydrogen bonding and as a result, the intersystem crossing process is suppressed. Moreover, when the dyes possess the electronic characteristics that correspond to n→π* states in addition to CT, their QYs are enhanced by protic solvents due to their hydrogen bond donating capacity [68,69]. Facts reported for methoxyflavones indicates some contribution of the n→π* transition with the dominant character of π→π* [33]. Hence it can be deduced that a hydrogen bond is formed between the carbonyl oxygen in 6MF and solvent molecules, which results in a reversion of the lowest-lying transition from n→π* to π→π* states in the alcohols. Since π→π* becomes the lowest-lying transition in the alcohols and express relatively high QYs as compared to hydrocarbon solvents. In view of this, it can be considered that high QY of 6MF in TFE is due to its hydrogen bond donating capacity (α), which is highest in comparison to the other polar solvents. The values of the oscillator strength (*f*) and transition dipole moment ($\Delta\mu_{12}$) were calculated from Eq. (S26) and (S27), respectively. It is observed that

Journal Pre-proof

the highest value of oscillator strength (0.847) and transition dipole moment (7.231 Debye) for 6MF were found to be in diethylether (DEE) solvent.

3.2. Evaluation of ground and excited-state dipole moments: Solvatochromic behavior of 6MF

To determine dipole moments μ_g and μ_e of 6MF, solvatochromic shift methods that associate spectral shift (Stokes shift) and solvent polarity as given by, Bilot-Kawski (Eqs. (S1 of Supplementary Information) and (S2)), Lippert-Mataga (Eq. (S11)), Bakhshiev (Eq. (S12)), Kawski-Chamma- Viallet (Eq. (S13)) and Reichardt correlation (Eq. (S20)), were taken into account. The solvent polarity functions $F_{B-K\,(1)}(\varepsilon, n)$, $F_{B-K\,(2)}(\varepsilon, n)$, $F_{L-M}(\varepsilon, n)$, $F_B(\varepsilon, n)$, and $F_{K-C-V}(\varepsilon, n)$ were calculated using Eq. (S3), (S4), (S17), (S18) and (S19), respectively, for solvents of increasing polarities (given in Table S1 of Supporting Information). Values of $\bar{\nu}_a - \bar{\nu}_f$ versus $F_{B-K\,(1)}(\varepsilon, n)$ and $\frac{\bar{\nu}_a + \bar{\nu}_f}{2}$ versus $F_{B-K\,(2)}(\varepsilon, n)$, $\bar{\nu}_a - \bar{\nu}_f$ versus $F_{L-M}(\varepsilon, n)$, $F_B(\varepsilon, n)$, E_T^N and $\frac{\bar{\nu}_a + \bar{\nu}_f}{2}$ versus $F_{K-C-V}(\varepsilon, n)$ were arranged in the graphs and the slopes, intercepts and correlation coefficients calculated from the linear plots are presented in Table 2. The linearly fitted plots obtained from the respective equations are depicted in Figs. 3(a-f)**.** Correlation coefficients (R) were determined by using the maximum number of solvents fitted to obtain good linearity and are within the range of (0.702 to 0.94).

The slopes $m_{B-K\,(1)}$, $m_{B-K\,(2)}$ extracted from Bilot–Kawski plots on the basis Eq.(S2 of Supplementary Information) and (S3) have been employed in Eq.(S7) and (S8) to calculate μ_g and μ_e, respectively. According to the calculations, values of dipole moments μ_g and μ_e comes out to be 1.00 D and 7.67 D, respectively. Values of μ_g and μ_e, the difference in dipole moments ($\Delta\mu$), dipole moment ratio (μ_e/μ_g), Onsager cavity radius (α) and the angle between dipole moment vectors (φ) of 6MF as figured by calculations are accumulated in Table 3. As can be seen in Table 3, significant differences were identified for the μ_e values of 6MF, depending on the type of solvatochromic relationship. Such deviations in the (μ_e) values are perhaps due to a

Journal Pre-proof

number of suppositions and approximations harnessed in these solvatochromic methods, as mentioned in the literature [55]. The difference in μ_g and μ_e states that the molecule has a substantial distribution of charge in singlet ES and therefore plays a vital role in the ICT process. The angle between two electronic state dipole moment vectors is calculated using μ_g and μ_e obtained from Bilot-Kawski solvent correlation (Eq. (S10)) and is found to be 0º. The angle between μ_g and μ_e shows that two electronic states are approximately parallel to each other. To exclude the possibility of interference of the specific solvent effect, dipole moment calculations of 6MF were also carried out in the set of non-polar and aprotic solvents (excluding alcohols). It is observed that GS dipole moment (μ_g) is 0.426 D and singlet ES dipole moment (μ_e) is 7.65 D. A noticeable change (increment) in μ_e suggests that the reallocation of charge in the ES is not due to specific solvent effects.

It has been investigated [59] that the dependence of dipole moments on the positional substituents of hydroxyquinoline (HQ) expressed that the CT character is more favorable when OH group is at 5, 6, 7 and 8 positions of HQ. Based on the report, it could be assumed that in 6MF, CT takes place between OCH_3 and C=O containing ring, as O atom of carbonyl acts as an electron donor and OCH_3 appears to be electron acceptor. However, considering the separation between the functional groups and native ring, OCH_3 group is located much far from the carbonyl (native) ring in the case of 6MF. As solvatochromism is directly related to $(\Delta\mu)^2$, therefore, 6MF also displays large Stokes shift values. The presence of negative and positive sites of interaction thus frames the examined molecule more responsive when interacting with the solvent in the singlet ES. Hence, large separation may lead to high dipole moment value for the 6MF too [56,59].

As depicted in Table 3, μ_g is 5.06 D and μ_e is 5.79 D computed (in the gas phase) using DFT (Gaussian 09W), specifying that the 6MF molecule is undoubtedly more polar in its ES than in its GS. The distinction between experimentally and computationally determined μ_g and

Journal Pre-proof

μ_e values can be comprehended by the fact that computational measurement belongs to specific molecules without any physical interaction between the solvent and the solute.

3.3. Computational studies by quantum chemical methods

In order to gain more information and insight into the molecular structure and electronic properties of 6MF, quantum chemical calculations were performed using HF/6-31G(d,p), DFT/B3LYP/6-31G(d,p) for GS and TD-DFT/B3LYP/6-31G(d,p) for ES (Table S2). The molecular geometry of 6MF was optimized by TD-DFT/6-31G(d,p) and is depicted in Fig. 4. Theoretically calculated spectral parameters of 6MF are summarized in Table 4. Both μ_g and μ_e were also estimated for 6MF are given in Table 4. Repeatedly the value of the ES dipole moment is higher than the value of the GS dipole moment, and this trend is in accordance with the experimental results. However, the difference is acclaimed in the experimental and computational values of GS and ES dipole moments. It is worth mentioning here that examples of the variations in dipole moments (theoretical) from the experimental results were found in previous reports based on the computational methods used in quantum chemistry [55-59].

Furthermore, the calculated frontier molecular orbitals, HOMOs and LUMOs and the corresponding orbital energies of 6MF are shown in Fig. 5, which show charge transfer character, were determined using TD-DFT/6-31G(d,p) level of theory. The positive and negative phases of HOMO-LUMO orbitals have been represented in red and green regions. 6MF framework consists of a 6-methoxy-2-phenyl-4-benzopyrone moiety. As can be seen from HOMO-LUMO orbital's that in HOMO electrons are confined to the 6-methoxy-4-benzopyrone moiety, whereas in LUMO they are located on all the groups. The low energy gap value indicates that the molecule becomes more responsive and has a simpler $\pi \rightarrow \pi^*$electronic transition. Efficient energy spacing in 6MF points out its high reactivity and low kinetic stability. The calculated electronic transition energies and corresponding oscillator strength (*f*) of the low-lying singlet excited states are given in Table 5. Typically, the S3 state seems the brightest one and the CI value is maximum corresponds to the HOMO−2 to the LUMO. The CI value, the contribution

Journal Pre-proof

of orbital transitions for other transitions, is comparatively small (Table 5). The slightly brighter state (S4) is corresponding to the HOMO-1 to the LUMO+1.

Global chemical reactivity descriptors (GCRD) help to describe chemical reactivity and stability of the molecules. GCRD parameters such as ionization potential (IP), electron affinity (EA), chemical hardness (η), softness (s), chemical potential (μ) and electron negativity (χ) of a molecule are collected in Table 6. IP and EA values indicate that 6MF is highly reactive and acts as an electron donor. In 6MF, carbonyl oxygen makes it electronegative, as electronegativity describes the ability of a molecule to attract electron density. The negative value of 'μ' also shows 6MF as highly reactive. Furthermore, chemical hardness (η), softness (s), and low energy gap describe that 6MF is moderately stable and can easily be polarised. Table 7 represents various parameters (viz dipole moments, HOMO, LUMO, oscillator strength, λ_{abs}, λ_{em}, etc.) in different solvents calculated by DFT and TD-DFT level of theories using Integral Equation Formalism Polarizable Continuum Model (IEFPCM) as a solvation model. The observed absorption and emission spectra obtained at TD-DFT/6-31G(d,p) level of theory level are given in Fig. 6 and demonstrating the vertical transition of 6MF. Calculations of electronic absorption spectra for lower energy $\pi \rightarrow \pi^*$ transitions show maximum absorption wavelengths at around 288 nm in the gas phase, 297 nm in less polar solvents and 299 nm in polar solvents, they nearly resembled to the experimental absorption spectra (290-306 nm) in respective solvents. Emission wavelengths are blue-shifted as compared to that achieved in solvents. Moreover, variation in dipole moment is also noticed while going from gaseous to solvent phases, i.e., higher values of dipole moments are observed in the solvent as compared to the gas phase. This is due to the intramolecular interactions between the solvent molecules and the solute. Furthermore, the theoretical values of oscillator strength and transition dipole moments are virtually identical to the experimentally observed values for most of the solvents.

Molecular polarizability, dipole moment, electronic structure and various molecular properties of the system are influenced by the atomic charges [70]. Mulliken atomic charge

Journal Pre-proof

distribution of 6MF was investigated using DFT/B3LYP/6-31G(d,p) and TD-DFT/B3LYP/6-31G(d,p) level of theory, in gas phase. The Mulliken charge values of 6MF are given in Table S3 of supporting information. The heteroatoms, carbon atoms and hydrogen atoms allied with negative charges are donor atoms, whereas atoms bearing positive charges indicate the acceptors. Oxygen (14O ~ −0.574) of carbon group is most negative, carbon (8C ~ 0.390) of carbonyl group is most positive, whereas hydrogen of methoxy group (29H ~ 0.130) bears more positive charge density than other hydrogen atoms. Thus charge transfer may occur from carbonyl to methoxy group. Mulliken atomic charge in GS is compared with ES charge distribution (Fig. 7). It is noticed that negative charges on atoms 3C, 7H, 13H, 24C, 25H, 26H and positive charges on the atoms 5C, 14O, 21H, 28C are somewhat increased upon excitation. This may result in the difference of μ_g and μ_e values and can lead to the stabilization of the ES structure of 6MF.

Further, to get more insight into the structure of 6MF, the molecular electrostatic potential (MEP) region is plotted using the TD-DFT method. MEP helps in understanding the influence of molecular structure in biological activities and also indicates the hydrogen bonding sites of molecular interaction. The MEP of 6MF was recognized in the region of −7.998e−2 kcal/mol (red) to 7.998e−2 kcal/mol (blue), as represented in Fig.8. In the plotted MEP surface red color region indicates a potentially negative region near the oxygen atom of the carbonyl group. The red region of the MEP map reveals that oxygen is electron-rich and consists lowest electrostatic potential energy of the molecule. As a result, oxygen has quite a high electron density and hydrogen is comparatively electron deficient in 6MF.

3.4. Multiple linear regression analysis

The results obtained from multiple linear regression (MLR) analysis proposed by Kamlet-Abboud-Taft and Catalan equations have been shown in Tables 8 and 9. Kamlet-Abboud-Taft parameters and Catalan solvent scales have confirmed that bathochromic shift in absorption and fluorescence spectra on increasing polarities are attributable to negative values for all the parameters in Eq. (S22) and (S23).

Journal Pre-proof

Kamlet-Abboud-Taft parameters have been calculated for all the solvents used in the study (except isooctane). The absorption and fluorescence wavenumbers of π→π* electronic transition in the gas phase of the 6MF molecule are 34172 cm^{-1} and 30804 cm^{-1}, respectively. In Kamlet-Abboud-Taft scale, large values of dipolarity/polarizability (π*) reflect its major contribution towards the bathochromic shift in absorption and solute-solvent interactions. It can be suggested that the changes in fluorescence spectra were mainly controlled by dipolarity/polarizability, which describes non-specific solute-solvent interactions. It is noteworthy that the electronic distribution of GS and singlet ES of 6MF differs considerably and that the singlet ES was more polarized than the GS. Hence the dipole moment of singlet ES of 6MF was found to be greater than the GS dipole moment. Whereas in specific solute-solvent interaction, hydrogen bond donor capacity (α) of solvent is observed to have a major effect in leading the π→π* electronic transitions in comparison to that of hydrogen bond acceptor capacity (β), since for absorption and fluorescence spectra, the absolute value of ***a*** is greater than the absolute value of ***b***. Thereby it can be proposed that the charge transfer in 6MF is mainly regulated by hydrogen bond donor ability of solvent and acceptor center (carbonyl moiety) of 6MF.

In terms of Catalan solvent scales, the correlation regression coefficient values of polarizability and dipolarity were proved to be higher than the acidity and basicity of solvents. This confirms that polarizability and dipolarity of solvents contribute significantly to the bathochromic shifts of absorption and fluorescence spectra and solute-solvent interactions. But the value of dipolarity is more than polarizability; therefore, the absorption, fluorescence and Stokes shift strongly depend on dipolarity of the solvents. It is assumed that the solvent acidity highly influences the absorption and fluorescence spectral shifts in comparison with solvent basicity. This indicates that the titled molecule has both of non-specific and specific interactions.

4. Conclusion

Journal Pre-proof

Photophysical properties of 6MF have been examined in a range of solvents of increasing polarities. Significant variations in the Stokes shift values with increasing solvent polarities for 6MF have shown that the bathochromic shift of absorption and fluorescence spectra can be due to dominating $\pi \rightarrow \pi^*$ transitions in methoxyflavones. This propounds that the titled molecule was more solvated in singlet ES than in GS. Calculations based on solvatochromic shift methods conclude that ES dipole moment is significantly higher than that of GS. Discussions also reveal the possibility of ICT process associated with the donor-acceptor moiety of 6MF. DFT and TD-DFT calculations demonstrate various photophysical properties (dipole moments, HOMO-LUMO, atomic charges, electrostatic potential) and endorse the experimental assumptions. MLR analysis proposed that solute-solvent interactions are regulated by solvent dipolarity/polarizability and hydrogen-bond donor ability of the solvents. 6MF was found to be more reactive due to efficient energy gap. Therefore, on the basis of the aforementioned results, 6MF may be considered as a promising candidate for fluorescent probes, luminescent materials and optoelectronic devices.

Author Statement

NF: Formulated the idea, performed the spectral measurements and simulations and written the paper. MSM: Formulated the idea, performed few simulations and finalized the paper. NP: Helped to NF in the discussion on the results. SP: Read the final version of the paper and provided a helpful discussion.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Journal Pre-proof

MSM is thankful of Prof. Keli Han, Dalian Institute of Chemical Physics, CAS, China for providing the Gaussian 09 facilities and author (NF) to all the members of Photophysics Laboratory, Department of Physics, Centre of Advance Study, DSB Campus, Kumaun University, Nainital and Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Delhi for their support. Part of this work is supported by the Science and Engineering Research Board (SERB) (EMR/2016/001110), DST, Govt. of India.

References

1. J. Braun, M. Tevini, Regulation of UV-protective pigment synthesis in epidermal layer of rye seedlings, Photochem. Photobiol. 57 (1993) 318-323.
2. W. Heller, G. Forkmann, Biosynthesis of flavonoids, in J.B. Haghorne (ed,), The flavonoids--advances in research since 1986, Chapman and Hall, London, 1994.
3. G. J. Smith, K. R. Markham, The dissipation of excitation energy in methoxyflavones by internal conversion, J. Photochem. Photobiol. A: Chem. 99 (1996) 97-101.
4. M. Sisa, S. L. Bonnet, D. Ferreira, J. H. Van der Westhuizen, Photochemistry of flavonoids, Molecules 15 (2010) 5196-5245.
5. S. V. Jovanovic, S. Steeden, M. Tosic, B. Marjanovic, M. G. Simicg, Flavonoids as antioxidants, J. Am. Chem. Soc. 116 (1994) 4846-4851.
6. M. Samsonowicz, E. Regulska, M. Kalinowska, Hydroxyflavone metal complexes – molecular structure, antioxidant activity and biological effects, Chem.-Biol. Interact. 273 (2017) 245–256.
7. O. S. Wolfbeis, M. Leiner, P. Hochmuth, Absorption and fluorescence Spectra, pKa values, and fluorescence lifetimes of monohydroxyflavones and monomethoxyflavones, Ber. Bunsenges. Phys. Chem. 88 (1984)759-767.
8. T. J. Mabry, K. R. Markham, M. B. Thomas, The systematic identification of flavonoids, Chapter 4, Springer-Verlag, New York, Heidelberg, Berlin, 1970.

Journal Pre-proof

9. E. de Rijke , P. Out, W. M. A. Niessen, F. Ariese, C. Gooijer, U. A. T. Brinkman, Analytical separation and detection methods for flavonoids, J. Chromatogr. A 1112 (2006) 31–63.
10. M. Voicescu, S. Ionescu, F. Gatea, Photophysical properties of some flavones probes in homogeneous media, J. Fluoresc. 24 (2014) 75–83.
11. S. Kumar, S. K. Jain, R. C. Rastogi, An experimental and theoretical study of excited-state dipole moments of some flavones using an efficient solvatochromic method based on the solvent polarity parameter $E^{T}_{N.}$ Spectrochim Acta. A. 57:(2001) 291–298.
12. A. Zhu, B. Wang, J. O. White, H. G. Drickamer, The effect of pressure on the excited-state intramolecular proton transfer of polyhydroxyflavone, J. Phys. Chem. B. 108 (2004) 895-898.
13. A. Sytnik, I. Litvinyuk, Energy transfer to a proton-transfer fluorescence probe: Tryptophan to a flavonol in human serum albumin, Proc. Natl. Acad. Sci. USA 93 (1996) 12959-12963.
14. S. Mondal, S. Basu, D. Mandal, Ground- and excited-state proton-transfer reaction of 3-Hydroxyflavone in aqueous micelles, Chem. Phys. Lett. 479 (2009) 218–223,
15. S. M. Dennison, J. Guharay, P. K. Sengupta, Intramolecular excited-state proton transfer and charge transfer fluorescence of a 3-hydroxyflavone derivative in micellar media, Spectrochim. Acta A 55 (1999) 903-909.
16. A. S. Klymchenko, A. P. Demchenko, Probing AOT reverse micelles with two-color fluorescence dyes based on 3-Hydroxychromone, Langmuir 18 (2002) 5637-5639.
17. T. Shyamala, A. K. Mishra, Ground and excited- state proton transfer reaction of 3-Hydroxyflavone in Dimyristoylphosphatidylcholine liposome membrane, J. Photochem. Photobiol. A 80 (2004) 309-315.
18. J. Guharay, R. Chaudhuri, A. Chakrabarti, P. K. Sengupta, Excited state proton transfer fluorescence of 3-hydroxyflavone in model membranes, Spectrochim. Acta A 53 (1997) 457-462.

Journal Pre-proof

19. A. Banerjee, P. K. Sengupta, Encapsulation of 3-hydroxyflavone and fisetin in β-cyclodextrins: Excited state proton transfer fluorescence and molecular mechanics studies, Chem. Phys. Lett. 424 (2006) 379-386.

20. L. Tormo, A. Douhal, Caging anionic structure of a proton transfer dye in a hydrophobic nanocavity with a cooperative H-bonding, J. Photochem. Photobiol. A: Chem. 173 (2005) 358-364.

21. S. Protti, A. Mezzetti, C. Lapouge, J. P. Cornard, Photochemistry of metal complexes of 3-hydroxyflavone: towards a better understanding of the influence of solar light on the metal–soil organic matter interactions, Photochem. Photobiol. Sci., 7 (2008) 109–119.

22. M. D. Engelmann, R. Hutcheson, I. F. Cheng, Stability of ferric complexes with 3-Hydroxyflavone (Flavonol), 5,7-Dihydroxyflavone (Chrysin), and 3',4'-Dihydroxyflavone, J. Agric. Food Chem. 53 (2005) 2953-2960.

23. L. Mira, M. T. Fernandez, M. Santos, R. Rocha, M. H. Florencio, K. R. Jennings, Interactions of flavonoids with iron and copper ions: a mechanism for their antioxidant activity, Free Rad. Res. 36 (2002) 1199–1208.

24. M. Thompson, C. R. Williams, G. E. P. Elliot, Stability of flavonoid complexes of copper (ii) and flavonoid antioxidant activity, Anal. Chim. Acta. 85 (1976) 375-381.

25. S. M. Z. Al-Kindy, K. H. Al-Hinai, F. O. Suliman, H. J. Al-Lawati, A. Pillay, Development of a selective fluorimetric technique for rapid trace determination of zinc using 3-hydroxyflavone, Arab. J. Chem. 4 (2011) 147–152.

26. N. Fatma, M. S. Mehata, N. Pandey, S. Pant, Flavones fluorescence-based dual response chemosensor for metal ions in aqueous media and fluorescence recovery. J. Fluoresc. (2020)1-14. Doi:10.1007/s10895-020-02540-z

27. P. K. Mandal, A. Samanta, Evidence of ground-state proton-transfer reaction of 3-hydroxyflavone in neutral alcoholic solvents, J. Phys. Chem. A. 107 (2003) 6334-6339.

Journal Pre-proof

28. P. K. Sengupta, M. Kasha, Excited state proton-transfer spectroscopy of 3-hydroxyflavone and quercetin, Chem. Phys. Lett. 68 (1979) 382-385.

29. S. A.-Beg, S. M. Ormson, R. G. Brown, P. Matousek, M. Towrie, E. T. J. Nibbering, F. V. R. Neuwahl, The ultrafast measurements of excited state intramolecular proton transfer (ESIPT) in room temperature solutions of 3-hydroxyflavone and derivatives. J. Phys. Chem. A. 105 (2001) 3709–3718.

30. T. Walle, Methoxylated flavones, a superior cancer chemopreventive flavonoid subclass? Semin Cancer Biol. 17 (2007) 354–362.

31. B. Skarzynski, Spektrographische untersuchung von flavonfarbstoffen. Biochem. Z. 301, 150 (1939).

32. O. S. Wolfbeis, M. Begum, H. Geiger, Fluorescence properties of hydroxy- and methoxyflavones and the effect of shift reagents. Zeitschrift Für Naturforschung B, 39 (1984) 231–237.

33. M. Christoff, V. G. Toscano, W. J. Baader, Influence of methoxy substitution on flavonoid photophysics: a steady state and laser flash photolysis study, J. Photochem. Photobiol. A: Chem. 101 (1996) 11-20.

34. H. Cao, X. Liu, N. P. Ulrih, P. K. Sengupta, J. Xiao, Plasma protein binding of dietary polyphenols to human serum albumin: A high performance affinity chromatography approach, Food Chem. 270 (2019) 257–263.

35. A. Shafaei, K. Croft, J. Hodgson, M. C. Boyce, Simultaneous quantitative analysis of polyphenolic compounds in human plasma by liquid chromatography tandem mass spectrometry, J. Sep. Sci. 42 (2019)2909–2921.

36. M. Shahid, F. Subhan, N. Ahmad, R. D. E. Sewell, The flavonoid 6-methoxyflavone allays cisplatin-induced neuropathic allodynia and hypoalgesia, Biomed. Pharmacother. 95 (2017) 1725–1733.

37. Y. Marcus, The properties of solvents, John Wiley & Sons, New York, 1998.

Journal Pre-proof

38. J. R. Lakowicz, Principles of fluorescence spectroscopy, third ed., Plenum Press, New York, 2006.

39. L Bilot, A. Kawski, Zur Theorie des einflusses von losungsmitteln auf die elektronenspektren der molecule, Z. Natuforsch. 17A (1962) 621–627.

40. L. Bilot, A. Kawski Der EinfluB des losungsmittels auf die elektronenspektren lumineszierender molecule, Z. Natuforsch. 18A (1963) 10.

41. A. Kawski, in: J.F. Rabek (Ed.) Progress in photochemistry and photophysics. vol. 5, CRC Press, Boca Raton 1992, p. 1 (Ann. Arbor, Boston).

42. E. Lippert, Dipole moment and electron structure of excited molecule, Z. Natuforsch. A 10 (1955) 541–5

43. E. Lippert, Spectroscopic determination of the dipole moment of aromatic compounds in the first excited singlet state, Z. Elektrochem. 61 (1957) 962–75

44. N. Mataga, N. Kaifu, M. Koizumi, Solvent effects upon fluorescence spectra and the dipole moments of excited molecules, Bull. Chem. Soc. Japan, 29 (1956) 465–70

45. N. G. Bakhshiev, Universal intermolecular interactions and their effect on the position of the electronic spectra of molecules in two component solutions, Opt. Spectrosk. 16 (1964) 821–32

46. A. Kawski, P. Bojarski, Comments on the determination of excited state dipole moment of molecules using the method of solvatochromism, Spectrochim. Acta A 82 (2011) 527–528

47. A. Kawski, Der wellenzahl von elecktronenbanden lumineszie- renden molekùle, Acta Phys. Pol. 29 (1966) 507–18

48. A. Kawski, On the estimation of excited-state dipole moments from solvatochromic shifts of absorption and fluorescence spectra, Z. Natuforsch. 57a (2002) 255–62.

49. A. Chamma, P. Viallet, Détermination du moment dipolaire d'une molécule dans un état excité singulet: application á l'indole, au benzimidazole et á l'indazole, C. R. Acad. Sci. Paris 270 (1970) 1901–1904.

Journal Pre-proof

50. M. Ravi, A. Samanta, T. P. Radhakrishan, Excited state dipole moments from an efficient analysis of solvatochromic Stokes shift data, J. Phys. Chem. 98 (1994) 9133–6.

51. P. Suppan, Excited state dipole moments from absorption/ fluorescence solvatochromic ratios, Chem. Phys. Lett. 3 (1983) 272–275.

52. K. Aggarwal, J. M. Khurana, Effect of hydroxyl group on the photophysical properties of benzo[a]xanthenes – Solvatochromic studies and estimation of dipole moment, J. Photochem. Photobiol. A: Chem. 276 (2013) 71– 82.

53. R. Rautela, N. K. Joshi, H. C. Joshi, N. Tewari, S. Pant, Solvatochromic study of 2-hydroxy-4-methylquinoline for the determination of dipole moments and solute–solvent interactions, J. Mol. Liq. 154 (2010) 47–51.

54. N. Tewari, N. K. Joshi, R. Rautela, R. Gahlaut, H. C. Joshi, S. Pant, On the ground and excited state dipole moments of dansylamide from solvatochromic shifts of absorption and fluorescence spectra, J. Mol. Liq. 160 (2011)150–3.

55. N. Pandey, R. Gahlaut, P. Arora, N.K. Joshi, H.C. Joshi, S. Pant, Study of dipole moments of some coumarin derivatives, J. Mol. Struct. 1061 (2014) 175-180.

56. A. M. Asiri, T. R. Sobahi, O. I. Osman, S. A. Khan, Photophysical investigation of (D-p-A) DMHP dye: Dipole moments, photochemical quantum yield and fluorescence quantum yield, by solvatochromic shift methods and DFT studies, J. Mol. Struct. 1128 (2017) 636-644.

57. M. K. Patil, M.G. Kotresh, S.R. Inamdar, A combined solvatochromic shift and TDDFT study probing solute-solvent interactions of blue fluorescent Alexa Fluor 350 dye: Evaluation of ground and excited state dipole moments, Spectrochim. Acta A. 215 (2019) 142-152.

58. A. Al Sabahi, S. N. Al Busafi, F.O. Suliman, S. M. Al Kindy, Photophysical and theoretical studies on the solvatochromic effects and dipole moments evaluation of substituted 1-phenyl-3-naphthyl-5- (4-ethyl benzoate)-2-pyrazoline, J. Mol. Liq. 307 (2020) 112967.

Journal Pre-proof

59. M. S. Mehata, A. K. Singh, R. K. Sinha, Experimental and theoretical study of hydroxyquinolines: hydroxyl group position dependent dipole moment and charge-separation in the photoexcited state leading to fluorescence, Methods Appl. Fluoresc. 4 (2016) 045004.

60. M. S. Mehata, A. K. Singh, R. K. Sinha, Investigation of charge-separation/change in dipole moment of 7-azaindole: Quantitative measurement using solvatochromic shifts and computational approaches, J. Mol. Liq. 231 (2017) 39–44.

61. M. J. Frisch, et al., Gaussian 09, Revision C.01, Gaussian Inc., Wallingford CT, 2010.

62. GaussView, Version 5, Ray Dennington, Todd Keith and John Milam, Semichem Inc.,Shawnee Mission KS, 2009.

63. M. S. Mehata, Y. Yang, Z. J. Qu, J. S. Chen, F. J. Zhao, K. L. Han, Spin mixed charge transfer states of iridium complex Ir(ppy)3: transient absorption and time-resolved photoluminescence, RSC Adv. 5 (2015) 34094–34099.

64. P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. 136 (1964) B864–B871.

65. C. C. J. Roothan, New developments in molecular orbital theory, Rev. Mod. Phys. 23 (1951) 69–89.

66. P. Zhou, K. Han, Unraveling the detailed mechanism of excited-state proton transfer, Acc. Chem. Res. 51 (2018) 1681−1690.

67. R. E. Stratmann, G. E. Scuseria, M. J. Frisch, An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules, J. Chem. Phys. 109 (1998) 8218–8224.

68. Y. Niko, S. Kawauchi, G. Konishi, Solvatochromic pyrene analogues of prodan exhibiting extremely high fluorescence quantum yields in apolar and polar solvents, Chem. Eur. J. 19 (2013) 9760 – 9765.

69. B. M. Beşer, Y. Onganer, M. Arik, Photophysics and photodynamics of Pyronin Y in n-alcohols, Luminescence. (2018) 1-7

Journal Pre-proof

70. S. A. Khan, A. M. Asiri, S. H. Al-Thaqafy, H. M. Faidallah, S. A. El-Daly, Synthesis, characterization and spectroscopic behavior of novel 2-oxo-1,4-disubstituted-1,2,5,6-tetrahydrobenzo[h]quinoline-3-carbonitrile dyes, Spectrochim. Acta Part A: Molec. Biomol. Spectrosc. 133 (2014) 141-148.

Journal Pre-proof

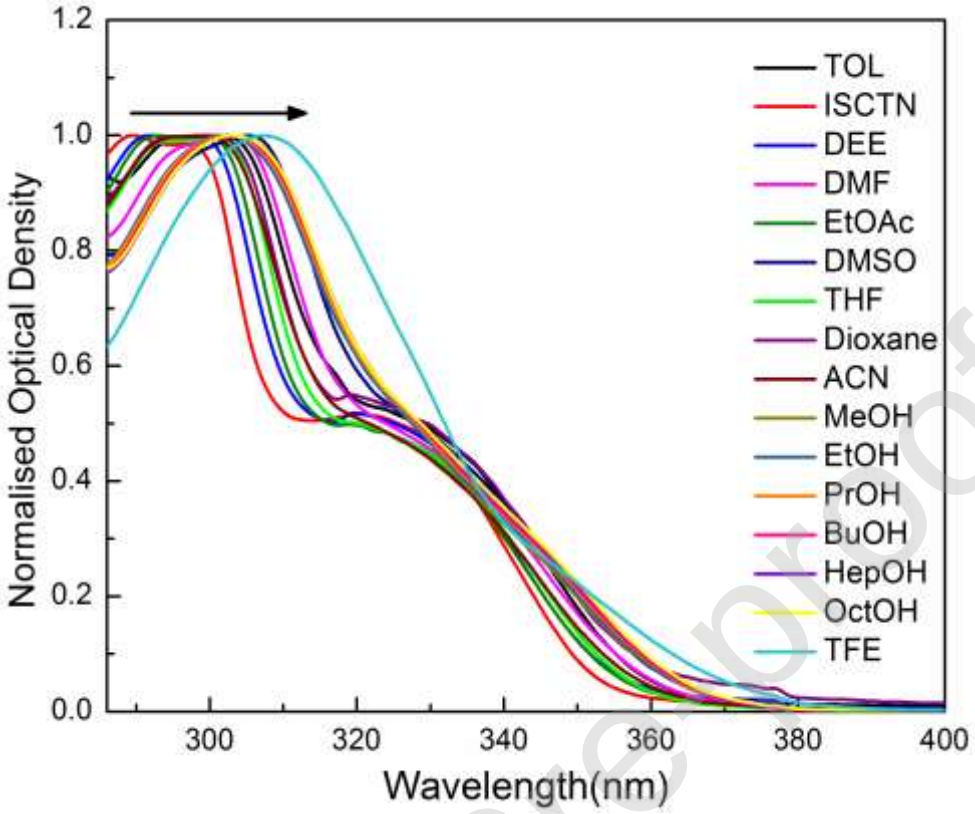


Fig. 1. UV-Visible absorption spectra of 6-methoxyflavone (10^{-5} M) in different solvents at room temperature.

Journal Pre-proof

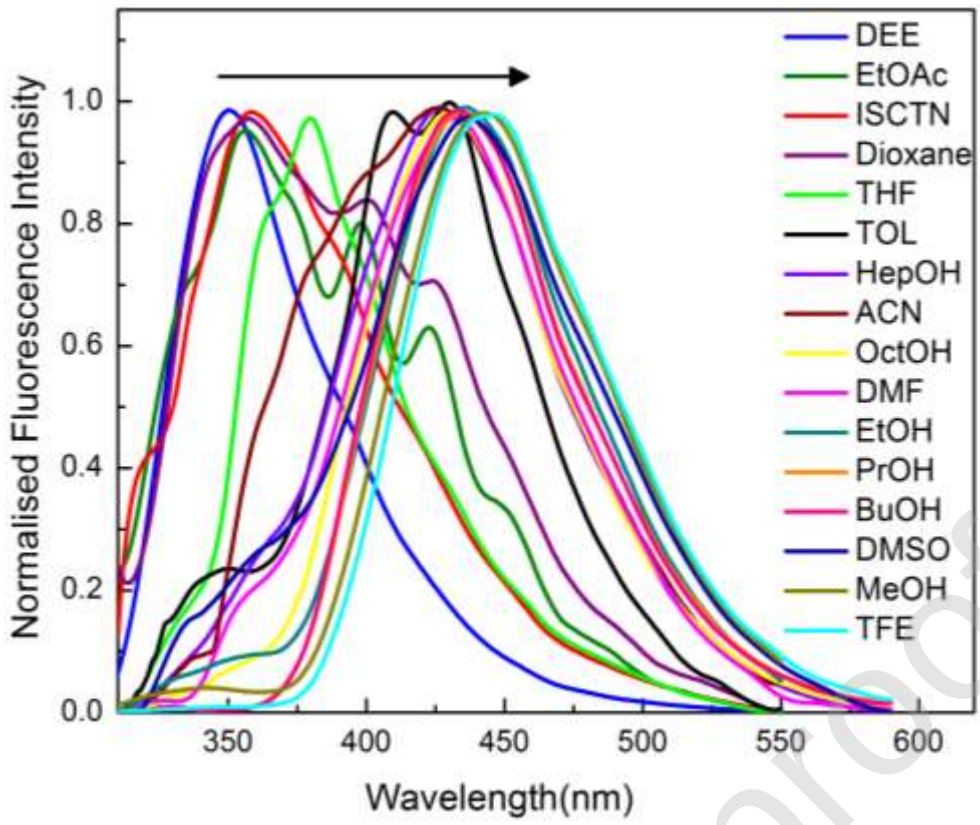
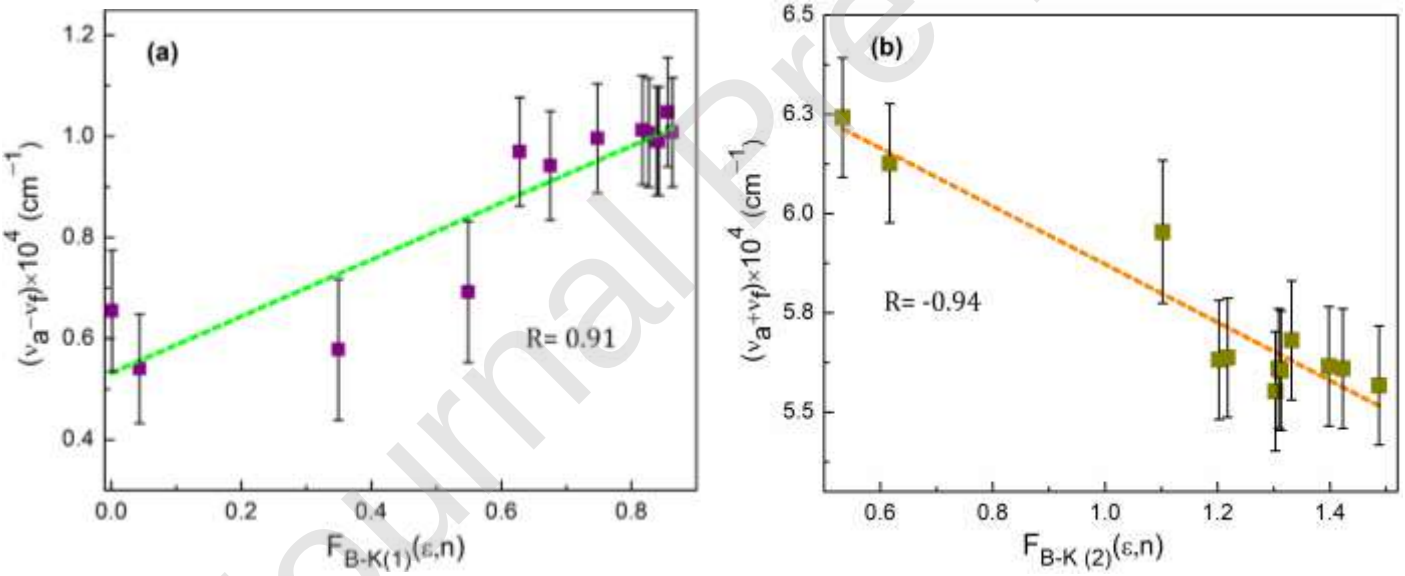


Fig. 2. Fluorescence spectra of 6MF (10^{-5} M) in different solvents at room temperature.



Journal Pre-proof

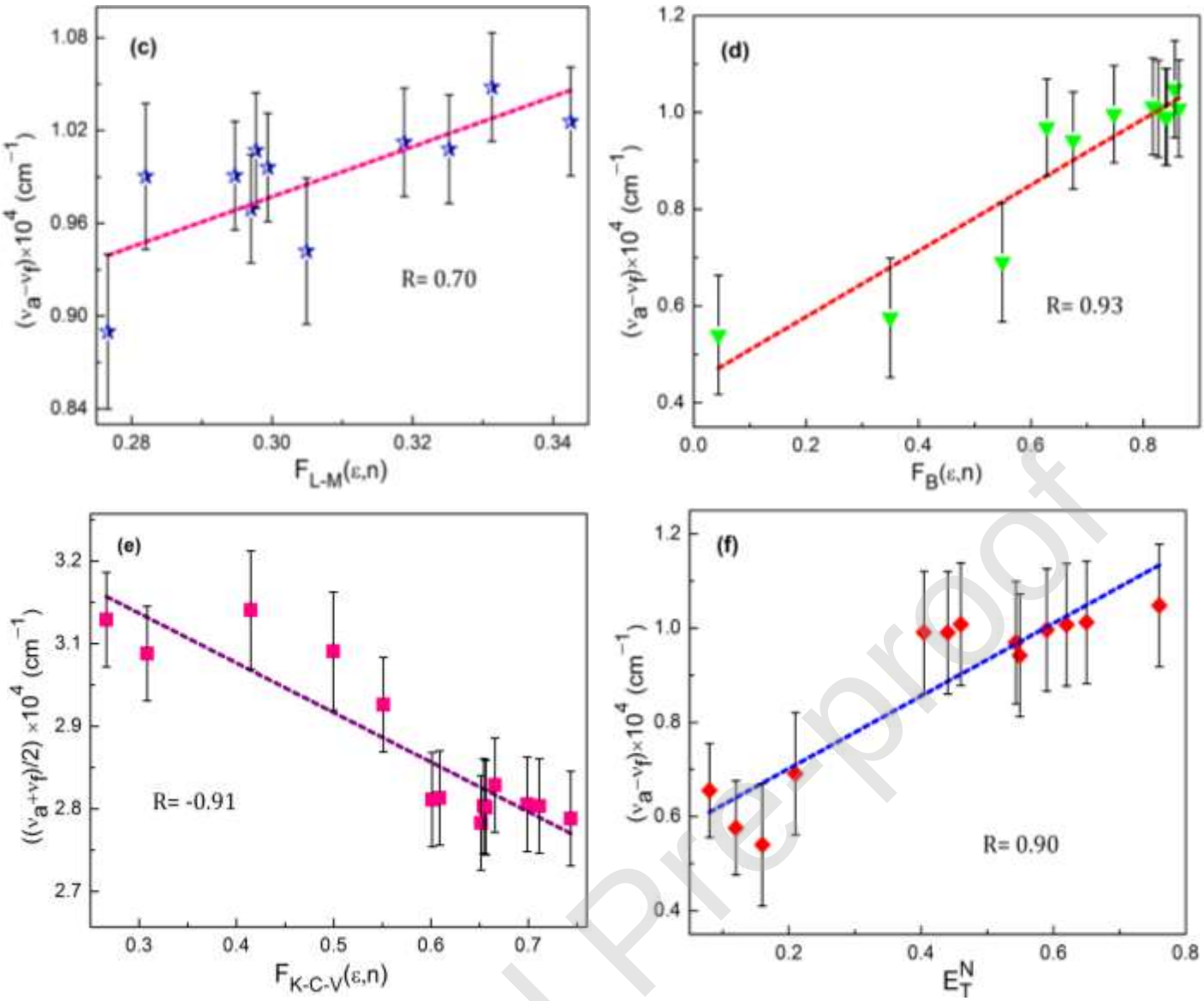


Fig. 3. Solvatochromic linear plots of Bilot-Kawski (a-b), Lippert-Mataga (c), Bakhshiev (d), Kawski-Chamma-Viallet (e) and Riechardt (f) correlations for 6MF in different solvents.

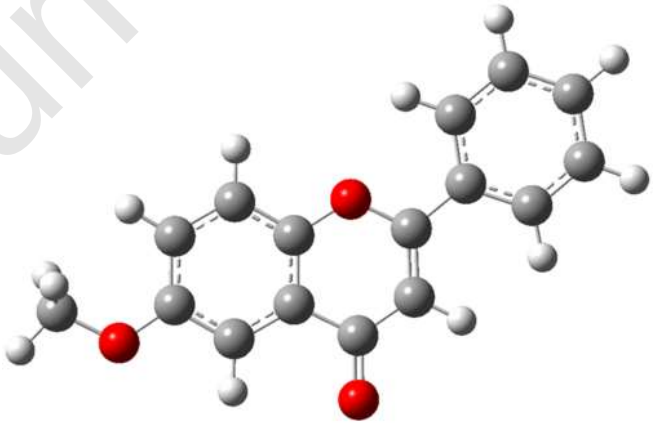


Fig. 4. The optimized geometrical structure of 6MF.

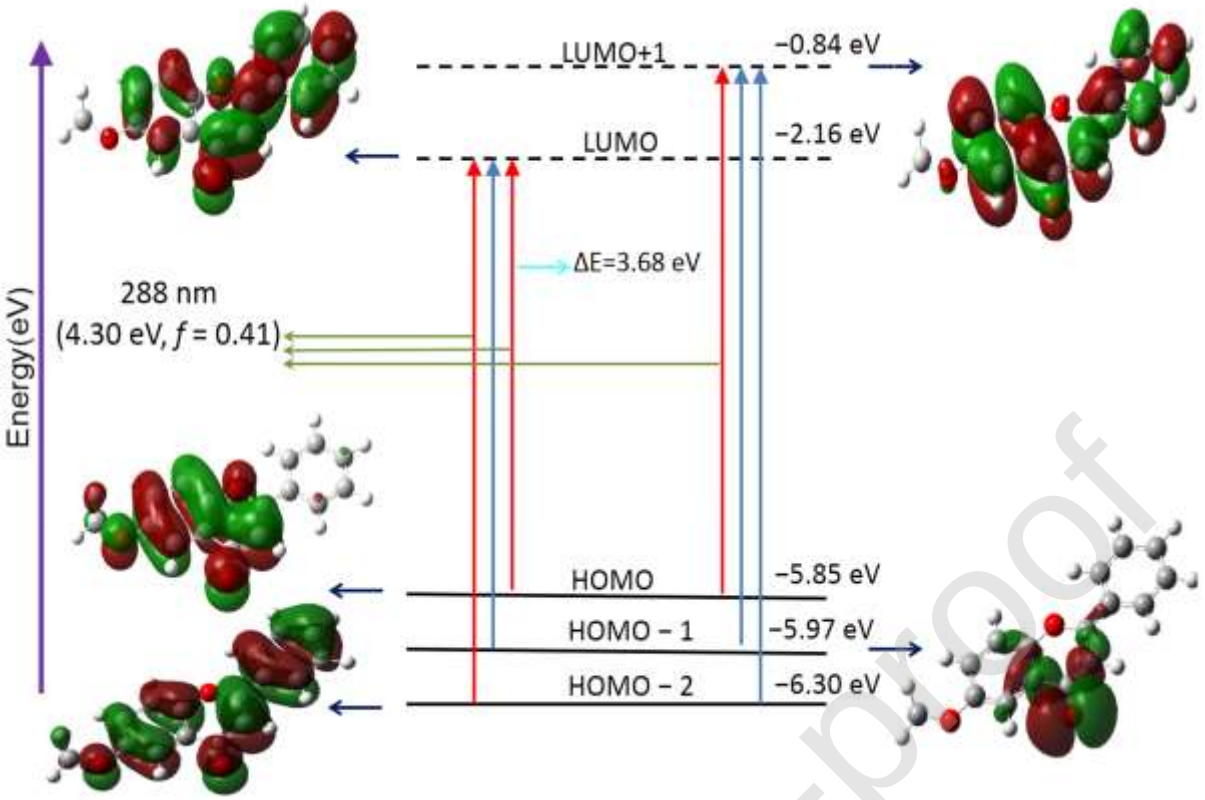
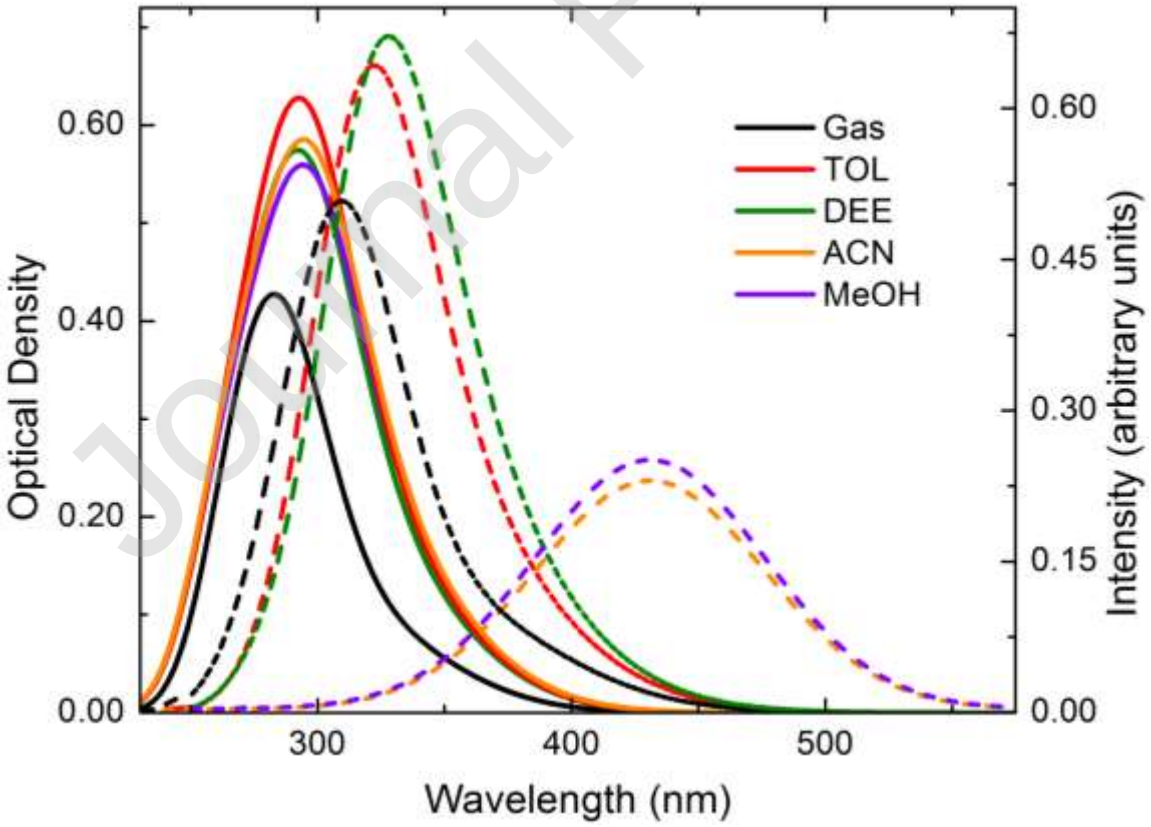


Fig. 5. Molecular orbital diagram of 6MF with HOMO-LUMO density plot and ΔE (HOMO−LUMO) calculated at (TD-DFT/B3LYP/6-31G (d,p)) level of theory.



Journal Pre-proof

Fig. 6. Theoretically calculated vertical transitions/absorption (solid line) and emission (dotted line) spectra of 6MF using TD-DFT in various solvents.

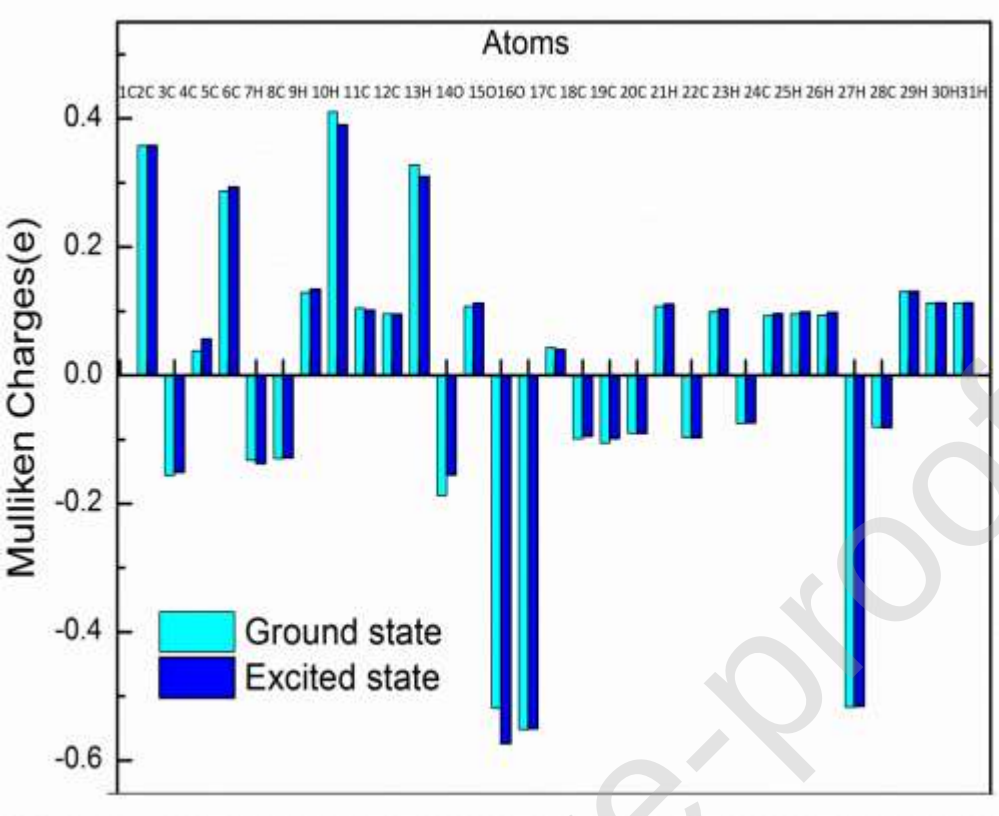


Fig. 7. Comparison of Mulliken atomic charges of 6MF in ground and excited states.

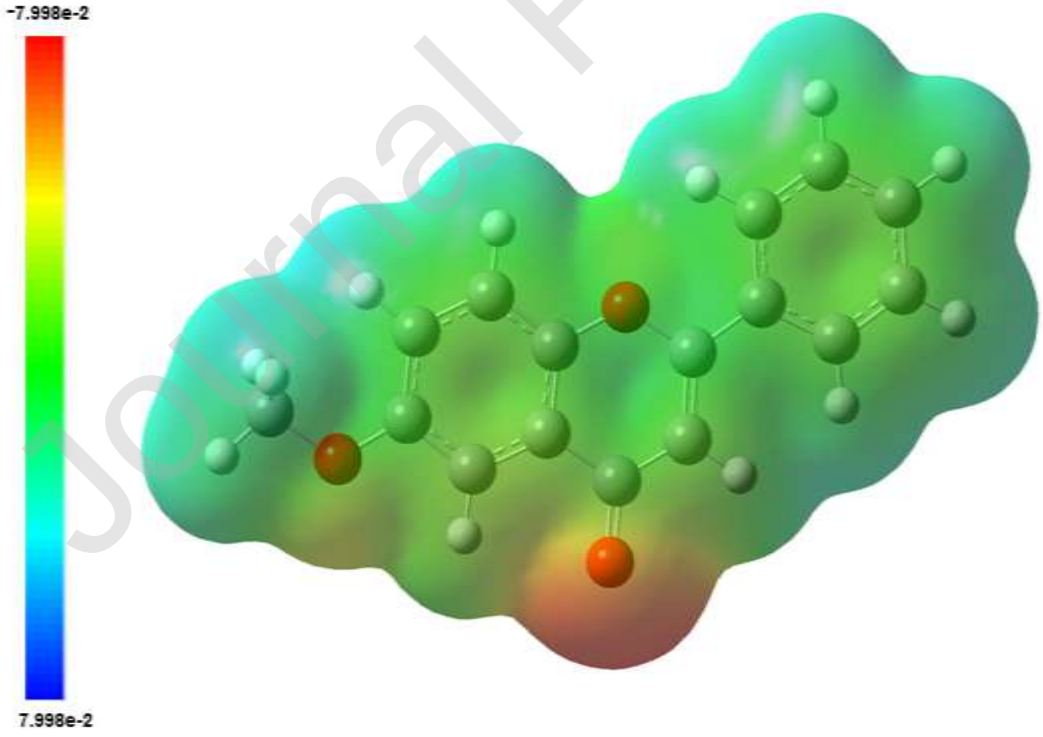


Fig. 8. Molecular electrostatic potential surface of 6MF.

Journal Pre-proof

Table 1. Spectroscopic data of 6MF in different solvents at room temperature

Solvents	λ_a (nm)	ε	$\bar{\nu}_a$ (cm^{-1})	λ_f (nm)	$\bar{\nu}_f$ (cm^{-1})	$\bar{\nu}_a - \bar{\nu}_f$ (cm^{-1})	$\bar{\nu}_a + \bar{\nu}_f$ (cm^{-1})	$\frac{\bar{\nu}_a + \bar{\nu}_f}{2}$ (cm^{-1})	Φ^a	f^b	$\mu_{12}{}^c$
ISCTN	290	1.94	34482	358	27932	6549	62415	31207	0.028	0.517	5.63
DIOXN	300	2.22	33333	358	27932	5400	61266	30633	0.017	0.332	4.59
TOL	302	2.38	33112	410	24390	8722	57502	28751	0.016	0.725	6.80
DEE	292	4.30	34246	351	28490	5756	62736	31368	0.027	0.846	7.23
ETAC	300	6.08	33333	357	28011	5322	61344	30672	0.019	0.293	4.31
THF	301	7.58	33222	380	26315	6906	59538	29769	0.092	0.408	5.09
TFE	306	8.55	32679	446	22421	10258	55101	27550	0.328	0.686	6.66
OctOH	303	10.30	33003	429	23310	9693	56313	28156	0.096	0.559	5.98
HepOH	304	12.00	32894	426	23474	9420	56368	28184	0.116	0.549	5.94
BuOH	303	17.84	33003	434	23041	9961	56044	28022	0.136	0.523	5.79
PrOH	302	20.80	33112	434	23041	10071	56154	28077	0.144	0.536	5.85
EtOH	302	25.33	33112	435	22988	10124	56101	28050	0.147	0.449	5.35
MeOH	303	33.10	33003	444	22522	10480	55525	27762	0.219	0.333	4.62
ACN	299	37.50	33444	428	23364	10080	56809	28404	0.016	0.600	6.16
DMF	303	38.25	33003	433	23094	9908	56097	28048	0.032	0.486	5.58
DMSO	305	47.24	32786	437	22883	9903	55670	27835	0.057	0.372	4.90

[a] Quantum yield was obtained by using Eq. (S25 of Supplementary Information)

[b] Oscillator strength calculated using Eq. (S26)

[c] Transition dipole moment calculated using Eq. (S27)

Journal Pre-proof

Journal Pre-proof

Table 2. Linear plot data for 6MF from Bilot –Kawski, Lippert Mataga, Bakhshiev, Kawski – Chamma –Viallet and Reichardt methods.

Method	Slope(m)	Intercept	Correlation coefficient(R)	No. of data points
Bilot -Kawski	$m_{B-K\,(1)} = 5621.00$	5313.44	0.90	13
Bilot -Kawski	$m_{B-K\,(2)} = -7314.29$	66035.61	0.94	12
Lippert Mataga	$m_{L-M} = 16221.51$	4906.14	0.70	11
Bakhshiev	$m_{B} = 6806.50$	4416.55	0.93	12
Kawski -Chamma - Viallet	$m_{K-C-V} = -8409.62$	33838.00	0.91	14
Reichardt	$m_{E_T^N} = 8100.26$	5258.98	0.90	12

Table 3. Ground and excited state dipole moments of 6MF (μ in Debye), the angle between dipole moments (**φ** in degree) and Onsager radius ('a'in Å).

μ_g^a	μ_e^b	α^c	μ_g^d	μ_e^d	$\mu_{e(L-M)}^e$	$\mu_{e(B)}^f$	$\mu_{e(K-C-V)}^g$	$\Delta\mu^h$	$\frac{\mu_e{}^i}{\mu_g}$	φ^j
5.06	5.79	4.3	1.0	7.67	12.32	8.34	8.21	4.34	7.63	0

[a] calculated at DFT/B3LYP/6-31G(d,p) level using G09 software,
[b] calculated at TD-DFT/B3LYP/6-31G(d,p) level using G09 software,
[c] Onsager radius calculated using Eq. (S24).
[d] obtained using Bilot-Kawski method (Eq. (S7) and (S8)),
[e] calculated using Lippert-Mataga method (Eq. (S14)),
[f] obtained using Bakhshiev method (Eq. (S15)),
[g] calculated using Kawski-Chamma-Viallet method (Eq. (S16)),
[h] difference in dipole moment calculated using Reichardt method,
[i] ratio of ground and excited-state dipole moments obtained using Eq. (S9)
[j] Angle between the ground and excited-state dipole moments obtained using Eq. (S10)

Journal Pre-proof

Table 4. Theoretically calculated spectral parameters in the ground and excited states of 6MF using different theories.

Ground State							
Theory& Basis set				μ_g **(Debye)**			
HF/6-31G(d,p)				5.39			
DFT/B3LYP/6-31G(d,p)				5.06			
Excited State							
Theory/Basis Set	μ_e **(Debye)**	**HOMO (eV)**	**LUMO (eV)**	**ΔE (eV)**	***f***	**Excitation energy(eV)**	λ_{abs} **(nm)**
TD-DFT/B3LYP/6-31G (d,p)	5.79	−5.85	−2.16	3.68	0.506	4.30	288
TD-DFT/TZVP	5.39	−6.26	−2.19	4.07	0.407	4.20	294

Table 5. Calculated electronic transition energies and corresponding oscillator strengths of the low-lying singlet excited states and dipole moment of 6MF using TD-DFT/6-31G(d,p) level of theory.

Electronic Transition	**Energy nm (eV)**	**Oscillator strength**	**Contribution**	**CI%**
$S_0 \rightarrow S_1$	361 (3.43)	0.000	H−1→L	95.62
			H→L+1	02.36
$S_0 \rightarrow S_2$	335 (3.69)	0.070	H−2→L	09.46
			H→L	87.75
$S_0 \rightarrow S_3$	288 (4.30)	0.415	H−2→L	84.05
			H→L	07.29
			H→L+1	05.80
$S_0 \rightarrow S_4$	269 (4.60)	0.205	H−1→L	02.62
			H−1→L+1	95.22
$S_0 \rightarrow S_5$	265 (4.67)	0.014	H−3→L	94.11

Journal Pre-proof

H−2→L+1	02.21

Table 6. Global reactivity descriptors for 6MF calculated at TD-DFT/B3LYP/6-31G(d,p) level of theory.

Parameters (in eV)	Value
Ionisation potential(IP) = −HOMO	5.85
Electron affinity(EA)= −LUMO	2.16
Electronegativity(χ)= $\frac{IP+EA}{2}$	4.01
Chemical Potential(μ) = −χ	−4.01
Chemical Hardness(η) = $\frac{IP-EA}{2}$	1.84
Chemical Softness(s) = $\frac{1}{2\eta}$	0.27

Table 7. Spectral parameters of 6MF in the ground and excited-state based on quantum chemical calculations in the gas phase and different solvents.

Solvents	Ground State DFT/6-31G(d,p)	Excited State TD-DFT /6-31G(d,p)								
	μ_g (Debye)	μ_e (Debye)	HOMO (eV)	LUMO (eV)	ΔE (eV)	λ_{abs} (nm)	λ_{em} (nm)	Excitation energy (eV)	*f*	Transition Dipole moment
Gas	5.06	5.79	−5.85	−2.16	3.68	288	312	4.30	0.415	3.94
Toluene	5.95	6.85	−5.95	−2.31	3.63	297	322	4.17	0.601	5.88
Diethylether	6.40	7.37	−6.00	−2.24	3.75	297	327	4.16	0.562	5.51
Acetonitrile	7.10	7.32	−5.68	−2.21	3.39	299	427	4.14	0.560	5.52
Methanol	7.09	7.31	−5.68	−2.21	3.39	299	427	4.14	0.555	5.47

Journal Pre-proof

Journal Pre-proof

Table 8. Estimated coefficients (ν_0, $a\alpha$, $b\beta$, $c\pi^*$), their errors and correlation coefficients (R) for multi-linear correlation analysis of ν_a, ν_f, $\Delta\nu$ of 6MF as a function of Kamlet-Abboud-Taft solvent scale (Eq.(S22)).

Parameters	**ν_0**	**$a\alpha$**	**$b\beta$**	**$c\pi^*$**	**R**
ν_a	34172 ± 283	$-(425\pm133)$	$-(204\pm241)$	$-(1206\pm360)$	0.80
ν_f	30804 ± 1449	$-(2934\pm688)$	$-(2144\pm1300)$	$-(6592\pm1878)$	0.86
$\Delta\nu$	3373 ± 1414	2528 ± 671	2026 ± 1268	5310 ± 1833312	0.83

Table 9. Estimated coefficients (ν_0, aSA, bSB, cSP, dSdP) their errors and correlation coefficients (R) for multi-linear correlation analysis of ν_a, ν_f, $\Delta\nu$ of 6MF as a function of Catalan -four-parameter solvent scale (Eq.(23)).

Parameters	**ν_0**	**aSA**	**bSB**	**cSP**	**dSdP**	**R**
ν_a^*	35612 ± 321	$-(824\pm143)$	$-(103\pm139)$	$-(398\pm259)$	$-(2785\pm403)$	0.96
$\nu_f^{\#}$	26540 ± 706	$-(1194\pm253)$	(408 ± 166)	$-(3573\pm470)$	$-(305\pm847)$	0.97
$\Delta\nu$	5655 ± 5661	4923 ± 2341	$-(384\pm2064)$	8531 ± 5150	9388 ± 7886	0.80

*except for ISCTN, DEE, HepOH, OctOH, TFE

except for ISCTN, DEE, Dioxane, EtOAc, THF

Journal Pre-proof

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344363046

Explicit finite element analysis of single pass rolling using ABAQUS CAE

Article · January 2020

DOI: 10.32438/WPE.0602119

CITATIONS
0

READS
2

2 authors:

Sanjay Kumar
Delhi Technological University
90 PUBLICATIONS **4,100** CITATIONS
SEE PROFILE

Pranjal Shiva
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Design and Fabrication of wing of a high payload Aerial vehicle View project

Single-Cell Western Blotting View project

All content following this page was uploaded by Pranjal Shiva on 24 September 2020.

The user has requested enhancement of the downloaded file.

WEENTECH Proceedings in Energy ISSN:2059-2353

Available online at www.weentechpublishers.com

WEENTECH Proceedings in Energy 6 (2) (2020) 124-132

International Conference on Recent Advances in Mechanical Engineering, Department of Mechanical Engineering,
Delhi Technological University, New Delhi, India

Explicit finite element analysis of single pass rolling using ABAQUS CAE

Sanjay Kumar, Pranjal Shiva*

Department of Mechanical Engineering, Delhi Technological University, Delhi-110042, India

Corresponding author's mail: shiva.pranjal@gmail.com

Abstract

The paper presents a finite element analysis of micro-alloyed steel and Aluminium 2124 which are subjected to the rolling pressure during the manufacturing process using a commercial software package - ABAQUS CAE. The plasticity properties of the materials are obtained from an experimental stress-strain curve. The conditions similar to the rolling process are imposed on the model, wherein the roller is given different angular velocities. Contact properties are also imposed between the plate and roller. The analysis is carried out using an explicit solver as it is more consistent and computationally less expensive than an implicit solver. Further, Von-Mises stress and deformation contours in the y-axis are requested as output. Relative deformations are compared for three different roller speeds. Also, kinetic energy, total energy, internal energy, and external work plots are extracted and plotted for both the materials. Therefore, the research presented in the paper is a novel method that reduces the need of conducting various physical experiments to validate the material's response.

Keywords: Rolling process; Finite Element Method; Explicit Analysis; Micro-Alloyed steel; Al2124.

Copyright © 2020 Published by WEENTECH Ltd. All Peer-review process under responsibility of the scientific committee of the International Conference on Recent Advances in Mechanical Engineering, Department of Mechanical Engineering, Delhi Technological University, New Delhi, India, RAME2020

Abbreviations

https://doi.org/10.32438/WPE.0602119
Manuscript History
Receipt of completed manuscript: 10 June 2020
Receipt of Revised Manuscript: 16 July 2020
Date of Acceptance: 20 July 2020
Online available from: 18 September 2020

S=Stress
U=Displacement
U2= Displacement in Y-direction
R=radius
2-D= two dimensional
MPa= Mega Pascal

1. Introduction

Rolling is a type of manufacturing process wherein the metal stock is forced to pass between two rollers to reduce the cross-sectional area (or thickness) and subsequently increase the length of the material. The rollers used for the metal forming process rotate in opposite directions [1], [2]. This research paper illustrates a method for analyzing the rolling process using Finite Element Analysis, performed using ABAQUS CAE software package [3]. The materials analyzed in the research are micro-alloyed steel and Aluminium 2124. Steel has a wide variety of applications ranging from automotive to power generation industries whereas Aluminium is extensively used in aircraft industries due to light-weight and good response to fatigue loading [4]. The primary objective of the research is to identify the changes observed in the behavior of forces and energy of a thick steel plate, subjected to rolling, and also compare the relative deformation with an increase in the roller's angular velocity.

2. Material composition and properties

The various properties of both materials used for finite element analysis are entered from [5]:

Table 1 Composition of Aluminium 2124

Element	Weight Percentage (%)
Copper	5.12
Zinc	1.13
Manganese	0.592
Iron	0.565
Magnesium	0.454
Silicon	4.09
Titanium	0.112
Tin	0.112
Lead	0.044

Table 2 Composition of Micro-Alloyed steel

Element	Weight Percentage (%)
Carbon	0.12
Manganese	1.6
Phosphorous	0.025
Sulphur	0.015
Copper	0.5
Nickel+Chromium+Molybdenum	0.5
Niobium+Titanium+Vanadium	0.15

Table 3. Plastic properties of Aluminium 2124

Yield Stress (MPa)	Plastic Strain
85	0
85	0.00369
106.8	0.00389
116	0.00423
121	0.00565
126	0.00924
137	0.01156
142	0.0137
153	0.0185
158	0.02
162	0.0222
165	0.0247
166	0.0256
173	0.028

Table 4. Plastic properties of Micro-alloyed steel

Yield Stress (MPa)	Plastic Strain
482.94	0
482.94	0.004507
492.92	0.008687
505.73	0.008948
518.61	0.01295
526	0.01632
532.86	0.019674
545.83	0.02749
554.3	0.03391
567.14	0.04524
576.76	0.05807

3. Modelling

3.1 Part:

The model contains two parts – plate and roller.

Plate: It is modelled as a 2-D deformable shell-type object with dimensions as mentioned in the Fig.1and Fig.2.

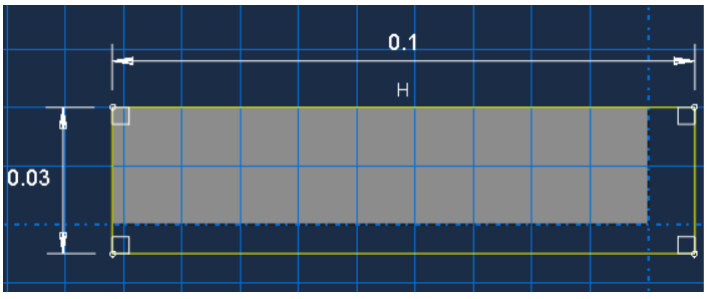


Fig. 1 Dimensions of plate

Roller: Owing to symmetry, a 90-degree quarter arc with a radius of 0.17 m is sketched as an analytical rigid part type.

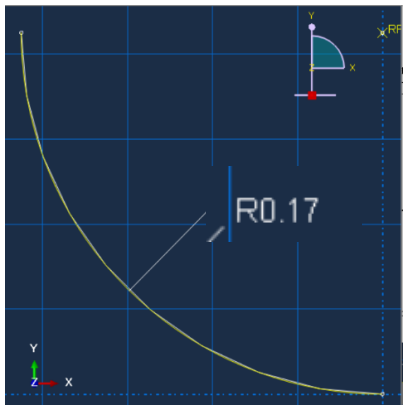


Fig. 2 Roller geometry

3.2 Material properties:

As the material is undergoing non-linear deformation, both the elastic and plastic properties of Micro-alloyed steel (Table 3) and Aluminium 2124 (Table 4) are directly imported in the software as a stress-strain curve and appropriate corrections are applied. The density and Poisson's ratio are also entered subsequently. Furthermore, the plate is assigned a homogeneous solid section using the section assignment tool.

3.3 Assembly:

The individual instances (plate and roller) are assembled in the assembly module, as shown in Figure 3. Upon investigation, it was found that the plate first touches the roller at a point located at a circumferential distance of 0.053 m from the bottom point of the roller. This constraint is used to define an initial configuration before the start of the rolling process.

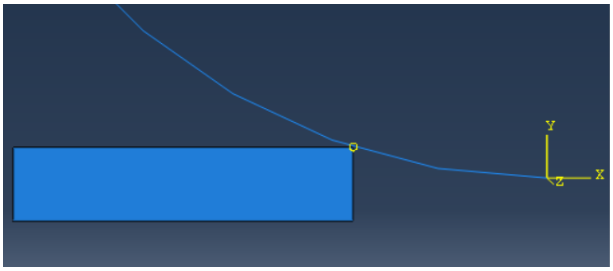


Fig. 3 Assembled geometry

3.4 Step definition:

For the first analysis, an explicit dynamic step is created having a time period of 0.1s, which is sufficient to capture the entire rolling process. Additionally, the mass scale factor for the plate is entered as 1600.

3.5 Contact pairs:

The tangential behavior type interaction property is defined as having a coefficient of friction of 0.4. Subsequently, a surface-to-surface contact is imposed between the roller and the

material to be rolled. This is done by selecting the roller (rigid part) as the master surface and plate as the slave surface.

3.6 Loads and Boundary conditions:

The roller is given three different angular velocities – 3.14 rad/s, 6.28rad/s, and 9.42 rad/sec to compare the change in deformation with velocities. The translational movement of roller is fixed in both x & y directions by imposing boundary conditions on the reference point. Furthermore, the plate is given an initial velocity of 1.01 m/sec in the x-direction.

4 Results and Discussions

The following plots are obtained from the above simulation for Aluminium 2124 and micro-alloyed steel:

Figure 4 depicts the deformation of micro-alloyed steel in y-direction at a roller speed of 6.28 rad/s. Moreover, as shown in figure 5, the maximum deformation (8.3 mm) is achieved after a time interval of 0.09s, 0.06s, and 0.035s for angular velocities of 3.14 rad/s, 6.28 rad/s, 9.42 rad/s.

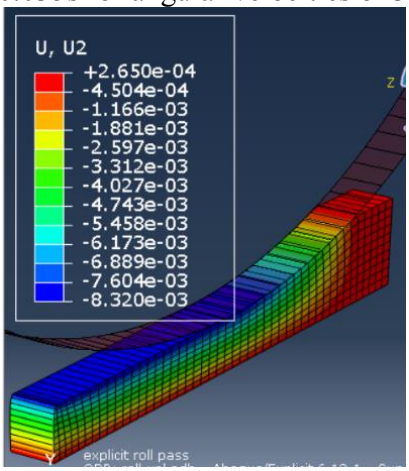


Fig. 4 Deformation contours in Y-axis for micro-alloyed steel at 6.28 rad/s

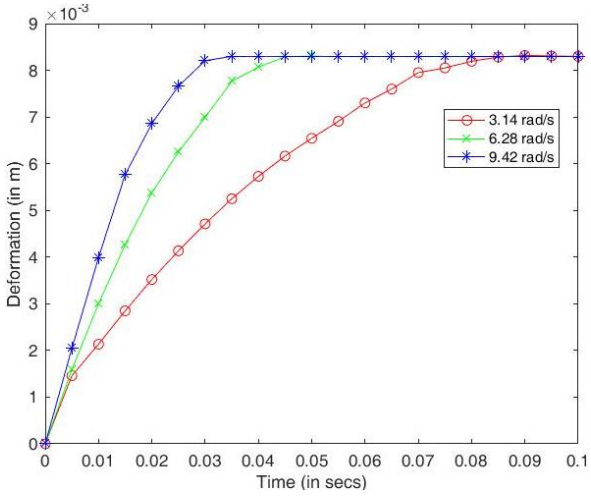


Fig. 5 Deformation vs time plot for different roller speeds for Micro-alloyed steel

Figure 6 shows the von mises stress developed in the Micro-alloyed steel at a roller speed of 6.28 rad/s. The maximum von mises stress developed for all three speeds is 576.8 MPa.

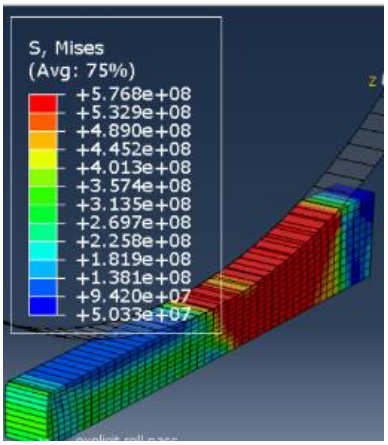


Fig. 6 Von-mises stress contours for Micro-alloyed steel at 6.28 rad/s

Figure 7 depicts the deformation of Aluminium 2124 in y-direction at a roller speed of 6.28 rad/s. Moreover, as shown in figure 8, the maximum deformation (8.3 mm) is achieved after a time interval of 0.085s, 0.045s, and 0.03s for angular velocities of 3.14 rad/s, 6.28 rad/s, 9.42 rad/s.

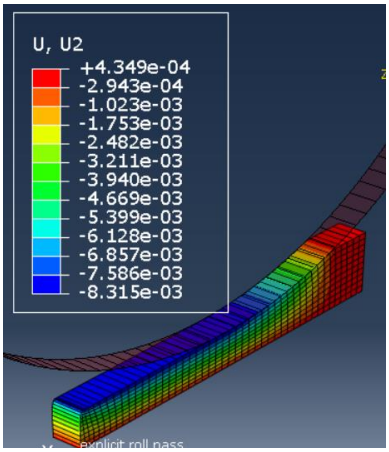


Fig. 7 Deformation contours in Y-axis for Aluminium 2124

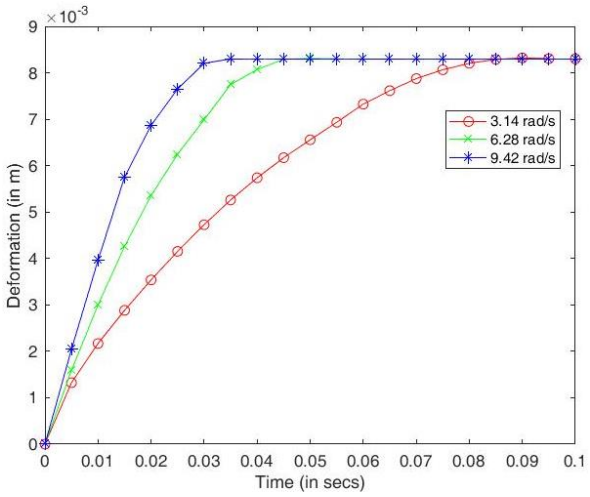


Fig. 8 Deformation vs time plot for different roller speeds for Aluminium 2124

Figure 9 shows the von mises stress developed in the Aluminium 2124 at a roller speed of 6.28 rad/s. The maximum von mises stress developed for all three speeds is 173 MPa.

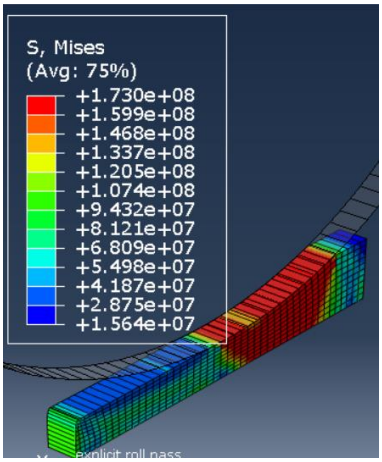


Fig. 9 Von-mises stress contours for Aluminium 2124

Finally, Figure 10 & 11 show the plot of different energies such as external work, internal energy, kinetic energy & total energy, produced during the rolling process. It is also seen that these energies produced in Micro-alloyed steel are higher than Aluminium 2124.

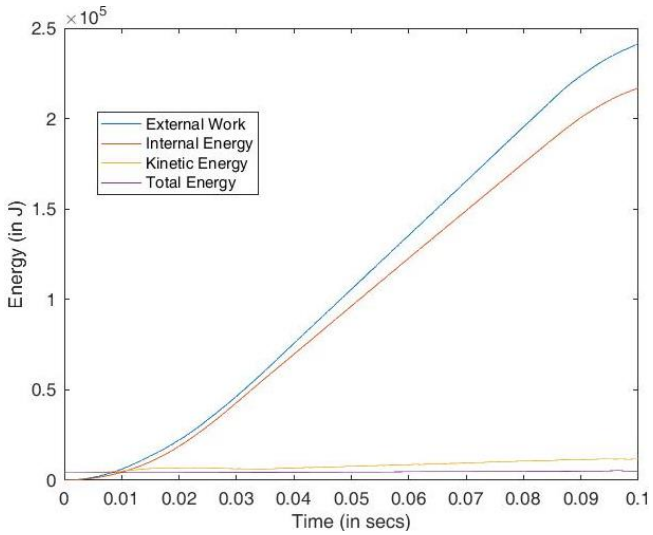


Fig. 10 Energy vs time plot for Aluminium 2124

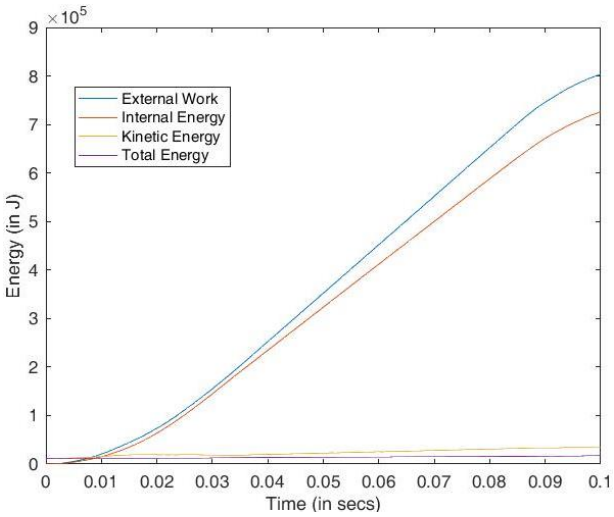


Fig. 11 Energy vs time plot for Micro-alloyed steel

5 Conclusions and inferences

In this research two different materials, Micro-alloyed steel and Aluminium 2124 have been rolled and the effect of roller speeds on the deformation of materials as well as the induced Von-Mises stress are studied. The total deformation was observed to be constant at 8.3 mm for all roller speeds and materials, indicating that the deformation is independent of material properties and roller speeds. On the other hand, the Von-Mises stress generated in Micro-Alloyed steel is more than that in Aluminium 2124. Finally, the energies induced in materials are also evaluated. It is found that with an increase in roller speeds, the time required for reduction in thickness is reduced.

ORCID Id of authors

Shiva, Pranjal: https://orcid.org/0000-0002-1157-7046
Kumar, Sanjay: https://orcid.org/0000-0001-9310-1262

References

[1] Vimalanathan, Srinivasan, and Laine Mears. Finite Element Simulation of Ring Rolling Process. No. 2010-01-0270. SAE Technical Paper, 2010.
[2] Guo, Lianggang, He Yang, and Mei Zhan. "Research on plastic deformation behaviour in cold ring rolling by FEM numerical simulation." Modelling and simulation in materials science and engineering 13.7 (2005): 1029.
[3] Min, Wang, Yang He, Zhi-Chao Sun, Liang-gang Guo, and Xin-zhe Ou. "Dynamic explicit FE modeling of hot ring rolling process." Transactions of Nonferrous Metals Society of China 16, no. 6 (2006): 1274-1280.
[4] Utsunomiya, Hiroshi, Yoshihiro Saito, Tomoaki Shinoda, and Ichiro Takasu. "Elastic–plastic finite element analysis of cold ring rolling process." Journal of Materials Processing Technology 125 (2002): 613-618.
[5] Pndya, Nivesh, Sehgalm D.K., Pandey R.K, M.Tech Thesis," Determination of the fracture properties of metallic alloys using miniature specimen technique and finite element simulation" I.I.T, Delhi 2010.Wang, T.W. (2007) A case study of gas pipe explosion. Journal of Chinese army police force academy, 23, 55-56.
[6] Han, Y.L., Chen, LZ. (2008) A review on explosion and the prevention in residential buildings. Journal of disaster prevention and mitigation engineering. 28, 84-88.

© 2020 by the authors. Licensee WEENTECH Publishers, Coventry, United Kingdom. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/)

View publication stats

Fabricating and Assembling Acoustic Metamaterials and Phononic Crystals

Christabel Choi, Shubhi Bansal, Niko Münzenrieder, Sriram Subramanian**

C. Choi, Dr. S. Bansal, Prof. S. Subramanian

Department of Computer Science, Faculty of Engineering, University College London, London WC1E 6BT, UK.

E-mail: shubhi.bansal@ucl.ac.uk, s.subramanian@ucl.ac.uk

Prof. N. Münzenrieder

Faculty of Science and Technology, Free University of Bozen-Bolzano, Bozen, 39100, Italy

Keywords: metamaterials, acoustics, fabrication

Acoustic Metamaterials (AMM) and Phononic Crystals (PC) have the potential to unfold a new wave of disruptive technologies to radically transform human interactions, sensory communications, and beyond. Although essential, cultivating a deep understanding of fundamental theory and design principles is insufficient alone, in the practical advancement of AMMs and PCs. Equally important is the physical realization of these artificial structures for tangible prototyping and experimental investigation, however such aspects are seldom discussed in literature. This review critically examines fabrication and assembly approaches for AMMs and PCs, with a tight-coupling of theoretical and experimental considerations. Crucial parameters like operating frequency, materials, and geometry for efficient structural implementation are addressed. Herein, fabrication methods for specific structure-types are categorized under *single-step fabrication* including printing and machining, and *multi-step fabrication* like microfabrication and molding. Various *assembly* techniques are proposed, such as for ordering colloidal assemblies or fastening components without adhesives. This framework uncovers innovative designs, e.g. origami-based structures with

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/adem.202000988.

This article is protected by copyright. All rights reserved

Accepted Article

conductive coatings, only accessible if fabrication and assembly aspects form an integral part of the initial design phase. By establishing a greater understanding and awareness of these methods, a host of undiscovered pathways, opportunities and research gaps are revealed, supporting a fresh paradigm for innovation.

Accepted Article

1. Introduction

Acoustic Metamaterials (AMMs) and Phononic Crystals (PCs) have garnered significant attention in recent years, as part of the collective driving force towards creating intelligent acoustic devices. Advancements in these fields have greatly enhanced the way we manipulate sound waves through transmission, reflection, refraction, absorption, diffraction, or attenuation. In the past decade, AMMs and PCs have enabled novel applications such as acoustic lensing,[1–3] cloaking,[4] levitation,[5] and holography.[6–8] While these exotic structures have been well-explored through theoretical and numerical analysis,[9–13] their physical realization is an important topic that is rarely discussed.

Considering the ubiquity of sound and the powerful capabilities of AMMs and PCs, the impact of these acoustic structures could be phenomenal. Across the full acoustic frequency spectrum, practical applications such as noise cancellation,[14] underwater detection,[15] medical imaging,[16] and energy harvesting,[17,18] could benefit key sectors in our society like healthcare, well-being, environmental sustainability and security. Moreover, AMMs and PCs can help to usher in next-generation technologies for personalized, immersive multi-sensory[19–22] experiences. The manipulation of sound can enrich the way we communicate and interact with our surroundings, not simply through audio, but also through tactile sensations. In the future, AMMs and PCs could be used in virtual reality (VR) setups,[23] compact wearable devices, and dynamic mid-air volumetric displays[24] that are controllable and capable of providing haptic feedback.[25] Beyond the notion that AMMs and PCs can replace phased arrays, they could readily complement one another for

This article is protected by copyright. All rights reserved

more precise control. In commercial devices, AMM and PC functionalities could even be combined together in different ways, e.g. transmissive and sound absorptive structures, for improved performance. To unlock the full potential of AMMs and PCs, it is therefore vital to ensure that practical, physical realization is pursued alongside theoretical investigation in the development of viable acoustic designs.

Building an AMM or PC requires some form of fabrication or assembly, or both. Fabrication refers to the technologies and processes used to manufacture an object, while assembly refers to the strategic amalgamation of parts for a constructive purpose. AMMs and PCs can also be assembly-free, i.e. the complete structure can be directly and monolithically fabricated. More often than not, extensive design requirements can make building physical prototypes a challenging task. Desired properties, like perfect absorption or negative refraction, are mainly governed by different mechanical, geometric, and material characteristics. Hence, any discrepancies arising from imperfections in fabricated structures can negatively influence their acoustic performance. Moreover, dynamic AMM and PC designs can increase building complexity by incorporating stimuli-responsive structures and actuation mechanisms.[26,27] Prototyping multiple iterations for optimization can consume enormous time, labor, and monetary resources, which we believe can be alleviated by considering relevant physical and experimental aspects from the initial design phase.

In this review, we critically discuss various fabrication and assembly methods to build AMMs and PCs. We first establish popular structure-types that form the building blocks of AMMs and PCs, and discuss how structural properties and device operation can point towards appropriate fabrication techniques. Subsequently, we classify viable fabrication techniques into the categories of *single-step* and *multi-step* fabrication. Under single-step fabrication, we firstly analyze three-dimensional (3D) printing and machining technologies appropriate for AMMs and PCs operating in the audible (20 Hz – 20 kHz) or ultrasonic (20 kHz – 100 MHz) regimes. Then, hybrid manufacturing processes and their implications are briefly examined for more efficient structural

This article is protected by copyright. All rights reserved

Accepted Article

implementation. Under multi-step fabrication, we see a range of approaches for fabricating nanoscale, heterogeneous, or soft-matter based designs which mainly range from ultrasonic to hypersonic (> 100 MHz) operating frequencies. These approaches involve microfabrication, molding and casting, and microfluidic-based wet chemical techniques. Thereafter, we present several assembly methods which encompass the use of adhesives, self-assembly, and fitting techniques to strategically construct AMM and PC assemblages in random or ordered arrangements. We conclude with a comprehensive outlook on further applications and future prospects. Our discussions of the relevant techniques, along with underutilized materials and research gaps, thereby provide a roadmap promoting the development of advanced AMMs and PCs.

2. Types of Structures

AMMs and PCs are discussed together as they often share similar fabrication approaches and acoustic applications. Both are composed of unit cells, with each unit cell consisting of elements unique to their design. In operation, these elements interact with the impinging sound waves. The unit cells in a PC are always periodic, and the size of their elements are on an order of the operating wavelength. However, the unit cells in an AMM can be either periodic or non-periodic, and their elements are smaller than the operating wavelength (i.e. sub-wavelength).[28] Planar AMMs and PCs, known as acoustic metasurfaces[29,30] have been gaining traction in recent years. The unit cells in these two-dimensional (2D) devices are structured in arrays, and their appealing thin, compact design offers greater versatility as compared to their bulk counterparts for wave engineering applications, especially for long wavelengths. For example, phase discontinuities can be engineered at the interface between media for targeted acoustic focusing.[31]

Traditionally, AMMs have operated through locally resonating elements,[32] as seen in colloidal suspensions **(Figure 1a)**,[33–35] Helmholtz resonator **(Figure 1b)**,[36–38] and membrane-type **(Figure 1c)**,[39,40] structures. PC lattices **(Figure 1d)**,[41,42] and certain AMMs,[3] rely on the mutual interaction between neighboring periodic elements within a matrix. Bandgaps that forbid acoustic

This article is protected by copyright. All rights reserved

Accepted Article

transmission are formed through multiple scattering and interference of propagating waves. Other types of AMMs may not operate through resonance, but by coiling space in labyrinthine structures **(Figure 1e)**,[43,44] or wave-guiding using origami-based structures **(Figure 1f)**.[45–47] AMM design generally revolves around the parameters of effective mass density (ρ) and bulk modulus (K),[32] where they can be singly-negative (either $\rho < 0$ or $K < 0$) or doubly-negative (both $\rho < 0$ and $K < 0$), depending on the type of structure.[48,49] While theoretically defining such parameters, we must determine the feasibility of building such structures. Numerical predictions may characterize a design for a certain scale, which may not be easily achievable experimentally. For instance, reconfigurable origami barriers for traffic noise reduction[50] were scaled down to one-seventh the original size, promoting greater experimental feasibility and efficiency in terms of power and cost.

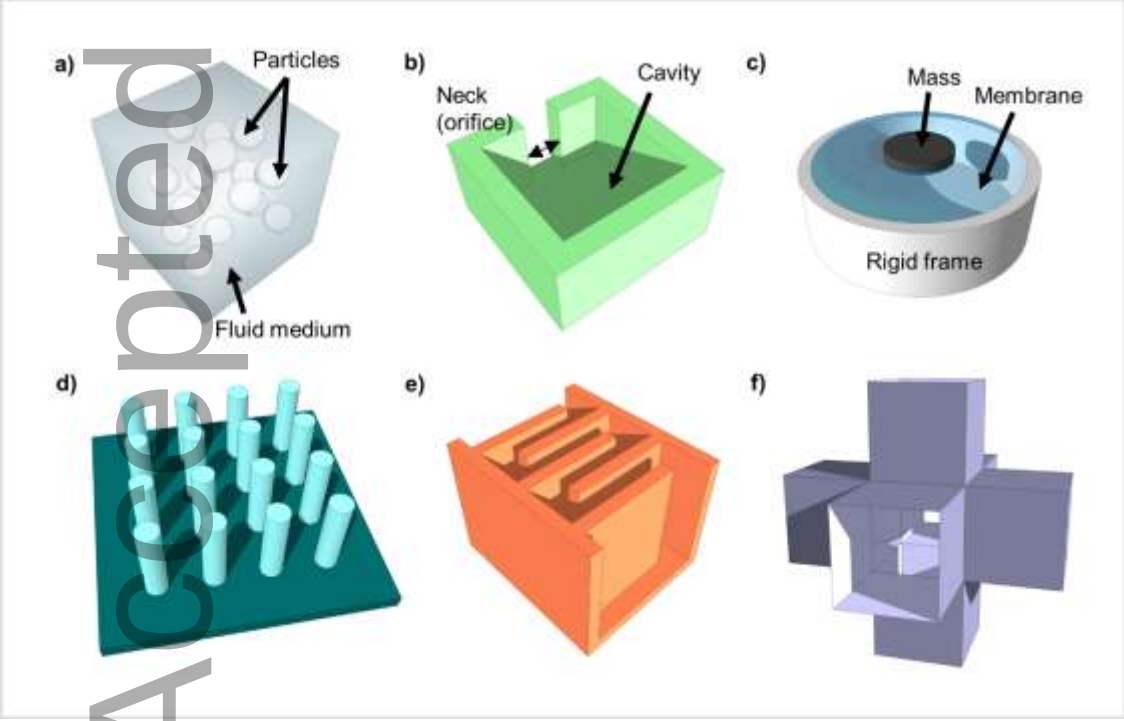


Figure 1. Typical AMM and PC structures. Figure illustrates general examples of six common structures found in Acoustic Metamaterial (AMM) and Phononic Crystal (PC) designs. **a)** Colloidal metamaterial consisting of spherical particles within a fluid medium. **b)** Helmholtz resonator structure with a cavity and a narrow orifice. **c)** Membrane-type structure with a mass at the center of the membrane, surrounded by a rigid frame. Adapted with permission.[51] Copyright 2014, Springer

This article is protected by copyright. All rights reserved

Accepted Article

Nature Limited. **d)** Lattice structure consisting of an array of cylindrical pillars. **e)** Labyrinthine space-coiling structure. **f)** Deformable origami-based waveguide. Adapted under terms of the CC-BY license.[45] Copyright 2016, The Authors, published by AAAS.

The six common AMM and PC structures shown in **Figure 1** represent very broad categories of designs. In particular, there has been a growing interest in lattice-type acoustic structures that allow the topologically-protected propagation of sound waves, where propagation is forbidden in the bulk but permitted on the surface. Various topological phases have been reported, such as quantum spin Hall effect topological insulators (TI),[52,53] acoustic valley Hall TIs,[54,55] and Weyl PCs.[56,57] Higher-order Wannier-type TIs[58–61] can exhibit unique boundary states in different dimensions, such as edge, corner, and hinge states. Acoustic TIs enable unusual capabilities, such as immunity to backscattering and ultra-broadband transmission.[62] Gradient refractive index (GRIN)[63–65] AMMs and PCs also tend to assume lattice configurations, and they are typically designed by spatially varying parameters such as the lattice constant, or element-size. The elements are non-resonant, and devices can be broadband.

Another class of AMMs is acoustic metafluids,[66–68] which have anisotropic mass densities and acoustic elastic properties resembling fluids, with extraordinary broadband capabilities. These structures provide the ability to cloak objects, which is valuable for acoustic stealth and sensing applications. Lastly, for hypersonic acoustic manipulation, nanopatterned surface PCs with enhanced surface wave confinement have been developed for gigahertz wave generation, and sophisticated nanoscale sensing techniques.[69–72] Besides the selected structures discussed in this review, the covered building methods are applicable to a myriad of other reported theoretical designs in literature.

Accepted Article

This article is protected by copyright. All rights reserved

When choosing appropriate fabrication and assembly methods, it is essential to draw connections between the way they operate and the physical characteristics of the design. By assessing the specific building methods for these structures, we identify key features that define their usability and suitability. Subsequently, this knowledge can be extended towards developing and realizing other acoustic designs. For instance, labyrinthine structures impose path differences for delayed sound propagation to achieve phase or amplitude modulation. Physically, they are characterized by complex internal geometries. If they use high operating frequencies, elements are likely to be in the size-range of millimeters or centimeters. To facilitate acoustic transmission or reflection, a material with a high acoustic impedance is preferred. With an understanding of available methods, we can deduce that a high-resolution 3D printing technology like Polyjet printing, would be a suitable choice. With Polyjet printing, rigid spiraling or meandering features can be captured with high accuracy. If similar geometric complexity is found in other devices e.g. Helmholtz resonators, the same printing technology can be applied with confidence.

Through our analysis of different structure-types, we examine fabrication methods for AMMs and PCs and categorize them into *single-step* and *multi-step* techniques. Single-step fabrication can replicate elements with high precision and accuracy, compatible with structures consisting of ordered and repeating elements. Multi-step fabrication is applied for structures that are highly heterogeneous, complex, or consist of various minute features on the micro or nanoscale.

3. Single-Step Fabrication

In single-step fabrication, only one process (besides post-processing) is undertaken to fabricate either a full AMM or PC, or parts of it for further assembly. 3D printing and machining are the most standard techniques for this purpose, ideal for producing designs involving Helmholtz resonators,[73,74] labyrinthine structures,[43,44,75,76] and periodic lattices.[77,78] These techniques are accessible, cost-effective, and simple to use. Here, rapid prototyping is possible through Computer-Aided Design (CAD), compatible with both printing[79] and Computer Numerical Control (CNC)

This article is protected by copyright. All rights reserved

Accepted Article

machining technologies. Geometric modifications for multiple prototypes can be easily done by changing the design file, simplifying the fabrication of gradient or broadband designs with elements of varying shapes or sizes.

3.1. Printing technologies

Three dimensional (3D) printing is one of the most popular ways to fabricate AMMs and PCs because of its unparalleled manufacturing flexibility. The inherent limitations and operational style of a 3D printer can affect the type of materials, geometries, size, and resolution of the prints, including post-processing requirements. We propose a classification of 3D-printed AMM and PC structures, into *microstructures* and *macrostructures*, to both categorize selected 3D printing technologies, and emphasize their unique strengths and limitations. Direct Ink Writing (DIW), Stereolithography (SLA), Digital Light Processing (DLP), and Polyjet printing are high-resolution technologies to support *microstructures*, while Fused Deposition Modelling (FDM) and powder-bed fusion techniques are versatile and cost-effective choices for *macrostructures*. Here, 'structures' refers to unit cells or elements.

AMM and PC *microstructures* are typically designed with features corresponding to wavelengths between centimeters and micrometers. AMMs and PCs have unit cells less than or equal to the wavelength of the chosen operating frequency, therefore elements interacting with high-frequency sound waves are structurally smaller (and vice versa) i.e., ultrasonic (from 20 kHz to 10 MHz) designs would use wavelengths from 1.7 cm down to 34.3 μm. For fabricating *microstructures*, we would generally select a high-resolution printer for a well-defined print with small physical features, and a smooth surface finish to prevent unwanted sound attenuation and reflections.

Technologies like DIW allow a broad range of materials to be printed using customized inks such as graphene-aerogels,[80] liquid metal,[81] and shape-morphing liquid crystal elastomers.[82] Dimensions of printed parts typically fall within the range of 100 μm–10 cm, but fine deposition nozzles can have diameters of 10 μm for higher resolution printing. DIW technology can fabricate

This article is protected by copyright. All rights reserved

Accepted Article

ultrasonic ceramic-based PCs as tight scaffolds of micron-scale rods **(Figure 2a)**[41,83] for sound damping (100 kHz – 800 kHz) and acoustic focusing (2 MHz – 12 MHz) applications. Here, each extruded line corresponds to the width of one rod. To date, DIW has rarely been used to fabricate AMMs, and non-ceramic materials are yet to be applied to print PCs.

The benefit of high-resolution printing, like SLA (e.g. 25 μm – 100 μm, Formlabs Form 3), is the ability to capture subtle geometric differences between elements and unit cells. A special kind of GRIN device is known as a Luneburg lens, which allows beam focusing in multiple directions.[84,85] An ideal spherical geometry would require a radially varying index profile from the center, which is challenging to fabricate as the element-size is not purely homogeneous throughout the structure. SLA technology was first used to demonstrate a 3D Luneburg lens prototype for 8 kHz,[63] while Polyjet printing later enabled the experimental realization of a flattened design for ultrasonic (40 kHz) beam steering with an index ranging from 1 to 1.82, and a periodicity of 2 mm.[84] High-resolution SLA printing also allowed the direct fabrication of Helmholtz resonators with varying geometries in a single device for rainbow-trapping applications. Forty gradient resonators with different cavity sizes were fabricated to mimic a human cochlear **(Figure 2b)**, inducing the formation of bandgaps corresponding to different frequencies (1 kHz – 10 kHz) for sound filtering functionalities.[74] Another design used flat compact panels of asymmetric Helmholtz resonators for perfect broadband absorption (absorption coefficient, $\alpha = 1$) between 300 Hz and 1000 Hz.[86]

Vat-polymerization techniques can replicate identical repeating unit cells as part of complex lattice configurations, such as third-order TIs. An anisotropic rhombohedron TI lattice was directly printed using SLA, consisting of a combination of coupled cylindrical resonators and connecting waveguides, where bulk corner states were observed.[60] DLP technology is similar to SLA, but uses light projection instead of a laser beam to fully illuminate, rather than trace, each print slice for UV-polymerization. As a result, DLP allows shorter processing times.[87] Using DLP technology, a topological pyrochlore lattice **(Figure 2c)** which demonstrated a third-order hierarchy of

This article is protected by copyright. All rights reserved

Accepted Article

topological states was fabricated.[61] The localization of states in three separate dimensions was reported, which had implications for the stability and control of these modes.

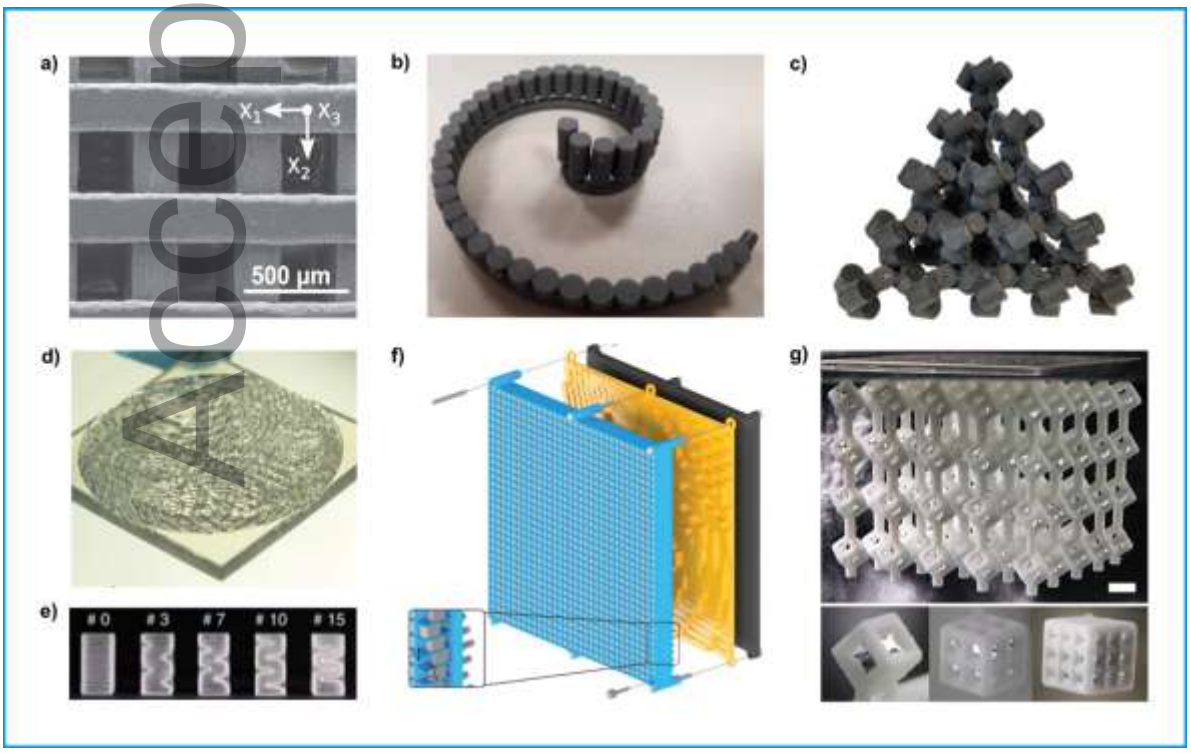
Although both SLA and DLP are mostly limited to thermoset photopolymer materials such as epoxy resins or acrylic, adding nanoparticles to the resin can modify properties such as mechanical strength or electrical conductivity. It has also been proposed that SLA could substitute complicated microfabrication and molding techniques such as soft lithography, demonstrated through printable polydimethylsiloxane (PDMS, Sylgard-184) microdevices.[88] While out-of-plane thicknesses are limited for soft lithography processes, 3D printing can fabricate soft and deformable 3D architectures, potentially allowing the incorporation of microfluidic channels or pneumatic systems. Here, inspiration can be drawn from the field of soft robotics, where DLP-printed elastomers have been effectively used for pneumatically controlled grippers and actuators.[89] An integrated and fully automated setup is also possible, showcased in an entirely microfluidic-based soft robot, fabricated via a combination of embedded printing and soft lithography.[90] Using similar design principles, DIW could instead be used in tandem with SLA, and adapted for the integrated actuation of AMM and PC devices.

It is noteworthy that a major drawback of vat-polymerization techniques is the inability for continuous multi-material printing. Moreover, standard clear resins for SLA require post-processing for optical transparency, and incomplete or improper curing of the resin may alter the print geometry (e.g. bending), affecting the performance of the AMM or PC. Polyjet[63,91] printing presents itself as an attractive alternative, with its ability to directly print multi-material composite structures using photopolymers, exhibiting an extremely smooth surface finish. Surface roughness for a glossy finish can go as low as 0.5 μm (RMS). Unfortunately, Polyjet printing also suffers from certain limitations. The photosensitivity of the materials may cause them to become brittle and degrade in UV-light over time. Water-soluble support material used for overhangs (e.g. SUP706) may also be challenging to remove in narrow cavities or channels, even with water-jetting.

This article is protected by copyright. All rights reserved

Accepted Article

Nevertheless, Polyjet or Multijet printing can easily produce highly transparent prints **(Figure 2d)**[5,6] using standard materials such as Stratasys' acrylate-based VeroClear™. Optical transparency is beneficial for structures with intricate internal geometries, as it provides additional visibility and imaging capabilities. This is especially useful for reconfigurable designs which rely on variable physical states, permutations, or combinations, of the involved elements and unit cells. For example, **Figure 2e** shows five out of sixteen space-coiling meta-bricks which can be strategically assembled into a metasurface for any desired diffraction-limited spatial sound profile.[5] Every brick is different, and their optical transparency allows them to be visibly distinguished. Rather than assembling unit cells, another design consisted of a fixed array of transparent Helmholtz resonators that derived their tunability from precisely controlling the water level within each cavity.[92] Changing the volume of the cavities allowed phase and amplitude modulation. For future research, optical transparency could play a central role in the development of optoacoustic AMMs, where the manipulation of both visible light and sound could be optimized in a combined fashion.



Accepted Article

This article is protected by copyright. All rights reserved

Figure 2. 3D-printed AMMs and PCs. a) SEM image of Direct Ink Writing (DIW) printed PC. Reproduced with permission.[83] Copyright 2014, AIP Publishing LLC. **b)** 40 SLA-printed rainbow-trapping Helmholtz resonators in a spiral, mimicking a human cochlear. Reproduced under terms of the CC-BY license.[74] Copyright 2019, The Authors, published by MDPI. **c)** Assembled topological pyrochlore lattice printed by Digital Light Processing (DLP). Reproduced under terms of the CC-BY license.[61] Copyright 2020, The Authors, published by AAAS. **d)** Polyjet printed transmission hologram using VeroClear™ material. Length of sides, 5 cm. Reproduced with permission.[6] Copyright 2016, Springer Nature Limited. **e)** Space-coiling meta-bricks. Numbers denote different coiled channel lengths corresponding to different phase shifts in units of $\pi/8$. Reproduced under terms of the CC-BY license.[5] Copyright 2017, The Authors, published by Springer Nature Limited. **f)** Reflective spatial sound modulator (SSM) metasurface. The 1024 SSM reflectors were printed by Stereolithography (SLA), using the standard grey resin. Diameter of cylindrical base, 1.9 mm. The frame (blue) and imprinting plate (yellow) were printed in polylactide (PLA) by Fused Deposition Modelling (FDM). High-end FDM allowed printed walls of 0.6 mm for the frame. Reproduced under terms of the CC-BY license.[31] Copyright 2020, The Authors, published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **g)** Selective Laser Sintering (SLS) printed hydrophobic frame for bubble formation. Scale bar, 5 mm. Below, additional pillars at cube face to produce bubble sizes at 10 mm, 7.5 mm, and 5 mm, respectively [from left to right]. Reproduced with permission.[93] Copyright 2019, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Low-frequency structures are mostly designed for the audible regime from 20 Hz to 20 kHz, and elements are in the range of millimeters to centimeters. Very high-resolution printers are not necessary for *macrostructures*, which makes FDM an obvious choice to print low-frequency AMMs and PCs.[73,75] Newer models can even produce layer thicknesses down to ~ 0.1 mm. FDM printers are generally affordable and simple to operate, making them popular for both industry and

This article is protected by copyright. All rights reserved

Accepted Article

consumer use. They mainly produce hard prints through the use of thermoplastic filaments, which are useful for physical boundaries with high acoustic impedances to facilitate the reflection or transmission of airborne acoustic waves. The three most common and accessible materials used for AMMs and PCs are acrylonitrile butadiene styrene (ABS),[94,95] polylactide (PLA),[31,96,97] and polycarbonate.[98,99]

Due to layer-by-layer extrusion, FDM prints are often characterized by a rough layered texture. Post-processing can significantly improve surface finish, using methods such as chemical vapor treatment, ultrasonic surface rolling, or polishing.[100,101] For example, an ABS sample with a 0.1 mm layer height from the printed wall was polished down using acetone vapor to an improved surface roughness of 0.8 μm.[102]

Macrostructures can also be fabricated via inexpensive powder-bed fusion techniques such as Selective Laser Sintering (SLS) and Selective Laser Melting (SLM).[103–106] Recently, broadband acoustic transmission at the water-to-air interface was demonstrated by SLS-printed Nylon (FS3300PA) structures **(Figure 2g)**.[93] The metasurface allowed almost perfect underwater transmission at resonance, by creating architected bubble formations using the intrinsically hydrophobic printed frame. Despite the metasurface exhibiting numerous overhangs, powder-bed fusion prints can remain highly stable during printing without needing support material, and waste powder can also be reused. This example reflected a smart coupling of theoretical design, and sensible fabrication and material choices. Besides Nylon, other powdered raw materials including metals such as aluminum alloys, stainless steel and titanium can also be obtained.[107,108] Printing entire geometries with metal eliminates unnecessary secondary processes such as welding, especially for complex features.

A reconfigurable ultrasonic (40 kHz) reflective AMM was recently reported, consisting of both *micro* and *macrostructures* **(Figure 2f)**.[31] The planar metasurface could enable acoustic focusing as well as multiplane holography. It consisted of small reflector elements (cylindrical base, 1.9 mm)

This article is protected by copyright. All rights reserved

Accepted Article

that were fabricated by SLA and Polyjet printing (for 16×16 array), while the larger imprinting plates were FDM-printed using PLA. These imprinting plates are encoded with the desired collective phase response, subsequently positioning the reflectors at different heights to induce specific phase delays. Here, we see how different printing technologies can complement one another to optimally fabricate different geometric elements.

3.2. Machining

Subtractive manufacturing has been around for centuries, and many involved processes are still relevant today, even for metamaterials and PCs. Machining is a negative i.e. subtractive process, whereby the resulting object is sculpted from the initial work piece, such as a piece of acrylic or slab of metal. In comparison to 3D printing, there is greater control over the properties of the raw material being machined, such as mechanical strength and hardness. Machining can also deal with high-volume or heavy-duty parts at a lower cost. Like conventional 3D printing, the covered machining processes are mainly useful for operating frequencies within the audible (20 Hz – 20 kHz) and ultrasonic (> 20 kHz) regimes.

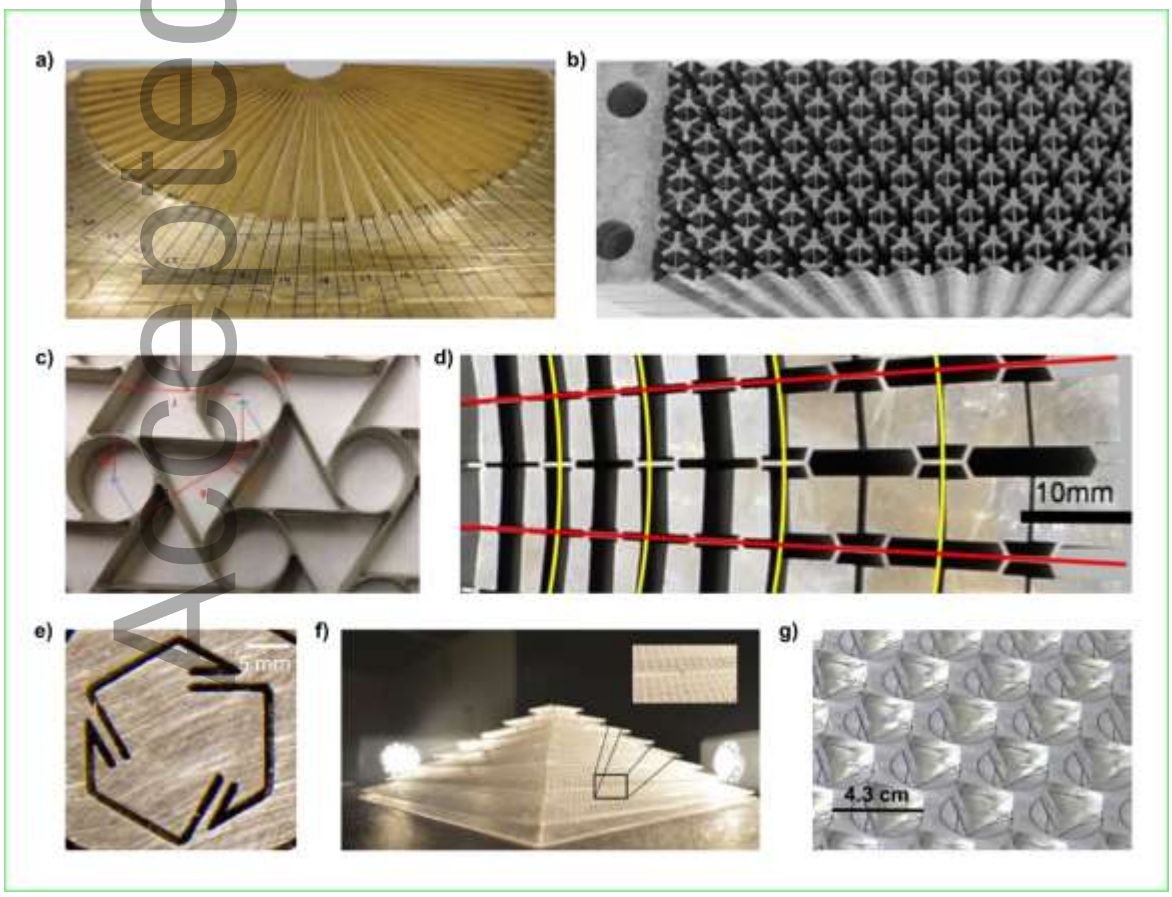
Milling and drilling are popular multipoint mechanical machining processes that can either be automatic or manual, typically using cylindrical rotating cutting-heads. Minimum feature size is dependent on the size of the tool-head. These technologies have been used to fabricate AMMs exhibiting hyperbolic dispersion, often characterized by arrays of perforations or grooves, such as hyper-lenses for imaging applications **(Figure 3a)**.[109–111] Cavities can also be cut into different materials, seen in an AMM with a drilled array of ultrasonic aluminum Helmholtz resonators,[36] and a valley topological PC with milled acrylic channel-cavity unit cells.[112] The 1 mm × 1 mm cylindrical necks of the Helmholtz resonators demonstrated how features can be easily machined to high accuracy and precision. Although CNC mills are good at making perforations and surface contours, they generally have trouble with internal corners. Hence unlike 3D printing, 3D structures composed of machined parts often require a secondary assembly step. However, milling can

This article is protected by copyright. All rights reserved

Accepted Article

produce structures with an exceptionally high-quality surface finish. A milled surface can exhibit a surface roughness as low as 0.2 µm.

Water Jet Machining (WJM) is a mechanical machining process that cuts samples using pressurized water between 200 MPa – 400 MPa. Both 2D and 3D structures can be machined, from materials like glass, wood, or metal, with thicknesses up to ~ 230 mm. However, before using WJM, it is important to ensure that the AMM or PC material is corrosion-resistant. WJM can precisely cut soft materials into repeating patterns, as demonstrated by an elastomeric chiral lattice[113], similar in design to an aluminum lattice beam for broadband vibration attenuation **(Figure 3c)**.[114] **Figure 3b** shows a foam-like metallic PC[115] built for acoustic focusing, consisting of fifteen 5 mm thick plates cut by WJM. Each slab was machined with intricate honeycomb lattices, and later stacked and glued. This simple method of assembly would be further covered in a later section (section 5).



Accepted Article

This article is protected by copyright. All rights reserved

Figure 3. Machined structures. AMM and PC structures fabricated via mechanical and thermal machining processes. **a)** Acoustic magnifying hyper lens with 36 brass fins fabricated by CNC milling. Fins have a radius of 2.7 cm to 21.8 cm. Reproduced with permission.[111] Copyright 2009, Springer Nature Limited. **b)** Foam-like metallic PC lattice fabricated by Water Jet Machining (WJM). The width of interconnections are 0.5 mm and the sides of a regular hexagon are 6.445 mm. Reproduced with permission.[115] Copyright 2013, AIP Publishing LLC. **c)** Aluminum chiral metamaterial beam fabricated by WJM. Each circular inclusion has a diameter of 15 mm, and wall thicknesses are at 0.5 mm. The vertical thickness is 25.4 mm. Reproduced with permission.[114] Copyright 2014, Elsevier B.V.. **d)** Aluminum broadband acoustic cloak fabricated by Electrical Discharge Machining (EDM). Reproduced with permission.[116] Copyright 2017, American Physical Society. **e)** Stainless steel chiral microstructure. Scale bar, 2 mm. Reproduced with permission.[78] Copyright 2014, Nature Springer Limited. **f)** An omnidirectional acoustic cloak. The sides of the square base have an approximate length of 34.3 cm, while the vertical height can extend from 5.7 cm to 11.4 cm. Reproduced with permission.[4] Copyright 2014, Nature Springer Limited. **g)** Acrylic prisms cut to build a sonic crystal. Reproduced with permission.[54] Copyright 2017, Springer Nature Limited. (e, f, g) are structures fabricated by the direct laser-cutting technique.

Electrical Discharge Machining (EDM) is an unconventional thermal machining process that removes hard metals through electric sparks within a dielectric medium. Unlike CNC mills, EDM can better machine sharp internal corners and high-aspect ratio features with a minimum feature size of ~ 0.25 mm, highly beneficial for AMMs with deep, narrow inclusions or thin walls. EDM mostly machines conductive materials. High-precision EDM has been utilized to fabricate metafluids,[67] which can exploit negative refraction using prisms,[66] and support underwater cloaking . The acoustic cloak[116] in **Figure 3d** was made from a block of aluminum, consisting of five layers of unit cells with the smallest rib thickness of 0.35 mm. The cloak had a height of

This article is protected by copyright. All rights reserved

Accepted Article

50 mm and exhibited broadband performance over the frequencies of 9 kHz – 15 kHz. Other examples include planar AMMs and PCs for lensing and wavefront manipulation, such as a holey GRIN PC and a hexagonal aluminum matrix for generating underwater plane waves from cylindrical waves within the range of 15 kHz – 23 kHz.[65,117]

Another thermal machining process is direct laser-cutting, popular for fabricating origami and kirigami-based mechanical metamaterials.[118,119] Similar techniques are employed to fabricate foldable and tunable AMM designs,[45] where quality and cut-rate relies on the power, material thickness, and type of laser. Unlike WJM, structures are typically 2D, and material thicknesses cannot exceed ~ 25 mm. In particular, conventional CO_2 lasers mostly cut non-metal material thicknesses < 12 mm.

Direct laser-cutting is a highly flexible technique that can cut out both patterns and separate components. Sophisticated shapes were outlined on sheets of metal **(Figure 3e)**[78] and polymer membranes (10 μm – 300 μm thick)[120] for lensing and imaging applications. Other than intricate shapes, numerous laser-induced circular holes that were perforated into acrylic sheets have been used to construct an omnidirectional acoustic cloak for airborne sound **(Figure 3f)**.[4] Another interesting application is the sculpting of different side-surfaces of a Weyl PC lattice with a woodpile configuration.[121] The geometric surface terminations were laser-cut, enabling surface acoustic waves (SAW) to be either positively or negatively refracted at different interfaces. Moreover, besides directly patterning AMMs and PCs, a high volume of identical parts can be machined. A topological insulator (TI) was assembled from hundreds of laser-cut acrylic prisms, which behaved as rotating anisotropic scatterers **(Figure 3g)**.[54] Not only were acoustic valley hall edge modes observed, but the tunable TI possessed internal degrees of freedom as the scatterers modulated the frequency bandgaps.

In the cases where both machining and printing can fabricate a certain structure, we can evaluate the trade-offs to choose between them, depending on the application and available resources. For

This article is protected by copyright. All rights reserved

Accepted Article

example, printing a complex metallic structure via SLM would result in diminished overall surface quality. Conversely, machining such a complex structure could necessitate a further assembly step which requires more time and monetary resources.

3.3. Hybrid Manufacturing

Hybrid manufacturing leverages on the strengths of different single-step processes via integrated fabrication approaches. Additive-subtractive processes can reduce time and cost, improve tolerances, and increase surface quality, particularly for metal parts. Recently a novel 6-axis hybrid set-up was reported,[122] using a robotic arm compatible with changeable printing and machining heads and a heated print bed. The integrated set-up allowed multi-plane processing with minimal material wastage. Moreover, multi-axis printing and machining with more degrees of freedom could reduce the need for support structures, lessening the burden of post-processing. Subtractive-subtractive hybrid micromachining systems have combined Micro-EDM (μEDM) with microelectrochemical machining (μECM) for shaping and finishing microscale cavities.[123] Such micromachining processes could both rival or complement complicated microfabrication techniques. Hybrid manufacturing can maximize the quality and production of AMMs and PCs, also encouraging the future integration of microelectromechanical systems (MEMS) for miniaturization and dynamic high-frequency operation in the megahertz range and beyond.

4. Multi-Step Fabrication

Single-step techniques are insufficient to fabricate all kinds of AMMs and PCs. In the instances where the structures are highly heterogeneous or consist of various minute features on the micro or nanoscale, we require multiple processes, besides printing and machining.

4.1. Microfabrication

Traditionally used in the semiconductor industry, microfabrication can fabricate micro or nanoscale structures not easily accessible via other technologies. Here, the maximum size of in-plane structures is restricted by the substrate's planar-dimensions, while out-of-plane structures are

This article is protected by copyright. All rights reserved

Accepted Article

typically limited to thicknesses < 500 μm. Microfabrication involves multiple processes, like deposition and lithography. This technique has fabricated high-frequency (ultrasonic and hypersonic regime) AMMs, and mainly PCs,[124] for specific applications such as SAW manipulation,[125] often used in microfluidic applications like sensing. SAW-based devices typically use acoustic frequencies ranging from 5 MHz to 159 MHz, sometimes going higher into the gigahertz range.

In microfabrication, nanometer layers can be easily deposited onto substrates, which is not possible using conventional high-resolution 3D printers such as SLA and Polyjet printing. Using thin-film deposition, compact, heterogeneous metasurfaces could be engineered with high surface quality and low tolerances. Depending on the substrate quality, an RMS roughness in the nanometer range is assuredly achievable. Surface modifications of material layers can be used to exhibit advanced properties such as enhanced adhesion, hydrophobicity, and conductivity.

We also see dry plasma etching processes like reactive ion etching (RIE) and deep reactive ion etching (DRIE) being used for AMM and PC fabrication.[37,64] These processes can be isotropic or anisotropic, producing very little undercut, helpful for making high aspect ratio features. For example, a large array of Helmholtz resonators was etched into a 1 mm thick silicon wafer using DRIE, with each neck width at 50 μm.[37] This 2D inertial AMM was developed for sound attenuation (up to 18 dB), in the ultrasonic regime. A vapor release etch that provides isotropic etching using xenon difluoride (XeF_2) can also be used to release fabricated structures from substrates like polysilicon.[126] In a recent paper, a molecular etching technique was used to etch through 2 mm of steel to fabricate a meta-lens that could produce super-oscillation wave-packets, demonstrating acoustic tweezing and acoustic imaging. The lens, shown in **Figure 4a**, used an operating frequency of 1 MHz.[127]

Accepted Article

This article is protected by copyright. All rights reserved

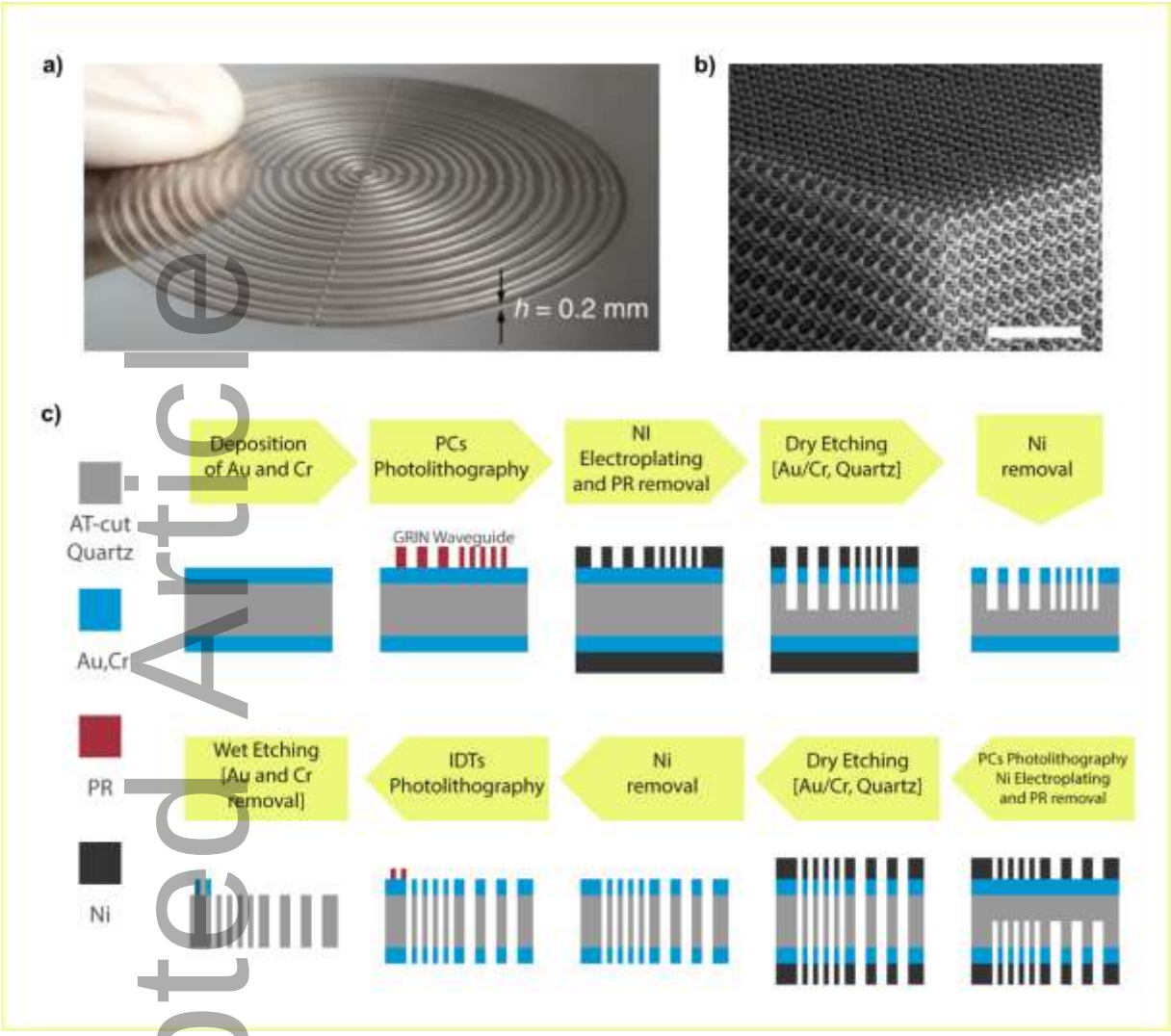


Figure 4. Structures fabricated using microfabrication methods. a) Ultrasonic meta-lens, having a thickness ~ 0.13π of the operating frequency. Reproduced under terms of the CC-BY license.[127] Copyright 2019, The Authors, published by Springer Nature Limited. **b)** SEM image of an AMM (with tapered truss elements) fabricated by two-photon lithography. Scale bar, 200 μm. Reproduced with permission.[128] Copyright 2016, American Physical Society. **c)** Fabrication process flow for gradient refractive index phononic crystal (GRIN PC) lens. 'PR' refers to photoresist, and 'IDT' refers to interdigital transducers. 'Au', 'Cr' and 'Ni' are the chemical symbols for gold, chromium, and nickel, respectively. Adapted with permission.[64] Copyright 2014, Elsevier B.V..

Accepted Article

This article is protected by copyright. All rights reserved

A combination of wet and dry etching techniques were utilized to etch graduated apertures through an entire 80 μm thick quartz substrate, to produce a GRIN PC lens for Lamb wave-focusing and wave-guiding **(Figure 4c)**.[64] GRIN elements are ordered with varying periodicity, which might seem challenging to fabricate on a micron-scale. However, in microfabrication, lithography masks can easily transfer the entire pattern, enabling the parallel fabrication of all the elements at the same time, as opposed to previously discussed single-step serial processing techniques. By incorporating microfabrication processes alongside printing or machining GRIN lenses, the minimum perforation size of the index profile could be reduced, potentially enhancing their performance.

Microfabrication patterning processes come in many forms including conventional photolithography, soft lithography, two-photon lithography, electron-beam lithography (EBL), and focused ion-beam (FIB) structuring. Soft lithography is a type of micro-molding and embossing technique often used for microfluidic device fabrication. The entire process may involve EBL to pattern the mold.[129] Soft lithography is highly compatible with elastomeric materials like Polydimethylsiloxane (PDMS), which was seen in the realization of a bubble PC.[130] The PC consisted of periodic air cavities embedded within soft matrix layers that could be stacked. It was found that hybridization bandgaps relied on the lattice constant and size of cavities, whilst the Bragg gap depended on the distance between layers. Two-photon lithography is a special method that exposes the photoresist in volume elements using a controlled laser to pattern 3D nanoscale structures. **Figure 4b** shows an ultrasonic micro-lattice fabricated using this technique, that operates through the formation of hybridization bandgaps.[128]

With the emergence of surface PCs operating in the hypersonic regime from 1 GHz to 100 GHz, powerful techniques such as EBL[131] are required for precise nanopatterning down to a few nanometers. EBL is entirely maskless as it utilizes a focused beam of electrons to scan and pattern a resist layer on the substrate, allowing greater flexibility in processing complex features and designs. This method was utilized to fabricate an array of pillars (top diameter, 170 nm) in a surface PC for

This article is protected by copyright. All rights reserved

Accepted Article

frequencies within 2 GHz – 7 GHz.[71] To directly structure the substrate, FIB technology can not only pattern, but also micromachine via local etching or milling. Micro or nanoscale structures can also be grown using this technique.[132,133] FIB is promising for hypersonic PC fabrication, demonstrated for a 2D freestanding silicon membrane operating at a frequency of 33 GHz.[134] However, such particle-based processes are currently better suited for low-volume production due to high production costs and slow processing times.

In recent years, surface PCs have also played a major role in advanced nanometrology research for reliable characterization, mass-sensing applications, and nondestructive testing of nano-systems.[135–137] Schemes involving optoacoustic excitation of surface PCs offer unprecedented sensitivity for detecting even picometer-range displacements, using technologies such as extreme ultraviolet (EUV) tabletop laser sources.[72]

Apart from being able to fabricate complex AMM and PC designs, microfabrication can provide the means for active control and tunability. Integrated thin-film circuits can be patterned on flexible[138] or rigid substrates for electronic or electromagnetic actuation capabilities, thereby creating opportunities for dynamically reconfigurable acoustic devices. Although printing technologies like DIW can fabricate circuits, high quality insulators are generally difficult to print. Further miniaturization of high-frequency AMM and PC designs via microfabrication approaches can pave the way for future integration into commercial devices, which has not yet been attempted.

4.2. Molding and Casting

Molding and casting is a viable method to fabricate soft and deformable structures for AMMs and PCs. It is usually a two-step process, where the mold is initially fabricated, prior to the casting and curing of the desired material. The mold may be fabricated using the previously discussed technologies such as 3D printing. Only certain materials are suitable for casting, such as resin, or silicone, which can conform into any shape and cure into a rubber-like solid. This method is advantageous, whereby elastomers or mixtures with altered properties (e.g. magnetic, conductive)

This article is protected by copyright. All rights reserved

Accepted Article

can be molded into functional shapes, and even be actuated for actively tunable AMMs and PCs. For instance, Fe_3O_4 particles were embedded into a silicone mixture using a two-roll mill, to cast a membrane that was actively responsive to magnetic fields.[39] Another magnetically tunable AMM **(Figure 5a)**[49] was casted into the hollow channels of a water-soluble sacrificial lattice mold, to produce a structure that could switch between singly-negative or doubly-negative operation modes. The mold was fabricated using projection microstereolithography, a type of printing technology modified for microfabrication.

3D-printed molds[139] can effectively cast elastomeric structures, whereby mechanical buckling would modulate the stiffness of deformable elements to open and close bandgaps.[140] Such acoustic switching behavior was demonstrated in entirely-soft 3D helical structures **(Figure 5b)**. Here, Polyjet printed molds casted the vinylpolysiloxane elastomers **(Figure 5c)**.[141] Using FDM printed molds, magnetoactive sharkskin-inspired AMMs were fabricated and magnetically configured into acoustic logic gates **(Figure 5d)**.[142] The same limitations in 3D printing technologies apply when printing a mold, directly affecting the casted structure. For example, a 3D-printed mold with a poor surface finish would likely cast an object with a non-uniform surface profile. The degree of adhesion of the cured material with the mold is another aspect to consider. Moreover, conventional 3D-printed molds are unlikely to provide sufficient resolution for AMMs and PCs in the megahertz and gigahertz range. Soft lithography is more suitable in such cases.

Accepted Article

This article is protected by copyright. All rights reserved

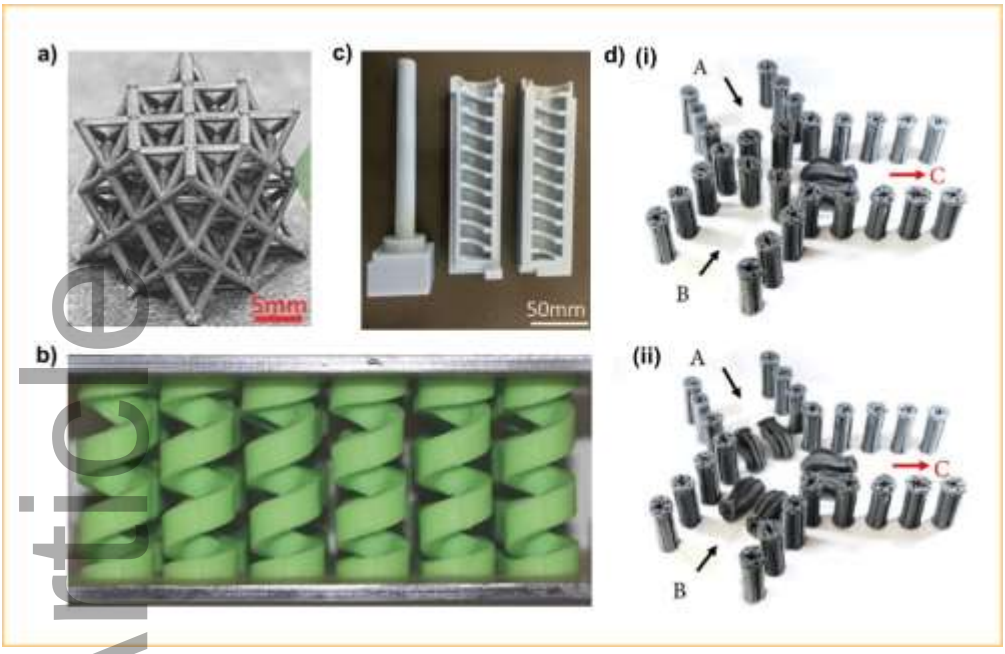


Figure 5. Different structures fabricated via molding and casting techniques. a) Magnetoactive AMM casted from water-soluble mold, fabricated by projection microstereolithography. Reproduced with permission.[49] Copyright 2018, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **b)** Un-deformed vinylpolysiloxane elastomeric helices for controlled acoustic switching. When deformed, height is 40 mm. Reproduced with permission.[141] Copyright 2015, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **c)** Polyjet printed molds used to cast helical 3D structures in (b). **d)** Sharkskin-inspired magnetically reconfigurable acoustic structures being used to construct acoustic logic gates. Each pillar has a diameter of 1.5 cm, and a height of 4.25 cm. 'A' and 'B' denote inputs, while 'C' is the output, in red. Figure shows an **(i)** AND gate and **(ii)** OR gate. Reproduced under terms of the CC-BY license.[142] Copyright 2020, The Authors, published by AAAS.

We also see unique fabrication strategies utilizing simple objects for molding, demonstrated in the experimental realization of a non-periodic AMM exhibiting dipolar resonance and a negative mass density. In this example, circular petri dishes casted suspended tungsten carbide (WC) beads (diameter, 397 μm or 500 μm) within epoxy resin.[143] The casting and curing process was repeated

This article is protected by copyright. All rights reserved

Accepted Article

multiple times, with each layer of resin containing a random dispersion of WC particles (2% – 10% volume fraction). Repetitive casting and curing presents a new opportunity to fabricate layered AMM and PC designs, exploiting the versatility of silicones and resins with each layer characterized by different properties.

4.3. Microfluidics and Wet-Chemical Techniques

Microfluidics deals with the precise control of fluids in volumes of micro/nano liters or less. A possible application of microfluidic techniques is the fabrication of porous microspheres, which can manipulate acoustic wave propagation within a fluid medium. There is a class of soft AMMs, designed to consist of a random dispersion of porous micro-particles, which behave as slow soft Mie-resonators. These suspended resonators form an AMM that exhibits a negative bulk modulus and macroscopic isotropy. A possible fabrication technique used to create the microspheres is high internal phase emulsion (HIPE) templating.[144] This procedure was carried out by a microfluidic device for in-situ polymerization, producing soft porous PDMS microspheres **(Figure 6b)**.[34,145] Besides microspheres, HIPE templating was also recently employed to fabricate a soft, porous GRIN metasurface **(Figure 6a)**.[146] The broadband ultrasonic metasurface could shape 3D wavefronts underwater, allowing acoustic focusing (200 kHz) and the generation of vortices (150 kHz).

Another way to fabricate porous microspheres for acoustic resonance is via the sol-gel method, a type of wet-chemical technique. For instance, silica xerogel beads were synthesized by means of digital microfluidics, where the precursor and catalyst reacted together within the microchannel to form droplets within a continuous stream of fluorinated oil. The beads were washed and dried in ethanol, followed by plasma oxidation for random distribution into a carbopol polymer matrix **(Figure 6c)**.[35] The resulting colloidal metamaterial exhibited frequency bands with a negative acoustic refractive index.

Accepted Article

This article is protected by copyright. All rights reserved

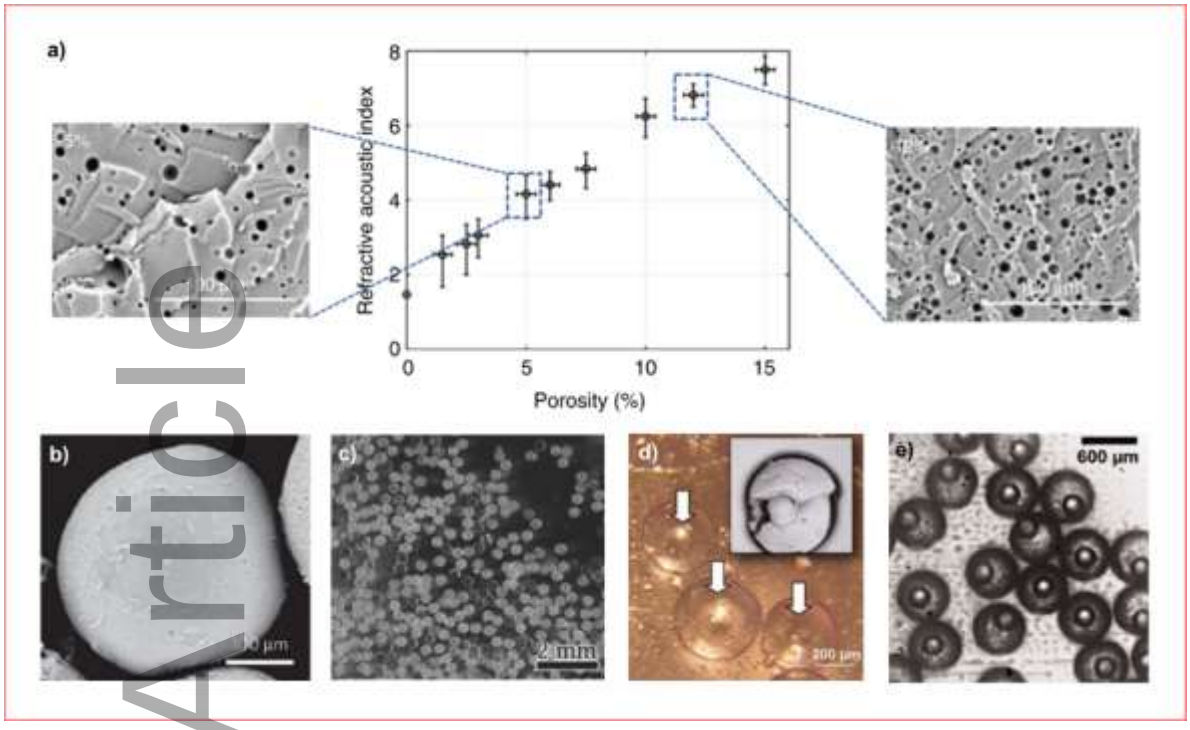


Figure 6. Different structures fabricated using microfluidic techniques. a) Acoustic refractive index measured against porosities of silicone rubber. SEM images show porosities of 5% (left) and 12% (right). The GRIN metasurface was fabricated by high internal phase emulsion (HIPE) templating**.** Reproduced under terms of the CC-BY license.[147] Copyright 2019, The Authors, published by Springer Nature Limited. **b)** Polydimethylsiloxane (PDMS) microspheres fabricated by HIPE templating. Reproduced with permission.[34] Copyright 2015, Springer Nature Limited. **c)** Silica xerogel beads fabricated via the sol-gel method, dispersed within a carbopol matrix. Reproduced with permission.[35] Copyright 2015, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **d)** Core-shell particles fabricated by the first method, through temperature-controlled gelation, and the **e)** second method, through in-situ UV polymerization and a complexation reaction. Reproduced with permission.[148] Copyright 2012, Wiley Periodicals, Inc.

Apart from porous micro-particles, core-shell particles can also act as locally resonant mechanical oscillators. Two decades ago, a locally resonant sonic crystal was fabricated for the first time, with the use of periodically arranged core-shell particles, having a lattice constant of 1.55 cm.[149] It was

This article is protected by copyright. All rights reserved

Accepted Article

observed that bandgap phenomena using silicone-coated lead spheres only existed for soft, rather than rigid coatings. A few years later, coated spheres distributed either randomly or periodically, in a matrix were investigated analytically, where the proposed composite structures exhibited low-frequency (375 Hz - 1335 Hz) bandgaps at resonant frequencies.[150] More recently, based on the same analytical model, core-shell microspheres for a non-periodic AMM with a negative mass density were experimentally demonstrated.[148] This time, the core-shell particles were significantly reduced in size for operation in the ultrasonic regime. Due to the miniaturization of these composite particles, fabrication was not a straightforward task.

Two different methods, both via microfluidic manipulation, were presented.[148] The first method **(Figure 6d)** used temperature-controlled gelation to encapsulate solid polymethylmethacrylate (PMMA) core particles. Agarose droplets were dispensed into the microchannel using silicone oil as a carrier fluid to form the hydrogel coating. The second method **(Figure 6e)** instead used a liquid core made from ethylene glycol dimethacrylate (EGDA) droplets, which solidified through in-situ UV polymerization after encapsulation. The shell-coating gelation was carried out using a complexation reaction by calcium cross-linking. Whilst gelation-based approaches are proficient in microsphere synthesis, it can be a time-consuming process. In the discussed examples, the successful cross-disciplinary integration of technologies was showcased by utilizing fluid-mechanical processes to build acoustic structures.

5. Arrangement and Assembly Methods

Theoretical models can determine the configurations or assemblies in which AMMs and PCs are put together, and there are several approaches to achieve them physically. Bonding through the use of adhesives, such as glues and epoxies, is an accessible method to fabricate heterogeneous (e.g. membrane-type structures) or homogenous[44] structures. For example, Popa and Cummer[151] developed a non-reciprocal and highly nonlinear AMM consisting of a thin piezoelectric membrane (PZM) bonded between Helmholtz cavities, as an actively controlled unidirectional metamaterial

This article is protected by copyright. All rights reserved

Accepted Article

(isolation factor > 10 dB). The Helmholtz resonators were tuned to different frequencies to create asymmetry, while a nonlinear electronic circuit drove the PZM that reversibly converted electrical and acoustic energy. Another design modified a Helmholtz cavity by gluing a polyvinylidene fluoride (PVDF) patch with magnetic blocks onto an aluminum membrane. Acrylic plates were mechanically fastened together using bolts to construct the resonator for sound insulation **(Figure 7a)** and energy harvesting applications.[18]

A major advantage of using adhesives is the ability to bond different types of materials together. Other membrane-type AMMs have bonded material layers such as polycarbonate and porous fiberglass wool **(Figure 7b)**,[40] polyurethane and plastic,[26] or latex rubber on aramid fiber sheets.[152] Besides membranes, this assembly technique was also utilized over a range of other AMM and PC design types. These include examples like reconfigurable origami modules and lattices,[46,47] and a PC plate with arrays of cylindrical piezoelectric resonators.[153] It is noteworthy that the handling and alignment of small structures can be challenging when carried out manually. Even for larger structures, the presence of glue and misalignment can impact acoustic performance. For this very reason, the beam-focusing experimental frequency (81 kHz) for the stacked foam-like metallic PC[115] discussed in an earlier section (section 3.2, **Figure 3b**) deviated from the theoretical value (71.2 kHz). In addition, bonding nanoscale AMM and PC structures operating in the hypersonic regime (> 100 MHz) would require other specialized techniques like anodic or vacuum bonding as an alternative to conventional adhesives.

Particularly for 3D printed structures, adhesives can also help to reduce the burden of post-processing. The 3D Luneburg lens **(Figure 7c)**[63] discussed in an earlier section (section 3.1) was printed as two separate hemispheres, before being bonded together via UV irradiation. A main reason for the bonding step was to minimize the use of support structures. In an isotropic spherical space-coiling device[44], the entire ball was composed of seven glued parts, each with identical labyrinthine channels **(Figure 7e)**. Here, the primary purpose of using adhesives was to allow sound

This article is protected by copyright. All rights reserved

Accepted Article

waves to enter the structure and freely propagate through the channels. However in doing so, the water-soluble support material could be more easily removed in each part before the complete AMM was assembled. This strategy is a simple and effective method to bring together 3D-printed parts to form AMM and PC structures with complex internal geometries.

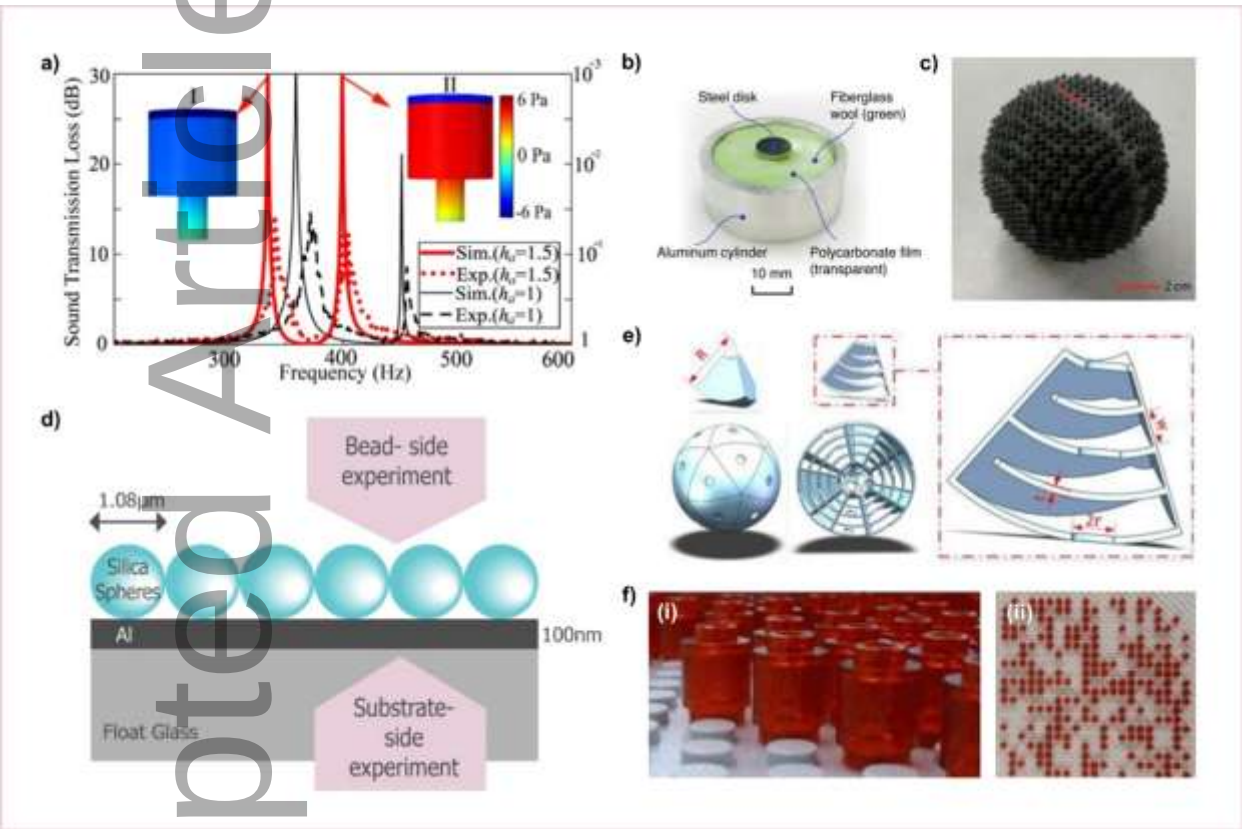


Figure 7. Assembled AMM and PC structures. a) Simulated and experimental results of sound transmission loss in dB (left axis) and the acoustic transmission factor (right axis) showing sound insulation effects. Rigid masses have thicknesses, h_a = 1.5 mm and 1 mm. Reproduced with permission.[18] Copyright 2018, IOP Publishing Ltd. **b)** Layered membrane-type AMM for absorption and broadband insulation. Reproduced with permission.[40] Copyright 2018, AIP Publishing LLC. **c)** Acoustic Luneburg lens. Reproduced under terms of the CC-BY license.[63] Copyright 2018, The Authors, published by Springer Nature Limited. **d)** Monolayer of silica spheres assembled using the wedge-shaped cell convective technique. The aluminum (Al) film coating has a thickness of 100 nm. Reproduced under terms of the CC-BY license.[154] Copyright

This article is protected by copyright. All rights reserved

Accepted Article

2018, The Authors, published by IOP Publishing Ltd on behalf of Deutsche Physikalische Gesellschaft. **e)** Isotropic spherical space-coiling device, R = 35 mm. Magnified labyrinthine part shows channel, where t = 1 mm, r = 3 mm, and w = 5 mm. Reproduced with permission.[44] Copyright 2017, AIP Publishing LLC. **f) (i)** Lego® bricks assembled onto acrylonitrile butadiene styrene (ABS) baseplate via interference fitting. **(ii)** A random arrangement of 210 bricks, which maintained a low-frequency bandgap of 2.2 kHz. Reproduced with permission.[155] Copyright 2015, AIP Publishing LLC.

Rather than being randomly dispersed as discussed in the previous section (section 4.3), microspheres can be ordered or disordered in a certain manner, through several self-assembly strategies. One such strategy is acoustic levitation. This scheme was showcased in a PC created from a colloidal suspension of polystyrene spheres in water.[33] Piezo-ceramic transducers levitated the spheres to assume different lattice configurations, excited at the frequencies of 2.25 MHz, 3.75 MHz, and 5.25 MHz. Besides demonstrating contactless tunability, few reconfigurable PCs had periodicities corresponding to such high frequencies. However, when levitating spherical objects through the use of a standing wave field, the size is often limited to no more than half of the wavelength. Furthermore, levitated objects are ideally symmetrical. It is challenging to levitate larger sizes and different shapes, requiring other techniques such as twin-traps and vortices.[156–158]

If not assembled within a liquid medium, self-assembly can be utilized at the air-water interfaces. Such designs typically operate through contact resonance of the spheres with the substrate. The Langmuir-Blodgett technique is a prevailing method for the fabrication of molecular monolayers, adapted for the monolayer deposition of spheres. It was demonstrated in the assembly of disordered polystyrene microspheres (1.02 μm, diameter) over a soda-lime glass substrate coated with aluminum, with the monolayer subsequently removed by a micro-contact printing technique.[159] Sound attenuation of SAW waves was achieved at ~ 240 MHz, through Hertzian contact resonance

This article is protected by copyright. All rights reserved

Accepted Article

with the spheres. A different AMM employed a similar set-up using a silicon substrate, to observe the hybridization phenomena of the fundamental flexural Lamb wave mode.[160] Another self-assembly method is the wedge-shaped cell convective technique, implemented to build an AMM for gigahertz imaging capabilities, also exploiting contact mechanics **(Figure 7d)**.[154] These examples show how self-assembly techniques can transform random colloidal dispersions into functional AMMs or PCs for acoustic applications.

Interference fitting, or press-fitting, is a method of fastening objects together by means of friction. Lego® is a well-known example of simple and robust interference fitting, which was exploited for fabricating a PC.[155] Cylindrical bricks were assembled and fitted into different topological patterns without the use of adhesives, onto an ABS platform (**Figure 7f**). Definite bandgaps were observed from 1.8 kHz to 2.6 kHz and 5.5 kHz to 11 kHz. This versatile scheme was again seen in the fabrication of a porous AMM, where rigid plates were inserted into melamine foams for sound absorptive properties.[161] The topological pyrochlore lattice[61] discussed in an earlier section (section 3.1, **Figure 2c**) used a different technique called snap-fitting. Snap-fitting involves the momentary deflection of the protruded feature of one object as it is inserted into and interlocked with the depressed feature of another object. Due to this deflection, this method is not ideal for constant reassembly, risking damage to the components.

Purely through strategic arrangement, soda cans have proven to be surprisingly effective acoustic elements (Helmholtz resonators) that can exhibit graphene-like dispersion or achieve super lensing when placed in a honeycomb configuration.[3,162] Graphene-like Dirac dispersion was achieved with a double soda can lattice resonating at separate frequencies. The second sub-lattice was composed of resonant defects which behaved as subwavelength cavities, mimicking the tight-binding governed Hamiltonians of graphene. It is remarkable that such low-cost and accessible commonplace objects can be collectively combined using smart theoretical design and strategic assembly for meaningful acoustic wave manipulation.

This article is protected by copyright. All rights reserved

Accepted Article

6. Conclusion and Outlook

Our discussions of the enlisted fabrication and assembly techniques took us through different examples and strategies for building AMMs and PCs. It often involves a delicate balance between specific building requirements (e.g. geometry, materials) and the limitations of each method, keeping in view the factors of accessibility, time, cost, and quality. The operating frequency of the device can also significantly widen or narrow down the scope of suitable methods. **Figure 8** illustrates fabrication processes and technologies appropriate for designing AMMs and PCs which generally occupy specific operating frequency ranges within the phononic spectrum.

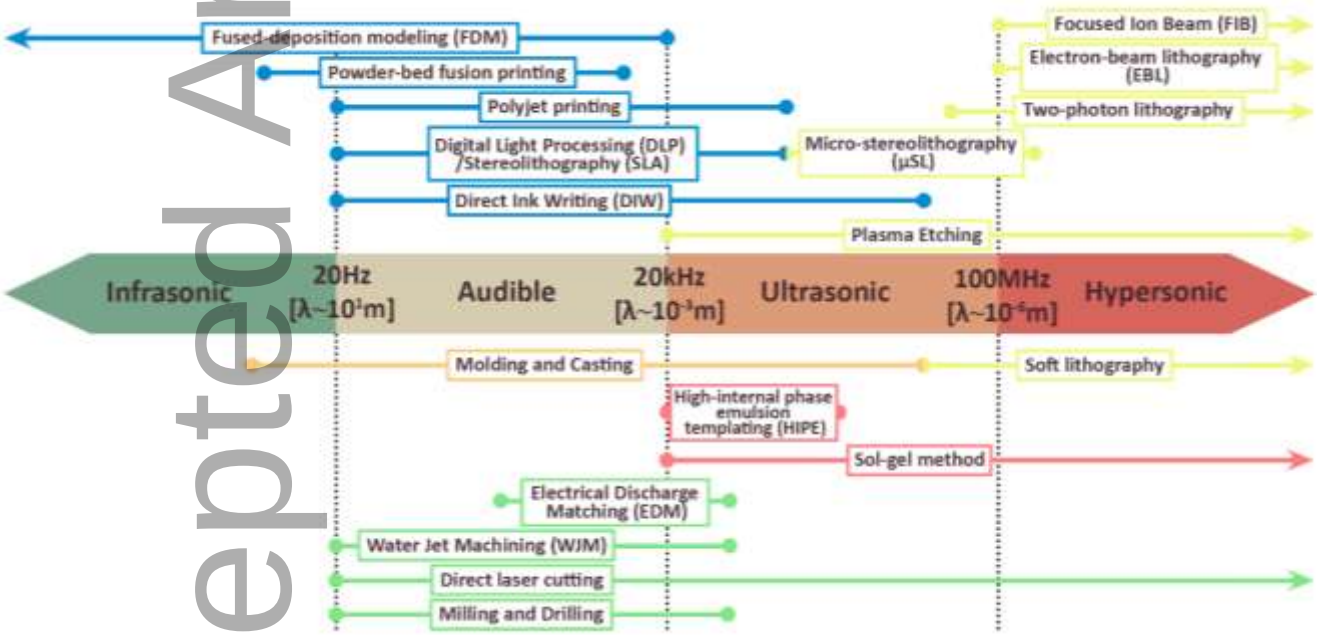


Figure 8. Phononic spectrum showing fabrication methods for various operating frequency ranges for AMMs and PCs. The color scheme for the lines and boxes denote the type of fabrication method. Blue refers to '3D printing', green refers to 'Machining', yellow refers to 'Microfabrication', orange refers to 'Molding and Casting', and coral pink refers to 'Microfluidics and Wet Chemical Techniques'. λ refers to the order of the wavelength of the listed frequency.

Based on the examined techniques, we also synthesize a broad visual map highlighting examples from recent literature dealing with building active and passive, micro and macro AMM and PC

This article is protected by copyright. All rights reserved

Accepted Article

designs **(Figure 9)**. Colored blocks marked with crosses denote possible untapped combinations, uncovering and inspiring future opportunities for research.

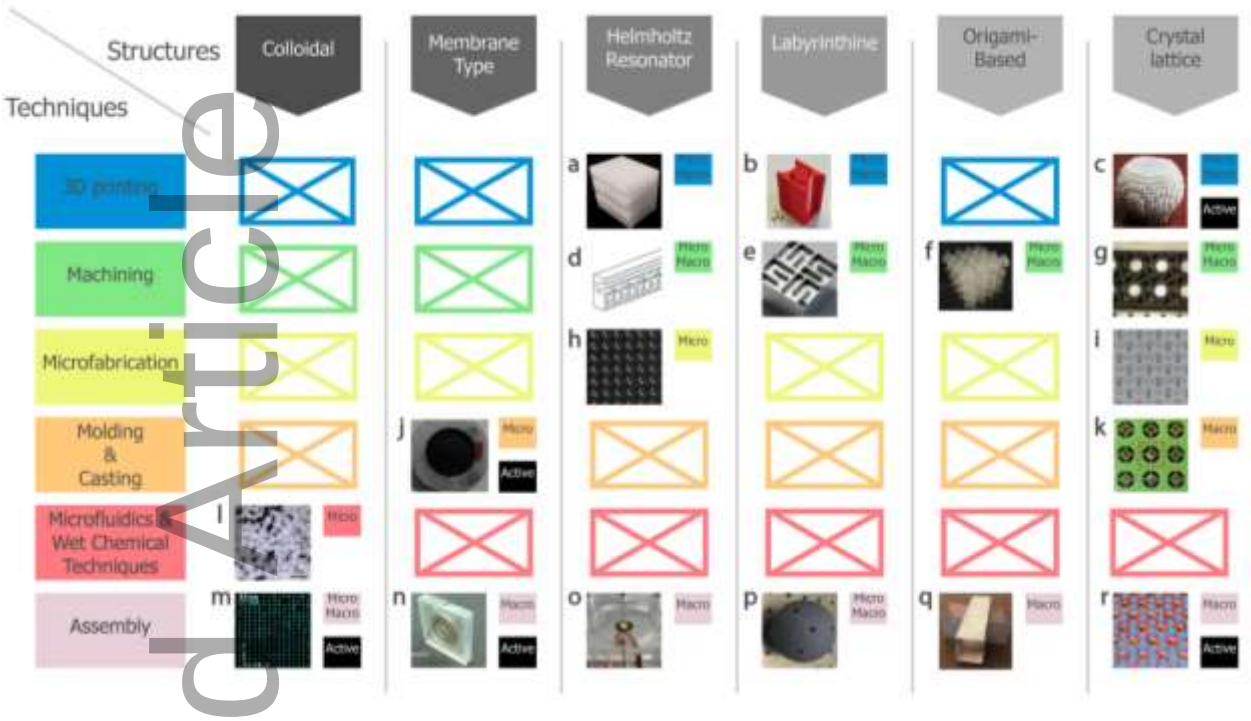


Figure 9. A visual map of techniques and six highlighted AMM and PC structure-types. 3D printing of **a)** Helmholtz resonators,[38,73,74,86] (Image reproduced under terms of the CC-BY license.[86] Copyright 2017, The Authors, published by Springer Nature Limited.) **b)** labyrinthine structures,[1,5,43,44,75,94,96,97,163,164] (Image reproduced with permission.[75] Copyright 2013, AIP Publishing LLC.) and **c)** crystal lattice structures.[41,61,104,105,165,63,77,83,84,93,95,98,99] (Image reproduced with permission.[84] Copyright 2020, AIP Publishing LLC.) **Machining** of **d)** Helmholtz resonators,[36] **e)** labyrinthine structures,[76] (Reproduced with permission.[76] Copyright 2013, AIP Publishing LLC.) **f)** origami-based structures,[45] (Adapted under terms of the CC-BY license.[45] Copyright 2016, The Authors, published by AAAS.) and **g)** crystal lattice structures.[54,58,65,109,110,112–115,117] (Image reproduced with permission.[113] Copyright 2013, Elsevier B.V..) **Microfabrication** of **h)** Helmholtz resonators,[37] (Reproduced with permission.[37] Copyright 2013, AIP Publishing LLC.) and **i)** crystal lattice structures.[42,64,71,125,126] (Image reproduced with permission.[71] Copyright 2016, American Physical Society.) **Molding & Casting**

This article is protected by copyright. All rights reserved

Accepted Article

of **j)** membrane-type structures,[39] (Reproduced with permission.[39] Copyright 2014, AIP Publishing LLC.) and **k)** crystal lattice structures.[49,140,141] (Image reproduced with permission.[140] Copyright 2014, American Physical Society.) **Microfluidics & Wet Chemical Techniques** for **l)** colloidal[34,35,148] structures. (Image reproduced with permission.[34] Copyright 2015, Springer Nature Limited.) **Assembly** of **m)** colloidal,[33] (Reproduced with permission.[33]) **n)** membrane-type structures,[18,26,39,40,151,152], (Image reproduced with permission.[151] Copyright 2014, Springer Nature Limited.) **o)** Helmholtz resonators,[18,151] (Image reproduced with permission.[18] Copyright 2018, IOP Publishing Ltd.) **p)** labyrinthine structures,[44,76] **q)** origami-based structures,[45–47] (Image reproduced with permission.[47] Copyright 2019, American Physical Society.) and **r)** crystal lattice structures.[3,33,153,159,160,162] (Image reproduced under terms of the CC-BY license.[162] Copyright 2017, The Authors, published by Springer Nature Limited.) To the right-hand side of the images, small colorful text boxes denote 'microstructure' or 'macrostructure' classification while small black boxes indicate active structures. Larger colorful boxes with crosses show possible unexplored combinations for future research.

There are still many unexplored paradigms that can offer increasingly creative solutions to build AMMs and PCs, as researchers continue to push for unusual, ground-breaking acoustic devices. For example, based on the discussed single-step fabrication techniques we have observed that despite being able to print an impressive catalogue of materials, Direct Ink Writing (DIW) has been under-utilized in fabricating acoustic devices. DIW has fabricated ceramic-based PCs and graphene-aerogels, providing a clear indication of its ability to print porous structures. This technique could help develop printable colloidal particles with high porosity. Moreover, DIW can print liquid crystal elastomers (LCE), useful for responsive and programmable actuation. Periodic lattices, membrane-type or origami-based LCE structures could morph or deform on-demand, actively modulating the transmission of acoustic waves. Alternatively, Stereolithography (SLA) or Digital

This article is protected by copyright. All rights reserved

Accepted Article

Light Processing (DLP) printing can help to integrate channels for actuating foldable structures using fluidic or pneumatic control. Subsequently, automated microfluidic logic[166] could be applied for active computational designs.

Besides being able to cut out elaborate shapes and patterns as a thermal machining process, a conventional CO_2 laser can also be used to create graphene foam. Laser induced graphene (LIG) has been showcased recently to not only be able to act as a sound source, but also as an acoustic sensor, with thicknesses in the range of microns.[167] By using LIG coatings, AMMs and PCs could exploit the same capabilities in combination with directing or focusing sound. In addition to graphene-based coatings, the integration of thin-film electronics can support smarter, compact, and portable metasurfaces. Another interesting application is the synthesis of metallic nanoparticles using Micro-Electrical Discharge Machining (μ-EDM). If these particles are ferromagnetic, they could be collectively suspended within a fluid medium as tunable assemblies in response to external magnetic stimuli.

Under multi-step fabrication, microfabrication techniques can be further developed for higher frequency AMMs in the megahertz to terahertz range which are limited in current literature. Hybrid micromachining and microfabrication processes can help to achieve this with the fabrication of microelectromechanical (MEMS) and nanoelectromechanical (NEMS) based designs. Two-photon lithography has enabled the direct fabrication of 3D nanostructures. This method could be adapted to create colloidal particles of both spherical and non-spherical shapes, which could later be self-assembled in different configurations. Additionally, two-photon lithography, together with other processes such as sputtering, can fabricate novel electrical, thermal, or chemo-mechanical[168] actively tunable micro-lattices to control bandgap formation. By drawing inspiration from mechanical metamaterial[169] fabrication, deformable polymer micro-lattices are also possible through 3D laser micro-printing.[170] Moreover, lithography-based folding approaches, or 4D

Accepted Article

This article is protected by copyright. All rights reserved

printing with shape memory polymers may be adapted to create millimeter to micrometer scale origami or labyrinthine structures[171] for acoustics.

Molding and casting with sacrificial templates is a powerful method that could cast complex architectures which are difficult to print directly. Microfluidic-based elastomeric or hydrogel[172] networks can be fabricated for building soft, controllable AMMs and PCs. Hydrogel-based structures with embedded nanoparticles[173] can also be further developed for impedance matching and enhanced acoustic performance for applications like underwater cloaking. Besides using high internal phase emulsion (HIPE) templating for fabricating porous structures, salt can be printed and sintered as a template before leaching, to cast micro-scaffolds with ordered porosity.[174] Endless possibilities await the burgeoning fields of AMMs and PCs. As they continue to mature and transition from theory to reality, alongside advancements in fabrication, they are set to transform the way we live, experience, and interact with sound.

Acknowledgements

Christabel Choi and Shubhi Bansal contributed equally to this work. The authors would like to thank Mr. Tom Williams and Mr. Eimontas Jankauskis for their support in creating the figures and illustrations. This work was supported by the EU's H2020 program through their ERC Advanced Grant (number 787413), and the Royal Academy of Engineering through their Chairs in Emerging Technology Program. The authors declare no competing interests.

References

[1] B. Yuan, Y. Cheng, X. Liu, *Appl. Phys. Express* **2015**, *8*, 027301.

[2] Y. X. Shen, Y. G. Peng, F. Cai, K. Huang, D. G. Zhao, C. W. Qiu, H. Zheng, X. F. Zhu, *Nat. Commun.* **2019**, *10*, 1.

[3] N. Kaina, F. Lemoult, M. Fink, G. Lerosey, *Nature* **2015**, *525*, 77.

[4] L. Zigoneanu, B.-I. Popa, S. A. Cummer, *Nat. Mater.* **2014**, *13*, 352.

[5] G. Memoli, M. Caleap, M. Asakawa, D. R. Sahoo, B. W. Drinkwater, S. Subramanian, *Nat.*

This article is protected by copyright. All rights reserved

Accepted Article

Commun. **2017**, *8*, 14608.

[6] K. Melde, A. G. Mark, T. Qiu, P. Fischer, *Nature* **2016**, *537*, 518.

[7] Y. Zhu, J. Hu, X. Fan, J. Yang, B. Liang, X. Zhu, J. Cheng, *Nat. Commun.* **2018**, *9*, 1632.

[8] J. Prat- Camps, G. Christopoulos, J. Hardwick, S. Subramanian, *Adv. Mater. Technol.* **2020**, 2000041.

[9] D. Zhao, Y.-T. Wang, K.-H. Fung, Z.-Q. Zhang, C. T. Chan, *Phys. Rev. B* **2020**, *101*, 054107.

[10] D. Lee, M. Kim, J. Rho, *J. Phys. Condens. Matter* **2019**, *31*, 375901.

[11] V. E. Gusev, O. B. Wright, *New J. Phys.* **2014**, *16*, 123053.

[12] A. Souslov, B. C. Van Zuiden, D. Bartolo, V. Vitelli, *Nat. Phys.* **2017**, *13*, 1091.

[13] Y. Zhao, H. Y. Dong, S. Zhao, S. Min, J. Cheng, B. Li, F. Chi, S. Liu, *J. Appl. Phys.* **2019**, *126*, 065103.

[14] R. Ghaffarivardavagh, J. Nikolajczyk, S. Anderson, X. Zhang, *Phys. Rev. B* **2019**, *99*, 024302.

[15] J. Han, S. Tang, R. Wang, W. Wang, *AIP Adv.* **2019**, *9*, 065201.

[16] C. Shen, Y. Xie, N. Sui, W. Wang, S. A. Cummer, Y. Jing, *Phys. Rev. Lett.* **2015**, *115*, 254301.

[17] S. Qi, M. Oudich, Y. Li, B. Assouar, *Appl. Phys. Lett.* **2016**, *108*, 263501.

[18] X. Zhang, H. Zhang, Z. Chen, G. Wang, *Smart Mater. Struct.* **2018**, *27*, 105018.

[19] O. Georgiou, H. Limerick, L. Corenthy, M. Maksymenko, S. Frish, J. R. Kim, M. Perry, Jö. Müller, M. Bachynskyi, in *Conf. Hum. Factors Comput. Syst. - Proc.*, Association For Computing Machinery, New York, NY, USA, **2019**, pp. 1–9.

[20] T. Chen, J. Liu, K. Zhu, T. Waliczky, in *SIGGRAPH Asia 2017 VR Showcase, SA 2017*, Association For Computing Machinery, Inc, New York, New York, USA, **2017**, pp. 1–2.

[21] J. Albouys-Perrois, J. Laviole, C. Briant, A. M. Brock, in *Conf. Hum. Factors Comput. Syst. - Proc.*, Association For Computing Machinery, New York, New York, USA, **2018**, pp. 1–14.

Accepted Article

This article is protected by copyright. All rights reserved

[22] C. T. Vi, D. Ablart, E. Gatti, C. Velasco, M. Obrist, *Int. J. Hum. Comput. Stud.* **2017**, *108*, 1.

[23] G. Memoli, L. Chisari, J. P. Eccles, M. Caleap, B. W. Drinkwater, S. Subramanian, in *Conf. Hum. Factors Comput. Syst. - Proc.*, Association For Computing Machinery, New York, New York, USA, **2019**, pp. 1–14.

[24] R. Hirayama, D. Martinez Plasencia, N. Masuda, S. Subramanian, *Nature* **2019**, *575*, 320.

[25] M. A. Norasikin, D. Martinez Plasencia, S. Polychronopoulos, G. Memoli, Y. Tokuda, S. Subramanian, in *31st Annu. ACM Symp. User Interface Softw. Technol. - UIST '18*, ACM Press, New York, New York, USA, **2018**, pp. 247–259.

[26] G. Ma, X. Fan, P. Sheng, M. Fink, *Proc. Natl. Acad. Sci.* **2018**, *115*, 6638.

[27] Y. Li, S. Wang, Q. Peng, Z. Zhou, Z. Yang, X. He, Y. Li, *Nanoscale* **2019**, *11*, 16384.

[28] M. R. Haberman, M. D. Guild, *Phys. Today* **2016**, *69*, 42.

[29] B. Liang, J. Cheng, C.-W. Qiu, *Nanophotonics* **2018**, *7*, 1191.

[30] B. Assouar, B. Liang, Y. Wu, Y. Li, J.-C. Cheng, Y. Jing, *Nat. Rev. Mater.* **2018**, *3*, 460.

[31] J. Prat- Camps, G. Christopoulos, J. Hardwick, S. Subramanian, *Adv. Mater. Technol.* **2020**, 2000041.

[32] G. Ma, P. Sheng, *Sci. Adv.* **2016**, *2*, e1501595.

[33] M. Caleap, B. W. Drinkwater, *Proc. Natl. Acad. Sci.* **2014**, *111*, 6226.

[34] T. Brunet, A. Merlin, B. Mascaro, K. Zimny, J. Leng, O. Poncelet, C. Aristégui, O. Mondain-Monval, *Nat. Mater.* **2015**, *14*, 384.

[35] S. Raffy, B. Mascaro, T. Brunet, O. Mondain-Monval, J. Leng, *Adv. Mater.* **2016**, *28*, 1760.

[36] N. Fang, D. Xi, J. Xu, M. Ambati, W. Srituravanich, C. Sun, X. Zhang, *Nat. Mater.* **2006**, *5*, 452.

[37] W. N. Yunker, C. B. Stevens, G. T. Flowers, R. N. Dean, *J. Appl. Phys.* **2013**, *113*, 024906.

[38] X. Jiang, Y. Li, B. Liang, J. Cheng, L. Zhang, *Phys. Rev. Lett.* **2016**, *117*, 034301.

[39] X. Chen, X. Xu, S. Ai, H. Chen, Y. Pei, X. Zhou, *Appl. Phys. Lett.* **2014**, *105*, 071913.

Accepted Article

This article is protected by copyright. All rights reserved

[40] X. Wang, X. Luo, H. Zhao, Z. Huang, *Appl. Phys. Lett.* **2018**, *112*, 021901.

[41] T. Grabec, M. Koller, P. Sedlák, A. Kruisová, B. Román-Manso, M. Belmonte, P. Miranzo, H. Seiner, *Wave Motion* **2020**, *92*, 102417.

[42] N. Wang, J. M.-L. Tsai, F.-L. Hsiao, B. W. Soon, D.-L. Kwong, M. Palaniapan, C. Lee, *J. Microelectromechanical Syst.* **2012**, *21*, 801.

[43] S. K. Maurya, A. Pandey, S. Shukla, S. Saxena, *Sci. Rep.* **2016**, *6*, 33683.

[44] X. F. Fu, G. Y. Li, M. H. Lu, G. Lu, X. Huang, *Appl. Phys. Lett.* **2017**, *111*, 251904.

[45] S. Babaee, J. T. B. Overvelde, E. R. Chen, V. Tournat, K. Bertoldi, *Sci. Adv.* **2016**, *2*, e1601019.

[46] H. Fang, X. Yu, L. Cheng, *Smart Mater. Struct.* **2018**, *27*, 095007.

[47] Y. Zhu, F. Fei, S. Fan, L. Cao, K. Donda, B. Assouar, *Phys. Rev. Appl.* **2019**, *12*, 034029.

[48] S. A. Cummer, J. Christensen, A. Alù, *Nat. Rev. Mater.* **2016**, *1*, 16001.

[49] K. Yu, N. X. Fang, G. Huang, Q. Wang, *Adv. Mater.* **2018**, *30*, 1706348.

[50] M. Thota, K. W. Wang, *J. Appl. Phys.* **2017**, *122*, 154901.

[51] G. Ma, M. Yang, S. Xiao, Z. Yang, P. Sheng, *Nat. Mater.* **2014**, *13*, 873.

[52] C. He, X. Ni, H. Ge, X. C. Sun, Y. Bin Chen, M. H. Lu, X. P. Liu, Y. F. Chen, *Nat. Phys.* **2016**, *12*, 1124.

[53] Y.-G. Peng, C.-Z. Qin, D.-G. Zhao, Y.-X. Shen, X.-Y. Xu, M. Bao, H. Jia, X.-F. Zhu, *Nat. Commun.* **2016**, *7*, 13368.

[54] J. Lu, C. Qiu, L. Ye, X. Fan, M. Ke, F. Zhang, Z. Liu, *Nat. Phys.* **2017**, *13*, 369.

[55] Z. Tian, C. Shen, J. Li, E. Reit, H. Bachman, J. E. S. Socolar, S. A. Cummer, T. Jun Huang, *Nat. Commun.* **2020**, *11*, 1.

[56] F. Li, X. Huang, J. Lu, J. Ma, Z. Liu, *Nat. Phys.* **2018**, *14*, 30.

[57] H. He, C. Qiu, L. Ye, X. Cai, X. Fan, M. Ke, F. Zhang, Z. Liu, *Nature* **2018**, *560*, 61.

[58] H. Xue, Y. Yang, F. Gao, Y. Chong, B. Zhang, *Nat. Mater.* **2019**, *18*, 108.

Accepted Article

This article is protected by copyright. All rights reserved

[59] X. Zhang, B.-Y. Xie, H.-F. Wang, X. Xu, Y. Tian, J.-H. Jiang, M.-H. Lu, Y.-F. Chen, *Nat. Commun.* **2019**, *10*, 5331.

[60] H. Xue, Y. Yang, G. Liu, F. Gao, Y. Chong, B. Zhang, *Phys. Rev. Lett.* **2019**, *122*, 244301.

[61] M. Weiner, X. Ni, M. Li, A. Alù, A. B. Khanikaev, *Sci. Adv.* **2020**, *6*, eaay4166.

[62] Z. Zhang, Y. Tian, Y. Cheng, Q. Wei, X. Liu, J. Christensen, *Phys. Rev. Appl.* **2018**, *9*, 034032.

[63] Y. Xie, Y. Fu, Z. Jia, J. Li, C. Shen, Y. Xu, H. Chen, S. A. Cummer, *Sci. Rep.* **2018**, *8*, 16188.

[64] M.-J. Chiou, Y.-C. Lin, T. Ono, M. Esashi, S.-L. Yeh, T.-T. Wu, *Ultrasonics* **2014**, *54*, 1984.

[65] S. A. R. Kuchibhatla, P. Rajagopal, *NDT E Int.* **2019**, *102*, 304.

[66] K. Li, Y.-G. Peng, X.-F. Zhu, J.-T. Zhang, H.-J. Lv, S.-C. Liu, *Appl. Phys. Lett.* **2014**, *104*, 043505.

[67] A. Zhao, Z. Zhao, X. Zhang, X. Cai, L. Wang, T. Wu, H. Chen, *Appl. Phys. Lett.* **2017**, *110*, 011907.

[68] A. N. Norris, *J. Acoust. Soc. Am.* **2009**, *125*, 839.

[69] L. Shao, S. Maity, L. Zheng, L. Wu, A. Shams-Ansari, Y.-I. Sohn, E. Puma, M. N. Gadalla, M. Zhang, C. Wang, E. Hu, K. Lai, M. Lončar, *Phys. Rev. Appl.* **2019**, *12*, 014022.

[70] S. L. Yeh, Y. C. Lin, Y. C. Tsai, T. Ono, T. T. Wu, *Ultrasonics* **2016**, *71*, 106.

[71] D. Yudistira, A. Boes, B. Graczykowski, F. Alzina, L. Y. Yeo, C. M. Sotomayor Torres, A. Mitchell, *Phys. Rev. B* **2016**, *94*, 094304.

[72] D. Nardi, M. Travagliati, M. M. Murnane, H. C. Kapteyn, G. Ferrini, C. Giannetti, F. Banfi, *IEEE Sens. J.* **2015**, *15*, 5142.

[73] Y. Li, X. Jiang, B. Liang, J. Cheng, L. Zhang, *Phys. Rev. Appl.* **2015**, *4*, 024003.

[74] L. Zhao, S. Zhou, *Sensors* **2019**, *19*, 788.

[75] Y. Xie, A. Konneker, B.-I. Popa, S. A. Cummer, *Appl. Phys. Lett.* **2013**, *103*, 201906.

[76] T. Frenzel, J. David Brehm, T. Bückmann, R. Schittny, M. Kadic, M. Wegener, *Appl. Phys.*

Accepted Article

This article is protected by copyright. All rights reserved

Lett. **2013**, *103*, 061907.

[77] E. R. Fotsing, A. Dubourg, A. Ross, J. Mardjono, *Appl. Acoust.* **2019**, *148*, 322.

[78] R. Zhu, X. N. Liu, G. K. Hu, C. T. Sun, G. L. Huang, *Nat. Commun.* **2014**, *5*, 5510.

[79] S. A. M. Tofail, E. P. Koumoulos, A. Bandyopadhyay, S. Bose, L. O'Donoghue, C. Charitidis, *Mater. Today* **2018**, *21*, 22.

[80] J. Zhao, Y. Zhang, X. Zhao, R. Wang, J. Xie, C. Yang, J. Wang, Q. Zhang, L. Li, C. Lu, Y. Yao, *Adv. Funct. Mater.* **2019**, *29*, 1900809.

[81] J. W. Boley, E. L. White, G. T.-C. Chiu, R. K. Kramer, *Adv. Funct. Mater.* **2014**, *24*, 3501.

[82] A. Kotikian, R. L. Truby, J. W. Boley, T. J. White, J. A. Lewis, *Adv. Mater.* **2018**, *30*, 1706164.

[83] A. Kruisová, H. Seiner, P. Sedlák, M. Landa, B. Román-Manso, P. Miranzo, M. Belmonte, *Appl. Phys. Lett.* **2014**, *105*, 211904.

[84] L. Zhao, E. Laredo, O. Ryan, A. Yazdkhasti, H. T. Kim, R. Ganye, T. Horiuchi, M. Yu, *Appl. Phys. Lett.* **2020**, *116*, 071902.

[85] C. M. Park, S. H. Lee, *Appl. Phys. Lett.* **2018**, *112*, 074101.

[86] N. Jiménez, V. Romero-García, V. Pagneux, J.-P. Groby, *Sci. Rep.* **2017**, *7*, 13595.

[87] R. L. Truby, J. A. Lewis, *Nature* **2016**, *540*, 371.

[88] N. Bhattacharjee, C. Parra-Cabrera, Y. T. Kim, A. P. Kuo, A. Folch, *Adv. Mater.* **2018**, *30*, 1800001.

[89] D. K. Patel, A. H. Sakhaei, M. Layani, B. Zhang, Q. Ge, S. Magdassi, *Adv. Mater.* **2017**, *29*, 1606000.

[90] M. Wehner, R. L. Truby, D. J. Fitzgerald, B. Mosadegh, G. M. Whitesides, J. A. Lewis, R. J. Wood, *Nature* **2016**, *536*, 451.

[91] V. Peri, M. Serra-Garcia, R. Ilan, S. D. Huber, *Nat. Phys.* **2019**, *15*, 357.

[92] Z. Tian, C. Shen, J. Li, E. Reit, Y. Gu, H. Fu, S. A. Cummer, T. J. Huang, *Adv. Funct. Mater.*

Accepted Article

This article is protected by copyright. All rights reserved

2019, *29*, 1808489.

[93] Z. Cai, S. Zhao, Z. Huang, Z. Li, M. Su, Z. Zhang, Z. Zhao, X. Hu, Y. Wang, Y. Song, *Adv. Funct. Mater.* **2019**, *29*, 1906984.

[94] R. Ghaffarivardavagh, J. Nikolajczyk, R. Glynn Holt, S. Anderson, X. Zhang, *Nat. Commun.* **2018**, *9*, 1349.

[95] S. Ebrahimi-Nejad, M. Kheybari, *Mater. Res. Express* **2018**, *5*, 105801.

[96] S.-D. Zhao, A.-L. Chen, Y.-S. Wang, C. Zhang, *Phys. Rev. Appl.* **2018**, *10*, 054066.

[97] C. Zhang, X. Hu, *Phys. Rev. Appl.* **2016**, *6*, 064025.

[98] O. R. Bilal, A. Foehr, C. Daraio, *Adv. Mater.* **2017**, *29*, 1700628.

[99] K. H. Matlack, A. Bauhofer, S. Krödel, A. Palermo, C. Daraio, *Proc. Natl. Acad. Sci.* **2016**, *113*, 8386.

[100] Z. Wang, Z. Liu, C. Gao, K. Wong, S. Ye, Z. Xiao, *Surf. Coatings Technol.* **2020**, *381*, 125122.

[101] J. S. Chohan, R. Singh, K. S. Boparai, R. Penna, F. Fraternali, *Compos. Part B Eng.* **2017**, *117*, 138.

[102] Y. He, G. Xue, J. Fu, *Sci. Rep.* **2015**, *4*, 6973.

[103] S. Tol, F. L. Degertekin, A. Erturk, *Addit. Manuf.* **2019**, *29*, 100780.

[104] M. Reynolds, S. Daley, *Smart Mater. Struct.* **2014**, *23*, 045030.

[105] L. D'Alessandro, R. Ardito, F. Braghin, A. Corigliano, *Sci. Rep.* **2019**, *9*, 8039.

[106] A. Bergamini, M. Miniaci, T. Delpero, D. Tallarico, B. Van Damme, G. Hannema, I. Leibacher, A. Zemp, *Nat. Commun.* **2019**, *10*, 1.

[107] J. H. Martin, B. D. Yahata, J. M. Hundley, J. A. Mayer, T. A. Schaedler, T. M. Pollock, *Nature* **2017**, *549*, 365.

[108] J.-Y. Lee, J. An, C. K. Chua, *Appl. Mater. Today* **2017**, *7*, 120.

[109] V. M. García-Chocano, J. Christensen, J. Sánchez-Dehesa, *Phys. Rev. Lett.* **2014**, *112*,

Accepted Article

This article is protected by copyright. All rights reserved

144301.

[110] J. Zhu, J. Christensen, J. Jung, L. Martin-Moreno, X. Yin, L. Fok, X. Zhang, F. J. Garcia-Vidal, *Nat. Phys.* **2011**, *7*, 52.

[111] J. Li, L. Fok, X. Yin, G. Bartal, X. Zhang, *Nat. Mater.* **2009**, *8*, 931.

[112] Z. Tian, C. Shen, J. Li, E. Reit, H. Bachman, J. E. S. Socolar, S. A. Cummer, T. Jun Huang, *Nat. Commun.* **2020**, *11*, 1.

[113] E. Baravelli, M. Ruzzene, *J. Sound Vib.* **2013**, *332*, 6562.

[114] R. Zhu, X. N. Liu, G. K. Hu, C. T. Sun, G. L. Huang, *J. Sound Vib.* **2014**, *333*, 2759.

[115] A.-C. Hladky-Hennion, J. O. Vasseur, G. Haw, C. Croënne, L. Haumesser, A. N. Norris, *Appl. Phys. Lett.* **2013**, *102*, 144103.

[116] Y. Chen, M. Zheng, X. Liu, Y. Bi, Z. Sun, P. Xiang, J. Yang, G. Hu, *Phys. Rev. B* **2017**, *95*, 180104.

[117] Y. Chen, G. Hu, *Phys. Rev. Appl.* **2019**, *12*, 044046.

[118] J. L. Silverberg, A. A. Evans, L. McLeod, R. C. Hayward, T. Hull, C. D. Santangelo, I. Cohen, *Science* **2014**, *345*, 647.

[119] Y. Tang, G. Lin, S. Yang, Y. K. Yi, R. D. Kamien, J. Yin, *Adv. Mater.* **2017**, *29*, 1604262.

[120] H. Tang, Z. Chen, N. Tang, S. Li, Y. Shen, Y. Peng, X. Zhu, J. Zang, *Adv. Funct. Mater.* **2018**, *28*, 1801127.

[121] H. He, C. Qiu, L. Ye, X. Cai, X. Fan, M. Ke, F. Zhang, Z. Liu, *Nature* **2018**, *560*, 61.

[122] L. Li, A. Haghighi, Y. Yang, *J. Manuf. Process.* **2018**, *33*, 150.

[123] Z. Zeng, Y. Wang, Z. Wang, D. Shan, X. He, *Precis. Eng.* **2012**, *36*, 500.

[124] R. H. Olsson III, I. El-Kady, *Meas. Sci. Technol.* **2009**, *20*, 012002.

[125] Y. Bourquin, R. Wilson, Y. Zhang, J. Reboud, J. M. Cooper, *Adv. Mater.* **2011**, *23*, 1458.

[126] M. Ghasemi Baboly, C. M. Reinke, B. A. Griffin, I. El-Kady, Z. C. Leseman, *Appl. Phys. Lett.* **2018**, *112*, 103504.

Accepted Article

This article is protected by copyright. All rights reserved

[127] Y. X. Shen, Y. G. Peng, F. Cai, K. Huang, D. G. Zhao, C. W. Qiu, H. Zheng, X. F. Zhu, *Nat. Commun.* **2019**, *10*, 1.

[128] S. Krödel, C. Daraio, *Phys. Rev. Appl.* **2016**, *6*, 064005.

[129] D. Qin, Y. Xia, G. M. Whitesides, *Nat. Protoc.* **2010**, *5*, 491.

[130] V. Leroy, A. Bretagne, M. Lanoy, A. Tourin, *AIP Adv.* **2016**, *6*, 121604.

[131] M. Sledzinska, B. Graczykowski, J. Maire, E. Chavez- Angel, C. M. Sotomayor- Torres, F. Alzina, *Adv. Funct. Mater.* **2020**, *30*, 1904434.

[132] Y. Chang, W. Lu, J. Guénolé, L. T. Stephenson, A. Szczpaniak, P. Kontis, A. K. Ackerman, F. F. Dear, I. Mouton, X. Zhong, S. Zhang, D. Dye, C. H. Liebscher, D. Ponge, S. Korte-Kerzel, D. Raabe, B. Gault, *Nat. Commun.* **2019**, *10*, 1.

[133] F. Porrati, S. Barth, R. Sachser, O. V. Dobrovolskiy, A. Seybert, A. S. Frangakis, M. Huth, *ACS Nano* **2019**, *13*, 6287.

[134] D. F. Goettler, M. F. Su, C. M. Reinke, S. Alaie, P. E. Hopkins, R. H. Olsson, I. El-Kady, Z. C. Leseman, *AIP Adv.* **2011**, *1*, 042001.

[135] B. Abad, J. L. Knobloch, T. D. Frazer, J. N. Hernández-Charpak, H. Y. Cheng, A. J. Grede, N. C. Giebink, T. E. Mallouk, P. Mahale, N. N. Nova, A. A. Tomaschke, V. L. Ferguson, V. H. Crespi, V. Gopalan, H. C. Kapteyn, J. V. Badding, M. M. Murnane, *Nano Lett.* **2020**, *20*, 3306.

[136] J. N. Hernandez-Charpak, K. M. Hoogeboom-Pot, Q. Li, T. D. Frazer, J. L. Knobloch, M. Tripp, S. W. King, E. H. Anderson, W. Chao, M. M. Murnane, H. C. Kapteyn, D. Nardi, *Nano Lett.* **2017**, *17*, 2178.

[137] J. N. Hernandez-Charpak, J. L. Knobloch, B. Abad Mayor, T. Frazer, H. Kapteyn, M. Murnane, H. Cheng, A. Grede, N. Giebink, T. Mallouk, P. Mahale, W. Chen, Y. Xiong, I. Dabo, V. Crespi, D. Talreja, V. Gopalan, J. Badding, in *Heal. Monit. Struct. Biol. Syst. XII* (Ed.: T. Kundu), SPIE, **2018**, p. 69.

Accepted Article

This article is protected by copyright. All rights reserved

[138] Y. Khan, A. Thielens, S. Muin, J. Ting, C. Baumbauer, A. C. Arias, *Adv. Mater.* **2020**, *32*, 1905279.

[139] R. L. Harne, Y. Song, Q. Dai, *Extrem. Mech. Lett.* **2017**, *12*, 41.

[140] P. Wang, F. Casadei, S. Shan, J. C. Weaver, K. Bertoldi, *Phys. Rev. Lett.* **2014**, *113*, 014301.

[141] S. Babaee, N. Viard, P. Wang, N. X. Fang, K. Bertoldi, *Adv. Mater.* **2016**, *28*, 1631.

[142] K. H. Lee, K. Yu, H. Al Ba'ba'a, A. Xin, Z. Feng, Q. Wang, *Research* **2020**, *2020*, 1.

[143] M. Duranteau, T. Valier-Brasier, J.-M. Conoir, R. Wunenburger, *J. Acoust. Soc. Am.* **2016**, *139*, 3341.

[144] T. Zhang, R. A. Sanguramath, S. Israel, M. S. Silverstein, *Macromolecules* **2019**, *52*, 5445.

[145] K. Zimny, A. Merlin, A. Ba, C. Aristégui, T. Brunet, O. Mondain-Monval, *Langmuir* **2015**, *31*, 3215.

[146] Y. Jin, R. Kumar, O. Poncelet, O. Mondain-Monval, T. Brunet, *Nat. Commun.* **2019**, *10*, 1.

[147] Y. Jin, R. Kumar, O. Poncelet, O. Mondain-Monval, T. Brunet, *Nat. Commun.* **2019**, *10*, 1.

[148] L. Lukyanova, L. Séon, A. Aradian, O. Mondain-Monval, J. Leng, R. Wunenburger, *J. Appl. Polym. Sci.* **2013**, *128*, 3512.

[149] Z. Liu, *Science* **2000**, *289*, 1734.

[150] Z. Liu, C. T. Chan, P. Sheng, *Phys. Rev. B* **2005**, *71*, 014103.

[151] B.-I. Popa, S. A. Cummer, *Nat. Commun.* **2014**, *5*, 3398.

[152] N. Sui, X. Yan, T.-Y. Huang, J. Xu, F.-G. Yuan, Y. Jing, *Appl. Phys. Lett.* **2015**, *106*, 171905.

[153] F. Casadei, T. Delpero, A. Bergamini, P. Ermanni, M. Ruzzene, *J. Appl. Phys.* **2012**, *112*, 064902.

[154] P. H. Otsuka, S. Mezil, O. Matsuda, M. Tomoda, A. A. Maznev, T. Gan, N. Fang, N. Boechler, V. E. Gusev, O. B. Wright, *New J. Phys.* **2018**, *20*, 013026.

[155] P. Celli, S. Gonella, *Appl. Phys. Lett.* **2015**, *107*, 081901.

[156] A. Marzo, M. Caleap, B. W. Drinkwater, *Phys. Rev. Lett.* **2018**, *120*, 044301.

Accepted Article

This article is protected by copyright. All rights reserved

[157] L. Cox, A. Croxford, B. W. Drinkwater, A. Marzo, *Appl. Phys. Lett.* **2018**, *113*, 054101.

[158] S. Inoue, S. Mogami, T. Ichiyama, A. Noda, Y. Makino, H. Shinoda, *J. Acoust. Soc. Am.* **2019**, *145*, 328.

[159] J. K. Eliason, A. Vega-Flick, M. Hiraiwa, A. Khanolkar, T. Gan, N. Boechler, N. Fang, K. A. Nelson, A. A. Maznev, *Appl. Phys. Lett.* **2016**, *108*, 061907.

[160] A. Khanolkar, S. Wallen, M. Abi Ghanem, J. Jenks, N. Vogel, N. Boechler, *Appl. Phys. Lett.* **2015**, *107*, 071903.

[161] G. Ji, Y. Fang, J. Zhou, X. Huang, *J. Appl. Phys.* **2019**, *125*, 215110.

[162] S. Yves, F. Lemoult, M. Fink, G. Lerosey, *Sci. Rep.* **2017**, *7*, 15359.

[163] H. Esfahlani, H. Lissek, J. R. Mosig, *Phys. Rev. B* **2017**, *95*, 024312.

[164] M. Yang, S. Chen, C. Fu, P. Sheng, *Mater. Horizons* **2017**, *4*, 673.

[165] F. Lucklum, M. J. Vellekoop, *Procedia Eng.* **2015**, *120*, 1095.

[166] F. G. Woodhouse, J. Dunkel, *Nat. Commun.* **2017**, *8*, 1.

[167] L. Q. Tao, H. Tian, Y. Liu, Z. Y. Ju, Y. Pang, Y. Q. Chen, D. Y. Wang, X. G. Tian, J. C. Yan, N. Q. Deng, Y. Yang, T. L. Ren, *Nat. Commun.* **2017**, *8*, 1.

[168] X. Xia, A. Afshar, H. Yang, C. M. Portela, D. M. Kochmann, C. V. Di Leo, J. R. Greer, *Nature* **2019**, *573*, 205.

[169] J. Bauer, L. R. Meza, T. A. Schaedler, R. Schwaiger, X. Zheng, L. Valdevit, *Adv. Mater.* **2017**, *29*, 1701850.

[170] T. Frenzel, M. Kadic, M. Wegener, *Science* **2017**, *358*, 1072.

[171] Y. Zhang, F. Zhang, Z. Yan, Q. Ma, X. Li, Y. Huang, J. A. Rogers, *Nat. Rev. Mater.* **2017**, *2*, 17019.

[172] H. N. Chan, Y. Shu, Q. Tian, Y. Chen, Y. Chen, H. Wu, *Mater. Horizons* **2016**, *3*, 309.

[173] Y. Gao, J. Song, S. Li, C. Elowsky, Y. Zhou, S. Ducharme, Y. M. Chen, Q. Zhou, L. Tan, *Nat. Commun.* **2016**, *7*, 12316.

Accepted Article

This article is protected by copyright. All rights reserved

[174] N. Kleger, M. Cihova, K. Masania, A. R. Studart, J. F. Löffler, *Adv. Mater.* **2019**, *31*, 1903783.

Biographies

Accepted Article

Christabel Choi is a Ph.D. student in metamaterials within the Department of Computer Science at University College London, UK. Under the supervision of Prof. Sriram Subramanian, her research focus is on the development and fabrication of actively reconfigurable acoustic metamaterials. At the University of Sussex, Falmer, UK, she received her BEng degree in Electrical and Electronic Engineering in 2018, before joining the Interact Lab from 2019 – 2020.

Shubhi Bansal is a postdoctoral research fellow at University College London, UK. She received her Ph.D. degree along with M.Sc. (Engineering) Degree in Nano Science and Engineering from Indian Institute of Science, India in 2018. She received her B.Tech. degree in Engineering Physics (with Major in Electronics and Minor in Material Science and Technology) from Delhi Technological University, India in 2013. She was a Research Associate at Indian Institute of Science, India from 2018 – 2019. She subsequently worked as a postdoctoral research fellow with the Interact Lab at the University of Sussex, Falmer, UK from 2019 – 2020.

This article is protected by copyright. All rights reserved

Accepted Article

Niko Münzenrieder is an associate professor for physics of matter at the Free University of Bozen-Bolzano, Italy, where he is working on thin-film technology as well as flexible and wearable electronics. He received a Diploma in physics from the Technical University of Munich, Germany, in 2008, and a PhD in electrical engineering from the Swiss Federal Institute of Technology (ETH) Zürich, Switzerland, in 2013, where he also worked as a postdoctoral research fellow in 2014. From 2015 – 2019 he was a Lecturer, later Senior Lecturer, with the Sensor Technology Research Centre at the University of Sussex, Brighton, UK.

Sriram Subramanian is a Royal Academy of Engineering Chair in Emerging Technologies at University College London (UK) where he leads a research group on designing interactive applications using acoustic metamaterials and phased arrays. Before joining UCL he was a Professor of Human-computer Interaction at the University of Sussex (2015 – 2020) and University of Bristol (2007 – 2015).

This article is protected by copyright. All rights reserved

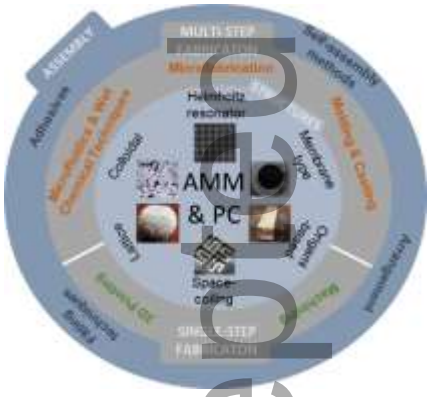
Table of Contents (ToC)

Promoting the physical realization of Acoustic Metamaterials (AMM) and Phononic Crystals (PC) is an essential step towards motivating the practical and industrial advancement of novel acoustic structures. Based on representative AMM and PC structure-types, the techniques to build them are discussed within the strategic categorization of relevant fabrication and assembly methods. This framework guides towards inspiring new acoustic designs.

C. Choi, S. Bansal*, N. Münzenrieder, S. Subramanian*

Fabricating and Assembling Acoustic Metamaterials and Phononic Crystals

ToC figure



Accepted Article

This article is protected by copyright. All rights reserved

Silicon
https://doi.org/10.1007/s12633-020-00734-5

ORIGINAL PAPER

Gate Oxide Variability Analysis of a Novel 3 nm Truncated Fin–FinFET for High Circuitry Performance

Mridul Prakash Kashyap [1] · Rishu Chaujar [1]

Received: 17 June 2020 / Accepted: 23 September 2020
© Springer Nature B.V. 2020

Abstract
In this work, we examined the analog and circuitry amplifying capacity of our novel 3 nm Truncated Fin Junctionless bulk FinFET (n-type) with two different oxide thicknesses at this small scale of gate length. Both oxide widths of high K-material HfO_2 have their own individual benefits, due to high gate controllability as compared to conventional FinFETs having SiO_2 as a gate oxide. The device works best with Tox = 1 nm in terms of power amplification. When tested with this width of gate oxide, we end up with increase of 59.18%, 7.22% and 12.11 times at corresponding peak values of Unilateral power gain (G_u), IP3 and f_{max}. This actually evident the enhanced performance of TF-FinFET for A.C applications at this high range of frequency of the input signal. When device tested at Tox = 1 nm, we end up with the increase of 45%, 21.45%, 16% and decrease of 65% at the corresponding peak values of Intrinsic delay (t_i), Transconductance (g_m), Drain current (I_D) and OFF-state current (I_{OFF}). These results of simulation also showed the compatibility of TF-FinFET in terms of high-performance analog application. After these analyses, we can expect a strong potential for wide variety of applications to high-speed System on chip from this device.

Keywords Junctionless · Truncated fin (TF) · FinFET · GaAs substrate · Unilateral power gain · f_{max}

1 Introduction

Due to the ever-more widespread use of CMOS circuitry in modern wireless communication systems, a thorough knowledge of FinFET transistor has become more mandatory. In order to model and simulate a circuit using these FinFET, the device has to be characterized in the respective laboratory. Even though the MOS transistors are generic devices, with some small variations from one to another manufacturer [1]. Each process produces a unique device that has to be analyzed and characterized independently to obtain authentic and scalable models to accurately simulate a circuit of these devices [2–4].

The lowering of the cost and chip area due to total integration of RF building blocks, which makes the Si-based designs a better choice for implementation. Moreover, the advancement in performance of CMOS technology for integrated power amplification in frequency range of gigahertz have also been reported through this investigation. The enhancement of RF power amplifier is quite important in terms of modern wireless communication systems [5–7]. Many techniques have already been innovated to have efficiency enhancement in power amplification designs. From linearity point of view, improved techniques are used to do cancellation of non-linear terms in CMOS power amplifiers. As, the design of transistors are also crucial to control the dc supply consumption to ultimately improve the overall efficiency [8].

In this TF-FinFET, the I_{OFF} is lying in the order of 10^{-10} A/um and hence has the less discharging ability from battery perspective, which is good for high performance. In TF-FinFET structure, SiO_2 is replaced with HfO_2 as gate oxide. Due to this replacement, the TF-FinFET has enhanced amount of I_{ON} and also good depletion of the channel at off-state (i.e., at $V_{GS} = 0$). Here in our design, the structure although is having a look like GAA FinFET structure, but it works as an extension of 3-gated FinFET [9]. As the transistors are compared on the basis of their gate oxide thicknesses. Both oxide widths of high K-material HfO_2 on the gates have their own individual benefits, due to high gate controllability as

✉ Rishu Chaujar
chaujar.rishu@dtu.ac.in

Mridul Prakash Kashyap
mridul.prakash2012@gmail.com

[1] Applied Physics Department, Delhi Technological University, New Delhi, India

Published online: 06 October 2020

compared to conventional FinFETs having SiO_2 as gate oxides. From this investigation, we get to know that device have excellent performance of analog and power amplification when simulated at Tox of 0.4 nm and 1 nm respectively.

2 Device Structure and Simulation Methodology

In Fig. 1 (a) and (c) shows the TF-FinFET and simplified small signal circuit model used in this investigation for terminology aspects respectively, and (b) shows the more detailed two port structural parameters of TF-FinFET. In this device, the fin is 6 nm tall above the STI and one-third of its length is dedicated to the truncatedness of this overall device. The lower fin (W_{fl}) is only one-third of the width of fin (i.e., 3 nm) and thinner than upper fin by 2 nm at the same time. The upper fin (W_{fu}) is two-third of remaining height (H_f) of the fin and having 3 nm as its width [11, 12]. The fin is wrapped in such a way that, the thickness of gate oxide remained as same on the overall surface area of gate region. Substrate width is also adjusted in accordance with the width of lower fin (i.e., 1 nm), while keeping GaAs as the substrate material [13]. This truncatedness of the fin is dedicated to enhance the analog performance of our device and along with the doping profile, it is also giving excellent performance in terms of power amplification. All the structural parameters of our device are shown by Table 1.

This replacement of HfO_2 with SiO_2 as material of gate oxide enhanced the gate controllability (i.e., by giving higher driving current) [14]. Both Devices have same device length of 25 nm, same source/drain length of 11 nm, same device height of 9 nm and same gate length (Lg) of 3 nm. The value of electrode metal (Molybdenum) workfunction is taken as 4.55 eV for source, drain and gate metal electrode. The following physical models used under the drift-diffusion formulation are modification of local density approximation for quantum correction. All the simulation work is performed on Silvaco Atlas TCAD EDA tools [15–19]. Models such as conc. Dep. Mobility and Re-combination model (which indicates the recombination of majority (i.e., electrons) and minority charge carriers in the device), Fermi-Dirac statistic (which describes the distribution of electrons as majority charge carriers over energy states in a system consisting of many identical particles that obeys Pauli exc. Principle), hydrodynamic energy transport model (which adds continuity equations for the carrier temperature, treat mobilities and impact ionization coefficients as function of temperature instead of function of local electric fields) and Band to Band tunnelling model (which accounts for the dynamic variation in electric field at the tunnelling junction and the carrier movement through the bands which shouldn't able to move, when assumed classically).

The Doping concentration of both the simulated n-type devices having different widths of gate oxides kept as same which are 1.0×10^{20} (cm^{-3}) for source, 1.0×10^{18} (cm^{-3}) for channel, 1.0×10^{19} (cm^{-3}) for drain and 1.0×10^{15} (cm^{-3}) for substrate. Doping profile of the devices kept as simple as possible [20]. This is also useful, if the device will further be considered as the idea of fabrication. Moreover, the remaining structural parameters which are used during the simulation are given in Table 1.

3 Simulated Results and Discussion

In this investigation, all the simulations are performed for n-type devices with different gate oxide thicknesses. As, the device with Tox = 0.4 nm works best in terms of analog and low leakage perspective as compared to the device with Tox = 1 nm. So firstly, we are going to show some characteristics of both the devices by ramping the gate voltage.

Figure 2(a) showed that the device at Tox = 0.4 nm is quite good from the device at Tox = 1 nm in terms of input

Table 1 Structural Parameters for Device Simulation

Parameters	Dimensions of Simulated Device
Gate length (L_g)	3 nm
Fin height (H_f)	6 nm
Lower fin width (W_{fl})	1 nm
Upper fin width (W_{fu})	3 nm
Effective oxide thickness (T_{ox})	0.4,1 nm
Gate Electrode Thickness ($T_{elec.}$)	0.05 nm
Gate metal workfunction (Molybdenum)	4.55 eV
Substrate width (W_s)	1 nm
Contact Resistivity	53.4 nΩ.m

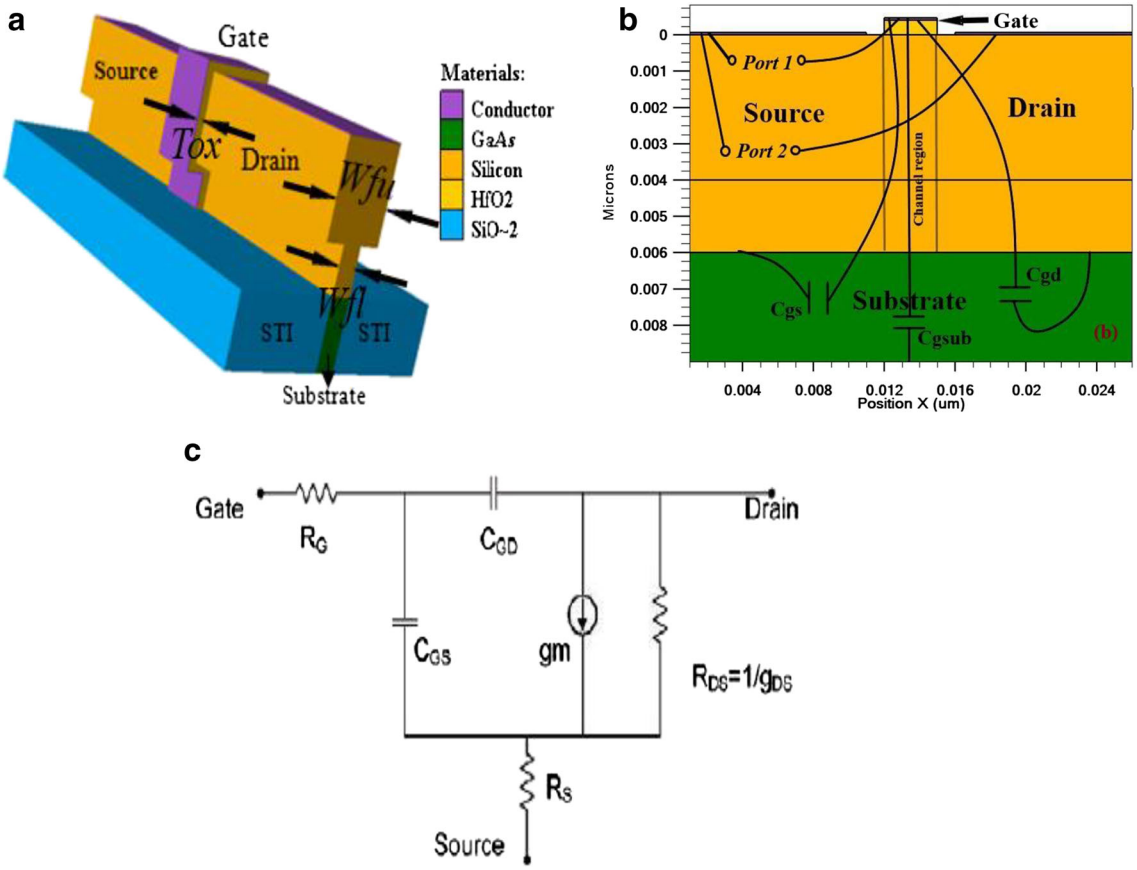


Fig. 1 (**a**) A 3-D view of JL bulk TF-FinFET. (**b**) 2-D cross-sectional two port view of TF-FinFET. (**c**) Circuitry view of TF-FinFET used for both thicknesses of gate oxide

characteristics obtained at $V_{DS} = 0.2$ V and frequency of input signal at 1GHz. In case of device with Tox = 0.4 nm, the saturation current (i.e., at $V_{GS} = 1.6$ V) comes out to be 0.129 mA/fin and 0.11 mA/fin for the device with Tox = 1 nm. Moreover, the I_{OFF} of the devices has majorly different values. As it stood 3.94×10^{-10} for Tox = 0.4 nm and 1.12×10^{-9} for the devices having Tox = 1 nm. Here we seen the great performance in case of Tox = 0.4 nm over the device having Tox = 1 nm. This is due to good gate controllability, as gate has more control over the movement of majority charge carriers in the channel region in this case.

As depicted from Fig. 2(b), the variance of gate-gate capacitance (C_{gg}) on gate voltage (V_G) at $V_{DS} = 0.2$ V and applied gate signal frequency of 1Ghz. The nature of C_{gg} analyzed by the graphical representation is mainly due to the accumulation of electron charge carriers in the vicinity of gate region increases and hence the parasitic (Cgg = Cgs + Cgd) capacitance rises. If seen from the view of gate oxide thickness, C_{gg} stood minimum for Tox = 0.4 nm and maximum for Tox = 1 nm at their respective peak values (i.e., $V_{GS} = 1.6$ V). As f_T is defined as the frequency at which current gain (ration of drain to source and gate to source current) is close to unity. We can easily depict from Fig. 2(a) that the current started to saturate from $V_{GS} = 0.4$ V and onwards. That is how we obtained a peak in Fig. 2(b) at $V_{GS} = 0.4$ V. f_T is maximum for Tox = 1 nm and minimum for Tox = 1 nm because of better current saturation in this case which can be depicted from 1st fig. of this series.

Figure 2(c) denotes the input conductance (Transconductance) of the simulated devices. It is proportional to the linear gain of a device. Here as observed, g_m of TF-FinFET at Tox = 0.4 nm leads to the g_m of the device at Tox = 1 nm from $V_{Gs} = 0.4$ V and continues as it is up to $V_G = 1.6$ V. This lead is due to better gate controllability over channel region in case of TF-FinFET at Tox = 0.4 nm. TF-FinFET gives more incremental change in I_D at Tox = 0.4 nm as compared to the TF-FinFET at Tox = 1 nm from same input gate voltage. *IP*3 (third-order intercept point) shows the derived input power at which the first and third harmonic will get equal. It is a unique quantity by which itself, one can serve it as a means of differentiating the linearity of non-identical devices. It is shown by Fig. 2(c) and can be seen that the device having Tox = 1 nm is far better than the device having Tox = 0.4 nm in terms of linearity at $V_{GS} = 0.3$ V and 1.3 V, which ultimately declined towards the superior amplification capacity of the device [21]. Given below are some of the equations on which the particular quantities are obtained during the simulation.

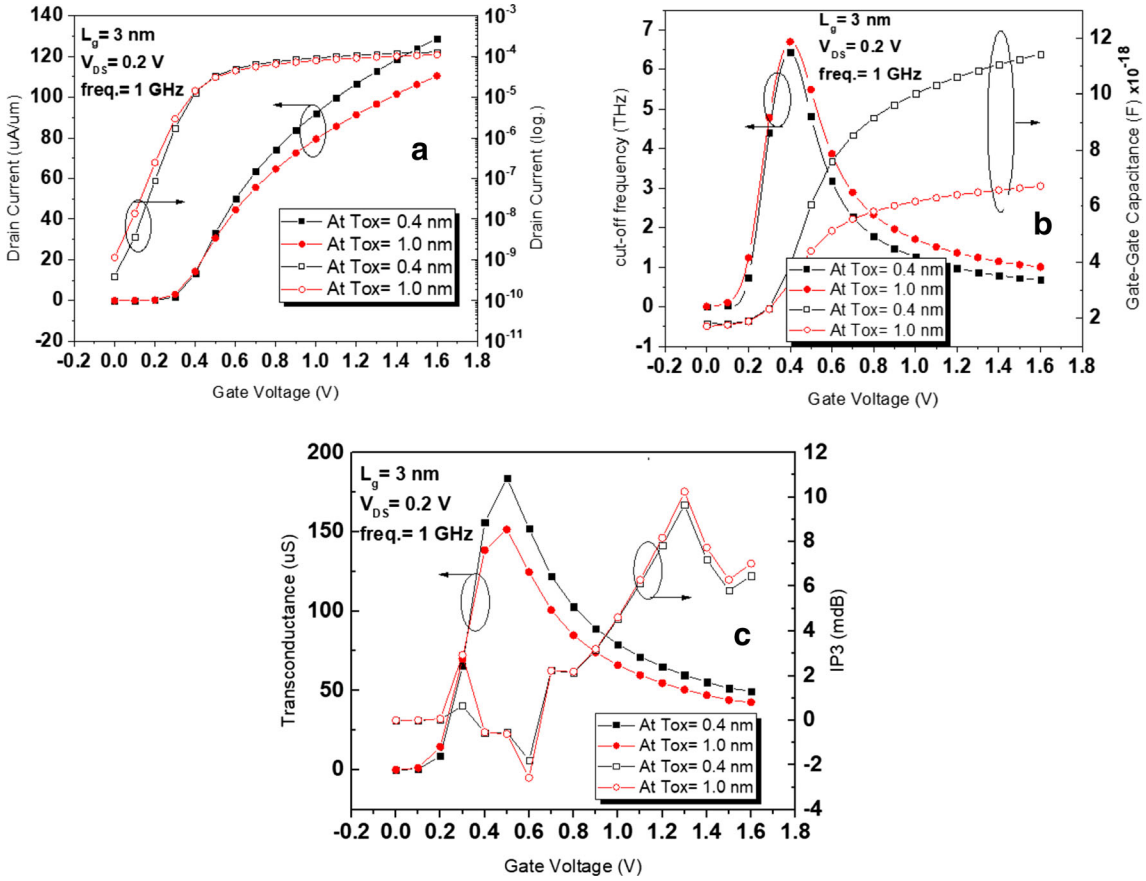


Fig. 2 (**a**) Dependency of Drain current (I_D) on V_{GS}. (**b**) Dependency of cut-off frequency and gate capacitance on V_{GS}. (**c**) Dependency of Transconductance and IP3 (third-order intercept point) on V_{GS} at $V_{DS} = 0.2$ V and frequency of input signal of 1GHz for all three cases in TF-FinFET

$$f_T = \frac{g_m}{2\pi C_{gg}} \tag{1}$$

$$f_{max} = \frac{f_T}{\sqrt{4R_g\left(g_{ds} + 2\pi f_T C_{gd}\right)}} \tag{2}$$

$$G_u = \frac{1}{2}\,\frac{\left|\frac{S_{21}}{S_{12}} - 1\right|^2}{k\left|\frac{S_{21}}{S_{12}}\right| - Re\left(\frac{S_{21}}{S_{12}}\right)} \tag{3}$$

Where,

$$k = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{12}.S_{21}|} \tag{4}$$

(and) $\Delta = S_{11}S_{22} - S_{21}S_{12}$

$$G_{ms} = \frac{|S_{21}|}{|S_{12}|} \tag{5}$$

$$IP3 = (2/3)\left(\frac{g_m}{g_m{}''(R_s)}\right) \tag{6}$$

$$t_i = \frac{C_{gg}.V_D}{I_{on}} \tag{7}$$

Figure 3 showed the variation of intrinsic delay (t_i) of TF-FinFET with different applied Tox on the gate of it. As

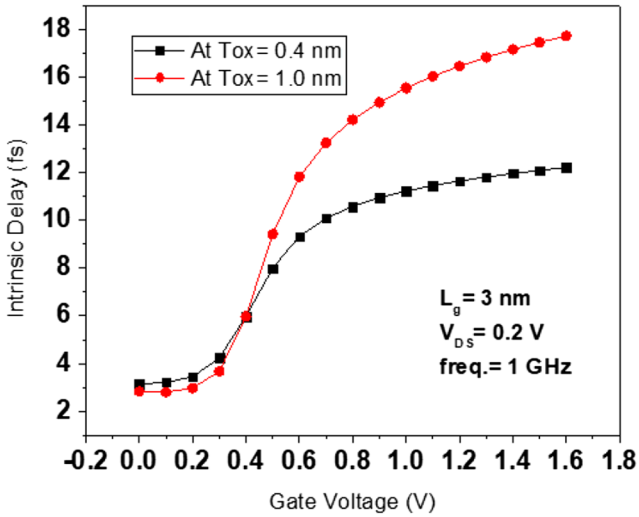


Fig. 3 Dependency of Intrinsic Delay (t_i) on V_{GS} at $V_{DS} = 0.2$ V and frequency of input signal of 1GHz for both cases of TF-FinFET

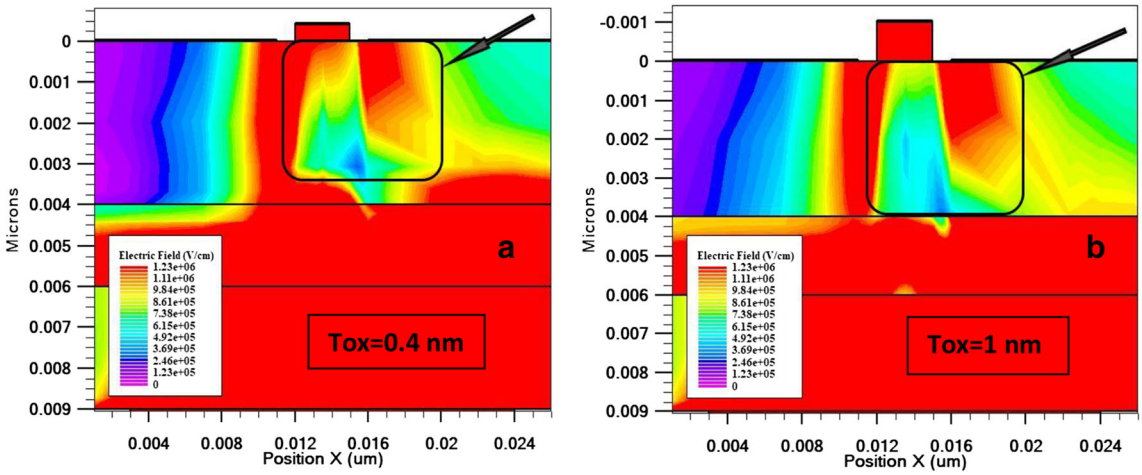


Fig. 4 (**a**) Dependency of Electric Field on V_{GS} = 0.5 V and V_{DS} = 0.5 V at Tox = 0.4 nm. (**b**) Dependency of Electric Field on V_{GS} and V_{DS} in TF-FinFET at Tox = 1.0 nm

intrinsic delay majorly depends upon the value of C_{gg} of the particular device due to very less difference of I_{ON} in both the cases of different Tox which can easily be depicted from eq. (7). This plot is simulated at V_{DS} = 0.2 V and 1GHz of applied signal frequency has a maximum value of intrinsic delay (t_i) of 12.2 fsec. and 17.7 fsec. For Tox = 0.4 nm and Tox = 1 nm respectively at V_{GS} = 1.6 V. As intrinsic delay denotes the time interval taken by the device to give its output after applying input to it in which the device with Tox = 0.4 nm is far better than the device with Tox = 1 nm. This is due the more input power consumption in case Tox = 0.4 nm as compared to device with Tox = 1 nm. This phenomenon can easily be analyzed from the input energy consumption and delay inverse relationship of the MOSFETs, and here our device can be assumed as the tri-gated MOSFET [22].

Figure 4 showed the trend of electric field in our proposed device TF-FinFET at 2 different gate oxide widths, with the help of contour plots if seen cross-sectionally after bisecting them from their mid-point. This analysis is performed at V_{DS} = 0.5 V, V_{GS} = 0.5 V and at 1 GHz frequency of applied input signal by keeping highest and lowest values of electric field as 1.23×10^6 V/cm and 0 V/cm respectively. As we can see by the area showed by an arrow, indicates large amount of electric field between gate-drain region in case of Tox = 1 nm. Tox = 0.4 nm large current leakage as compared to the device with Tox = 1 nm.

Figure 5. (a) showed the dependency of Drain-Source Resistance (R_{DS}) on particular values of V_{GS} and V_{DS} as shown. The highest and lowest value set of R_{DS} obtained on varying frequency of applied input signal are (2.92, 2.52) KΩ and (2.99, 2.72) KΩ for the device at Tox = 0.4 nm and Tox = 1 nm respectively. The device at Tox = 1 nm is far better than device with Tox = 0.4 nm as shown in respective figure. Moreover, higher R_{DS} is desirable at the output port to have good power amplification. The main reason behind enhancement of R_{DS} in case of Tox = 1 nm is the low amount of g_{ds}, due to more decremental change in output current. For the delay point of view, occurred during the switching of transistors. As the switching of transistors depends on charging-discharging of capacitors assumed at the respective port.

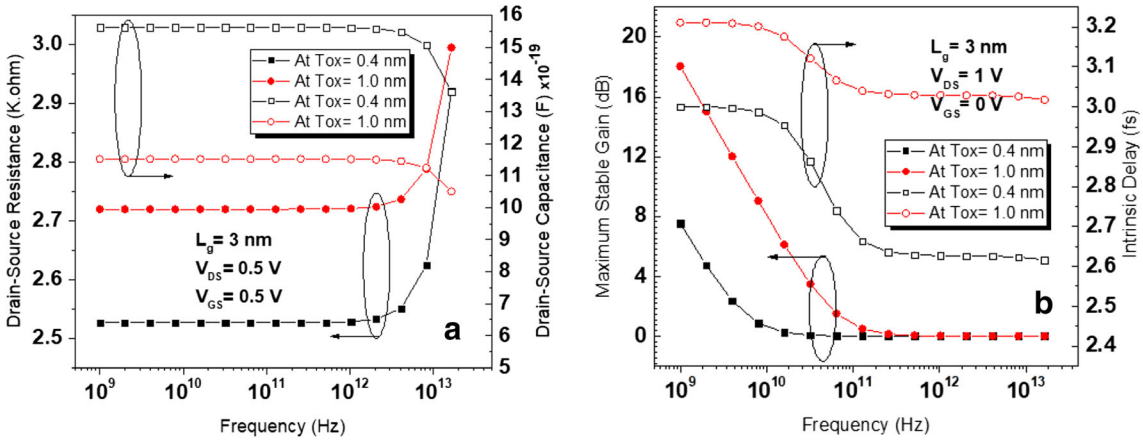


Fig. 5 (**a**) Variance of Drain-Source Resistance and capacitance w.r.t frequency at both Tox. (**b**) Variance of Maximum Stable Gain and Intrinsic Delay w.r.t frequency at both applied Tox

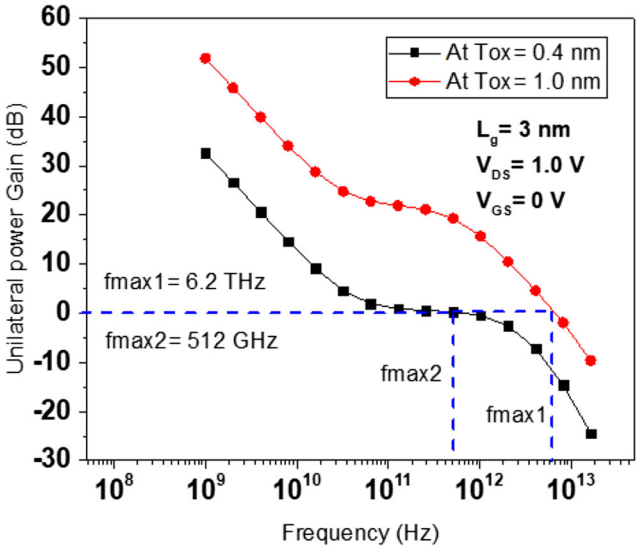


Fig. 6 Variation of Unilateral power Gain (G_u) w.r.t frequency of applied signal at $V_{DS} = 1$ V and $V_{GS} = 0$ V

Less capacitance would ultimately give speed to a device at the time of switching. So, in these aspects too, the device with Tox = 1 nm stood bigger than the device at Tox = 0.4 nm at this high frequency range of input signal in the figure.

Figure 5. (b) depicted the variation of Maximum Stable Gain (G_{ms}) and Intrinsic Delay at specified values of V_{DS}, V_{GS} and w.r.t frequency of input signal applied. One can easily analyze that the device with Tox = 1 nm has higher values of G_{ms} as compared to the device with Tox = 0.4 nm at starting frequency of 1GHz (i.e., 18.04 and 7.57 respectively). Moreover, the values of G_{ms} at last point of analysis are 44×10^{-4} and 3.7×10^{-4} which are still greater than zero for Tox = 1 nm and Tox = 0.4 nm. As we can see from eq. (5), the through mechanism (S_{21}) should be larger than feedback mechanism (S_{12}) to have higher stability gain (G_{ms}). This phenomenon is also declined towards the device with Tox = 1 nm. Apart from this, the same fig. Also indicated the variation of Intrinsic Delay (t_i) on the same applied parameters. Through this, the device with Tox = 1 nm stood more slower (i.e., has more delay) than the device with Tox = 0.4 nm, due to the corresponding delay values w.r.t frequency of applied signal [23].

Figure 6 evident the Unilateral power Gain (G_u) variation with respect to variation in frequency of applied input signal, for the device with different Tox applied to it. The same can be easily analyzed by eq. (3), which is performed at $V_{DS} = 1$ V and $V_{GS} = 0$ V. As, unilateral power gain indicates the maximum operable frequency (f_{max}) of the device by depicting this plot at the frequency on which, the gain would become close (i.e., extrapolated) to zero. Moreover, the results of this technique are satisfactory, when the device (amplifier) is operating at this VHF frequency with a narrow bandwidth from 1 GHz to few THz. Here we can say that the device with Tox = 1 nm has less internal feedback mechanism (i.e., S_{12} is small) as compared to the device with Tox = 0.4 nm [24, 25]. Moreover, f_{max} for the device of Tox = 1 nm found much greater than the device with Tox = 0.4 nm as shown in respective plot.

All the observation (i.e., Gains) are evaluated on the basis of Fig.7. (a) and (b) for Tox = 0.4 nm and Tox = 1 nm respectively. Here, the major differences between the two smith-charts are based on S_{21} and S_{22}, however the changes can also be seen in S_{11} and S_{12} which are quite minor comparatively.

3.1 Evidence of the Novelty of TF-FinFET

Figure 8. shows the calibration of input characteristics of the 30 nm (i.e., gate length) fabricated device as reference, where "experimental" signifies the actual results of fabricated device analyzed for SiO_2 spacer of respective paper [10] and "simulated" signifies the simulated results of the same device on our

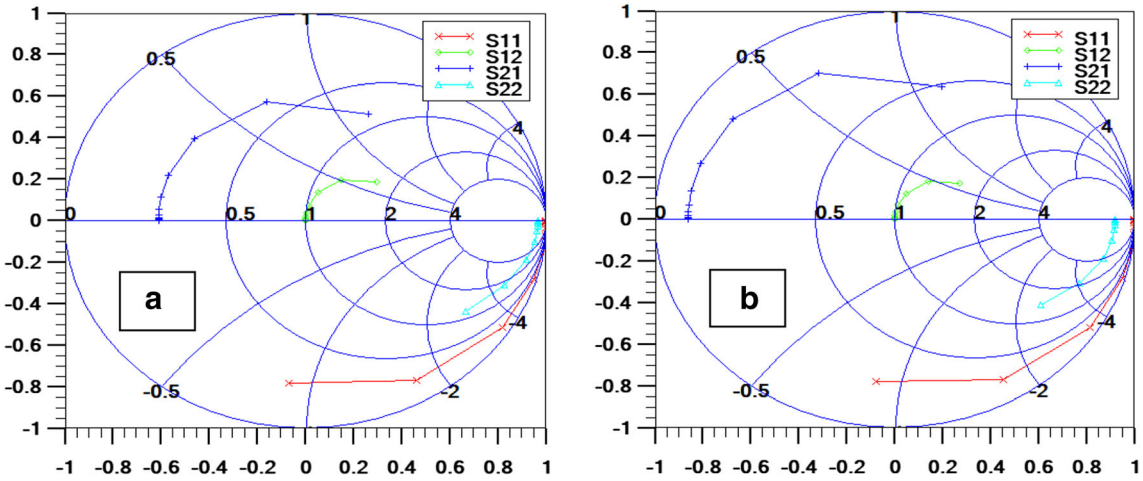


Fig. 7 (**a**) Variance of scattering parameters at Tox = 0.4 nm. (**b**) Variation of scattering parameters at Tox = 1 nm, At $V_{GS} = 0.3$ V and $V_{DS} = 0.4$ V for both

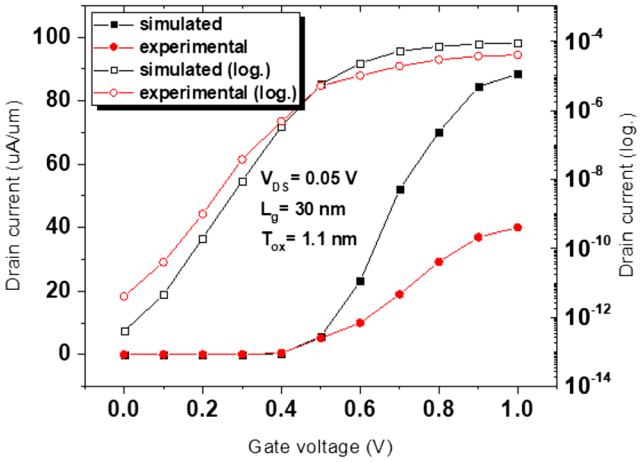


Fig. 8 Variation of Drain current (I_D) w.r.t applied Gate voltage at V_{DS} = 50 mV

Silvaco EDA TCAD tools. Model parameters which are used in the simulation have been calibrated according to the experimental results by [10].

4 Conclusion

In this investigation, the TF-FinFET with different widths of gate oxides can be used for analog as well as amplification purposes. The DC and AC performance of the devices stood excellent, as claimed by this simulation. As, the maximum cut-off frequency for the device with Tox = 0.4 nm and 1 nm analyzed as 6.44 and 6.70 THz respectively w.r.t V_{GS}. Furthermore, it gave maximum operable frequency (f_{max}) as 512 GHz and 6.2 THz for the device with Tox = 0.4 nm and Tox = 1 nm respectively as analyzed from Unilateral Power Gain (G_u). If seen from amount of I_{ON} and I_{OFF} perspective, both the devices worked well. So, only by adjusting the gate oxide thickness, device can work for multiple purpose or as per requirement. This device has great potential for wide variety of applications to high-speed ICs.

Acknowledgments The authors are grateful to Microelectronics Research Lab, Department of Applied Physics, Delhi Technological University for supporting this research work.

References

1. Kumar A, Tripathi MM, Chaujar R (2017) Investigation of parasitic capacitances of In2O5Sn gate electrode recessed channel MOSFET for ULSI switching applications. Microsystem Technol 23(12): 5867–5874. https://doi.org/10.1007/s00542-017-3348-2
2. Kumar A, Gupta N, Chaujar R (2016) Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behavior. Microsystem Technol 22(11):2665–2671. https://doi.org/10.1007/s00542-015-2554-z
3. Kumar A, Tripathi MM, Chaujar R (2018) Reliability issues of In2O5Sn gate electrode Recessed Channel MOSFET: impact of Interface trap charges and temperature. IEEE Trans Electron Devices 65(3):860–866. https://doi.org/10.1109/TED.2018.2793853
4. Kranti A, Armstrong GA (2007) Design and optimization of FinFETs for ultra-low-voltage analog applications. IEEE Trans Electron Devices 54(12):3308–3316. https://doi.org/10.1109/TED.2007.908596
5. Madan J, Chaujar R (2017) Numerical simulation of N+ source pocket PIN-GAA-tunnel FET: impact of Interface trap charges and temperature. IEEE Trans Electron Devices 64(4):1482–1488. https://doi.org/10.1109/TED.2017.2670603
6. Gupta S, Nandi A (2017) Effect of air spacer on analog performance of underlap tri-gate FinFET. Superlattice Microst 109:693–701. https://doi.org/10.1016/j.spmi.2017.05.053
7. Su PH, Li Y (2018) Analysis of in-line process parameters of the Unity gain frequency of HKMG bulk FinFET devices. IEEE Electron Device Lett 39(3):335–338. https://doi.org/10.1109/LED.2018.2791436
8. Thirunavukkarasu V, Jhan YR, Liu YB, Wu YC (2015) Performance of inversion, accumulation, and Junctionless mode n-type and p-type bulk silicon FinFETs with 3-nm gate length. IEEE Electron Device Lett 36(7):645–647. https://doi.org/10.1109/LED.2015.2433303
9. Zheng P, Connelly D, Ding F, Liu TJK (2015) Simulation-based study of the inserted-oxide FinFET for future low-power system-on-Chip applications. IEEE Electron Device Lett 36(8):742–744. https://doi.org/10.1109/LED.2015.2438856
10. Sachid AB, Chen MC, Hu C (2017) Bulk FinFET with low-κ spacers for continued scaling. IEEE Trans Electron Devices 64(4):1861–1864. https://doi.org/10.1109/TED.2017.2664798
11. Bolognesi C, Dvorak MMW, Yeo P, Xu XG, Watkins S InP/GaAsSb/InP double HBTs: a new alternative for InP-based DHBTs. IEEE Trans Electron Devices 48:2631–2639. https://doi.org/10.1109/16.960389
12. Cho, Seongjae & Kang, In & Kamins, Theodore & Park, Byung-Gook & Harris, James. (2011). Silicon-compatible compound semiconductor tunnelling field-effect transistor for high performance and low standby power operation. Appl Phys Lett. 99. https://doi.org/10.1063/1.3670325
13. Cho S, Kim K, Park B-G, Kang I (2011) RF performance and small-signal parameter extraction of Junctionless silicon nanowire MOSFETs. IEEE Trans Electron Devices 58:1388–1396. https://doi.org/10.1109/TED.2011.2109724
14. Bhattacharya M, Jogi J, Gupta R, Gupta M Scattering parameter based modelling and simulation of symmetric tied-gate InAlAs/InGaAs DG-HEMT for millimeter-wave applications. Solid-State Electron - Solid State Electron 63:149–153. https://doi.org/10.1016/j.sse.2011.05.025
15. Singh R, Aditya K, Parihar S, Chauhan Y, Vega R, Hook TB, Dixit A Evaluation of 10nm Bulk FinFET RF Performance - Conventional vs. NC-FinFET. IEEE Electron Device Lett. https://doi.org/10.1109/LED.2018.2846026
16. Sanchez Esqueda I Confinement Effects on Radiation Response of SOI FinFETs at the Scaling Limit. IEEE Electron Device Letters, pp 1–4. https://doi.org/10.1109/LED.2017.2661239
17. Zhang J, Niu G, Cai W, Wang W (2019) Intermodulation linearity characteristics of 14-nm RF FinFETs. IEEE Trans Electron Devices 66:1–7. https://doi.org/10.1109/TED.2019.2912516
18. Madan H, Saripalli V, Liu H, Datta S (2012) Asymmetric tunnel field-effect transistors as frequency multipliers. IEEE Electron Device Lett 33:1547–1549. https://doi.org/10.1109/LED.2012.2214201
19. ITRS 2.0 Publication, "International Technology Roadmap for Semiconductors, 2015," Available online at http://www.itrs2.net/

20. PH Drive, S Clara, (2017) Atlas User's Manual, Device Structure Editing and Simulation Software. Pp. 1–1715. Available: www.silvaco.com
21. Cho S, Kim H, Jhon H, Kang I, Park B-G, Harris J (2013) Mixed-mode simulation of nanowire Ge/GaAs Heterojunction Tunnelling field-effect transistor for circuit applications. IEEE J Electron Dev Soc 1:48–53. https://doi.org/10.1109/JEDS.2013.2256458
22. Palanichamy DV (2016) Modelling and simulation of unilateral power gain for GaN/AlGaN HEMT. Int J Adv Eng Global Technol 4:2105–2109
23. Kumar A, Tripathi MM, Chaujar R (2017) Investigation of parasitic capacitances of In2O5Sn gate electrode recessed channel MOSFET for ULSI switching applications. Microsyst Technol 23:5867–5874. https://doi.org/10.1007/s00542-017-3348-2
24. Hofmann AJL et al (2019) Dipolar Doping of Organic Semiconductors to Enhance Carrier Injection. *Phys. Rev. Appl.* 12(6):1
25. Berg M et al (2019) Imaging Atomically Thin Semiconductors Beneath Dielectrics via Deep Ultraviolet Photoemission Electron Microscopy. Phys Rev Appl 12(6):1

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-020-02566-y

ORIGINAL RESEARCH

Improved ant colony optimization for achieving self-balancing and position control for balancer systems

Rupam Singh[1] · **Bharat Bhushan**[1]

Received: 11 April 2020 / Accepted: 18 September 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
The balancer systems represent feedback in loop-based underactuated system which is electromechanical, multivariate, and nonlinear. This paper develops a self-balancing controller using an improved ant colony optimization (ACO) to optimize the proportional integral derivative controller (PID) controller. The proposed controller achieves self-balancing control for a ball on the plate by controlling the plate inclination angle. Initially, the modelling of the ball balancer system is achieved with the help of a two degree of freedom (2DoF) ball balancer system controlled by a PID controller. Further, ACO is employed to autonomously evaluate the condition of a process and find the optimal tuning parameters for the PID controller. The transition probability of the ACO is revised to improve the response and convergence speed of the algorithm resulting in an improved ACO. The developed control schemes were applied with the 2DoF ball balancer model both in simulation as well as for the real-time operation. The results depicted the performance of the proposed control scheme by analysing the characteristics such as transient response and steady-state error. Further, stability analysis has been done for the developed control schemes using describing function method for multiple frequencies. The results depicted the superiority of the improved ACO based PID controller over the conventional PID controller.

Keywords Self-balancing control · Ball balancer setup · Proportional integral derivative control · Ant colony optimization

1 Introduction

The approximation of underactuated nonlinear systems through automatic decision development and intelligent control methods (Murray et al. 2003) is an issue that appears in many problems (Nelles 2001) and can be tackled through various approaches. The diversity and complexity of these systems have led researchers in the field to analyse the action of various linear, nonlinear, model-free, passivity, and intelligent controllers. These controllers are focussed at achieving self-balancing control and steady-state operation for various systems like inverted pendulum (Boubaker 2012; Moness et al. 2020), the twin rotor multi input multi output system (TRMS) (Chalupa et al. 2015), the ball beam system (Andreev et al. 2002; Nowopolski 2013), hovercraft (Aranda et al. 2006), Furuta pendulum (Acosta 2010), and ball and plate system (Awtar et al. 2002). In general, linear controllers offer a simple way of designing closed-loop control for these systems. Various linear controllers were available in the literature to solve the problems in underactuated systems (de Jager 1998; Olivares and Albertos 2013; Aguilar-Avelar and Moreno-Valenzuela 2015). But the complicated nonlinear dynamics of underactuated systems affect the capability of providing a plausible solution and limits the generalized applications of control laws. Besides, the linearization of the nonlinear systems is also carried which heavily affect the speed of system response. This motivated the development of several nonlinear control techniques to deal with the problems in underactuated systems. Lots of nonlinear controllers like Lagrangian, lambda method, and backstepping controller (Chang 2008; Choukchou-Braham et al. 2014; Rudra et al. 2017), have been evolved in the last few years. These controllers faced some drawbacks during the early stage of their development with the limitations of lambda method in dealing with external load and back stepping controllers lagging due to additional feedback. Apart from the above controllers, the adaptive iterative learning methods

✉ Rupam Singh
singhrupam99@gmail.com

Bharat Bhushan
bharat@dce.ac.in

[1] Department of Electrical Engineering, Delhi Technological University, Shahbad Daulatpur, Delhi 110042, India

Published online: 29 September 2020

like Kalman particle filter are also developed to achieve the nonlinear control (Li 2020). The major draw back of the particle filter approach is its dependency on multiple calculations which exponentially increase with the number of state variables (Sutharsan et al. 2012).

These issues in the conventional nonlinear techniques were overcome by carrying out feedback linearization (Aguilar-Avelar and Moreno-Valenzuela 2016; Moreno-Valenzuela and Aguilar-Avelar 2018), and partial feedback linearization (Spong 1994) for underactuated mechanical systems. This method transformed the nonlinear system into an equivalent linear system by cancelling the nonlinearities through feedback control. However, the problem with lack of robustness is considered as a major drawback while adapting the procedure of feedback linearization approach and the partial feedback linearization. To achieve robustness while controlling, the sliding mode approach is considered as a reasonable solution (Cheng and Ho 2017). In Ye et al. (2020), the sliding mode controller is operated with a neural network based two wheeled mobile robot to enhance the control framework by accelerating the convergence speed of the tracking error. But the switching surface behaviour of the sliding mode controller makes it insensitive to external disturbances and parametric variations. Another approach in the line of nonlinear controllers for underactuated systems is the passivity-based control approach (Shiriaev et al. 1999; Ortega et al. 2002). This aims at passivating the system with a storage function, which has a minimum at the desired balance point. This has a disadvantage with differential feedback as it cannot amplify the measurement noises.

Considering the drawbacks while controlling the underactuated systems, an efficient control strategy capable of achieving steady-state operation for the above-mentioned systems needs to be developed. To achieve this, a ball and plate system is chosen as a substantial benchmark problem for underactuated systems. Generally, the ball balancer system is a platform where the ball moves on the top of a plate which is connected with the shaft of an electric system. The ball can be positioned around a fixed plate area by controlling the motors in the electrical system. This control operation is performed by developing a feedback controller. Conventionally, exchanging controllers is developed depending on the non-straight examination of the ball and plate framework (Tian et al. 2006). Simultaneously, depending upon the Lyapunov dependability hypothesis, the back-venturing controllers were developed to accomplish an explicit control law and achieved the control execution rapidly (Ker et al. 2007). Further, the proportional integral derivative (PID) controller which is nonlinear has been considered for one-to-one control of the ball on the plate system in classical control (Sun and Li 2012). The reaction of PID is upgraded, when structured dependent on generalized Kalman–Yanukovych–Popov lemma (GKYPL) strategy, and contrasted a typical PID concerning the relentless state reaction (Mochizuki and Ichihara 2013). An inventive disturbance observer utilizing adaptive disturbance rejection control (ADRC) strategy is intended to maintain a strategic distance from destructive grating consequences for the plate for adjusting of the ball (Wang et al. 2014), and examination among resonant and repetitive controller has been made dependent on linear quadratic regulator (LQR) strategy (Da Silveira Castro et al. 2014). Further, disturbance rejection controllers (Pinagapani et al. 2018) and metaheuristic optimization algorithms (Ali et al. 2019) were illustrated on ball and plate systems to achieve desired tracking performance. The action of sliding mode controller (SMC) (Hammadih et al. 2016; Bang and Lee 2018) for achieving self-balancing control and the development of fractional-order sliding mode control (Das and Roy 2017) which overcomes the chattering phenomenon of conventional SMC with enhanced efficiency has been widely discussed in (Debono and Bugeja 2015; Basci and Can 2017; Kao and Ho 2018). The model predictive controllers were also widely employed with ball balancer systems because of their advantages with time varying reference (Bang and Lee 2019). But, the major drawback with these conventional techniques is that, they resulted in large settling time and peak overshoot. In addition, the position of the ball can likewise be controlled through a machine vision framework. The picture handling algorithms for the same framework was linked and implemented on an field-programmable gate array (FPGA) appliance to encounter continuous requirements of the real-world (Ho et al. 2013). B&P demonstrated by the computer-generated laboratory utilizing 3D java re-enactment pursues the predefined direction for the ball and accomplished point to point controllability (Fabregas et al. 2015). The full state feedback input linearizing control law was built up and could control divergent rapports and additionally link among the co-ordinate axis (Hoover and St. Amand 2012).

Besides, the self-balancing control, trajectory tracking and ball position control of the ball and plate system were achieved by various intelligent and hybrid controllers such as fuzzy (Rad et al. 2003; Fan et al. 2004; Zhang and Yuan 2010), fuzzy cerebellar model articulation controller (Moreno-Armendariz et al. 2010), and particle swarm optimization based fuzzy-neural controller (Dong et al. 2011). In advanced control, a Fuzzy controller with online learning control has been demonstrated to an open-loop unstable system (Rad et al. 2003). A novel control scheme has been designed using fuzzy with a genetic algorithm for tracking, planning, and supervisory control of ball motion (Fan et al. 2004). In another approach, the supervisory fuzzy controller has been proposed based on a single input rule module and visual servo control for trajectory tracking and motion control of the ball on the plate (Ming et al. 2006). A new hybrid controller has been designed with a sliding mode

controller tuned by the supervisory fuzzy for position control (Moarref et al. 2008). For the study on the multivariable and complex B&P system, a touchscreen, and a rotary pneumatic cylinder are used as a replacement for of a camera and step motor, correspondingly and controlling has been carried on fuzzy & state observer (Zhang and Yuan 2010). Fuzzy visual control of the Mamdani-type model has been applied to control the movement of a ball on the plate (Moreno-Armendariz 2010). Further, fuzzy with adaptive integral control improves basic fuzzy response and almost eliminates steady-state error (Pattanapong and Deelertpaiboon 2013). An indirect adaptive control strategy is presented using hierarchical fuzzy CMAC neuronal networks (Moreno-Armendariz et al. 2010). Recently linearized sliding mode controller which measures the beam angle from ball position has been implemented (Hammadih et al. 2016) and fractional-order sliding mode controller provides improved tracking accuracy and faster response to, less chattering and excellent efficiency over simple sliding mode controller (Roy et al. 2016; Das and Roy 2017). All though there are numerous control strategies available in the literature for achieving self-balancing control with balancer systems, the PID controller is widely adapted in practical engineering applications. The major advantages with the PID controller are its simple structure, high reliability, and good stability. But the key drawback while dealing with traditional PID controllers is their problem with parameter tuning.

There are many techniques available in the literature to achieve the tuning of PID parameters. These techniques involve various intelligent techniques like fuzzy (Wang et al. 2017) and neural network (Wang et al. 2020), self-tuning algorithms (Abdo et al. 2014), genetic (Zhang et al. 2009; Feng et al. 2018) and evolutionary algorithms (Hou 2017). However, these techniques are focused on saving the better generations resulting in local optimum rather than the global optima. Further, the problems due to weight adjustment of the intelligent controllers, low memory, premature convergence, weak local search, and high computational efforts for genetic and other evolutionary algorithms has resulted in the need for an optimal multi-objective approach for solving combinatorial optimization problems (Chang et al. 2012; Angelov 2016). One such method is ant colony optimization (ACO) technique. The ACO is a probabilistic method which has previously been applied to find the optimal PID parameters (Hsiao et al. 2004; Chiha et al. 2012). The main concept of ACO was originally proposed by Dorigo (Dorigo et al. 2008). It is based on the behavior of foraging ants that find the shortest path to a food source through a positive feedback mechanism (pheromones). This resulted in the application of ACO to various combinational optimization problems such as traveling salesperson problem (Tsai 2004; Angus and Hendtlass 2005), quadratic assignment problem (Tseng and Liang 2006), vehicle routing (Negulescu et al. 2008), and job shop scheduling (Heinonen and Pettersson 2007). Besides, a wide literature is available where the ACO is implemented with multi-objective path planning using different strategies (Ning et al. 2018; Masoumi et al. 2019). These studies identified the advantages of ACO in efficiently solving path tracking and other discrete problems, fewer memory constraints, ability to rapidly discover good solutions, robustness, easy implementation, and the algorithm does not need the prior knowledge of the controlled system (Nguyen et al. 2017). These advantages of ACO over the classical and conventional optimization techniques has motivated the authors in adopting ACO for tuning of PID controller in self-balancing and position control problems of a ball balancer system. Although the standard ACO showed poised performance, the problems with convergence, low accuracy, falling into local extremum, and insufficient ergodicity remained as the major shortcomings. These drawbacks were overcome in the literature by hybridizing the ACO algorithm with different evolutionary and metaheuristic algorithms for various applications. In Che et al. (2020), ACO was hybridized with particle swarm optimization (PSO) to achieve optimal path planning for autonomous under water vehicle. In the process of hybridization, it is identified that, the probability distribution of ACO changes for every iteration with the action of PSO. Further, the process has some drawbacks with uncertain convergence time, difficult theoretical analysis, and is highly dependent on sequences of random decisions. In light of the above disadvantages, this work proposes a dynamic pheromone update to achieve optimal performance of the ACO algorithm in tuning the PID parameters for a balancer system. The development of the controller is carried out by modelling the ball balancer system with a PID controller. Numerical simulations and experiments were carried to assess the action of the proposed controller. Further, describing function-based stability analysis is developed to assess the stability of the controlled system.

Further sections of the paper are organized as follows: Sect. 2 discusses the modelling of the ball balancer system; Sect. 3 develops an optimized PID controller for achieving self-balancing and position control. The simulation results and the stability analysis using the describing function method are discussed in Sect. 4. Section 5 discusses the real-time experimental analysis of the ball balancer system with a developed controller and the research is concluded in Sect. 6.

2 Ball balancer system

A ball on the plate system is a typical benchmark problem for achieving, balancing control, ball position tracking, and visual servo control. The typical representation of a ball and plate system is given in Fig. 1.

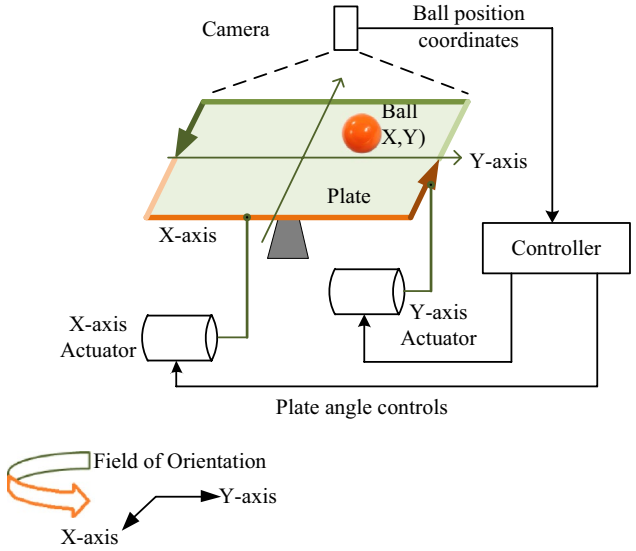


Fig. 1 Typical representation of ball and plate system

This is an underactuated system with access to four DoF controlled by two actuators. Hence its operation is referred to as a 2DoF ball balancer. The other parts of the system consist of a digital camera that captures the images of X and Y coordinates for the ball movement on the plate. Further, a vision algorithm computes the reads the ball coordinates from the image and provides information to the controller (Mochizuki and Ichihara 2013) for adjusting the angle of the plate through the X and Y servo motors to make the ball track the time-varying reference.

2.1 Mathematical modelling of plant model

Figure 2 depicts the block representation of transfer function modelling for the ball balancer system in an open loop. Plant $W_s(s)$ represents the dynamics between the servo input voltage and resulting load angle, whereas plant $W_{ss}(s)$ represents the angle of the servo load gear and dynamics of balls position. As the operation of the 2DoF ball balancer is dependent on the control of two rotary servo base units and the arrangement of the plate is symmetrical to these servo devices, it is anticipated that both the device dynamics are same. Hence the 2DoF ball balancer can be modelled as two decoupled ball and beam systems.

The overall transfer function of 1-d ball Balancer is given by

$$W(s) = W_{ss}(s)W_s(s) \tag{1}$$

where $W_{ss}(s) = \frac{A(s)}{\beta_l(s)}$, $W_s(s) = \frac{\beta_l(s)}{V_m(s)}$

2.2 Calculation of motion equation using the first principle

The free body diagram in Fig. 3 depicts a ball and plate system. By considering the equation of motion, the ball movement concerning the plate angle can be formulated. As per the law of inertia, the forces exerted on the ball regarding the angle of plate is given by:

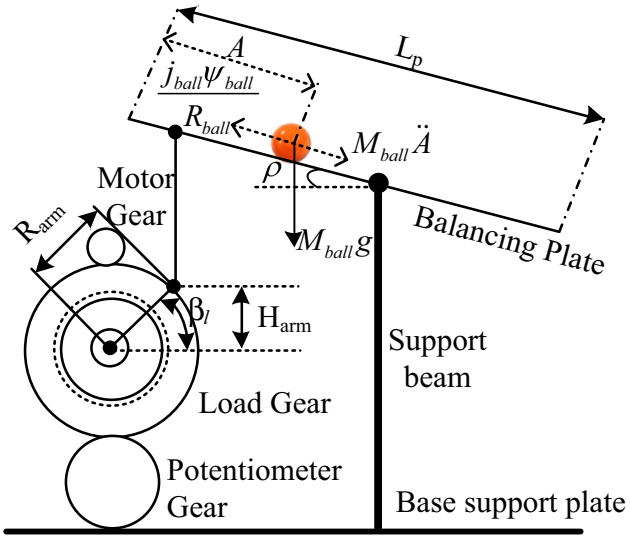


Fig. 3 Free body representation of the ball balancer system

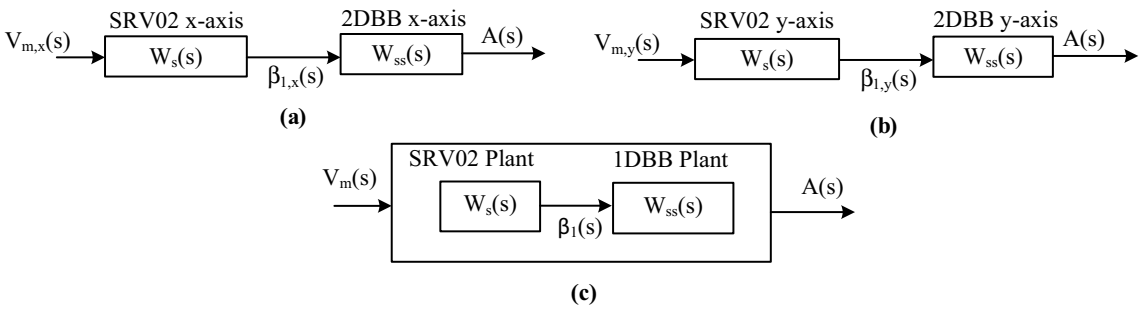


Fig. 2 Open-loop block diagram of 2DoF ball balancer system **a** x-axis servo, **b** y-axis servo, **c** 1D open-loop block diagram

$$M_{ball}\ddot{A}(t) = \sum F = F_{x,t} - F_{x,r} \tag{2}$$

Here the mass of the ball is given by M_{ball}, A depicts the displacement of the ball, $F_{x,r}$ is the force of inertia of ball, lastly $F_{x,t}$ is the translational force due to gravity. The friction damping and viscous constants are ignored.

The force $F_{X,t}$ in the direction along with the plate (X) due to gravity is formulated as:

$$F_{x,t} = M_{ball}G\sin\rho(t) \tag{3}$$

By turning the ball, the force $F_{x,e}$ is formulated as:

$$F_{x,e} = \frac{\tau_{ball}}{R_{ball}} \tag{4}$$

where τ_{ball} is the torque exerted on the ball and R_{ball} is the radius of the ball.

Hence the motion of the ball and place systems is formulated by the linear acceleration as given in Eq. 5.

$$\ddot{A}(t) = \frac{M_{ball}G\sin\rho(t){R_{ball}}^2}{M_{ball}R_{ball}^2 + j_{ball}} \tag{5}$$

where the inertia of the ball is given by j_{ball}

2.3 Calculation for servo angle

Here, the behavior of the ball and plate system in terms of dynamic variables: motion and time are represented by the position of the ball about the angle of servo load. This relation between the servo angle and the beam is given by Eq. 6:

$$\sin\rho(t) = \frac{2R_{arm}\sin\beta_l(t)}{L_p} \tag{6}$$

where L_p is plate length and R_{arm} is the distance between the coupled joint and the shaft of the SRV02 output gear.

To find the dynamic variables for representing the movement of the ball (equation of motion) which corresponds to the servo angle θ_l, the dynamic variables are linearized about $\theta_l = 0$. Considering the servo and plate angle relationship, the mathematics are given as:

$$\ddot{A}(t) = \frac{2M_{ball}GR_{arm}{R_{ball}}^2}{L_p(M_{ball}R_{ball}^2 + j_{ball})}\sin\beta_l(t) \tag{7}$$

To linearise the equation of motion, the sine function $sin\beta_l$ is approximated as β_l and the final equation of movement is for 1D ball balancer is given as:

$$\ddot{A}(t) = \frac{2M_{ball}G\beta_l(t)R_{arm}{R_{ball}}^2}{L_p(M_{ball}R_{ball}^2 + j_{ball})} \tag{8}$$

3 Balancing controller design

The controlling model for the x-axis of the SRV02 integrated with the ball balancer system is illustrated in Fig. 4. The ball balancer block diagram explains the controlling in two loops. The first one for the SRV02 motor model and the second one is the 1D ball balancer. The inner loop describes the position control for the servo motor and estimates the needed voltage to track the desired angle of the load.

Control techniques designed and evaluated for the balancing and controlling of the ball balancer system are:

- Proportional Integral Derivative Controller (PID)
- Ant Colony Optimisation-PID (ACO-PID)

3.1 PID control

For the implementation of the Proportional-integral-derivative controller, the inner loop values (servo proportional gain ($K_{PG,\ servo}$)) remains same, whereas, PID has been executed on the external shell to add the location of the pole as far as there is decay in terms of the time constant (T_p). The entire demonstrating for the controller is clarified by the square outline in Fig. 5.

An ideal PID controller produces an output based on the error e between the setpoint and process value, as shown in (9).

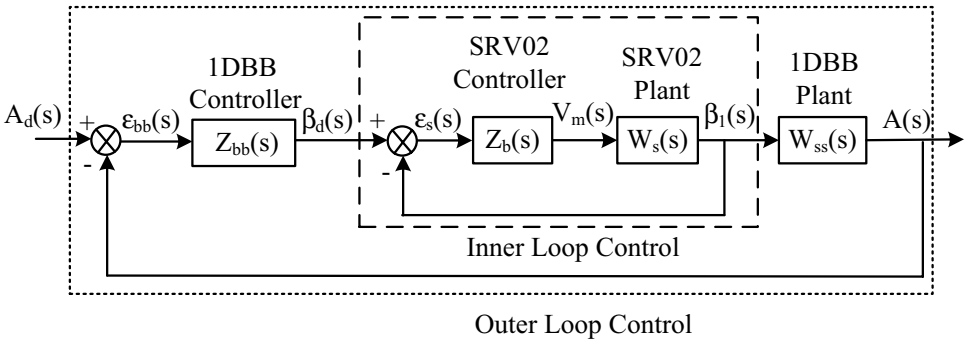


Fig. 4 Full block-based representation for ball balancer setup

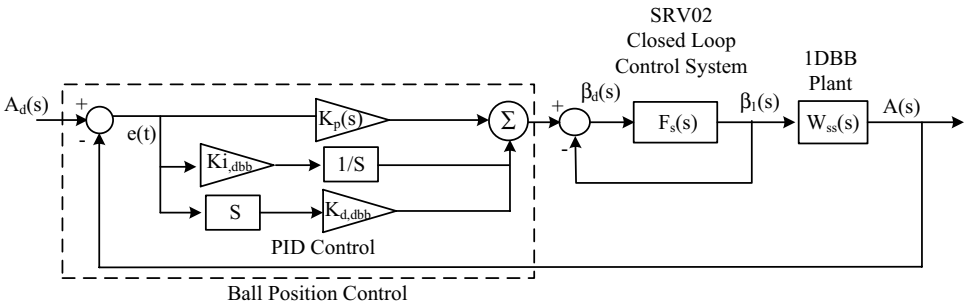


Fig. 5 PID control through block representation

$$e(t) = A_d(s) - A(s) \tag{9}$$

where $A_d(s)$ is the desired or setpoint and $A(s)$ is the process value.

Further, the representation of the PID controller in the time domain is required to calculate the initial operating gains which are given by:

$$\beta_d(t) = K_{p,dbb}\big(A_d(t) - A(t)\big) + K_{d,dbb}\left(\frac{d}{dt}A_d(t) - \frac{d}{dt}A(t)\right) + K_{i,dbb}\int A_d(t) - A(t)dt \tag{10}$$

where $K_{p,dbb}, K_{d,dbb}$, and $K_{i,dbb}$ are proportional, derivative, and integral gains respectively.

To calculate the gain values for achieving the control of ball balancer, the closed-loop equation of the outer loop system (no servo dynamics) is calculated from Fig. 5 or by considering the Laplace of (10) as follows:

$$\beta_d(s) = \left(K_{p,dbb} + \frac{K_{i,dbb}}{s} + K_{d,dbb}s\right)\big(A_d(s) - A(s)\big) \tag{11}$$

As no servo dynamical forces are acting i.e., the estimated and required load point are equal, this gives the closed-loop equation by substituting the outer loop controller (11) with the one-dimensional ball balancer system as follows:

$$\frac{A(s)}{A_d(s)} = \frac{K_{MG}(K_{p,dbb}s + K_{i,dbb} + K_{d,dbb}s^2)}{s^3 + K_{MG}(K_{p,dbb}s + K_{i,dbb} + K_{d,dbb}s^2)} \tag{12}$$

where K_{MG} is the Model constant gain.

Further, from the third-order prototype characteristic equation in (13), the PID gains can be calculated as follows:

$$(s^2 + 2\zeta\omega_n s + \omega_n^2)(s + p_0) \tag{13}$$

where ω_n is the natural frequency of the system, ζ is the damping ratio, and p_0 corresponds to the pole location.

When expanded, the third-order characteristic equation in (13) can be written as:

$$s^3 + \big(2\zeta\omega_n + p_0\big)s^2 + \big(\omega_n^2 + 2\zeta\omega_n p_0\big)s + \omega_n^2 p_0 \tag{14}$$

Concerning (14), the third-order characteristic equation of closed-loop equation in (12) is given by:

$$s^3 + K_{MG}K_{p,dbb}s + K_{MG}K_{i,dbb} + K_{MG}K_{d,dbb}s^2 \tag{15}$$

Equating (14) and (15), the observations made are as follows:

$$K_{MG}K_{i,dbb} = \omega_n^2 p_0 \tag{16}$$

$$K_{MG}K_{p,dbb} = \omega_n^2 + 2\zeta\omega_n p_0 \tag{17}$$

$$K_{MG}K_{d,dbb} = 2\zeta\omega_n + p_0 \tag{18}$$

Further by solving (16)–(18), the PID control gains can be calculated as follows:

$$K_{p,dbb} = \frac{\omega_n(\omega_n + 2\zeta p_0)}{K_{MG}} \tag{19}$$

$$K_{i,dbb} = \frac{\omega_n^2 p_0}{K_{MG}} \tag{20}$$

$$K_{d,dbb} = \frac{2\zeta\omega_n + p_0}{K_{MG}} \tag{21}$$

The Eqs. (16)–(18) aid in calculating the gain values of the PID controller. This process of manually calculating the PID gain values generates a large error, specifically while operating under various parametric and external disturbances. Hence, the need for automatic tuning of PID gains arises, which is achieved using ACO explained in the further sections.

3.2 Ant colony optimization (ACO)

Before the design of ACO for tuning of the PID controller, the selection of objective functions with desired specifications and constraints is a crucial aspect. For a ball balancer system discussed above, its operation is dependent on the control of two rotary servo base units as the arrangement

of the plate is symmetrical to the servo devices. Hence by controlling the plate angle, the position of the ball on the plate can be controlled to follow a predefined path (reference trajectory). Further, the setup involves a vision-based control which tracks the coordinates of the ball on the plate to control the angle of the plate. From this observation, it is identified that by measuring the error between the reference trajectory and captured ball position coordinates, the plate angle can be controlled. Hence the choice of the objective function to optimize the operation of a controller is based on the error between measured and desired trajectories of the ball position. Generally, an error can be termed as an objective function by formulating it as an integral of the square, time, and absolute. All these functions express error as an objective function by evaluating its integral over a fixed period (theoretically up to infinity such that the output achieves settling time). In this research, the error is expressed as an integral time absolute error (ITAE) as it increases the weights on the error of the system response to achieve quick settling time than other error functions. The objective function obtained through ITAE might produce an initially sluggish response which will aid in avoiding sustained oscillations throughout settling time. The expression of error as an objective function with IAE is given as follows:

$$J = \int_0^\infty w|e(t)|dt \tag{22}$$

where w is the weight for the error of the system response and $e(t)$ is the error between the measured and desired position of the ball on the plate of the ball balancer system.

Further, the proposed methodology aims at controlling the ball balancer system with a PID controller where the proportional, integral, and derivative gains are optimized simultaneously. This is achieved through the ant colony optimization method. The generalized expression for the basic principle of ACO is given by:

$$ACO = \{b_i(t), n, edge(i,j), d_{ij}, \tau_{ij}(t), tabu_k\} \tag{23}$$

where $b_i(t)$ is the total number of ants in a particular location i at time t, n corresponds to total locations, m defines the number of ants in all the locations which are mathematically given as follows:

$$m = \sum_{i=1}^{n} b_i(t) \tag{24}$$

Further, $edge(i,j)$ is the path traveled by the ants between two locations i and j, d_{ij} is the distance between the locations, τ_{ij} is the pheromone released by the ants between two locations, and $tabu_k$ corresponds to the details of ant k which include location and path traveled by k at time t, and forbids k in vising those locations again.

The probability P that an ant k will choose a path from the current node (location) i to a potential node j at time t can be calculated per (25). The sum in the denominator is that of the values of all forward nodes available to ant k at time t. The constants α and β can be altered for weighting (Varol and Bingul 2004).

$$P_{ij}^k(t) = \frac{\left[\tau_{ij}(t)\right]^\alpha \left[\eta_{ij}(t)\right]^\beta}{\sum_{k \in allowed_k} \left[\tau_{ij}(t)\right]^\alpha \left[\eta_{ij}(t)\right]^\beta} \tag{25}$$

where $allowed_k = n - tabu_k$ corresponds to the next node that k can choose, α corresponds to the heuristic coefficient of pheromone which indicates the intensity of pheromone, η_{ij} is the visibility, and β corresponds to the expected heuristic coefficient which indicates the importance of visibility. The visibility is defined as:

$$\eta_{ij}(t) = \frac{1}{d_{ij}} \tag{26}$$

As an ant completes one cycle for all the nodes, the corresponding residual information should be updated. Thus, at after time t, the pheromone of the ant for one cycle is updated at $(t + a)$ as follows:

$$\begin{cases} \tau_{ij}(t+a) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}(t+a) \\ \Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}^k(t) \end{cases} \tag{27}$$

$$\Delta\tau_{ij}^k = \begin{cases} \frac{Q}{L_k} & if\ ant\ k\ uses\ edge(i,j)\ between\ time\ t\ and\ t+a \\ 0 & otherwise \end{cases} \tag{28}$$

where $\rho \in (0,1)$ corresponds to pheromone updating coefficient, and $\Delta\tau_{ij}(t)$ corresponds to change in pheromone, which initially is 0 ($\Delta\tau_{ij}(0) = 0$). Further, $\Delta\tau_{ij}^k(t)$ determines the quantity of pheromone released by ant k while traveling from node i to j, Q is constant, and L_k is the length of the path traveled by ant k.

By relating the generalized expression of ACO with PID parameter tuning, few differences regarding optimization goal and pheromone update strategy are observed. The tuning of PID parameters using ACO requires a path irrespective of the closed or shortest path to achieving optimal performance index. Further, while tuning PID, the pheromone is released at each node and the pheromone update is based on the objective strategy but not on the path length.

Based on the above observations, three parameters of PID controller $K_{p,dbb}$, $K_{i,dbb}$, and $K_{d,dbb}$ that needs to be optimized are considered to have five significant digits each.

The arrangement of these parameters in an XY plane to resemble the generalized expression of ACO is shown in Fig. 6. As per the observations regarding the PID controllers in the literature, the 5 significant bits of all the parameters are arranged as follows: two bits of $K_{p,dbb}$ are set before the decimal point and the remaining bits are set after the decimal point, similarly, one each digit of $K_{i,dbb}$, and $K_{d,dbb}$ are set before the decimal point and the remaining bits are set after the decimal point. For any condition, if an ant starts at the origin and travels any node on L_{15}, it completes one cycle. The value od PID parameters for this cycle or path can be calculated using:

$$\begin{cases} K_{p,dbb} = y_{1,j} \times 10^{1} + y_{2,j} \times 10^{0} + y_{3,j} \times 10^{-1} + y_{4,j} \times 10^{-2} + y_{5,j} \times 10^{-3} \\ K_{i,dbb} = y_{6,j} \times 10^{0} + y_{7,j} \times 10^{-1} + y_{8,j} \times 10^{-2} + y_{9,j} \times 10^{-3} + y_{10,j} \times 10^{-4} \\ K_{d,dbb} = y_{11,j} \times 10^{0} + y_{12,j} \times 10^{-1} + y_{13,j} \times 10^{-2} + y_{14,j} \times 10^{-3} + y_{15,j} \times 10^{-4} \end{cases} y \tag{29}$$

where $y_{i,j}$ corresponds to ant position on the line L_i at node j.

During this process, the insufficient ergodicity, long search time, and easy stagnation of solution space are observed. Hence, the ACO process is optimized by adapting some updates.

3.3 Improved ACO

3.3.1 Updating the local pheromone

During the local pheromone update, the pheromone on a particular node is updated constantly resulting in a reduced probability of next ant choosing the same node. Here, the pheromone evaporation coefficient plays a major role, as large pheromone coefficient reduces the global search ability of the algorithm, and a smaller coefficient will improve the global search ability (Zou and Qian 2019). This resulted in the reduced convergence speed of the algorithm. To overcome this, the pheromone coefficient is adjusted dynamically to adjust the pheromone update at different iterations. The pheromone update at each node for different iterations is given by:

$$\tau_{ij} new = (1 - \xi)\tau_{ij} + \xi\lambda(x) \tag{30}$$

where ξ corresponds to the coefficient of pheromone evaporation, function $\lambda(x)$ varies for different iterations which are given as:

$$\lambda(x) = \begin{cases} -0.0001x \in \left[1, \frac{N}{5}\right) \\ 0.0005x \in \left[\frac{N}{5}, \frac{N}{3}\right) \\ 0.001x \in \left[\frac{N}{3}, 4\frac{N}{5}\right) \\ 0.002x \in \left[\frac{4N}{5}, N\right) \end{cases} \tag{31}$$

where x corresponds to iteration number, and N corresponds to the total number of iterations. Initially, negative feedback information is provided to reduce the difference

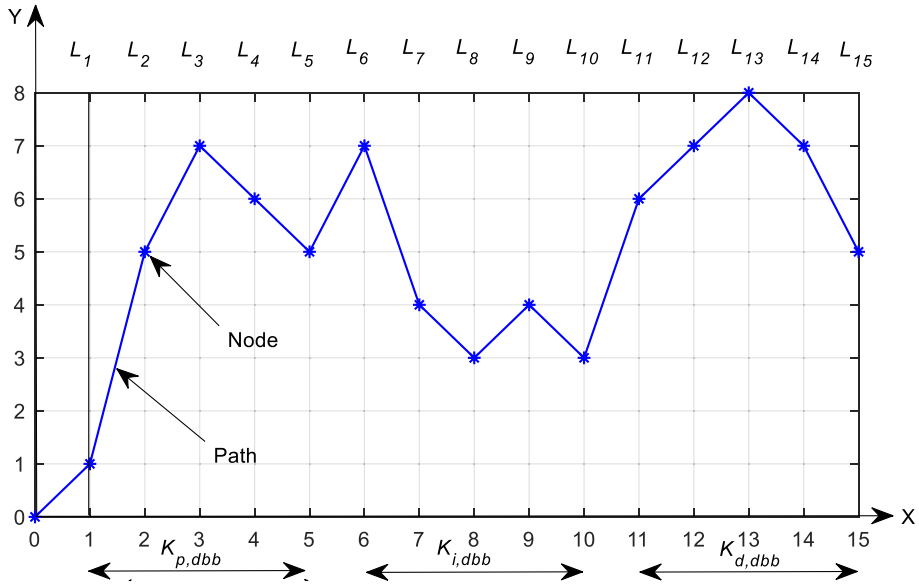


Fig. 6 XY plane resembling node and path for tuning of PID Controller parameters

in pheromone of the local optimal solution and worst-case solution. Further, the value of $\lambda(x)$ is enlarged to ensure search speed and rapid convergence.

3.3.2 Updating global pheromone

During the global pheromone update, the pheromone is updated after the ants complete one cycle. This is based on the objective function J of the PID parameter tuning problem. For any case, if the path searched or traveled by an ant k is better than the previous optimal solution $J^k < J_{best}$, the intensity of pheromone is strengthened, and vice versa. Further, an additional pheromone is added to enhance the optimal solution and to attract the ants in the next cycle. Besides, an iterative approach is adapted to avoid quick convergence of the search process. This dynamically adjusts the pheromone evaporation coefficient during the global pheromone update. This process is mathematically given by:

$$\tau_{ij}(t+a) = \omega(1-\rho)\tau_{ij}(t) + \frac{Q}{J} \tag{32}$$

where ω corresponds to variation coefficient. For $\omega = 2$, $J^k < J_{best}$, and for $\omega = 0.5$, $J^k > J_{best}$.

3.3.3 Bounded optimization range:

Long search time due to random initial position reflects the complexity of the ACO algorithm. This uneven randomization or initial position leads the search to easily achieve local optimal solution resulting in slow convergence. Hence, a bounded range is specified to the parameters to improve the optimization efficiency. This is achieved by calculating the initial parameters of the PID with the help Eqs. (19–21). The calculated parameters are extended over a range to specify the boundary for optimizing the problem.

3.3.4 Adjusting transition probability:

Further, to avoid premature convergence, the transition probability is adjusted. This is achieved by considering that the travel time of all the ants between the node of the line L_u to the node of the line L_{u+1} is equal. In this process, the ants determine the travel path with the help of pheromone and visibility. The difference in the objective function for nodes i and j on the line L_u is given by:

$$\Delta J_{ij} = J_i - J_j \tag{33}$$

According to transition probability $p_{ij}^k(j)$, the ant k travels from node i to node j when $\Delta J_{ij} > 0$, and searches its neighborhood for a better solution when $\Delta J_{ij} < 0$. Hence, the transition probability is given by:

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^{\alpha}[\Delta J_{ij}]^{\beta}}{\sum_{k\in allowed_k}[\tau_{ik}(t)]^{\alpha}[\Delta J_{ik}]^{\beta}} & j \in allowed_k) \\ 0 & otherwise \end{cases} \tag{34}$$

3.4 Improved ACO for PID parameter optimization

The optimal solution is connected with an objective function J during the optimization process. Conventionally, the integral square error, integral absolute error, and integral time absolute error were used for formulating the objective function. But the drawbacks due to slow response and large oscillation time has resulted in the development of new objective functions. Here, the general objective function is added with the square term of the control input to achieve fast and stable operation. The updated objective function J of (22) is given by:

$$J = \int_0^{\infty} (w_1|e(t)| + w_2u^2(t))dt + w_3t_u) \tag{35}$$

where w_1, w_2, and w_3 correspond to the weights, and t_u corresponds to the system rise time.

To avoid overshoot, an overshoot is added with the objective function J, and when $e(t) < 0$, the objective function is given as:

$$J = \int_0^{\infty} \big(w_1|e(t)| + w_2u^2(t)dt + w_4|y(t) - y(t-1)|\big)dt + w_3t_u \tag{36}$$

where w_4 corresponds to weight $\gg w_1$

The steps for achieving parameter tuning of the PID controller are as follows:

Step 1 Initial PID parameters and boundary conditions are calculated using (19–21).

Step 2 Define the number of ants, and a 1-dimensional array of 15 significant for each ant to store the value of nodes it travels. Further, these 15 significant bits are divided between the three parameters of the PID controller with the first 5 digits corresponding to $K_{p,dbb}$, next 5 digits corresponding to $K_{i,dbb}$, and last 5 digits corresponding to $K_{d,dbb}$.

Step 3 Define the initial and maximum cycle numbers, set the pheromones of each node, and determine the path considering an initial position for placing the ants.

Step 4 Increment the cycles.

Step 5 Determine the transition probability according to (34) and select the next node where the ant will arrive.

Step 6 Once, all the ants choose their next node, the pheromone is updated as per (30).

Step 7 As per the path traveled by an ant, the parameters of PID are calculated using (29). Further, the objective function is obtained from (35) and (36), and the corresponding PID parameters and the best objective function are stored.

Step 8 Once all the ants complete one cycle, the global pheromone is updated as per (32).

Step 9 The steps 4–8 are repeated until the intended target is achieved.

The block diagram of the ACO-PID controller is given in Fig. 7.

4 Simulation analysis

The numerical simulation of the 2DoF ball balancer model discussed in Sect. 2 is developed using MATLAB/ Simulink software. Since the arrangement of the plate on the two servo units is symmetrical, the action of the controller on one servo unit has an impact on the action of the controller of the other servo unit. This operates them in a coupled environment irrespective of the aspect that both the controllers are designed in a decoupled environment. Initially, the system is operated to control the position of the ball on a plate in a square trajectory. This is achieved by controlling the ball balancer with improved ACO-PID by providing a square input signal with a frequency of 0.08 Hz and amplitude of 3 as the reference trajectory. Initially, the values of the PID controller are calculated by following the procedure discussed in Sect. 3.1. Further, the improved ACO-PID operates by measuring the error between the desired and measured ball position to optimize the values of the PID controller as per the process discussed in steps 1–9 od Sect. 3.4. The results obtained are then compared with the action of the classical PID controller on the same simulation operating for the same trajectory to assess the performance of the improved ACO-PID controller. The results of the ball position, servo angle, and voltage of the ball balancer system for the cation of improved ACO-PID controller and the classical PID controller are shown in Fig. 8.

Figure 8a shows the results of the position of the ball on the x-axis for both PID and the improved ACO-PID controller. The results identify that the range for the position of the ball on the x-axis is 0.01–5.015 cm. Hence, the lowest difference between the initial and final positions shows the effectiveness of the controller implemented. In this case, the improved ACO-PID has a minimum final position and reaches the desired value in very little time. Further, in Fig. 8b, the servo angle response of the ball on the x-axis shows that variation in the control angle provided by the servo motor. The minimum control angle defines the accuracy of a controller in achieving balancing control for the ball balancer system. In this case, the improved ACO-PID has less control angle when compared with the classical PID control. This will move the plate very slowly while balancing the ball and helps in achieving a stable response for the system. The results in Fig. 8c shows the variation of the servo input voltage with the action of the controller. For the improved ACO-PID control it is observed that the servo units are operated at relatively low voltage than the PID controller and settle

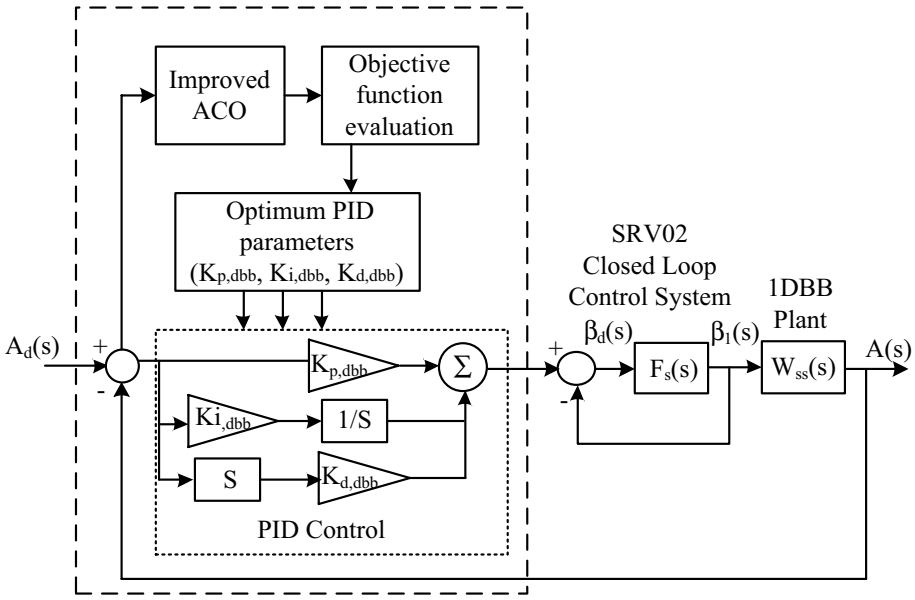


Fig. 7 Block representation of ACO-PID Controller for ball position control

Position Control using improved ACO-PID

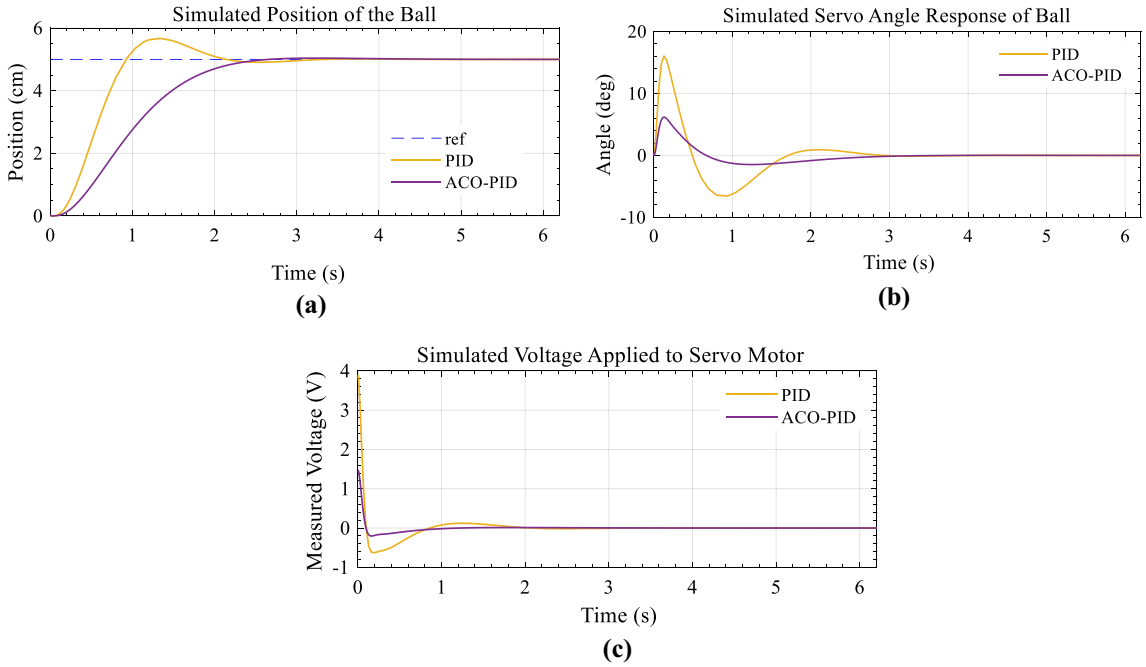


Fig. 8 **a** Position of the ball on the x-axis, **b** servo angle response of ball on the x-axis, **c** input voltage applied to the servo motor for the x-axis

Table 1 Performance parameters for various controllers on ball balancer using Simulink

Controllers	Peak time (t_p) (s)	Settling time (t_s) (s)	Peak overshoot (M_p) (%)	Steady-state error (e_{ss}) (cm)
Proportional integral derivative	1.35	3.5	13.5	0.00472
Improved ACO-PID	2.5	3.01	0.827	0.000648

down at nearly 0.1 V. This is due to the fact that, as the improved ACO-PID achieves minimum position control and balancing angle the speed of the servo motors is minimized to the least possible value such that the ball remains positioned on the plate while following the desired path. Hence, the improved ACO-PID has better performance when compared with the PID controller. Further, to assess the performance of the improved ACO-PID controller, the time domain specifications are calculated and the results are shown in Table 1. From the result, it is observed that the peak overshoot of PID is 13.5%, which causes huge oscillations and made the ball difficult to balance on the plate. On the other side as per the results observed in the graph, it's observed that improved ACO-PID has a fine response to peak overshoot 0.827% and shows the excellent balance of ball on the plate without any oscillation.

Further, the root mean square error (RMSE) value of PID and improved ACO-PID controller actions on the operation of ball balancer for square trajectory are calculated in terms of position and plate angle parameter. The results are tabulated as shown in Table 2.

Table 2 Root mean square error (RMSE) for position and angle during simulation results

Controllers	RMSE	
	Position	Angle
Proportional integral derivative	2.1455	3.3255
Improved ACO-PID	1.7586	0.5931

Moreover, to assess the action of the developed controller with the techniques available in the literature, the error in position and variation in plate angle after the application of the improved ACO-PID controller are compared in Table 3.

Apart from the time domain specifications, the stability of the system with the developed controller is assessed by performing the describing function analysis.

4.1 Describing function method

Describing function technique is expanded adaptations of frequency response strategy and used to investigate,

analyses, and anticipate nonlinear system behavior by creating a limit cycle (Slotine and Li 1991). To represent nonlinear elements by describing function a sinusoidal input of frequency (f) and amplitude (A) is considered whereas output generalized by Fourier series from (37).

$$f(t) = \frac{a_0}{2} \sum_{n=1}^{\infty} a_n \cos(nft) + b_n \sin(nft) \tag{37}$$

The describing function gain is calculated by the fundamental component of the series as its gain relies on amplitude. So, when the oscillation is anticipated for the system, its amplitude is determined and results in the desired sustained oscillation. The describing function of a nonlinear element is a complex ratio of its fundamental component to the sinusoid input, i.e.,

$$N(A,f) = \frac{Me^{i(ft+\varnothing)}}{Ae^{ift}} = \frac{M}{A} e^{i\varnothing} \tag{38}$$

4.2 Defining describing function for ball balancer system

Describing function is calculated by combining MATLAB script and Simulink modelling for dead zone nonlinearity approximation. The output voltage from the actuator dynamics of the ball balancer is fed to the dead zone model for the generation of describing function gain. Actuator dynamics gives the relation between ideal voltage (V_i) and input motor voltage (V_m) after going through the data acquisition board and amplifier limitation. This model is designed by a numerical integration method and implements the describing function definition with the MATLAB script file. Simulink model provides a better understanding of calculation and for stability analysis. Different nonlinearity like saturation, backlash, relay, and quantizer, etc. can describe in place of a dead zone for testing purposes.

4.3 Describing function analysis for PID and improved ACO-PID

This approximation is done to analyze the ball balancer response for PID and the improved ACO-PID controller. The results after running the combination of code and Simulink model to generate describing function are shown in Fig. 9. Figure 9a, b shows the magnitude and phase of the describing function gain vs amplitude and frequency of the input signal of PID controller and Fig. 10a, b for improved ACO-PID controller.

Stability analysis for PID response based on amplitude

(1) The describing function gain in Fig. 9a, starts at zero when the amplitude is at minimum values and gradually increases to stabilize from this point.
(2) We can observe that the zero gain increases as the input amplitude increases. When describing function gain of amplitude varies between zero to 2, for which the system oscillates, then for same time the amplitude of system starts settles.
(3) If the system were to change so that the effective gain decreased then in this region of the frequency response curve, the describing function gain would decrease.
(4) If the ball balancer system gains increase, the amplitude decreases, and the describing function gain increases.

Stability analysis for improved ACO-PID response based on amplitude

(5) From Fig. 10a, it can be observed that when the describing function gain is at zero, the amplitude is at the maximum value. Here in this case the describing function gain settles down for 0.23.
(6) It was observed that the input amplitude slightly increases when the gain of ball balancer increases. The system oscillates between zero to 0.23, where the amplitude of the system increases.

Table 3 Comparative analysis for position error and angle variation

Method applied	Position error	Angle variation
State observer with fuzzy logic controller (Zhang and Yuan 2010)	17cm	15.46deg
PSO based fuzzy-neural (Dong et al. 2011)	0.55cm	–
Feedback Linearization (Ho et al. 2013)	0.9cm	1.66deg
Multiple controllers (KASSEM et al. 2015) a. Sliding Mode controller b. Linear quadratic controller c. Fuzzy logic controller	a. 0.05 cm b. 0.1 cm c.0.1 cm	–
Fractional order sliding mode control (Das and Roy 2017)	0.00026 cm	–
Disturbance observer enhanced PD controller (Pinagapani et al. 2018)	0.0045 cm	–
$H-\infty$ controller (Umar et al. 2019)	0.95 cm	–
Developed improved ACO-PID controller	$1.2e-06$ cm	0.91deg

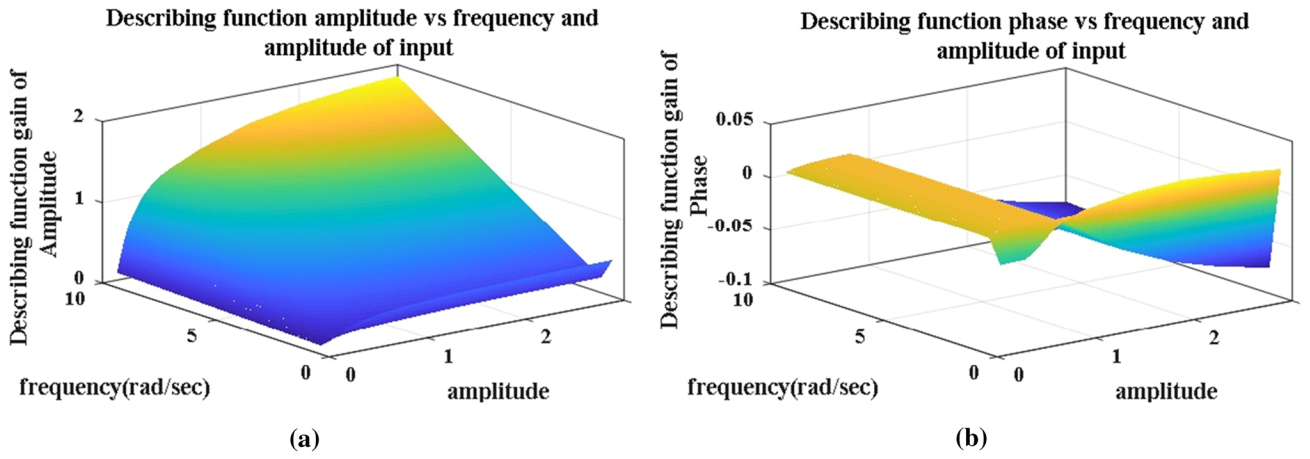


Fig. 9 The magnitude and phase of the describing function plotted as functions of the amplitude and frequency of the input for PID response stability validation

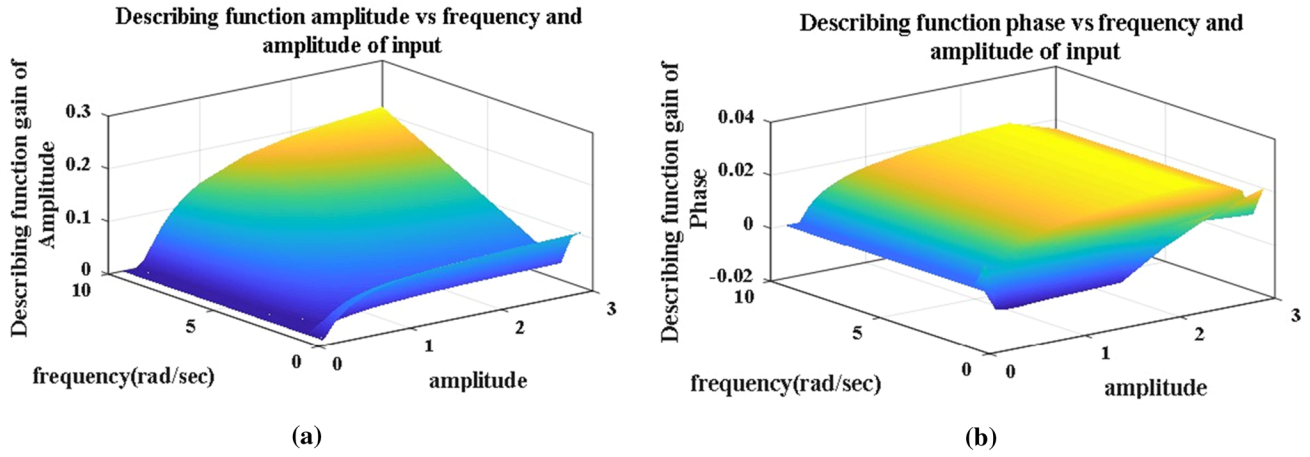


Fig.10 The magnitude and phase of the describing function plotted as functions of the amplitude and frequency of the input for improved ACO-PID response stability validation

(7) If the ball on plate moves slowly or stops almost and causes effective gain decrement, then the frequency response curve would decrease.
(8) If system gain increases, the amplitude increases which causes an increase of describing function gain as well.

Stability analysis for PID and improved ACO-PID response based on phase

It can be observed from the PID phase response, Fig. 9b and improved ACO-PID phase response, Fig. 10b that describing function gain is a function of amplitude and does not dependent on the frequency. From Fig. 9b, the describing function phase plot described as a function of amplitude and frequency with very small values (near to zero) for PID response. On the other side, from Fig. 10b, the describing function phase as a function of amplitude and frequency is observed to be zero, especially for the practical condition. This is due to improved ACO-PID controller design as it only compliances with theoretical expectations.

5 Real-time analysis

The modelling of the 2DoF ball balancer system discussed in Sect. 2 is further realized by a Quanser 2DoF ball Balancer system shown in Fig. 11. This system aids in assessing the real-time operation of the developed controller for a ball balancer system.

The construction of the ball balancer system has a digital camera with a vision computing algorithm, and a square metal plate fixed at the center through a gimble joint to tilt in 2-dimensional direction, X, and Y-axis. A Faulhaber DC

micromotor series 2338 motor (MicroMo Electronics, Inc. 2015) along with a potentiometer and tachometer is used for balancing the system in both directions. The main objective of the system is to balance the ball without falling off the quadrangular metallic plate. Initially, the system is calibrated by manually rotating the load gear of each servo to the position at 0 degrees. Further, the camera is calibrated to capture a 440×440 image with an offset height of 40 and a width parameter of 100. This ensures that the plate and ball of the system are in flycap view (Quanser 2013). Further, the RGB (red, green, and blue) values are measured across the flycap view to identifying the position of the ball. At first, a controlled electric signal related to the inclination of the plate is taken from hardware and then specified to the Data Acquisition device (DAQ) to analyze the real-time motion of the ball with respect to voltage. Once the signal is acquired it is sent to the controller and the signal coordinates captured through the camera are combined with it and the output is moved forward to amplify power by amplifier through the DAQ for further signal alteration which is further processed to the hardware. The experimental framework of the 2DoF ball balancer is depicted by Fig. 12.

The hardware study is initiated with the same parameters as the simulation analysis. Further, the developed improved ACO-PID controller and the classical PID controller are implemented on the setup, and the output responses of each controller are identified. The ball position, plate angle, and operating voltage of the servo mechanism are observed as shown in Fig. 13.

From the results in Fig. 13a it is observed that the improved ACO-PID is provides stabilized operation for the ball balancer system by balance the ball on the plate with almost zero oscillation. The results in Fig. 13b show that the plate stabilization angle of improved ACO-PID controller is maintained between $-10°$ and $12.5°$, whereas the PID controller has a stabilization angle varying between $-30°$ and $35°$. This identifies that the improved ACO-PID has smooth operation even in the presence of external disturbances that affect the movement of the ball during real-time experiments. Further, the voltages in Fig. 13c define the action of improved ACO-PID in stabilizing the plate angle and balancing the ball on the plate. Besides, the time-domain characteristics of the improved ACO-PID controller and the classical PID controller are calculated to assess their operation in the real-time experiment. As of Table 4, the response of improved ACO-PID is better

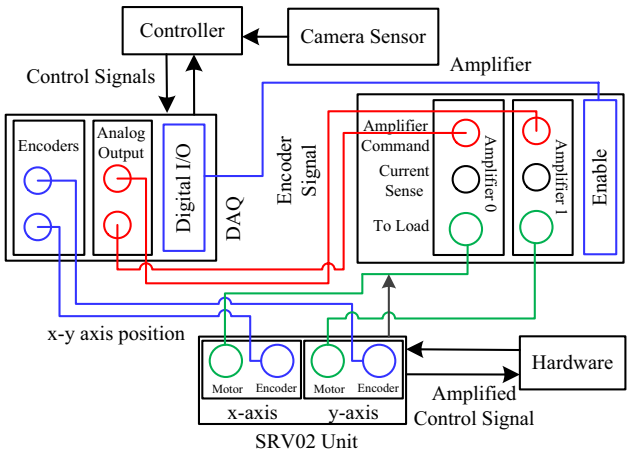
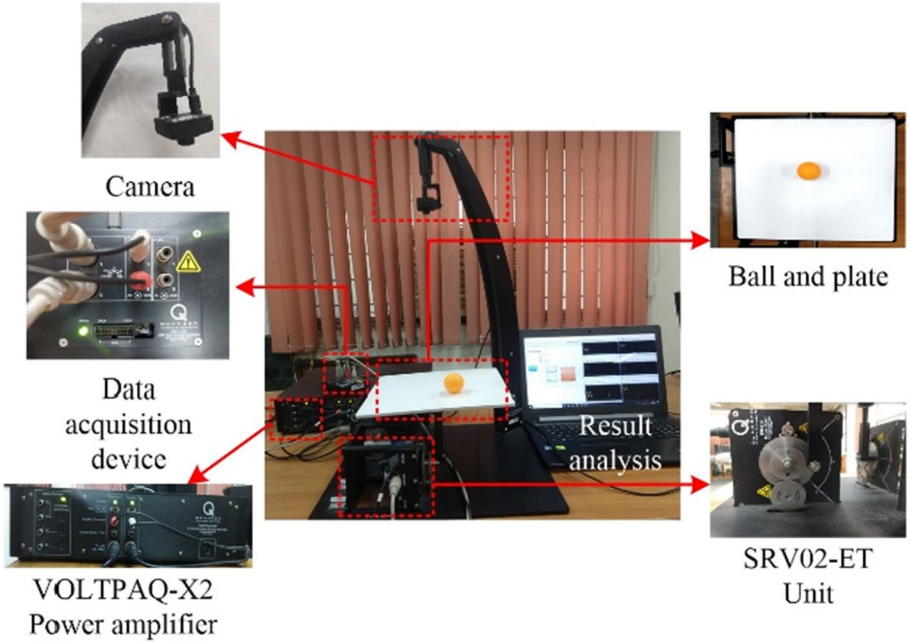


Fig. 12 Framework of ball balancer connection to data acquisition to the amplifier

Fig. 11 Ball balancer setup



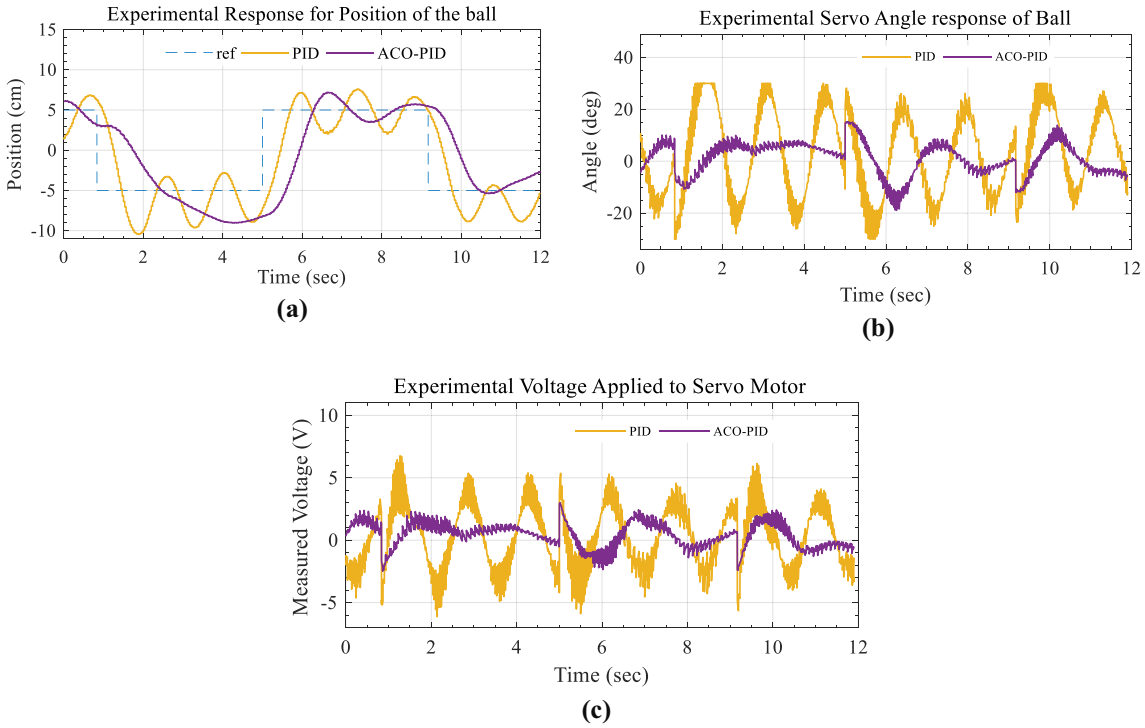


Fig. 13 **a** Real-time position of the ball on the x-axis, **b** real-time servo angle response of ball on the x-axis, **c** input voltage applied to the servo motor for the x-axis

Table 4 Performance parameters for various controllers on ball balancer in real-time

Controllers	Peak time (t_p) (s)	Settling time (t_s) (s)	Peak overshoot (M_p) (%)	Steady-state error (e_{ss}) (cm)
Proportional integral derivative	2.4	2.76	22.9	1.151
Improved ACO-PID	1.62	2.09	12.4	0.487

than the PID controller for Peak time, settling time, and peak overshoot.

Similarly, the root mean square error (RMSE) value for the action of PID and improved ACO-PID controller on the operation of real-time ball balancer setup for square trajectory are calculated in terms of position and plate angle parameter. The results are tabulated as shown in Table 5.

Table 5 Root mean square error (RMSE) for position and angle during simulation results

Controllers	RMSE	
	Position	Angle
Proportional integral derivative	5.1894	3.4049
Improved ACO-PID	3.5680	2.7517

6 Conclusion

In this paper, the self-balancing and position control of two degree of freedom balancer system is achieved by developing an improved ant colony optimized technique to tune the parameters of proportional integral derivative control. The results from simulations and real-time experiments conclude that the developed method has noteworthy performance improvement within the framework of the traditional control structure. The results from the implemented control techniques are validated on the bases of time response analysis. Further, the describing function method generalizes the fact that system response is independent of frequency in case of describing function phase analysis for PID and improved ACO-PID. The presented controller has the adaptability, and satisfactory control performances on the ball balancer system when tested for simulation and real-time experiments. The results identified that the improved ACO-PID shows the best response

in terms of low RMSE, reduced settling time, peak time, steady-state-error, and peak overshoot.

References

Abdo MM, Vali AR, Toloei AR, Arvan MR (2014) Stabilization loop of a two axes gimbal system using self-tuning PID type fuzzy controller. ISA Trans 53:591–602. https://doi.org/10.1016/j.isatra.2013.12.008

Acosta JÁ (2010) Furuta's pendulum: a conservative nonlinear model for theory and practise. Math Probl Eng. https://doi.org/10.1155/2010/742894

Aguilar-Avelar C, Moreno-Valenzuela J (2015) A composite controller for trajectory tracking applied to the Furuta pendulum. ISA Trans 57:286–294. https://doi.org/10.1016/j.isatra.2015.02.009

Aguilar-Avelar C, Moreno-Valenzuela J (2016) New feedback linearization-based control for arm trajectory tracking of the Furuta pendulum. IEEE/ASME Trans Mechatron 21:638–648. https://doi.org/10.1109/TMECH.2015.2485942

Ali HI, Jassim HM, Hasan AF (2019) Optimal nonlinear model reference controller design for ball and plate system. Arab J Sci Eng 44:6757–6768. https://doi.org/10.1007/s13369-018-3616-1

Andreev F, Auckly D, Gosavi S et al (2002) Matching, linear systems, and the ball and beam. Automatica 38:2147–2152. https://doi.org/10.1016/S0005-1098(02)00145-0

Angelov PP (2016) Handbook on computational intelligence. World Scientific, Singapore

Angus D, Hendtlass T (2005) Dynamic ant colony optimisation. Appl Intell 23:33–38. https://doi.org/10.1007/s10489-005-2370-8

Aranda J, Chaos D, Dormido-Canto S et al (2006) Benchmark control problems for a non-linear underactuated hovercraft: a simulation laboratory for control testing. IFAC Proc 7:463–468. https://doi.org/10.3182/20060621-3-ES-2905.00080

Awtar S, Bernard C, Boklund N et al (2002) Mechatronic design of a ball-on-plate balancing system. Mechatronics 12:217–228. https://doi.org/10.1016/S0957-4158(01)00062-9

Bang H, Lee YS (2018) Implementation of a ball and plate control system using sliding mode control. IEEE Access 6:32401–32408. https://doi.org/10.1109/ACCESS.2018.2838544

Bang H, Lee YS (2019) Embedded model predictive control for enhancing tracking performance of a ball-and-plate system. IEEE Access 7:39652–39659. https://doi.org/10.1109/ACCESS.2019.2907111

Basci A, Can K (2017) Position control of a ball beam experimental setup based on sliding mode controller. Int J Appl Math Electron Comput 1:29–35. https://doi.org/10.18100/ijamec.2017SpecialIssue30467

Boubaker O (2012) The inverted pendulum: a fundamental benchmark in control theory and robotics. In: International conference on educational e-learning and innovation

Chalupa P, Přikryl J, Novák J (2015) Modelling of twin rotor MIMO system. In: Proceeding 2015 20th international conference on process control PC 2015, pp 314–319. https://doi.org/https://doi.org/10.1109/PC.2015.7169982

Chang DE (2008) Some results on stabilizability of controlled lagrangian systems by energy shaping. IFAC Proc 41:3161–3166. https://doi.org/10.3182/20080706-5-KR-1001.00537

Chang Y-H, Chang C-W, Tao C-W et al (2012) Fuzzy sliding-mode control for ball and beam system with fuzzy ant colony optimization. Expert Syst Appl 39:3624–3633. https://doi.org/10.1016/j.eswa.2011.09.052

Che G, Liu L, Yu Z (2020) An improved ant colony optimization algorithm based on particle swarm optimization algorithm for path planning of autonomous underwater vehicle. J Ambient Intell Humaniz Comput 11:3349–3354. https://doi.org/10.1007/s12652-019-01531-8

Cheng C-C, Ho C-H (2017) Design of adaptive sliding mode controllers for mismatched perturbed systems with application to underactuated systems. In: 2017 36th Chinese control conference (CCC). IEEE, pp 1329–1336

Chiha I, Liouane N, Borne P (2012) Tuning PID controller using multiobjective ant colony optimization. Appl Comput Intell Soft Comput 2012:1–7. https://doi.org/10.1155/2012/536326

Choukchou-Braham A, Cherki B, Djemaï M, Busawon K (2014) Underactuated mechanical systems from the Lagrangian formalism. Analysis and control of underactuated mechanical systems. Springer International Publishing, Cham, pp 15–33

Da Silveira CR, Flores JV, Salton AT, Pereira LFA (2014) A comparative analysis of repetitive and resonant controllers to a servo-vision ball and plate system. IFAC Proc 19:1120–1125. https://doi.org/10.3182/20140824-6-ZA-1003.01074

Das A, Roy P (2017) Improved performance of cascaded fractional-order SMC over cascaded SMC for position control of a ball and plate system. IETE J Res 63:238–247. https://doi.org/10.1080/03772063.2016.1258336

de Jager B (1998) Control of under actuated systems: a monocycle example. IFAC Proc 31:61–66. https://doi.org/10.1016/s1474-6670(17)40006-1

Debono D, Bugeja M (2015) Application of sliding mode control to the ball and plate problem. In: Proceedings of the 12th international conference on informatics in control, automation and robotics. SCITEPRESS—Science and and Technology Publications, pp 412–419

Dong X, Zhao Y, Xu Y et al (2011) Design of PSO fuzzy neural network control for ball and plate system. Int J Innov Comput Inf Control 7:7091–7103

Dorigo M, Maniezzo V, Colorni A (2008) Toward a complex adaptive intelligence community—central intelligence agency. Cia 26:1–13. https://doi.org/10.1109/3477.484436

Fabregas E, Chacón J, Dormido-Canto S et al (2015) Virtual laboratory of the ball and plate system. IFAC-PapersOnLine 48:152–157. https://doi.org/10.1016/j.ifacol.2015.11.229

Fan X, Zhang N, Teng S (2004) Trajectory planning and tracking of ball and plate system using hierarchical fuzzy control scheme. Fuzzy Sets Syst 144:297–312. https://doi.org/10.1016/S0165-0114(03)00135-0

Feng H, Yin C-B, Weng W et al (2018) Robotic excavator trajectory control using an improved GA based PID controller. Mech Syst Signal Process 105:153–168. https://doi.org/10.1016/j.ymssp.2017.12.014

Hammadih ML, Al HK, Boiko I (2016) Interpolating sliding mode observer for a ball and beam system. Int J Control 89:1879–1889. https://doi.org/10.1080/00207179.2016.1161235

Heinonen J, Pettersson F (2007) Hybrid ant colony optimization and visibility studies applied to a job-shop scheduling problem. Appl Math Comput 187:989–998. https://doi.org/10.1016/j.amc.2006.09.023

Ho M-T, Rizal Y, Chu L-M (2013) Visual servoing tracking control of a ball and plate system: design, implementation and experimental validation. Int J Adv Robot Syst. https://doi.org/10.5772/56525

Hoover RC, St. Amand J (2012) Design of an approximate control law using input-state linearization for the ball on a plate system. vol 4 Dyn Control Uncertainty, Parts A B 203. https://doi.org/https://doi.org/10.1115/IMECE2012-88418

Hou Y-Y (2017) Design and implementation of EP-based PID controller for chaos synchronization of Rikitake circuit systems. ISA Trans 70:260–268. https://doi.org/10.1016/j.isatra.2017.04.016

Hsiao Y-T, Chuang C-L, Chien C-C (2004) Ant colony optimization for designing of PID controllers. In: 2004 IEEE international conference on robotics and automation (IEEE Cat. No.04CH37508). IEEE, pp 321–326

Kao S-T, Ho M-T (2018) Second-order sliding mode control for ball-balancing system. In: 2018 IEEE conference on control technology and applications (CCTA). IEEE, pp 1730–1735

Kassem A, Haddad H, Albitar C (2015) Commparison between different methods of control of ball and plate system with 6DOF stewart platform. IFAC-PapersOnLine 48:47–52. https://doi.org/10.1016/j.ifacol.2015.09.158

Ker CC, Lin CE, Wang RT (2007) Tracking and balance control of ball and plate system. J Chinese Inst Eng Trans Chin Inst Eng A/Chung-kuo K Ch'eng Hsuch K'an 30:459–470. https://doi.org/10.1080/02533839.2007.9671274

Li X (2020) Robot target localization and interactive multi-mode motion trajectory tracking based on adaptive iterative learning. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-01878-3

Masoumi Z, Van Genderen J, Sadeghi Niaraki A (2019) An improved ant colony optimization-based algorithm for user-centric multi-objective path planning for ubiquitous environments. Geocarto Int. https://doi.org/10.1080/10106049.2019.1595176

MicroMo Electronics, Inc. (2015) FAULHABER DC-Micromotors Series 2338, Clearwater-Florida

Ming B, Huiqiu L, Jintao S, Yantao T (2006) Motion control of ball and plate system using supervisory fuzzy controller. Proc World Congr Intell Control Autom 2:8127–8131. https://doi.org/10.1109/WCICA.2006.1713557

Moarref M, Saadat M, Vossoughi G (2008) Mechatronic design and position control of a novel ball and plate system. In: 2008 Mediterranean conference on control automation—conference proceedings, MED'08 1071–1076. https://doi.org/https://doi.org/10.1109/MED.2008.4602212

Mochizuki S, Ichihara H (2013) Generalized Kalman-Yakubovich-Popov lemma based I-PD controller design for ball and plate system. J Appl Math 2013:1–9. https://doi.org/10.1155/2013/854631

Moness M, Mahmoud D, Hussein A (2020) Real-time Mamdani-like fuzzy and fusion-based fuzzy controllers for balancing two-wheeled inverted pendulum. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-01991-3

Moreno-Armendariz ER, CAP-O (2010) Design and implementation of a visual fuzzy control in FPGA for the ball and plate system. In: Proceedings 2010 international conference on reconfigurable computational FPGAs, ReConFig 2010, pp 85–90. https://doi.org/10.1109/ReConFig.2010.69

Moreno-Armendariz MA, Perez-Olvera CA, Floriberto Ortiz Rodrıguez ER (2010) Indirect hierarchical FCMAC control for the ball and plate system. Neurocomputing 73:2454–2463. https://doi.org/10.1016/j.neucom.2010.03.023

Moreno-Valenzuela J, Aguilar-Avelar C (2018) Feedback linearization control of the furuta pendulum. pp 69–92

Murray RM, Astrom KJ, Boyd SP et al (2003) Future directions in control in an information-rich word. IEEE Control Syst Mag 23:20–33. https://doi.org/10.1109/MCS.2003.1188769

Negulescu SC, Kifor CV, Oprean C (2008) Ant colony solving multiple constraints problem: vehicle route allocation. Int J Comput Commun Control 3:366. https://doi.org/10.15837/ijccc.2008.4.2404

Nelles O (2001) Nonlinear system identification. Springer, Berlin

Nguyen DCH, Ascough JC, Maier HR et al (2017) Optimization of irrigation scheduling using ant colony algorithms and an advanced cropping system model. Environ Model Softw 97:32–45. https://doi.org/10.1016/j.envsoft.2017.07.002

Ning J, Zhang C, Sun P, Feng Y (2018) Comparative study of ant colony algorithms for multi-objective optimization. Information 10:11. https://doi.org/10.3390/info10010011

Nowopolski K (2013) Ball-and-beam laboratory system controlled by Simulink model through dedicated microcontrolled-Matlab data exchange protocol. Comput Appl Electr Eng 11:310–320

Olivares M, Albertos P (2013) On the linear control of underactuated systems: the flywheel inverted pendulum. In: 2013 10th IEEE international conference on control and automation (ICCA). IEEE, pp 27–32

Ortega R, Spong MW, Gomez-Estern F, Blankenstein G (2002) Stabilization of a class of underactuated mechanical systems via interconnection and damping assignment. IEEE Trans Automat Contr 47:1218–1233. https://doi.org/10.1109/TAC.2002.800770

Pattanapong Y (2013) Deelertpaiboon C (2013) Ball and plate position control based on fuzzy logic with adaptive integral control action. IEEE Int Conf Mechatron Autom IEEE ICMA 2013:1513–1517. https://doi.org/10.1109/ICMA.2013.6618138

Pinagapani AK, Mani G, Chandran KR, Pandian K (2018) Composite disturbance rejection control for ball balancer system. Proc Comput Sci 133:124–133. https://doi.org/10.1016/j.procs.2018.07.016

Quanser (2013) 2D Ball balancer user manual. 1.3:23

Rad AB, Chan PT, Lo WL, Mok CK (2003) An online learning fuzzy controller. IEEE Trans Ind Electron 50:1016–1021. https://doi.org/10.1109/TIE.2003.817690

Roy P, Das A, Roy BK (2016) Cascaded fractional order sliding mode control for trajectory control of a ball and plate system. Trans Inst Meas Control. https://doi.org/10.1177/0142331216663826

Rudra S, Barai RK, Maitra M (2017) Block backstepping control of the underactuated mechanical systems. Block backstepping design of nonlinear state feedback control law for underactuated mechanical systems. Springer, Singapore, pp 31–52

Shiriaev A, Ludvigsen H, Egeland O, Pogromsky A (1999) On global properties of passivity based control of the inverted pendulum. In: Proceedings of the 38th IEEE conference on decision and control (Cat. No.99CH36304). IEEE, pp 2513–2518

Slotine J-J, Li W (1991) Applied NoLinear Optimal Control

Spong MW (1994) Partial feedback linearization of underactuated mechanical systems. In: Proceedings of IEEE/RSJ international conference on intelligent robots and systems (IROS'94). IEEE, pp 314–321

Sun S, Li L (2012) The study of ball and plate system based on non-linear PID. Appl Mech Mater 187:134–137. https://doi.org/10.4028/www.scientific.net/AMM.187.134

Sutharsan S, Kirubarajan T, Lang T, Mcdonald M (2012) An optimization-based parallel particle filter for multitarget tracking. IEEE Trans Aerosp Electron Syst 48:1601–1618. https://doi.org/10.1109/TAES.2012.6178081

Tian Y, Bai M, Su J (2006) A non-linear switching controller for ball and plate system. Int J Model Identif Control. https://doi.org/10.1504/IJMIC.2006.011940

Tsai C (2004) A new hybrid heuristic approach for solving large traveling salesman problem*1. Inf Sci (Ny) 166:67–81. https://doi.org/10.1016/j.ins.2003.11.008

Tseng L-Y, Liang S-C (2006) A hybrid metaheuristic for the quadratic assignment problem. Comput Optim Appl 34:85–113. https://doi.org/10.1007/s10589-005-3069-9

Umar A, Haruna Z, Musa U et al (2019) Graphical user interface (GUI) for position and trajectory tracking control of the ball and plate system using H-infinity controller. Covenant J Inform Commun Technol 7:35–56. https://doi.org/10.20370/yhas-n460

Varol HA, Bingul Z (2004) A new PID tuning technique using ant algorithm. In: Proceedings of the 2004 American control conference, vol 3. IEEE, pp 2154–2159

Wang Y, Sun M, Wang Z et al (2014) A novel disturbance-observer based friction compensation scheme for ball and plate system. ISA Trans 53:671–678. https://doi.org/10.1016/j.isatra.2013.11.011

Wang Y, Jin Q, Zhang R (2017) Improved fuzzy PID controller design using predictive functional control structure. ISA Trans 71:354–363. https://doi.org/10.1016/j.isatra.2017.09.005

Wang J, Zhu Y, Qi R et al (2020) Adaptive PID control of multi-DOF industrial robot based on neural network. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-01693-w

Ye T, Luo Z, Wang G (2020) Adaptive sliding mode control of robot based on fuzzy neural network. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-020-01809-2

Zhang Z, Yuan D (2010) Modelling and control scheme of the ball–plate trajectory-tracking pneumatic system with a touch screen and a rotary cylinder. IET Control Theory Appl 4:573–589. https://doi.org/10.1049/iet-cta.2008.0540

Zhang J, Zhuang J, Du H, Wang S (2009) Self-organizing genetic algorithm based tuning of PID controllers. Inf Sci (Ny) 179:1007–1018. https://doi.org/10.1016/j.ins.2008.11.038

Zou Z, Qian Y (2019) Wireless sensor network routing method based on improved ant colony algorithm. J Ambient Intell Humaniz Comput 10:991–998. https://doi.org/10.1007/s12652-018-0751-1

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Improvement of Throughput using Dual Connectivity in Non-Standalone 5G NR Networks

Sakshi Dubey, Jasraj Meena
Department of Information Technology
Delhi Technological University
Delhi, India
sakshidubey515@gmail.com, jasrajmeena@dtu.ac.in

***Abstract*— Wireless communications are constantly evolving to facilitate better access to information for the improvement of business, education, and technology around the world. Hence to overcome the lack of spectrum new bands like millimeter wave (mmWave) bands are needed to be utilized. The mmWave bands offer a huge spectrum due to which substantial attention is being given to exploit this frequency range for deploying 5G cellular systems. 5G wireless technology is expected to deliver multi-Gbps data rates, vast network capacity, low latency, and more reliability. Despite the provision of the spectrum, the new mmWave bands are prone to blockages and obstacles. Hence, these bands must be deployed with the help of an architecture that can overcome this drawback and can prevent any disruption of service in case of blockage. This paper implements a non-standalone (NSA) architecture using modified dual connectivity of user equipment (UE) to Long Term Evolution (LTE) module and mmWave module of ns-3 for more robust deployment of 5G New Radio (NR) network. It also includes a detailed discussion upon the design and functioning of 5G NR in NSA mode with the help of a simulation setup. Finally, the performance of the 5G network in standalone (SA) and non-standalone modes are compared both graphically and numerically based on the throughput measurements.**

Keywords—5G; cellular; dual connectivity; millimeter-wave; new radio; non-standalone; ns-3; standalone

I. INTRODUCTION

The constant evolution of wireless communication and shortage of spectrum in the existing wireless services has led to the study of millimeter-wave (mmWave) bands as a frequency range that can provide a huge spectrum and multi-Gbps wireless services. Because of this ability, mmWave bands have gained a significant role in the deployment of Fifth Generation New Radio networks or 5G NR networks [1]. Despite this potential, the susceptibility of mmWave bands towards obstacles and weak radio propagation at high frequency makes these new bands require several layer-wise improvements and advancements in the protocol stack to get deployed adequately. Some of these advancements have already been implemented in the previous studies like modified channel models and adaptive antenna in the physical layer [2], [3]. And these advancements are implemented and tested with the help of ns-3. The network simulator-3 or ns-3 [4] already supports various network protocols and implements modules for simulating networks like 3GPP-LTE [5]. It also has a millimeter-wave (mmWave) module first implemented in [6]. So, this work aims to utilize this module and develop a non-standalone architecture for 5G networks.

To overcome the challenges faced by a typical standalone mmWave system and improve the overall throughput achieved, this paper suggests the design of non-standalone architecture (NSA) for 5G cellular networks. Hence, the existing mmWave and LTE modules of ns-3 are integrated with the help of the dual connectivity technique introduced in [7] with some modifications to implement an NSA architecture for the 5G NR network. So, this work is arranged as follows. Section II of this paper presents the evolution of the mmWave module by describing the related work done in previous years for the advancement of the mmWave module. Table I in this section discusses and compares the previous studies that led to the expansion of the existing module and enhancements presented by each of these studies in tabular form. Section III imparts information about the design and functioning of the proposed non-standalone framework at the Radio Resource Control (RRC) layer. Section IV evaluates the performance of the proposed architecture and compares it to the typical standalone architecture. Section V concludes the work done and highlights future work.

II. EVOLUTION OF MMWAVE MODULE

This module is inspired by the LTE-LENA module [5] of ns-3 and designed similarly. It has a customizable design and is entirely developed in C++. The functioning and the architecture of the module can be understood through its UML diagram provided in [6]. Further customizations are made in this module according to the requirements of the studies. As dual connectivity is implemented to avoid latency caused by the hard handover and to handle the issue of the outage of mmWave base station due to user mobility. Here, LTE base stations are represented as evolved Node Base or eNBs and NR base stations as mmWave eNBs or next-generation node base or gNBs. Also, to eliminate the drawbacks like the sensitivity of mmWave band towards obstacles and to stabilize the channel quality some of the improvements were implemented as multi-connectivity of user equipment (UE) to more than one cell and fast handover methods between cells [8], [9]. A

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

detailed description of evolution and enhancements made in the module

TABLE I. COMPARISON OF RELATED WORKS FORMULATING THE DEVELOPMENT OF MMWAVE MODULE IN NS-3

	Proposed Model	Novel Aspects	Configurational Enhancements	Results	Advantages
[6]	The first mmWave module for the ns-3 simulator is developed	The module comprises of 1) PHY layer, 2) MAC layer and 3) propagation and channel models	PHY consists of 1) TDD frame structure, 2) Radio characterization with channel variation supporting MIMO techniques, 3) Error model at receiver, 4) Interference model. MAC transfers the upper layer data to the PHY layer. Also conveys resource allocation & scheduling decisions	A new & customizable simulation module for mmWave 5G cellular networks is introduced consisting of fundamental implementation of PHY, MAC & mmWave devices	1) Spectrum scarcity can be eliminated by introducing new bands, 2) Researchers can flexibly analyze various setups using the different configuration of the simulation
[14]	A simulation module to evaluate end-to-end connectivity is developed for 5G mmWave cellular networks in ns-3	The module consists of MAC and PHY-layer configurations which can be interfaced with the core network module of the ns-3 LTE model with the help of SAPs and higher layer protocols for end-to-end simulation	PHY layer implements 1) TDD frame which allows flexible allocation of data and control channels for variable Transmission Time Interval (TTI), 2) Path loss model is merged with building obstacle model for realistic hindering MAC layer now contains 1) modified AMC, 2) two scheduler classes, 3) multi-process stop-and-wait HARQ	A customizable simulation module is introduced for the provision of evaluation of new 5G protocols custom mmWave MAC/PHY layer protocols can also be analyzed over the implementations of mmWave eNB, mmWave stack & channels	1) Millimeter-wave provides a huge amount of spectrum, low latency links, 2) the developed module provides an easy evaluation of new 5G protocols
[15]	3GPP channel model is presented for 6-100 GHz band for end-to-end 5G mmWave framework of ns-3	1) 3GPP channel model for ns-3 mmWave framework is introduced that can be further utilized for simulating NR technologies, and 2) MIMO beamforming architecture is introduced	Propagation Loss Model is extended to implement the 3GPP guidelines for calculating path loss and an alternative model is provided which uses the ns-3 Building model to position buildings in the simulation in a random or deterministic manner MIMO based beamforming calculation is implemented	An enhanced and improved channel model is introduced to accurately access and analyze 5G NR technology	1)The channel model covers broader frequency range, 2) Eliminates channel inconsistency due to random obstacles and mobility
[7]	The model consists of a dual connectivity framework and improved mechanisms for handover and path switching	1) Implementation of dual connectivity for better mobility management & faster handover process, 2) evaluation of the handover process for mmWave systems	In the dual connectivity (DC) architecture UE performs secondary cell handover (SCH) rather than performing hard handover (HH). The upper layers of mmWave & LTE frameworks are extended from the same layers of the ns-3 LTE module. PHY & MAC layers of mmWave module and LTE classes are modified for the proper functioning of framework and PDCP is the integration layer	The dual connectivity model performs better in terms of latency, packet loss, and throughput.	1) Improved end-to-end network performance, 2) reduction in latency due to fast switching with the help of SCH, and 3) multi-connectivity framework enhances the quality of service

are formulated in Table I. It also compares some previous works based on the parameters like their novel aspects, modifications made in the existing module, results of these modifications, and their benefits.

III. FRAMEWORK ILLUSTRATION FOR NON-STANDALONE MODE

The deployment of 5G NR network in non-standalone (NSA) mode refers to the system where 5G networks rely on the control plane of the existing LTE module for control functions whereas the 5G networks indulge in the user plane tasks only. In this work, the functioning of the NSA mode is designed with the help of dual connectivity technology. The dual connectivity (DC) mechanism is implemented in [7]. The DC protocol is implemented by extending the NetDevice module of ns-3 and developing the McUeNetDevice function.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

A layer-wise configuration of the DC protocol with PDCP (Packet Data Convergence Protocol) layer as the integration point of dual connectivity architecture can be studied in [10].

As required by this framework UE can maintain a simultaneous physical layer link with LTE and 5G networks

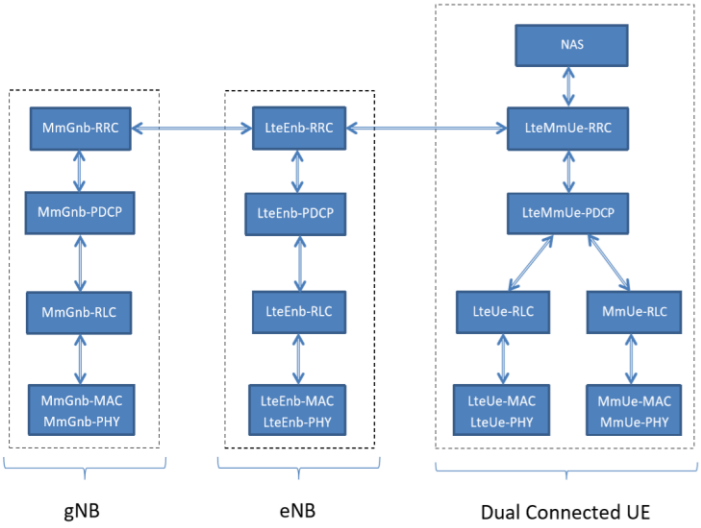


Fig. 1. Protocol stack for modified DC

with the help of the DC mechanism. So a modified DC mechanism is used in the proposed architecture that is inspired by the pre-existing mechanism of dual connected UE. The protocol stack representation of modified DC is shown in Fig. 1. Here, the dual connected UE is accommodating the protocol stacks of both the radio access technologies (RATs) with PDCP as the integration layer. The UE has established an enduring RRC connection with the LTE eNB and has appointed it as coordinator for selecting a suitable gNB and then initiating a connection with it. This helps LTE eNB to handle the control plane functions for gNB while gNB exclusively operating the user plane tasks and thus creating a non-standalone design for the execution of 5G NR networks. The eNB also functions as a selector and selects the best gNB in its vicinity with potential signal-to-interference-plus-noise-ratio or SINR. The process of NR addition and removal is executed at the Radio Resource Control (RRC) layer as shown in Fig. 2 and explained as follows.

A. *Attach or RRC connect*

Since the Radio Resource Control (RRC) layer of LTE module controls the functioning of the user plane of LTE as well as mmWave module so the addition and removal of gNB are initiated by this layer. RRC connection establishment takes place when UE sends *attach request* to the eNB. The *attach request* signals that UE is ready for the dual connectivity mechanism. And the Evolved Packet Core (EPC) checks if UE is authorized to connect to LTE eNB and mmWave gNB.

B. *Detach or RRC idle*

The state before the RRC connection establishment is *RRC idle* state. As the name signifies in this state RRC layer is idle and eNB does not serve any connection to any UE. When connection establishment encounters any failure or the purpose is accomplished the eNB releases the connection and returns to the *RRC idle* state.

C. *gNB Selection*

The selection of a potential gNB based on the SINR is termed as *gNB selection*. The LTE module is the centralized coordination system that performs all the necessary tasks including *gNB selection*, *gNB switching*, and *gNB config*. The procedure for *gNB selection* is stated below:

- The coordinator eNB transmits the sound referencing signals (SRS) in all the directions sweeping the angular space around it.
- These signals are decoded by C-RNTI (Cell Radio Network Temporary Identifier) [11] and then used by gNBs to estimate the channel.
- The gNBs collect all these signals by scanning through their directions.
- Now each gNB fills a record table mentioning its best SINR value for every UE that it can get by selecting the best direction to transmit.
- The coordinator eNB collects all these tables that it received via the X2 link to the gNBs and merges them into one table.
- After going through the entire table it selects the best gNB candidate based on the SINR of each entry.
- Hence, a connection to the most eligible gNB candidate of the table can be commenced.
- If the table contains K entries, for gNB_i, $i = 1,\ldots,K$, then the selected gNB should have maximum SINR, i.e.,

$$\max(SINR_i,\ldots\ldots,SINR_k) \quad (1)$$

D. *gNB Switching*

The switching between gNBs takes place if the SINR of the serving gNB degrades due to user mobility or outage, i.e., *gNB config failure* occurs. gNB switching is initiated according to the following guidelines. Assuming there exists three potential gNBs in the vicinity of the UE-eNB pair, i.e., gNB_i, gNB_j, gNB_k with $SINR_i$, $SINR_j$, $SINR_k$, then:

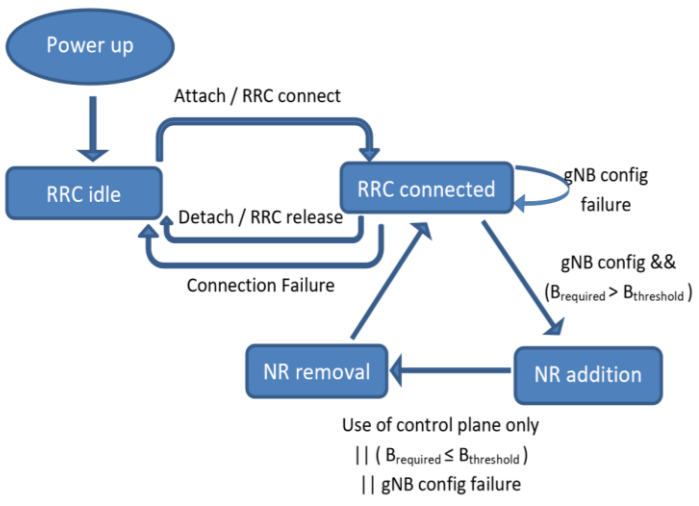


Fig. 2. Process of NR adition and removal at RRC layer

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

- If after a certain time interval the SINR of the selected gNB degrades, suppose gNB_i is the selected gNB and $SINR_i$ becomes less than the SINR of other gNBs which are not in the outage at that timestamp, i.e.,

$$\begin{aligned}&((SINR_i < SINR_j) \wedge (SINR_j > Outage_{th}))\\ &\vee ((SINR_i < SINR_k) \wedge (SINR_k > Outage_{th}))\end{aligned} \quad (2)$$

- Here $Outage_{th}$ is the SINR value which is set as a threshold to check whether the certain gNB has suffered an outage or is still in the vicinity. So if the serving gNB is still in the vicinity of the UE-eNB pair then its SINR should be greater than the $Outage_{th}$, i.e.,

$$SINR_i > Outage_{th} \quad (3)$$

- And if the SINR degrades and becomes lesser than the set threshold due to which *gNB config failure* occurs then a switch is performed to next potential gNB, i.e.,

$$SINR_i \leq Outage_{th} \quad (4)$$

E. gNB Config

This entire procedure of switching, selection, and maintenance of connection with the most suitable gNB candidate is termed as *gNB config*. The configuration of gNB covers the following cases:

- The first case only involves *gNB selection*. This is the case when the RRC connection is just established between the UE and the eNB.
- The second case involves *gNB switching* due to *gNB config failure* followed by *gNB selection*. In this case, the serving gNB is not capable to serve anymore.

F. NR Addition

When eNB enters *RRC connected* state it decides to set up a connection with the appropriate gNB which is termed as *NR addition*. It is needed when any one of the two situations **is** satisfied: 1) if user-plane functions or transmission of data packets is required or 2) if the bandwidth requirement exceeds the set threshold, i.e.,

$$B_{required} > B_{threshold} \quad (5)$$

Here, $B_{required}$ is the bandwidth required by UE and $B_{threshold}$ is the threshold bandwidth based on which *Enb-RRC* decides whether the addition of NR is needed or not.

G. NR Removal

NR removal takes place if any of these conditions hold: 1) when the purpose of NR addition is fulfilled and user plane functions are no longer required, 2) SINR of the connected gNB drops and switching to another gNB with better SINR is required, i.e., *gNB config failure* has occurred due to user mobility or link quality degradation, or 3) Bandwidth requirements can be accomplished with the help of LTE, i.e.,

$$B_{required} \leq B_{threshold} \quad (6)$$

After this state eNB returns to RRC connected state as it still serves the connection to UE but the connection to gNB is terminated.

IV. Performance Evaluation

The primary motive of this implementation is to evaluate

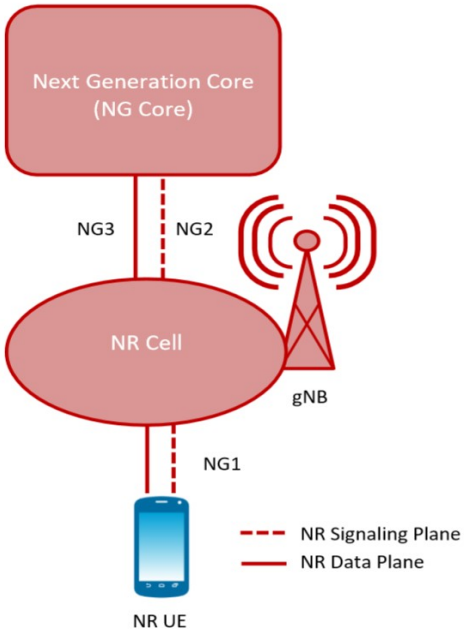


Fig. 3. Simulation model of 5G NR in SA mode

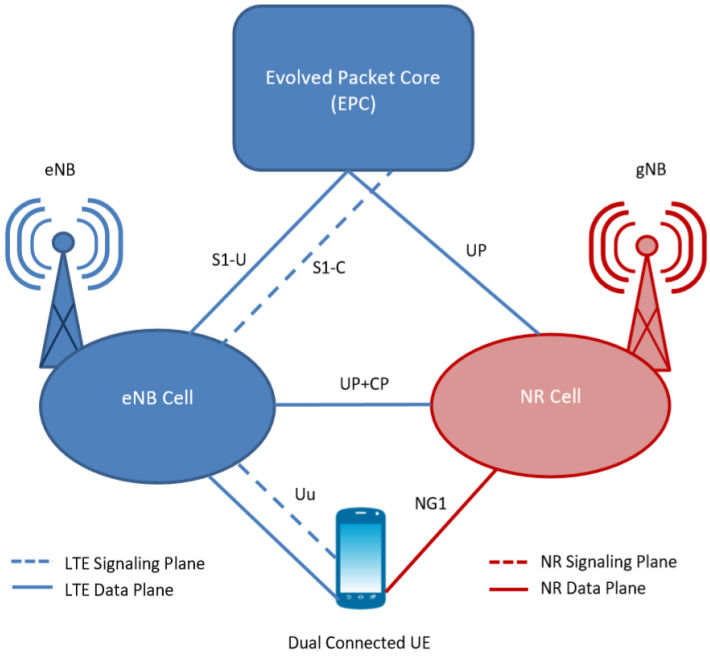


Fig. 4. Simulation model of 5G NR in NSA mode

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

the performance of NSA architecture for 5G NR networks

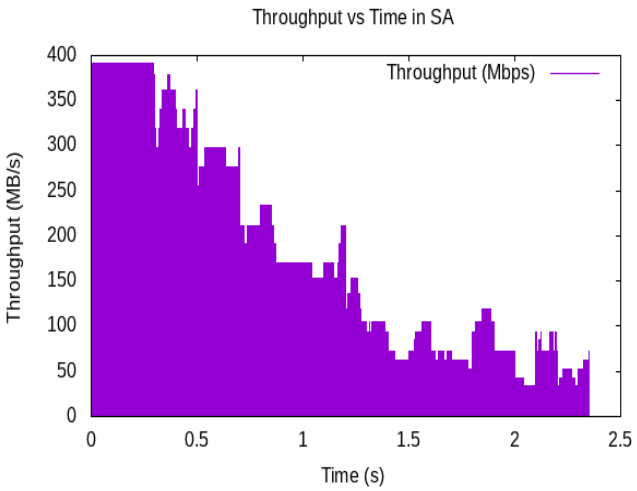


Fig. 5. A graph representimg the gradual diminution in throughput with respect to time for SA architecture of 5G networks

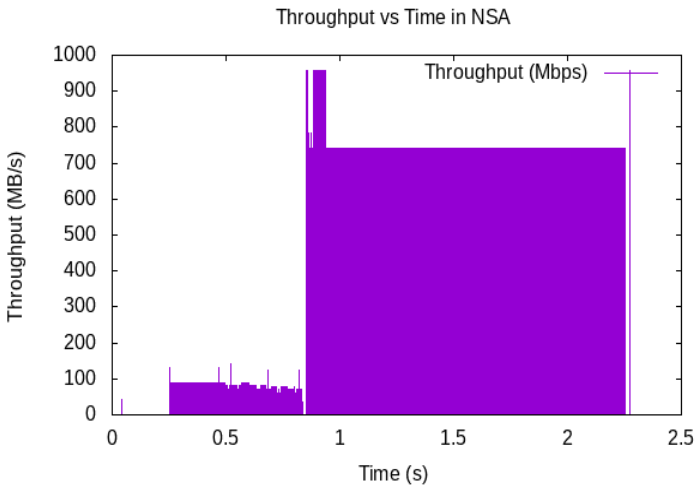


Fig. 6. A graph representing a constant curve for throughput with respect to time for NSA architecture

which relies on the LTE EPC to perform control plane tasks. This section simulates two scenarios one is SA and the other is the NSA 5G network. The simulation aims to compare the performance of the two architectures.

A. *StandaloneMode*

The SA architecture contains a gNB that serves the direct connection to a UE to deliver 5G services as shown in Fig. 3. Here, the gNB performs user (or data) plane as well as control (or signaling) plane functions. UE and gNB are assigned with the *Constant Position Mobility Model* [12] to simulate user mobility. The distance between the gNB and UE is updated with time. The parameters for simulation are provided in Table II. The distance of UE is updated after every 100ms. Since the UE constantly moves away from the connected gNB, the overall throughput decreases with time as shown in Fig. 5. And the SINR also reduces gradually as the distance between UE and gNB increases. The channel state assigned to the simulation is non-line-of-sight (NLOS) which is a type of radio transmission required in partially obstructed propagation

TABLE II. SIMULATION PARAMETERS FOR SA ARCHITECTURE

Parameters	Values	Description
distUpdateInterval	100.0 ms	Distance updation time
distInc	10 metres	Increment in UE-eNB distance
numGnb	1	No. of gNBs
numUe	1	No. of user equipments
distMin	15.0 metres	Minimum eNB-UE distance
distMax	250.0 metres	Maximum eNB-UE distance
channelState	NLOS	Non-Line-of-Sight path

conditions. The simulation of NLOS and LOS can be reviewed in the adaptive and advanced beamforming section of [13].

B. *Non-StandaloneMode*

The simulation scenario of NSA architecture is deployed with one LTE eNB, two gNBs in its vicinity, and a UE. The UE seeks an enduring connection with the eNB to access 5G services from the best gNB in its vicinity with the help of dual connectivity between LTE and mmWave module as shown in Fig. 4. Hence, the eNB acts as a coordinator and selects the data best gNB between the two to access 5G data services. Here the LTE cell performs all the signaling functions to manage the user plane of 5G NR cell. The switching judgment between the gNBs depends upon two factors: 1) the superiority of SINR among the gNBs and 2) if the serving gNB suffers an outage. The superiority of SINR decides which gNB should be added for the provision of better data transfer rates. And threshold bandwidth determines when to initiate dual connectivity between LTE and 5G NR to proceed to the NR addition state from the RRC connected state as presented in Fig. 2. The eNB manages the state of the user plane of the linked gNB and both of them communicate to set up the radio

TABLE III. SIMULATION PARAMETERS FOR NSA ARCHITECTURE

Parameters	Values	Description
gnbDist	200 metres	Distance between gNB 1 and 2
x2Latency	50	Latency on X2 interface (µs)
mmeLatency	100	Latency on MME interface (µs)
ueSpeed	10 m/s	Speed of UE
ueInitialPosition	80	Starting position of UE
ueFinalPosition	180	Final position of UE
symPerSf	24	Symbols per subframe
gnb1PosVector	(50,70,3)	Position of first gnb
gnb2PosVector	(150,70,3)	Position of second gnb
outageThreshold	-5	Outage threshold for gNBs
$B_{threshold}$	80 Mb/s	Threshold bandwidth

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

bearer and the UE is reported about the bearer by the *RRC Reconfiguration* message. The simulation parameters are stated in Table III. For simulating user mobility *Constant Velocity Mobility Model* of ns-3 is assigned to the simulation.

C. Results

The plots in Fig. 5 and Fig. 6 validate the performance improvement achieved by the proposed architecture. The plot in Fig. 5 depicts gradual decrement in throughput as UE travels away from the serving gNB. This happens due to the lack of any supporting connection to UE when the serving gNB suffers an outage. While the plot in Fig. 6 demonstrates a constant curve for throughput after a certain interval of time. This is achieved due to the enduring connection of UE with the LTE module of ns-3 when gNB suffers an outage, preventing any disruption of data service. It also performs gNB switching for better data transfer rates. For numerical assessment of the performances of the two architectures the rate at which data is transferred at the PHY layer is recorded. On simulating the two scenarios over ns-3 two trace files are generated depicting the traces of bytes transferred at the PHY layer with respect to the time in seconds. This helps in the average throughput calculation of the two simulation scenarios. The average throughput of NSA architecture is 512.6 Mbps while that of the SA architecture is 180 Mbps. And the peak data rates achieved can be approximated as 400 Mbps in SA whereas 1 Gbps in NSA. So, this demonstration illustrates the better performance of the proposed NSA architecture with the modification of the DC protocol.

V. Conclusion and Future Work

The NSA architecture presented in this paper is implemented with the help of the DC mechanism and mmWave module of ns-3. The proposed architecture has appropriately used mmWave technology to produce improved outputs in terms of data rates. It overcomes the highly susceptible nature of mmWave technology by taking leverage of DC protocol and customizing it accordingly. This architecture provides the dual connectivity of UE to LTE and 5G NR to prevent any disruption of connectivity as the UE is permanently connected to LTE eNB. The eNB also acts as the selector of the best gNB for better data rates. Thus, the NSA architecture is more efficient as compared to the typical SA architecture. The deployment of 5G NR in the real world is simpler in NSA mode as compared to the SA mode because in NSA mode it is developed with the support of fully established LTE architecture whereas a lot of effort is still needed for the provision of the end to end support for the 5G SA network.

Though the NSA architecture is a better alternative right now in comparison to partially developed SA architecture of 5G NR, a lot of improvement is needed to develop a better mobility module that accurately simulates the real-life user mobility scenario. And the development of end-to-end support for standalone mode of 5G NR network that remains as an upcoming interest.

References

[1] C. A. Professor, "COMPREHENSIVE SURVEY OF WIRELESS COGNITIVE AND 5G NETWORKS," *J. Ubiquitous Comput. Commun. Technol. (UCCT*, 2019, doi: 10.36548/jucct.2019.1.003.

[2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, 2014, doi: 10.1109/JPROC.2014.2299397.

[3] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014, doi: 10.1109/JSAC.2014.2328154.

[4] "ns-3 | a discrete-event network simulator for internet systems." [Online]. Available: https://www.nsnam.org/. [Accessed: 13-May-2020].

[5] N. Baldo, M. Miozzo, M. Requena-Esteso, and J. Nin-Guerrero, "An open source product-oriented LTE network simulator based on ns-3," *MSWiM'11 - Proc. 14th ACM Int. Conf. Model. Anal. Simul. Wirel. Mob. Syst.*, pp. 293–297, 2011, doi: 10.1145/2068897.2068948.

[6] M. Mezzavilla, S. Dutta, M. Zhang, M. R. Akdeniz, and S. Rangan, "5G MmWave module for the ns-3 network simulator," *MSWiM 2015 - Proc. 18th ACM Int. Conf. Model. Anal. Simul. Wirel. Mob. Syst.*, pp. 283–290, 2015, doi: 10.1145/2811587.2811619.

[7] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved Handover Through Dual Connectivity in 5G mmWave Mobile Networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2069–2084, 2017, doi: 10.1109/JSAC.2017.2720338.

[8] M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "An efficient uplink multi-connectivity scheme for 5G millimeter-wave control plane applications," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 10, pp. 6806–6821, 2018, doi: 10.1109/TWC.2018.2864650.

[9] M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Multi-connectivity in 5G mmWave cellular networks," *2016 Mediterr. Ad Hoc Netw. Work. Med-Hoc-Net 2016 - 15th IFIP MEDHOCNET 2016*, no. Dl, 2016, doi: 10.1109/MedHocNet.2016.7528494.

[10] M. Polese, M. Mezzavilla, and M. Zorzi, "Performance comparison of dual connectivity and hard handover for LTE-5G tight integration," *Proc. 9th EAI Int. Conf. Simul. Tools Tech. SIMUTools 2016*, vol. 3, pp. 118–123, 2016.

[11] S. Kucera and D. Lopez-Perez, "C-RNTI management for orthogonally-filled subframes in LTE heterogeneous networks," in *IEEE International Conference on Communications*, 2015, vol. 2015-September, pp. 1483–1488, doi: 10.1109/ICC.2015.7248533.

[12] "ns-3: ns3::Constant Position Mobility Model Class Reference." [Online]. Available: https://www.nsnam.org/doxygen/classns3_1_1_constant_position_mobility_model.html#a6c5b0c0e334ffa40cc3480a94c8bd526. [Accessed: 02-May-2020].

[13] M. Mezzavilla *et al.*, "End-to-end simulation of 5G mmWave networks," *IEEE Commun. Surv. Tutorials*, vol. 20, no. 3, pp. 2237–2263, 2018, doi: 10.1109/COMST.2018.2828880.

[14] R. Ford, M. Zhang, S. Dutta, M. Mezzavilla, S. Rangan, and M. Zorzi, "A framework for end-to-end evaluation of 5G mmWave cellular networks in ns-3," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1321, pp. 85–92, 2016, doi: 10.1145/2915371.2915380.

[15] M. Zhang, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Ns-3

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

implementation of the 3GPP MIMO channel model for frequency spectrum above 6 GHz," *ACM Int. Conf. Proceeding Ser.*, vol. Part F1283, pp. 71–78, 2017, doi: 10.1145/3067665.3067678.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:21:30 UTC from IEEE Xplore. Restrictions apply.

IPDroid: Android Malware Detection using Intents and Permissions

Kartik Khariwal *, Jatin Singh † and Anshul Arora‡
Discipline of Mathematics and Computing, Department of Applied Mathematics
Delhi Technological University
Delhi, India
*kartik.khariwal@gmail.com, †jatinsingh0710@gmail.com, ‡anshul15arora@gmail.com

***Abstract*—With the increasing popularity of Android smartphones over the years, the number of malware attacks on Android has also increased. Around 26 million malware samples were detected on the Android platform in the year 2018 and those samples pose serious threats such as financial loss and information leakage. Hence, stronger security solutions need to be developed to detect such threats. Several static Android malware detection techniques exist in the literature that analyze manifest file components such as permissions or intents. However, to the best of our knowledge, none of them have aimed to find the best set of permissions and intents combined that could give better accuracy. In this paper, we use Information Gain to rank the permissions and intents intending to find the best set of permissions and intents that can detect Android malware with better accuracy. We propose a novel algorithm to find the best set by applying several machine learning algorithms such as Random Forest, SVM, and Naive Bayes. The experimental results demonstrate that the best set consisted of 37 features, i.e., 20 intents and 17 permissions, and Random Forest classifier gave the best accuracy of 94.73%.**

***Index Terms*—Android Security, Malware Detection, Permissions, Intents, Static Solution.**

I. Introduction

Smartphones have become so popular nowadays as it can be said that almost everyone who has a pulse has a smartphone. They have become one of the most feasible and easiest ways to access the Internet. According to a recent study, the share of smartphones in the sales market is 50.71% and desktop sales hold a sharemarket of 45.33% [1]. Amongst all mobile OS, Android holds the highest share of 81% [2]. These reports depict the growing popularity of smartphones and it is expected to increase further in the upcoming years. The presence of feature-rich apps ranging from gaming to online banking to social media and location-based services makes smartphones more popular than desktop systems.

With the increasing demand for smartphones, the threat to them is also increasing. According to a report [3], 10 million smartphones were infected with malware through a "Samsung Update", which depicts how vulnerable today's smartphones are and how easy it is for the malware developers to target them. The most common mobile malware are Trojans, mobile bots, ransomware, etc. Mobile malware can pose threats like information leak, system damage, huge financial losses, etc.

Since Android has been the topmost choice for the users, undoubtedly it has become a major target for the applications (apps) based malware as well. According to the reports [4], around 26 million Android malware were detected in 2018. Moreover, around 45,000 Android devices were infected by a recent malware named xHelper in May 2019 [5]. The reasons for such a large number of malware attacks on the Android platform include (i) the considerable number of users, (ii) availability of third-party app markets, and (iii) open architecture of Android. Therefore, stronger and stealthier mechanisms need to be developed to detect Android malware.

A. Motivation

Several solutions exist in the literature that aim to detect Android malware and can be categorized into three types: static, dynamic and hybrid. Dynamic and hybrid solutions aim to analyze the run time behavior of the apps, which can be complex to analyze. Hence, we focus on static detection.

Permissions present within the Android manifest file is the widely used static detection feature followed by Java code and intents [6]. The proposed work in this paper focuses on analyzing manifest file components, hence, we consider permissions and intents for analysis and detection. An app can access the Android's resources with permission constraints, i.e., permissions must be granted to an app to access the resources. Examples of Android's resources include data access stored in the phone such as contacts or media files, and access to hardware components such as camera or microphone. Applications cannot get the permissions granted on their own, instead permissions required by the app need to be confirmed by the user. An Intent is a messaging object we can use to request an action from another app component. There are two types of Intents: Explicit and Implicit. If we want communication between the components of a same application, then we use the Explicit Intents. However, by using the Implicit Intents we can communicate between various apps.

Many detection mechanisms have used these static features for detection. For instance, works such as [7], [8], [9], [10], [11], [12], and [13] have used permissions, intents or some combination of manifest file components to detect Android malware. However, to the best of our knowledge, none of them have aimed to find the best set of permissions and intents combined that could give better accuracy. Hence, in this work, we aim to find the best set of permissions and intents combined that gives better detection results.

978-1-7281-6823-4/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

B. Contributions

The whole methodology is divided into three basic yet effective steps which are highlighted below:

1) First, for all the apps, we created a Bag-of-Words model for permissions and intents which helped in creating a matrix representation with each unique permission/intent represented as a column and apps as rows.
2) Then we calculated Information Gain (I.G.) for each unique permission and intent and thereafter we ranked them based on their I.G. score.
3) Further, to find the best set of permissions and intents, we proposed a novel algorithm to combine the set of permissions and intents based upon by their I.G. score and applied different machine learning models such as SVM, Random Forest & Naive Bayes.

C. Organization

The rest of the paper is organized as follows. We review the works related to Android malware detection in Section II. We discuss the proposed methodology in Section III. Section IV summarizes the results obtained from the proposed approach and Section V concludes with future work directions.

II. Related Work

In this section, we review the works that have been proposed in the literature for Android malware detection. We divide the related works into three categories: Static Detection, Dynamic Detection, and Hybrid Detection.

A. Static Detection

The static solutions aim to detect malicious apps without running or executing them on smartphones. Such solutions use static features for malware detection such as manifest file components, API calls/Java code.

1) Manifest File Based Detection: Few works in the literature such as [14], [15] aimed to find the dangerous permissions present within the benign apps. Sanz et al. [8] analyzed manifest file components like permissions, intents, hardware components, etc. for malware detection. Talha et al. [7] calculated the score of every permission based upon the ratio of number of malicious apps containing that permission, to the total number of malicious apps in the dataset and used these scores in detection. Moonsamy et al. [9] analyzed the dangerous permission patterns present within the malware apps. Wang et al. [10] ranked the permissions based upon the T-Test and Information Gain and used machine learning algorithms for malware detection. The authors in [16] identified dangerous permission combinations that could help in detecting malware. Recent works like [17], [18] and [19] have also analyzed manifest file components for detection. Idrees et al. [12] found the significant permissions and intents present in malware and normal apps and used them for detection. However, they did not discuss the best set of permissions and intents combined that could give better accuracy. Moreover, none of these works have aimed to find the best combined set of permissions and intents for malware detection, which is the aim of this work.

2) API Calls Based Detection: Wu et al. [13] used both API calls and the manifest file components for malware detection. Elish et al. [20] analyzed user-trigger dependence and sensitive APIs in malicious apps. Zhang et al. [21] constructed the dependency graphs of the API calls and classified the malicious apps into their corresponding Android malware families by applying similarity metrics. The authors in [22] proposed a model named *Apposcopy* that analyzed the data-flow and the control-flow properties of the apps for malware detection. However, sometimes, if the API calls used in the process are not connected to any permission then such API calls may prove to be noisy in the detection model. Hence, we have not used the API calls for malware detection.

B. Dynamic Detection

Static solutions do not execute the apps, hence, such solutions may not detect the apps that download malicious component at update time. Hence, dynamic solutions came into existence. We divide these solutions into two categories: OS-based detection and Network Traffic based detection.

1) OS-Based Detection: The authors in [23] analyzed the API and system calls to detect Android malware. Shabtai et al. [24] analyzed percentage of CPU usage, number of active processes, and the number of network traffic packets communicated etc., to detect malware. The authors in [25] studied the system calls of Android malware and observed whether the malicious behavior in the malware apps is initiated from the Java, JNI or native code execution. Extracting the dynamic features like system calls is complex and has high computational overheads as compared to static solutions. Hence, in this work, we propose a static solution.

2) Network Traffic Based Detection: Shabtai et al. [26] analyzed traffic patterns in normal apps and used machine learning classifiers to detect the deviations in those normal patterns. Wang et al. [27] applied NLP techniques on the HTTP headers to detect malware. Arora et al. proposed two models for detecting malicious apps based upon network traffic analysis [28], [29]. All such solutions can only be applied to the malware that generate some network traffic and fail for those which do not generate any traffic, for instance, a malware sending SMS in the background. Hence, we have not considered either an OS-based or network traffic based solution in this work.

C. Hybrid Detection

Some of the solutions aim to combine both static and dynamic features to propose a hybrid mechanism. The authors in [30] analyzed mix of static and dynamic features such as permissions, user activity logs, and system calls for detection. Riskranker [31] analyzed static features like permissions and dynamic features such as run-time Dalvik code loading for malicious apps detection. The authors in [32] and [33] proposed two different hybrid models by combining permissions with network traffic features. Since all the hybrid solutions make use of dynamic features, therefore, they also have the computational overheads of extracting the run time features, hence, we have focused only on static detection in this work.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

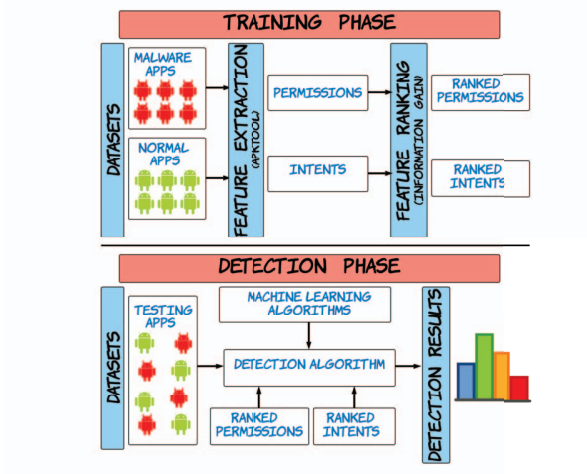


Fig. 1: Proposed Methodology

III. Proposed Methodology

In this section, we discuss the proposed methodology. Figure 1 shows the two phases involved in the methodology: Training and Detection phase. We explain each of the two phases in the following subsections:

A. Training Phase

We collected 1714 malware apps and 1414 benign apps for our experiments. We downloaded the benign dataset from the Google Play Store. Furthermore, we tested the benign dataset on the VirusTotal[1] just to be sure that the benign apps from the Play Store are not malicious. We got the malware dataset from three sources: Genome [34], Drebin [35], and Koodous [36]. Genome and Drebin dataset contained relatively older samples detected back around 2010-14. However, we got the recent samples detected around 2018-19 from Koodous. The training phase consists of the following sub-phases:

1) Features Extraction: Permissions and Intents are defined in the manifest file of an application. To extract these features, we used *Apktool* which extracts the manifest file from the app package(.apk file). Further, we wrote Python scripts to extract the required features, i.e., the permissions and the intents from all the normal and malware apps.

2) Bag-of-Words Model: After extracting the permissions and the intents from the manifest files, we used the Word embedding technique in Natural Language Processing (NLP) for creating two Bag-of-Words Model, one each for the permissions and intents. This model further helped in the matrix representation of the features; with each unique permission/intent represented as a column and applications as rows. The entries in the matrix, say M[A][B], are either 0 or 1. M[A][B] = 1 signifies that Ath app contains the Bth intent/permission, and M[A][B] = 0 signifies that Ath app does not contain the Bth intent/permission.

[1]www.virustotal.com

3) Features Ranking: After representing the permissions and intents in the matrix form, we calculated Information Gain (I.G.) for each unique permission and intent. Information Gain [37] determines which feature, amongst a given set of features, is the most distinguishing one, i.e., the feature that can better distinguish between various classes. Entropy measures the degree of uncertainty of the features and I.G. measures the reduction in uncertainty. For instance, say, there is a training set T that consists of *"n"* number of features F1, F2.., Fn, and there is a class label *C*. Then the entropy for the class C is given by equation 1.

$$H(C) = -\sum_{i} P(C_i) log_2(P(C_i)) \quad (1)$$

where $P(C_i)$ denotes the prior probability for the class C. Further, the uncertainty for class C for any feature F is given by equation 2.

$$H(C/F) = -\sum_{j} P(F_j) \sum_{i} P(C_i/F_j) log_2(P(C_i/F_j)) \quad (2)$$

where $P(C_i|F_j)$ is the posterior probability of the class label C given feature F. Finally, the I.G. of the feature F for the class C is calculated by the equation 3.

$$I.G.(C/F) = H(C) - H(C/F) \quad (3)$$

We calculate I.G. for each unique permission and intent using the above equations. Higher the I.G., better is the distinguishing feature. We rank the permissions and the intents in the decreasing order of the I.G. score. We rank the permissions and the intents separately. At the end of this phase, we get two ranked lists, one each of permissions and intents.

B. Detection Phase

We keep the testing and the training samples separately for our experiments. For each testing sample, we extract its permissions and intents. We use three machine learning algorithms [37]: Random Forest, Naive Bayes and Support Vector Machines (SVM). Algorithm 1 describes the procedure that returns the best set of permissions and intents separately that gives better accuracy than any other set of permissions and intents. The algorithm makes use of ranked permissions and intents obtained from the I.G. score. For the permissions case, initially, it selects the highest-ranked permission. We train machine learning algorithms on the selected permission and find the initial detection accuracy. $Best_{SetP}$ represents the best set of permissions. The top-ranked permission is inserted into $Best_{SetP}$. Further, we select the second-ranked permission, join this with top-ranked permission, train the machine learning algorithms on the joined set and find the accuracy. If the next accuracy obtained from the top two permissions is higher than the previous accuracy i.e., with first permission alone, only then we insert second-ranked permission into the $Best_{SetP}$. Else, we ignore the second-ranked permission, go to third-ranked permission and repeat

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

the procedure. After iterating through all the permissions, the algorithm returns the $Best_{SetP}$. We follow the same procedure to find the best set of intents ($Best_{SetI}$).

However, this work aims to find the best set of permissions and intents combined, hence, we further use the sets: $Best_{SetP}$ and $Best_{SetI}$. Algorithm 2 summarizes the procedure that returns the best set of permissions and intents combined. We merge the sets $Best_{SetP}$ and $Best_{SetI}$ to a single set represented by F_{List}. We further sort the permissions and intents present in F_{List} in decreasing order of I.G. score and the algorithm returns the best set of permissions and intents represented by $Best_{SetPI}$. We must note here that to find this best set $Best_{SetPI}$, we do not use the whole of permissions and intents. By applying Algorithm 1, we first find the best set of permissions and intents separately. And then we use these best sets, i.e., $Best_{SetP}$ and $Best_{SetI}$ to find the combined best set $Best_{SetPI}$.

Each machine learning classifier gives different $Best_{SetP}$ and $Best_{SetI}$, and different detection accuracy. We have used three machine learning classifiers: Random Forest, Naive Bayes, and SVM. Hence, for each classifier, we have different $Best_{SetPI}$. We discuss the detection results in the next section.

Algorithm 1 Returns Best Set of Permissions and Intents Separately

Input: Testing Dataset (T_{Set}), Ranked Permissions List (P_{List}), and Ranked Intents List (I_{List})
Output: Best Set of Permissions($Best_{SetP}$) and Intents($Best_{SetI}$)
Parameters Initialization: $Acc_{MaxP} = 0$, $Acc_{MaxI} = 0$, $P_{Set} \leftarrow \Phi$, $I_{Set} \leftarrow \Phi$, $Best_{SetP} \leftarrow \Phi$, $Best_{SetI} \leftarrow \Phi$
for each permission P_i in P_{List} **do**
    $P_{Set} \leftarrow P_{Set} \cup P_i$
    Train Machine Learning Classifiers on P_{Set} and find the detection accuracy(Acc_{Det})
    if $Acc_{Det} > Acc_{MaxP}$ **then**
        $Best_{SetP} \leftarrow Best_{SetP} \cup P_i$
        $Acc_{MaxP} \leftarrow Acc_{Det}$
    else
        $P_{Set} \leftarrow P_{Set} - P_i$
    end if
end for
for each intent I_j in I_{List} **do**
    $I_{Set} \leftarrow I_{Set} \cup I_j$
    Train Machine Learning Classifiers on I_{Set} and find the detection accuracy(Acc_{Det})
    if $Acc_{Det} > Acc_{MaxI}$ **then**
        $Best_{SetI} \leftarrow Best_{SetI} \cup I_j$
        $Acc_{MaxI} \leftarrow Acc_{Det}$
    else
        $I_{Set} \leftarrow I_{Set} - I_j$
    end if
end for
Return $Best_{SetP}$ and $Best_{SetI}$

Algorithm 2 Returns Best Set of Permissions and Intents Combined

Input: Testing Dataset (T_{Set}), Best Set of Permissions ($Best_{SetP}$), Best Set of Intents ($Best_{SetI}$)
Output: Best Set of Permissions and Intents Combined ($Best_{SetPI}$)
Parameters Initialization: $Acc_{Max} = 0$, $F_{List} \leftarrow Best_{SetP} \cup Best_{SetI}$, $F_{Set} \leftarrow \Phi$, $Best_{SetPI} \leftarrow \Phi$
Sort the features in F_{List} in decreasing order of I.G. score
for each feature F_i in F_{List} **do**
    $F_{Set} \leftarrow F_{Set} \cup F_i$
    Train Machine Learning Classifiers on F_{Set} and find the detection accuracy(Acc_{Det})
    if $Acc_{Det} > Acc_{Max}$ **then**
        $Best_{SetPI} \leftarrow F_{Set}$
        $Acc_{Max} \leftarrow Acc_{Det}$
    else
        Continue
    end if
end for
Return $Best_{SetPI}$

IV. Results and Discussion

In this section, we review the results obtained from our proposed model. After extracting the permissions and intents from all the apps, we got 815 unique intents and 280 unique permissions. Table I summarizes the top 20 ranked permissions and intents, sorted in decreasing order of I.G. score.

A. Permissions Analysis

First, we discuss the results if we use the permissions alone for detection. We apply Algorithm 1 to find the best set of permissions and detection accuracy. We use three machine learning classifiers, and we get different results for each of

TABLE I: Top Ranked Permissions and Intents.

Top 20 Permssions	Top 20 Intents
READ_SMS	BROWSABLE
READ_PHONE_STATE	VIEW
SEND_SMS	BOOT_COMPLETED
WRITE_SMS	SIG_STR
RECEIVE_SMS	DEFAULT
WAKE_LOCK	USER_PRESENT
WRITE_APN_SETTINGS	BATTERY_CHANGED_ACTION
RESTART_PACKAGES	INPUT_METHOD_CHANGED
READ_EXTERNAL_STORAGE	MY_PACKAGE_REPLACED
READ_LOGS	UMS_CONNECTED
CALL_PHONE	UMS_DISCONNECTED
WRITE_CONTACTS	LEANBACK_LAUNCHER
FOREGROUND_SERVICE	DAYDREAM
READ_CONTACTS	SEND
INSTALL_PACKAGES	PHONE_STATE
READ_APP_BADGE	QUICKBOOT_POWERON
MOUNT_UNMOUNT_FILE-SYSTEMS	SEARCH
ACCESS_NETWORK_STATE	TIMEZONE_CHANGED
DISABLE_KEYGUARD	ACTION_POWER_ DISCONNECTED
CHANGE_WIFI_STATE	SEND_MULTIPLE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

TABLE II: Detection Results with Permissions.

Classifier	Detection Accuracy (in %)	Best Set	Rank of Permissions Involved
SVM	92.32	16	1, 2, 3, 6, 7, 13, 15, 17, 21, 22, 24, 25, 26, 32, 37, 45
Naive Bayes	91.56	26	1, 3, 5, 7, 8, 10, 13, 15, 17, 18, 23, 25, 28, 32, 34, 43, 53, 56, 57, 60, 69,86,93,112,129,168
Random Forest	94.25	18	1, 2, 3, 6, 7, 10, 11, 13, 15, 17, 18, 20, 23, 25, 32, 62, 64, 130

the classifiers. Table II summarizes the results for each of the three classifiers.

We can interpret the results as follows: SVM gave the detection accuracy of 92.32% with the best set of 16 permissions. It is interesting to note here that this 16 set of permissions is not the same as the top 16 permissions ranked with I.G. score. The fourth column in Table II gives the rank of those 16 permissions that give the best accuracy with SVM. The top three permissions are present in the set. On adding the fourth permission to the set, the accuracy decreased. Hence, the fourth-ranked permission was discarded from the best set. We observe that with the case of SVM, the proposed detection algorithm went up to 45th ranked permission. On adding further permissions, the accuracy decreased. On comparing the SVM results with those obtained from Naive Bayes and Random Forest, we find that Naive Bayes gave its best accuracy of 91.56% on the best set of 26 permissions. However, the detection algorithm went up to 168th ranked permission to find the best set. Random Forest gave the best accuracy of 94.75% for the permissions case with set of 18 permissions.

B. Intents Analysis

Similar to permissions analysis, we review the results if we use the intents alone for detection. Again, we apply Algorithm 1 to find the best set of intents and detection accuracy. Table III summarizes the results obtained with intents. Similar to the results obtained with the permissions, we observe that Random Forest gave the highest accuracy amongst all the classifiers. However, Random Forest also used the highest number of intents amongst all classifiers in its best set, i.e., 24. Surprisingly the Naive Bayes and Random Forest classifiers went up to 794th ranked intent to give their best accuracy. There were a total of 815 unique intents, so these two classifiers applied Algorithm 1 to almost the entire list of intents. We also observe that permission as a feature gave better accuracy than the intents for our experiments.

C. Permissions and Intents Combined Analysis

Before analyzing the results obtained from the proposed Algorithm 2, we merge the best set of permissions and intents obtained from Algorithm 1 and check for detection accuracy obtained from the merged set. Table IV summarizes the detection accuracy for the merged best sets of permissions and intents. We note that Random Forest gave a better accuracy of 93.98% amongst other classifiers when we merge the best set of permissions and intents.

TABLE III: Detection Results with Intents.

Classifier	Detection Accuracy (in %)	Best Set	Rank of Intents Involved
SVM	86.31	15	1, 3, 4, 6, 9, 10, 11, 12, 15, 18, 19 21, 22, 50, 65
Naive Bayes	82.48	21	1, 2, 6, 9, 12, 14, 15, 18, 26, 28, 35, 39, 67, 75, 85, 90, 105, 120, 124, 780, 794
Random Forest	87.34	24	1, 3, 4, 5, 6, 9, 10, 12, 13, 15, 16, 18, 22, 34, 40, 54, 55, 56, 59, 65, 79, 84, 200, 794

TABLE IV: Detection Results with Merged Best Set of Permissions and Intents.

Classifier	Best Set Count			Accuracy (in %)
	Permissions	Intents	Merged	
SVM	16	15	31	91.94
Naive Bayes	26	21	47	92.45
Random Forest	18	24	42	93.98

Further, Table V summarizes the results obtained from the proposed Algorithm 2. We make two important observations from the results:

1) Random Forest gave a better accuracy of 94.73% from Algorithm 2. Moreover, if we compare the detection accuracy obtained from the best set of permissions (Table II), best set of intents (Table III), merged best set of permissions and intents (Table IV) and the combined set obtained from Algorithm 2 (Table V), we find that the proposed approach gives better accuracy as compared to all other cases for all the three classifiers.
2) On comparing Tables IV and V, we note that the best set count is also lesser from Algorithm 2 as compared to merging the best set of permissions and intents.

Therefore, we conclude that combining permissions and intents using the proposed approach is better than merging their best set.

D. Limitations and Discussion

In this subsection, we discuss the limitations and the false results obtained from the proposed Algorithm 2. We discuss the false results concerning the Random Forest classifier since it has the highest accuracy amongst other classifiers. Most of the false negatives, i.e., malware apps getting detected as normal, contain very few permissions and intents. For instance, many samples of the malware families named BaseBridge, AnserverBot, and DroidKungFu3 contained only one permission: WRITE_EXTERNAL_STORAGE, and two intents:

TABLE V: Detection Results with Permissions and Intents Combined (Algorithm 2).

Classifier	Best Set Count			Accuracy (in %)
	Permissions	Intents	Combined	
SVM	9	8	17	92.42
Naive Bayes	26	11	37	92.54
Random Forest	**17**	**20**	**37**	**94.73**

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

MAIN and LAUNCHER. Similarly most of the false positives, i.e., normal apps getting detected as malware also contain one permission: INTERNET, and two intents: MAIN and LAUNCHER. Hence, the proposed approach fails to detect those apps that contain very few permissions and intents. Therefore, in our future work, we will incorporate other features of manifest files as well like string and network addresses, hardware components, etc. for detection. Additionally, the approach also fails to detect some samples of malware like Pjapps and Geinimi. These malware families are stealthier in the sense that they use some kind of obfuscation technique like encrypting their communication with the remote server. Hence, the proposed approach may not detect stealthier samples with high accuracy. Moreover, since the proposed approach is a static solution, like other static solutions, it may not detect the update attacks, i.e., the apps that download malicious components at update time. Hence, one may include dynamic features as well with the manifest file components to detect stealthier apps and update attacks.

V. Conclusion

In this work, we proposed a novel static technique to detect Android malware using the intents and permissions. Initially, we ranked both the permissions and intents with the help of Information Gain and further we aimed to combine permissions and intents to find their best set that could give better accuracy using different machine learning algorithms. The experimental results demonstrated that the proposed methodology of combining permissions and intents gives better detection accuracy than permissions alone and intents alone. In our future work, we will incorporate other manifest file components as well like hardware components, for our analysis and detection.

References

[1] Desktop vs Mobile vs Tablet Market Share Worldwide, Available Online. https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet.

[2] Android dominates 81 percent of world smartphone market, Available Online. https://www.cnet.com/news/android-dominates-81-percent-of-world-smartphone-market/.

[3] Critical Warning Issued Regarding 10 Million Samsung Phone Updates, Available Online. https://www.forbes.com/sites/daveywinder/2019/07/05/critical-warning-issued-regarding-10-million-samsung-phone-updates/#77c532923cf8.

[4] Development of new Android malware worldwide from June 2016 to May 2019, Available Online. https://www.statista.com/statistics/680705/global-android-malware-volume/.

[5] 45,000 Android devices infected by new unremovable xHelper malware, Available Online. https://thenextweb.com/security/2019/10/30/45000-android-devices-infected-by-new-unremovable-xhelper-malware/.

[6] A. Feizollah et al., "A review on feature selection in mobile malware detection", Digital Investigation, vol. 13, pp. 22-37, 2015.

[7] K. Talha, D. Alper, and C. Aydin, "APK Auditor: Permission-based Android malware detection system", Digital Investigation, vol. 13, pp. 1-14, 2015.

[8] B. Sanz, I. Santos, C. Laorden, X. Ugarte-Pedrero, J. Nieves, P. Bringas, and G. Lvarez, "MAMA: manifest analysis for malware detection in android", Cybernetics and Systems, vol. 44, pp. 469-488, 2013.

[9] V. Moonsamy, J. Rong, and S. Liu, "Mining permission patterns for contrasting clean and malicious android applications", Future Generation Computer Systems, vol. 36, pp. 122-132, 2014.

[10] W. Wang et al., "Exploring Permission-Induced Risk in Android Applications for Malicious Application Detection", IEEE Transactions on Information Forensics and Security, vol. 9 , pp. 1869-1882, 2014.

[11] F. Idrees et al., "PIndroid: A novel Android malware detection system using ensemble learning methods", Computers & Security, vol. 68, pp. 36-46, 2017.

[12] F. Idrees, and M. Rajarajan, "Investigating the Android Intents and Permissions for Malware detection", 7th International Workshop on Selected Topics in Mobile and Wireless Computing, 2014.

[13] D. Wu et al., "DroidMat: Android Malware Detection through Manifest and API Calls Tracing", 7th Asia Joint Conference on Information Security, 2012.

[14] M. Grace, W. Zhou, X. Jiang, and A. Sadeghi, "Unsafe exposure analysis of mobile in-app advertisements", 5th ACM WiSec, 2012.

[15] W. Enck, M. Ongtang, and P. McDaniel, "On Lightweight Mobile Phone Application Certification", 16th ACM CCS, 2009.

[16] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang, "Hey, You, Get Off of My Market: Detecting Malicious Apps in Official and Alternative Android Markets", NDSS, 2012.

[17] T. Kim et al., "A Multimodal Deep Learning Method for Android Malware Detection Using Various Features", IEEE Transactions on Information Forensics and Security, vol. 14, 2019.

[18] A. Arora, S. Peddoju, and M. Conti, "PermPair: Android Malware Detection using Permission Pairs", IEEE Transactions on Information Forensics and Security, 2019.

[19] H. Zhua et al., "DroidDet: Effective and robust detection of android malware using static analysis along with rotation forest model", Neurocomputing, vol. 272, pp. 638-646, 2018.

[20] K. Elish et al., "Profiling user-trigger dependence for Android malware detection", Computers & Security, vol. 49, pp. 255-273, 2015.

[21] M. Zhang et al., "Semantics-Aware Android Malware Classification Using Weighted Contextual API Dependency Graphs", ACM CCS, 2014.

[22] Y. Feng, S. Anand, I. Dillig, and A. Aiken, "Apposcopy: Semantics based detection of android malware through static analysis", 22nd ACM SIGSOFT Symposium on Foundations of Software Engineering, 2014.

[23] V.M. Afonso et al., "Identifying Android malware using dynamically obtained features",Journal of Computer Virology and Hacking Techniques,vol. 11,pp.9-17,2015.

[24] A. Shabtai, U. Kanonov, Y. Elovici, C. Glezer, and Y. Weiss, "Andromaly: a behavioral malware detection framework for android devices", Journal of Intelligent Information Systems, vol. 38, pp. 161-190, 2011.

[25] A. Reina, A. Fattori, and L. Cavallaro, "A System Call-Centric Analysis and Stimulation Technique to Automatically Reconstruct Android Malware Behaviors", 6th European Workshop on System Security, 2013.

[26] A. Shabtai et al., "Mobile malware detection through analysis of deviations in application network behavior", Computers & Security, vol. 43, pp. 1-18, 2014.

[27] S. Wang, et al., "Detecting Android Malware Leveraging Text Semantics of Network Flows", IEEE Transactions On Information Forensics And Security, vol. 13, pp. 1096-1109, 2018.

[28] A. Arora, S. Garg, and S.Peddoju,"Malware detection using network traffic analysis in android based mobile devices", 8th IEEE NGMAST,2014.

[29] A. Arora, and S. Peddoju, "Minimizing Network Traffic Features for Android Mobile Malware Detection", 18th ACM ICDCN, 2017.

[30] A. Saracino, D. Sgandurra, G. Dini, and F. Martinelli, "MADAM: Effective and Efficient Behavior-based Android Malware Detection and Prevention", IEEE Transactions on Dependable and Secure Computing, vol. 15, pp. 83-97, 2018.

[31] M. Grace, Y. Zhou, Q. Zhang, S. Zou, and X. Jiang, "Riskranker: scalable and accurate zero-day android malware detection", 10th ACM Mobisys, 2012.

[32] A. Arora, and S. Peddoju, "NTPDroid: A Hybrid Android Malware Detector Using Network Traffic and System Permissions", 17th IEEE TrustCom, 2018.

[33] A. Arora, S. Peddoju, V. Chauhan, and A. Chaudhary, "Hybrid Android Malware Detection by Combining Supervised and Unsupervised Learning", 24th ACM MobiCom, 2018.

[34] Y. Zhou, and X. Jiang, "Dissecting android malware: Characterization and evolution", IEEE Symposium on Security and Privacy, 2012.

[35] D. Arp,M. Spreitzenbarth,M. Hubner,H. Gascon, and K. Rieck, "DREBIN: Effective and Explainable Detection of Android Malware in Your Pocket",NDSS,2014.

[36] Koodous Malware Dataset, "www.koodous.com".

[37] J. Han, P. Jian, and M. Kamber, "Data mining: concepts and techniques", Elsevier, 2011.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 06,2020 at 07:03:09 UTC from IEEE Xplore. Restrictions apply.

Resources, Conservation & Recycling 164 (2021) 105215

Contents lists available at ScienceDirect

Resources, Conservation & Recycling

journal homepage: www.elsevier.com/locate/resconrec

Full length article

Managing supply chains for sustainable operations in the era of industry 4.0 and circular economy: Analysis of barriers

Pravin Kumar [a], Rajesh Kr Singh [b,*], Vikas Kumar [c,d]

[a] Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi 110042
[b] Management Development Institute, Gurgaon, India 122007
[c] Bristol Business School, University of the West of England (UWE), UK
[d] Faculty of Accounting, Ton Duc Thang University, Ho Chi Minh City, Vietnam

ARTICLE INFO

Keywords:
Industry 4.0
Circular economy
Sustainability
Barriers
AHP
ELECTRE

ABSTRACT

Organizations are struggling to leverage emerging opportunities for maintaining sustainability in the global markets due to many barriers in the era of Industry 4.0 and circular economy. The main aim of this study is to analyze these barriers to improve the sustainability of a supply chain. Our study identifies the major criteria for sustainable operations and barriers that need to be overcome to achieve the objectives of sustainability through literature review and experts' opinions. An integrated approach comprising Analytic Hierarchy Process (AHP) and Elimination and Choice Expressing Reality (ELECTRE) is used to analyze these barriers and ensure the sustainable supply chain operations. Resource circularity, increasing profits from green products, and designing processes for resource and energy efficiency have been found to be as major sustainability criteria.

There are many barriers to the implementation of Industry 4.0. These barriers include but are not limited to, a lack of a skilled workforce that understands Industry 4.0, ineffective legislation and controls, ineffective performance framework, and short-term corporate goals. The study finds that ineffective strategies for the integration of industry 4.0 with sustainability measures, combined with a lack of funds for industry 4.0 initiatives, are just two of the major barriers. The findings of the study will help organizations to develop an effective and integrated strategic approach that will foster sustainable operations through the utilization of improved knowledge of Industry 4.0 and the circular economy.

1. Introduction

The increasing population, a growing economy, and an improved lifestyle of people have increased the exploitation of natural resources. Preston and Herron (2016) observed that the demand for these natural resources has been growing steadily. Due to this increasing demand for natural resources, organizations are facing many operational challenges. The shortage of resources is resulting in increased input costs and as a consequence products are less sustainable in the market (Preston and Herron, 2016). Most manufacturing organizations are still working on the traditional linear economy model (Singh et al., 2020). The conventional approach of reuse and recycling of materials is not much cost-effective and leading to wastage of valuable resources.

Global supply chains are facing extensive pressure from regulatory bodies to become sustainable across all operational activities (Vanalle et al., 2017). Environmental and social concerns are forcing organizations to transition from linear to circular processes to halt the unproductive disposal of consumer or end of life products. The Circular Economy (CE) approach consists of the 3Rs (Reduce, Reuse, and Recycle). The 3R's system is used as a tool to improve sustainability within operations (Sauve et al., 2016). CE enables more effective use and reuse resources by employing competent design and manufacturing processes.

Technology has emerged as an important driver of sustainable operations (Dubey et al., 2017). Organizations must exploit emerging technologies to make operations environmentally-friendly and economically feasible. Wang et al. (2016) observed that some Industry 4.0 based technologies can make operations sustainable. These are the Internet of Things (IoT), Cloud Manufacturing (CM), Cyber-Physical Systems (CPS), and Additive Manufacturing (AM) (Hozdie, 2015).

Using the combined approach of Industry 4.0 and CE, resources can be utilized multiple times for different purposes (Stock and Seliger,

* Corresponding author.
E-mail addresses: pravin.dce@gmail.com (P. Kumar), rajesh.singh@mdi.ac.in (R.K. Singh), vikas.kumar@uwe.ac.uk, Vikas.Kumar@tdtu.edu.vn (V. Kumar).

https://doi.org/10.1016/j.resconrec.2020.105215
Received 31 March 2020; Received in revised form 4 October 2020; Accepted 6 October 2020
Available online 10 October 2020
0921-3449/© 2020 Elsevier B.V. All rights reserved.

2016). Many Indian organizations have started using Industry 4.0 based technologies such as IoT, machine learning and blockchain that have been found to improve the process efficiency. For example, Tata Power Ltd has created a digital platform to aid its customers to better manage power consumption with real-time data and improve the efficiency of power plants. Voltas Ltd is using IoT based solutions to offer their customers a superior chiller maintenance service.

Lin (2019) observed that Industry 4.0 based technologies may help to foster a more suitable decision-making process and facilitate smart production for CE. Dubey et al. (2017) state that through effective integration of Industry 4.0 and sustainability measures, organizations can develop world-class sustainable manufacturing systems. It has been observed that the implementation and usage of digital technologies can overcome various challenges of CE (Singh et al., 2019). Many organizations, however, are not sufficiently able to exploit their potential for sustainable growth in a globalized market because they lack awareness of the basic principles that are the foundation of resource circularity and Industry 4.0.

Past studies have not made any attempt to pursue the analysis of the barriers for sustainable operations within the context of CE and Industry 4.0. Singh et al. (2019) and Dubey et al. (2017) have posited that very little literature on the integration of Industry 4.0 and CE is available and that most studies have analyzed the barriers that impede CE and Industry 4.0 separately.

With this in mind, this study addresses the following research objectives:

- Identification of criteria for sustainability in the context of CE.
- Identification of barriers for sustainable operations in the era of Industry 4.0 and CE.
- Mapping of these barriers of sustainable operations.

The main goal of the study is to identify the ranking of different barriers that impede sustainable operations of a supply chain in the context of Industry 4.0 and CE. A positive argument can then be made in favor of sustainable digitized manufacturing processes. The findings may also help industry professionals to exploit emerging business opportunities by managing barriers and challenges.

The organization of the remaining part of this paper is as follows: Literature review of industry 4.0 and sustainability parameters of CE is presented in Section 2. The research methodology comprising of AHP and ELECTRE is detailed in Section 3. The results are presented in Section 4 and the practical and theoretical implications of this study are discussed in Section 5. Finally, Section 6 provides the conclusions of the study including its limitations and potential directions for future research.

2. Literature review

Within the field of industrial research, two major trends have emerged in recent years: Industry 4.0 and the CE.

Industry 4.0 has led to a wave of technology that drives the digitization of operations (Jabbour et al., 2018a). Digital communication along the whole value chain, or the flow of the data, is termed 'digital threads' (Nanry et al., 2015), and these simplify the business processes.

CE is based on the closed-loop design of the production system and leads to an increase in the efficiency of resources. This study will map the barriers that hinder sustainable operations in an era of Industry 4.0 and CE. The literature review in the following sub-sections discusses different opportunities and challenges that arise in the emerging business environment of Industry 4.0 and CE.

2.1. Industry 4.0: Opportunities and challenges

The term "Industry 4.0″ was first coined at a trade fair organized in Hannover, Germany in 2011 (Kagermann et al. 2015). Industry 4.0 technologies can assist companies to achieve integrated, flexible, and diversified production systems that will lead to the production of customized products (Li et al., 2017; Thoben et al., 2017). Organizations can allocate resources efficiently on a real-time basis by employing Industry 4.0 technologies.

The IoT, cloud manufacturing, cyber-physical systems and AM are four major components of Industry 4.0 (Jabbour et al., 2018a). The cyber-physical system is the integration of cyberspace, physical systems, objects and machines or devices that form a network that obtains real-time data enabling the decision-making process for prioritization of production orders, maintenance requirements, and optimization of tasks (Lee et al., 2015). The IoT facilitates collection and transmission of data among different devices and objects (Li et al., 2017): Radio Frequency Identification (RFID), bar-coding and wireless sensors are just some technologies that form part of the IoT (Zang et al., 2017). Cloud manufacturing provides information concerning the availability and sharing of manufacturing resources (Liu and Xu, 2017).

With the application of IoT, companies collect and manage big data that enable effective decision making. For example, business applications for improving the response of mobile taxis (Lanza et al., 2015) and the reduction of energy consumption in manufacturing and service processes (Tao et al., 2016).

Cloud computing has various industrial applications such as monitoring the mobility of transportation management systems (Nowicka, 2016) and calculating carbon footprint (Singh et al., 2015). Artificial intelligence may improve the effectiveness of energy management (Zahraee et al., 2016) and Shivajee et al. (2019) observed that by using IT-enabled quality tools and monitoring their processes continuously, organizations were able to reduce their manufacturing conversion cost.

Many organizations are unable to implement Industry 4.0 technologies because of the poor leadership style (Shao et al., 2017). For successful implementation of Industry 4.0 technologies that promote the sustainable performance of organizations, senior management should be both inspiring and transformational (Politis, 2001). Senior management support is vital for making changes regarding processes and performance (Young and Jordan, 2008). The application of technologies should be strategically aligned with the long-term goals of an organization. Indeed, a lack of effective strategic integration of Industry 4.0 with sustainability goals may lead to the failure of an organization (Kahn et al., 2012). Absence of alignment between Industry 4.0 technologies and CE criteria has been observed as a source of difficulty for sustainable operations (Ball et al., 2018) and it may also result in the making of unsustainable manufacturing decisions.

Successful implementation of emerging technologies requires a skilled workforce and to this effect, employees should be trained to meet emerging challenges found within sustainable operations and Industry 4.0 technologies (Waibel et al., 2017; Jabbour and Jabbour, 2016; Sarkis et al., 2010). Successful implementation will help to reduce employees' resistance to change from a linear economy approach to one of CE. Organizational culture should motivate employees to remain updated about emerging technologies and future challenges (Jabbour and Jabbour, 2016). Many organizations lack the ability to take initiatives for upgrading technological processes which would make them sustainable and more competitive. This is due to a lack of awareness surrounding the potential of Industry 4.0 technologies and CE.

Jabbour et al. (2018a) also observed that organizations may not implement Industry 4.0 technologies because of the absence of senior management support, ineffective strategies, poor workforce skills, lack of funds for investment and, again, insufficient information regarding Industry 4.0. Several other market barriers hindering sustainable operations are lower cost of virgin materials, costly recycling process, lower profit margin of green products and poor market demand (Mont et al., 2017; Preston, 2012).

There are various constraints that hamper the technological ability of an organization to fulfill its sustainability potential. For example, altering a linear economy to CE is expensive and can take an inordinate

amount of time. Many organizations are not able to implement new technologies such as IoT, cloud computing and big data analytics due to fear of misinvestment (Ranta et al., 2018). Effective handling of barriers to sustainable manufacturing operations is crucial (Luthra et al., 2015).

Through a literature review and experts opinion, fifteen important barriers are identified within this study and these barriers are summarized in Table 1.

2.2. Circular economy and sustainable operations

Stahel and Reday (1976) were the first to introduce the term CE. CE comprises a regenerative system based on the philosophy of zero waste, the concept being that waste produced within an organization has the potential to be utilized as a valuable resource by another organization. Geng and Doberstein (2008) describe CE as "realization of closed-loop material flow in the whole economic system". Webster (2015) state that "CE is one that is restorative by design, and which aims to keep products, components, and materials at their highest utility and value, at all times".

Some of the most important components of CE within the literature are performance economy (Stahel, 2010), industrial ecology (Lifset and Graedel, 2001), cradle-to-cradle design (Braungart et al., 2007), product-service systems (Tukker, 2015), natural capitalism (Hawken et al., 2008), industrial symbioses (Chertow and Ehrenfeild, 2012), biomimicry (Benyus, 2003), circular flow of materials (Lieder and Rashid, 2016), industrial ecosystems (Jelinski et al., 1992), eco-efficiency (Haas et al., 2015), regenerative design (Lyle, 1994), zero emissions, industrial ecology (Graedel and Allenby, 1995), etc.

The cradle-to-cradle model uses a material reuse principle. In this model, industrial materials can be classified as technical or biological nutrients. Technical nutrients are considered as nontoxic, non-hazardous synthetic materials that can be utilized repeatedly, biological nutrients are organic materials that decompose without affecting the natural environment and biomimicry systems imitate the nature of the innovation in process and product design.

Using CE, the material value can be maintained for a long time; the sustainability of operations is an essential requirement of CE (Kazancoglu et al., 2018). Operations sustainability can be judged by measuring different kinds of emissions. For example, SO_2, NO, effluent and solid wastes, hazardous and toxic material consumption (Paulraj et al., 2017; Zhu et al., 2017). Sustainability is assured through the implementation of reverse logistics and other resource circularity practices within a supply chain (Pourjavad and Shahin, 2018; Sharma et al., 2017; Sarkis, 2003). Organizations should also develop a capability to improve resource efficiency, product quality and decreased scrap that will aid the sustainability of operations (Geng et al., 2017; Sharma et al., 2017). It is safer to produce environmentally-friendly products of consistent quality and regular use. The green-design of products minimizes energy consumption and waste disposal (Foo et al., 2018; Jabbour and Jabbour, 2016) but green products require green materials and degradable/reusable packaging materials to protect the environment (Wu et al., 2015). Papadopoulos and Giama (2007) emphasized the implementation of certification and accreditation based on the use of the 3Rs (recycling, reuse, and reduction) in supply chain operations. A sustainable supply chain not only boosts community welfare initiatives, but also decreases cost, increases market share and resource efficiency, sales, and profit margin and employees' motivation and satisfaction (Younis et al., 2016). The literature review identifies some of the major sustainability criteria in Table 2.

The proposed framework for mapping of barriers restricting sustainable operations of supply chains in the era of Industry 4.0 and CE is shown in Fig. 1. This framework summarizes four Industry 4.0 based technologies: Cyber-Physical Systems, IoT, Cloud manufacturing, and AM and will include some major barriers that block the implementation of these technologies as discussed in the previous section. CE can be summarized by using four fundamental principles: cradle-to-cradle,

Table 1
Barriers to sustainable operations of supply chains.

S. No.	Barriers for Sustainable Operations of supply chain	References	Remarks
B1	Risk of misinvestment	Müller et al. (2017), Moser et al. (2017), Ranta et al. (2018).	A large investment is required for procuring the Industry 4.0 technologies. Investors worry about the implications of misinvestment.
B2	Insufficient legislation & control	Kamble et al. (2018), Oesterreich and Teuteberg (2016), Raj et al. (2019), Muduli et al. (2013).	Insufficient legislation and control over the small and medium enterprises and ancillary companies have served to ensure that management is reluctant to utilize new technologies for sustainable operations.
B3	Insufficient strategy for integration of Industry 4.0 & C.E	Ball et al. (2018), Rajput and Singh (2019); Kahn et al. (2012), Kiel et al. (2017a, 2017b), Kamble et al. (2018); Jabbour et al. (2018b).	A well-defined strategy is required to understand why Industry 4.0 should be implemented and how implementation can be fulfilled. The strategy may overcome a major hurdle regarding the integration of Industry 4.0 technologies with the concept of the circular economy.
B4	Lack of skilled workforce	Waibel et al. (2017); Jabbour and Jabbour (2016); Sarkis et al. (2010); Erol et al. (2016), Müller and Voigt (2017), Kiel et al. (2017a, 2017b).	Successful application of new technology requires a specialized workforce. Lack of skills may be treated as a barrier that impedes this implementation.
B5	Lack of funding for Industry 4.0	Kulatunga et al. (2013); Erol et al. (2016), Kiel et al. (2017a, 2017b), Müller and Voigt (2017), PWC (2014), Mittal et al. (2018), Kumar et al. (2020)	The investment required for Industry 4.0 technologies is very high. Lack of funds is a barrier to the implementation of Industry 4.0 technologies.
B6	Ineffective performance framework	Sarkis (2012), José et al. (2017), Batista et al. (2018).	The ineffective performance framework displays a lack of planning and policy for implementation and measuring performance of the systems.
B7	Use of materials as energy	Braungart et al. (2007), Benyus (2003).	Technology is required for recycling waste materials and to utilize them as renewable materials. Some material can be also used as alternative sources of energy.
B8	Lack of waste management	Pourjavad and Shahin (2018), Sharma et al. (2017); Sarkis (2003).	Recycling, reuse and re-manufacturing are the major components of CE. Lack of waste management creates a hurdle in front of fulfilling the requirements of the CE.
B9	Poor resource/ infrastructure quality	Xu et al. (2018), Batista et al. (2018), Ghadimi et al. (2019), Geng et al. (2017), Sharma et al. (2017).	High quality of infrastructure and sufficient resources are required for successful implementation of industry 4.0 technologies.
B10			

(continued on next page)

Table 1 (*continued*)

S. No.	Barriers for Sustainable Operations of supply chain	References	Remarks
	Lack of Govt. support	Prakash and Barua (2015), Govindan et al. (2013), Moktadir et al. (2018).	There are a plethora of rules and regulations for waste minimization and eco-friendly manufacturing activities but the government fail to encourage process automation through Industry 4.0 technologies.
B11	Employees resistance to change	Jabbour and Jabbour (2016), Luthra et al. (2019), Kiel et al. (2017b), Haddud et al. (2017).	The employees may resist changes due to fear of loss of jobs: the automation may eliminate many activities that are normally performed manually.
B12	Insufficient market demand	Mont et al. (2017), Preston (2012), Lin et al. (2013).	Due to higher processing and manufacturing cost of green products, market demand is not sufficient. Thus, marketing of green, refurbished and recycled products is a major barrier for sustainable operations of the supply chains.
B13	Lack of management support	Turker and Altuntas (2014), Horváth and Szabó (2019), Ghadimi et al. (2019), Young and Jordan (2008), Jabbour et al. (2018a).	Due to lack of vision and also lack of finance, senior management is often reluctant to support activities for sustainable operations.
B14	Short term goals	Giunipero et al. (2012), Kambli et al. (2018), Moeuf et al. (2018).	In most cases, the management has short term goals and they require immediate output. Therefore, organizations are hesitant to invest financially in CE and Industry 4.0.
B15	Lack of awareness of Industry 4.0	Luthra and Mangla (2018), Jabbour et al. (2018a), Hofmann and Rüsch (2017).	Many organizations are not aware of the benefits to be gained from implementing Industry 4.0 and CE concepts, hence they lack motivation for the change.

biomimicry, industrial ecology and blue economy covering nine sustainability parameters.

Based on the literature review, the authors have identified the following gaps:

(i) Most of the literature available on the circular economy is based on the philosophy and benefits of CE. Further study is therefore needed to look at feasibility related issues in the implementation of CE.
(ii) Most of the studies have analyzed issues related to Industry 4.0 and the circular economy in isolation. Therefore, study integrating Industry 4.0 concepts with a CE is required.
(iii) The authors have determined the barriers that burden the implementation of green supply chains and advanced manufacturing technologies in isolation. This leads to further study of the barriers that impede sustainable operations in the current business environment of Industry 4.0 and CE.

Table 2
Sustainability parameters in the context of the circular economy.

S. No.	Sustainability Parameters	References	Remarks
C1	Resource Circularity	Figge et al. (2018), Stock and Seliger (2016), Lieder and Rashid (2016), Linder et al. (2017).	Resource circularity is based on the concept of cradle-to-cradle and biomimicry. It aids the elimination of waste.
C2	Cost-saving through product quality	Yeow and Sen (2006), Shivajee et al. (2019); Rust et al. (2002).	Improved quality reduces the cost of scrap, shutdown, inspection, and testing, etc. Thus, these costs can be lessened by improving the quality of the systems.
C3	Decreasing emission	Singh et al. (2015), Graedel and Allenby (1995), Paulraj et al. (2017), Zhu et al. (2017).	Decreasing emission is one of the criteria for evaluating the effectiveness of sustainable operations.
C4	Waste reduction & pollution monitoring	Heck and Rogers (2014), Geng et al. (2012).	Waste in the various forms (Waste materials, packaging materials, wastewater, and gaseous emission) is to be reduced and pollution must be monitored properly for sustainability.
C5	Process design for resource and energy efficiency	Towler (1996), Tao et al. (2016), Zahraee et al. (2016).	The circular economy leads to the highest utility of products, components, and energy since it is restorative and regenerative by design.
C6	Increasing profit from green products	Foo et al. (2018); Jabbour and Jabbour (2016).	For market survival of green products, the profit margin should be high. Reducing product cost leads to an increase in product demand.
C7	Improving green logistics	Kazancoglu et al. (2018), Zeng et al. (2017).	All green practices such as the use of green packaging materials, low emission of the vehicle, proper maintenance of the vehicle are a part of green logistics.
C8	Improved employees and community health	Younis et al. (2016), Moldavska and Welo (2017), Barnidge et al. (2011).	This is concerned with social sustainability. The proper care of employees and community health are part of corporate social responsibility.
C9	Improving green purchasing	Moktadir et al. (2018), Tseng et al. (2019), Yadav et al. (2020).	It means to purchase green materials, components, and sub-components and services for producing the goods and services.

3. Research methodology

In the literature, several multi-criteria decision-making (MCDM) tools have been used, all of which have advantages and disadvantages. Within this study, an AHP and ELECTRE integrated approach is implemented for the analysis. The AHP method is preferred by many researchers due to the possibility of applying both qualitative and quantitative criteria. It has effective traceability of the decision and assures quality using consistency indices (Tscheikner-Gratle et al., 2017). However, AHP suffers from rank reversals in the case of addition and removal of the criteria and alternatives within the decision model. This can be prevented by prohibiting the user from using the addition and removal of criterion and any alternatives from the decision model (Hodgett, 2016).

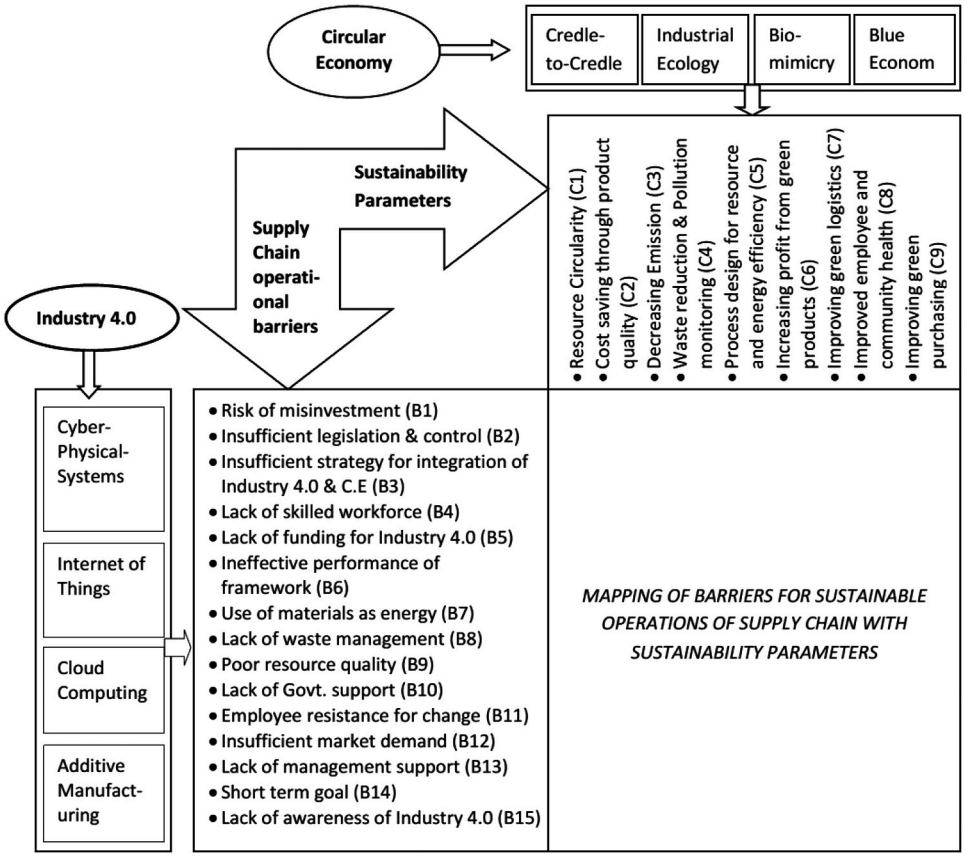


Fig. 1. Mapping of barriers that impede sustainable operations in the era of Industry 4.0 and CE.

Similarly, ELECTRE also has some advantages and disadvantages. According to Li and Wang (2007), the weakness of the normal ranking of ELECTRE is that it requires an additional threshold to be introduced. The ranking of the alternatives depends upon the size of the threshold, for which there exists no 'correct' value. On the other hand, Sabaei et al. (2015) mentioned that "the main advantage of the method is that ELECTRE can handle both quantitative and qualitative data for out-ranking alternatives with high uncertainty". ELECTRE is also less sensitive to changes in data compared to other methods. Considering the above facts, the combination of the AHP and ELECTRE model is found to be more accurate than the individual model of ranking. ELECTRE is used in this study.

The barriers that hamper sustainable operations of supply chains are mapped within this study to sustainability parameters of CE as shown in Fig. 1. The importance of the ranking of criteria for sustainable operations has been found using AHP. The weights during pairwise comparisons between criterion are based on the expert consensus that took place during the discussion. Finally, ELECTRE is used to find the rank of the barriers of sustainable operations of supply chains within the CE environment.

The research methodology is next discussed in two parts. The first part discusses the steps of AHP and the second part discusses the principles underpinning ELECTRE.

3.1. Analytic hierarchy process

AHP helps decision-makers to analyze the relative importance of criterion to implement effective decision-making (Saaty, 1980). AHP is found to be more advantageous in comparison to ANP because of lower pairwise comparisons regarding the lesser hierarchy problem. In this study, AHP has been used for ranking the parameters of CE. These criteria are compared with each other on the Saaty's (1980) nine-point rating scale. The calculation procedure of AHP is summarized below:

- Define the problem and objective.
- Divide the problems in the form of a hierarchy (from objective to the lower- level).
- Undertake pairwise comparison using Saaty's 9-point rating scale for each lower level.
- Check the consistency ratio.
- Estimate the relative rating of the component at each level.

Relative ranking of sustainability criteria received by AHP is used in the next stage of this study concerned with the ranking of barriers (Table 3).

3.2. Elimination and choice expressing reality (ELECTRE)

ELECTRE is used to rank the barriers for sustainable operations of supply chains in the context of Industry 4.0. Decision making can be undertaken by adopting AHP. MCDM, however, gives superior results when integrated with other decision-making tools (Kang and Park, 2014). Kumar and Singh (2020) used the integrated model of AHP and ELECTRE for evaluating alternative solutions for crop residue burning. Roy (1991) proposed the ELCTRE method in 1965 for ranking and outranking of alternatives. The applications of ELECTRE have also been reported in a plethora of articles that include its fuzzification (Kumar et al., 2017a, 2017b, 2016).

Using the ELECTRE process, the barriers are rated against the sustainability criterion on the five-point rating scale. These ratings are normalized and a weighted normalized matrix is formed. The concordance and discordance matrices are prepared from the weighted matrix based on the threshold, defined in the ELECTRE method. These concordance and discordance matrices are converted into the Boolean

Table 3
Pairwise Comparison of Criteria for Importance Rating.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Weights	Rank
C1	1	4	5	3	2	3	7	6	7	0.297	1
C2	0.25	1	2	3	0.5	0.5	4	5	5	0.131	4
C3	0.2	0.5	1	0.25	0.33	0.5	2	2	3	0.061	6
C4	0.33	0.33	4	1	0.5	0.33	3	2	3	0.091	5
C5	0.5	2	3	2	1	0.33	5	3	4	0.144	3
C6	0.33	2	2	3	3	1	2	2	2	0.16	2
C7	0.14	0.25	0.5	0.33	0.2	0.5	1	0.5	2	0.037	8
C8	0.17	0.2	0.5	0.5	0.33	0.5	2	1	2	0.047	7
C9	0.14	0.2	0.33	0.33	0.25	0.5	0.5	0.5	1	0.031	9

Number of comparisons = 36; Consistency Ratio CR = 7.5%; Principal Eigen value = 9.873; Eigenvector solution: 6 iterations, delta = 2.1E-8.

matrix and the elements of these Boolean matrices are multiplied, forming a Dominance matrix. These barriers of sustainable operations in the supply chain are ranked based on the Dominance matrix. The detailed calculation steps of ELECTRE (Shanian and Savadogo, 2006) are mentioned in the Appendix (Eq. (A1) to A12).

4. Results

The analysis within the study is based upon the expert opinions of employees within companies located in Delhi NCR, India. A large number of small and medium scale enterprises have established their plants in Delhi NCR. The national capital region, comprising the states of Delhi, some parts of Haryana, Uttar Pradesh and Rajasthan, produce on average 30% of the total auto components produced throughout the country. The primary industrial sector locations within the NCR are Delhi, Gurgaon, Manesar, Bhiwadi, Meerut, Ghaziabad, Faridabad, Kundli, Sonepat, Murthal, Gautam Budh Nagar, Bulandshahr, and Panipat.

Maruti Udyog Limited, also located in Delhi NCR, is the largest car manufacturer in India along with other major auto sector companies, including Honda Motors and Hero MotoCorp. This region is also home to many electronics companies such as Samsung Electronics and Oppo mobiles Pvt Ltd.

A team comprising five experts was put together to undertake pairwise comparisons of different criterion. Three experts were experienced in the industry and had more than ten years' experience regarding sustainability-related issues. The remaining two experts were academicians working in the area of sustainability.

Nine sustainability parameters of CE and fifteen barriers impeding sustainable operations have been identified within this study through literature review and experts' opinions. The proposed framework for analyzing barriers prohibiting sustainable operations is shown in Fig. 1. Initially, the pairwise comparison matrix of the criterion has been prepared based on a 9-point rating scale using AHP as discussed in the research methodology section. The importance ratings of criterion using pairwise comparison are shown in Table 3.

After finding the importance rating of the sustainability criterion, a decision matrix (Table 4) of different barriers hampering sustainable operations of supply chains using a five-point rating scale has been prepared using Eq. (A1).

Normalization of the decision matrix is required to ensure its elements are on the same scale, the scale being from 0 to 1. Normalization is undertaken using Eq. (A2). After normalization, the weighted normalized matrix is produced by multiplying the elements with the corresponding weights of the criterion using Eq. (A3). The concordance and discordance sets of criteria are segmented for all the barriers from the weighted normalized matrix using the conditions in Eqs. (A4) and (A5) respectively. Now, the concordance matrix is developed from the summation of the concordance criterion as mentioned in Eq. (A6). The discordance matrix is developed using Eq. (A7). The concordance matrix is converted into a Boolean Matrix E using the threshold as mentioned in Eqs. (A8) and (A9). The discordance matrix is converted into Boolean Matrix F using the threshold in equations (A10) and (A11). Finally, a dominance matrix G is formed by multiplying the corresponding elements of the Boolean Matrices E and F as shown in Table 5.

The Boolean value 1 indicates that the i^{th} barrier outranks the j^{th} barrier, here, 'i' indicates the row and 'j' indicates the column. For example, in row 2 of Table 5, barrier B2 outranks the barriers B7, B8, B10, B11, B12, and B13. It means that barrier B2 will be ranked higher than the barriers B7, B8, B10, B11, B12, and B13 for the importance. Similarly, all barriers are ranked. Based on the dominance matrix G, the ranks are provided to the barriers for their importance in sustainable operations of supply chains as shown in Table 6.

5. Discussion and implications

It has been observed that the barrier B4 that the lack of skilled workforce is ranked as the most important barrier in the way of achieving sustainable operations of supply chains. A skilled workforce is required to implement Industry 4.0 technologies and integrate them with the sustainability criteria of CE. Agyemang et al. (2019) observed that lack of expertize is one of the important barriers in the application of a circular economy but, they found it as 3rd important barrier. They have found the unawareness of employees is the top-ranked barrier, but in our analysis, it is a lower-ranked barrier. Ensuring that the employees are kept up to date with skills is a major driver for any new initiative, such as implementation of new technology in an organization. Jabour et al. (2018b) have also emphasized the use of technology for the

Table 4
Decision matrix (Rating of barriers for sustainable operations).

	WEIGHT	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
C1	**0.297**	4	4	5	5	4	5	3	3	3	4	4	3	3	4	3
C2	**0.131**	3	2	3	3	3	2	3	3	4	2	2	2	2	4	3
C3	**0.061**	3	3	3	3	2	3	2	5	2	3	4	2	2	3	2
C4	**0.091**	3	4	3	3	3	4	4	4	4	3	4	3	3	3	3
C5	**0.144**	3	4	4	5	3	5	4	4	4	3	4	3	3	3	3
C6	**0.160**	3	3	2	2	2	2	3	2	3	2	2	2	2	4	2
C7	**0.037**	2	2	4	2	3	2	2	3	2	3	3	2	2	2	2
C8	**0.047**	2	2	2	2	2	2	2	2	2	3	2	2	2	3	4
C9	**0.031**	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2

Table 5
Global Matrix (Dominance matrix).

	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15
B1	–	0	0	0	1	0	1	0	0	1	0	1	1	0	1
B2	0	–	0	0	0	0	1	1	0	1	1	1	1	0	0
B3	0	0	–	0	1	0	1	1	0	1	1	1	1	0	0
B4	0	0	0	–	1	1	1	1	0	1	1	1	1	0	1
B5	0	0	0	0	–	0	0	0	0	1	0	1	1	0	1
B6	0	0	0	0	1	–	1	1	1	1	1	1	1	0	1
B7	0	0	0	0	0	0	–	0	0	0	0	1	1	0	1
B8	0	0	0	0	0	0	0	–	0	0	0	1	1	0	1
B9	0	0	0	0	0	0	1	0	–	0	0	1	1	0	1
B10	0	0	0	0	0	0	0	0	0	–	0	1	1	0	1
B11	0	0	0	0	0	0	0	1	0	1	–	1	1	0	1
B12	0	0	0	0	0	0	0	0	0	0	0	–	1	0	0
B13	0	0	0	0	0	0	0	0	0	0	0	1	–	0	0
B14	1	1	0	0	1	0	1	0	1	1	0	1	1	–	1
B15	0	0	0	0	0	0	0	0	0	0	0	1	1	0	–

Table 6
Rank of the barriers for sustainable operations in the circular economy.

Barriers	Outranking Barriers	Rank
B1	B5, B7, B10, B12, B13, B15	4th
B2	B7, B8, B10, B11, B12, B13	4th
B3	B5, B7, B8, B10, B11, B12, B13	3rd
B4	**B5, B6, B7, B8, B10, B11, B12, B13, B15**	**1st**
B5	B10, B12, B13, B15	6th
B6	**B5, B7, B8, B9, B10, B11, B12, B13, B15**	**2nd**
B7	B12, B13, B15	7th
B8	B12, B13, B15	7th
B9	B7, B12, B13, B15	6th
B10	B12, B13, B15	7th
B11	B8, B10, B12, B13, B15	5th
B12	B13	9th
B13	B12	9th
B14	**B1, B2, B5, B7, B9, B10, B12, B13, B15**	**2nd**
B15	B12, B13	8th

implementation of CE. However, before using the new technology, awareness and training of the employees should be given priority.

The barriers B6 and B14, for example, ineffective performance framework and short-term goals, are jointly ranked as second important barriers. The performance framework should be designed in such a way that it evaluates the performance of the supply chains in the context of CE and Industry 4.0.

Govindan and Hasanagic (2018) have also considered the ineffective circular economy framework as a barrier. The framework should have detailed planning and policy of the management for the implementation of the circular economy. To make it effective, the management must review the plan and policy continuously as per the needs and feedback of the organization.

Most organizations focus on the short-term goal (B14), for example, earning profit by compromising product quality and process sustainability. The short-term goal of the management may be linked to the financial constraint of the organization. Financial constraint is one of the important barriers in the implementation of Industry 4.0 and CE. For longer survival of society and the country, a clear vision of economic and environmental sustainability is required. The government should frame policies to incentivize investment initiatives. The goal must be long term and include further expansion of the business activities incorporating the dimensions of CE.

The barrier B3 (ineffective strategy for integration of industry 4.0 with CE) is the third important barrier. In most cases, it has been observed that the purpose of industry 4.0 is to automate the processes for the optimization of their resource utilization and to increase productivity. Kirchherr et al. (2017) have also observed that a lack of policies for supporting the transition from a linear economy to CE is also a regulatory barrier. The new technology must be focused on the digitization of the processes and consider the sustainability dimensions of CE, for example, resource circularity.

The fourth important barriers for the sustainable operations of supply chains are B1 and B2 (Risk of misinvestment and insufficient legislation and control). The fear of the risk of misinvestment is due to a lack of vision from management and lack of awareness regarding the benefits of CE in the long-term. Insufficient legislation and control also lead to the reluctant attitude of the management towards sustainability.

The fifth important barrier is B11 (Employee resistance to change). Employees have the impression that new technologies may put their jobs at risk. To this end, employees should be made aware of long term benefits coming from such initiatives for the organization and be reassured that such initiatives will not pose any risk to their employment.

The resistance from the employees to such initiatives can be overcome by instigating appropriate motivation and encouraging employees to engage.

The remaining barriers are also important and need to be addressed properly. Most of these are the result of reluctance to instigate change by the management, lack of awareness of the employees, ineffective legislation and control of the supply chain activities. Therefore, management should involve its employees in all such initiatives for achieving goals of CE and Industry 4.0.

5.1. Practical implications

The findings of the study may help industry professionals to focus the efforts to digitize or automate the systems in the context of sustainability or resource circularity. Resource conservation is very important in the present context of the CE. It will also lead to saving the environment. The study also improves our knowledge of the various barriers in the sustainable operations of the entire supply chain. When the priority ranking of the different barriers are known, management should develop an effective action plan for sustainable operations in the present business environment of CE and Industry 4.0. Organizations need to integrate principles of CE and Industry 4.0 in their manufacturing processes for being competitive and sustainable. Performance frameworks should be designed considering the requirements of CE and Industry 4.0.

5.2. Theoretical implications

By highlighting the barriers to sustainable operations of the supply chain in context to Industry 4.0 and CE, researchers/academics may be motivated to explore and validate possible solutions to remove barriers in different contexts. The priority of different barriers will help academics and researchers to propose strategies for sustainable growth of organizations in the dynamic business environment. Researchers will also be able to explore the requirements of newer skills and knowledge as per the current requirement of the industry. The findings will also

help in proposing and validating a comprehensive performance framework for ensuring sustainable operations.

5.3. Implication for policymakers

Government policies play a crucial role in creating an environment for sustainable operations. Policies may differ from country to country and as such the findings of this study will be beneficial to the policymakers of developing economies like India so that they may formulate effective guidelines to promote the culture of CE and Industry 4.0. In developing countries, organizations are still reluctant to adopt sustainability and technological innovations. Policymakers should therefore formulate guidelines for honing the skills of their workforce to adopt new changes in the era of CE and Industry 4.0. The government should also think about incentivizing adoption of emerging technologies in operations. By adopting concepts of CE and Industry 4.0, organizations would be more sustainable and competitive in global markets.

6. Conclusion and future scope

In the present business environment of Industry 4.0 and CE, digitization of processes has become a critical success factor to remain competitive. To ensure sustainability of operations, digitization of processes requires linkage of Industry 4.0 technologies with CE dimensions. Organizations are facing many barriers that hinder exploiting the benefits of CE and Industry 4.0 and the biggest challenge for existing production systems is how to adopt these technological changes in an effective and integrated manner (Singh et al., 2019; Bhandari et al., 2019). Organizations need to adopt a holistic approach for handling these sustainability barriers in the present business environment.

In this study, nine-criteria of sustainability and fifteen barriers for sustainable operations are considered. An integrated approach comprising of AHP and ELECTRE has been used for ranking of sustainability criteria and different barriers. Resource circularity, increasing profits from green products, process designs that enable high efficiency are major sustainability criterion. It has been observed that major barriers in the way of sustainable operations of supply chains are a lack of skilled workforce, ineffective performance framework and short-term goals of an organization.

Most of the existing workforce is not familiar with the emerging technologies of Industry 4.0. Due to a lack of awareness and knowledge, organizations are not able to use them for meeting objectives of the CE. The short-term goals of management is another considerable barrier; it diverts managements' attention from integrating Industry 4.0 based technologies with sustainability criteria in performance evaluation. Organizations hesitate to make large monetary investments in industry 4.0 technologies and traditional performance frameworks are not relevant in the current business environment. Organizations should design a performance framework that considers the current business environment of Industry 4.0 and CE. Risk of misinvestment, insufficient legislation and control, and employees' resistance to change within the existing production system also work as barriers against the implementation of sustainability. Ranta et al. (2018) have observed that organizations have a fear factor regarding monetary investments in these technologies.

There are many emerging sustainable practices, such as life cycle assessment, design for disassembly/recycling, refurbishment/remanufacturing, design for modularity, bio-based materials, product labeling, energy management, community involvement and socially responsible consumption. To overcome barriers and achieve the objective of sustainable supply chain operations, organizations need to explore sustainable practices as per specific requirement. To ensure that these practices have maximum effect, judicious use of Industry 4.0 technologies should be thoroughly explored. As a future scope of the study, a framework for strategies to overcome barriers can be also proposed. This framework can be used to analyze impact of these strategies on the performance of an organization by using structural equation modeling or case studies.

Limitations: The analysis of this research is based on the opinion of five experts in the manufacturing industry in India. Therefore, findings of the study cannot be generalized because observations may differ from one industry to another. There may be minor variations in results due to changes in the experts' opinions. Additionally, the authors have taken the crisp value for analysis, which may not be able to capture vagueness in the decision making of the experts. The fuzzy approach may be applied in future research to validate the findings. For a generalization of the findings in the future, empirical studies may be also carried out across manufacturing organizations of different sectors.

Credit author statement

It is certified that all the three authors have contributed to this study. Introduction and LR have been done by Rajesh Kr Singh, Methodology and analysis have been done by Pravin Kumar, conclusion, implications etc. have been done by Vikas Kumar

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Authors are grateful to the editor and reviewers for their valuable comments and suggestions that have improved this study.

Appendix: Calculation steps in ELECTRE

(a) The barriers of sustainable operations of supply chains are rated with respect to the various criteria (sustainability parameters of CE) on a 5-point rating scale using Eq. (A1):

$$X = \left[x_{ij}\right] \tag{A1}$$

(b) The decision matrix prepared in step (a) is normalized using Eq. (A2):

$$R = \left[r_{ij}\right] \tag{A2}$$

$$\text{Where} \quad r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{m} x_{ij}^2}};$$

$i = 1, 2, \ldots n$ represents the specific criteria; and $j = 1, 2, 3, \ldots, m$ represents the specific barriers.

(a) Weighted normalized matrix is prepared using Eq. (A3).

$$V = [v_{ij}] \tag{A3}$$

Where $v_{ij} = r_{ij} \times w_i$

(b) The concordance and discordance sets of criteria are found for each barrier using the threshold as mentioned in Eqs. (A4) and (A5)

$$C_{kl} = \{I | v_{ik} \geq v_{il}\} \tag{A4}$$

$$D_{kl} = \{I | v_{ik} < v_{il}\} = I - C_{kl} \tag{A5}$$

Where C_{kl} is the set of concordance criteria and D_{kl} is the set of discordance criteria.

(c) The concordance matrix is prepared using Eq. (A6):

$$C = [c_{kl}], \text{and}$$
$$c_{kl} = \sum_{i \in C_{kl}} w_i \Big/ \sum_{i=1}^{n} w_i; \qquad \text{where} \quad 0 \leq c_{kl} \leq 1 \tag{A6}$$

Here, $c_{kl}$is a summation of the weights of concordance criteria for which the k^{th} barrier is preferred over the l^{th} barrier.

(d) The discordance matrix prepared using Eq. (A7):

$$D = [d_{kl}], \text{and}$$
$$d_{kl} = \frac{\max\limits_{i \in D_{kl}} |v_{ik} - v_{il}|}{\max\limits_{i \in I} |v_{ik} - v_{il}|}; \quad \text{where} \quad 0 \leq d_{kl} \leq 1 \tag{A7}$$

Here, the numerator shows the maximum difference distance between the weighted normalized elements of k^{th} and l^{th} barriers considering only those criteria for which the k^{th} barrier is inferior to the l^{th} barrier. Denominator shows the maximum difference distance between k^{th} and l^{th} barriers for all the criteria.

(a) The average value of c_{kl} is found using Eq. (A8)

$$\bar{c} = \sum_{k=1, k \neq l}^{m} \sum_{l=1, l \neq k}^{m} c_{kl} \,/\, m(m-1) \tag{A8}$$

(b) A Boolean Matrix E is prepared based on the values of c_{kl} and $\bar{c}$ as shown in Eq. (A9):

$$\begin{aligned} e_{kl} &= 1; \quad \text{for } c_{kl} \geq \bar{c} \\ &\quad 0; \quad \text{for } c_{kl} < \bar{c} \end{aligned} \tag{A9}$$

(c) The average value of d_{kl} is found using Eq. (A10)

$$\bar{d} = \sum_{k=1, k \neq l}^{m} \sum_{l=1, l \neq k}^{m} d_{kl} \Big/ m(m-1) \tag{A10}$$

(d) A Boolean Matrix F is calculated based on the values of d_{kl} and $\bar{d}$ as shown in Eq. (A11):

$$\begin{aligned} f_{kl} &= 0; \quad \text{for } d_{kl} \geq \bar{d} \\ &\quad 1; \quad \text{for } d_{kl} < \bar{d} \end{aligned} \tag{A11}$$

(e) The elements of the matrix E with the corresponding elements of matrix F are multiplied to get the aggregate dominance matrix G using Eq. (A12):

$$G = [g_{kl}]$$
$$\text{where} \quad g_{kl} = e_{kl} \times f_{kl} \tag{A12}$$

Based on the final dominance matrix, we can find the dominance of one alternative over another and we can map the outranking relationship of the

substitutes with respect to the other dominant substitutes.

References

Agyemang, M., Kusi-Sarpong, S., Khan, S.A., Mani, V., Rehman, S.T., Kusi-Sarpong, H., 2019. Drivers and barriers to circular economy implementation. Manag. Decis. 57 (4), 971–992.

Ball, C., Burt, G., De Vries, F., MacEachern, E., 2018. How environmental protection agencies can promote eco-innovation: the prospect of voluntary reciprocal legitimacy. Tech. Forecast. Soc. Change 129, 242–253.

Barnidge, E.K., Baker, E.A., Motton, F., Fitzgerald, T., Rose, F., 2011. Exploring community health through the sustainable livelihoods framework. Health Educ. Behav. 38 (1), 80–90.

Batista, L., Bourlakis, M., Smart, P., Maull, R., 2018. In search of a circular supply chain archetype – a content-analysis-based literature review. Prod. Plan. Control 29, 438–451.

Benyus, J.M., 2003. Biomimicry. HarperCollins, New York.

Bhandari, D., Singh, R.K., Garg, S.K., 2019. Prioritisation and evaluation of barriers intensity for implementation of cleaner technologies: framework for sustainable production. Res. Cons. Recy. 146, 156–167.

Braungart, M., McDonough, W., Bollinger, A., 2007. Cradle-to-cradle design: creating healthy emissions–A strategy for eco-effective product and system design. J. Clean. Prod. 15 (13–14), 1337–1348.

Dubey, R., Gunasekaran, A., Childe, S.J., Papadopoulos, T., Luo, Z., Wamba, S.F., Roubaud, D., 2017. Can big data and predictive analytics improve social and environmental sustainability? Tech. Forecast. Soc. Change. 144, 534–545.

Erol, S., Jäger, A., Hold, P., Ott, K., Sihn, W., 2016. Tangible industry 4.0: a scenariobased approach to learning for the future of production. Procedia CIRP 54, 13–18.

Figge, F., Thorpe, A.S., Givry, P., Canning, L., Franklin-Johnson, E., 2018. Longevity and circularity as indicators of eco-efficient resource use in the circular economy. Eco. Economics. 150, 297–306.

Foo, P.Y., Lee, V.H., Tan, G.W.H., Ooi, K.B., 2018. A gateway to realising sustainability performance via green supply chain management practices: a PLS–ANN approach. Exp. Sys. App. 107, 1–14.

Geng, R., Mansouri, S.A., Aktas, E., 2017. The relationship between green supply chain management and performance: a meta-analysis of empirical evidences in Asian emerging economies. Int. J. Prod. Eco. 183, 245–258.

Geng, Y., Doberstein, B., 2008. Developing the circular economy in China: challenges and opportunities for achieving' leapfrog development. The Int. J. Sust. Dev. World Eco. 15 (3), 231–239.

Geng, Y., Fu, J., Sarkis, J., Xue, B., 2012. Towards a national circular economy indicator system in China: an evaluation and critical analysis. J. Clean. Prod. 23 (1), 216–224.

Ghadimi, P., Wang, C., Lim, M.K., 2019. Resources, Conservation & Recycling Sustainable supply chain modeling and analysis: past debate, present problems and future challenges. Resour. Conserv. Recycl. 140, 72–84.

Giunipero, L.C., Hooker, R.E., Denslow, D., 2012. Purchasing and supply management sustainability: drivers and barriers. J. Purch. Supply Manag. 18 (4), 258–269.

Govindan, K., Hasanagic, M., 2018. A systematic review on drivers, barriers, and practices towards circular economy: a supply chain perspective. Int. J. Prod. Res. 56 (1–2), 278–311.

Govindan, K., Popiuc, M.N., Diabat, A., 2013. Overview of coordination contracts within forward and reverse supply chains. J. Clean. Prod. 47, 319–334.

Graedel, T.E., Allenby, B.R., 1995. Industrial Ecology. Prentice-Hall, Englewood Cliffs, NJ.

Haas, W., Krausmann, F., Wiedenhofer, D., Heinz, M., 2015. How circular is the global economy? An assessment of material flows, waste production, and recycling in the European Union and the world in 2005. J. Ind. Eco. 19 (5), 765–777.

Haddud, A., DeSouza, A., Khare, A., Lee, H., 2017. Examining potential benefis and challenges associated with the Internet of Things integration in supply chains. J. Manuf. Technol. Manag. 28 (8), 1055–1085.

Hawken, P., Lovins, A., Lovins, H.L., 2008. Natural capitalism: Creating the Next Industrial Revolution (Backbay).

Heck, S., Rogers, M., 2014. How Resource Scarcity is Driving the Third Industrial Revolution. Available at: https://www.mckinsey.com/business-functions/sustainability/our-insights/how-resource-scarcity-is-driving-the-third-industrial-revolution.

Hodgett, R.E., 2016. Comparison of multi-criteria decision-making methods for equipment selection. Int. J. Adv. Manuf. Technol. 85 (5–8), 1145–1157.

Hofmann, E., Rüsch, M., 2017. Industry 4.0 and the current status as well as future prospects on logistics. Comput. Ind. 89, 23–34.

Horváth, D., Szabó, R.Z., 2019. Driving forces and barriers of Industry 4.0: do multinational and small and medium-sized companies have equal opportunities? Tech. Forecast. Soc. Change 146, 119–132.

Li, H.F., Wang, J.J., 2007. An improved ranking method for ELECTRE III. In: 2007 international conference on wireless communications, networking and mobile computing, pp. 6659–6662. IEEE.

Jabbour, A.B.L., Jabbour, C.J.C., GodinhoFilho, M., Roubaud, D., 2018b. Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations. Ann. Oper. Res. 270 (1–2), 273–286.

Jabbour, A.B.L., Jabbour, C.J.C., Foropon, C., Godinho Filho, M., 2018a. When titans meet–Can industry 4.0 revolutionise the environmentally-sustainable manufacturing wave? The role of critical success factors. Tech. Forecast. Soc. Change. 132, 8–25.

Jabbour, C.J.C., Jabbour, A.B.L.S., 2016. Green human resource management and green supply chain management: linking two emerging agendas. J. Clean. Prod. 112, 1824–1833.

Jelinski, L.W., Graedel, T.E., Laudise, R.A., McCall, D.W., Patel, C.K., 1992. Industrial ecology: concepts and approaches. Proc. Nation. Acad. Sc. 89 (3), 793–797.

José, C., Jabbour, C., Mauricio, A.L., Lopes, A.B., 2017. Critical success factors and green supply chain management proactivity: shedding light on the human aspects of this relationship based on cases from the Brazilian industry. Prod. Plan. Control 28, 671–683.

Kahn, K.B., Barczak, G., Nicholas, J., Ledwith, A., Perks, H., 2012. An examination of new product development best practice. J. Prod. Inn. Manag. 29 (2), 180–192.

Kamble, S.S., Gunasekaran, A., Sharma, R., 2018. Analysis of the driving and dependence power of barriers to adopt industry 4.0 in Indian manufacturing industry. Comput. Ind. 101, 107–119.

Kang, D., Park, Y., 2014. Review based measurement of customer satisfaction in mobile service: sentiment analysis and VIKOR approach. Exp. Sys. App. 41 (4), 1041–1050.

Kazancoglu, Y., Kazancoglu, I., Sagnak, M., 2018. A new holistic conceptual framework for green supply chain management performance assessment based on circular economy. J. Clean. Prod. 195, 1282–1299.

Kiel, D., Arnold, C., Voigt, K.-I., 2017a. The inflence of the industrial internet of things on business models of established manufacturing companies – a business level perspective. Technovation 68, 4–19.

Kiel, D., Müller, J., Arnold, C., Voigt, K.-I., 2017b. Sustainable industrial value creation: benefis and Challenges of Industry 4.0. In: International Society for Professional Innovation Management (ISPIM). Vienna, pp. 1–21.

Kirchherr, J.W., Hekkert, M.P., Bour, R., Huijbrechtse-Truijens, A., Kostense-Smit, E., Muller, J., 2017. Breaking the Barriers to the Circular Economy. https://dspace.library.uu.nl/handle/1874/356517.

Kulatunga, A.K., Jayatilaka, P.R., Jayawickrama, M., 2013. Drivers and barriers to implement sustainable manufacturing concepts in Sri Lankan manufacturing sector. In: 11th Global Conference on Sustainable Manufacturing, pp. 171–176. https://doi.org/10.13140/2.1.2952.1927.

Kumar, R., Singh, R.K., Dwivedi, Y.K., 2020. Application of industry 4.0 technologies in SMEs for ethical and sustainable operations: analysis of challenges. J Clean Prod 275, 124063-124063.

Kumar, P., Singh, R.K., 2020. Selection of sustainable solutions for crop residue burning: an environmental issue in northwestern states of India. Envi. Dev. Sust. 1–35. https://doi.org/10.1007/s10668-020-00741-x.

Kumar, P., Singh, R.K., Kharab, K., 2017a. A comparative analysis of operational performance of Cellular Mobile Telephone Service Providers in Delhi working area using an approach of fuzzy ELECTRE. App. Soft Comput. 59, 438–447.

Kumar, P., Singh, R.K., Sinha, P., 2016. Optimal site selection for a hospital using a fuzzy extended ELECTRE approach. J. Manag. Anal. 3 (2), 115–135.

Kumar, P., Singh, R.K., Vaish, A., 2017b. Suppliers' green performance evaluation using fuzzy extended ELECTRE approach. Clean Tech. Env. Pol. 19 (3), 809-021.

Lanza, J., Sánchez, L., Muñoz, L., Galache, J.A., Sotres, P., Santana, J.R., Gutiérrez, V., 2015. Large-scale mobile sensing enabled internet-of-things testbed for smart city services. Int. J. Dist. Sen. Net. 11 (8), 785061.

Lee, J., Bagheri, B., Kao, H.A., 2015. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manuf. Let. 3, 18–23.

Li, G., Hou, Y., Wu, A., 2017. Fourth industrial revolution: technological drivers, impacts and coping methods. Chin. Geogr. Sci. 27 (4), 626–637.

Lieder, M., Rashid, A., 2016. Towards circular economy implementation: a comprehensive review in context of manufacturing industry. J. Clean. Prod. 115, 36–51.

Lin, K.Y., 2019. User experience-based product design for smart production to empower industry 4.0 in the glass recycling circular economy. Comput. Ind. Eng. 125, 729–738.

Lin, R., Tan, K., Geng, Y., 2013. Market demand, green product innovation, and firm performance: evidence from Vietnam motorcycle industry. J. Clean. Prod. 40, 101–107.

Linder, M., Sarasini, S., van Loon, P., 2017. A metric for quantifying product-level circularity. J. Ind. Eco. 21 (3), 545–558.

Liu, Y., Xu, X., 2017. Industry 4.0 and cloud manufacturing: a comparative analysis. J. Manuf. Sci. Eng. 139 (3), 034701.

Luthra, S., Mangla, S.K., 2018. Evaluating challenges to Industry 4.0 initiatives for supply chain sustainability in emerging economies. Process Saf. Environ. Protect. 117, 168–179.

Luthra, S., Garg, D., Haleem, A., 2015. Critical success factors of green supply chain management for achieving sustainability in Indian automobile industry. Prod. Plan. Cont. 26 (5), 1–24.

Luthra, S., Mangla, S.K., Yadav, G., 2019. An analysis of causal relationships among challenges impeding redistributed manufacturing in emerging economies. J. Clean. Prod. 225, 949–962.

Mittal, S., Khan Ahmad, M., Romero, D., Wuest, T., 2018. A critical review of smart manufacturing & Industry 4.0 maturity models: implications for small and mediumsized enterprises (SMEs). J. Manuf. Syst. 49, 194–214.

Moeuf, A., Pellerin, R., Lamouri, S., Tamayo-Giraldo, S., Barbaray, R., 2018. The industrial management of SMEs in the era of Industry 4.0. Int. J. Prod. Res. 56 (3), 1118–1136.

Moktadir, M.A., Ali, S.M., Rajesh, R., Paul, S.K., 2018. Modeling the interrelationships among barriers to sustainable supply chain management in leather industry. J. Clean. Prod. 181, 631–651.
Moldavska, A., Welo, T., 2017. The concept of sustainable manufacturing and its definitions: a content-analysis based literature review. J. Clean. Prod. 166, 744–755.
Mont, O., Plepys, A., Whalen, K., Nubholz, J.L.K., 2017. Business Model Innovation Fora Circular Economy Drivers and Barriers For the Swedish Industry – the Voice of REES Companies. https://lup.lub.lu.se/search/publication/833402ef-b4d4-4541-a10e-34d1e89d2146.
Moser, E., Huss, A.K., Liebrecht, C., Lanza, G., 2017. A portfolio theory approach to identify risk-efficient enablers of change in global production networks. Procedia CIRP 63, 768–773.
Muduli, K., Govindan, K., Barve, A., Geng, Y., 2013. Barriers to green supply chain management in Indian mining industries: a graph theoretic approach. J. Clean. Prod. 47, 335–344.
Müller, J., Maier, L., Veile, J., Voigt, K.I., 2017. Cooperation strategies among SMEs for implementing industry 4.0. In Digitalization in Supply Chain Management and Logistics: smart and Digital Solutions for an Industry 4.0 Environment. In: Proceedings of the Hamburg International Conference of Logistics (HICL), 23. epubli GmbH, Berlin, pp. 301–318.
Müller, J., Voigt, K.I., 2017. Industry 4.0—Integration Strategies for Small and Medium-sized Enterprises. In: Proceedings of the 26th International Association for Management of Technology (IAMOT) Conference. Vienna, Austria, pp. 14–18.
Nanry, J., Narayanan, S., Rassey, L., 2015. Digitizing the value chain: challenges remain for "Industry" 4.0, but the buzz is growing. In: McKinsey Quarterly. https://www.mckinsey.com/business-functions/operations/our-insights/digitizing-the-value-chain.
Nowicka, K., 2016. Cloud computing in sustainable mobility. Transp. Res. Proc. 14, 4070–4079.
Oesterreich, T.D., Teuteberg, F., 2016. Understanding the implications of digitisation and automation in the context of Industry 4.0: a triangulation approach and elements of a research agenda for the construction industry. Comput. Ind. 83, 121–139.
Papadopoulos, A., Giama, E., 2007. Environmental performance evaluation of thermal insulation materials and its impact on the building. Build. Envi. 42 (5), 2178–2187.
Paulraj, A., Chen, I.J., Blome, C., 2017. Motives and performance outcomes of sustainable supply chain management practices: a multi-theoretical perspective. J. Bus. Ethics. 145, 239–258.
Politis, J.D., 2001. The relationship of various leadership styles to knowledge management. Lead. Org. Dev. J. 22 (8), 354–364.
Pourjavad, E., Shahin, A., 2018. The application of Mamdani fuzzy inference system in evaluating Green supply chain management performance. Int. J. Fuzzy Syst. 20 (3), 901–912.
Prakash, C., Barua, M.K., 2015. Integration of AHP-TOPSIS method for prioritizing the solutions of reverse logistics adoption to overcome its barriers under fuzzy environment. J. Manuf. Syst. 37, 599–615.
Preston, F., 2012. A global redesign? Shaping the circular economy energy, environment and resource governance | March 2012. EERG BP, 2012/02.
Preston, M., Herron, J.P., 2016. Minerals and metals scarcity in manufacturing: the tickingtime bomb. In: PwC. Available at: http://www.pwc.com/gx/en/services/sustainability/publications/metal-minerals-scarcity.html.
PWC, 2014. Industry 4.0- opportunities and challenges of the industrial internet. https://www.pwc.nl/en/assets/documents/pwc-industrie-4-0.pdf.
Raj, A., Dwivedi, G., Sharma, A., de Sousa Jabbour, A.B.L., Rajak, S., 2019. Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: an inter-country comparative perspective. Int. J. Prod. Eco., 107546
Rajput, S., Singh, S.P., 2019. Connecting circular economy and industry 4.0. Int. J. Info. Manag 49, 98–113.
Ranta, V., Aarikka-Stenroos, L., Ritala, P., Mäkinen, S.J., 2018. Exploring institutional drivers and barriers of the circular economy: a cross-regional comparison of China, the US, and Europe. Res. Cons. Recy. 135, 70–82.
Roy, B., 1991. The outranking approach and the foundations of ELECTRE methods. Theo. Decis. 31 (1), 49–73.
Rust, R.T., Moorman, C., Dickson, P.R., 2002. Getting return on quality: revenue expansion, cost reduction, or both? J. Market. 66 (4), 7–24.
Saaty, T.L., 1980. The Analytical Hierarchy Process, Planning, Priority. Resource Allocation. RWS Publications, USA.
Sabaei, D., Erkoyuncu, J., Roy, R., 2015. A review of multi-criteria decision making methods for enhanced maintenance delivery. Procedia CIRP 37, 30–35.
Sarkis, J., 2003. A strategic decision framework for green supply chain management. J. Clean. Prod. 11, 397–409.
Sarkis, J., 2012. A boundaries and flows perspective of green supply chain management. Supply Chain Manag: Int. J. 17 (2), 202–216.
Sarkis, J., Gonzalez-Torre, P., Adenso-Diaz, B., 2010. Stakeholder pressure and the adoption of environmental practices: the mediating effect of training. J. Oper. Manag. 28 (2), 163–176.
Sauve, S., Bernard, S., Sloan, P., 2016. Environmental sciences, sustainable development and circular economy: alternative concepts for trans-disciplinary research. Envi. Dev. 17, 48–56.
Shanian, A., Savadogo, O., 2006. ELECTRE I decision support model for material selection of bipolar plates for polymer electrolyte fuel cells applications. J. New Mat. Elect. Syst. 9, 191–199.
Shao, Z., Feng, Y., Hu, Q., 2017. Impact of top management leadership styles on ERP assimilation and the role of organizational learning. Inf. Manag. 54 (7), 902–919.
Sharma, V.K., Chandna, P., Bhardwaj, A., 2017. Green supply chain management related performance indicators in agro industry: a review. J. Clean. Prod. 141, 1194–1208.
Shivajee, V., Singh, R.K., Rastogi, S., 2019. Manufacturing conversion cost reduction using quality control tools and digitization of real-time data. J. Clean. Prod. 237, 117678.
Singh, A., Mishra, N., Ali, S.I., Shukla, N., Shankar, R., 2015. Cloud computing technology: reducing carbon footprint in beef supply chain. Int. J. Prod. Econ. 164, 462–471.
Singh, R.K., Kumar, A., Garza-Reyes, J.A., de Sá, M.M., 2020. Managing operations for circular economy in the mining sector: an analysis of barriers intensity. Resour. Policy 69, 101752.
Singh, R.K., Luthra, S., Mangla, S.K., Uniyal, S., 2019. Applications of information and communication technology for sustainable growth of SMEs in India food industry. Res. Cons. Recy. 147, 10–18.
Stahel, W.R., Reday, G., 1976. The Potential For Substituting Manpower For Energy report to the Commission of the European Communities, Brussels.
Stahel, W., 2010. The Performance Economy. Springer.
Stock, T., Seliger, G., 2016. Opportunities of sustainable manufacturing in Industry 4.0. In: Proc. CIRP, 40, pp. 536–541.
Tao, F., Wang, Y., Zuo, Y., Yang, H., Zhang, M., 2016. Internet of Things in product lifecycle energy management. J. Ind. Inf. Integrat. 1, 26–39.
Thoben, K.D., Wiesner, S., Wuest, T., 2017. Industry 4.0" and smart manufacturing-a review of research issues and application examples. Int. J. Aut. Tech. 11 (1), 4–16.
Towler, G.P., 1996. Integrated process design for improved energy efficiency. Ren. Energy 9 (1–4), 1076–1080.
Tseng, M., Islam, S., Karia, N., Ahmad, F., 2019. A literature review on green supply chain management: trends and future challenges. Resour. Conserv. Recycl. 141, 145–162.
Tukker, A., 2015. Product services for a resource-efficient and circular economy-a review. J. Clean. Prod. 97, 76–91.
Turker, D., Altuntas, C., 2014. Sustainable supply chain management in the fast fashion industry: an analysis of corporate reports. Eur. Manag. J. 32, 837–849.
Vanalle, R.M., Ganga, G.M.D., Godinho Filho, M., Lucato, W.C., 2017. Green supply chain management: an investigation of pressures, practices, and performance within the Brazilian automotive supply chain. J. Clean. Prod. 151, 250–259.
Waibel, M.W., Steenkamp, L.P., Moloko, N., Oosthuizen, G.A., 2017. Investigating the effects of smart production systems on sustainability elements. Proc. Manuf. 8, 731–737.
Wang, S., Wan, J., Zhang, D., Li, D., Zhang, C., 2016. Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. Comput. Net. 101, 158–168.
Webster, K., 2015. The Circular Economy – A Wealth of Flows. Cowes: Ellen Mac Arthur Foundation Publishing, Isle of Wight.
Wu, K.J., Liao, C.J., Tseng, M.L., Chiu, A.S., 2015. Exploring decisive factors in green supply chain practices under uncertainty. Int. J. Prod. Econ. 159, 147–157.
Xu, L.D., Xu, E.L., Li, L., 2018. Industry 4.0: state of the art and future trends. Int. J. Prod. Res. 56 (8), 2941–2962.
Yadav, G., Luthra, S., Jakhar, S., Mangla, S.K., Rai, D.P., 2020. A framework to overcome sustainable supply chain challenges through solution measures of industry 4.0 and circular economy: an automotive case. J. Clean. Prod., 120112
Yeow, P.H., Sen, R.N., 2006. Productivity and quality improvements, revenue increment, and rejection cost reduction in the manual component insertion lines through the application of ergonomics. Int. J. Ind. Ergon. 36 (4), 367–377.
Young, R., Jordan, E., 2008. Top management support: mantra or necessity? Int. J. Proj. Manag. 26 (7), 713–725.
Younis, H., Sundarakani, B., Vel, P., 2016. The impact of implementing green supply chain management practices on corporate performance. Compet. Rev. 26 (3), 216–245.
Zahraee, S.M., Khalaji, A.M., Saidur, R., 2016. Application of artificial intelligence methods for hybrid energy system optimization. Renew. Sustain. Energy Rev. 66, 617–630.
Zeng, H., Chen, X., Xiao, X., Zhou, Z., 2017. Institutional pressures, sustainable supply chain management, and circular economy capability: empirical evidence from Chinese eco-industrial park firms. J. Clean. Prod. 155, 54–65.
Zhu, Q., Feng, Y., Choi, S.B., 2017. The role of customer relational governance in environmental and economic performance improvement through green supply chain management. J. Clean. Prod. 155 (2), 46–53.

Dr. Pravin Kumar is working as an Associate Professor in the Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi-110,042. He has more than 18 years of teaching and research experience. He obtained his Ph.D. degree in Supply Chain Management from IIT Delhi and M. Tech. in Industrial Management from IIT (Banaras Hindu University), Varanasi. His-research area is supply chain and operations management. He has authored a book on engineering economy and management, Wiley India Pvt. Ltd. and a book on Industrial Engineering and Management from Pearson Learning, India. He has also published more than 50 research papers in international journals and conferences. His-research articles have been published in the journals such as Applied Soft Computing; Annals of Operations Research; Clean Technology and Environmental Policy; Environment, Development and Sustainability; International Journal of System Assurance Engineering and Management, Journal of Modelling in Management, International Journal of Mobile Communications, International Journal of Business Excellence etc.

Email: pravin.dce@gmail.com; Mobile: +91–9,818,273,764

Dr. Rajesh Kr Singh is Professor in Operations Management Area at Management Development Institute (MDI), Gurgaon, India. His-areas of interest include Supply chain management, Industry 4.0, Circular Economy. Currently he is serving as Associate Editor

to International Journal of Consumer Studies (Wiley) and International Journal of Global Business and Competitiveness (Springer). He is also in Editorial Board of Journal of Supply Chain Management Systems and Vision-The Journal of Business Perspective. He has also worked as Guest Editor for special issues of Resources, Conservation and Recycling, Competitiveness Review, Journal of Modelling in Management, Administrative Sciences, Global Journal of Operations and Strategic Sourcing etc. He has about 180 research papers published in reputed international/national journals and conferences. He has published papers in journals such as Resources, Conservation and Recycling, Journal of Cleaner Production, Production planning and Control, Annals of Operations Research, International Journal of Production Research, Applied Soft Computing, Clean Technologies and Environmnet Policy, Industrial Management and Data Systems, Singapore Management Review, International Journal of Productivity and Performance Management, Benchmarking: An International Journal, Journal of Modelling in Management, Competitiveness Review: An International Business Journal.

Email-rksdce@yahoo.com, rajesh.singh@mdi.ac.in

Dr. Vikas Kumar is a Professor of Operations and Supply Chain Management and Director of Research at Bristol Business School, University of the West of England (UWE), UK. He is also a visiting Professor at Faculty of Accounting, Ton Duc Thang University, Ho Chi Minh City, Vietnam. He holds a PhD degree in Management Studies from Exeter Business School, UK and a Bachelor of Technology (first class distinction) degree in Metallurgy and Material Science engineering from NIFFT, India. He has published more than 200 articles in leading international journals and international conferences including the International Journal of Production Economics, International Journal of Production Research, Supply Chain Management: an International Journal, Expert System with Applications, Computers & Industrial Engineering, and Production Planning & Control. He serves on the editorial board for several international journals and has also acted as a guest editor for several leading journals. He has successfully secured funding in the excess of £1 million from various research agencies such as EPSRC, Innovate UK, British Academy, Newton Fund and British Council. His-current research interests include Sustainability, Industry 4.0, Circular Economy, Food Supply Chains, and Operational Excellence.

E mail- Vikas.Kumar@uwe.ac.uk; Vikas.Kumar@tdtu.edu.vn

Surface Topography: Metrology and Properties

PAPER

Mechanical and wear behaviour of Friction stir processed surface composite through Self-Assembled Monolayer Technique

To cite this article: Ravi Butola *et al* 2020 *Surf. Topogr.: Metrol. Prop.* **8** 045007

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 14.139.251.98 on 19/10/2020 at 07:57

Surface Topography: Metrology and Properties

PAPER

Mechanical and wear behaviour of Friction stir processed surface composite through Self-Assembled Monolayer Technique

Ravi Butola, Ranganath M Singari* and Qasim Murtaza*

Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

* Authors to whom any correspondence should be addressed.

E-mail: ranganath@dce.ac.in **and** qasimmurtaza@dce.ac.in

RECEIVED
9 June 2020

REVISED
13 September 2020

ACCEPTED FOR PUBLICATION
29 September 2020

PUBLISHED
6 October 2020

Keywords: friction stir processing, wear resistance, self-assembled monolayer, surface composite, micro-hardness.

Abstract

In the present research study, Friction stir processing (FSP) has been utilized to prepare nano surface composites, AA7075 based matrix was reinforced with B_4C nanoparticles (size <30 nm). The aim of this study is to form a thin layer of B_4C over the surface of AA7075 based matrix material through Self-Assembled Monolayer (SAM) technique followed by Friction stir processing. The major advantage of SAM is to minimize the quantity of B_4C nanoparticles used in the preparation of nano surface composites. Additionally, this research also investigates the effect of tool rotation speed of Friction stir processing on mechanical and wear properties of processed nano surface composite. The results observed a uniform dispersion of nanoparticles in the processed nano surface composite and an improved value of microhardness with maximum value was found to be 185 Hv of the sample processed at 1200 rpm, compared to base metal. For the constant load, as FSP tool rotation speed increases, wear resistance increases from 1000 to 1200 rpm and decreases slightly for 1400 rpm. Scanning Electron Microscope (SEM) micrograph, tensile test and Field Emission Scanning-Electron Microscope (FESEM) fractography image used to study microstructure and the mechanical properties of processed nano surface composite. The x-ray Diffraction (XRD) showed the presence of B_4C nanoparticles. Processed nano surface composites can be used for aircraft and automobile industry applications.

1. Introduction

Metal matrix nano surface composites are one of the upcoming fields involving advanced material engineering. One of the most commonly used metal matrix composites (MMCs) is of Aluminium. Aluminium matrix composites (AMCs) consist of properties that are distinct and preferred over aluminium alloys in many applications. Commonly used reinforcement particles are silicon carbide, boron carbide (B_4C), and zirconium silicate. MMC may be produced using many techniques like powder metallurgy, stir casting, compo casting, squeeze casting, Laser Surface engineering and so on. Nowadays, Aluminum alloys matrix nano surface composite reinforced with B_4C nanoparticles are being developed, which are extensively used in applications like aircraft and automobile industry, due to their lightweight, high strength to weight ratio and wear resistance.

Surface modification methods have also been proved to improve tribological performance of machine constituents during modification. It involves the deposition of a unique surface coating of tougher material such as TiN, CrN, etc. Surface hardening can be achieved by various methods such as thermal treatment, heat hardening or mechanical treatment, such as shot peening and thermal spray. Friction Stir Processing (FSP) is a solid-state plastic deformation method that may also be used to achieve surface hardening. It is also used for the advancement of mechanical, tribological and corrosion properties of metals and alloys. The surface specific properties that can be improved are hardness, corrosion resistance and wear resistance [1]. The relatively soft nature of aluminium alloys leads to greater wear, even at lower loads. Hence, surface hardening techniques are used to improve the wear resistance of aluminium alloy components.

© 2020 IOP Publishing Ltd

Several other routes are also possible such as, squeeze casting process [2], Stir casting [3–5], but the most economical way of refining the surface property of the material is preparing composites through Friction Stir Processing. This has the added advantage over other techniques that no residual stress is developed while it also helps the inclusion of new phase materials into the surface layer.

Various researchers have reported the utilization of methanol and acetone in the preparation of a surface composite of Aluminium alloys using Friction Stir Processing (FSP) [6–8], although in the present research work we are using boron carbide nanoparticles blend with Di-ethyl ether (DEE) [9]. Self-assembled monolayer (SAM) is generally used for the adsorption on different surfaces [10]. In the present research work we have used SAM technique for adsorption of boron carbide nanoparticles on the surface of AA7075 based matrix material. The optimization of adhesion property is a major advantage of using SAM. The surface composite using nanoparticles enhance the surface properties of Aluminium alloy matrix material. A reaction between Aluminum alloy and Di-methyl ether (DME) was presented by Torbjorn Fangstrom *et al* [11], using the comparative examination of density functional theory and EPR results. Density functional theory at B3LYP and BP86 levels helped setup and locate the stationary points by discussing the reaction between Al alloy and Di-methyl ether. Though both the methods provided the same results qualitatively, the barriers differed slightly because of the methods. SAM multilayer thin films, along with corrosion resistance, were investigated by Duhua Wang *et al* [12], using a four different alkyl-saline on the Aluminium alloy 2024-T3 substrate surface. It was observed from the results that the self-assembling of the alkyl-saline compounds was modified by the properties of the intricate surface of the base metal.

In recent years, FSP has gained a lot of popularity among researchers and a wide variety of research work has been done on the FSP. Friction Stir Processing (FSP) being a relatively new and emerging technology has also generated interest in the industry, like aerospace and automobile sector. Many well-established companies are researching and utilizing FSP technique to fabricate stronger composites. Mishra *et al* [13], published the first ever paper on Friction Stir Processing that practised the method to initiate a high strain rate in commercial aluminum alloys. It was established that superplasticity can now be achieved by FSP in a thick section aluminum sheet. Chaudhary *et al* [14], studied the consequences using FSP on different alloys. The difference in the properties of Friction stir processed alloys was observed as the processing was executed at different angles and speeds.

Nowadays research done on Friction Stir Processing involves studying the effects of various process parameters on the homogeneous dispersion of reinforcement or the mechanical and tribological behaviour of surface composites. It has been found in literature, that resultant composites generally show better microstructure and mechanical properties. Elangovan *et al* [15], investigated the influence of axial force and tool pin shape of the five pin profiles used, square tool pin produced defect-free, good quality FSPed zone, irrespective of applied axial force. Mahamani [16], investigated that the tribological behaviour of AA6061-TiB_2/ZrB_2 *in situ* composite using a pin on disc setup. It was concluded that the effect of reinforcement ratio, normal load and sliding speed on wear properties of fabricated composite. It was also found that the test condition and reinforcement ratio offer significant effect on wear behaviour of composite. Bisadi *et al* [17], studied the influence of tool rotational speed and a traverse rate of FSP on particle distribution and microstructural enhancement. It was found that with the increase in speed of rotation a more uniform dispersion of TiB2 particles could be obtained. Even the microhardness values of hence prepared composite increased by rising rotational and traverse speeds. The tensile test resulted in the increased value for yield strength by two times compared to the base metal. Adilokht *et al* [18], have reported the microstructural and wear behaviour of a hybrid surface composite formed by SiC and MoS_2 particles embedded into a matrix metal of an Al alloy A356. The composite was produced through FSP. Advancement in wear resistance was observed in the sample produced. Within the nugget zone a uniform dispersion of secondary particles was observed by the micrographs while a MoS_2 rich mechanically mixed layer (MML) on the top of the worn out a surface. Due to severe plastic deformation and it improve tribological behaviour of hybrid composite. Mardi *et al* [19], investigated the wear characteristics of Mg-6 Al alloy reinforced with alumina nanoparticles. It has been concluded that wear rates show a decline with decreasing sliding speeds irrespective of loading. Aruri *et al* [20], have reported the influence of tool rotational speed on mechanical and wear behaviour of aluminum based surface hybrid composites manufactured by FSP. The microhardness value showed a fall in value an increment in rotational speed while an increase in hardness of Al–SiC/Al_2O_3 prepared specimen was an outcome of pinning effect of a hard SiC and Al_2O_3 particles. Ramesh *et al* [21], investigated a method to increase the surface hardness of a surface composite using FSP. The matrix material was AA7075-T651 and the reinforcement used was B_4C. The average hardness of FSPed samples was observed to be more than that of its base value by 50%. The increment in hardness value was observed as a result of a fine scattering of reinforcement particles and the fine grain size of the matrix metal. Rana *et al* [22], discussed the key points in manufacturing of AA7075-T651/B_4C surface composite using different combinations of tool rotational speed, a tool traverse rate and number of passes. The average hardness of the

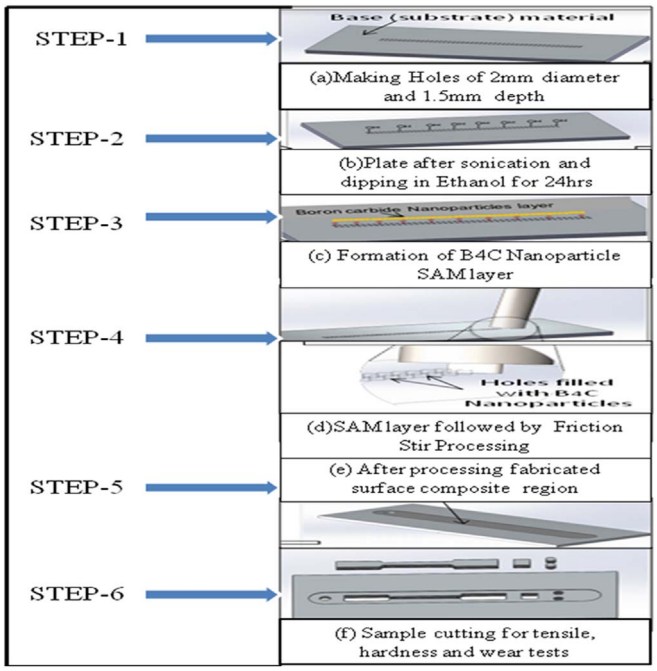


Figure 1. Schematic representations of SAM technique followed by Friction stir processing.

FSPed sample was higher than that of base metal by 1.5 to 2 folds. Resistance to wear almost doubled. In another research Panneerselvam *et al* [23], focuses on tribological properties of hybrid composite fabricated using a stir casting technique synthesised with B_4C and zirconium silicate. Composite samples were prepared using three different wt.% 3%, 6% and 9%. Optical micrographs and SEM were employed to investigate wear behaviour and wear mechanism mode.

Abrahams *et al* [24] showed improved mechanical properties of 5005-H34 and AA7075-T651 using different tool profile and investigated properties such as microhardness and microstructure under different process parameters. It was observed that increasing traverse speed caused an increase in hardness and mechanical properties as the precipitate free zones reduced at the grain boundary. Traverse speed was found to have more influence than rotational speed on microhardness and microstructure. Kurt *et al* [25] prepared composite using SiC particles paired with pure aluminium following Friction Stir Processing. The results presented as a trend of higher rotating and traverse rates leading to a better dispersion of SiC particles. The microhardness of fabricated surface composite was increased by three times as compared to that of base metal. Shamsipur *et al* [26] investigated Ti/SiC surface nano-composite layer, prepared by FSP technique over Ti plates by incorporating SiC particles into it. A uniform scattering of nano particles was found only after double pass FSP. Yuvaraj *et al* [27] fabricated 5083 aluminium alloy with reinforced layers of B_4C (micro and nano sized particles) by using friction stir processing. It was observed that the wear properties of the Al5083 alloy were enhanced by the addition of the nano B_4C particles in comparison with micro B_4C particles and the wear resistance of a nano surface composite layer was greater than that of without reinforced Al5083 alloy.

Using high concentrations of B_4C nanoparticles is not feasible, moreover, it has been reported that for using B_4C as reinforcement in metal matrix composites the best results were observed using 6–8 wt% Harichandran *et al* [28]. Increasing the reinforcement concentration beyond this resulted in a decrease in wear resistant and tensile strength properties. Thus, the novelty of the present research is to minimize the utilization of nanoparticles reinforcement materials that causes obtained better properties in comparison to other reinforced technique such as groove filling and a paste applying method. A more new technique prominent in the research field for polymers and chemicals is the formation of Self-Assembled Monolayers (SAM) has been used for fabrication of nano surface composite. A Self-assembled monolayer is the deposition of a single layer of molecules on a chemically active solid surface reported by Bommarito *et al* [29]. The head group is the part which is bonded over the substrate surface while the tail group forms the chain of a SAM film explained by Smith *et al* [30]. Hence, there has been no such report of the formation of SAM on the surface of Al alloy as matrix material followed by FSP technique to fabricate nano surface composite.

The main aim of the current research work is the formation of a thin layer over the surface of AA7075 based matrix material and reinforced boron carbide nanoparticles through Self-assembled monolayers (SAM) technique followed by Friction Stir Processing

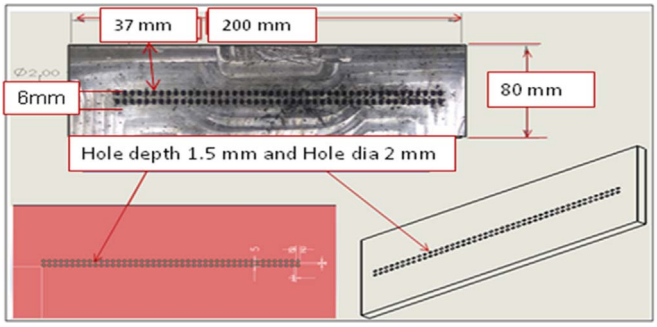


Figure 2. Holes over the surface of AA7075-T6 with dimensions.

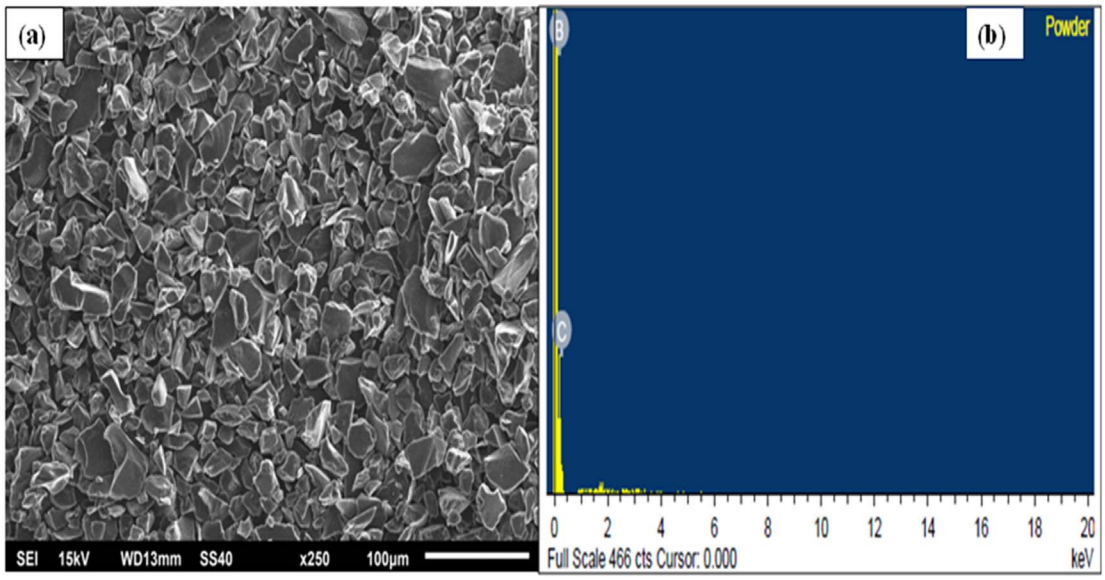


Figure 3. SEM and EDX of Boron carbide nanoparticles.

Table 1. The Chemical composition of AA7075-T6.

Alloy	Zn	Mg	Cu	Fe	Mn	Ti	Cr	Si	Al
Wt.%	2.273	2.4	1.3	0.18	0.04	0.08	0.18	0.06	Balance

to prepare nano surface composite and obtained the better mechanical and wear properties.

In figure 1(a) Holes on the base (substrate) material of surface followed by (b) After sonication of base (substrate) material and dipping in methanol for 24 h (c) A formation of a boron carbide nanoparticle SAM layer on the holes (d) The fabrication of the nano surface composite using FSP method (e) Friction stir processed plate (f) Sample cut out from the FSPed plate for testing.

2. Experimentation procedure

2.1. Material and method

In this study, AA7075-T6 is selected as matrix material and the matrix (base plate) dimension was (200 × 80 × 6) mm. At first two lines of 45 holes each, was machined utilizing CNC vertical milling machine on a center of the base plate, making an aggregate of 2 lines and 90 holes. The length of each line was 145 mm. The dimension of the holes was 2 mm dia and 1.5 mm deep as shown in figure 2. The separation between the two nearby holes was 2 mm. B_4C nanoparticles of 99.9% purity with the particle size of (<30 nm) were used as reinforcement. The SEM micrograph of the nanoparticles and Elemental composition are shown in figure 3. The SEM micrograph reveals the morphology of nanoparticles. The chemical composition of the substrate material is shown in table 1. Further processing of material gets halted due to foreign undesired particles and grease so sonication process had to be performed Ayad *et al* [31]. To perform the sonication a laboratory detergent was

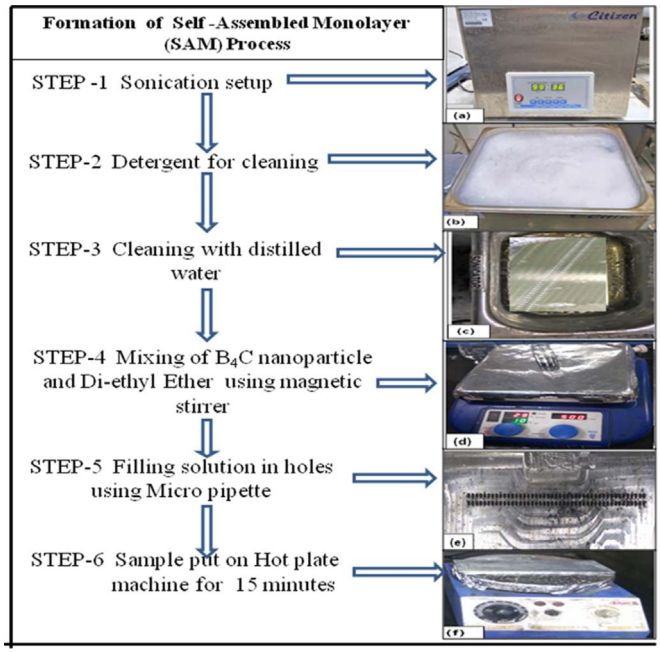


Figure 4. Formation of Self-assemble monolayer process (a) Sonicator setup (b) Detergent for cleaning (c) Cleaning with distilled water (d)Solution of B_4C in Di-ethyl Ether (e) Filling of a solution in holes using Micro-pipette (f) Sample put on a hot plate machine setup.

used and poured in the sonication bath along with the plates. After performing 30 min of sonication with detergent sonication was done using distilled water.

The Self-assembled monolayer (SAM) formation had to be performed on sonicated plates but before that the surface had to be active to adsorption the substrate. For this we kept the plates in an ethanol solution overnight Awad *et al* [32]. In figure 4(b) sonication process has been completed with a 5% chemical detergent solution in distilled water. In figure 4(a) the sonicator was switched on for periods of 30 min and left to soak for 15 min intervals, repeated twice. In figure 4(c) using distilled water with similar cycles of sonication process and repeated thrice. In figure 4(d) we took 0.276 gm of B_4C in a 10 ml of Di-ethyl ether (DEE), the solution was kept slightly dilute for easy flow into the holes, for constant stirring to form a colloidal solution a magnetic stirrer was used**.** After the plates with minimum moisture exposure were brought for SAM formation. To prevent precipitation of particles continuous stirring of the solution is required. After this, we put the prepared sample as shown in figure 4(e) on a hot processing stage at around a 150 °C temperature to bind the nanoparticle boron carbide materials in holes of the plate for 15 min shown in figure 4(f).

The FSP was performed in two stages. In the primary stage, for covering the slits a tool not equipped with a pin had been put to use, this mixed the aluminum matrix and prevented the particles from rushing outwards during the actual process. The force value during the capping pass reached a maximum value of 600 N. For the subsequent stage, a set of values for process parameters having significance according to previous research work had been put to use tabulated in table 2. Traverse speed being constant for the processes was chosen a value of 40 mm min^{-1}. The tool was made of H13 Tool steel and a hardness of 60–65 HRC. It was a two sided tool pin with a complete length of 117 mm. The diameter of the tool shank was 19.95 mm. The length of the tool pin (triangular in shape) was 3.5 mm while it's side measured 6.25 mm similar types of tool pin profile were also used in a previous report [33, 34]. After the capping passes using a pinless tool to cover the nanoparticle layer, and then the tool with the triangular pin was used to provide the stirring pass. Only a single stirring pass was performed over each set of twin adjacent rows of holes.

Table 2. Process parameters of FSP.

Process parameters	Value taken
Tool shoulder	Flat
Tool shoulder dia	19.95 mm
Tool pin profile	Triangular
Tool pin length	3.5 mm
Tilt angle	2° constant
Tool rotational speed (rpm)	1000, 1200, 1400
Reinforcement	B_4C (<30 nm)

As appeared in figure 5(a) shows the FSP experimental setup and in figures 5(b), (c) shows the FSP tool dimension and schematic design of the FSP, wherein the shoulder dia across an element of sheet thickness has been looked into for Friction stir processing arrangements, the pattern utilizing least-square surmised shows the shoulder breadth as ~2.2 times the

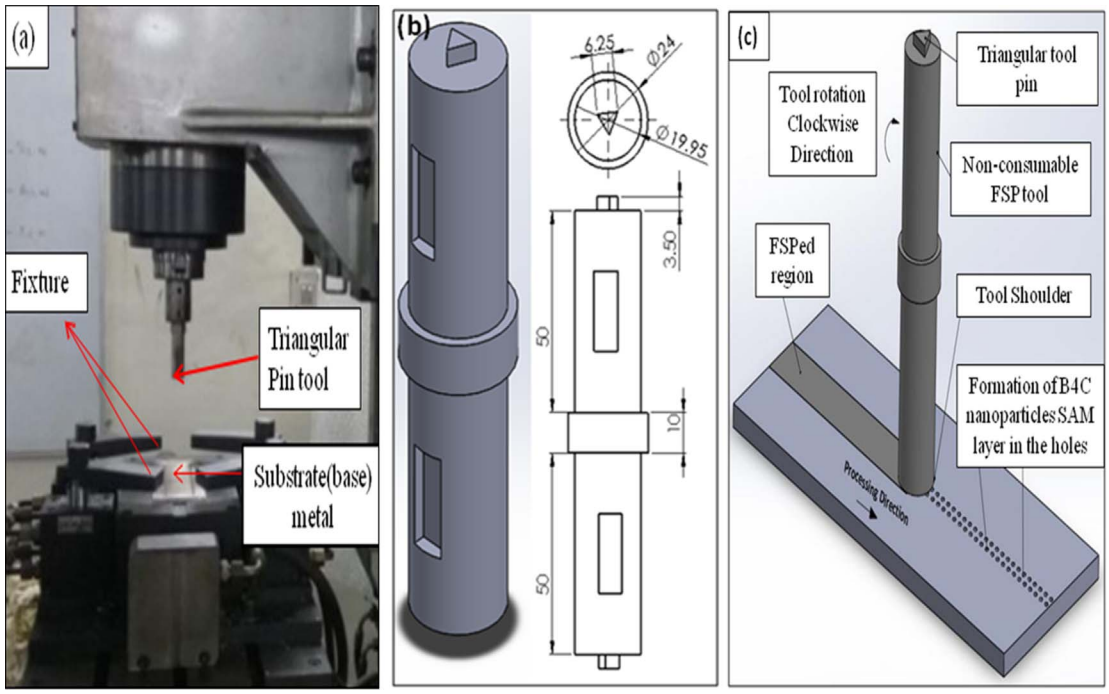


Figure 5. (a) FSP Machine setup (b) Tool dimensions (C) A schematic representation of FSP.

sheet thickness in addition to a consistent of 7.3 mm has been reported by Liu *et al* [35]. It is notable that with increasing thickness, more vitality input is required and in this manner, a bigger shoulder breadth is important to generate the heat.

2.2. Wear test experimentation

Wear tests performed in this research work consisted of a pin on disk tribometer setup (DUCOM) as per ASTM G99-04 standard. In figure 6 schematic designs (a) A front view (b) A top view shows the exact picture of the experimental Pin on disc setup. Wear samples were cylindrical pin shape with 10 mm diameter cut by wire EDM from the processed zone. The disc was EN-24 steel hardened to 60HRC. Sliding speed was taken 1 m s^{-1} and normal loads varying from 20, 30, 40 and 50 N were preferred for conducting the test with a sliding distance of 1000 m. A constant track diameter 30–60 mm was utilized in all the wear tests. The circular samples were weighed by an electronic weighing machine and disc roughness is approximately 0.2 μm. In table 3 present the Pin on Disk test condition. The worn surface composite samples were subjected to SEM analyses.

3. Results and discussion

3.1. Microhardness

Hardness tests were performed on all the processed and unprocessed samples. Six readings were taken for each sample. A Vickers scale was utilized for the hardness. Hardness is of two kinds, it is called macro-hardness when the testing is performed under load states of above 10 N. At the point when it is performed underneath 10 N, it is called micro-hardness. The load taken was 0.300 N and dwell time was 20 s. The average hardness was determined to utilize six readings on a similar load. The apparatus utilized for the testing was Fischer's microhardness tester, which uses the Vickers scale for measurement. The normal and the extreme hardness for the surface composites have been seen as more than that of the base metal. The disparity in microhardness is superfluous at ambient temperature as compared to a higher temp. The reduction of microhardness of the given tool is because of additional tool wear at elevated temperatures [36]. When the fabrication of nanocomposite using TIO_2 as reinforcement by friction stir processing was performed under an ambient environment, a slight nanoparticles agglomeration was obtained at 1400 rpm, due to heat input exceeding and causing more tool wear and insufficient material flow and strain rate. Therefore, the performance of the tool shows decrease with consecutive increase in tool rotational speed has been explained elsewhere [37, 38]. In figure 7, represents the comparison of hardness values for the nano surface composite samples and the base metal. The pattern is to such an extent that hardness increases with an increase in tool rotational speed until 1200, after which it shows a diminishing pattern. At higher tool rotational speed causes more grain growth during dynamic recrystallization and decreases the microhardness. Further particle distribution and spacing between the particles affects the microhardness of material as on increasing rotational

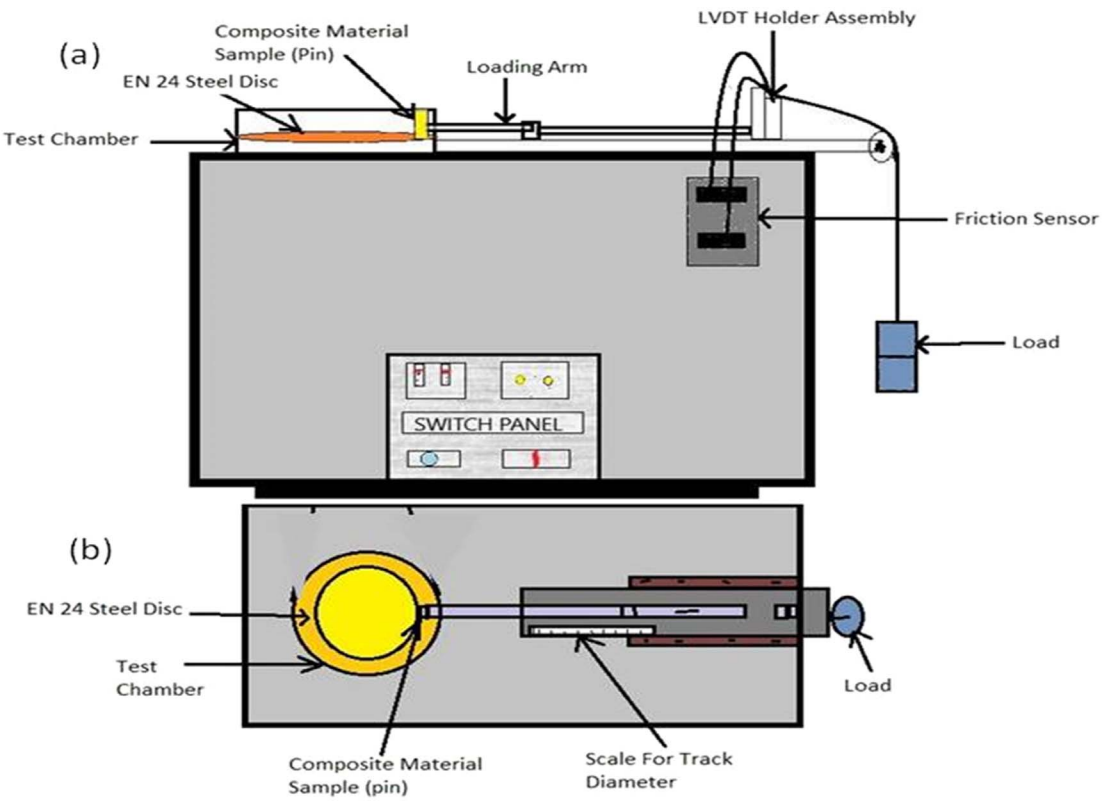


Figure 6. High Temperature Pin on disc setup (a) A front view (b) A top view.

Table 3. Pin on Disk test condition.

S. N	Test condition	Value taken
1.	RPM	319, 382, 478, 637
2.	Load (N)	20, 30, 40, 50
3.	Track diameter (mm)	30–60
4.	Pin diameter, Height (mm)	10
5.	Disc diameter (mm)	100
6.	Disc surface roughness (Ra)	0.2 μm
7.	Sliding distance (m)	1000
8.	Temperature	30°

speed the inter particle distance increases leading to decrease in microhardness [39].

The maximum average hardness was observed to be 185 Hv, which was obtained from the sample processed at 1200 rpm. This was followed by 182 Hv, obtained from the sample processed at 1400 rpm, and then 170 Hv which was obtained from the sample processed at 1000 rpm. The least average hardness for obvious reasons was of the base alloy, numerically of the order of 102 Hv. Due to the uniform distribution of nanoparticles at tool rotational speed 1000 to 1200 rpm, results obtained increases in microhardness and wear properties, but after tool rotational speed 1200 to 1400 rpm microhardness and wear properties has slightly decreased due to bad material flow and nanoparticles agglomeration. When the heat input exceeds a certain level and causes more tool wear and insufficient material flow, so tool wear is related to both tool geometry and high tool rotational speed. Similar results have been reported by Molla Ramezani *et al* [37].

3.2. SEM micrograph analysis

After the fabrication of the nano surface composite using FSP, the samples were cut in the perpendicular direction from the processed and unprocessed region and then polished in different grades of emery paper. The nano surface composite samples cut orthogonal to the processed region were polished following dry and wet polishing techniques. Scanning Electron Microscope (SEM) model (JSM-6510LV, JEOL JAPAN) was used to examine the homogeneous distribution of reinforcement particles. In figure 8(a) represents the SEM micrograph of the base metal surfaces and in figures 8(b), (c) shows the micrograph of the FSPed nano surface composite sample. The homogeneous distribution of nanoparticles can be seen in the stir zone of the processed region in the SEM micrograph. Sharp geometries like square, triangle, and threaded pin tool wear cause more tool wear and cause insufficient material flow at the higher rotational speed and some other process parameters cause more tool pin wear, due to bad material flow the agglomeration of the nanoparticles as shown in figure 8(d). One of the authors investigated the tool pin wear of several rotational speed and constant traverse speed and he had concluded that minimum tool pin wear occurred with lower tool rotational

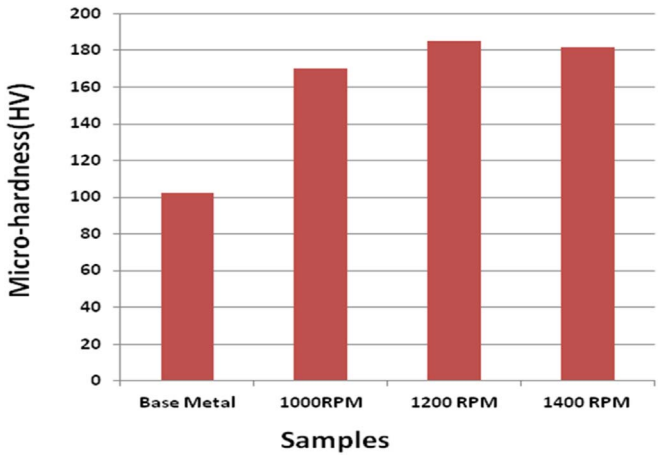


Figure 7. The Micro-hardness profile of the base metal and FSPed sample at different rpm.

speed. At the lowest tool rotational speed, the tool pin wear was nearly one-fourth of the higher tool rotation speed. It has been reported by Fernandez and L E Murr [40]. This ensures improvement in microhardness and other characteristics compared to the base metal. The light gray particles can be seen in the processed samples in figures 8(b)–(d) are B_4C nanoparticles as a similar result has been explained by Butola *et al* [41]. He had performed the same experiment in the aluminium alloy as a matrix metal and made (2 × 1.5) mm long groove, then he had observed a significant increment in microhardness and grain refinement [41]. The main aim of the formation of the Self-assembled monolayer (SAM) technique is to minimize the concentration of nanoparticles and obtained better properties as compared to the other technique. In figures 9(a), (b) shown the EDX results of the base metal and the FSPed sample at the stir zone. EDX result shows the presence of B_4C nanoparticle in figure 9(b).

3.3. Evaluation of the tensile test

Friction stir processing was carried out on the AA7075-T6 at three different tool rotation speed viz. 1000 rpm, 1200 rpm, and 1400 rpm. After the fabrication of the nano surface composites, the samples were cut orthogonal to the processed region using a wire EDM setup. The tensile sample prepared according to the ASTM standard E8M and the sample before and after the tensile test is present in figures 10(a)–(c). As the rotational speed of the FSP was incremented, it was observed that the microhardness value of the sample increased as shown in figure 11. Subsequently, a decrease in the percentage elongation of the sample was observed. Yuvaraj *et al* [27] have reported that percentage elongation increases when FSP was performed without particle one pass. However, the presence of hard nanoparticle in the current study contributes to decreasing the ductility of the nano surface composites. The Ultimate tensile strength (UTS) of the nano surface composite samples also decreased in a similar manner due to the uniform distribution of B_4C nanoparticles, which increases the microhardness, causes a decrease in grain size in the surface composite. These results are supported by Butola *et al* [41].

Suri *et al* [42] have reported the effect of FSP on the mechanical properties of Al2024 alloy at 1200 rpm and 1800 rpm. The results pertaining to the increase and decrease in microhardness and UTS with an increment in the tool rotation speed are in-line with the ones seen in our study. Balamurugan and Mahadevan [43] investigated the effects of process parameters on the mechanical and corrosion behaviour of friction stir-claded AZ31B-magnesium alloy. They observed a uniform increase in microhardness of the FSPed sample at 500 rpm, 710 rpm, and 1000 rpm. Moustafa [44] investigated the impact of various tool rotational speeds on AA2024/Al_2O_3 and concluded that the rotational tool speed of 900 rpm improved the tensile strength after that increasing tool rpm, it shows a lower the tensile strength, and a number of passes play a significant role in improving the microhardness of the sample. The effect of tool rotation speed on the UTS of a friction stir weld stir cast AA6061–T6/AlNp composite was reported by Murugan and Kumar [45] who inferred that as a tool rotation speed increases the strength of the weld decreases and hence the value of UTS also decreases.

3.4. Fractographs analysis after a tensile test

Field emission scanning electron microscope (FESEM) model (ZeissSigma300) was used to examine the fracture point of a base metal and FSPed samples processed at different tool rotational speed. After the tensile test was carried out in UTM setup, the next stage was to cutting the sample according to the dimension of the FESEM arrangement. In figure 12(a) FESEM fractography image with 10 μm shows that the morphologies of the fracture point of base metal has a

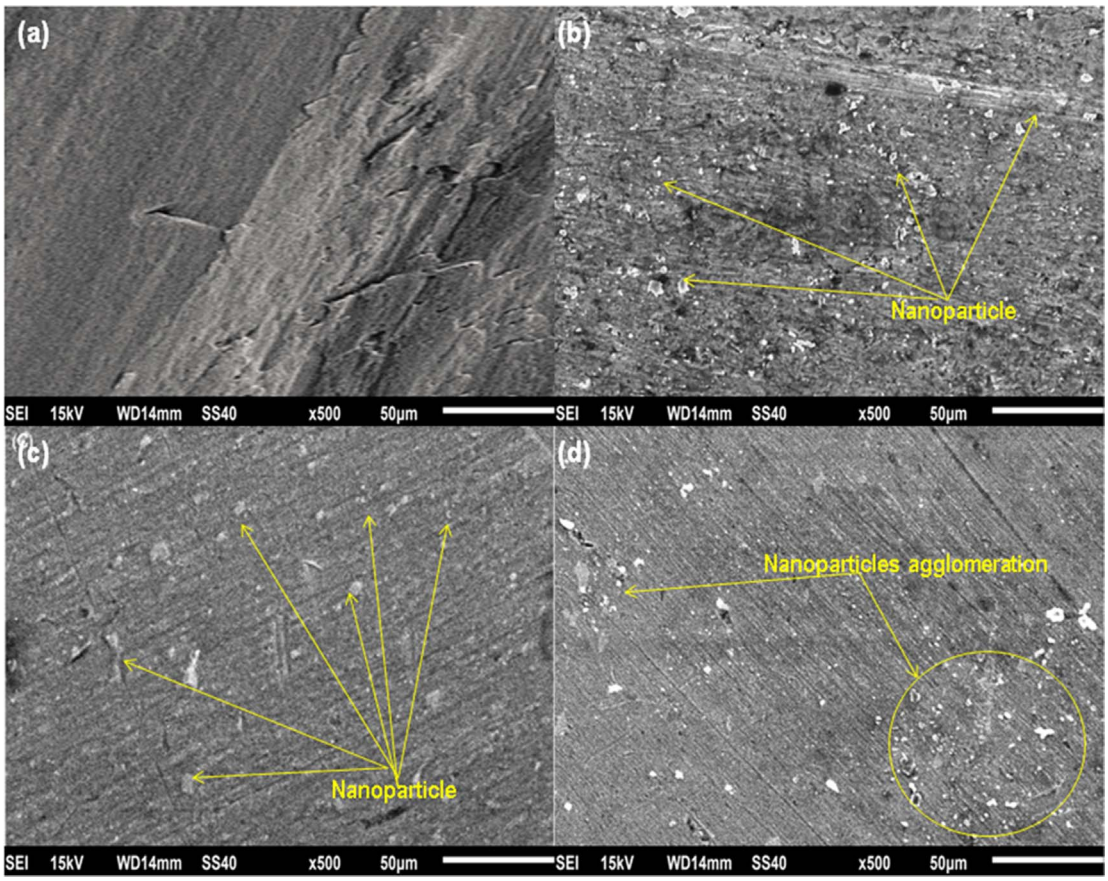


Figure 8. SEM images of (a) substrate (base) metal and, (b) FSPed sample at 1000 rpm (c) FSPed sample at 1200 rpm (d) FSPed sample at 1400 rpm.

tiny number of dimples which indicate the maximum elongation percentage. In figure 12(b) shows the fractography image with 20 μm of the fracture point of the FSP sample processed at 1000 rpm. It shows the large dimples and cleavage facets which increase the microhardness of the nano surface composite. In figure 12(c) FESEM fractography image with 10 μm shows the large dimples and a wide crater, which increases the microhardness of nano surface composite processed at 1200 rpm. In figure 12(d) the large dimples and crack formation occurred due to brittle nature of the nano surface composite are shown, which decrease the strength and ductility of the FSP sample processed at 1400 rpm.

3.5. XRD analysis

X-ray diffraction (XRD) analysis was used to determine the presence of crystalline phase into the material. X-ray diffraction (BRUKAR D8 ADVANCE) was utilized to identify the phase of B_4C nanoparticles into a nano surface composite. The XRD pattern obtained is shown in figures 13(a), (b). In figure 13(a) shows the XRD pattern of the substrate (base) metal sample. It clearly shows the presence of B_4C nanoparticles in the XRD pattern analysis of the FSPed sample in figure 13(b). The B_4C nanoparticles were shown between 40–45° at 2 theta angle. In FSPed the nano surface composite nanoparticles can be clearly seen in the SEM micrograph also, it is also supported by a XRD pattern. The highest Al peak was present and further it was shifted towards increasing the angle. A lower peak were shown in figure 13(b) at 40–45 degree, which was observed due to less concentration of nanoparticle using Self-assembled monolayers (SAM) technique followed by FSP method. It was observed that only two phases are present, i.e. Al and B_4C and no other diffraction peaks were detected before and after processing. On the other hand, Al peak intensity was increased by Friction stir processing of nano surface composite.

3.6. Wear analysis

Samples prepared by FSP technique have been exposed to wear tests to distinguish the wear resistance capacities of the nano surface composite samples and compare them with the base material just as one another to recognize the optimum tool rpm represented in figure 14. The samples have additionally been exposed to various load conditions.

For each of the loads i.e. 20 N, 30 N, 40 N and 50 N, the maximum wear has been observed in the substrate (base) metal followed by the FSPed sample at tool speeds of 1000 rpm, 1400 rpm and 1200 rpm. This means that wear resistance is a maximum for the

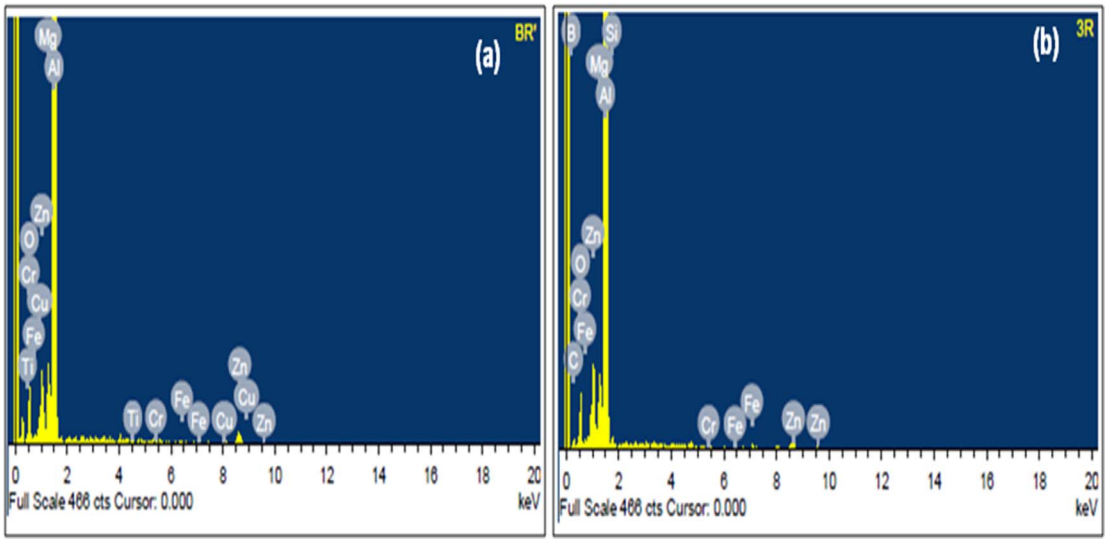


Figure 9. EDX results of (a) substrate (base) metal and, (b) The FSPed sample.

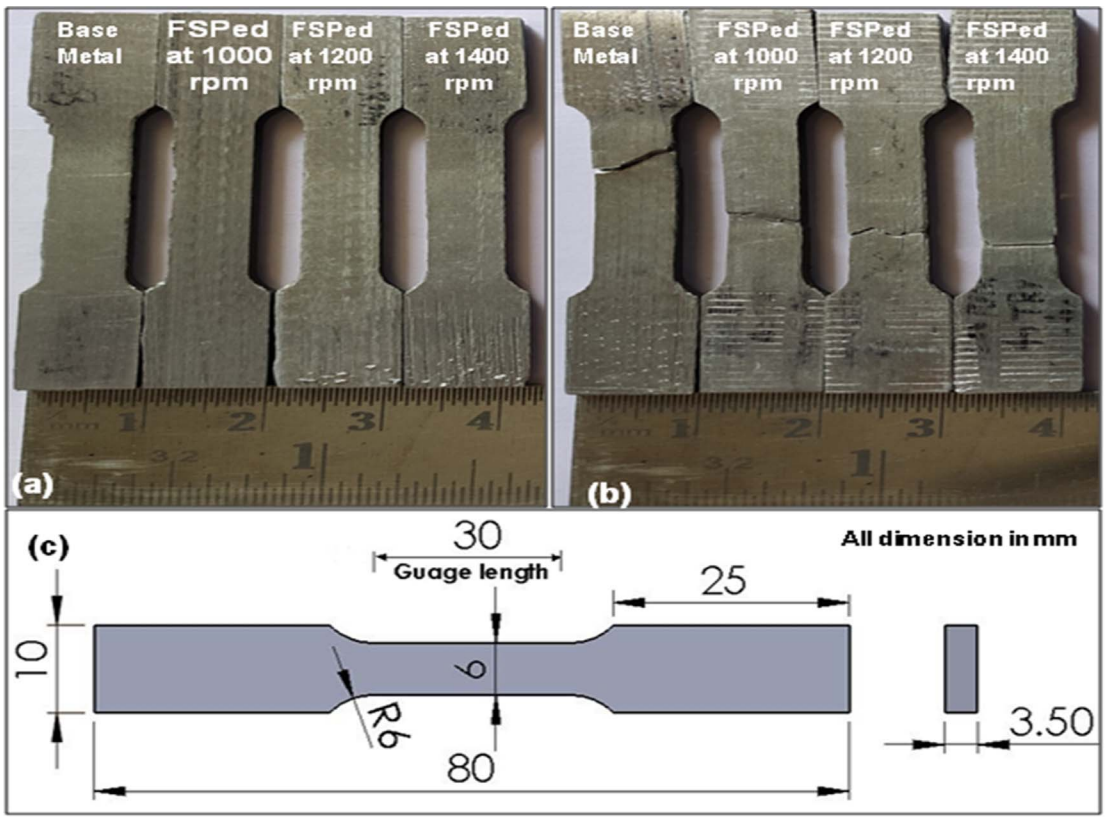


Figure 10. The tensile test sample of (a) Base metal and the FSPed sample before fracture, (b) The fracture a sample after tensile test and (c) Dimension of the Sample as per ASTM standard.

FSPed sample at a tool rotation speed of 1200 rpm, followed by the FSPed sample at 1400 rpm and thereafter 1000 rpm. This follows the same pattern as the values from microhardness, which explains this phenomenon. As load increases wear increases for the same material. Wear resistance has improved by up to 42.6%, the maximum wear resistance improvement in terms of percentage has been seen in the sample processed at 1200 rpm and subjected to 20 N load i.e. a wear value of 0.0032 grams as compared to 0.0075 grams of the substrate (base) metal at the same load in the FSPed sample compared to the substrate (base) metal.

Figure 15 represents the SEM images of the worn out surface of tested samples. The deeper groove due to high material removal has been seen in the

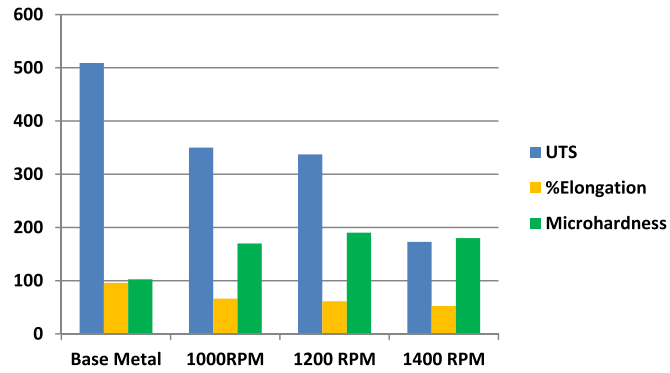


Figure 11. The effect of tool rpm on the Ultimate tensile strength, %Elongation and Microhardness.

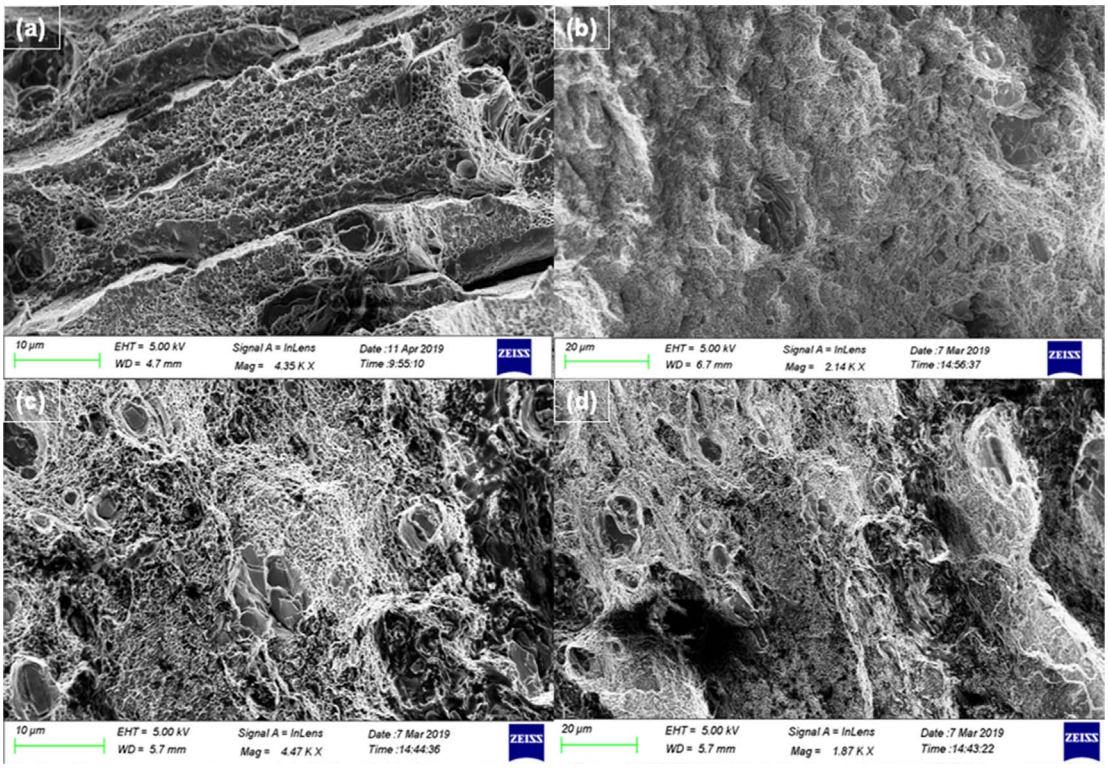


Figure 12. Fractographs image of (a) Base metal (b) The FSPed sample at 1000 rpm (c) FSPed sample at 1200 rpm (d) The FSPed sample at 1400 rpm.

figure 15(a), base metal. While comparatively smaller scratches and carbides are seen on the other three FSPed samples process at 1000 rpm 1200 rpm and 1400 rpm. The finer and thinner scratches in figures 15(b)–(d) represents the difficulty in material removal due to homogeneous distribution of boron carbide nanoparticles in the FSPed sample. As also seen from the wear results in the figure 14 previous discussions, the base metal has the highest wear rate after that the wear rate has decrease in the FSPed samples at 1000 rpm, 1200 rpm and 1400 rpm. In figures 15(a)–(d) the wear tracks and carbides can be seen. It was found that the main wear mechanism is adhesive and abrasive. These results are supported by [46, 47]. At high tool rotation speed at 1400 rpm causes more heat and more material flow and, the tool wear effects in the material flow play a dominant role and causes poor material flow and the results nanoparticles agglomeration and make the wear properties decreases.

4. Conclusions

In the present study, Friction stir processed surface composite has been processed through Self-assembled monolayer Technique. This work has underscored on the impacts of tool rotational speed in the FSP procedure on the mechanical and wear resistance of the prepared surface composite. The following conclusions have been derived from the test results.

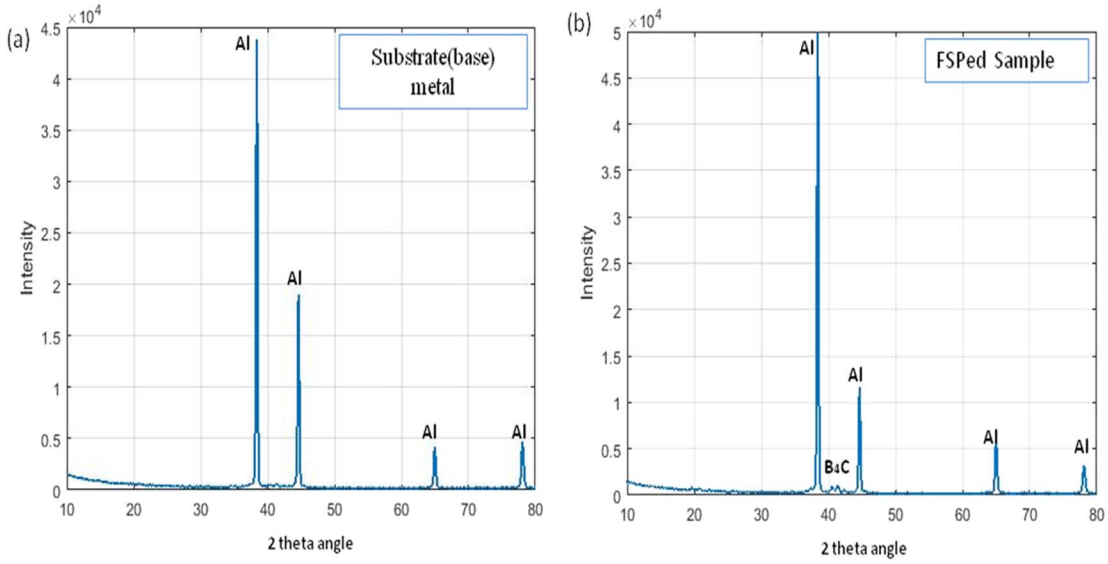


Figure 13. XRD pattern of (a) substrate (base) metal and (b) The FSPed sample.

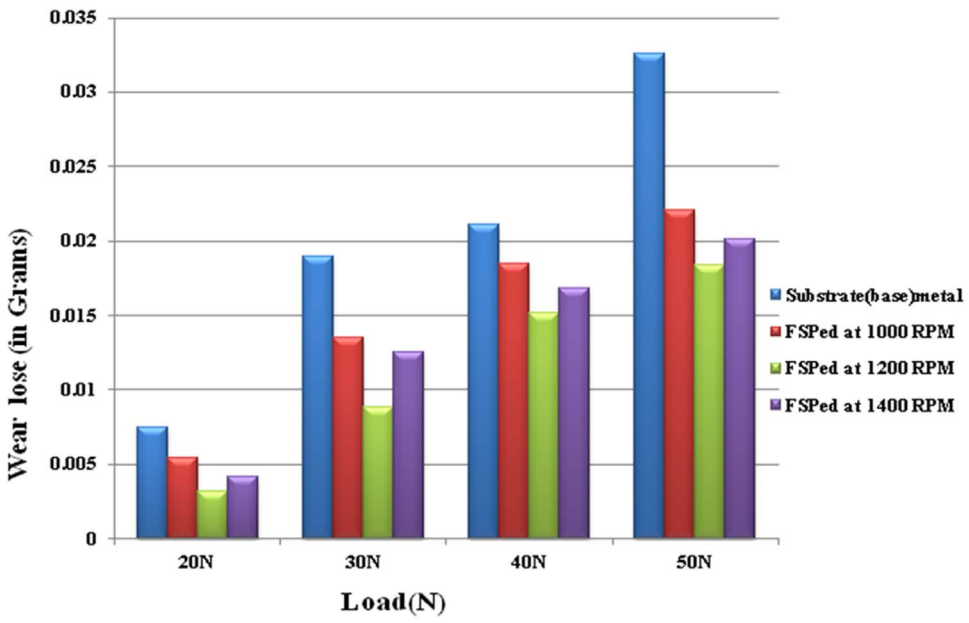


Figure 14. Comparison of wear rate at different loads for the four the processed and unprocessed sample.

1. Noteworthy, uniform distribution of nanoparticles has been found in the nano surface composite inside the stirred zone is achieved by using FSP methodology that causes exceptional plastic deformation and recrystallization due to frictional heating and were indicative of the presence of B_4C nanoparticle.
2. The average hardness for the surface composites has been found to be more than that of the substrate (base) metal. The trend is such that hardness increases with an increase in tool rotation speed upto 1200, after which it shows a decreasing trend.
3. FESEM fractography images, EDS composition and XRD pattern confirms the presence of boron carbide nanoparticles as reinforcement in nano surface composite.
4. The micrograph of the FSPed sample at 1000, 1200 and 1400 rpm reinforced with boron carbide nanoparticles exhibit the homogeneous distribution of reinforcement particle. The FSPed sample has obtained higher hardness 170 Hv, 185 Hv and 182 Hv, respectively as compared to the base metal (BM) hardness (102Hv).
5. For the 20, 30, 40 and 50 N load, as tool rotation speed increases from 1000 rpm to 1200 rpm, the

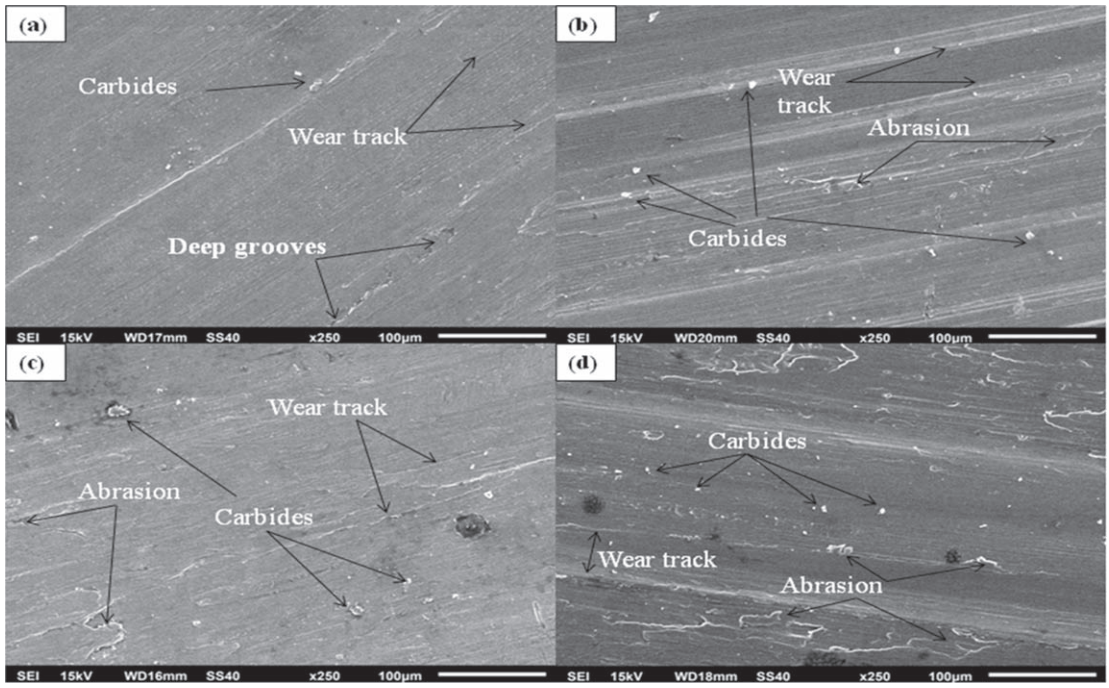


Figure 15. Worn-out surface morphology of (a) Substrate metal (b) FSPed at 1000 rpm (c) FSPed at 1200 rpm (d) FSPed at 1400 rpm.

wear resistance increases, after which a decrease in wear resistance is observed at 1400 rpm due to bad material flow and nanoparticle agglomeration.

6. Wear resistance has improved by upto 42.6% in the FSPed sample compared to the substrate (base) material. The maximum wear resistance improvement in terms of percentage has been seen in the sample processed at 1200 rpm and subjected to 20 N load.

Acknowledgments

The authors acknowledge Laser Spectroscopy Laboratory (DTU) for access to physical apparatus and fabricators and heartily send regards to Dr Mohan S Mehta, for assisting in processing of Self-assembled monolayers for experimental work.

ORCID iDs

Ravi Butola https://orcid.org/0000-0002-6581-4131
Ranganath M Singari https://orcid.org/0000-0003-4981-3857
Qasim Murtaza https://orcid.org/0000-0001-8784-2304

References

[1] Behnagh R A, Shen N, Abdollahi M and Ding H 2016 Ultrafine-grained surface layer formation of aluminum alloy 5083 by friction stir processing *Procedia CIRP* **45** 243–6
[2] Kumar P A, C M H, Pariyar A, Perugu C S, Kailas S V, Garg U and Rohatgi P 2020 Friction stir processing of squeeze cast A356 with surface compacted graphene nanoplatelets (GNPs) for the synthesis of metal matrix composites *Materials Science & Engineering* A **769** 138517
[3] Butola R, Pandey C P, Shukla A and Walia R S 2019 Effect on the mechanical properties of aluminum based hybrid metal matrix composite using stir casting method *Mater. Sci. Forum* **969** 253–9
[4] Coyal A, Yuvaraj N, Butola R and Tyagi L 2020 An experimental analysis of tensile, hardness and wear properties of aluminium metal matrix composite through stir casting process *SN Applied Sciences* **2** 892
[5] Hynes N R J *et al* 2020 Mechanical and tribological characteristics of boron carbide reinforcement of AA6061 matrix composite *J Braz. Soc. Mech. Sci. Eng.* **42** 155
[6] Rathee S, Maheshwari S, Siddiquee A N and Srivastava M 2018 Distribution of reinforcement particles in surface composite fabrication via friction stir processing, suitable strategy *Mater. Manuf. Processes* **33** 262–9
[7] Mishra R S, Ma Z Y and Charit I 2003 Friction stir processing: a novel technique for fabrication of surface composite *Materials Science and Engineering.* A **341** 307–10
[8] Yuvaraj N, Aravindan S and Vipin 2017 Comparison studies on mechanical and wear behavior of fabricated aluminum surface nano composites by fusion and solid state processing *Surface & Coatings Technology* **309** 309–19
[9] Chen X, Luais E, Darwish N, Ciampi S, Thordarson P and Gooding J J 2012 Studies on the effect of solvents on self-assembled monolayers formed from organophosphonic acids on indium tin oxide American Chemical Society*Langmuir* **28** 9487–95
[10] Hoque E, De Rose J, Bhushan B and Mathieu H 2008 Self-assembled monolayers on aluminum and copper oxide surfaces: surface and interface characteristics, nano tribological properties, and chemical stability In*Applied Scanning Probe Methods IX, Nano Science and Technolgy* ed M Tomitori, B Bhushan and H Fuchs (eds) (Berlin, Heidelberg: Springer)
[11] Torbjorn F, Adam K, Erikssona L A and Sten L 1998 The reaction between aluminium and dimethyl ether comparative

study of density functional theory and EPR results *J. Chem. Soc. Faraday Trans.* **94** 777–82
[12] Wang D, Ni Y, Huo Q and Tallman D E 2005 Self assembled monolayer and multilayer thin films on aluminum 2024-T3 substrates and their corrosion resistance study *Thin Solid Films* **47** 1177–85
[13] Mishra R S, Mahoney M W, McFadden S X, Mara N A and Mukherjee A K 2000 High strain rate superplasticity in a friction stir processed 7075 al alloy *Scripta Mater.* **42** 163–8
[14] Chaudhary A, Dev A K, Goel A, Butola R and Ranganath M S 2018 The mechanical properties of different alloys in friction stir processing: a review *Materials Today: Proceedings.* **5** 5553–62
[15] Elangovan K, Balasubramanian V and Valliappan M 2008 Influences of tool pin profile and axial force on the formation of friction stir processing zone in AA6061 aluminium alloy *Int. J. Adv. Manuf. Technol.* **38** 285–95
[16] Mahamani A 2016 Dry sliding wear and friction characteristics of AA6061-TiB_2/ZrB_2 *in situ* metal matrix composites at room temperature *Int. J. of Surface Science and Engineering* **10** 570–84
[17] Bisadi H and Abasi A 2011 Fabrication of Al7075/TiB_2 surface composite via friction stir processing *American Journal of Materials Science* **1** 67–70
[18] Alidokht S A, Abdollah Z, Soleymani S and Assadi H 2011 Microstructure and tribological performance of an aluminium alloy based hybrid composite produced by friction stir processing *Mater. Des.* **32** 2727–33
[19] Mardi K B, Dixit A R, Mallick A and Reddy A R 2018 Effect of Al_2O_3 nanoparticles on tribological behaviour of Mg-6Al alloy-based nanocomposite *Int. J. Surface Science and Engineering* **12** 402–18
[20] Aruri D, Adepu K and Bazavada K 2013 Wear and mechanical properties of 6061-T6 aluminum alloy surface hybrid composites [(SiC + Gr) and (SiC + Al_2O_3)] fabricated by friction stir processing *Journal of Materials Research and Technology* **2** 362–9
[21] Ramesh R and Murugan N 2013 Microstructure and metallurgical properties of aluminium 7075—T651 alloy/B4C 4% vol. Surface composite by friction stir processing *Advanced Materials Manufacturing & Characterization* **3** 301–6
[22] Rana H G, Badheka V J and Kumar A 2016 Fabrication of Al7075/B4C surface composite by novel friction stir processing (FSP) and investigation on wear properties *Procedia Technology* **23** 519–28
[23] Panneerselvam T, Kandavel T K, Arun K S and Dinesh kumar V 2018 Tribological study on hybrid reinforced aluminium-based metal matrix composites *Int. J. Surface Science and Engineering* **12** 449–66
[24] Abrahams R, Mikhail J and Fasihi P 2019 Effect of friction stir processing parameters on the mechanical properties of 5005-H34 and 7075-T651 aluminium alloys *Materials Science and Engineering* **751** 363–73
[25] Kurt A, Uygur I and Cete E 2011 Surface modifications of aluminium by friction stir processing *J. Mater. Process. Technol.* **211** 313–7
[26] Shamsipur A, kashani-Bozorg S F and Hanzaki Z A 2012 Fabrication of Ti/Sic surface nano-composite layer by friction stir processing *International Journal of Modern Physics, Conference Series* **5** 367–74
[27] Yuvaraj N, Aravindan S and Vipin 2015 Fabrication of Al5083/B4C surface composite by friction stir processing and its tribological characterization *Journal of Materials Research and Technology* **4** 398–410
[28] Harichandran R and Selvakumar N 2016 Effect of nano/micro B_4C particles on the mechanical properties of aluminium metal matrix composites fabricated by ultrasonic cavitation-assisted solidification process *Archives of Civil and Mechanical Engineering* **16** 147–58
[29] Bommarito G M and Pocius A V 1998 An electrochemical study of the changes in the passivation of an aluminum alloy surface induced by the presence of a self-assembled monolayer *Thin Solid Films* **327–329** 481–5
[30] Smith R K, Lewis P A and Paul S W 2004 Patterning self-assembled monolayers *Prog. Surf. Sci.* **75** 1–68
[31] Ayad A M, Konstantakopoulou M, Gibson A G and George K 2016 Preparation of Novel Nanocomposites Using the Ultra-sonication Technique *2016 International Conference for Students on Applied Engineering (ICSAE) (Newcastle upon Tyne, UK, 20-21 Oct. 2016)* (IEEE) 383–91
[32] Awad S B and Nagarajan R 2010 *Development in Surface Contamination and Cleaning* (Amsterdam: Elsevier) 225–80
[33] Butola R, Murtaza Q and Singari R M 2019 CNC turning and simulation of residual stress measurement on H13 tool steel *Advances in Computational Methods in Manufacturing. Lecture Notes on Multidisciplinary Industrial Engineering* ed R Narayanan, S Joshi and U Dixit (Singapore: Springer) 337–48
[34] Butola R, Ranganath M S and Murtaza Q 2019 Fabrication and optimization of AA7075 matrix surface composites using Taguchi technique via friction stir processing (FSP) *Eng. Res. Express* **1** 025015
[35] Liu P, Shi Q, Bian X, Xu S and Ren G 2009 Microstructure of a novel Al-based amorphous reinforced aluminum metal matrix composite *Hanjie Xuebao/Transactions of the China Welding Institution* **30** 13–6
[36] Sahlot P *et al* 2017 Quantitative wear analysis of H13 steel tool during friction stir welding of u-0.8% Cr-0.1% Zr alloy *Wear* **378** 82–9
[37] Ramezani N M R, Davoodi B, Aberoumand M and Hajideh M R 2019 Assessment of tool wear and mechanical properties of Al 7075 nanocomposite in friction stir processing (FSP) *Journal of the Brazilian Society of Mechanical Sciences and Engineering* **4** 182
[38] Khodabakhshi F, Gerlich A P, Simchi A and Kokabi A H 2015 Cryogenic friction-stir processing of ultrafine-grained Al-Mg-TiO_2 nanocomposites *Materials Science & Engineering* A **620** 471–82
[39] Azizieh M, Kokabi A H and Abachi P 2011 Effect of rotational speed and probe profile on microstructure and hardness of AZ31/Al_2O_3 nanocomposites fabricated by friction stir processing *Mater. Des.* **32** 2034–41
[40] Fernandez G J and Murr L E 2004 Characterization of tool wear and weld optimization in the friction-stir welding of cast aluminum 359 + 20% SiC metal-matrix composite *Mater. Charact.* **52** 65–75
[41] Butola R, Murtaza Q and Ranganath M S 2020 Formation of self-assembled monolayer and characterization of AA7075-T6/B_4C nano-ceramic surface composite using friction stir processing *Surface Topography: Metrology and Properties* **8** 025030
[42] Suri A *et al* 2017 Impact and tensile testing of Al2024 alloy processed by friction stir processing *Procedia Engineering* **173** 679–85
[43] Ganesa B K and Mahadevan K 2015 Investigation on the effects of process parameters on the mechanical and corrosion behaviour of friction stir-claded AZ31B magnesium alloy *Arab. J. Sci. Eng.* **40** 1647–55
[44] Moustafa E 2017 Effect of multi-pass friction stir processing on mechanical properties for AA2025/Al_2O_3 *Materials* **10** 1053
[45] Murugan N and Kumar B A 2013 Prediction of tensile strength of friction stir welded stir cast AA6061-T6/AlNp composite *Material Design* **51** 998–1007
[46] Tyagi L, Butola R and Jha A K 2020 Mechanical and tribological properties of AA7075-T6 metal matrix composite reinforced with ceramic particles and aloevera ash via friction stir processing *Material Research Express* **7** 066526
[47] Butola R, Tyagi L, Kem L, Ranganath M S and Murtaza Q 2020 Mechanical and wear properties of aluminium alloy composites: a review *Manufacturing Engineering. Lecture Notes on Multidisciplinary Industrial Engineering* ed V Sharma *et al* (Singapore: Springer) (https://doi.org/10.1007/978-981-15-4619-8_28)

Journal of Intelligent Information Systems
https://doi.org/10.1007/s10844-020-00623-8

Modeling information diffusion in online social networks using a modified forest-fire model

Sanjay Kumar[1,2] · **Muskan Saini**[3] · **Muskan Goel**[4] · **B. S. Panda**[2]

Received: 12 July 2020 / Revised: 22 September 2020 / Accepted: 22 September 2020 /
Published online: 12 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Information dissemination has changed rapidly in recent years with the emergence of social media which provides online platforms for people worldwide to share their thoughts, activities, emotions, and build social relationships. Hence, modeling information diffusion has become an important area of research in the field of network analysis. It involves the mathematical modeling of the movement of information and study the information spread pattern. In this paper, we attempt to model information propagation in online social networks using a nature-inspired approach based on a modified forest-fire model. A slight spark can start a wildfire in a forest, and the spread of this fire depends on vegetation, weather, and topography, which may act as fuel. On similar lines, we labeled users who haven't joined the network yet as $Empty$, existing users as $Tree$, and information as $Fire$. The spread of information across online social networks depends upon users-followers relationships, the significance of the topic, and other such features. We introduce a novel $Burnt$ state to the traditional forest-fire model to represent non-spreaders in the network. We validate our method on six real-world data-sets extracted from Twitter and conclude that the proposed model performs reasonably well in predicting information diffusion.

Keywords Information diffusion · Forest-fire model · Nature-inspired algorithm · Online social networks · Twitter

1 Introduction

Many complex systems like biological, communication and social networks can be modeled as graphs (Newman 2010). These systems constitute a large number of nodes with a complicated interaction among its members (Wang et al. 2012; Barabási 2016). Social networks are online resources that link people and help in the spread of information (Bakshy et al. 2012). Due to the rapid advancements of the internet and mobile networks, online

✉ Sanjay Kumar
sanjay.kumar@dtu.ac.in

Extended author information available on the last page of the article.

social networks (OSNs) like Twitter, Facebook, Siena Weibo have become very popular and connects billions of people worldwide. These platforms excel as tools for people to share news, trending topics, ideas, and opinions. Hence, online social networks have brought people together, enhance communication speed and generate a massive amount of data in a few minutes. Some actions of a few numbers of people may lead to a large scale spreading of information. It takes only a few clicks for information to reach from one corner of the world to another. Twitter, a micro-blogging website is one such network. As compared to other networks, Twitter is more diverse and densely connected. These properties make Twitter an engrossing network for research. Social networks like Twitter play an essential role in the analysis of the spread of information. Modeling information diffusion on these networks has many applications like finding trending topics (Mashiach and Sharma 2020), finding influential users (Kumar and Panda 2020), devising marketing strategies, identifying opinion leaders (He et al. 2020; Kimura et al. 2013), and many more. In the past, information diffusion models like susceptible-infected-recovered (SIR), susceptible-exposed-infected-recovered (SEIR), and other similar epidemic models have been used Guille et al. (2013) and Cai et al. (2012). These models try to model the information diffusion on the network and allowing the estimation of the spread of the information with certain limitations. Recently, many researchers have made proposals in the field of information diffusion simulating natural processes like spring damper, bee colonies, etc Cai et al. (2012) and Sankar and Kumar (2016).

A common hazard in forests is forest-fire, which occurs due to human-made or natural reasons. Broad research has demonstrated that the fundamental variables influencing the spread of forest-fires are forest flammability, climate, and landscape (Hawley 1926). Many studies were conducted in past to model the spread of wildfire in forests in form of computer simulations and algorithms (D'Ambrogio et al. 2016; Kanga and Singh 2017). It is an important area of research aiming to help mangers to quickly anticipate the spread of fires, giving a decision premise to work out viable fire extinguishing plans in advance. This natural phenomenon can be used to model many real time problems efficiently. The traditional forest-fire model has numerous applications in applied mathematics like cell automata, and community detection (Pattanayak et al. 2019; Wang and Liu 2011).

In this paper, we assume information spreading in an online social network is analogous to the spread of fire in a deep forest. We explore the usage of the forest-fire algorithm for modeling the spread of information on real-life data-sets of Twitter based on user's tweets around a given hashtag. The contributions of our paper are as follows:

1. We successfully extend the usage of forest-fire model on complex networks to study information diffusion and predict influenced people in online social networks.
2. We introduce a *Burnt* node status in the traditional forest-fire model to categorize the infected population into spreaders, users who are responsible for spreading the information, and non-spreaders, users who keep information to themselves and do not spread it further.
3. We model the information spread for multiple hashtags including the information regarding the global pandemic of COVID-19. The proposed algorithm performs reasonably well on real-world tweet data-sets of Twitter.

The rest of the paper is organized into the following sections: Section 2 explores the traditional models in the field of information diffusion. Section 3 discusses the methodology used and our proposed algorithm. Section 4 presents the description of the datasets used to validate the model. In Section 5, we visualize the proposed algorithm on parts of the Twitter

network. In Section 6, we discussed the results obtained for various real-world datasets. Section 7 concludes the paper and discusses the contribution of the paper in the field of information diffusion.

2 Backgrounds and related works

The process of information diffusion involves the flow of information from one part of a network to another. Information Diffusion in online social networks is an interesting area of research (Bakshy et al. 2012; Chakraborty et al. 2018). Several attempts have been made in the past to study the dynamics of the information diffusion process (Guille et al. 2013). Modeling information propagation can assist in predicting the extent of information spread. The models can be broadly divided into two categories, as shown on Fig. 1: Explanatory models and Predictive Models (Guille et al. 2013). Explanatory Models try to reiterate the path of information. Predictive models, on the other hand, predict the diffusion process in any network. Predictive Models can be further categorized into graph-based and non-graph-based models. Graph-based models include threshold models like Linear Threshold (LT) and cascading models like Independent Cascades (IC). These models work on the assumption that a static graph exists for the network and, thus, try to model the diffusion process. Non-Graph based include epidemic models like SI, SIR, SIS, SIRS Model, and Partial Differential Equation (PDE) Model (Zhou et al. 2006). These approaches do not assume the presence of a static graph. Epidemicmodeling is a popular approach in the field of information diffusion modeling (Daley and Gani 2001). Such models represent the spread of information in a manner similar to the spread of a disease in a community. In the epidemic spread, infected people come in contact with susceptible people and pass on the infection.

Similarly, information spreads from one person to another Stai et al. (2018). Several attempts have been made to study the dissemination of information using traditional epidemic models like the Susceptible-Infected model, Susceptible-Infected-Recovered model.

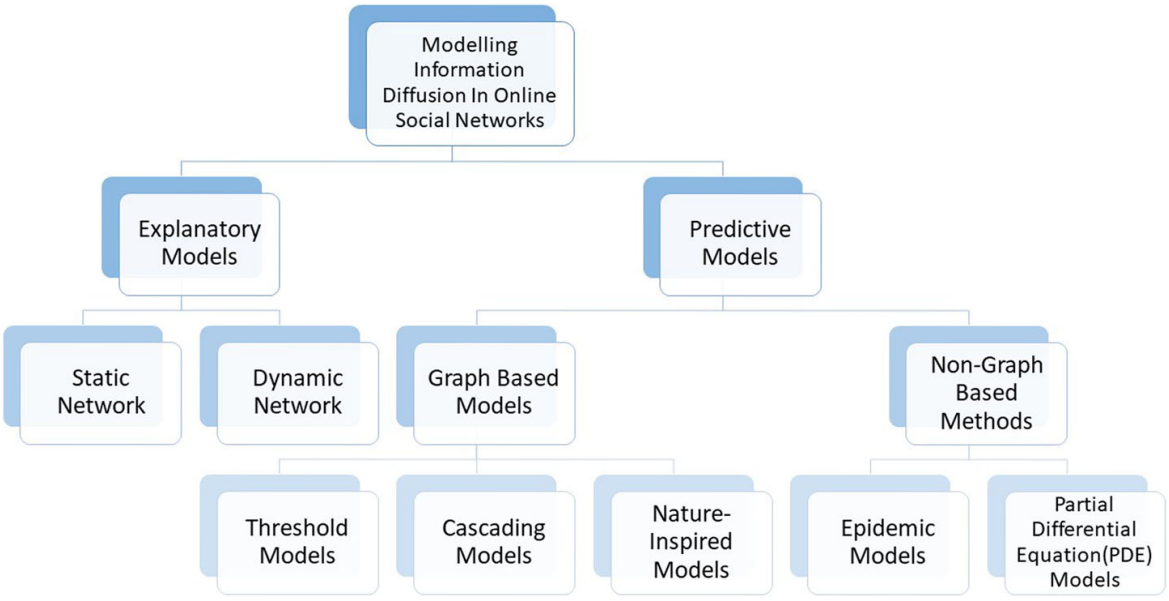


Fig. 1 Information diffusion models

Also, variants of these traditional models like SEIR (Biswas et al. 2014), SHIR (Liu et al. 2016), and others have been developed in the past. Recent work proposed a model named CISIR (competitive information susceptible infected recovered) for studying information dissemination in online social networks (He and Liu 2020). Some of the models were developed, especially for Twitter. Saito et al. (2015) observed information diffusion data and devised a generic algorithm to detect changes in parameter values of an information diffusion model. Hoang and Mothe (2018) proposed a predictive model for information diffusion on twitter based on user-based, content-based, and time-based features to evaluate if a post will be retweeted or not. Ding and Li (2018a) presented a topologically biased random walk for the diffusion process in multiplex networks displaying how topological properties affect the diffusion process. In the recent past, attempts have been made to model information diffusion using nature-inspired algorithms. Various nature-inspired algorithms like a firefly, cuckoo search algorithm, forest-fire, and similar have been used to solve problems related to social networks. Yang et al. (Ding and Li 2018b) discusses challenges and problems like scalability, parameter tuning, etc. that occur while dealing with nature-inspired algorithms. Forest-fire Model has been used to generate graphs (Rui et al. 2018) and investigate interactions in social networks (Fischer et al. 2013; Indu and Thampi 2019). We attempt to use the forest-fire algorithm to study information diffusion on Twitter and modify it to find the actual spreaders of information. Hu et al. (2018) gave an overview of different perspectives to view the information diffusion process. These perspectives are macroscopic, microscopic, and interaction between macroscopic and microscopic. The macroscopic perspective focuses on the extent to which information spreads in a network, whereas a microscopic perspective deals with individual nodes i.e., how a particular node will act upon a piece of information. The perspective of the interaction between macroscopic and microscopic levels deals with the relations between two levels. In this paper, we deal with both macroscopic and microscopic levels. We model the extent of information spread and also discover the nodes who may spread the information further and who may not.

3 Methodology

In this section, we describe the forest-fire algorithm and its relevance for modeling information diffusion. Firstly, we describe an algorithm to use forest-fire model online social networks to model the extent of information diffusion i.e., the total number of infected people. Secondly, we propose a modification in the traditional forest-fire model by adding a *Burnt* node to divide the infected population into spreaders and non-spreaders. This helps in identifying the people who actually spread the information. Further, we define the parameters on which the information spread depends.

3.1 Forest-fire algorithm

The traditional forest-fire model defined for a cell automata consists of three states: *Empty*, *Tree* and *Fire*. The model has two probabilities p and f where p is the probability with which new trees originate in a forest, and f is the probability with which any tree catches fire. Whenever a cell consisting of a tree catches fire, its adjacent cells also catch fire, and then they further spread the fire. We argue the same idea can be applied to graph datasets to study information diffusion in online social networks. Thus, in the case of graphs, whenever

a node spreads information, its neighbors can also spread the information. The model can be applied to social network graphs like twitter as well. We present below Algorithm 1 based on the forest-fire model to simulate information diffusion in the Twitter network. In this algorithm, all the current Twitter users are considered as $Tree$. New users join the twitter network every day, leading to an $Empty$ to $Tree$ transformation. When a user tweets or retweets on a topic, he is given the status $Fire$. Also, all his followers are considered informed on the topic and are thus converted into $Fire$. The followers can further retweet a tweet spreading the information further. The fire keeps on spreading and soon covers the entire forest. Similarly, information unfurls over the whole network. In the algorithm, we calculate the probability f_u of a user posting a tweet on a topic considering various factors like user activity, and topic significance where f_0 is the threshold value for the probability f. The Algorithm 1 explains how forest-fire algorithm can be applied on graphs of social networks like twitter to study the extent of information diffusion. The input to the algorithm is the social network graph (G) and output is the list of infected or informed nodes (I) about the particular topic or subject.

Algorithm 1 Forest-fire algorithm for modeling information diffusion.

Input: Social Network, $G = (V, E)$
Output: List of informed nodes (I)

```
1:  procedure ForestFire(G)
2:      Initialize I = {}
3:      Initialize status of each node
4:      for each node u in G do
5:          if status [u] = Tree then
6:              if f_u ≥ f_0 then
7:                  status [u] = Fire
8:                  I = I ∪ {u}
9:              end if
10:         end if
11:         if status[u] = Fire then
12:             for each neighbour v of u do
13:                 status[v] = Fire
14:                 I = I ∪ {v}
15:             end for
16:         end if
17:     end for
18:     Return I
19: end procedure
```

The detailed explanation of the steps involved in the Algorithm 1 is as follows:

Step-1 The status of each node is initialized as per line 3 of Algorithm 1. The users who had posted a tweet on a topic are initialized as *fire*, and all others are labeled as *tree*.

Step-2 Each node of the graph is then iterated. If the status of node (u) is *tree*, then the tweet probability, f_u is calculated for node u using the Eq. No. (2). If $f_u \geq f_0$, where f_0 is the threshold value, then the user is considered to have posted a tweet

on the topic and thus, is labeled as *fire*. The node (u) is then added to the list of informed nodes (I). This step explains lines 5-8 of Algorithm 1.

Step-3 If the status of the node is *Fire*, then all of its neighbors are also labeled as *fire* and added to the list of infected nodes. This is just like a forest-fire, when a tree is burning, all of its neighboring trees also catch fire. This step covers lines 9-12 of Algorithm 1.

3.2 Modified forest-fire (MFF) algorithm

In this section, we propose a modified forest-fire (MFF) algorithm for online social networks. In the forest-fire Algorithm 1, once a tree catches fire, it spreads it to its neighbors, which spread it further. In the case of social networks, all people do not spread the information further. Thus, infected or informed people can be classified into spreaders and non-spreaders. To model the non-spreaders, we introduce a new *Burnt* node. This can be correlated with the forest-fire phenomenon as some trees may get entirely burnt without spreading the fire further. Whether a tree may spread the fire further depends on the composition of the tree, climate, and vegetation of the area. Similarly, whether a user may pass on the information further depends on user characteristics and various other factors like user activity, user interests, topic significance. Thus, we calculate the retweet probability r_u, the probability with which a user retweets a tweet considering various characteristics of users and interests on tweets. The algorithm presented in 2 is written for the Twitter platform but can be easily extended to other online social networks like Facebook, Instagram, and others. We have created a correlation between forest-fire and information diffusion on the Twitter network. The algorithm partitions the entire population into 4 categories: *Empty*, *Tree*, *Fire*, *Burnt*

(i) *Empty* represents the users who have not yet joined the network. Twitter is a dynamic network, with new users joining the network everyday. This can be considered analogous to a forest where new trees originate every now and then.
(ii) *Tree* refers to users that have no knowledge about the topic.
(iii) *Fire* refers to those users that are informed and are actively involved in spreading the information further by either tweeting or retweeting on that topic.
(iv) *Burnt* refers to the nodes that are infected but do not spread the information further. A user v can become infected in two ways, either a user u whom v follows publishes a tweet on a particular topic or user v becomes aware of the topic from some outside source. It is important to note that *Burnt* nodes are not spreading the information further.

There are three types of probabilities that come into play in this algorithm p , f_u and r_u.

(i) p is the probability that a new user joins the network i.e. *Empty* to *Fire* transformation.
(ii) f_u is the probability that an existing user will post a tweet related to a topic i.e. transformation from *Tree* to *Fire* state.
(iii) r_u is the probability that a user v retweets a tweet that the person u he follows tweeted.

Algorithm 2 describes the modified forest-fire algorithm and explains its applicability on online social networks to model information diffusion. In the algorithm, f_0 and r_0 are the threshold probability values for the probabilities f_u and r_u, respectively. The input to the Algorithm 2 is the social network graph (G). The output is the list of spreader nodes (S) and non-spreader nodes (NS).

Algorithm 2 Modified forest-fire algorithm.

Input: Social Network, $G = (V, E)$
Output: List of spreaders (S) and non-spreaders (NS)

```
1:  procedure ModifiedForestFire( G )
2:      Initialize S = {}, NS = {}
3:      Initialize status of each node
4:      for each node u in G do
5:          if status[u] = Tree then
6:              if f_u ≥ f_0 then
7:                  status [u] = Fire
8:                  S = S ∪ {u}
9:              end if
10:         end if
11:         if status [u] = Fire then
12:             for each neighbour v of u do
13:                 if status [v] = Tree then
14:                     if r_u ≥ r_0 then
15:                         status [v] = Fire
16:                         S = S ∪ {v}
17:                     else
18:                         status [v] = Burnt
19:                         NS = NS ∪ {v}
20:                     end if
21:                 end if
22:             end for
23:         end if
24:     end for
25:     Return S and NS
26: end procedure
```

The detailed explanation of the steps involved in the Algorithm 2 is as follows:

Step-1 The status of each node is initialized as either *tree* or *fire*. The users who had posted a tweet on a topic are labeled as *fire*, and all others are labeled as *tree*. The list of spreaders (S) and non-spreaders (NS) is also initialized. This explains lines 2 and 3 of the Algorithm 2.

Step-2 It explains lines 4-8 of the algorithm. Each node of the graph is iterated. If the status of the node (u) is *tree*, then the tweet probability, f_u is calculated for that node using the Eq. No. (2). If $f_u \geq f_0$, where f_0 is the threshold value, then the user is considered to have posted a tweet on the topic and thus, is labeled as *fire*. The node (u) is added in the list of spreaders.

Step-3 If the node's status is *fire*, then it means that the user has posted a tweet or retweet on a topic, then each of its neighbor or follower (v) is considered. If the status of node v is *tree*, then the retweet probability, r_u is calculated for that follower using the Eq. No. (8). If $r_u \geq r_0$, where r_0 is the threshold value, the follower (v) is considered to have re-tweeted the person's tweet or retweet he follows and thus, v is labeled as *fire* and is added to the list of spreaders. If $r_u < r_0$, it means that the person (v) did not retweet the tweet and did not spread the information further.

Therefore, user v is labeled as *Burnt* and is added to the list of non-spreaders. This explains lines 9-17 of the Algorithm 2.

Step-4 After iterating all the nodes of the graph, the list of spreaders and non-spreaders is returned as output.

3.3 Defining probabilities

In this subsection, we study the probabilities p, f_u and r_u and enumerate the factors on which these probabilities depend. After scrutinizing the features of tweets and users, we model these probabilities and use them in the modified forest-fire model to study the rate of spread of information.

3.3.1 Probability of a new user joining the network (p)

Online social networks have become so popular and advanced in the last decade. They are still expanding and reaching people worldwide. Any new user joining the social network can act as a *Tree* to spread information. Hence, p represents the probability that state changes from *Empty* to *Tree*.

3.3.2 Probability of an existing user tweeting about the topic (f_u)

The entire forest area could burst into flames ignited by the slightest spark. Similarly, in online social networks, the users who tweet about the topic act as a spark in the system. These are the users who bring information to the network. After analyzing Twitter statistics, we found that whether or not a user tweets about a topic largely depends on the user's general activity and the importance of the topic.

(i) **User Activity** (UA): The general activity of a user can be determined by taking into consideration the total number of tweets or retweets posted by a user over a certain period of time. All the tweets/retweets posted by the user from the account registration period are considered. Its high value specifies that the user is highly active on Twitter, and thus, there is a high possibility that our user will tweet about a topic. Depending on the account activity, we assign a range of values to users. We compute the user's activity (UA) as follows:-

$$UA = \frac{\text{Total number of tweets or retweets made by the user}}{\text{Number of days from account registration period till date}} \tag{1}$$

(ii) **Topic Significance** (TS): In general, users tweet about the more popular and trending topics. Depending upon the significance of the topic of the tweet, we provide them ranks or scores (TS), as shown in Fig. 2. Higher the score, higher is the importance of topic and thus, higher is the probability that any user will tweet about it. We divide the topics into following categories and then give them scores according to their importance.

(a) *Global and Local Interest* : A topic can be of global interest like COVID-19 with multiple users from diverse corners of the world sharing and tweeting about the same or of local interest like elections in an area attracting the interests of only locals residing in or involved with the area. Topics with global affairs attract a lot of attention worldwide and hence, tend to get shared a lot as compared to topics of local interests.

Topic Significance (Rank)	Global Interest	Local Interest
Trending	4	3
Non - Trending	2	1

Fig. 2 Assigning ranking to tweets on the basis of their topic significance

(b) *Trending and Non-Trending* : If the content is related to hot news recently or say has a trending topic like #MeToo movement, COVID-19 Pandemic, etc. The users tend to get more captivated by it and may post their status on that particular topic. On the other hand, non-trending topics like #Travel, #Tuesday are just like regular contents that are less popular and may get less recognition.

Thus, global and trending topics are given highest score and local and non-trending topics are given the least score.

The probability f_u is calculated as the weighted summation of above discussed features like user activity, topic interest as shown in Fig. 3. Tweet probability (f_u) can be represented mathematically as:

$$f_u = w_1 \times UA + w_2 \times TS \tag{2}$$

Here, w_1 and w_2 are the constant weights associated with the above mentioned features namely User Activity and Topic Significance respectively.

$$w_1 + w_2 = 1 \tag{3}$$

3.3.3 Probability of an existing user retweeting a tweet (r_u)

To properly understand how information flows in online social networks, we first need to calculate the probability by which any user will retweet since retweeters (people who

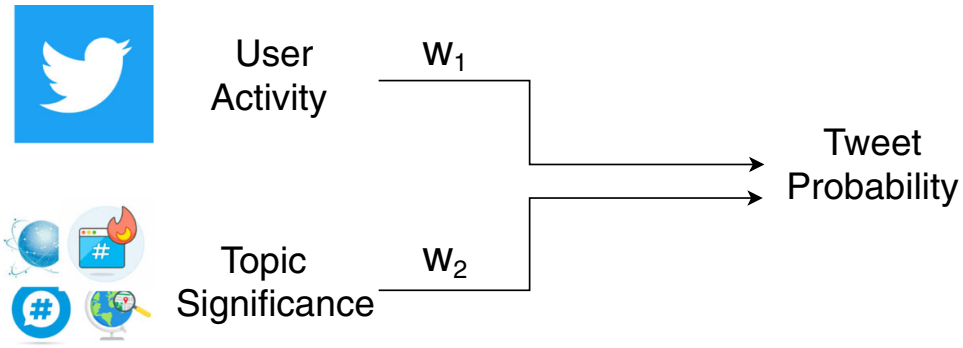


Fig. 3 Calculation of tweet probability (f_u)

re-tweet or re-post another person's tweet) play a significant role in disseminating the information over the entire network. Several works have been done in past as well to study the retweet phenomenon (Kuang et al. 2014; Nesi et al. 2018). We analyzed some important features of users listed below which can help us in answering that with what probability a particular user may retweet a particular tweet.

(i) **IsMentioned (IM)**:- One of the most popularly used functions on Twitter is the @mention function. The @mention function is used to 'tag' users in the status updates. Mentioning a user on Twitter is an effective way to grab a user's attention quickly and hence, there is a very high probability that the user will re-tweet or interact with it in some way or other. Recursive partitioning procedure models (RPART) also labelled mentions count as the most correlated feature with the action of retweeting (Nesi et al. 2018). Here, we take IsMentioned score as a binary variable.

$$\text{IsMentioned (IM)} = \begin{cases} 1, & \text{if user is mentioned in the tweet} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

(ii) **Similarity Score (SS)** Relationships between the users are the backbone of social media services. One of the most significant features that describe a relationship is how similar the two users are in terms of their interests, activities and behaviours. More similar the two users/accounts are, higher is the possibility of one retweeting another's content and spreading it in the network. We consider some user-specific features to calculate similarity score. We use Jaccard similarity index to calculate these measures or features. The formula to find Jaccard similarity index for a given measure m for two user accounts can be re-written as:

$$A_m(X, Y) = \frac{X_m \cap Y_m}{X_m \cup Y_m} \tag{5}$$

Here, A_m (X_m,Y_m) represents the Account similarity between two user accounts X and Y for a given feature or measure m. X_m and Y_m are the values of the measure under consideration. $X_m \cap Y_m$ represents the number of values of corresponding measure common in both accounts and $X_m \cup Y_m$ represents the total number of values of corresponding measure in either of the accounts. We consider five different features to find the similarity between two users X and Y using Jaccard similarity index. These features as follows :-

(a) *Followings List (A_{Fg})*: By intuition, users who share common interests must be following some common accounts. Hence, we find the "following" list of both the users and find the number of common accounts they are following.

(b) *Hashtags Used (A_{Ht})*: Hashtags have become the new fashion to grab more crowd's attention towards any topic on social media platforms. Users use hashtags to highlight some relevant keywords or phrases while posting content (or tweet). Clicking on these words containing hashtags take you to the other tweets labeled with the same hashtags. We can clearly say that hashtags define the interests and ideas of our users. Evidently, if two users have common hashtags, we can infer that they are interested in similar topics. We extracted hashtags of both the users, more the common hashtags and keywords in user's tweets, similar their interests, tend to be.

(c) *Languages used (A_{Lg})*: The online social networks constitute people from diverse backgrounds and cultures spread out geographically. Only those users

who can understand each other's languages will take interest in another's content (tweet) and may retweet it. Hence, we found the common languages used by both the users by extracting and analysing languages used in tweeting/retweeting by the user on the platform.

(d) *Location of Account (A_{Lc})*: Location can be also considered in calculating similarity score as people from same geographical boundaries may share some common interests, say related to local activities, although not necessarily. Providing your profile location is an optional feature in Twitter. So, we took this factor in account only for those users who have provided their location.

(e) *Followers List (A_{Fw})*: Let us consider a user who tweets about a topic. We can assume that the followers of our user can share some common interests with our user and are similar since similar people tend to follow each other. Furthermore, when these followers of our user are also following some other user, we can infer that they too may appear similar in some aspects. Hence, the two users are similar and hence, we can use followers list of both the users to find the common followers as one of the similarity characteristics.

Finally, similarity score is calculated as the weighted summation of above five measures as also shown in Fig. 4. The similarity score (SS) can be represented mathematically as:

$$SS = w_{Fg} \times A_{Fg} + w_{Ht} \times A_{Ht} + w_{Lg} \times A_{Lg} + w_{Lc} \times A_{Lc} + w_{Fw} \times A_{Fw} \quad (6)$$

Here, w_{Fg}, w_{Ht}, w_{Lg}, w_{Lc} and w_{Fw} are constant weights associated with above-discussed features, namely the Followings List, Hashtags Used, Languages Used,

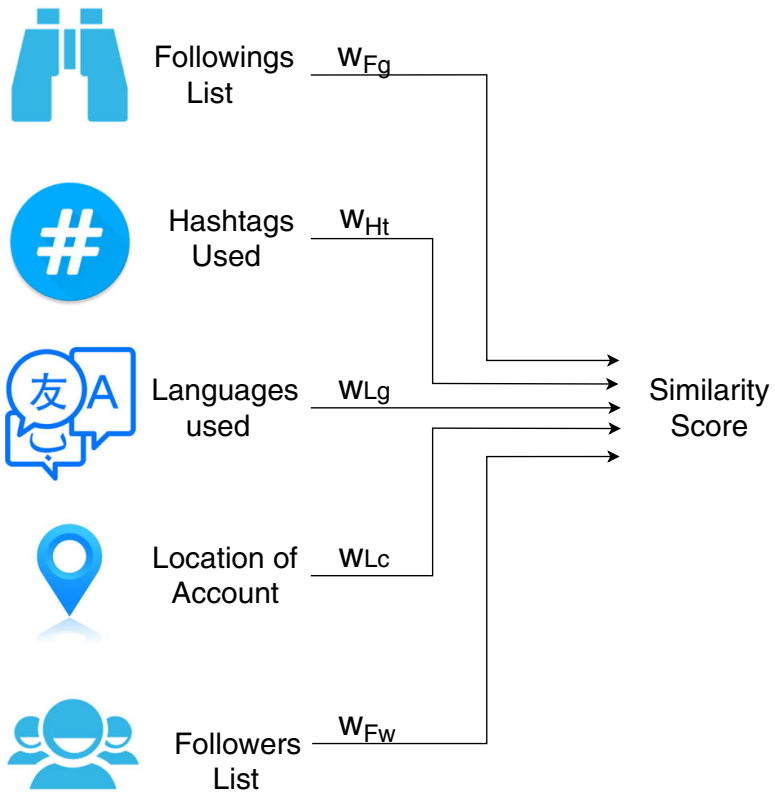


Fig. 4 Calculation of similarity score between users

Location of the Account, and Followers List features respectively. The sum of all these weight constants is one.

$$w_{Fg} + w_{Ht} + w_{Lg} + w_{Lc} + w_{Fw} = 1 \tag{7}$$

(iii) **User Activity (UA)**: The general activity of the user gives us the idea of how much active a user is on Twitter. A high value of user activity indicates that our user is much engaged with the platform and thus, there is a high possibility that our user will be retweeting another person's tweet. We use Eq. No. (1) to calculate the user activity.

(iv) **Topic Significance (TS)** The huge amount of retweeting happens just when the content is appealing, which we name as the topic significance. Users will in general give more consideration to those tweets with attractive and popular contents, that is, with high evaluation of significance of topic. Depending upon the significance of the topic of the tweet, we gave them scores (TS) using Fig. 2. Higher the score, higher is its probability to get retweeted or dispersed further in the online network.

Finally, retweet probability (r_u) of any user is just a weighted summation of the scores of four features listed above as shown in Fig. 5. It can be represented mathematically as:

$$r_u = w_{IM} \times IM + w_{SS} \times SS + w_{UA} \times UA + w_{TS} \times TS \tag{8}$$

, where w_{IM}, w_{SS}, w_{UA}, and w_{TS} are constants weight associated with IsMentioned, similarity score, user activity, and topic significance, respectively. The sum of all these weight constants is one.

$$w_{IM} + w_{SS} + w_{UA} + w_{TS} = 1 \tag{9}$$

4 Datasets description

We test our proposed model for the Twitter network. We perform the results validated for six real-time Twitter datasets on the topics: Corona Virus (Smith 2020), FIFA World Cup

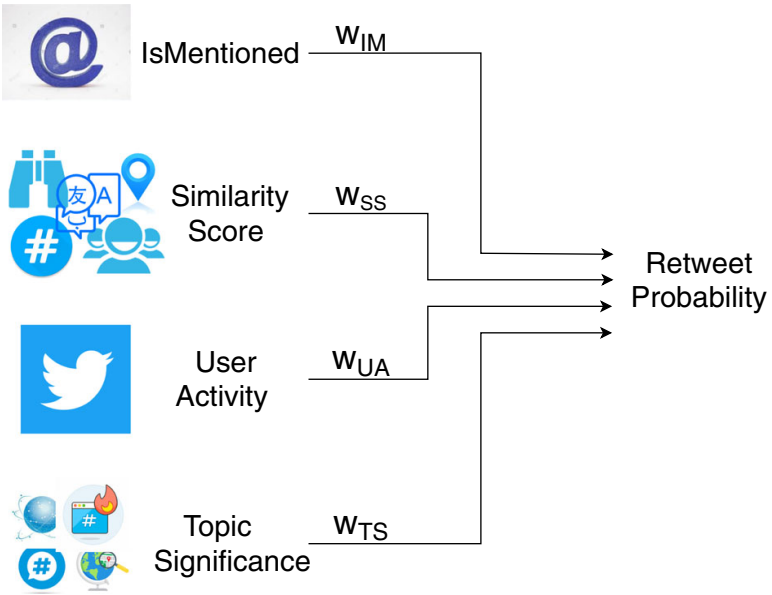


Fig. 5 Factors on which retweet probability (r_u) depends

(Rituparna 2018), NBA Finals (Pesic 2018), Game of Thrones S8 (de Abreu 2019), MeToo movement (Ramirez 2018) and Coachella Festival (Pesic 2019). "Coronavirus" or "COVID-19" refers to the global pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It started in December 2019 in Wuhan, China, and has been declared a public health emergency of international concern by the World health organization (WHO). As of April 2020, around 210 countries in the world have reported cases of COVID-19.

"FIFA World Cup" refers to the international football tournament in which men's national teams of 32 countries competed with each other. It was hosted by Russia from 14 June 2018 to 15 July 2018.

"NBAFinals" refers to the championship series of the National Basketball Association's 2017–18 season and conclusion of the season's playoffs. The dataset contains the tweets captured during the 3rd game of the 2018 NBA Finals between Cleveland Cavaliers and Golden State Warriors.

"Me Too" movement was a global campaign against sexual assault. It was first started by Tarana Burke in 2006 to empower girls who had faced sexual harassment. This movement gained popularity when a hashtag #metoo started creating a surge on social media platforms and women started opening up regarding their issues. The movement took hold in India around 2018 when women began sharing stories about harassment at the workplace and various aspects of their life. From actors to politicians to professionals to writers, everyone started calling out their experiences. In the past, women had mostly remained silent on such issues. This movement gave them the courage to open up and speak for themselves.

"Game of Thrones" refers to the American drama television series created by David Benioff and D. B. Weiss based on George R. R. Martin's series of novels. The series was launched in April 2011. Eight seasons have been broadcasted till now, with the latest one being launched in April 2019. The series has a huge fan following and active viewership.

"Coachella 2019" refers to the Coachella Valley Music and Arts Festival, an annual music and arts festival organized at the Empire Polo Club in Indio, California, in the Coachella Valley in the Colorado Desert. It was Coachella's 20th anniversary in 2019.

Table 1 presents the hashtags used to collect tweets of a topic and the duration for which the tweets were collected for each dataset.

We adopt the following pre-processing steps on the collected data:

i) The datasets that we utilize have many fields or tokens like user_id, status_id, created_at, screen_name, text_of_tweet, hashtags, language, account_created_at, location, retweet_count, followers_count, url, etc. We extracted required columns like followers count, re-tweet count, timestamp, etc. from the collected data based on the hashtag.
ii) Based on each topic or hashtag, tweets are segregated into time intervals using the created_at or timestamp attribute. We split tweets according to observed time intervals, like on an hourly basis or daily basis.
iii) For each time interval, we calculate the total tweets count, re-tweet count, and followers count using cumulative sum. Then, we compute the number of spreaders as the number of people who tweeted or re-tweeted and the number of non-spreaders as the total number of followers of people who tweeted or re-tweeted minus the number of followers who re-tweeted.

Table 2 presents the total number of tweets, retweets, and followers count for each dataset. Here, followers count refers to the total sum of followers of the users who tweeted on a topic.

Table 1 Datasets description

Topic	Hashtags used	Duration
Coronavirus Pandemic	#covid19 #coronavirus #coronavirusoutbreak #coronavirusPandemic	March 13, 2020 00:00 am to 23:59 pm IST
FIFA World Cup 2018	#FIFA #WorldCup #FIFA2018	July 2, 2018 to July 7, 2018
NBA Finals 2018	#NBAFinals	June 7, 2018 01:13 am to 01:58 am UTC
MeToo Movement	#metoo #SexualAssault #MarchForOurLives	February 20, 2018 19:05 pm to 19:55 pm IST
Game of thrones S8	#GameOfThrones #WinterIsComing #GoTS8 #GoT #TheMadKing	April 14, 2019 to May 14, 2019
Coachella festival 2019	#coachella #coachella19 #coachella2019	April 7, 2019 to April 23, 2019

5 Experimental setup and visualization

We generated graphs using user-follower data mined using Twitter API and then applied modified forest-fire (MFF) on the graphs. We perform the experimental results on a personnel system with configuration : 8 GB RAM and Intel(R) Core(TM) i5-8250U CPU @1.60 GHz (8 CPUs), 1.8GHz processor.

We visualized the obtained graphs using the Gephi Tool, which is an open-source network analysis, exploration, and visualization software package written in Java on the NetBeans platform. Figure 6 gives a visual understanding of how information is being propagated in a toy network. It represents a step-by-step illustration of our proposed model on

Table 2 Number of tweets, retweets and followers count

Topic	Number of tweets	Number of retweets	Total followers count
Coronavirus Pandemic	300273	694338	42819166666
FIFA world Cup 2018	242876	41927	2441853742
NBA finals 2018	19986	31439	748874447
MeToo Movement	1000	12	7620496
Game of Thrones S8	504726	12499	186646420
Coachella Festival 2019	390276	412949	22794218542

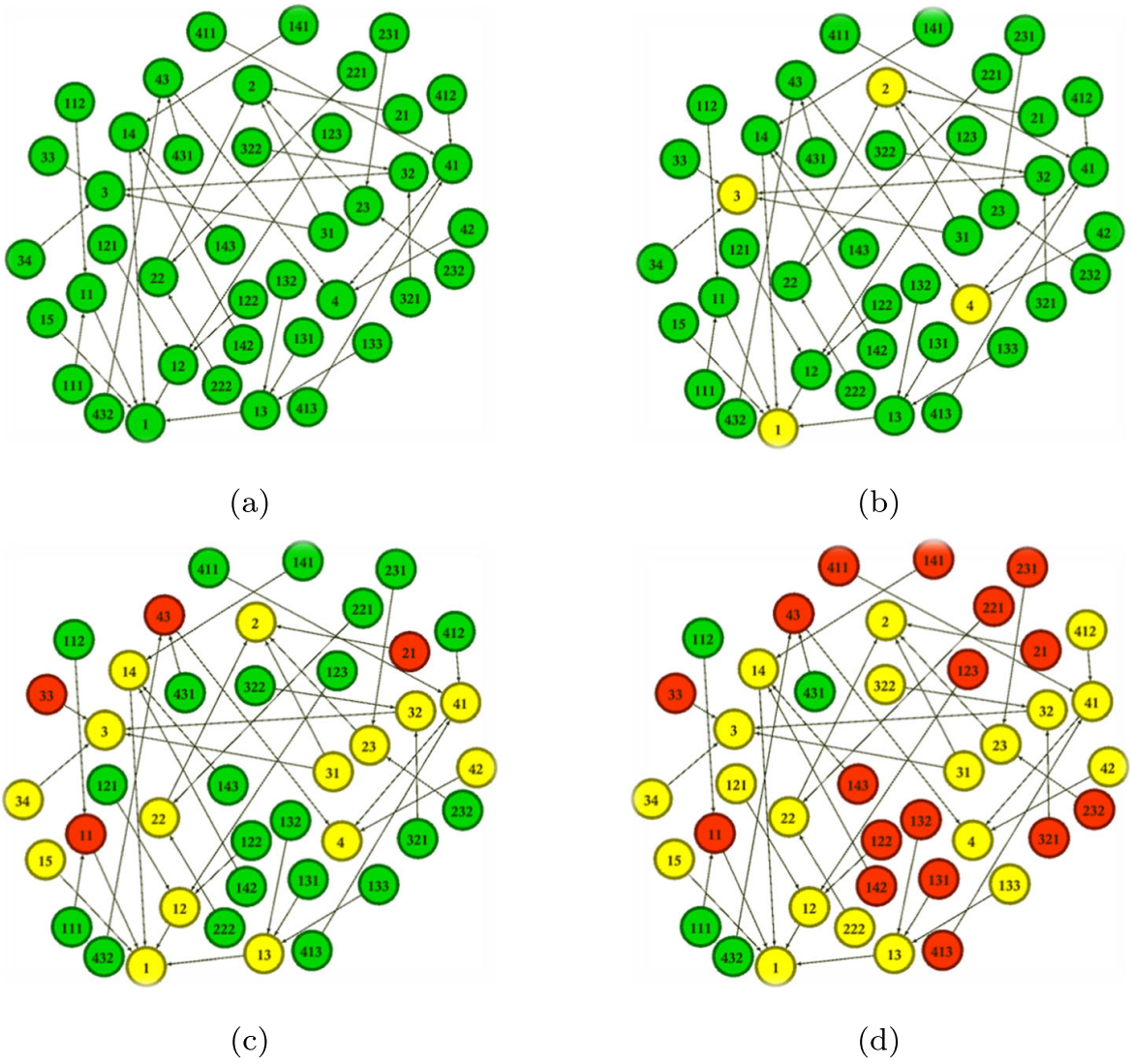


Fig. 6 Modified Forest-fire (MFF): **a** All trees are initially green means all the nodes have no knowledge about the topic or information **b** Some people have posted a tweet on a particular topic and are colored in yellow, **c** Some of the followers choose to spread the information and are colored in yellow. However, some followers decide not to spread the information further and hence, are considered burnt and colored red, and **d** The final stage of the simulation. As the fire runs over in the entire forest, the majority of the nodes become either red or yellow

the part of the Twitter network graph generated. Here, the green nodes represent *Tree*, the yellow nodes represent *Fire* and the red nodes represent *Burnt*. The red nodes are the users who are infected but choose not to spread the information further and thus, are kept in the category of *Burnt*, whereas yellow nodes are the users who are infected and are actively spreading the information further like a fire in a forest and are thus, labeled as *Fire*. It is evident from the figure that with time most of the nodes turn either yellow or red i.e., information reaches a large group of users worldwide. We present the simulation of the proposed algorithm on the toy network in the following four steps:-

Step 1: Initially, all nodes in the network are green representing trees in the forest. These nodes represent the set of Twitter users i.e. the entire network. These nodes have no knowledge about the topic.

Step 2: Then some of the nodes catch fire and are represented by yellow color. This means that some nodes have posted a tweet on that topic, thus starting the information diffusion process. These nodes are the spreaders of information on that topic.

Step 3: Now, the neighbours or followers of the spreader nodes also learn about the information. Some of them further spread the information by retweeting the tweet. These nodes have now become spreaders and hence, are colored yellow. Some of the followers choose not to spread the information further and hence, are considered burnt and colored red. These are the non-spreaders.

Step 4: With the passage of time, this fire keeps on spreading further with green nodes turning yellow or red i.e. conversion of most of the users from *Tree* to *Fire* or *Burnt*. This implies that information has reached to most of the users with time in the network. It is due to the highly inter-connected nodes that make fire spread swiftly in such a dense network of users worldwide. The extent of information spread depends on the rate at which users post tweets and retweets.

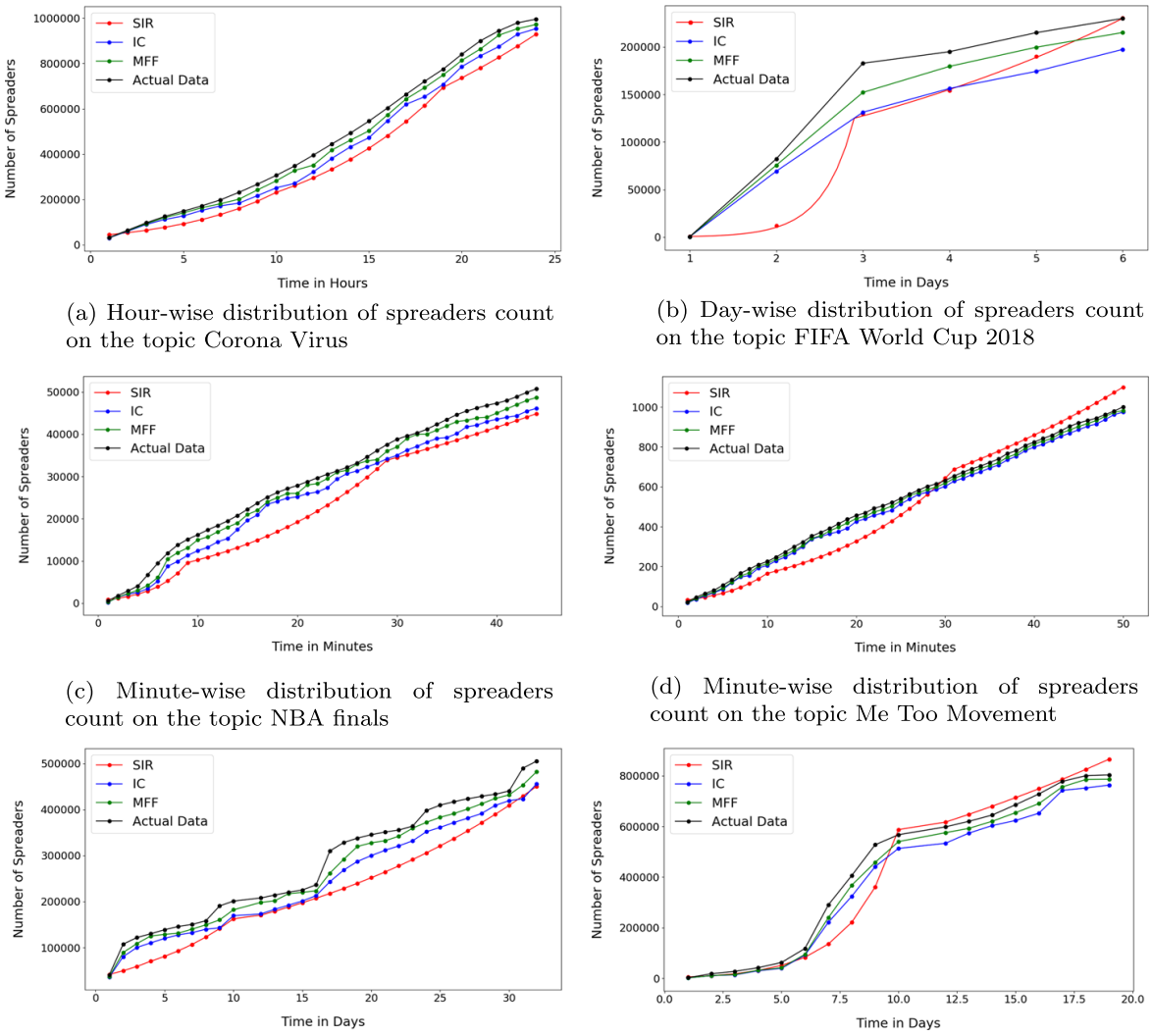
6 Results and discussion

We analyzed information spread for six real-time datasets extracted from Twitter. The tweets were clustered into several groups with respect to time (minutes, hours, days) depending on the datasets. A graph is generated based on the user-followers data extracted from Twitter. The modified forest-fire (MFF) model is applied to the generated graph. Since the data collected ranges over a few days, the number of new users joining the network is very low as compared to the twitter population of more than 300 million. Thus, the population can be considered constant during our time of research. Therefore, the probability p with which new users join the network can be considered zero. The probability f_u of any user posting a tweet on a topic is calculated as per Eq. No.(2). The factors user activity and topic significance were assigned equal weights. Thus, $w_1 = 0.5$ and $w_2 = 0.5$. The weights assigned to various factors involved in calculating similarity score in the (6) are $w_{Fg} = 0.3$, $w_{Ht} = 0.25$, $w_{Lg} = 0.25$, $w_{Lc} = 0.1$, $w_{Fw} = 0.1$. The retweet probability r_u with which a user retweets a tweet is calculated using the Eq. No. (8). Here, different measures are assigned different weights, $w_{IM} = 0.25$, $w_{SS} = 0.3$, $w_{UA} = 0.2$, and $w_{TS} = 0.25$. These weight constants are obtained by performing experiments on the Twitter data which gives optimal results.

The modified forest-fire (MFF) algorithm is applied on the Twitter graph to distinguish between spreaders and non-spreaders users. At $t = 0$, those users who had posted a tweet about the given topic were labeled as *Fire*. There can be some users who can get information from outside sources and post a tweet with probability f_u greater than or equal f_0. Here, we consider the value of threshold, i.e, $f_0 = 0.5$. Thus, we also considered the case of self-infection. The followers of an infected user are segregated into spreaders and non-spreaders based on the retweet probability (r_u). For each follower of the user who posted a tweet or retweet, we calculated the retweet probability using the formula mentioned in (8). We consider those followers for whom the retweet probability is greater than a minimum threshold (r_0) of 0.5 as spreaders nodes; otherwise, nodes are labeled as non-spreaders. The overall informed or infected count is the sum of spreaders and non-spreaders. The count of spreader nodes, non-spreader nodes and total informed nodes is plotted against time. The experimental values are found to be very close to actual values. For each topic or

hashtag, the actual data is the cumulative sum of the number of tweets and re-tweets over time. Spreaders are the people who actively take part in information diffusion process by spreading the information. In the case of twitter, these are the people who post tweets on a particular topic or retweet the tweets of users whom they follow. We compare the spreaders count predicted by our model with other classical models in literature. The models used for comparison are Independent Cascade (IC) Model and Susceptible-Infected-Recovered (SIR) Model. Both of these are prominent models used to study information propagation and are known to produce satisfactory results. We utilize these two models because both of these predict the active spreaders in a network.

Figure 7 illustrates the spreaders count predicted by the modified forest-fire (MFF) model, independent cascades (IC) model, and susceptible-infected-recovered (SIR) model along with the actual spreaders count as calculated from the tweets data. The actual data comprises of all the tweets and retweets posted by users. For each topic or hashtag, the actual data is the cumulative sum of the number of tweets and re-tweets over time. We considered people who tweeted or re-tweeted on a topic or hashtag as spreaders of information. Thus for spreaders, the actual data comprises of the number of tweets and retweets.



(a) Hour-wise distribution of spreaders count on the topic Corona Virus

(b) Day-wise distribution of spreaders count on the topic FIFA World Cup 2018

(c) Minute-wise distribution of spreaders count on the topic NBA finals

(d) Minute-wise distribution of spreaders count on the topic Me Too Movement

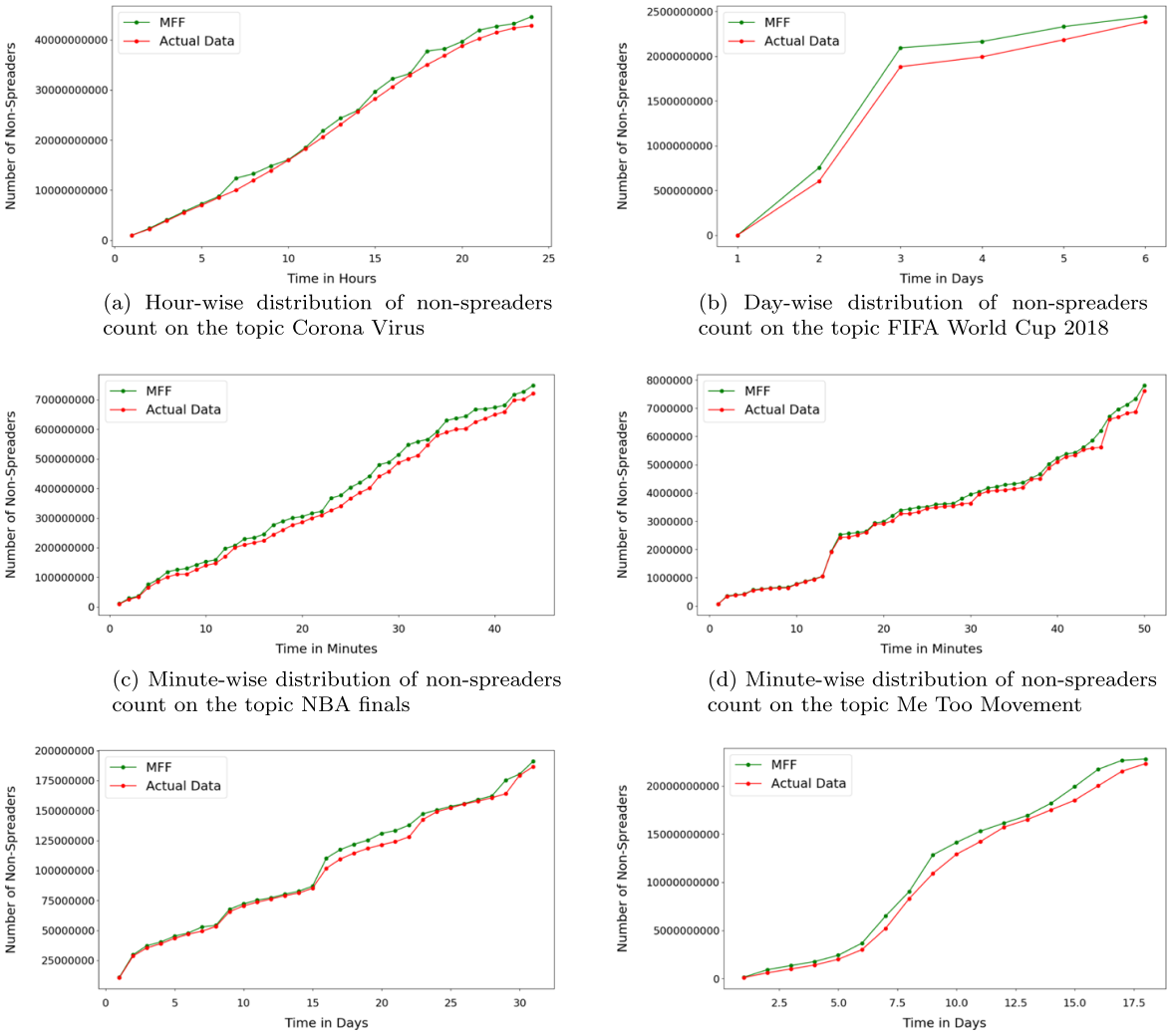
(e) Day-wise distribution of spreaders count on the topic Game of Thrones

(f) Day-wise distribution of spreaders count on the topic Coachella 2019

Fig. 7 Comparison of distribution of spreader nodes with respect to time according to various algorithms for various topics

Non-spreaders are the people who have knowledge of a topic but choose not to spread the information further. These are the followers of people who tweeted on a topic. These people chose not to retweet further. Thus, for non-spreaders, actual data is the difference of followers count and retweet count of the users tweeted on a topic. The total infected count is the sum of spreaders and non-spreaders. For the total infected count, we took the sum of tweets, retweets, and followers count. The model has been validated for six datasets as described in Table 1. Figure 7a illustrates the hour-wise spreaders count for the topic Corona Virus Pandemic. It can be seen that our MFF model predicts values close to the actual data followed by IC Model and SIR Model. Figure 7b shows the day-wise spreaders count for the topic FIFA World Cup 2018. It can be observed that SIR Model predicts very low spreaders count initially but later on, the count increased rapidly reaching close to the actual data. The MFF model predicts better results throughout the duration. The FIFA World Cup data was collected from July 2, 2018 to July 7, 2018. There were no football matches on July 4 and 5. The frequency of tweets was high during the match or on the day of the match. Thus, after day 3, there is a drop in tweet count. As the infected count is a cumulative sum of tweets over days, there is a plateau in the graph after day 3. Figure 7c represents the minute-wise spreaders count for the topic NBA Finals. Here, SIR Model underestimates the spreaders count. Our MFF Model gives better results than IC Model and SIR Model. Figure 7d illustrates the minute-wise spreaders count for the topic Me Too Movement. The SIR model underestimates the spreaders count initially and then overestimates it. The IC model and Modified Forest-fire model predicted results close to the actual data. Figure 7e shows the day-wise spreaders count for the topic Game of Thrones Season 08. It can be observed that MFF Model was able to fit the actual data better than IC Model and SIR Model. Figure 7f illustrates the day-wise spreaders count for the topic Coachella Music and Arts Festival. The SIR model initially underestimated the spreaders count but in later stages, gave optimal results. The MFF Model gave satisfactory results throughout. It can be observed from Fig. 7, that our model (MFF) predicted results which are close to the actual data. The SIR model produces varying results, sometimes overestimating the spreaders population and sometimes underestimating it. It can be seen from the graphs, that our model performed better than IC and SIR model as the values of spreaders count predicted were close to the actual values. Thus, it can be concluded that our model predicts active spreaders efficiently and can be used for modeling for information diffusion in real-life scenario.

Non-spreaders are people who have knowledge of the information but do not pass it on to others. The classical information diffusion models like IC and SIR model do not consider non-spreaders as infected. They only take in to account the users who spread the infection. However, our proposed MFF model considers non-spreaders in the information diffusion process. The number of non-spreaders with respect to time are also plotted for the six datasets as depicted in Table 1. Figure 8 shows the plot of number of non-spreaders vs. time duration either in minutes, hours or days for various topics. It can be seen from the results obtained on all the topics mentioned, that the non-spreaders count predicted by our model is in harmony with the actual count calculated from the tweets data. Here, the actual data is the difference of followers count and retweet count of the users who tweeted on a topic. It is observed from the results that the non-spreaders count is much higher than the spreaders count, indicating that only a few people spread the information further as compared to a large number of people who have knowledge about the information. It can be deduced that the actions of a few people can lead to large-scale dissemination of information. Thus, segregating spreaders and non-spreaders is useful to predict information diffusion.



(a) Hour-wise distribution of non-spreaders count on the topic Corona Virus

(b) Day-wise distribution of non-spreaders count on the topic FIFA World Cup 2018

(c) Minute-wise distribution of non-spreaders count on the topic NBA finals

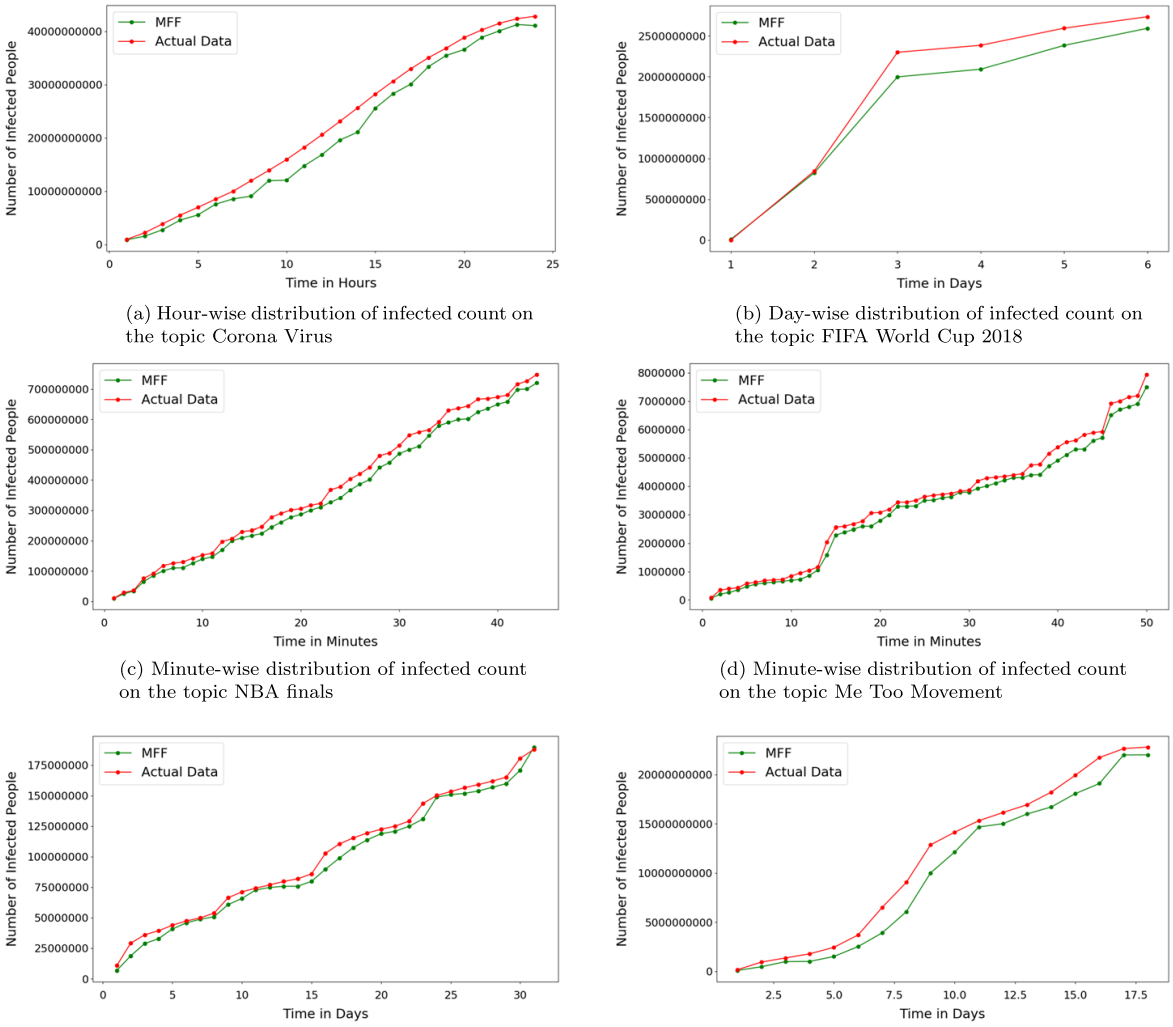
(d) Minute-wise distribution of non-spreaders count on the topic Me Too Movement

(e) Day-wise distribution of non-spreaders count on the topic Game of Thrones

(f) Day-wise distribution of non-spreaders count on the topic Coachella 2019

Fig. 8 Distribution of non-spreader nodes with respect to time for various topics

Figure 9 represent the plot of the total infected or informed count with respect to time for various topics. Here, the infected or informed count includes the number of spreaders and non-spreaders. These figures display the extent of information diffusion as predicted by our proposed model (MFF) in comparison to actual data. The actual data comprises of a number of users who posted a tweet/re-tweet on a particular topic and their respective followers count. Figure 9a shows the plot of infected users with respect to time in hours for the topic Corona Virus or COVID-19. COVID-19 is a global topic affecting the entire world. Thus, the infected count increased rapidly with time. Figure 9b represents the distribution of infected nodes related to the topic FIFA world Cup 2018 with respect to time in days. FIFA is a global topic bringing together football enthusiasts all over the world, thus attracting many tweets. Figure 9c illustrates the distribution of infected users against time in minutes for the topic of the NBA Finals. The infected count increased rapidly with each minute during the match. Figure 9d represents the plot of infected users related to the topic of Me Too Movement with respect to time in minutes. The infected count was low initially but then increased gradually. Figure 9e represents the plot of infected users related to the topic of Game of Thrones Season 8 with respect to time in days. The infected count kept on increasing with time. Figure 9f represents the plot of infected users related to the topic of



(a) Hour-wise distribution of infected count on the topic Corona Virus

(b) Day-wise distribution of infected count on the topic FIFA World Cup 2018

(c) Minute-wise distribution of infected count on the topic NBA finals

(d) Minute-wise distribution of infected count on the topic Me Too Movement

(e) Day-wise distribution of infected count on the topic Game of Thrones

(f) Day-wise distribution of infected count on the topic Coachella 2019

Fig. 9 Distribution of infected or informed nodes with respect to time for various topics

Coachella music and arts festival with respect to time in days. The infected count increased with a low rate initially, that is, a week before the festival but increased rapidly as the festival inched close. The number of tweets was high during the festival. It can be observed from the plots that our proposed MFF algorithm models the process of information diffusion efficiently. The predicted results very well fit the actual data. Also, it can be deduced from the plots that total infected count is much higher than the actual spreaders, thus, validating the fact that even a few people can lead to massive dissemination of information. Our proposed model considers the infected or informed people as the sum of spreaders and non-spreaders. This is because, in the case of online social networks, people can be infected i.e., aware of information even if they are not spreading further. Hence, our model intends to efficiently predict information diffusion in the case of online social networks like Twitter.

7 Conclusion

In this paper, we modeled information diffusion in online social networks using a nature-inspired model named as modified forest-fire model. We exhibited the flow of information in a network using the modified forest-fire model. We divided the entire population into

four categories: *Empty*, *Tree*, *Fire*, and *Burnt*. We also segregated the informed population into spreaders and non-spreaders. The study on the spreading of information in real-world data-sets is similar to as illustrated by our model, and thus, information spreads like fire in a network. We performed analysis on a densely connected and popular network of Twitter by using different hashtags. The modified forest-fire model resembles the real-world spreading of information diffusion in highly complex networks like Twitter.

References

Bakshy, E., Rosenn, I., Marlow, C., Adamic, L. (2012). The role of social networks in information diffusion.

Barabási, A.L. (2016). Network science. Cambridge University Press.

Biswas, M.H.A., Paiva, L.T., De Pinho, M.D.R. (2014). A SEIR model for control of infectious diseases with constraints. *Mathematical Biosciences & Engineering*, *11*(4), 761.

Cai, G., Wang, R., Qiang, B. (2012). Online social network evolving model based on damping factor. *Procedia Computer Science*, *9*, 1338–1344.

Chakraborty, A., Dutta, T., Mondal, S., Nath, A. (2018). Application of graph theory in social media. *International Journal of Computer Sciences and Engineering*, *6*, 722–729.

D'Ambrogio, A., Gaudio, P., Gelfusa, M., Luglio, M., Malizia, A., Roseti, C., Zampognaro, F., Giglio, A., Pieroni, A., Marsella, S. (2016). Use of integrated technologies for fire monitoring and first alert. In *2016 IEEE 10th International Conference on Application of Information and Communication Technologies (AICT)* (pp. 1–5): IEEE.

Daley, D.J., & Gani, J. (2001). Epidemic modelling: an introduction (Vol. 15). Cambridge University Press.

de Abreu, L.F. (2019). Game of Thrones S8 (Twitter) 7th April 2019 - 28th May 2019, US, Version 1. https://www.kaggle.com/monogenea/game-of-thrones-twitter.

Ding, C., & Li, K. (2018). Topologically biased random walk for diffusions on multiplex networks. *Journal of Computational Science*, *28*, 343–356.

Ding, C., & Li, K. (2018). Topologically biased random walk for diffusions on multiplex networks. *Journal of Computational Science*, *28*, 343–356.

Fischer, A., Korejwa, A., Koch, J., Spies, T., Olsen, C., White, E., Jacobs, D. (2013). Using the forest, people, fire agent-based social network model to investigate interactions in social-ecological systems. *Practicing Anthropology*, *35*(1), 8–13.

Guille, A., Hacid, H., Favre, C. (2013). Predicting the temporal dynamics of information diffusion in social networks. arXiv:1302.5235.

Guille, A., Hacid, H., Favre, C., Zighed, D.A. (2013). Information diffusion in online social networks: a survey. *ACM Sigmod Record*, *42*(2), 17–28.

Hawley, L.F. (1926). Theoretical considerations regarding factors which influence forest fires. *Journal of Forestry*, *24*(7), 756–763.

He, D., & Liu, X. (2020). Novel competitive information propagation macro mathematical model in online social network. *Journal of Computational Science*, *41*, 101089.

He, Q., Wang, X., Mao, F., Lv, J., Cai, Y., Huang, M., Xu, Q. (2020). CAOM: A community-based approach to tackle opinion maximization for social networks. *Information Sciences*, *513*, 252–269.

Hoang, T.B.N., & Mothe, J. (2018). Predicting information diffusion on Twitter–Analysis of predictive features. *Journal of Computational Science*, *28*, 257–264.

Hu, Y., Aiello, M., Hu, C. (2018). Information diffusion in online social networks: a compilation. *Journal of Computational Science*, *28*, 204–205.

Indu, V., & Thampi, S.M. (2019). A nature-inspired approach based on Forest Fire model for modeling rumor propagation in social networks. *Journal of Network and Computer Applications*, *125*, 28–41.

Kanga, S., & Singh, S.K. (2017). Forest fire simulation modeling using remote sensing & GIS.International Journal of Advanced Research in Computer Science 8 (5).

Kimura, M., Saito, K., Ohara, K., Motoda, H. (2013). Learning to predict opinion share and detect anti-majority opinionists in social networks. *Journal of Intelligent Information Systems*, *41*(1), 5–37.

Kuang, L., Tang, X., Guo, K. (2014). Predicting the times of retweeting in microblogs. Mathematical Problems in Engineering 2014.

Kumar, S., & Panda, B.S. (2020). Identifying influential nodes in Social Networks: Neighborhood Coreness based voting approach. Physica A: Statistical Mechanics and its Applications, pp. 124215.

Liu, Y., Diao, S.M., Zhu, Y.X., Liu, Q. (2016). SHIR Competitive information diffusion model for online social media. *Physica A: Statistical Mechanics and its Applications*, *461*, 543–553.

Mashiach, L.T., & Sharma, A. (2020). Selecting user posts related to trending topics on online social networks. U.S. Patent 10,,535,106.

Nesi, P., Pantaleo, G., Paoli, I., Zaza, I. (2018). Assessing the reTweet proneness of tweets: predictive models for retweeting. *Multimedia Tools and Applications*, *77*(20), 26371–26396.

Newman, M.E.J. (2010). *Networks: An Introduction*, (p. 18). New York: Oxford University Press.

Pattanayak, H.S., Sangal, A.L., Verma, H.K. (2019). Community detection in social networks based on fire propagation. *Swarm and Evolutionary Computation*, *44*, 31–48.

Pesic, P. (2018). Tweets during Cavaliers vs Warriors 3rd game of the 2018 NBA Finals #NBAFInals, Version 24. https://www.kaggle.com/xvivancos/tweets-during-cavaliers-vs-warriors.

Pesic, P. (2019). Coachella 2019 Tweets, Version 2. https://www.kaggle.com/pdp2600/coachella-2019-tweets.

Ramirez, V. (2018). MeToo Dataset. Retrieved from https://data.world/bikthor/metoo.

Rituparna (2018). FIFA World Cup 2018 Tweets, Version 4. https://www.kaggle.com/rgupta09/world-cup-2018-tweets.

Rui, X., Hui, S., Yu, X., Zhang, G., Wu, B. (2018). Forest fire spread simulation algorithm based on cellular automata. *Natural Hazards*, *91*(1), 309–319.

Saito, K., Ohara, K., Kimura, M., Motoda, H. (2015). Change point detection for burst analysis from an observed information diffusion sequence of tweets. *Journal of Intelligent Information Systems*, *44*(2), 243–269.

Sankar, C.P., & Kumar, K.S. (2016). Learning from bees: an approach for influence maximization on viral campaigns. *PloS One*, *11*(12), e0168125.

Smith, S. (2020). Coronavirus (covid19) Tweets-Tweets using hashtags associated with Coronavirus, Version 13. https://www.kaggle.com/smid80/coronavirus-covid19-tweets.

Stai, E., Milaiou, E., Karyotis, V., Papavassiliou, S. (2018). Temporal dynamics of information diffusion in twitter: Modeling and experimentation. *IEEE Transactions on Computational Social Systems*, *5*(1), 256–264.

Wang, X.F., Li, X., Chen, G.R. (2012). *Network Science: an Introduction*, (pp. 10–15). Beijing: Higher Education Press.

Wang, J., & Liu, X. (2011). The improvement of computer algorithm for forest fire model based on cellular automata. In *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)* (pp. 2567–2570): IEEE.

Zhou, T., Fu, Z., Wang, B.H. (2006). Epidemic dynamics on complex networks. *Prog Natural Science*, *16*, 452–457.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Sanjay Kumar[1,2] · Muskan Saini[3] · Muskan Goel[4] · B. S. Panda[2]

Muskan Saini
muskansaini257@gmail.com

Muskan Goel
muskangoel210499@gmail.coms

B. S. Panda
bspanda@maths.iitd.ac.in

1 Department of Computer Science and Engineering, Delhi Technological University, Main Bawana Road, New Delhi, 110042, India

2 Computer Science and Application Group, Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, 110016, India

3 Microsoft India, Hyderabad, Telangana, 500032, India

4 Microsoft India, Banglore, Karnataka, 560025, India

Multi-objective Reinforcement Learning based approach for User-Centric Power Optimization in Smart Home Environments

Saurabh Gupta*, Siddhant Bhambri†, Karan Dhingra*, Arun Balaji Buduru* and Ponnurangam Kumaraguru*
*Indraprastha Institute of Information Technology, Delhi, India
{saurabhg,karan19025,arunb,pk}@iiitd.ac.in
†Delhi Technological University, Delhi, India
siddhant_bt2k16@dtu.ac.in

arXiv:2009.13854v1 [cs.AI] 29 Sep 2020

***Abstract*—Smart homes require every device inside them to be connected with each other at all times, which leads to a lot of power wastage on a daily basis. As the devices inside a smart home increase, it becomes difficult for the user to control or operate every individual device optimally. Therefore, users generally rely on power management systems for such optimization but often are not satisfied with the results. In this paper, we present a novel multi-objective reinforcement learning framework with two-fold objectives of minimizing power consumption and maximizing user satisfaction. The framework explores the trade-off between the two objectives and converges to a better power management policy when both objectives are considered while finding an optimal policy. We experiment on real-world smart home data, and show that the multi-objective approaches: i) establish trade-off between the two objectives, ii) achieve better combined user satisfaction and power consumption than single-objective approaches. We also show that the devices that are used regularly and have several fluctuations in device modes at regular intervals should be targeted for optimization, and the experiments on data from other smart homes fetch similar results, hence ensuring transfer-ability of the proposed framework.**

***Keywords*-power management; multi-objective reinforcement learning (MORL); sequential decision-making**

I. Introduction

The amount of power consumption in households (residential) is among the top three in world electricity consumption [1], and is ever increasing with the increase in demand of smart homes and IoT (Internet of Things) devices. According to the United States Department of Energy (DoE), the average household consumes 90 million units of power a year, and much of that power is wasted [2]. Habits like leaving lights on when we leave rooms, forgetting to turn off televisions or computers when not in use, etc., are major reasons behind such wastage [3]. Therefore, there is a need for power controllers that can take actions like turning devices on and off, or changing devices' modes of operation on behalf of users to achieve a goal like optimized consumption.

In the past, researchers have used traditional reinforcement learning for several power optimization tasks. For example, [4] proposed a model-free constrained RL approach for online power management. [5] presented another similar algorithm that requires no prior information of the workload and dynamically adapts to the environment to achieve autonomous power management. [6] proposed an RL based technique that performs simultaneous online management of both performance and power consumption. The authors applied RL in a realistic laboratory testbed to find the optimal policy. None of these techniques towards power optimization are used for smart home power management, and they do not consider user satisfaction while finding optimal policies.

However, power management in a smart home is a problem that needs to solve two tasks with different rewards simultaneously: minimize power consumption and maximize user satisfaction. It is important for a power controller to consider user preferences as well, i.e., the goal of minimal power consumption must be achieved but not at the expense of user satisfaction. The scenario can be formulated as a MORL problem where sequential decision making is required with multiple objectives.

Our contribution: In this paper, we propose, for the first time, a novel multi-objective reinforcement learning (MORL) approach for power management inside a smart home with two objectives: minimize power consumption and maximize user satisfaction. In a MORL problem, an action on the environment results in multiple rewards. The agent (power controller) learns an optimal policy from these rewards using a variation of Q-learning [7]. Since the objectives are contrasting, there is a trade-off between the two, and based on their importance, optimization priorities are set. We use an overall reward function to incorporate these optimization priorities, which is a weighted sum of the two rewards R_E representing power consumption, and R_U representing user satisfaction. We specifically focus on the weighted-sum method [8] for multi-objective optimization and compare the results with single objective strategies.

We evaluate our proposed methods on the Smart* data set for sustainability[9]. The data samples include device-level real-world power consumption values in several smart homes, named as A, B, C, ..., H recorded at every 30 minutes. We show the effectiveness of our approach on data from smart home A, and demonstrate transferability of

experiments on smart homes B and C. We use Q learning with individual objectives (single policy single objective approaches) as a baseline reference to compare the proposed single policy multi-objective approaches. We also define a metric "clash rate" for evaluating user satisfaction in the predicted policy at each episode.

The remainder of this paper is organized as follows. Section II gives some background of traditional and multi-objective reinforcement learning (MORL). Section III explains our problem formulation followed by our algorithm to solve the optimization problem in Section IV. The experiments and results are presented in Section V-VII. We conclude our work in Section VIII.

II. Background

In this section, we discuss traditional reinforcement learning with Q-learning, an algorithm widely used to solve traditional RL problems. Then we introduce the concepts of multi-objective reinforcement learning (MORL) and how it differs from the traditional RL.

A. Traditional Reinforcement Learning

Traditional reinforcement learning [10] mimics the natural learning style of trial-and-error by interacting with an environment (static or dynamic) and receiving feedback based on an action. The components of reinforcement learning are:

- An Agent;
- A finite state space S;
- A set of available actions for the agent A;
- A reward function $R : S \times A \rightarrow R$.

The agent's objective is to maximize its average long-term reward. It is achieved by learning a policy π, which is a mapping between the states and the actions. In our problem, one goal is to minimize the power consumption of a smart home, and the other is to maximize user satisfaction. But, in a traditional reinforcement learning setting, the two goals are independent. An agent can either minimize power consumption, or it can maximize user satisfaction.

Q-learning [7] is a widely known algorithm used to solve sequential decision-making RL problems. In each step, on the successful execution of every action a, the environment yields a reward R, which indicates the value of a state transition. The issued reward can be positive or negative. The agent keeps a value function $Q^{\pi}(s, a)$ for each state-action pair. Learning to act in the environment will make the agent choose actions to maximize long-term rewards. Based on this value function, the agent decides its immediate action. The Q-value for each state-action pair is initially chosen during the problem formulation, and later, it is updated with each taken action and its issued reward. The value function is given by the following Bellman equation:

$$Q^{\pi}(s, a) = R(s, a) + \gamma \max_{a} Q^{\pi}(s', a) \tag{1}$$

where $R(s, a)$ is the reward issued after taking action a in state s, s' is the successive state of s, and γ is the discount factor used due to the different influences of future rewards on the present value.

The optimal state-action value function is defined as:

$$Q^{*}(s, a) = max_{\pi} Q^{\pi}(s, a) \tag{2}$$

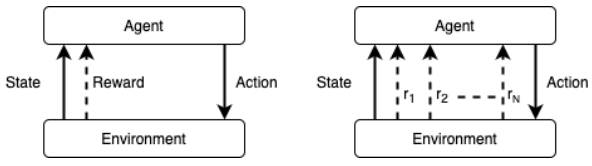
When $Q^{*}(s, a)$ is obtained, the optimal policy π^{*} can be computed by:

$$\pi^{*}(s) = \operatorname*{argmax}_{a} Q^{*}(s, a)$$

B. Multi-objective Reinforcement Learning

Reinforcement learning is a machine learning paradigm that helps with sequential decision making under several uncertainties and aims to achieve a single long-term objective. However, due to the complex requirements of real-world control systems, often times, there are two (or more) conflicting objectives. For example, in our case of smart home power management system the controller has two goals: i) to minimize energy consumption of the smart home, ii) to maximize user comfort by moving to states preferred by the user. In reinforcement learning, problems of this nature having more than one conflicting objectives are called multi-objective reinforcement learning problems (MORL).

MORL is different from tradtional RL in that there are two or more objectives to be optimized simultaneously by the learning agent. [11] provides an architecture for a MORL problem, where reward is provided for the learning agent at each step. Figure 1 shows the difference between architectures of traditional RL (Figure 1(a)) and multi-objective RL (Figure 1(b)). In MORL (Figure 1(b)), there are N objectives and $r_i (1 \leq i \leq N)$ is the i^{th} reward signal provided by the environment. The architecture illustrates a single agent that has to find an optimal policy for a set of multiple objectives simultaneously. The objectives can be conflicting, as in our case, or they can be independent as well.



(a) Traditional RL (b) Multi-Objective RL

Figure 1. Architecture of Reinforcement Learning.

For each objective i $(1 \leq i \leq N)$ and a stationary policy π, there is a corresponding state-action value function Q_i^{π}, which satisfies equation 1.

Let the combined value function for MORL is:

$$MQ^{\pi}(s, a) = [Q_1^{\pi}(s, a), Q_2^{\pi}(s, a), ..., Q_N^{\pi}(s, a)]$$

where $MQ^{\pi}(s,a)$ is a vector and it also satisfies the Bellman equation (1). Then the optimal state-action function will be given as:

$$MQ^{*}(s,a) = \max_{\pi} MQ^{\pi}(s,a) \tag{3}$$

and the optimal policy π^{*} can be obtained by:

$$\pi^{*}(s) = \operatorname*{argmax}_{a} MQ^{*}(s,a) \tag{4}$$

MORL is a combination of multi-objective optimization methods and RL techniques to solve sequential decision making problems with multiple conflicting objectives. We will justify why we formulate smart home power management as a MORL problem in the next section.

III. Problem Formulation

The case of a smart home power management system is a multi-objective problem with two objectives, viz., minimizing power consumption and maximizing user satisfaction. Ideally, a controller will try to reduce the power consumption as much as it can, given an optimization goal. The trivial solution for the controller will be to turn off all the devices that operate in the smart home. However, this state might not be desirable by the user. Therefore, it is important for a controller to consider user preferences as well. Hence, the goal of minimal power consumption must be achieved by establishing a trade-off with user satisfaction, and not at it's expense.

Based on the importance of an objective function, optimization priorities must be ensured while designing the policies. After appropriately expressing the preferences, we have to design an efficient algorithm that can solve the sequential decision making problems based on observed state transition data.

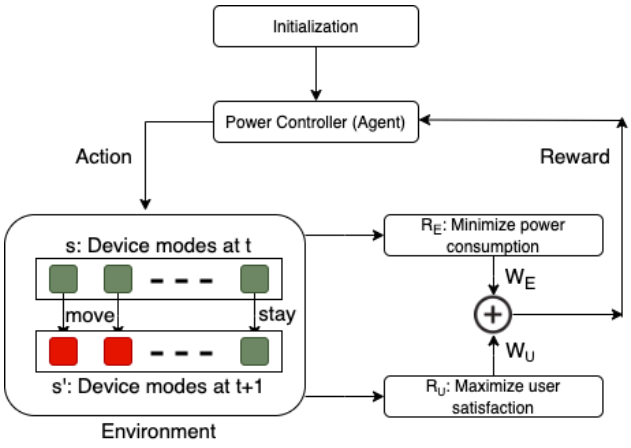


Figure 2. Illustration of Power controller. Detailed explanation of each component is available in the text.

A. Environment

Smart homes usually have smart meters to measure the power consumption for each device operating within it. The power consumption values for every device is independent, and take a fixed number of discrete values. This is because each device operates only in a fixed number of modes and their power consumption in a specific mode remains the same. For example, a simple furnace has only two modes, ON and OFF. In OFF mode, it consumes no power while in ON mode it consumes x (say) units of power. We are assuming that the consumption x remains the same and no degradation of device happens over time, hence causing more energy consumption. Similarly, a washing machine can have three modes of operation, viz, standby, washer and dryer. Let's assume a smart home has N devices. A state in an environment is a vector of energy consumption values (in whatever device modes they are in) of these N devices, as depicted by blocks s and s' in Figure 2.

Let's assume the number of device modes each one operates in is given by a set $D = \{nD_1, nD_2, ..., nD_N\}$

$$Total\ number\ of\ states\ in\ state-space = \prod_{i=1}^{N} nD_i$$

Increasing the number of devices, or just their modes of operation can lead to state-space explosion. Therefore, in our techniques we choose devices selectively and use data processing (explained in Section V) to avoid state-space explosion.

B. Power Controller (Agent)

The Agent is a power controller that can change the mode of operation of any of the N devices, consequently changing the energy consumption value. For example, a power controller can turn off the furnace, if it is on, or switch the washing machine to dryer mode from some other mode. However, an agent can also choose not to do anything. Therefore, the Agent can perform either of the two actions, i.e., $MOVE$ or $STAY$, on each device in a state s, to move the environment to a state s'. For example, in Figure 2, the controller changes the mode of operation of device 1 and device 2, from $green$ to red by $MOVE$ action, and chooses to keep the Nth device it its current state by $STAY$ action.

C. Reward

Whenever the proposed agent takes an action on the environment, a reward is calculated on the basis of the state chosen by the controller and the ground truth state from the Smart House Dataset. Since there are two distinct objectives, we formulate the reward functions to incorporate the power consumption and user satisfaction. Every update of the state-action value function (Equation 1) is dependent on the reward. Therefore, by integrating the optimizations in reward

function, the agent learns the trade-off between optimization priorities for an optimal state. First we introduce both rewards separately and then we combine them to form a single reward, as shown in Figure 2.

1) Minimizing Power Consumption: Let's say the power consumption in the predicted and ground truth state is $P_{s'}$, and P_s respectively. The reward, R_E is given as:

$$R_E = \frac{\sum_d -(P_{s'} - P_s)}{D} \tag{5}$$

which is the average difference of power consumed by the D devices between predicted state and the desired state. As the agent always tries to maximize the reward, thus we negate the sum in order to achieve state which consumes less power than what the user had chosen. By negating, the state with the least electricity consumption becomes goal state for the power controller.

2) Maximizing User Satisfaction: To model user behavior, we compute the Euclidean distance between the predicted state and the ground truth state. The reward

$$R_U = \frac{\sum_d |s' - s|}{D} \tag{6}$$

where $d \in D$ $devices$.

3) Overall reward: We take a weighted combination of both the rewards, R_E (Equation 5) and R_U (Equation 6), and define overall reward as:

$$R = W_E * R_E + W_U * R_U \tag{7}$$

where W_E and W_U are the weights to manipulate the optimization priorities of the two objectives. These weights are treated as hyper parameters during experimentation.

D. Evaluation

The evaluation of power controller's performance is two-fold due to the multi-objective nature of the optimization problem. The reward R_E represents negation of power consumption, therefore, a policy with more positive R_E value is desired. Hence, as we increase the number of iterations, the value of R_E should increase.

Similarly, the reward R_U represents the likelihood that the next state predicted by the controller, S_U^{pred}, matches to the next state that user prefers, S_U^{real}. However, to evaluate R_U, we introduce a term called "clash rate" to get a device level view of clashes. To calculate clash rate:

$$clash\ rate = \sum (S_U^{real} == S_U^{pred}) \tag{8}$$

where "==" is an element wise comparison that assigns 1 if values do not match and 0 otherwise, and returns an array with 1's and 0's.

For example, let us say the user wants next state to be $S_U^{real} = [D_1, D_2, D_3, D_3, D_1]$, where $D_i s$ represent the device modes at this state. Now, the controller takes an action on the environment to change its state to $S_U^{pred} = [D_2, D_1, D_3, D_1, D_1]$. The clash rate in this case is 3, as the device modes at index 0, 1, and 3 do not match (assume the vectors are indexed starting from 0). As we increase the iterations to train the power controller more, the clash rate should decrease.

IV. Solutions

MORL approaches can be divided into two groups based on the number of policies to be learned [12]:single policy and multiple policy approaches. In our case, the objectives are contrasting, and the availability of data allows us to create a sufficiently good representation of the environment. Therefore, we focus on a single policy approach to solve it.

The aim of single policy approaches is to obtain the best policy which satisfies the optimization priorities as set by the designer, or defined by the application domain. Therefore, based on varying optimization priorities we implemented four variations of a single policy algorithm to find an optimal policy for our two-fold objectives of minimum power consumption and maximum user satisfaction. A single policy approach to solve MORL problems is to formalize an objective function $TQ(s, a)$, which can represent overall preferences in optimization. The approach is very similar to Q-learning with a few modifications, as shown in Algorithm 1. The objective function $TQ(s, a)$ is given as the summation of Q-values for all the objectives, and is given as:

$$TQ(s, a) = \sum_{i=1}^{N} Q_i(s, a) \tag{9}$$

Algorithm 1 Single Policy Approach to solve MORL

```
K: the maximum number of episodes
N: the number of objectives
Initialize TQ(s, a);
Initialize Q_i(s, a), ∀(i < N);
for each episode j ranging from 1 to K do
    Fetch s_0^j, s_1^j from samples;
    Choose a using TQ(s, a) policy using ϵ_greedy approach;
    Take action a, s';
    Compute reward [r_1, r_2,..., r_N ] based on s_1^j, and s';

    for i = 1, 2, ..., N do
        Q_i(s, a) ← Q_i(s, a) + α[r_i + γ max_a' Q_i(s', a') − Q_i(s, a)];
    end for
    Compute TQ(s, a);
    s ← s';
end for
```

As discussed in Section III-C, we incorporate the optimization priorities using weights W_E and W_U in the reward

function. Since $Q(s, a)$ is dependent on the reward function, and $TQ(s, a)$ on $Q(s, a)$, any change in the weight values in equation 7 will result in a change of values in $TQ(s, a)$. They are defined as:

A. Single Policy Single Objective

As a baseline reference, we implement single policy approach with single objectives. Recall equation 7, the overall reward is defined as the sum of two rewards, one for minimizing power consumption and other for maximizing user satisfaction. Therefore, for single policy with one objective taken at a time has two cases:

1) Power Consumption Minimization: To implement this, we give 100% optimization priority to power consumption, and set W_E and W_U to 1 and 0, respectively in equation 7.

2) User Satisfaction Maximization: To implement this, we give 100% optimization priority to user satisfaction, and set W_E and W_U to 0 and 1, respectively in equation 7.

B. Single Policy Multi Objective

The goal of the power controller is to achieve a multi-objective optimization. Therefore, we consider two cases:

1) Equal weights: The case where both objectives are equally as important, and the power controller tries to optimize both. Both W_E and W_U are set to 1. Based on the policy calculated by the power controller, the action with the maximum summed values is chosen to be executed.

2) Weighted Sum: The weighted sum approach is proven to be effective with multiple objectives in the past. [13] used it to combine seven vehicle overlapping objectives, and [14] used it with a combination of three objectives, viz., degree of the crowd in an elevator, the waiting time, and the number of start-ends. The approach modifies equation 9 as:

$$TQ(s, a) = \sum_{i=1}^{N} W_i Q_i(s, a)$$

In our case, the $W_i s$ are W_E and W_U, and we experiment with different values of both to get the best results.

V. Environment Setup

We evaluate the proposed solutions in Section IV using the Smart* data set for sustainability [9]. As a baseline reference, we consider Q-learning with single objective of power consumption minimization and user satisfaction maximization. We plot the reward and clashes for all four proposed algorithms to contrast the results. In this section, we first briefly explain the data set, then the environment design, and finally the experiments and results.

A. Smart* Data Set for Sustainability

The data set includes real power consumption readings of multiple devices such as furnace, fridge, washing machine, etc., inside smart homes collected over regular intervals of 30 minutes. Each device has sensors attached to them to record the power consumption after regular time intervals inside seven smart homes[1]. For our experiments, we used data from smart home A collected over a period of three years. We also use data from smart home B, and C to show transferability of the framework. Note that data from each smart home is similar in nature. The only difference is in the type and the number of devices used to collect data.

B. Designing the Environment

The data set has power consumption values from more than 20 devices for each smart home. We have considered only 5 devices from a smart home: furnace, washing machine, fridge, heater, and kitchen lights. The reasons to do so are:

- In a real world scenario, a user does not want the controller to operate on all of the devices in their smart home.
- Formalizing an optimization problem with only a top few devices with maximum power consumption is more realistic and helpful than taking all the possible devices and constraints into consideration.
- For simplicity of experimentation.

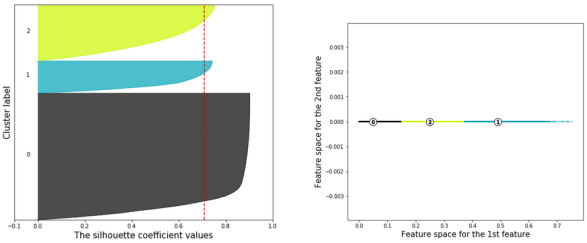
1) Data Processing: In our data set, the power consumption reading for each device took many distinct unique values. For example, the power consumption values for Furnace has $17,000$ unique entries, and that of Fridge is $16,000$. However, a lot of these values are very close and differ only at $4^{th}/5^{th}$ decimal place representing a data collection glitch. Since we chose 5 devices, a **state** in this environment is represented by a vector of size $[D_1^1, ..., D_i^k, ..., D_2^5]_{1\times5}$ where D_i^k represents the device D^k in mode i, and its value is given as the power consumption by device k in mode D_i.

The size of the state space is the cross product of all the unique values taken by each device. Therefore by this convention, if we consider only the furnace and the fridge, the size of state space will be 272 million ($16,000\times17,000$). With such a big state space, the problem becomes very complex to solve, and therefore, to avoid the state space explosion, we cluster the energy consumption values of each device separately to find a fixed **number of modes of operation** for each device. Intuitively, in real-life, a furnace cannot have 17,000 modes of operation. Therefore, finding device modes with clustering seems to be a fair assumption to make.

2) Clustering to assign the modes of operation for each device: We wanted to find cluster centers of power consumption values for each device individually, which can represent different modes of operation. The modes of operation can

[1] http://traces.cs.umass.edu/index.php/Smart/Smart

be readily available from manufacturer's end but they might not be ideal for our case. For example, suppose a washing machine consumes x power in standby mode, y in wash mode and z in dry mode, and the values, y and z are very close. The manufacturer can say that modes y and z are different, but we have similar readings for the two states, and hence, does not affect our objective. Therefore, we cluster the readings such that each device mode represents a significant amount of change in power consumption from one mode to another. The clustering helps us reduce the state space to a very good extent.



(a) Silhouette analysis for K-means clustering with k = 3 for Duct Heater

(b) Distribution of consumption values of Duct Heater with k = 3

Figure 3. Clustering results for Duct Heater from smart home A.

First we performed silhouette analysis [15] to find the optimal number of clusters for each device. We vary k, the number of clusters from 2 to 6 assuming it is rare that a device has more than 6 modes of operation. Figure 3(a) shows the silhouette plot for various clusters using Duct Heater's electricity consumption data for $k = 3$. The clusters are well formed with coefficient value more than the threshold, as can be seen in Figure 3(b). The plot is shown for $k = 3$ as it yielded the best silhouette score. Similar experiments are performed with remaining 4 devices.

After clustering, the optimal number of clusters for the chosen devices, Furnace, Washing Machine, Fridge, Duct Heater, and Kitchen Lights is 2, 3, 3, 3, and 5, respectively. Clustering reduces the size of state space from 272 million+ $(16,000 \times 17,000)$ to just 270 $(2 \times 3 \times 3 \times 3 \times 5)$ preventing the state space explosion.

VI. Experiments

The first objective is to minimize the total power consumption, and the second objective is to maximize user satisfaction. Our algorithm takes into account 4 hyper-parameters: learning rate (α), discount factor (γ), exploration rate (ϵ), reward prioritization weights W_E and W_U. As a baseline, we use Algorithm 1 with single objectives. Note that if we run the algorithm with a single objective, it becomes the traditional Q-learning algorithm. The clash rate as defined in Equation 8 will be maximum for single policy with power consumption minimization objective, and minimum with user satisfaction maximization objective. However, with multi-objectives, the cash rate should be between the two. The overall reward is given as the weighted sum, therefore, reward will be maximum for multi-objective approach. We implemented the solutions discussed in Section IV as:

A. Single Policy Single Objective

The overall reward has two weighted terms, W_E and W_U representing power consumption and user satisfaction, respectively. For the first set of experiments, we focus only on optimizing a single objective by initialising W_E and W_U as $(1, 0)$ for power consumption minimization objective, and $(0, 1)$ for user satisfaction maximization objective. Hence, in single policy single objective Q-learning formulation, our agent only receives the reward R_E in the former case and the reward R_U in the latter.

We experimented with more than 100 combinations of α and γ with $\alpha, \gamma \in (0, 1]$ to find the best hyperparameters. The agent calculates average total reward R and the $clashrate$ for every combination of our hyperparameters over a total of 463 unique states episodes learned over 300 epochs. We decayed the value of ϵ by a factor of 1.4 every 20 epochs. The set of parameters which gives us the highest average reward and least number of clashes is chosen. The hyperparameters shown in Table I achieve the best results when our aim is to minimize the average number of clashes to meet each objective individually.

Table I
Hyper-parameter values for Single Policy Single-Objective Q-learning

Objective/Hyper-parameters	**Power Consumption Minimization**	**User Satisfaction Maximization**
α	0.4	0.9
γ	0.1	0.05
ϵ	0.1	0.1
W_E	1	0
W_U	0	1

B. Single Policy Multi-Objective

We divide the experiments for multi-objective approaches into two approaches as discussed in Section IV: Equal Weights and Weighted-Sum. As shown in Line 11 of Algorithm 1, the update function for the Q-values is different than the normal Q-learning formulation. The equation for the Q-value update is given as:

$$Q_i(s,a) += \alpha(r_i + \gamma \max_{a'} Q_i(s',a') - Q_i(s,a)) \quad (10)$$

For equal weights approach, the weights W_E and W_U have been assigned the same value of 1 representing equal priority for both objectives. For Weighted-Sum approach, we perform experiments by taking approximately $2,300$ combinations of α, γ,ϵ, W_E and W_U with their values within the range (0,1]. The best hyperparameters for the multi-objective approaches are listed in the Table II.

Table II
HYPER-PARAMETER VALUES FOR SINGLE POLICY MULTI-OBJECTIVE Q-LEARNING

Approach/Hyper-parameters	**Equal Weights Approach**	**Weighted-Sum Approach**
α	0.9	0.9
γ	0.1	0.1
ϵ	0.05	0.07
W_E	1	0.3
W_U	1	0.1

VII. RESULTS

A. Average Power

To compare the four algorithms proposed to find an optimal policy, we ran them for equal number of epochs using the best hyperparameters obtained for each. Each epoch has 463 training steps and 450 validation steps, we plotted the average power for each epoch for comparison. Figure 4 shows the average reward for each algorithm.

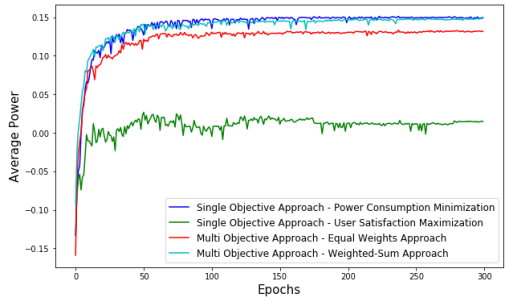


Figure 4. Average Power vs No. of Episodes for different approaches.

The average power is maximum for power satisfaction minimization because the policy is getting reward based on the predicted state and user's next state, and it is lowest for user satisfaction maximization due to the fact that if we move from a high power state to a low power state, it will hurt user's satisfaction, which is indeed the desired behavior. The plots for multi-objective approaches always end up between the two single objective ones, representing the trade-off between the two contrasting objectives.

B. Average Number of Clashes

Figure 5 shows the combined clash rate for all four algorithms. The experimental parameters are kept same as the previous section. Note that for power consumption minimization the clash rate is highest because no weightage is given to user satisfaction. If we deploy an agent with such policies, the user will get agitated and they will try to override agent's actions, rendering it useless. While on the other hand, an agent with user satisfaction maximization policies will not be helpful in optimizing power consumption. However, if we look at clash rates for multi-objective techniques, they lie between the two single objective approaches, and this clash rate can be adjusted using weights based on user preferences.

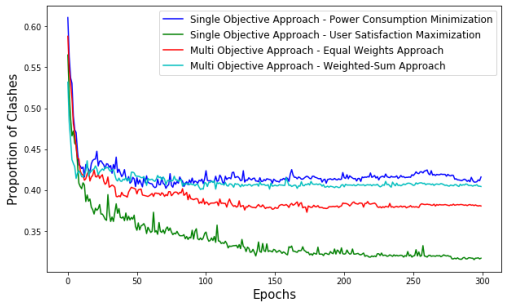
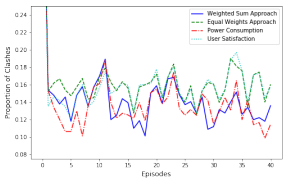


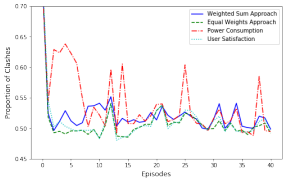
Figure 5. Average Number of Clashes vs No. of Episodes using different approaches.

C. Appliance-wise clashes

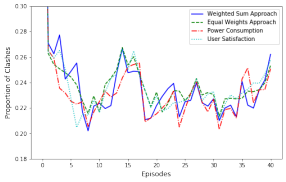
We calculate the average number of clashes for each of the five appliances and plot them separately to see the behavior of proposed approaches. Figure 6 shows the clash rate for each device. The experimental parameters are kept same as Section VII-A.



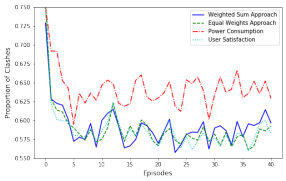
(a) Furnace



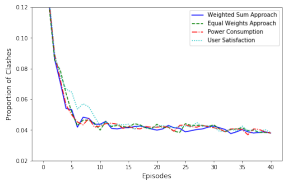
(b) Heater



(c) Lights



(d) Refrigerator



(e) Washer

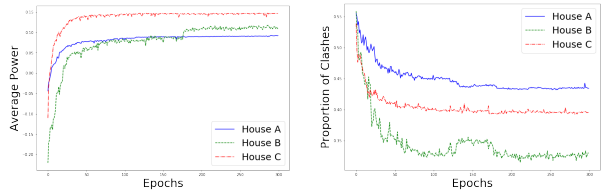
Figure 6. Appliance wise clash rate.

The results for appliances with three device modes is consistent with the overall results with two exceptions of furnace (two device modes) and kitchen lights (5 device modes).

For furnace, the power consumption minimization approach does not behave as expected. The reason could be the irregular usage and collection of data, as a furnace is used only during colder seasons and the data we used for experiments is collected over a span of three years. For lights, all algorithms fetch nearly same results. The reason can be because lights are used for prolonged times, and there are not many fluctuations in lights' modes of operation. Therefore, the clash rate coincides for user satisfaction, power consumption, and a combination of the two. Hence, Furnace and Lights have very little to contribute to the overall optimization. The results, therefore, suggest that devices that are used regularly and with several fluctuations in device modes at regular intervals should be targeted for optimization.

1) Transferability on other smart homes consumption data: To show that the proposed framework can be applied to power consumption data of multiple smart homes, we choose the best algorithm (weighted sum approach) and run it for smart homes B, and C from the same Smart* data discussed in Section V-A.

Figure 7 shows that the rewards increase and the clash rate decreases with increase in number of episodes. Figure 7(a) and Figure 7(b) show that the behavior is similar on all three smart homes data.



(a) Rewards (b) Clash Rate

Figure 7. Average reward and Clashes with Weighted Sum approach smart homes A, B, and C.

VIII. Conclusion

In this paper, we present a novel multi-objective reinforcement learning technique for power consumption optimization with contrasting objectives of minimizing power consumption and maximizing user satisfaction. We show that both objectives, when considered together, achieve the best optimal policy. Our experimental results show that the proposed multi-objective techniques establish a trade-off between the two objectives. The optimal policy achieves better user satisfaction than power optimization policies and achieves better power consumption than user maximization policies. We show that the devices used regularly in smart homes should be the ones targeted for such optimization purposes. Finally, we also show that the experiments can be performed with other smart home data set to achieve similar results.

References

[1] IEA, "World electricity final consumption by sector, 1974-2017," 2019. [Online]. Available: https://www.iea.org/data-and-statistics/charts/world-electricity-final-consumption-by-sector-1974-2017

[2] R. P. S. Center, "Energy data facts," 2019. [Online]. Available: https://rpsc.energy.gov/energy-data-facts

[3] Constellation, "Energy-saving strategies for smart homes," 2019. [Online]. Available: https://blog.constellation.com/2019/01/22/smart-home-energy-saving-strategies/

[4] Y. Tan, W. Liu, and Q. Qiu, "Adaptive power management using reinforcement learning," in *2009 IEEE/ACM International Conference on Computer-Aided Design-Digest of Technical Papers*. IEEE, 2009, pp. 461–467.

[5] H. Shen, Y. Tan, J. Lu, Q. Wu, and Q. Qiu, "Achieving autonomous power management using reinforcement learning," *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, vol. 18, no. 2, pp. 1–32, 2013.

[6] G. Tesauro, R. Das, H. Chan, J. Kephart, D. Levine, F. Rawson, and C. Lefurgy, "Managing power consumption and performance of computing systems using reinforcement learning," in *Advances in Neural Information Processing Systems*, 2008, pp. 1497–1504.

[7] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[8] I. Y. Kim and O. De Weck, "Adaptive weighted sum method for multiobjective optimization: a new method for pareto front generation," *Structural and multidisciplinary optimization*, vol. 31, no. 2, pp. 105–116, 2006.

[9] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, J. Albrecht *et al.*, "Smart*: An open data set and tools for enabling research in sustainable homes," *SustKDD, August*, vol. 111, no. 112, p. 108, 2012.

[10] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[11] C. Liu, X. Xu, and D. Hu, "Multiobjective reinforcement learning: A comprehensive overview," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 3, pp. 385–398, 2014.

[12] P. Vamplew, R. Dazeley, A. Berry, R. Issabekov, and E. Dekker, "Empirical evaluation methods for multiobjective reinforcement learning algorithms," *Machine learning*, vol. 84, no. 1-2, pp. 51–80, 2011.

[13] D. C. K. Ngai and N. H. C. Yung, "A multiple-goal reinforcement learning method for complex vehicle overtaking maneuvers," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 509–522, 2011.

[14] F. Zeng, Q. Zong, Z. Sun, and L. Dou, "Self-adaptive multi-objective optimization method design based on agent reinforcement learning for elevator group control systems," in *2010 8th World Congress on Intelligent Control and Automation*. IEEE, 2010, pp. 2577–2582.

[15] P. J. Rousseeuw, “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53 – 65, 1987.

Performance Evaluation of a Pulse Detonation Engine using Vortex Generators as obstacles in the combustion chamber

[1]Mukesh Yadav*, [2]Avhishek Jha, [3]Manish Bista, [4]BB Arora

[1]mukesh7541@gmail.com, [2]avhishekjha7@gmail.com, [3]manishbista4444@gmail.com, [4]bbarora@dce.ac.in

[1234]Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

Abstract

Propulsion based engines have been known for their exemplary thrust but also for lower fuel efficiency and speed of operation which is limited to subsonic and the level of intermittent vibrations produced. Pulse Detonation Engine (PDE) is similar to a Pulse Jet Engine (PJE) except for the fact that the flame formed inside the combustion chamber of PDE travels at supersonic speed, whereas in PJE, it is subsonic. The supersonic speed of flame front in PDE is because of the detonation wave formed during the combustion of the fuel mixture which is in the order of 2000 m/s and is based on constant volume combustion. A process called Deflagration to Detonation Transition (DDT) can be used for which turbulence has to be created inside the combustion chamber. The obstacles used in this study were vortex generators and the CAD model of the engine and the obstacles assembly was made in SOLIDWORKS 19.0 software. ANSYS (FLUENT) 16.0 was used for the simulation of the combustion inside the combustion chamber using the Species Transport Model for the performance evaluation.

Keywords: Pulse Detonation Engine, Ansys Fluent, DDT, Supersonic combustion

Nomenclature

1-D	One Dimensional
PDE	Pulse Detonation Engine
CJ	Chapman Jouget
ZND	Zeldovich-von Neumann-Doring
PDRE	Pulse Detonation Rocket Engine
SFC	Specific Fuel Consumption
T/W	Thrust-to-Weight
DDT	Deflagration to Detonation Transition
Dt	Diameter
VG	Vortex Generator

INTRODUCTION

The Pulse Detonation Engine (PDE) is a propulsion system that uses a very high rate of release of energy to produce thrust from the detonation waves [1] [24]. To get higher efficiency results than the current steady-state detonation engines, the use of simple detonation physics of combustion is done which helps in eliminating the use of high-pressure pumps or compressors leading to a major

improvement over the propulsion systems which are air-borne [6]. Any engine having the ability to be operated over a range of flight velocity such as Mach 0- Mach 5 does not exist currently [23].

There are many shortcomings of using PDE and for this reason; many studies have been performed by multiple organizations and various research facilities to overcome them [7]. The University of Michigan and US Naval Ordnance Test station carried out various experiments in the 1960s but were not able to perform a successful detonation because of the unsuitable implementation of the augmentation device of DDT leading to a halt of PDEs considering no future of such engine for in-flight applications. However, Helman in 1986 performed various experiments at the US Naval Postgraduate School which again brought the attention back to the PDEs. He focused on improving the frequency and the specific impulse on which the engine operates. The use of ethylene-air mixture was done, and from the results obtained, high operating frequencies of 150Hz with a high specification range of 1000-1400 seconds were observed.

Many single-pulse and multi-pulse detonation experiments were conducted. In the single pulse detonation, only the detonation wave initiation and their propagation with blowdown process are done [4]. Whereas, in the initiation of multi-cycle experiments, the requirement of additional purging with multiple filling processes was required. The use of single-pulse was done to find out the required mixture of fuel and oxidizer alongside to validate the various concepts used in the combustion process such as it would provide the insight of the more complex initiation of the multi-cycle process [5] [27].

The measurement of the performance of PDE is done based on the impulse generated by the process involved in the detonation initiation. However, it is observed that these methods fail as they are not accurate for the inlet conditions of the engine, and also the purging and the refilling are ignored. Either direct detonation or DDT can be used to attain detonation [28]. It was observed through various attempts that the detonation was limited to only the single pulse experiments due to which most experiments conducted in PDE use the DDT process. There is a correlation on the usage of fuel mixture, dimensions, and the roughness of the surface of the tube for a detonation to be achieved properly [24].

Sinibaldi [8] carried out an experimental study revealing that the ignitor should be used from the head end of the tube along with the equivalence ratio for the C2H4-O2-N2 mixture which greatly affects the length of DDT. A 7.5 cm of the tube length was found to be minimum for the mixture and an equivalence ratio of 1.2 was obtained. DDT length was increased significantly with the decrease of the equivalence ratio of .75 which proved that DDT length can be more than the length of the detonation tube. Hinkey[9] conducted various tests with H2-O2 mixtures which showed that various devices which could augment the DDT process with a shorter length.

To augment the DDT process, obstacles such as the Shchelkin spiral [10] [29] device was used, introduced by Kirill Ivanovich Shchelkin, a Russian physicist. The results obtained after the use of the obstacle showed that the DDT length was reduced by almost three times [30]. This led to the study and discovery of other obstacles as well. The drawback of using such obstacles was that there was a loss in the total pressure with a decrease in propulsive efficiency. Cooper [11] found out that there was a decrease of 65% length of DDT with a 25% decrease in the impulse.

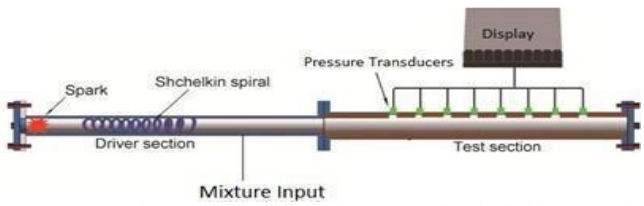


Figure 1. General PDE experimental setup with Shchelkin spiral [12]

The use of a nozzle is done at the end of the detonation tube for the improvement of the performance of the engine because it utilizes the internal energy of the products of the detonation exhausts, but it is a very challenging step as PDE is very unsteady [2] [11].

US Naval Postgraduate School carried out experiments focusing on increasing the overall efficiency from the conversion of thermal energy into kinetic energy [16] [18]. This process was conducted with a variable area of nozzle and various tests were conducted on flow rates to observe the characteristics of the flow. There was a great effect on the pressure with the injection at a variable flow rate of mass. Kailash Nath [12] later performed experiments on nozzle effects with various shapes and calculated their effects in inlet pressure along with the thrust of multi-cycle PDE. Thrust was found to be better in diverging, converging, and straight nozzles rather than a converging-diverging nozzle. From the experiments, it was also seen that the acceleration of flame and DDT along with the propagation of detonation is also affected by the structure as well as the roughness of the wall inside the detonation tube. There is a reduction of the velocity inside the detonation tube as compared to the smooth walls [12] [13]. Various experimental data suggest the effect of the composition of various mixtures of fuels on PDE performance. High detonation velocity is achieved for the heterogeneous mixture which has a high ratio of equivalence [12] [13].

The propagation of the wave is also affected by the diameter of the tube. From the computational as well as the experimental analysis and simulation, it was found that the size of the detonation cell is dependent on the initial pressure, composition of the mixture, temperature, and the diameter of the tube which is also called critical diameter [14]. Detonation waves transform from the ignition tube towards the main tube if the critical diameter is more than the tube diameter [24] [5] [11] [15]. There are many experimental studies that show the effects of the varying cross-sectional areas of the tube. From the studies, the prediction of the shock of the propagating wave and its flow characteristics were predicted.

METHODOLOGY

In this study, the computational analysis of a Pulse Detonation engine is carried out. In the initial step, the model of the engine, along with the obstacles, is made in the SOLIDWORKS 19.0 software. For the simulation of the engine, the CAD model of the setup was then imported in ANSYS 16.0 software. The setup was simulated considering a one-step chemical reaction model for the hydrogen-air mixture.

It has been observed that, if only deflagration based initiation is done, the achievement of constant propagation of flame is difficult since the self-propagation of deflagration flame tends to accelerate unceasingly thus becoming very unstable. But, with the help of suitable boundary conditions, a velocity of supersonic speed causing detonation phenomenon can be achieved from the subsonic speed of deflagration [19]. This can be done with the help of the long tube of the combustion chamber as it takes a certain time to change from deflagration to detonation phase.

For this reason, the detonation will be achieved through Deflagration to Detonation Transition (DDT) in this work.

CFD SOLVER

ANSYS FLUENT is used as the solver for the CFD where the convergence of the solution is obtained after defining the physics associated with the problem.

Since we need to solve the governing equations for energy, mass, momentum, and species transport with the help of a finite volume discretization method, a density-based solver is used as it is the most compatible for solving the required equation. To converge the solution, there need to be several iterations performed [20].

ONE DIMENSIONAL PROPAGATION OF WAVE IN CONSTANT AREA TUBE

A tube length of 2m and a diameter of .076m is selected after a thorough literature review study[17]. At a distance of 0.5 m from the tube head end, DDT initiation is placed.

A grid size of 0.1 m is used to get the best estimation of the properties such as flow inside the PDE tube as shown in Figure 2.

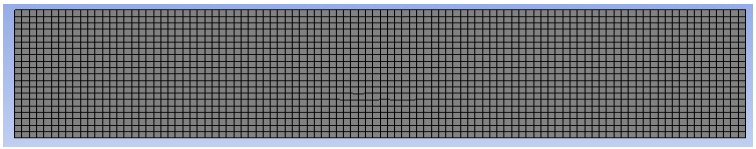


Figure 2. A two-dimensional mesh of an ideal PDE tube

Tables 1 and 2 show the condition for the unburned gas and ignition region respectively. [3] [21] [22] [25].

Table 1 Properties of the unburned gas mixture initially

Input Parameters	Values
Original Pressure	1 atm
Original Temperature	300 K
Mass Fraction of H2	2.852 %
Mass Fraction of O2	22.635 %
Mass Fraction of N2	74.512 %
Mass Fraction of H20	0.000%

Table 2 Properties for the ignition region

Input Parameters	Values
Original Pressure	30 atm
Original Temperature	2500 K
Mass Fraction of H2	0.000 %
Mass Fraction of O2	0.000 %
Mass Fraction of N2	73.42 %
Mass Fraction of H20	25.238 %

Table 3 Specifications of the PDE engine

The dimension of the PDE tube	
Inner Dia.	68mm
Length	2000mm

Wall thickness	4mm
Vortex Generator	
Cross-Sectional Area	340mm^2
Numbers of Vortex generator used	12 (evenly spaced at an angle of 60 degrees)
Blockage ratio	0.549
The length of Chambers	
Mixing Length	250mm
Ignition	250mm
Detonation	1200mm
Nozzle	300mm

Designing the Vortex Generator

Vortex generator acts as a passive flow control device by delaying flow separation point on the flow surface thereby leading to efficient lift and power characteristics. VGs help to control and alter the boundary layer over a surface. Vortex generator produces a swirl in between the high energy flow and low energy flow, filling the low energy level with high energy flow, hence delaying the flow separation.
The increment in output parameters such as lift coefficient and electrical power is dependent on the geometry of the vortex generator, angle of incidence (ψ) i.e. the angle between the wedge and the incoming flow, the angle between pairs of generators (έ), spacing between a pair of generators(s), the distance between generators (a), chordwise position(x/c) along the airfoil (or) blade and arrangement of the generator for creating vortices.

Determining the shape of Vortex Generators

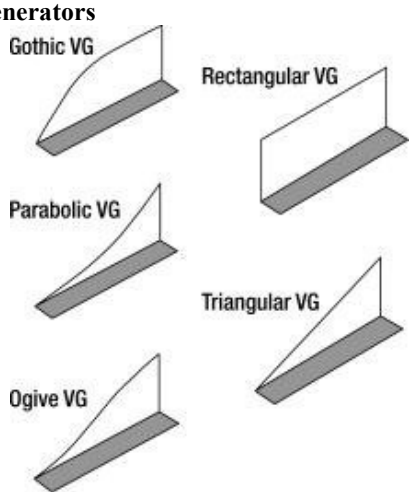


Figure 3. Various shapes of vortex generators.

From various studies conducted on VGs and their shapes, at various angles of attacks, we found out that Gothic VGs are well suited for generating an optimal amount of thrust. At 10° Angle of Attack, there are some changes in the drag value where the rectangular vortex generator gives the small reduction in drag coefficient and gothic VGs give high drag.

Determining the Reynolds Number of flow

Before designing the VGs, we needed to determine what Reynolds Number will we be operating at. For this, we had to approximate or know the speed of airflow in the PDE. Equation 1 below shows how to calculate this value.

$$Re = \frac{\rho V x}{\mu}$$

Variable Meaning	Value at Sea Level
ρ = Air density [kg/m^3]	1.205
V = Stall speed [m/s]	-
x = Chord length [m]	-
μ = dynamic viscosity [kg/m-s]	1.983*10^-5

Figure 4. The equation for the Reynolds Number Calculation

Determining the length of VGs

As per the various research works published, the length of your VGs should be around 5-8% of the chord length of the flow surface. In our case, the flow surface was the entire detonation chamber which was 1200mm in length. Hence, we decided to select an optimal value of 75mm.

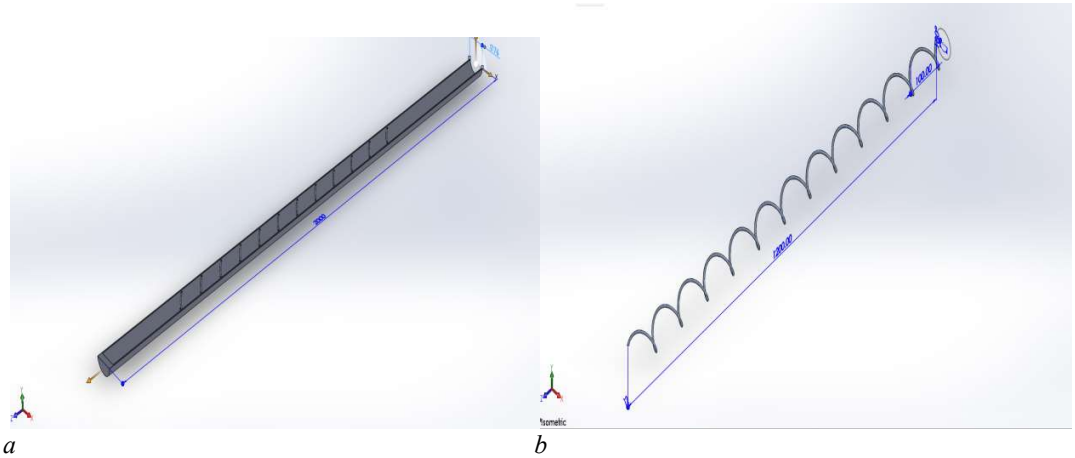
The VGs should be placed just in front of the laminar to the turbulent transition of the boundary layer on the flow surface. But, our case did not match with other conditions where the flow is over a wing or an airfoil. Hence, we decided to place VGs at an even distance to one another in a spiral fashion throughout the detonation chamber.

Determining the height of your VGs

The shape and size of PDE are similar to a narrow tube and theoretically, the height of the boundary layer in a narrow tube is equal to the thickness of the narrow tube. Hence, we decided to use the height of VGs at 68mm.
Each of these small elements creates a swirling wake that places energy in the boundary layer of the wing.

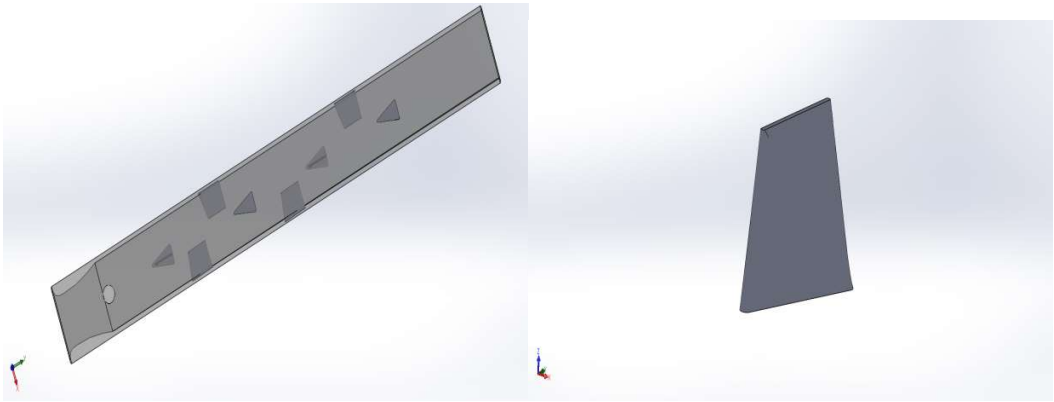
RESULTS

CAD MODELS



a *b*

Figure. 5(a) Isometric View of the combustion chamber with Shchelkin Spiral (b) Isometric View of Shchelkin Spiral



a *b*

Figure.6 (a) Isometric View of PDE with Vortex Generator; (b) Isometric view of Vortex generator

SIMULATION

ANSYS FLUENT software was used to carry out the simulation with the *k-ε* model used. The use of the Species transport model was done. The number of iterations performed was 300 to initialize the solution as shown in Figure 7.

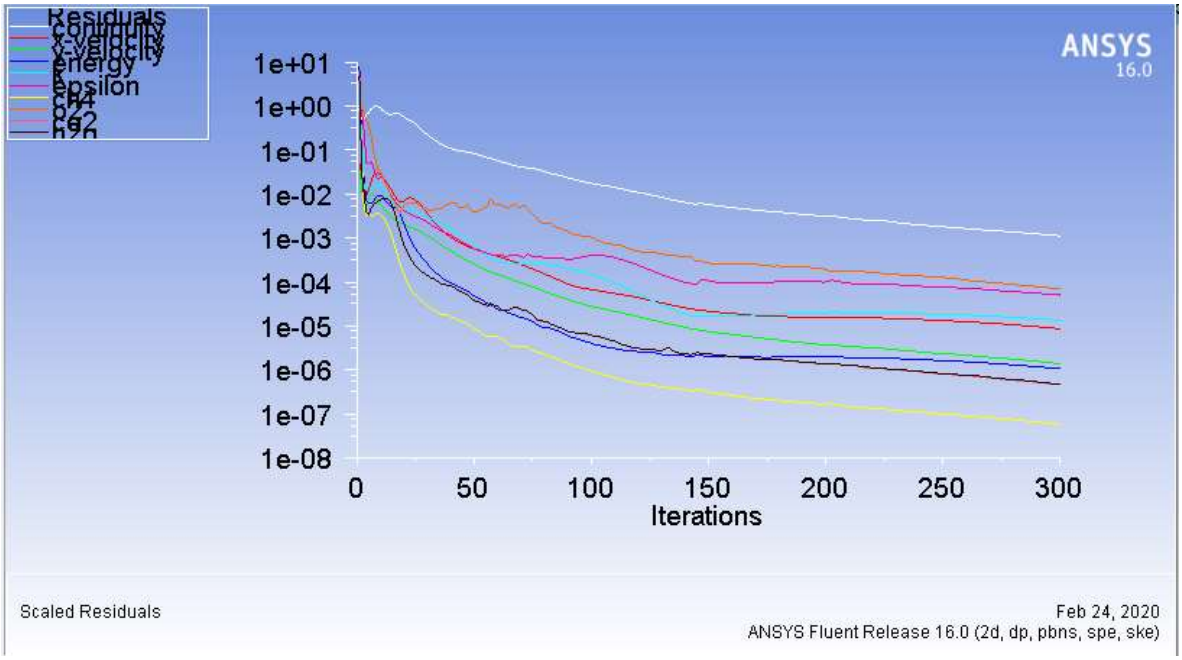


Figure 7. Total number of Iterations

PRESSURE

Figure 8. Initial conditions for the ideal PDE tube with pressure contour

It is observed that the initial detonation pressure rise occurred at .2m in the tube from the front end which shows the highest rate of the reaction being occurred at that location.

TEMPERATURE

The sharp increase in the temperature can be seen to be 2270K initially which afterward takes a constant value of 2200K similar to the pressure distribution. The temperature is observed at .350m in the tube from the head as shown in the figure below.

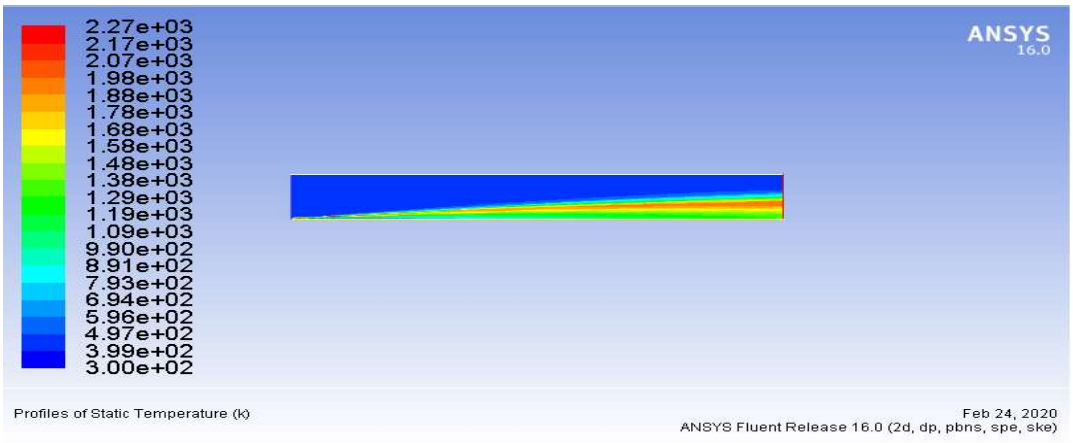


Figure 10. ZND Temperature profile

This verifies that the hydrogen-mixture can be used with just a one-step chemical reaction to emulate the ZND behavior of the model.

CJ VELOCITY

To calculate the CJ velocity, the average wave velocity is done at different locations inside the tube by using the formula of displacement to the total time elapsed. The displacement values at the peak pressure waves were taken from the distance of .25m to .6m inside the tube from the head end which was calculated to be approximately 2246 m/s and is similar to the results obtained by Vizcaino[22].

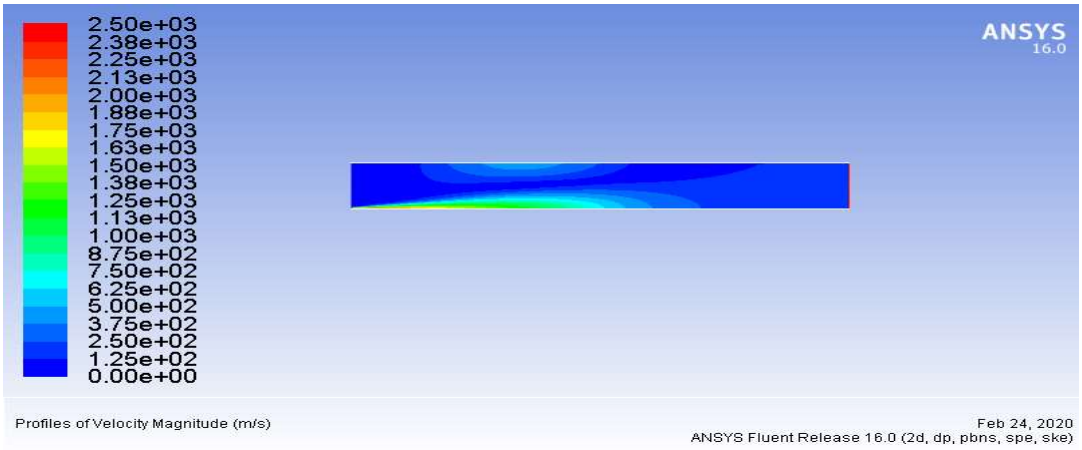
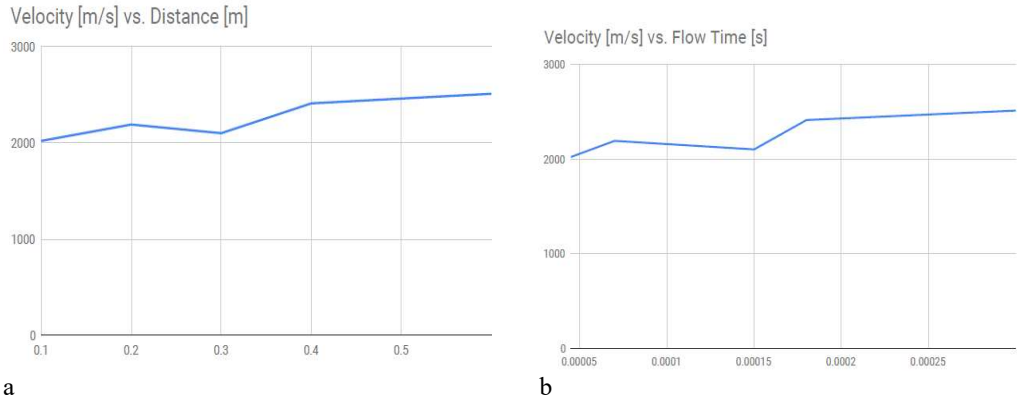


Figure 11. Wave velocity profile

Table 4 Wave Velocity Measurement

Distance [m]	Flow Time [s]	Velocity [m/s]
0.1	0.000045	2020
0.2	0.00007	2190
0.3	0.00015	2100
0.4	0.00018	2410
0.6	0.00030	2510
Average		2246 m/s



a b

Figure12(a). Change of velocity wrt time 12(b) Change of velocity wrt tube length

CONCLUSION AND FUTURE RECOMMENDATIONS

The use of a two-dimensional pulse detonation engine was done in the work to simulate one-dimensional wave propagation. A lean hydrogen-air mixture was used as the fuel species in the species transport model. The verification of C-J and the ZND model was done which showed that they could be simulated even with the use of a vortex generator as the obstacle. From the results, it was seen that

DDT happened around.6m in the chamber, proving the Vortex generator as a suitable turbulence inducer inside the chamber. The Optimum Blockage ratio of 54.9% is achieved in single pulse detonation. On introducing Blockage Ratio, there is a successful coupling of flame and shock fronts to produce a detonation wave.

For future works, it is recommended to use other designs of the chamber to study their effects in the combustion and the detonation. Also, multiple cycle detonations can be performed for accurate and real-time visualization of the combustion.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. BB Arora for providing workspace and the necessary equipments to carry out the research.

AUTHOR'S CONTRIBUTION

The contribution of the authors to this work is equivalent. All authors have read and approved the final manuscript.

FUNDING

Not applicable

AVAILABILITY OF DATA AND MATERIALS

All data generated or analyzed during this study are included in this published article with appropriated citations.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

1. Wu, Y., Lee, J. H. (April 2015). "Stability of spinning detonation waves", *Combustion and Flame*.
2. Yungster, S. (2003). "Analysis of Nozzle Effects on Pulse Detonation Engine Performance".
3. Glassman, I., Yetter, R. A. (2008). *Combustion, Fourth Edition.* Elsevier Inc.
4. Srihari, P., Mallesh, M. A. (2015). "Numerical Study of Pulse Detonation Engine with One-step Overall Reaction Model". *Defence Science Journal*
5. Ebrahimi, H. a. (2002). "A Numerical Simulation of a Pulse Detonation Engine with Hydrogen Fuels".
6. T., B., G., P. (1995). "*Pulse Detonation Engine Theory and Concepts".* The American Institute of Aeronautics and Astronautics, Inc.
7. Helman, D. S. (June 1986). "Detonation Pulse Engine". *AIAA Paper*
8. Lam, M., Tillie, D., T., L., B., M. (2004). "*Pulse Detonation Engine Technology: An Overview".* The University of British Columbia.
9. Sinibaldi, J. B. (July 2001). "Ignition Effects on Deflagration-to Detonation Transition Distance in Gaseous Mixtures". *AIAA Paper*

10. Hinkey, J. B. (July 1995). “Shock Tube Experiments for the Development of a Hydrogen-Fueled Pulse Detonation Engine”. *AIAA Paper*

11. Shchelkin, K. (1940) “*Soviet Journal of Technical Physics”.*

12. Cooper, M. a. (July 2002). “The Effect of Nozzles and Extensions on Detonation Tube Performance”. *AIAA Paper*

13. Kailasanath, K. (2009-631). “Research on pulse detonation combustion systems -a status report”. *AIAA J.*

14. Kailasanath, K. (n.d.). “*A review of research on pulse detonation engines. Report”.* Naval research laboratory.

15. Nicholls, J. A., Wilkinson, H. R., Morrison, R. B., “Intermittent Detonation as a Thrust-Producing Mechanism”. Jet Propulsion, Vol. 27, No. 5, 1957

16. Krzycki, L. (1962). “*Performance Characteristics of an Intermittent-Detonation Device”.* U.S. Naval Ordnance Test Station, China Lake, CA: NAVWEPS Report 7655.

17. Endo, T., & Fujiwara, T. (2002). “A Simplified Analysis on a Pulse Detonation Engine Model”. *Trans. JApan Soc. Aero. Space Sci. Vol 44, No.146*, 217-222.

18. Ma, F. (2003). “*Thrust Chamber Dynamics and Propulsive Performance of Airbreathing Pulse Detonation Engines”.* Pennsylvania: The Pennsylvania State University.

19. Garg, A., Dhiman, A. (October 2014). “Innovative Trends in Pulse Detonation Engine, its Challenges and Suggested Solutions”. *Journal of Basic and Applied Engineering Research, Volume 1, Number 4*.

20. Lee, J. H. (2008). “*The Detonation Phenomenon”.* McGill University: Cambridge University Press.

21. Gopalakrishnan, N. (2017). “*Numerical Simulation of Flow in Fluid Valves in Rotating Detonation Engines”.*

22. Lim, W. H. (2010-12). “*Gasdynamic Inlet Isolation in Rotating Detonation Engine”.* Monterey, CA: Naval Postgraduate School.

23. Vizcaino, J. (2013). “*Investigation of Pulse Detonation Engines; Theory, Design and Analysis”.* Daytona, Beach, FL: Embry-Riddle Aeronautical University, Scholarly Commons.

24. M. L. Coleman (2001). “Overview of pulse detonation propulsion technology” Chapter 70, April.

25. Wang Y. Q., Fan W., Yan C.J., et. al. (2009). “Performance measurement of a kerosene oxygen Pulse detonation rocket engine at frequencies of 35-40 hz.” AIAA

26. B.W. Knox, D.J. Forliti, et. al. (2011). “A Comparison of Fluidic and Physical Obstacles for Deflagration-to-Detonation Transition” AIAA.

27. Kailashnath K. (2006). “Liquid Fuelled detonations in tubes” Journal of Propulsion and Power

28. D. Valiev, V. Bychkov, V. Akkerman, C. K. Law, and L.-E. Eriksson (2010). “Flame acceleration in channels with obstacles in the deflagration-to-detonation transition,” Combustion and Flame, vol. 157, no. 5, pp. 1012–1021.

29. A. V. Gaathaug, K. Vaagsaether, and D. Bjerketvedt, “Experimental and numerical investigation of DDT in hydrogen-air behind a single obstacle”

30. T. H. New, P. K. Panicker, F. K. Lu, et.al (2006), “Experimental Investigations on DDT Enhancements by Shchelkin Spirals in a PDE”

Taylor & Francis
Taylor & Francis Group

Polymer-Plastics Technology and Materials

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lpte21

Review – polymeric materials for energy harvesting and storage applications

Sarita S Nair , Sujeet K Mishra & D. Kumar

To cite this article: Sarita S Nair , Sujeet K Mishra & D. Kumar (2020): Review – polymeric materials for energy harvesting and storage applications, Polymer-Plastics Technology and Materials, DOI: 10.1080/25740881.2020.1826519

To link to this article: https://doi.org/10.1080/25740881.2020.1826519

Published online: 02 Oct 2020.

Submit your article to this journal

Article views: 7

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lpte21

POLYMER-PLASTICS TECHNOLOGY AND MATERIALS
https://doi.org/10.1080/25740881.2020.1826519

Taylor & Francis
Taylor & Francis Group

Check for updates

Review – polymeric materials for energy harvesting and storage applications

Sarita S Nair[a], Sujeet K Mishra[b], and D. Kumar[a]

[a]Department of Applied Chemistry & Polymer Technology, Delhi Technological University, Delhi, India; [b]Ramjas College, University of Delhi, Delhi, India

ABSTRACT

Advanced polymeric materials have attracted tremendous attention in recent years as the promising materials for the development of next generation energy technology. These polymeric materials have witnessed impressive electrochemical performance improvement in energy conversion and storage devices with the advances in understanding and design of advanced-nanostructured materials such as carbon-based nanomaterials, metal-organic frameworks. This review focuses on key strategies involved in the technological usage of polymeric materials as well as their composites for the designing and fabrication of energy devices such as solar cells, supercapacitors, batteries and fuel cells along with the remaining challenges and their future perspective.

ARTICLE HISTORY
Received 30 May 2020
Revised 30 August 2020
Accepted 17 September 2020

KEYWORDS
Advanced polymeric materials; hybrids; composites; energy

1. Introduction

Nowadays, the major issue concerning humans sustainable development is the energy shortage due to the depletion of existing fossil fuels and environmental pollution. With the tremendous increase in global population in the last two decades, the demand for energy supply and storage has also increased. Against this background, therefore, one of the major concerns for scientific and industrial community is to develop clean, efficient, and low-cost energy conversion and storage technologies such as solar cells, batteries, fuel cells, electrochemical supercapacitors, and so on. The performance of an energy conversion and storage device mainly depends on the type of electrode material. Currently, the electrode material for the energy storage systems has been made and investigated with carbon species,[1] metal-based inorganic compounds[2] and conducting polymers.[3] Although, carbon-based electrodes such as activated carbon, graphene, carbon nanotubes, etc. possess high conductivity and stability with excellent cycling stability and high power density, their energy density are usually low due to the limitation of energy storage mechanism. Metal/metal compound-based electrode materials have problems of high cost, limited natural abundance, toxicity and environmental issues relating to wastes and synthesis using extraction methods that release toxic materials. However, they are still being used as electrodes in supercapacitors, batteries, and fuel cells due to their high activity and good intrinsic electrochemical properties. Conducting polymers (CPs) have high potential in energy storage devices as they show high specific capacitance and have faster kinetics than most inorganic batteries. CPs have been mainly utilized for energy storage applications rather than metal-based or inorganic compounds based on electronic devices due to the well-known advantages of polymeric materials, such as good stability, excellent processability, high flexibility, low cost, and easy device fabrication.

CPs are the organic polymers, which hold on to the physical and mechanical properties of a plastic along with the conductivity comparable to that of a metal. These materials are also called intrinsically conducting materials as the conductivity has been induced in these materials intrinsically by certain structural modification such as doping them with oxidizing (*p*-doping) or reducing (n-doping) agents forming analogous polymeric radical cation or anion, respectively. Uniqueness in the structure of these conducting polymers which bestow them their electrical conductivity is the alternating single and double bonds acknowledged as the conjugated π-electron system, stretching all along the length of the polymeric chain with the intention that the electron density does not reside on a particular atom but it is delocalized over the entire chain.[1–4] During the last 20 years, there has been spectacular progress in the synthesis of hybrid materials based on conducting polymers driven by the development of novel synthetic procedures and by the large variety of applications. The combinations of conducting polymers with carbon materials and metal compounds have been utilized to

CONTACT D. Kumar dkumar@dce.ac.in Department of Applied Chemistry & Polymer Technology, Delhi Technological University, Delhi 110042, India

© 2020 Taylor & Francis

form flexible, lightweight, high-capacitive, and low-toxic devices with excellent performance. These advanced polymeric materials undoubtedly constitute one of the most promising classes of new materials and are attractive from technological perspectives as well, for example, in energy conversion and storage, electronics, catalysis, and optics. It has been very recent that comprehensive studies have been intended at realizing wearable and stretchable organic electronics.[5] Advanced polymeric materials particularly conjugated polymers allow tuning of their mechanical property chemically facilitating structural deformation without affecting their electronic functioning considerably. Some excellent reviews and books in the literature can give detailed perspective of recent advances in molecular designing and engineering of advanced polymeric materials incorporating new intrinsic mechanical and physical characteristic like fiber shaping to be woven in to textile for truly integrated solution.[6,7]

Based on the fundamentals of polymer science, this review will explore the use of advanced polymers in electronic and optoelectronic devices with a view to linking the physical properties of these polymers to the functionality of the technological applications.

2. Application in energy conversion and storage

2.1. Solar energy harvesting (solar and photoelectrochemical cells)

Solar cells are the devices accomplished by converting radiative solar energy into more useful electrical energy. Conventional solar cells were built from inorganic silicon-based materials wrapping almost 86% of solar cell market worldwide. Though the efficiency of such conventional solar cell has attained superlative mark of 20–25%; however, cost factor allied with such a device limits its utility and prospective to be economically feasible and as concerned photovoltaic technology is too complex, hazardous, and rigid. Finding alternative earth-abundant solar cell materials beyond Si has been at the leading edge of research for many years. The high absorption coefficient of conjugated polymers along with robust deposition on flexible substrates by simple inexpensive practices such as electrochemical deposition, printing, roll to roll coating makes them possible candidates to be used in a combination with n-type inorganic semiconductors (typically TiO_2, ZnO, and CdSe) to form simple and inexpensive hybrid organic–inorganic solar cells.[8–10] The *p*-type semiconductor hybrid polymer acts both as an electron donor and as a hole transporter. These potential substitutes of the conventional silicon photovoltaics have theoretical achievable efficiencies predicted to be as high as 12%,[9] but until now the performances obtained for real cells remain well below these expectations (1–4%). Although, the performance of organic solar cell has not yet reached a face of commercialization however, encouraging advancement made over last few years have potentially increased the possibility to go far ahead commencing the contemporary state-of-art.

2.1.1. Normal cell

Standard device fabricated on glass/PET substrate consists of transparent electrode of indium tin oxide as anode, spin-coated layers of polymer/fullerene mixture (this absorber layer is accountable for the photovoltaic activity of the cell) and evaporated metal electrode as cathode in a flat multilayer geometry. From material point of view in the current scenario donor polymer is the main absorber of the solar radiation. However, new focus is also shifting to the light-absorbing acceptors too. Recently, a new class of non-fullerene-based electron acceptors having sufficient absorption property within the solar-terrestrial spectrum has been designed. Non-fullerene based OPVs with *p*-type conjugated polymers as the donor and n-type small molecules as the acceptor have revealed stable progress over the time reaching efficiencies of over 15% in 2019.[11] Organic acceptor based on di-cyan substituted quinacridone derivatives (DCN-nCQA) i.e., DCN-8CQA with appropriate LUMO level of −4.1 and band gap of 1.8 eV demonstrated a power conversion efficiency (PCE) of 1.57% and also remarkable absorption in the region around 650–700 nm.[12] Other novel acceptor that contributes to light absorption includes the recently investigated perylene diimide derivatives (PDI).[13] Mechanistic pathway concerning solar energy harvesting by now has been extensively detailed in numerous literature accounts, which include the photon absorption by the electron donor leading to an excited electronic state i.e. bound electron-hole pair (exciton), followed by diffusion of this exciton within its existence to the interface between electron donor and acceptor for dissociation to free charge carriers and subsequent charge transport within the organic semiconductor to their respective electrodes and thus creating photovoltaic effect. Organic photovoltaic's (OPV) is currently represented by the bulk heterojunction (BHJ) solar cells based on poly (3-hexyl thiophene) (P3HT) donor polymer and fullerene (1-(3-methoxycarbonyl) propyl-1-phenyl [6, 6] C61) (PCBM) acceptor. To fabricate viable devices using polymer fullerene composite blend two architectures have been much exploited.

First is the donor-acceptor bilayer planar device whose prospects are narrowed down by the short diffusion length of the exciton (which is few tenth of nanometers), which is also a major reason limiting the necessary film thickness for optical absorption in such devices. The other is the single-step solution-processed intimately mixed composite active layer (BHJ). Later being the most important keystone recognized in the trail of polymer solar cell where nanoscale phase separation between donor/acceptor interface aid in the exciton dissociation and efficient charge transport through the bulk photoactive layer.[14] BHJ architecture consists of homogenously dispersed donor and acceptor interfaces throughout the matrix of active layer so as to effectively split exciton to free charge carriers which is thereby transported to respective electrodes through the bicontinuous-percolated networks. The headstream driving charge transfer and transport in the organic photovoltaic is the energetic offset between the donor and acceptor levels which produce what is called an integral potential or open circuit voltage (V_{oc}). However, a part of this energy is also exploited in overcoming columbic exciton binding energy (which is limiting the dissociation of photogenerated exciton). V_{oc}, one of the factors determining solar cell efficiency is related to energy difference between the highest occupied molecular orbital (HOMO) of the donor polymer and the lowest unoccupied molecular orbital (LUMO) of fullerene acceptor in case of ohmic contacts. In order to promote increase in efficiency of the OSCs new materials other than P3HT & PCBM are to be explored. A setback with the P3HT is its higher band gap around 1.9 eV restricting its absorbance only up to 650 nm of solar radiation and thus harvesting only 22.4% of solar photons accessible.[15] Improvement of all of the above-mentioned complications of light absorption profile, charge mobility, optical band gap, ionization potential (HOMO level) and electron affinity (LUMO level), absorber layer morphology pointed out from the studies on existing model and materials seeks creation and conception of next generation high efficiency organic solar cells.

A scope to achieve larger amount of solar absorption is by switching to lower band gap materials which can offer greater room for improving the efficiency of OSCs.[16] A synthetic approach to reduce the band gap of these materials is by preparing competent copolymer having alternatively fused heterocycles of electron-rich (donor) and deficient (acceptor) entity incorporated in to the main polymeric backbone, this repeating unit structure maximize performance. Common donor groups are dithiophene, fluorene, cyclopentadithiophene whereas acceptor molecules include diketo-pyrrolo-pyrrole (DPP), benzothiadiazole to name a few. Using this method, a series of low band gap (LBG) donar polymers for bulk heterojunction cells have been prepared. A copolymer of definite interest is based on cyclopentadithiophene family, namely, poly[2,6-(4,4-bis-(2-ethylhexyl)-4 H-cyclopenta [2,1-b;3,4-b0]-dithiophene)-alt-4,7-(2,1,3 benzothiadiazole)] (PCPDTBT) with a LBG of 1.45 eV .[17] Heeger's group attempted to increase the efficiency & alter the morphology of PCPDTBT thin film with the help of dithiol additive, PCE increased from 2.8 to 5.5%.[18] Recently, Scharber et al. presented a qualitative description of the morphological effect of processing additive on PCPDTBT. They proposed that use of octane dithiol strongly suppresses the charge transfer exciton emission (loss channel in conventional solar cell) by fine tuning the nano morphology of photoactive layer.[19]

More recently, Rogers et al. pointed out that additives not only act to change the general spatial distribution of donor and acceptor component in the BHJ active layer as was previously stated but it is also influential for controlling the internal order within the individual phases and also qualitatively assessed direct correlation between the device performance and polymorph, correlation length, and orientation of PCPDTBT crystallites.-[20] The only drawback of PCPDTBT is the rather high-lying HOMO level (–5.2 eV) that does not allow V_{oc} higher than 600–700 mV when blended with PCBM. Therefore, narrowing the band gap of the polymer without compromising on the efficient charge separation as well as the high V_{oc} is a major barrier while heading for improving efficiency of OSCs. In order to maintain high V_{oc}, tuning of band edges via vigilant structural & molecular design for attaining deep-buried HOMO energy level is imperative. For instance, applying stronger electron-withdrawing groups to the backbone of polymers has been found to be effective in lowering the HOMO of polymers, which directly affects the V_{oc} of PSCs. A new LBG semiconducting polymer with a synergetic combination of LBG and low-lying HOMO level based on benzodithiophene and thieno thiophene co-polymer derivative i.e poly[4,8-bis-substituted-benzo [1,2-b:4,5-b0]dithiophene-2,6-diyl-alt-4-substituted-thieno[3,4-b]thiophene-2,6-diyl]PBDTTT (afterward referred as PBDTTT-E) was synthesized and found to be a promising material for solar energy harvesting. Blending this polymer with methano-fullerene yielded a PSC with a PCE of 5.6%, open circuit voltage (V_{oc}) of 0.6 V & fill factor (FF) of 65% that is to say among highest values reported for PSC based on LBG polymer.-[21] Further using the same polymer system, Chen et al. attempted to alter the HOMO of the derived polymer by adding different electron-withdrawing functional groups step by step. They substituted the alkyloxy group on the carbonyl of thieno [3, 4-b] thiophene

unit with an alkyl group (afterward referred to as PBDTTT-C). Remarkably, alkyl group not only resulted in lower HOMO level and also a lower LUMO level, both LUMO & HOMO level were lowered by 0.1 eV in PBDTTT-C. They have also shown that addition of more than one electron-withdrawing group is effective in further lowering of HOMO level in PBDTTT. So, they modified PBDTTT-C with fluorine atom (afterward referred to as PBDTTT-CF). On moving from PBDTTT to PBDTTT-C further to PBDTTT-CF a V_{oc} as high as 0.76 was observed which together combined with high short circuit current (J_{sc}) and FF, a PCE of 6.58% for PBDTTT-C & 7.73% for PBDTTT-CF were achieved.[22,23]

An extensive structural optimization of the polymer belonging to this class of family lead to the development of new novel polymer poly({4,8-bis [(2-ethylhexyl) oxy] benzo [1,2-*b*:4,5-b′] dithiophene-2,6-diyl}{3-fluoro-2-[(2-ethylhexyl) carbonyl] thieno[3,4-*b*]thiophenediyl}) (PTB7) (Figure 1a), further promoted via mixed solvent approach (dichlorobenzene/1,8 diiodoctane), first polymer solar cell showing PCE above 7% has been realized. The photovoltaic (PV) parameters of the aforementioned device and effect of mixed solvents on these parameters are summarized in Table1. PTB7 polymer blended with PCBM shows strong absorption covering range from 300 to 800 nm as shown in Figure 1b.[24]

Figure 2 shows the structure of some donor conjugated polymers discussed in this section. Grafting 3-butyloctyl side chains to benzo[1,2-b:4,5-b 0] dithiophene unit at C4 and C8 positions, resulted in a polymer P1 with high hole mobility (10^{-2} cm^2 Vs^{-1}) and low-lying HOMO energy level (5.22 eV). P1 blended with $PC_{61}BM$, demonstrated a V_{oc} of 0.78 V and a high PCE of 4.8%. P2 with electron-withdrawing fluorine inserted in to thieno [3, 4-b] thiophene (Figure 2d) exhibited a significantly improved V_{oc} of 0.86 V, as compared to that of P1. So, replacement of the alkoxy chains on the benzo[1,2- b:4,5-b 0] dithiophene unit with less electron-donating alkyl chains allowed to increase the V_{oc} of these polymers by 0.1 V but without increasing their band gap energy which remains around 1.64 eV.[25]

Recently, Sonar et al. have reported a novel versatile donor-acceptor-based solution processable low band gap (1.5 eV)-based polymer semiconductor (poly{3,6-dithiophene-2-yl-2,5-di(2-octyldodecyl)-pyrrolo[3,4-c]pyrrole-1,4-dione-alt-naphthalene}) (PDPP-TNT) (Figure 2e) synthesized by Suzuki coupling using condensed diketo-pyrrolo-pyrrole (DPP) as acceptor moiety with fused naphthalene donor building block in the polymer backbone. OPV device using PDPP-TNT as donor with an appropriate HOMO energy level (5.29 eV) and $PC_{71}BM$ as acceptor have shown a J_{SC} of 11.8 mA cm^{-2}, V_{OC} of 0.76 V, FF of 0.52, resulting in a PCE of 4.7%.[26] For in-depth acquaintance, a review focusing on the conjugated polymer structure–property relationship for OPVs can be referred.[27] Recently, efforts have reappraised the need for the bilayer architecture for solar device by utilization of a pair of two orthogonal solvents ortho-dichloro benzene and dichloromethane to form sequential layers of P3HT and PCBM, respectively, with the PV performance challenging today's BHJ devices. Dichloromethane (DCM) solvent-casted PCBM overlayer did no alteration to the morphology of P3HT underlayer revealed of the AFM images (Figure 3). Highest FF in bi-layer device was achieved by balancing the hole and electron transit time with 4 times thicker P3HT under layer than the PCBM over layer.[28]

Parallel with the material development various fabrication techniques and device structures had also been explored. During the chase behind strategies to increase the PCE of OSCs, study of microscopic morphology of the active layer also sounds mandatory. The active layer in case of organic semiconductors, unlike silicon semiconductors have to deal with the recombination of free charge carriers even before its extraction from the cell and hence, a photocurrent loss mechanism. So, sound knowledge of morphology is inevitable when aiming

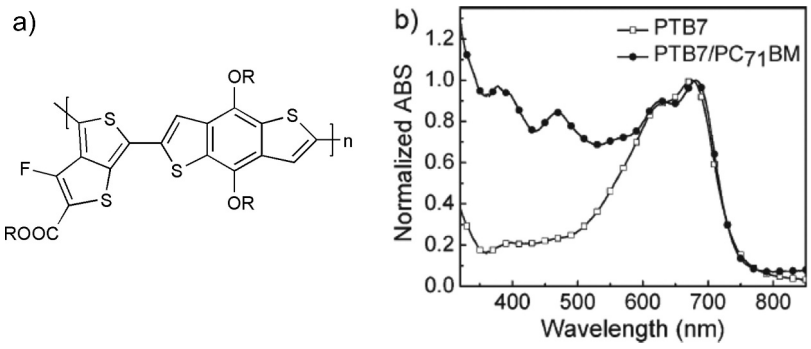


Figure 1. a) Structures of PTB7 and PC71BM. b) Absorption spectra of PTB7 pure polymer film and blend film PTB7/PC71BM in 1:1.5 weight ratio.[24] Copyright 2006, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

Table 1. Device photovoltaic parameters of (i) dcb only, (ii) dcb with 3% dio, (iii) cb, and (iv) cb with 3% dio as solvent(s).

SOLVENTS	V_{OC} (V)	J_{SC} (MA CM^{-2})	FF (%)	PCE (%)
DCB	0.74	13.95	60.25	6.22
DCB+DIO	0.74	14.09	68.85	7.18
CB	0.76	10.20	50.52	3.92
CB+DIO	0.74	14.50	68.97	7.40

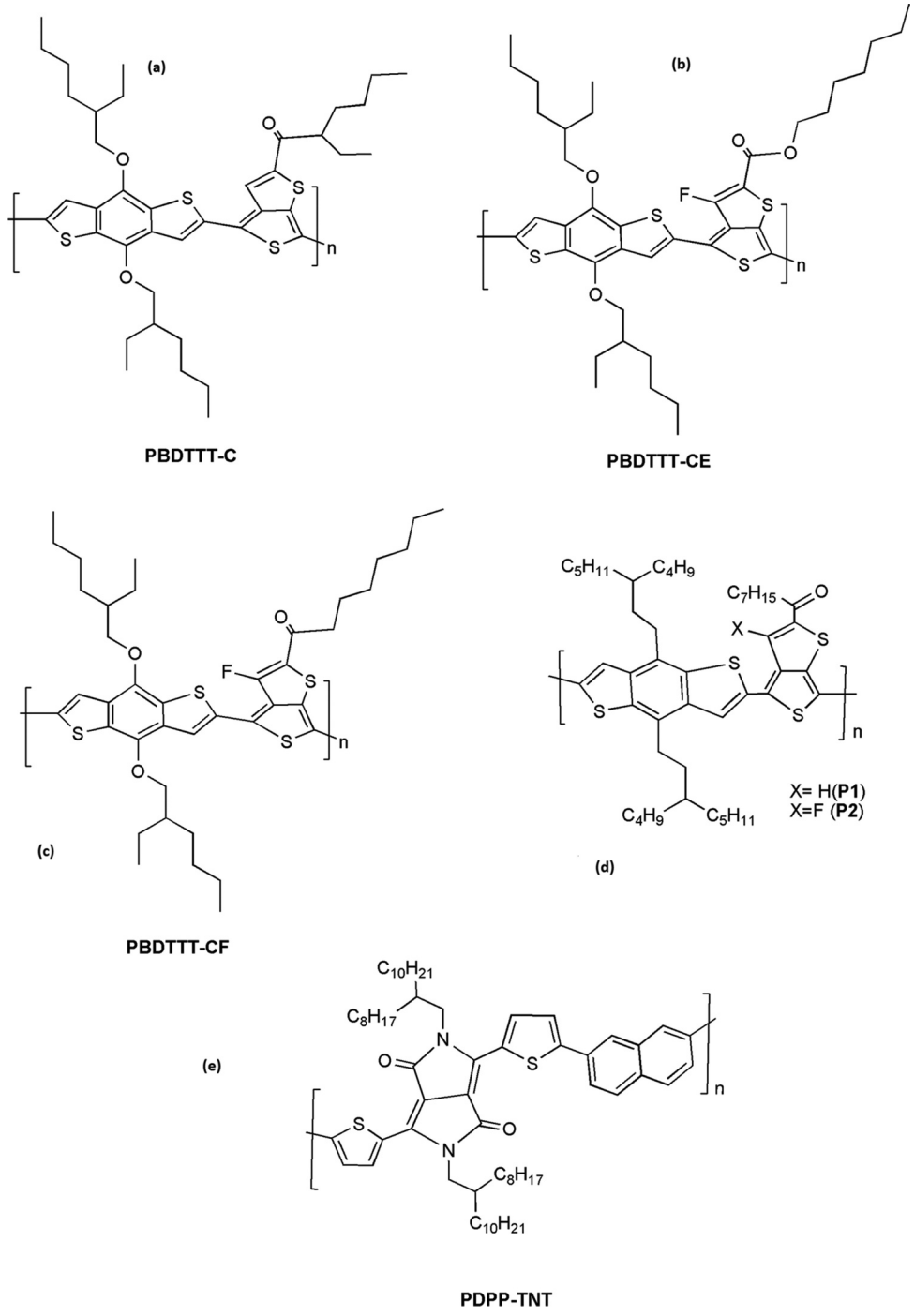


Figure 2. Structures of some donor conjugated polymers.

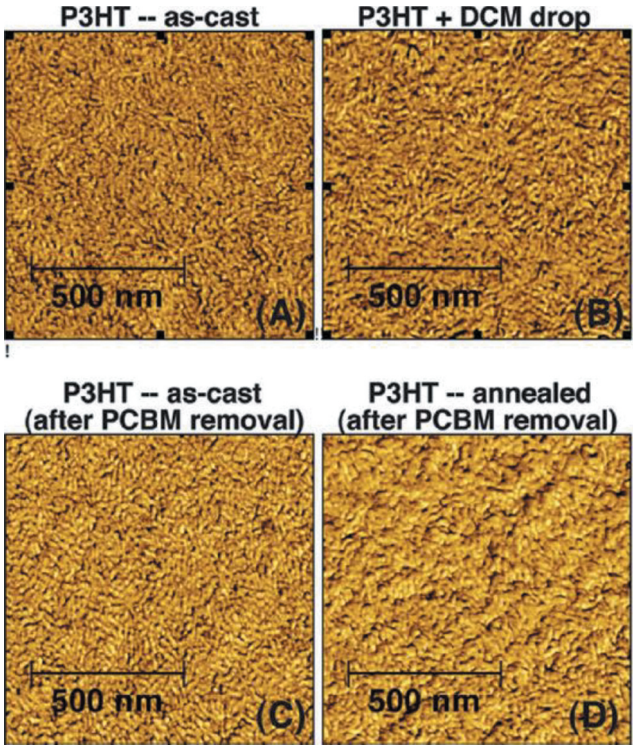


Figure 3. AFM tapping-mode phase images (1 μm × 1 μm) of (a) an-cast P3HT film. The nanocrystalline domains have an average diameter of~ 14 nm, and the rms surface roughness is 1.55 nm. (b) an as-cast P3HT film onto which a drop of DCM solvent has been spun. The size of the nanocrystalline domains and surface roughness are identical within the error to the as-cast film shown in panel A. (c) an as-cast P3HT film onto which a PCBM layer had been spun from DCM and then subsequentlyremoved by soaking the bilayer in CH. The size of the P3HT nanocrystalline domains is identical within the noise to the films shown in panels A and B. (d) a P3HT film onto which a PCBM layer had been spun from DCM with the bilayer annealed at 150°C for 20 min following removal of the PCBM overlayer by soaking in CH.[28] Copyright 2009, American chemical society.

a highly efficient organic photovoltaic. Solution processing fabrication technique allows sufficient freedom to control phase separation and molecular self-organization during solvent evaporation and/or film treatment.[29] Composition dependent morphology study by simultaneous grazing incidence small-angle X-ray scattering (GISAXS) and wide-angle X-ray diffraction (GIWAXD) techniques explained the optimum PCBM concentration of 40–50% for active layer thickness of the order of 90–140 nm, in favor of coexisting P3HT chain crystallization and PCBM aggregation. Figure 4 shows the GISAXS 2-D profiles of P3HT/PCBM films with different compositions of PCBM in the blend.[30]

A more efficient tool to manipulate the morphology of the active layer is by using controlled post annealing or thermal treatment. A twofold purpose, starting with a forced reorganization of film while one or all components of the hetrojunction can crystallize and long-standing stable reproducible morphology of active layer is been offered.[31] It has been shown by Kim and coworkers that the crystallinity of P3HT can be increased by thermal annealing, via formation of crystallites with the conjugated chain parallel to the substrate (a-axis orientation). The improved crystallinity in turn increases the photon absorption, hole mobility, and hindered the shunt pathways within hetrojunction. On comparing annealing sequences, post annealing after the aluminum (Al) electrode deposition resulted in substantial increase of interfacial contact area and rougher interface morphology and improvement of device performance.[32] However, with thermal annealing nanoscale morphology can also deteriorate as the domain size of donor or acceptor phase exceeds the exciton diffusion length forbidding efficient charge transfer. Higher temperature thermal annealing may perhaps degrade or oxidize the active layer. Solvent annealing also encouraged the polymer self-organization for improving the nanoscale morphology by managing the rate at which residual solvent evaporates. Park et al. studied the effect of annealing solvents solubility on the device performance. Interestingly, they observed greater increase in Jsc, FF, and PCE for poor-solvent-annealed devices than for the good solvent annealed one which is attributed to the optimized phase separation in the active layer and hence a balance carrier mobility and minimum exciton loss.-[33] The facile method of solvent or solvent mixture was exploited by several researchers to control the morphology.[34] Though the solvent used and system studied were different however, the foremost aim remained same i.e., to significantly recover the crystallinity of components in the mixture which otherwise have been suppressed in the presence of the other component while maintaining large interfacial area.

Previously, efforts were diverted to directly increase the crystallinity of a particular component in the mixture such as the new procedure by Berson et al. to produce P3HT nanofibers in *p*-xylene and blending these concentrated solution of nanofibers with PCBM as a mixture for active layer in PSC exhibited a PCE of 3.6%.[35] Kim et al. tried to optimize the solar cell based on P3HT nanowire via solution crystallization in DCM solvent. A systematic study of the aged blend solution signified an increase of both light absorption and charge balance achieving PCE of 3.23% with 60 h aged solution-processed PSC which is further promoted to 3.94% by insertion of a pure donor phase between PEDOT: PSS and active layer.[36] Somewhat related approach to control the nanomorphology of P3HT/PCBM blends

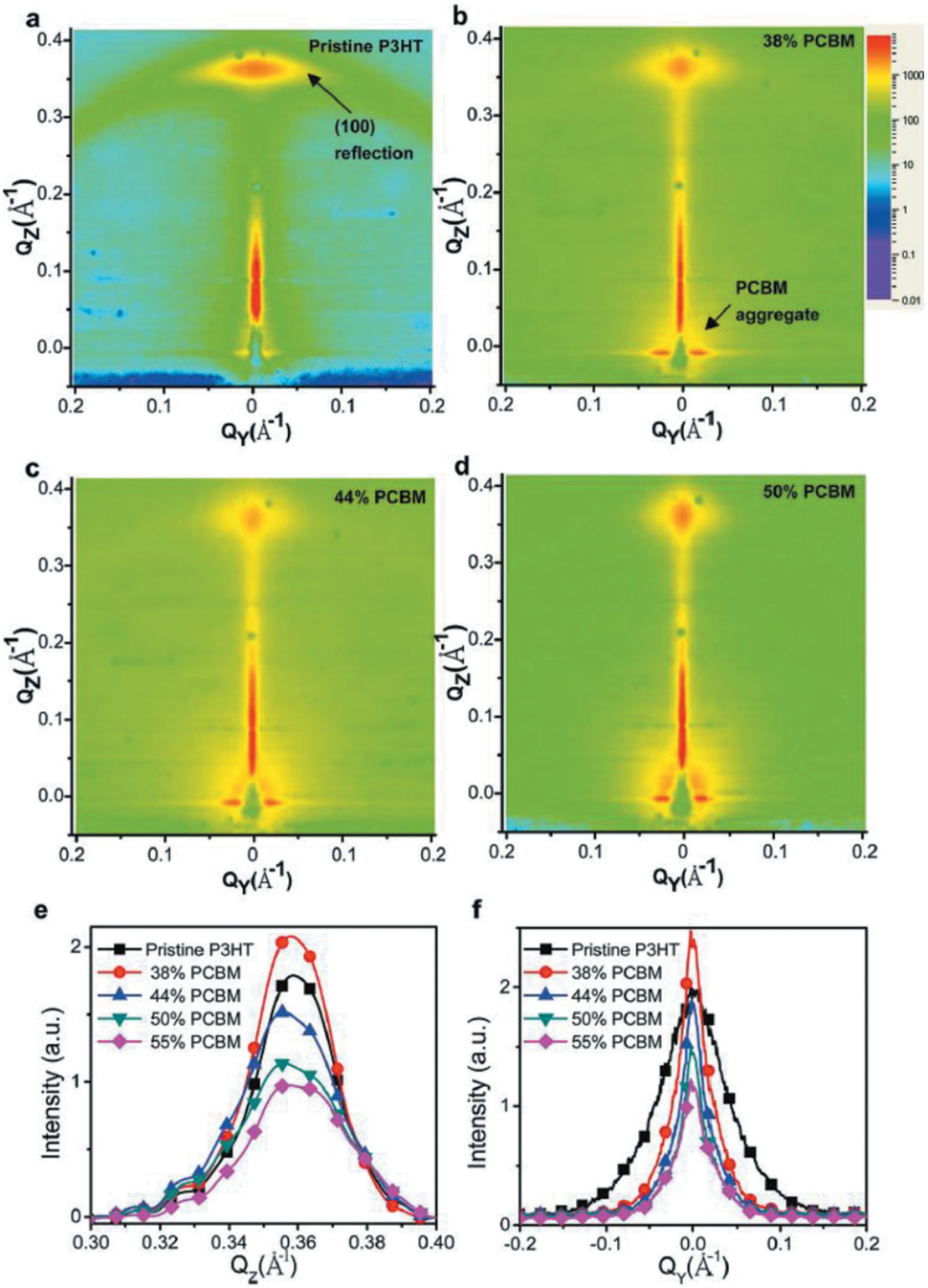


Figure 4. GISAXS patterns of the P3HT film (a), and the P3HT/PCBMfilms with 38%(b), 44%(c), and 50%(d) of PCBM.GISAXS profiles taken along QY = 0 (e) and QZ = 0.36A ° −1 (f) for the P3HT 100 reflection.[30] Copyright 2010, American Chemical Society.

involves the addition of non-solvents into the P3HT/PCBM solution for instance addition of marginal solvent acetonitrile to chlorobenzene accelerated the aggregation & self-assembly of P3HT nanowire with PCBM which is not disassembled by additional ultrasonication at ambient temperature achieving PCE of 4.09%.[37] Again, Kim and coworkers investigated in-situ formation of self-organized interpenetrating P3HT nanowire morphology in bulk hetrojunctions through the application of dual process of marginal solvent treatment (such as cyclohexanone) and subsequent mild thermal annealing, which leads to a substantial increase of electron and hole mobility to 1.6×10^{-3} cm^2 $V^{-1}s^{-1}$ and 1.4×10^{-3} cm^2 $V^{-1}s^{-1}$ respectively.[38]

2.1.2. Inverted cell

Fundamentally, indium tin oxide (ITO) has the ability to collect both holes as well as electrons as the work function of ITO (4.5–4.7) lies in between those of HOMO & LUMO values of most common organic photovoltaic materials. So,

polarity of ITO can be switched depending on its contact functional material. For hole extraction a high work function material layer such as PEDOT: PSS will work forming ohmic contact with *p*-type donor polymer. For electron extraction, a contact with a low work function metal will do such as spin coating with a cesium carbonate (Cs_2CO_3) layer by Huang et al. .[39] In inverted solar cell, ITO will function as cathode unlike as an anode in normal cell architecture. While, the high work function metal deposited on the top after the active layer will behave as anode. This architecture possibly overcomes the interface instability associated with normal architecture which is due to the degrading effect of acidic PSS doped PEDOT on ITO and thus, indium would diffuse in to the polymer layer. This architecture can also keep away the device from degradation which would take place in normal architecture due to greater sensitivity of low work function metal cathode on top to the oxygen and moisture present in air .[40,41]

It is apparent that functioning of a polymer BHJ in an "Inverted mode" where, electrons are extracted from the transparent electrode and holes from the reflective electrode generally necessitates tuning of the work function via interfacial modifiers .[42] Rider et al. reported, that hybrid film of cationic polythiophene [poly [3-(6-pyridiniumylhexyl) thiophene bromide] ($P3PHT^+Br^-$)] and anionic $(PEDOT: PSS)^-$ was found to modify the work function of ITO to a level appropriate for application as a cathode in inverted P3HT:PCBM solar cell. Inverted PSC with a layer-by-layer (LBL) assembled (P3HT/PEDOT: PSS) interface revealed device performance analogous to other interfacial materials whilst, displays long-term stability in excess of 500 h .[43] In an inverted architecture modification of ITO with electron injection layer is evident to serve as cathode. Functional layers of metal oxides such as TiO_2, ZnO etc. are more common and some alkali metal compound such as $CsCO_3$ can also serve the purpose. However, because of its easy deliquesce, it may not work in the long run. Recently, solution processible aluminum-doped zinc oxide (AZO) has been used as electron injection layer in inverted organic photovoltaics .[44] Steim et al. incorporated polyoxyethylene tridecyl ether interfacial layer between ITO/TiO_x layer & raised PCE of inverted organic photovoltaic to 3.6% .[45] Schmidt et al. studied the transient characteristics of inverted solar cell with titania layers upon illumination with different wavelengths to investigate on the significantly improved device characteristics upon illumination with UV light. Even though the short circuit current increased slightly, however, the most striking increase in FF is observed with UV exposure. They evidenced this behavior as a consequence of reversible oxygen adsorption and desorption. UV exposure facilitates desorption of oxygen leading to improved device characteristics .[46] Nanostructures of ZnO, TiO_2 offers additional hetrojunctions at the interfaces of donor and nanostructure promoting exciton dissociation. Mor et al. illustrated a vertically oriented TiO_2 nanotube film. P3HT & PCBM is infiltrated in to this nanotube in an inverted solar cell achieving PCE of 4.1% .[47] More recently, a substantial increase in device performance and operational stability in solution-processed inverted organic photovoltaic is demonstrated by introducing electrodeposited ZnO interlayer. Optimized device structure with a WO_3 interlayer on the opposite electrode achieved a remarkable maximum PCE of 4.91% as well as good operational stability in nitrogen and air .[48] A novel hole collecting layer of sulfonated poly (diphenyl amine) (SPDPA) with environmental stability was employed along with titania TiO_2 as an electron-collecting layer to construct an inverted solar cell with increased air stability. PCE was 3.91% for this device and was ascertained to maintain a PCE of 2.82% after 400 h of air storage. Ultra violet photoelectron spectroscopy studies indicated that SPDPA improves the barrier height, especially at the interface of metal deposited on organic film .[49] A straightforward technique for unstructured fabrication of nanoripples on ZnO thin films was developed and these nanostructured ZnO films have been employed as a hole-blocking layer in inverted solar cell. PCE of inverted OSCs with ~70 nm high nanoripples (3.2%) was higher than the relatively flat ZnO surface by a factor of 2. This may be because, existence of nanoripples on ZnO results in a higher contact area between ZnO and active layer .[50] In the recent past, Wang et al. fabricated an inverted organic solar cell devoid of any electron selective layer. A post-processing light soaking for 2 h with an AM 1.5 G solar simulator found to improve the PCE of the as formed device from 1.46% to 4.10% irreversibly which was tentatively attributed to the change in the surface composition of the blend layer .[51] Recently, Kim et al. have uncovered a peculiar dependence of V_{oc} of an inverted solar cell on the annealing sequence. They pointed that thermal annealing in inverted cell above 100°C after top metal electrode deposition leads to a lowering in V_{oc} as compared to analogous unannealed device or devices annealed before metal electrode deposition. However, they pointed that this lowering can also be reversed when the top electrode is replaced. They manifested that this variation in V_{oc} with the annealing sequence is the result of change in interface between photoactive layer and top electrode and not a reflection of structural changes in the bulk active layer .[52]

2.1.3. Tandem cell

Undoubtedly, a renowned breakthrough while heading toward an efficient organic solar cell ahead of the

Schockley-Queisser limitation was the stacking of several sub cells in series identified as "Tandem cell." A first demonstration of such a polymer tandem cell using two identical bulk hetrojunction sub cells was given by Kwano et al., where the two sub cells were separated using the sputter deposited ITO .[53] Following their work, Hadipour et al. for the first time fabricated tandem cell using different BHJ sub cells combining low band gap and larger band gap polymer having complementary absorption range. The optical cavity in the bottom cell was optimized so that, overlap of transmission from bottom cell and absorption of low band gap top cell is maximum .[54] It is worthwhile to cite that Kim et al. demonstrated a high-efficiency solution-processed organic tandem cell with 38% performance improvement versus the best single junction device. They employed PCPDTBT: PC70BM hetrojunction for bottom cell and a P3HT:PCBM blend for top cell complementing each other perfectly and thereby achieving phenomenal efficiency of 6.5% .[55] A more comprehensive report on organic tandem cell can be found elsewhere .[56,57] We shall be focusing on the latest advancement in this dynamic field. An efficient organic photovoltaic tandem cells consisting of P3HT:bis[60] PCBM and P3HT:[70]PCBM active layer and an original transparent conductive oxide (TCO) interlayer sandwiched between LiF & MoO_3 layers serving a recombination layer is presented. These tandem cells demonstrated J_{sc} almost as large as the J_{sc} simulated assuming lossless transport across the interlayer, PCE of 5.16% and the highest FF reported for the organic tandem cells .[58] Recently, inverted tandem cell utilizing a larger band gap P3HT: PCBM as front cell serially connected with a lower band gap (poly(dithienothiophene-co- dialkoxybenzothiadiazole) (PDTTDABT): PCBM back cell was fabricated by employing a three-layered solution-processed recombination structure consisting of vanadium (V) oxide, silver (Ag) nanoparticles and Zn(II) oxide (Figure 5b). This tandem cell yielded a PCE of 1% which is mainly limited by low internal quantum efficiency of the back cell ensuing poor current matching with the front cell resulting in a low J_{sc}. A similar tandem cell with P3HT: PCBM in the back cell yielded 1.4% efficiency led by relative increases in J_{sc} & $V_{oc,}$ evaluated of the current density-voltage characteristic (Figure 5c) of the fabricated device. However, none of the trialed recombination layer yielded perfect serial connection and thereby leading to low efficiency .[59]

More recently, to simplify the fabrication of organic solution-processed tandem cell architecture a sticky interconnect using graphene oxide/PEDOT through direct adhesive lamination process (gluing) was produced, which helped in eliminating the constraint imposed by orthogonal solubility during solution processing. Addition of insulating graphene oxide also increased the conductivity of PEDOT by altering its resonating ground state and morphology through strong π-π interaction and hydrogen bonding .[60]

The presence of undesirable materials such as oxygen and moisture can cause damage to energy devices leading to decline in device performance and a poor long-term stability. To counter such effects researchers have focused on the development of artificial self-healing materials for energy harvesting and storage devices .[61] A coumarin functionalized polyisobutylene (PIB) tri-arm star polymer-based self-healing polymer was developed by Fraust et al. as a smart material for the fabrication of solar cells. This polymeric material on irradiation with UV light shows an excellent light-induced self-

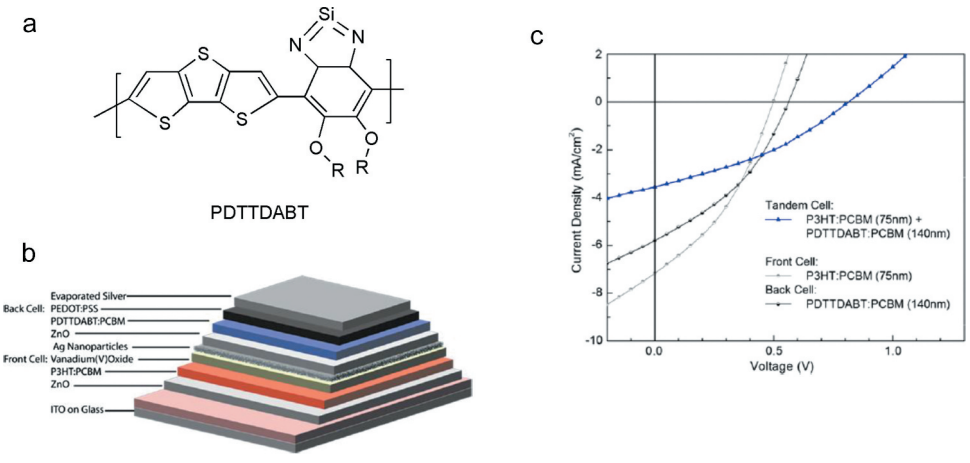


Figure 5. (a) Structure of PDTTDABT, (b) The layer stack of the best performing inverted tandem solar cell (c) J–V curve of the best performing tandem cell using a larger band gap polymer in the front cell (P3HT), and a lower band gap in the back cell (PDTTDABT).[59] Copyright 2011, Elsevier Science Ltd.

healing ability by performing a reversible cross-linking reaction. The solar cells generated using cross-linked coating of PIB showed a resistance toward oxygen and moisture resulting in efficiency retention of upto 85% even after 20 days in comparison to 10% efficiency retention for non-PIB encapsulated solar cells .[62]

2.2. Supercapacitors

Energy storage represents a vital prospect to meet the increasing global energy needs. Batteries have dominated the energy storage market till date. However, they possess lower power density, lower number of charge-discharge cycles, and higher rates of charging and discharging. These limitations can be overcome with the use of supercapacitors, which are considered to have the greatest potential in the field of energy storage devices.

In 1745 a new possibility opened up for capacitive charge storage and further in 1978 "Supercapacitor" was devised. Inevitably capacitors consist of a set of parallel conducting plates split by an insulator with high dielectric constant, accommodating capacitive charges in return of voltage variation amid them. Supercapacitors, also known as ultra-capacitors or electrochemical capacitors, are a transitional power source wedged between rechargeable batteries and electrolytic capacitors. A supercapacitor comprises mainly of two electrodes, electrolyte, and separator, where the electrodes are immersed in the electrolyte, and play a critical role in the property of a supercapacitor. They are in fact two capacitors connected in series, integrated via electrolytic media and propose a higher power density than batteries and higher energy density than conventional capacitors. These can be fully charged/discharged only in a few seconds, resulting in very higher power density uptake or delivery (10 kW/kg) .[63] Supercapacitors can be divided into two types based on the difference of the storage mechanism viz. electrochemical double layer capacitors (EDLCs) and pseudocapacitors. In EDLC, the energy storage mechanism is based on the electrostatic interaction between ions on the large specific surface area of active electrode materials and electrolytes. In 1980, standard oil company of Ohio (SOHIO) introduced their Electrochemical Double Layer Capacitor (EDLC) called "Maxcap." Subsequently, there have been numerous advances in EDLCs with reduced size and much increased capacitance .[64,65]

For EDLC the Double layer capacitance is a result of electrostatic (Non Faradaic) charge storage in form of counterbalancing ions of surface or charges at the electrode electrolyte interface. On the other hand in case of pseudocapacitors, the energy storage mechanism is via quick and reversible Faradaic redox reaction on the surface and bulk at the vicinity between active electrode materials such as metal oxide or conducting polymers and electrolyte ions.

Electrode materials play a decisive role on the performance of the pseudocapacitors, thus in recent years, the researchers have focussed primarily on the development of electrode materials that can improve the electrochemical performance of supercapacitors via synergistic effects. Conducting polymer (CP)-based electrode materials have gained much importance in supercapacitors because of their unique advantages including good conductivity, high energy density, flexibility, relatively cheap, and ease of fabrication[66–69] .

Supercapacitor devices made exclusively from conducting polymers are categorized in to three: Type I (symmetric) employing identical *p*-dopable polymer for both electrodes. Type II (asymmetric) employing two dissimilar *p*-dopable polymers with a different sort of electroactivity. Type III (symmetric) employs same polymer for both electrodes with the *p*-doped form as the positive electrode and the n-doped form as the negative electrode .[70] CPs potential in this application, however, has not been fully exploited yet, due to some limitations of CPs, most importantly related to their relatively low cycling stability and limited mobility of anions in bulk CP layers. Hybridization and composites with different inorganic/organic materials aim to overcome these limitations. In order to perk up the properties of CP-based supercapacitors, researchers have attempted to synthesize CP-based binary composites such as CP-metal oxides, CP-carbon materials and the CP-based ternary composite such as composite made of CP-carbon-metal oxide as supercapacitor electrode active materials to improve electrochemical performance of supercapacitors via synergistic effects. In most of the available examples, both components can contribute with their intrinsic electroactivity but the real synergy stems from other aspects of hybrid formation. Importantly, in the case of appropriately designed composites, the specific capacitance of the hybrid is not just a linear combination of that of the constituents, but in optimal cases the specific capacity of the hybrid can be higher than that of its components. The composites usually have smaller charge transfer resistance than the pure polymer, which further enhances the performance in electrochemical capacitors. The ternary composite as prepared by Cho et al.[71] (RuO_2 Nanoparticle/PEDOT:PSS/Graphene) exhibited a high conductivity of 1570 S cm^{-1} and high energy density of 73 Wh/Kg with a cyclic retention of 81.5% after 1000 cycles. Xia and coworkers[72] formulated a composite electrode with core shell architecture on graphite foam (Graphite Foam + Co_3O_4/PEDOT–MnO_2 core/shell nanowire

arrays). The use of metal oxide and polymers with graphite foam ensured the highest energy density of 9.8 Wh/Kg and the electrode performed for 20000 cycles with 90.2% retention capacity.

Besides these configurations hybrid supercapacitor systems too exist in which the positive electrode of conducting polymer is combined with negative-activated carbon electrode. While persuading the provisions of technology applications in different means, composite electrode materials have been thought about. A core-shell Ppy/PANI composite prepared via in-situ chemical oxidative polymerization of aniline on the surface of Ppy nanotubes inferred good electrochemical performance in acidic and neutral electrolyte. Ppy/PANI composite demonstrated a specific capacitance of 416 F g^{-1} in 1 M H_2SO_4 electrolyte and 291Fg^{-1} in 1 M potassium chloride (KCl) electrolyte .[73] To overcome the technical difficulties of corrosion and leakage allied with the practice of liquid electrolyte, a Ppy based all solid state redox supercapacitors using poly (methyl methacrylate) (PMMA) based gel electrolyte was investigated. Cells clarified the capacitive behavior with the larger value of capacitance 15.3–22.5 mF cm^{-2} (equivalent to single electrode specific capacitance of 120–178 F g^{-1} of Ppy) and moderately lower values of internal resistance due to the flexible nature of the gel electrolytes. A comparative study also indicated that the capacitive behavior of redox cell is resistant to the type and size of the cations of the salts used in the gel electrolytes .[74] Comprehensive study of an asymmetric (ITO/Steel) and symmetric (ITO/ITO, Steel/Steel) supercapacitor devices fabricated of nanometric films of PEDOT indicated highest value of specific capacitance and remarkable electrochemical stability for asymmetric ITO/Steel configuration analogous to the one that is measured for advanced PEDOT-inorganic hybrid composites. The exceptional characteristics of the asymmetric supercapacitors have been attributed to the ultraporous structure of the ultrathin films deposited on ITO, which held in undisturbed following oxidation–reduction processes, and also to the different electroactivities of the two electrodes.

Significant influence of substrate on the morphology and topography of the polymer film was also spotted. The morphology of films deposited on steel consisted of compact and relatively dense packing of sticks. While onto ITO layer polymer grew homogeneously as porous network of ultrathin sticks of PEDOT .[75] Usage of assorted carbonaceous substances for support of active materials not only resulted in an increased specific capacitance of active materials; however, also eased the cycle degradation problems. Surfacing of the hot point research on composite materials, an m-CSA/PANI composite was judged over as electrode for electrochemical supercapacitor. PANI nanoparticles was grown on the surface of carbon spheres aggregations, with functional groups on the surface of m-CSA acting as "structural directors" for chemical polymerization of aniline-forming PANI nanoparticle aggregation. PANI content of less than 60 wt. % exhibited improved electrochemical performance and a maximum specific capacitance 435 F g^{-1} was obtained .[76]

Coalescence of PANI with assorted porous carbon materials has been substantial to sustain the stability of PANI as well as to maximize the capacitance value. For instance, specific capacitance as high as 900 F g^{-1} at a current density of 0.1 A g^{-1} was obtained on direct introduction of PANI to OMC (ordered mesoporous carbon) .[77] It is worth noting that the mixing process does affect the electrochemical properties of PANI/OMC composite as electrodes for supercapacitors. OMC/PANI composite synthesized in-situ using OMC as a starting material exhibited the highest specific capacitance of 747 F g^{-1} at a current density of 0.1Ag^{-1}. This value is more than twice the value compared to the composite prepared via direct physical mixing. Such composite also displayed excellent rate capability, attributed to the high degree of dispersion of PANI on carbon matrix and the contact of PANI with electrolyte as well as the double fixing effects of surface and mesopore of OMC on PANI .[78] Recently, new group of porous carbon materials, i.e., calcium carbide-derived carbons (CCDCs), in which non-carbon elements are selectively etched from the carbide via reaction with chlorine at elevated temperature have been utilized as electrode material with polyaniline modification for supercapacitors. The supercapacitive behaviors of the CCDC/PANI composite materials showed higher specific capacitances of 713.4 F g^{-1} at 1 mVs^{-1} than the as grown CCDC electrodes (154 F g^{-1}) and also offered higher stability than the conducting polymers. Additionally, coin supercapacitor retained 80.1% capacitance even after 1000 cycles .[79] Moreover, scrutinizing the performance of carbon composite (activated carbon (AC) and 5% carbon black (CB)) electrodes for electrochemical double layer supercapacitor devices following PEDOT joining, brought out that supplementary pseudocapacitance provided by PEDOT enhanced the capacitance of the cell to a maximum value at 5% w/w PEDOT. PEDOT also acted as a binder for AC replacing conventional insulating binder such as polyvinylidene fluoride (PVDF). This proved to be beneficial for capacitors used in low-frequency applications, which still hold immense array of relevance such as automotive energy storage .[80]

Superior variety of composite from nanotubular material and conducting polymer in the doped state outlined in a three dimensional volumetric supercapacitor utilized complementary properties of both components for overcoming cycling instability of electron-conducting

polymers (ECPs). For finest application of ECPs in electrochemical capacitors, a particular electrode composition with 20 wt % of CNTs and a vigilant choice of the potential range is essential. Asymmetric capacitors built with MWCNT/Ppy and MWCNT/PANI as negative and positive electrodes, respectively, supplied capacitance up to 320 F g^{-1}.[81] Intending toward modifying the attributes of the capacitors through the optimization of their components, and packaging in the direction of portable systems applications, graphene platelets have been functional. Graphene platelets a new kind of carbon material with two-dimensional (2D) honeycomb lattice structure have been applied profitably in juxtaposition with aniline monomer to produce highly conductive nanocomposite material for functioning as an electrode material meant for high power supercapacitors. A specific capacitance between 300 and 500 F g^{-1} at a current density of 0.1Ag^{-1} was observed with graphene–PANI nanocomposite materials .[82]

Li et al.[83] exemplified on the utilization of graphene subunit incorporated polymeric material or the transitional intermediates between traditional polymer and grapheme for energy storage purposes represented in Figure 6. These graphenal polymers contain a large quantity of graphene subunits along with numerous pores in it which facilitate fast and efficient electron as well as ion transport channels when used as electrochemical material in batteries and supercapacitors. The chemical functional groups present on polymeric organic fractions in these graphenal polymers amplify the electrochemical performance of the supercapacitors by the integration of pseudocapacitive chemical groups for supercapacitor applications.

Optimized novel nanocomposite with nanofibrillar structures comprising interconnected tetrapropylammonium-manganese oxide (TPA-MO) and polypyrrole exhibited enhanced charge capacity, and a maximum specific capacitance of 870Fg^{-1}, if the molar ratio of TPA-MO and PPy approached 1:1. Such a buildup is acclaimed to better material indulgence on account of reasonable microstuctural parameters particularly, specific surface area which is remarkably affected by the molar ratio of reactants and post-synthesis calcination temperature .[84] Polyaniline electrochemically deposited on tantalum sheet was used as matrix to embed hydrous ruthenium (IV) oxide (RuO_2) particles proposing a composite electrode for supercapacitors. PANI due to its porous morphology provided a larger surface area to growing RuO_2. RuO_2/polyaniline electrode revealed high specific capacitance (474 F g^{-1}), ideal cyclical stability and smaller charge transfer resistance (2.24 Ω) .[85] Recently, Ag has been incorporated in to PANI via one-step electrodeposition method to prepare ultra high porous composite for use as pseudocapacitor electrode. The Ag/PANI composite demonstrated ultrahigh specific capacitance of 420 F g^{-1} when measured in 0.5 M Na_2SO_4 at 10 mVs^{-1}which is about 2-fold higher when compared with pure PANI. Greater electrical conductivity and rapid mass transport of electrolytes (i.e., Na^+ and H^+) through the macropores of the Ag/PANI composite was held responsible for such superior capacitive behavior. In addition 94% of composites original capacitance was retained even after 2000 cycles .[86] A novel integrated composite made up of PANI strands, anatase titanium dioxide (TiO_2) nanoparticles, and GO (graphene oxide) sheets featured as a potential material for

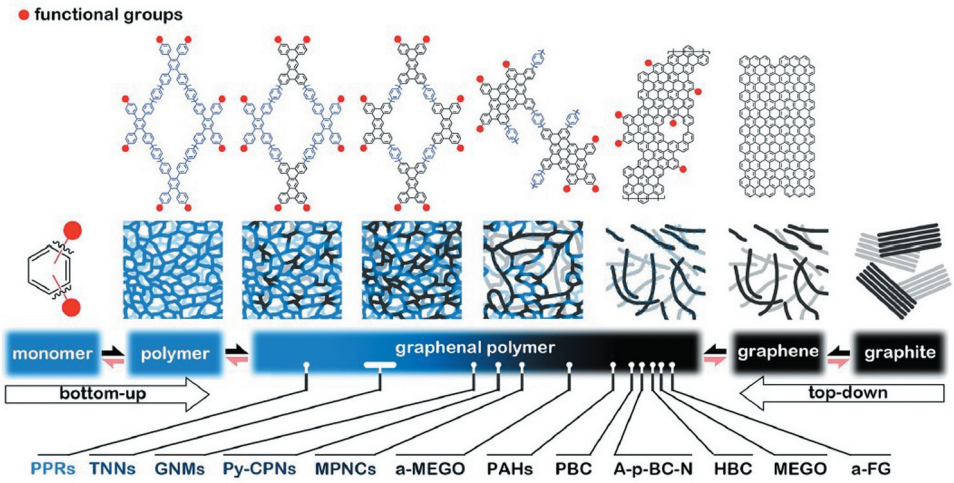


Figure 6. Schematic illustration showing the concept of graphenal polymers that appears as a traditional intermediate lying between the traditional polymers and graphene. Two ways of synthesis bottom up with monomers and top down approach with graphite as the starting material.[83] Copyright 2014, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

supercapacitor or other energy source system. The insertion of the GO and TiO_2 enhanced the electrode conductivity and stability. The intimate and cross binding of the components created a synergic effect and provided superficial electron transport amongst components, ensuing high specific capacitance (1020 F g^{-1} at 2 mV s^{-1}, 430 F g^{-1} at 1 mV s^{-1}) and long cycle life (over 1000 times) for the PANI/TiO_2/GO composite .[87]

One dimensional nanostructured conducting polymers were recognized to be promising electrode materials for supercapacitors. PANI nanofibers prepared by novel procedure using the fiber template route via binary oxidant system have been reported as the electrode material for super capacitors with fast faradic process supplying high specific capacitance value of 428 F g^{-1}. This porous nanostructure not only presented a large liquid–solid interfacial area and possibility for the electrolyte incorporation and removal however, also provided a short and straightforward conduit for ion transport .[88] A material with 3D arrayed nanowire architecture such as PANI nanowire arrayed electrodes synthesized via anodic deposition technique by means of membrane-template synthesis route can be decisive as electrode materials of a supercapacitor to provide high power density. Encouraging outcomes of PANI nanowire arrayed electrodes such as excellent specific capacitance of 1142 F g^{-1}, and good rate capability at a large current density, recommended it for construction of next generation electrochemical supercapacitor. These excellent capacitive abilities is because of better utilizing rate of PANI, owing to smaller diameter of the nanowires. The distinctive 3-D array structure not only lessened the electrolytes diffusion resistance however, also reinforced the ion transportation feature for extremely rapid charge/discharge reactions .[89] For the record, an asymmetric pseudocapacitor cell (Type II), consisting of a graphite anode customized with a novel conducting polymer, poly(5,12-dihydrothieno[3′,4′:2,3][1,4]dioxocino[6,7-b]quinoxaline) (pTDQ) and a graphite cathode modified with PEDOT, was primed. Supercapacitive performance investigation of this n/p type asymmetric pseudocapacitor presented highest specific capacitance value of 213.7 F g^{-1} in the presence of 0.1 M tetrabutylammonium tetrafluoroborate (Bu_4NBF_4) as supporting electrolyte, in conjunction with a specific energy value of 31.7 Wh kg^{-1} .[90]

Provided that the most preeminent line of attack to increase the specific energy of the ECP-based supercapacitors at a constant specific capacitance was to employ ECPs with distinctive operating potentials for the negative and positive electrodes and in a way increase the working potential window. So, a hybrid asymmetric electrochemical supercapacitor was assembled by employing two different nanostructured conducting polymers, i.e., polypyrrole (Ppy) and self-doped polyaniline (SDPA) (Figure 7a). Broadened working potential window up to 1.3 V of the hybrid Ppy/SDPA supercapacitor in aqueous media resulted in an increased charge-discharge time and presented columbic efficiency of approximately 95% (Figure 7b). The supercapacitor substantiated high specific energy, specific power, and maximum power values of 22.87 Wh kg^{-1}, 570 W kg^{-1} and 45,193 W kg^{-1} respectively, at a current density of 5 mA cm^{-2} in 1 M potassium chloride (KCl). As proposed supercapacitor demonstrated good electrochemical stability as it retained 89% of its capacitance at a higher current density (40 mA cm^{-2}) .[91] Self-contained fully flexible symmetric cell consisting of two identical Ppy/Nafion polymer layers

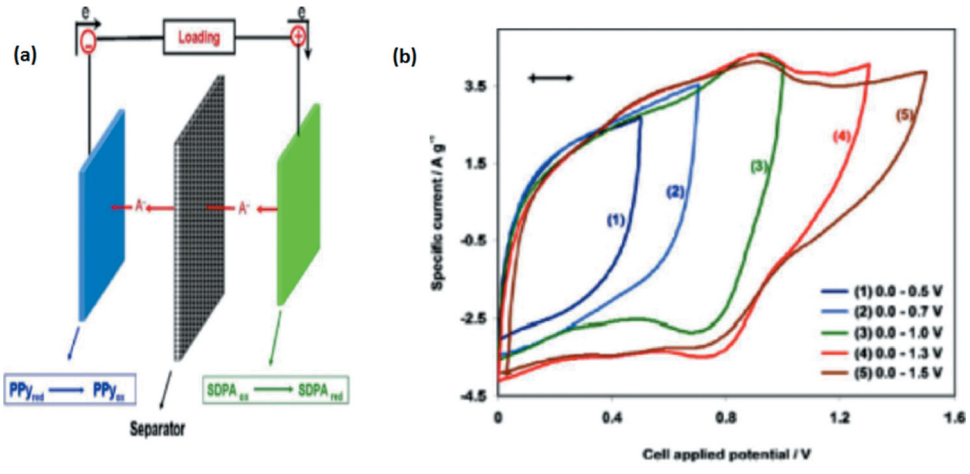


Figure 7. (a) Schematic view of the hybrid PPy/SDPA supercapacitor with SDPA as the positive electrode, PPy as the negative electrode and AGM as a separator. (b) Cyclic voltammograms of the optimized hybrid PPy/SDPA supercapacitor at different working potential windows at a scan rate of 20 mV s^{-1} in 1 M KCl.[91] Copyright 2012, Elsevier Science Ltd.

separated by Au/PVDF/Au membrane was pursued that displayed a capacitance of 318 F g^{-1}at 10 mVs^{-1}. The fully flexible symmetric cell retained about 25% of its specific capacitance (30 F g^{-1} of capacitance) past 5000 cycles in aqueous medium demonstrating utilization of conducting polymer prepared on flexible substrate laying basis of wearable electrode for supercapacitors .[92]

A new class of porous materials known as Metal-organic frameworks (MOFs) having high porosity and high specific surface areas along with CPs can act as a high performance supercapacitors. Wang et al.[93] developed a flexible supercapacitor consisting of cobalt-based MOFs (ZIF-67) and electrochemically polymerized PANI coated on a carbon cloth. The synergistic effect of EDLC capacitance of MOF and the pseudocapcitance of PANI in PANI-ZIF-67-carbon cloth resulted in an ultrahigh areal specific capacitance of 2146 $mFcm^{-2}$ at 10 mVs^{-1} in comparison to 1.47 $mFcm^{-2}$ of PANI-carbon cloth.

2.3. Batteries

Previously, batteries could portray well as energy storage devices. Steps forward to achieve light weight batteries is to exploit the novel materials such as conducting polymers foreseeing improvement in discharge capacity. Many conducting polymers suggestive of PANI, [94] Ppy, [95] PTh, [96,97] PEDOT, [98] and their composites[99] are favored to be used as electrode in rechargeable batteries alongside lithium electrode. The possibility of commercial PANI rechargeable batteries was suggested after the first successful demonstration of chemically synthesized PANI powder as a cathode in dry (due to low amount of electrolyte used) rechargeable batteries. Addition of graphite and acetylene black to PANI powder doped with perchlorate ion and subsequent pressing of mixture improved the conductivity of cathode and hence the battery efficiency. Selected alloy powder added to the negative electrode (mixture of zinc powder, magnesium oxide, zinc oxide, and carboxy methyl cellulose (CMC) pressed to cylindrical shape) prevents the evolution of hydrogen gas. As assembled battery revealed an open circuit voltage (OCV) of 1.64 V, charge storage capacity of 125.43 $mAhg^{-1}$ and displayed columbic efficiency greater than 90% during the first 100 charge-discharge cycles.[100]

It has been gathered that conducting polymer composite such as PANI/TiO_2 offers better charge-discharge characteristics up to 50 cycles and better cyclability when used as cathode material in rechargeable battery comprising of zinc container as anode, cellulose acetate as the separator and polyvinyl sulfate and CMC as the solid polymer electrolyte (SPE) in comparison to conducting polymer alone .[101] Tin dioxide-Ppy (SnO_2-Ppy) nanocomposite synthesized by a one-pot in-situ oxidative chemical polymerization method have also shown improved cyclability and enhanced electrode kinetics for lithium/tin (Li/Sn) alloying/de-alloying credited to its 1-D nanostructure and supported dispersion of SnO_2 in Ppy matrix .[102]

The foremost foundation for the preparation and the possible application of the PEDOT/molybdenum disulfide (MoS_2) nanocomposite as the cathode electrode to increase the specific capacity in rechargeable lithium battery has been demonstrated employing PEDOT nanoribbons interspersed between the layers of MoS_2 prepared via soft chemistry route of intercalation. Open Circuit Voltage (OCV) of this composite coupled to lithium metal anode gave a higher value of 2.62 V than that observed for pure MoS_2 (2.25 V). The discharge capacity was also found to be significantly enhanced from 40 mAh g^{-1} for pristine MoS_2 to 100 mAh g^{-1} for the nanocomposite which is attributed to the enhanced 'bi-dimensionality' and also to the greater structural disorder in the composite .[103] In comparison to PEDOT, rationalization of the better performance of sulfur analogue of PEDOT with thioether and polythiophene structure as cathode in lithium ion battery was attempted. Poly[3,4-(ethylenedithio)thiophene] (PEDTT) showed discharge specific capacity between 425$mAhg^{-1}$ and 550$mAhg^{-1}$ and a low discharge voltage between 2.0 V and 2.3 V. Tentative electrode reaction with the formation of thio-ether cation was proposed, which imparted multi-electron redox reaction, high discharge specific capacity, high charge voltage, and low discharge voltage to the as assembled battery .[104] In view of increasing the batteries capacity, an aqueous polyaniline rechargeable battery was accessed by means of electro-polymerized PANI on highly porous-reticulated vitreous carbon (RVC) facilitating large amount of polymer deposition on electrode surface. Test regarding the battery performance revealed a less than 5% reduction in the open circuit voltage past 70 cycles and 121 mA hg^{-1} capacities .[105]

Endeavors were considered to explore the viability of a prototype cell made of poly(1-napthol) in methyl cyanide as a primary move for all polymer rechargeable battery. A rechargeable polymer/polymer cell with an electromotive force (EMF) window of at least 1.3 V was constructed from de-doped PANI (reduced form) as positive electrode and doped poly(1-napthol) (oxidized form) as negative electrode. The cell conferred an open circuit potential (OCP) of 1.4 V, a specific capacity of 150 Ah kg^{-1} and an energy density of 113 W h kg^{-1} at 1.0 V cut off. The cell lost 5% of its capacity after 1 week due to self-discharge. The charging/discharging assessments (0.0–1.4 V) showed a slow deterioration in capacity, i.e., up to 15% after 100 cycles. Such battery provides the benefit of high power density (9 kW kg^{-1} at OCP), owing to the wide spread surface area of the

polymer layer in contrast to conventional metal electrode .[106] In an extended study, Wang et al. first coupled a novel anode material i.e. chemically synthesized functionalized polyterthiophene poly (3′-styryl -4,4″-didecyloxyterthiophene) (poly(OC10DASTT)) to a cathode consisting of Ppy doped with hexafluoro phosphate (PF_6^-) and obtained a discharge capacity of 39.1 $mAhg^{-1}$.[107] Advancing with a vision that electro polymerization can result in a more intimate substrate/polymer interaction the same author extended their study to electro-synthesized functionalized polyterthiophenes on nickel/copper (Ni/Cu) coated fabric as anode producing a discharge capacity of 45.2 $mAhg^{-1}$.[108]

A prime report of superior construction and charge/discharge attributes of an organic radical battery employing TEMPO (2, 2, 6, 6-tetramethylpiperidine-1-oxyl) bearing polythiophene as cathode material shown in Figure 8a was set by Aydin and his coworkers. The battery had 79 Ah kg^{-1} initial specific charge capacity which is 87% of its theoretical capacity and undertake charge and discharge cycles between 2.5 and 3.8 V with merely 23% loss past 50 cycles shown in Figure 8b, favoring the use of this material alongside the lithium (Li) anode in rechargeable battery delivering output voltage of 3.6 V .[109]

However, the first attempt at the use of organic radicals as an electrode active species for energy storage system in

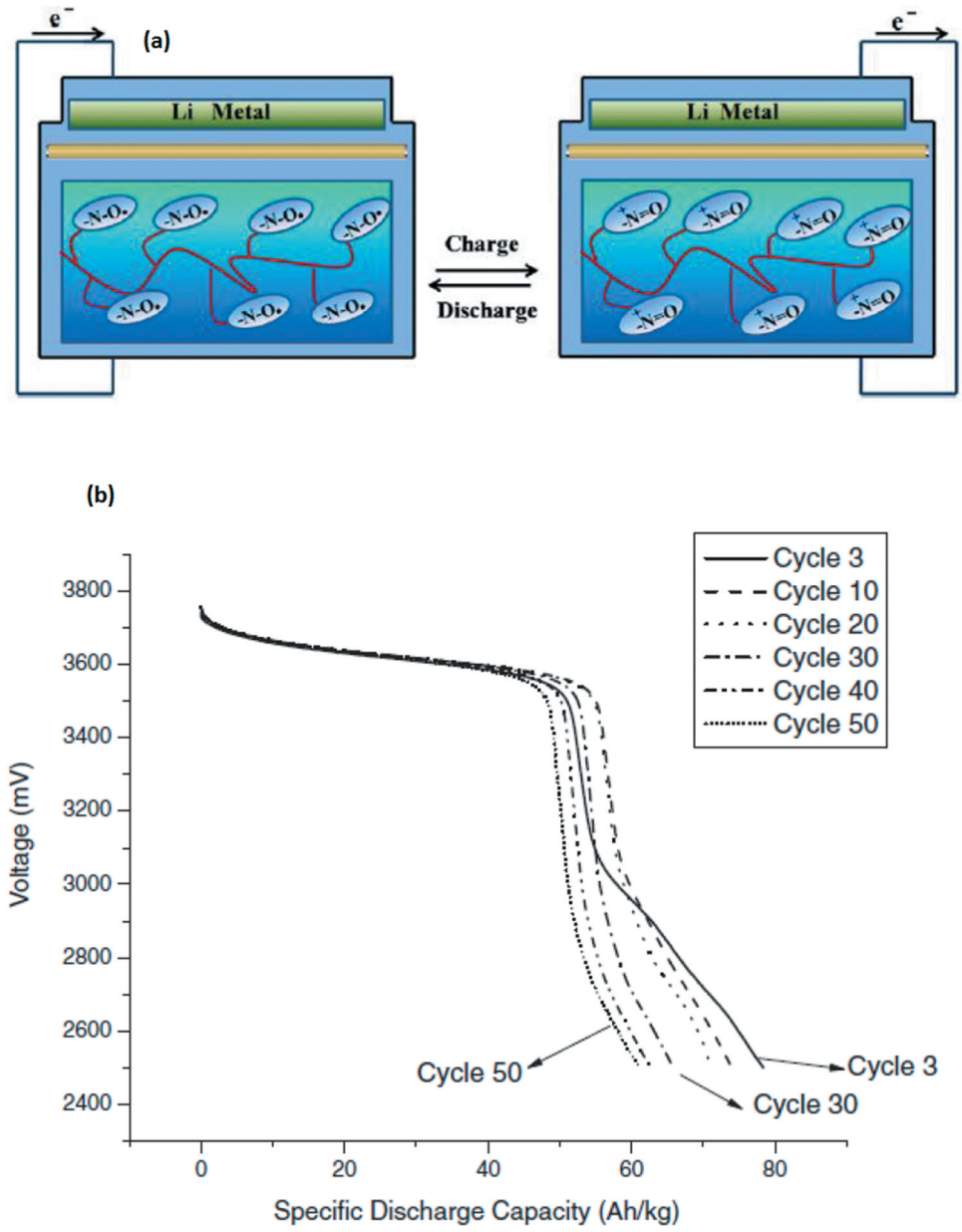


Figure 8. Charging and discharging processes of the cell constructed and Voltage vs. constant current specific discharge capacity at 0.05 mA discharge rate.[109]Copyright 2011, Elsevier science ltd.

lithium battery was made by Nakahara's group once they applied poly(2,2,6,6-tetramethylpiperidinyloxy- 4-yl methacrylate) (PTMA), a polymer with pendant TEMPO radicals, as the organic polyradical cathode-active material in rechargeable batteries and as produced batteries exhibited an average discharge voltage of 3.5 V, a discharge capacity of 77 Ah Kg^{-1} corresponding to 70% of the theoretical value, in a way hopeful to unlock an innovative field of use in favor of plastic batteries .[110]

A method to awaken the act of nonmetal based energy storage devices is likely to use composite electrode materials of conductive polymers, for example Ppy, deposited as thin layers on an appropriate large surface area substrate. A resourceful nanostructured high-surface area electrode material for energy storage applications composed of cellulose fibers, of algal origin individually coated with a 50 nm thin layer of Ppy was initiated. Outcomes confirmed thus far the highest ever reported charge capacities and charging rates for an all polymer paper-based battery. The composite conductive paper material revealed to have a specific surface area of 80 m^2g^{-1} and batteries derived from this material was capable of being charged with currents as high as 600 mA cm^{-2} with only 6% loss in capacity over 100 succeeding charge and discharge cycles .[111]

To strengthen the electrochemical performance of lithium ion batteries (LIBs), no stone has been left unturned. Miscellaneous carbon materials with different micro/macrostructures have been explored. Ppy polymer was transformed to continuous nanofibers (CNFs) by electro spinning. Electrospun multi-phase nanofibers acquired from polyacrylonitrile (PAN)/30 wt% Ppy bi-component precursor once used as anodes to assemble lithium-ion half cells leaving out any polymer binder or conductive additive exhibited high reversible capacity, improved cycle performance, and relatively good rate capability. Also, the apparent fibrous morphology was maintained over 50 charge/discharge cycles .[112] New results pertaining to the covalent functionalization of CNT with PVK and further utilization of this composite in lithium rechargeable cells as positive electrode along with $LiPF_6$ electrolytic solution has also been stated. PVK functionalized SWCNTs and MWCNTs in lithium cells displayed higher specific discharge capacity of 45 and 115 mA h g^{-1} respectively.[113] Recent studies have shown that the introduction of MWCNTs to PANI prepared by an in-situ oxidative chemical polymerization method has presented the polymer, a good electrochemical characteristic as cathode material in rechargeable lithium batteries. An apparent reason for the enhancement by MWCNT inclusion is the increased number of active sites for faradic reaction along with larger specific capacitance than pure PANI and also an enhanced electrical conductivity. Therefore, PANI/MWCNT composite has higher columbic efficiency, large discharge capability and cycle ability .[114] First step toward the integration of all polymer battery in to wearable and other textile structures have been implemented by Wang et al. Evaluating the performance of a fiber battery consisting of Ppy/PF_6 cathode and Ppy/PSS anode indicated capacities of 10 mA h g^{-1} over 30 cycles .[115] Same author attempted to enhance the performance of the fiber battery through the use of SWCNTs as anode material and PVDF hollow fiber membrane as the separator. Two-fold increase in capacity (approximately 20 mAh g^{-1}) was obtained than those found in their previous work on the fiber batteries consisting of a Ppy/PF_6 cathode, Ppy/PSS anode in the 'flooded cells'. They observed decreased capacity for the fiber battery with polypyrrole-hexafluorophosphate (Ppy/PF_6) as cathode and SWCNTs as anode when in the form of 'dry cell' than that of the 'flooded cell' attributing the limiting movement of cations and anions to the lack of electrolyte .[116] Cathode electronic conductivity could be improved by planning composite of sulfur and conductive material such as conducting polymers. Conducting macromolecule Ppy was used as stable backbones to provide electric conducting channels and long polysulfide lines were combined to polypyrrole as side chain in order to offer high specific capacity. Charge–discharge experimental results identified that the polymer had specific capacity of 515 mAh g^{-1} at the first cycle and 452 mAh g^{-1} at the 20th cycle .[117] Additionally, conductive material with high-surface area and porous structure is intended so as, to take up and hold polysulfides effectively. Specific discharge capacities of 1222 $mAhg^{-1}$ and 570 mAh g^{-1} were delivered in the first and 20th cycles, respectively by exploiting a nanowired sulfur/Ppy (S/Ppy) composite with a network structure as cathode. Composite was synthesized via heat treatment method in high energy lithium sulfur secondary batteries .[118] Morphology of Ppy presents a crucial role to stimulate the electrochemical performance. Liang et al. proposed tubular Ppy conductive matrix for supporting sulfur (S) electrode's electrochemical cycling performance and stability in lithium-sulfur secondary battery. S was identified to penetrate into hollow tubular Ppy (T-Ppy) fibers by the capillary forces during co-heating process, perceive out of Brunauer-Emmett-Teller (BET) surface area measurements (Figure 9). The decreasing surface area and pore volume with increasing sulfur content (inset of Figure 9) in the Ppy composite supported the sublimed S incorporation in tubular Ppy structure. The improved cycling stability of S/T-Ppy composite electrode was suggested to be a combined effect improved conductivity and favorably dispersed nano-sized S and generated lithium polysulfide within the T-Ppy matrix. .[119]

Novel synthetic technique, saving energy via simple one-step ball milling process excluding heat treatment

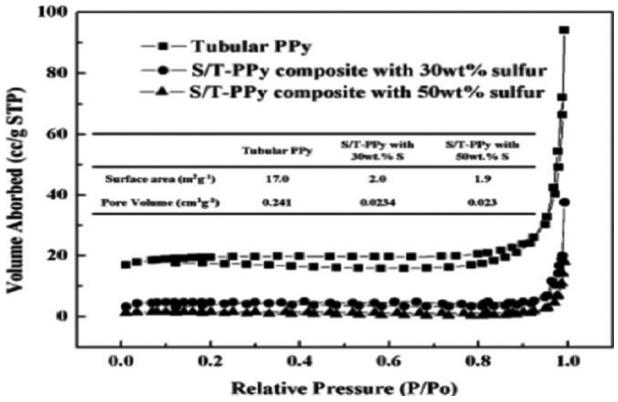


Figure 9. Nitrogen sorption isotherms of tubular PPy and S/T-PPy composites, insets shows the surface areas and pore volumes.[119] Copyright 2011, Elsevier Science Ltd.

was adopted to synthesize sulfur/Ppy binary composite so as to be build up a high energy lithium-sulfur secondary batteries. Dissolution and collapse of cathode material was delayed by the porous nature of the Ppy matrix. Subsequently developed branched composite can withstand the mechanical pressure encountered by the volume changes upon charge-discharge cycles leading to enhancement of charge transfer, delivering discharge capacity of 1320 mAh g^{-1} .[120] Ppy also materialized to be an excellent contender as an active element within a bioelectric battery. Bioelectric battery implanted in to the human body depends on the oxygen in the internal body fluid to produce the electrical energy. Ppy doped with a biological polyelectrolyte (dextran sulfate, an anticoagulant) as cathode and bio-adsorbable magnesium alloy (AZ61) as anode, combined within a commonly used biological media offered an energy density of ~790 Wh Kg^{-1}, sufficient enough to drive some implantable devices requiring low power densities such as pace maker or biomonitoring system .[121] A remarkable stretchable battery electrode was projected by Wang et al., based on buckled biocompatible Ppy-pTSA as the cathode material in magnesium (Mg) batteries. The fabrication method included the relaxation of the pre-strained Ppy-pTSA deposited Au sputter coated (poly (styrene-block-isobutylene-block-styrene)) (SIBS) substrate shown in Figure 10. This material electrode endured 2000 stretching cycles with 30% strain obtaining cell voltage of 1.2 V .[122]

2.4. Polymer electrolyte membranes fuel cell

Fuel cell is an ideal primary energy conversion device which has attracted increasing attention in recent decades due to their propensity to continuously convert chemical energy stored in fuels (methanol, ethanol, hydrogen, and formic acid) into electric energy with high efficiency and without emission of any harmful greenhouse gases such as carbon dioxide, nitrous oxide, and sulfur dioxide. The ease of portability and environment friendliness has led to its wide range of applications in areas, particularly where an assured uninterruptible electrical supply is required. Figure 11 shows a comparison between the main energy conversion devices in the market today.

2.4.1. Working of a fuel cell

Fuel cells consist of two electrodes; a cathode, where reduction occurs, an anode, where oxidation occurs, and an electrolyte to transmit the current between the electrodes. In comparison to common battery materials, the fuel cells have higher energy conversion efficiency due to the greater chemical energy stored in hydrogen and several hydrocarbon fuels. The fuel cells can operate as long as fuel and oxidant are continuously supplied from the external supply. It can generate electricity without combustion of any substance by merely combining oxygen and hydrogen to produce water and heat. The net reactions occurring in a fuel cell in acidic and alkaline media are:

In acidic electrolyte

Anode: $H_2 \rightarrow 2H^+ + 2e^-$

Cathode: $\frac{1}{2} O_2 + 2 H^+ + 2e^- \rightarrow H_2O$

In basic electrolyte

Anode: $H_2 + 2OH^- \rightarrow 2 H_2O + 2e^-$

Cathode: $\frac{1}{2} O_2 + H_2O + 2e^- \rightarrow 2OH^-$

The overall basic reaction of a H_2-O_2 fuel cell is

$H_2 + \frac{1}{2} O_2 \rightarrow H_2O$ $E° = 1.229$ V

For this reaction $\Delta G° = -235.76$ kJ/mol and $\Delta H° = -285.15$ kJ/mol. So, there is no thermodynamic limitation on its efficiency. The role of electrode in fuel cell is quite imperative as it not only conducts electrons from or to the reaction sites but also ensures a stable interface between the gaseous reactant molecules and the electrolyte. Fuel cells can be classified into different categories depending on the types of electrodes, its mode of operation and fuel used, namely;

(i) Based on the type of electrolyte: (a) Phosphoric acid Fuel cell, (b) Polymer electrolytic membrane fuel cell (PEM), (c) molten carbonate fuel cell, (d) solid oxide fuel cell, and (e) alkaline fuel cells

(ii) Based on Types of Fuel and oxidant: (a) Ammonia – air fuel cell (b) direct methanol fuel cell (c) Hydrogen (pure)-Oxygen (pure) fuel cell (d) Synthesis gas-air fuel cell and (e) Hydrogen-rich gas-air fuel cell

Among the various fuel cell systems, PEM fuel cell has gained interest due to its application in small scale products such as automotive, portable, and stationary power generation equipments. The main advantages of PEM fuel cell include quick start-up, high efficiency, low

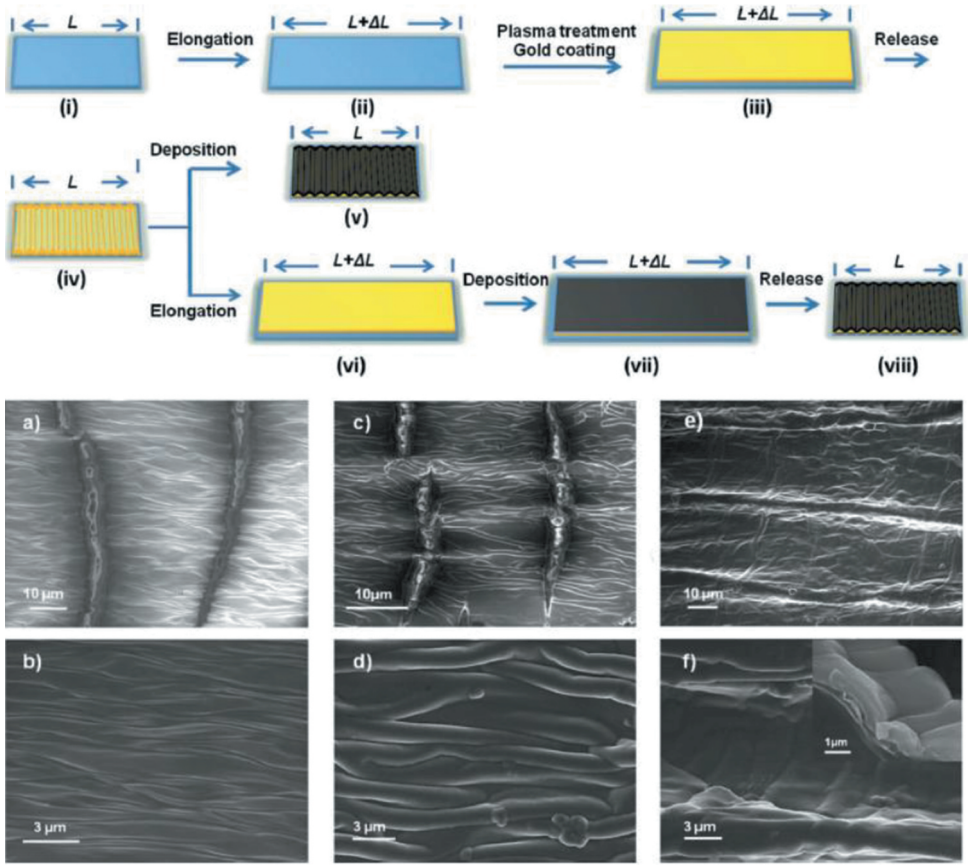


Figure 10. Schematic procedures used to prepare buckled Au or PPy- pTS films on SIBS substrate and FESEM images of these products. Buckled PPy-pTS fabrication comprises surface treatment, Au sputter coating, relaxation of the pre-strained SIBS substrate and PPy electrodeposition. 2D buckledPPy-pTS film was electrodeposited on the re-elongated Au coated substrate, obtained after the relaxation of the substrate. FESEM images: (a) and(b), buckled Au film (iv); (c) and (d), buckled PPy- pTS film (v); (e) and (f), 2D buckled PPy- pTS film (viii). Inset, side view of PPy- pTs film showing the trivial buckles.[122]Copyright 2011, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

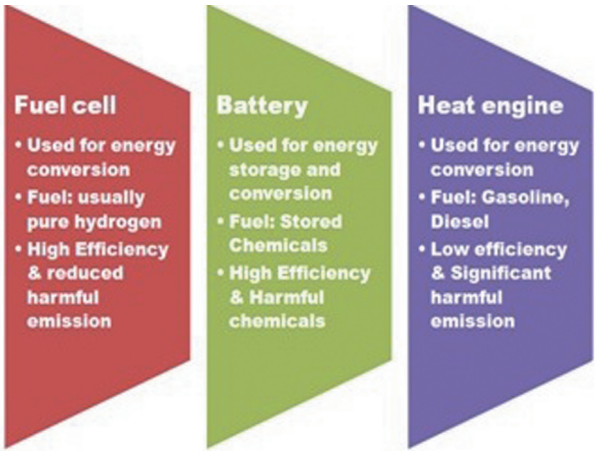


Figure 11. Comparison between fuel cell, battery and heat engine.

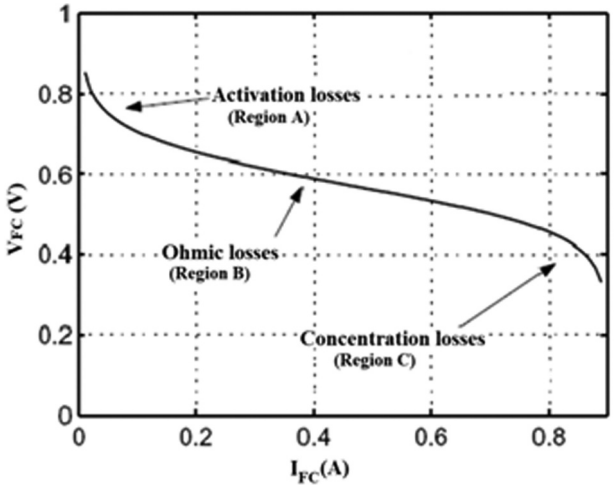


Figure 12. Typical fuel cell polarization curve.

operating temperature, and low CO_2 emissions. The fuel cell I–V characteristic curve (also known as Polarization curve) can be divided into three major regions (shown in Figure 12) depicting three different types of losses namely; (a) activation losses (Losses due to the rate of reaction), (b) ohmic losses (losses due to resistance) and (c) concentration losses (mass transport loses which are

due to the water generated from the reaction blocking the channels and leading to the limited diffusion of reactant gases) .[123,124]

The point at the boundary of regions B and C is known as **maximum power density point or knee/ optimum point**. Loading of the fuel cells above this point leads to the shifting of the operating point right of the region C causing a sudden collapse of the fuel cell voltage to zero resulting in no power drawn from the cell. Further operations in this region C may damage the fuel cell. Region B of the characteristic curve is the most common region for the operation of fuel cell.

The key components of a PEM fuel cell are biplates, gas diffusion layer, catalyst, and polymer electrolyte membrane. The polymer electrolyte membrane plays a pivotal role in overall functioning and performance of the fuel cell. Hence, in this regard, research has been going on to discover novel materials with exceptional properties that can be used in the device fabrication.

Nafion®, a copolymer of tetrafluoroethylene monomer having pendant side chains of perfluorinated vinyl ether perfluorosulfonic acid terminated groups, is the state-of-the-art PEM, and is currently being used in fuel cells [Figure 13]. The Teflon-like backbone of Nafion® makes it highly durable under oxidative and reductive environments while its fully hydrated membrane (Nafion®117) exhibits a proton conductivity of 0.10 S cm^{-1} and current density of 750mAcm^{-2} with 150 mV voltage loss .[125]

Nafion® membrane has some major drawbacks such as high permeability of methanol, poor conductivity, drying at temperatures above 90°C, and poor performance under low humidification, etc. The proton conductivity of Nafion® membrane can be increased by increasing the degree of sulfonation but this can lead to swelling of the membrane, causing the mechanical strength to decrease .[126] To overcome these drawbacks of Nafion® membranes and a possible replacement of it, various membranes have been explored in the literature, without much success .[127] Further, various nanocomposites of Nafion® with metallic blocking layers or inorganic nano-fillers have been proposed with combined attractive properties of organic polymer such as proton conductivity, specific chemical reactivity, flexibility, ductility and ease of processing, and the properties such as mechanical, thermal, and chemical stability of inorganic backbone with reduced fuel permeability .[128]

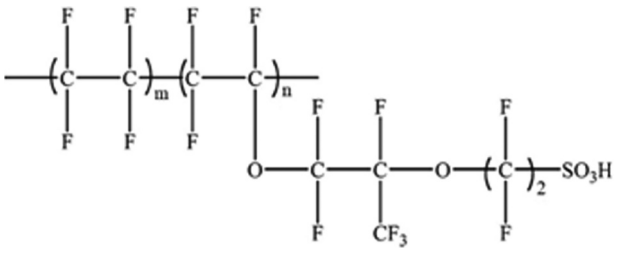


Figure 13. Chemical structure of commercial Nafion® membrane, where m and n represent molar composition and do not imply a sequence length.[125].

Some of the commonly used nanofillers include metal oxides, [129] Silica, [130] ionic liquid, [131] heteropoly acids, [132] graphene, and graphene oxide, [133,134] and CNTs[135] with polymer matrices such as Nafion, sulfonated poly(ether ether ketone), poly(ether sulfone), and poly(benzimidazole) .[136,137]

3. Conclusion

This report provides an overview of the recent advances of polymeric materials in terms of their synthesis approaches, their physical, chemical, optical, and electronic properties, morphology engineering, and development of new and advanced technologies for energy storage, and harvesting. Having outshined the emerging phase up to now and crossing over to the growth phase, there remain plentiful opportunities for innovation; however, sustainability still remains the key guiding thought. The key challenges for the high-throughput polymer-based production technologies for their commercialization and its future research perspective are discussed in detail in order to reach the phase of energy excellence.

Notes on contributors

Sarita S Nair did her B.Sc. and M.Sc in Chemistry from University of Delhi, Delhi. She did her Doctorate in Chemistry on "Development and efficiency studies of organic solar cells" from Delhi Technological University, Delhi. She is keenly interested and highly motivated in the field of smart materials, organic electronics particularly organic solar cells, material characterization, synthesis and characterization of conjugated polymers and nanotechnology. She has to her credit 9 publications in international journals of high repute. She has more than 4 years of experience in teaching and 2 years of industrial experience.

Sujeet Kumar Mishra is currently working as an Assistant Professor in the Department of Chemistry, Ramjas College, University of Delhi. He has received his Ph.D. degree from CSIR-National Physical Laboratory and Delhi Technological University. He has done his Masters in Chemistry from the University of Delhi. He has received a Merit Scholarship award from the Delhi government and has qualified CSIR-NET (JRF) and GATE exams in Chemical Sciences. He has published more than 15 research publications in various reputed international journals. He has also authored various online modules for the SWAYAM platform of MHRD, Govt. of India. He has highly enriched research experience in different areas of nanomaterial sciences and biosensors.

D. Kumar born on Nov 5, 1965 did his Doctorate on "Electrochemical, optical and thermal studies on conducting polyanilines" from University of Delhi, Delhi, India. At present, Dr. Kumar is working as professor in the Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi. He has published over 100 papers in the journals of international repute in the areas of conducting polymers, sensors, conductive adhesives, smart hydrogels, helical materials, organic solar cells, toughening of thermosetting polymers, self healing and blast mitigating polymer coatings.

References

[1] Kumar, D.; Sharma, R. C. Advances in Conductive Polymers. *Eur. Polym. J.* 1998, *34*, 1053–1060. DOI: 10.1016/S0014-3057(97)00204-8.

[2] Heeger, A. J.;. Semiconducting and Metallic Polymers: The Fourth Generation of Materials. *Angew. Chem. Int. Ed.* 2001, *40*, 2591–2611.

[3] Nair, S. S.; Mishra, S. K.; Kumar, D. Recent Progress in Conductive Polymeric Materials for Biomedical Applications. *Polym. Adv. Technol.* 2019, *30*(12), 1–22. DOI: 10.1002/pat.4725.

[4] MacDiarmid, A. G.;. Synthetic Metals: A Novel Role for Organic Polymers. *Angew. Chem. Int. Ed.* 2001, *40*, 2581–2590. DOI: 10.1002/1521-3773(20010716)40:14<2581::AID-ANIE2581>3.0.CO;2-2.

[5] Wang, M.; Baek, P.; Akbarinejad, A.; Barker, D.; Travas-Sejdic, J. Conjugated Polymers and Composites for Stretchable Organic Electronics. *J. Mater. Chem. C.* 2019, *7*, 5534–5552.

[6] Peng, H.; Sun, X.; Weng, W.; Fang, X. *Polymer Materials for Energy and Electronic Applications*; Academic Press, Elsevier Science Publishing Co Inc: United States, 2016.

[7] Weng, W.; Chen, P.; He, S.; Sun, X.; Peng, H. Smart Electronic Textiles. *Angew. Chem. Int. Ed.* 2016, *55*, 6140–6169.

[8] Huynh, W. U.; Dittmer, J. J.; Alivisatos, P. Hybrid Nanorod-polymer Solar Cells. *Science*. 2002, *295*, 2425–2427.

[9] Xu, T.; Qiao, Q. Conjugated Polymer–inorganic Semiconductor Hybrid Solar Cells. *Energy Environ. Sci.* 2011, *4*, 2700–2720.

[10] Zhao, L.; Lin, Z. Crafting Semiconductor Organic–inorganic Nanocomposites via Placing Conjugated Polymers in Intimate Contact with Nanocrystals for Hybrid Solar Cells. *Adv. Mater.* 2012, *24*, 4353–4368.

[11] Fu, H.; Wang, Z.; Sun, Y. Polymer Donors for High-Performance Non-Fullerene Organic Solar Cells. *Angew. Chem. Int. Ed. Engl.* 2019, *58*(14), 4442–4453.

[12] Zhou, T.; Jia, T.; Kang, B.; Li, F.; Fahlman, M.; Nitrile-substituted, W. Y. QA Derivatives: New Acceptor Materials for Solution-processable Organic Bulk Heterojunction Solar Cells. *Adv. Energy Mater.* 2011, *1*, 431–439.

[13] Kamm, V.; Battagliarin, G.; Howard, I. A.; Pisula, W.; Mavrinskiy, A.; Li, C.; Müllen, K.; Laquai, F. Polythiophene:perylenediimide Solar Cells – The Impact of Alkyl-substitution on the Photovoltaic Performance. *Adv. Energy Mater.* 2011, *1*, 297–302.

[14] Halls, J. J. M.; Walsh, C. A.; Greenham, N. C.; Marseglia, E. A.; Friend, R. H.; Moratti, S. C.; Holmes, A. B. Efficient Photodiodes from Interpenetrating Polymer Networks. *Nature*. 1995, *376*, 498–500.

[15] Smestad, G. P.; Krebs, F. C.; Lampert, C. M.; Granqvist, C. G.; Chopra, K. L.; Mathew, X.; Takakura, H. Reporting Solar Cell Efficiencies in Solar Energy Materials and Solar Cells. *Sol. Energy Mater. Sol. Cells*. 2008, *92*, 371–373.

[16] Berger, P. R.; Kim, M. Polymer Solar Cells: P3HT: PCBM and Beyond. J Renew Sustain Energy. 2018, *10*, 013508–013515.

[17] Mu¨hlbacher, D.; Scharber, M.; Morana, M.; Zhu, Z.; Waller, D.; Gaudiana, R.; Brabec, C. J. High Photovoltaic Performance of a Low-bandgap Polymer. *Adv. Mater.* 2006, *18*, 2884–2889.

[18] Peet, J.; Kim, J. Y.; Coates, N. E.; Ma, W. L.; Moses, D.; Heeger, A. J.; Bazan, G. C. Efficiency Enhancement in Low-bandgap Polymer Solar Cells by Processing Withalkane dithiols. *Nat Mater*. 2007, *6*, 497–500.

[19] Scharber, M. C.; Lungenschmied, C.; Egelhaaf, H. J.; Matt, G.; Bednorz, M.; Fromherz, T.; Gao, J.; Jarzabd, D.; Loi, M. A. Charge Transfer Excitons in Low Band Gap Polymer Based Solar Cells and the Role of Processing Additives. *Energy Environ. Sci.* 2011, *4*, 5077–5083.

[20] Rogers, J. T.; Schmidt, K.; Toney, M. F.; Kramer, E. J.; Bazan, G. C. Structural Order in Bulk Heterojunction Films Prepared with Solvent additives. *Adv Mater*. 2011, *23*, 2284–2288.

[21] Liang, Y.; Wu, Y.; Feng, D.; Tsai, S. T.; Son, H. J.; Li, G.; Yu, L. Development of New Semiconducting Polymers for High Performance Solar Cells. *J. Am. Chem. Soc.* 2009, *131*, 56–57.

[22] Chen, H. Y.; Hou, J.; Zhang, S.; Liang, Y.; Yang, G.; Yang, Y.; Yu, L.; Wu, Y.; Li, G. Polymer Solar Cells with Enhanced Open-circuit Voltage and Efficiency. *Nat. Photonics*. 2009, *3*, 649–653.

[23] Hou, J.; Chen, H. Y.; Zhang, S.; Chen, R. I.; Yang, Y.; Wu, Y.; Li, G. Synthesis of a Low Band Gap Polymer and Its Application in Highly Efficientpolymer Solar Cells. *J. Am. Chem. Soc.* 2009, *131*, 15586–15587.

[24] Liang, Y.; Xu, Z.; Xia, J.; Tsai, S. T.; Wu, Y.; Li, G.; Ray, C.; Yu, L. For the Bright Future—bulk Heterojunction Polymer Solar Cells with Power Conversion Efficiency of 7.4%. . *Adv. Mater.* 2010, *22*, 135–138.

[25] Wakim, S.; Alem, S.; Li, Z.; Zhang, Y.; Tse, S. C.; Lu, J.; Ding, J.; Tao, Y. New Low Band Gap Thieno [3,4-b] thiophene-based Polymers with Deep HOMO Levels for Organic Solar Cells. *J. Mater. Chem.* 2011, *21*, 10920–10928.

[26] Sonar, P.; Singh, S. P.; Li, Y.; Ooi, Z. E.; Ha, T. J.; Wong, I.; Soha, M. S.; Dodabalapur, A. High Mobility Organic Thin Film Transistor and Efficient Photovoltaic Devices Using Versatile Donor–acceptor Polymer Semiconductor by Molecular Design. *Energy Environ. Sci.* 2011, *4*, 2288–2296.

[27] Chochos, C. L.; Choulis, S. A. How the Structural Deviations on the Backbone of Conjugated Polymers Influence Their Optoelectronic Properties and Photovoltaic Performance. *Prog. Polym. Sci.* 2011, *36*, 1326–1414.

[28] Ayzner, A. L.; Tassone, C. J.; Tolbert, S. H.; Schwartz, B. J. Reappraising the Need for Bulk Heterojunctions in Polymer-fullerene Photovoltaics: The Role of Carrier Transport in All-solution-processed P3HT/PCBM Bilayer Solar cells.J. *Phys. Chem. C*. 2009, *113*, 20050–20060.
[29] Gusain, A.; Faria, R. M.; Miranda, P. B. Polymer Solar Cells—Interfacial Processes Related to Performance Issues. *Front. Chem.* 2019. DOI: 10.3389/fchem.2019.00061.
[30] Chiu MY, Jeng US, Su MS, Wei KH. Morphologies of Self-organizing Regioregular Conjugated Polymer/fullerene Aggregates in Thin Film Solar Cells. *Macromol.* 2010, *43*, 428–432.
[31] Yang, X.; Loos, J. Toward High-performance Polymer Solar Cells: The Importance of Morphology Control. *Macromol.* 2007, *40*, 1353–1362.
[32] Kim, H.; So, W. W.; Moon, S. J. The Importance of Post-annealing Process in the Device Performance of Poly (3-hexylthiophene): Methanofullerene Polymer Solar Cell. *Sol. Energy Mater. Sol. Cells*. 2007, *91*, 581–587.
[33] Park, J. H.; Kim, J. S.; Lee, J. H.; Lee, W. H.; Cho, K. Effect of Annealing Solvent Solubility on the Performance of Poly(3-hexylthiophene)/methanofullerene Solar Cells. *J. Phys. Chem. C*. 2009, *113*, 17579–17584.
[34] Nair, S. S.; Kumar, D.; Majumdar, A. Effect of Co-solvents on the Photovoltaic Performance of an Inverted Organic Solar Cell. *Polym. Eng. Sci.* 2015, *55* (6), 1382–1388.
[35] Berson, S.; Bettignies, R. D.; Bailly, S.; Guillerez, S. Poly (3-hexylthiophene) Fibers for Photovoltaic Applications. *Adv. Funct. Mater.* 2007, *17*, 1377–1384.
[36] Kim, J. S.; Lee, J. H.; Park, J. H.; Shim, C.; Sim, M.; Cho, K. High-efficiency Organic Solar Cells Based on Preformed Poly(3-hexylthiophene) Nanowires. *Adv. Funct. Mater.* 2011, *21*, 480–486.
[37] Kim, B. G.; Kim, M. S.; Kim, J. Ultrasonic-assisted Nanodimensional Self-assembly of Poly-3-hexylthiophenefor Organic Photovoltaic Cells. *ACS Nano*. 2010, *4*, 2160–2166.
[38] Kim, J. H.; Park, J. H.; Lee, J. H.; Kim, J. S.; Sim, M.; Shima, C.; Cho, K. Bulk Heterojunction Solar Cells Based on Preformed Polythiophene Nanowires via Solubility-induced Crystallization. *J. Mater. Chem.* 2010, *20*, 7398–7405.
[39] Huang, J.; Xu, Z.; Yang, Y. Low-work-function Surface Formed by Solution-processed and Thermally Deposited Nanoscale Layers of Cesium Carbonate. *Adv. Funct. Mater.* 2007, *17*, 1966–1973.
[40] Jorgensen, M.; Norrman, K.; Krebs, F. C. Stability/degradation of Polymer Solar Cells. *Sol. Energy Mater. Sol. Cells*. 2008, *92*, 686–714.
[41] Zhang, F.; Xu, X.; Tang, W.; Zhang, J.; Zhuo, Z.; Wang, J.; Wang, J.; Xua, Z.; Wang, Y. Recent Development of the Inverted Configuration Organic Solar Cells. *Sol. Energy Mater. Sol. Cells*. 2011, *95*, 1785–1799.
[42] Li, G.; Chu, C. W.; Shrotriya, V.; Huang, J.; Yang, Y. Efficient Inverted Polymer Solar Cells. *Appl. Phys. Lett.* 2006, *88*, 253503–253505.
[43] Rider, D. A.; Worfolk, B. J.; Harris, K. D.; Lalany, A.; Shahbazi, K.; Fleischauer, M. D.; Brett, M. J.; Buriak, J. M. Stable Inverted Polymer/fullerene Solar Cells Using a Cationic Polythiophene Modified Pedot: Pss Cathodic Interface. *Adv. Funct. Mater.* 2010, *20*, 2404–2415.
[44] Stubhan, T.; Oh, H.; Pinna, L.; Krantz, J.; Litzov, I.; Brabec, C. J. Inverted Organic Solar Cells Using a Solution Processed Aluminum-doped Zinc Oxide Buffer Layer. *Org. Electron.* 2011, *12*, 1539–1543.
[45] Steim, R.; Choulis, S. A.; Schilinsky, P.; Brabec, C. J. Interface Modification for Highly Efficient Organic Photovoltaics. *Appl. Phys. Lett.* 2008, *92*, 093303/1-3.
[46] Schmidt, H.; Zilberberg, K.; Schmale, S.; Flügge, H.; Riedl, T.; Kowalsky, W. Transient Characteristics of Inverted Polymer Solar Cells Using Titanium Oxide Interlayers. *Appl. Phys. Lett.* 2010, *96*, 243305/1-3.
[47] Mor, G. K.; Shankar, K.; Paulose, M.; Varghese, O. K.; Grimes, C. A. High Efficiency Double Heterojunction Polymer Photovoltaic Cells Using Highly Ordered TiO_2 Nanotube Arrays. *Appl. Phys. Lett.* 2007, *91*, 152111/1-3.
[48] Schumann, S.; Campo, R. D.; Illy, B.; Cruickshank, A. C.; McLachlan, M. A.; Ryan, M. P.; Riley, D. J.; McComb, D. W.; Jones, T. S. Inverted Organic Photovoltaic Devices with High Efficiency and Stability Based on Metal Oxide Charge Extraction Layers. *J. Mater. Chem.* 2011, *21*, 2381–2386.
[49] Li, C. Y.; Wen, T. C.; Lee, T. H.; Guo, T. F.; Huang, J. C. A.; Lind, Y. C.; Hsu, Y. J. An Inverted Polymer Photovoltaic Cell with Increased Air Stability Obtained by Employing Novel Hole/electron Collecting Layers. *J. Mater. Chem.* 2009, *19*, 1643–1647.
[50] Lim, D. C.; Shim, W. H.; Kim, K. D.; Seo, H. O.; Lim, J. H.; Jeong, Y.; Kim, Y. D.; Lee, K. H. Spontaneous Formation of Nanoripples on the Surface of ZnO Thin Films as Hole-blocking Layer of Inverted Organic Solar Cells. *Sol. Energy Mater. Sol. Cells*. 2011, *95*, 3036–3040.
[51] Wang, J. C.; Lu, C. Y.; Hsu, J. L.; Lee, M. K.; Hong, Y. R.; Perng, T. P.; Horng, S. F.; Meng, H. F. Efficient Inverted Organic Solar Cells without an Electron Selective Layer. *J. Mater. Chem.* 2011, *21*, 5723–5728.
[52] Kim, J.; Guan, Z.; Shu, A. L.; Kahn, A.; Loo, Y. L. Annealing Sequence Dependent Open-circuit Voltage of Inverted Polymer Solar Cells Attributable to Interfacial Chemical Reaction between Top Electrodes and Photoactive Layers. *Langmuir*. 2011, *27*, 11265–11271.
[53] Kawano, K.; Ito, N.; Nishimori, T.; Sakai, J. Open Circuit Voltage of Stacked Bulk Heterojunction Organic Solar Cells. *Appl. Phys. Lett.* 2006, *88*, 073514/1-3.
[54] Hadipour, A.; de Boer, B.; Wildeman, J.; Kooistra, F. B.; Hummelen, J. C.; Turbiez, M. G. R.; Wienk, M. M.; Janssen, R. A. J.; Blom, P. W. M. Solution-processed Organic Tandem Solar Cells. *Adv. Funct. Mater.* 2006, *16*, 1897–1903.
[55] Kim, J. Y.; Lee, K.; Coates, N. E.; Moses, D.; Nguyen, T. Q.; Dante, M.; Heeger, A. J. Efficient Tandem Polymer Solar Cells Fabricated by All-solution Processing. *Science*. 2007, *317*, 222–225.
[56] Sista, S.; Hong, Z.; Chen, L. M.; Yang, Y. Tandem Polymer Photovoltaic Cells—current Status, Challenges and Future Outlook. *Energy Environ. Sci.* 2011. *4*, 1606–20,

[57] Shi, Z.; Bai, Y.; Chen, X.; Rui Zeng, R.; Tan, Z. Tandem Structure: A Breakthrough in Power Conversion Efficiency for Highly Efficient Polymer Solar Cells. *Sustainable Energy Fuels*. 2019, *3*, 910–934.
[58] Sakai, J.; Kawano, K.; Yamanari, T.; Taima, T.; Yoshida, Y.; Fujii, A.; Ozaki, M. Efficient Organic Photovoltaic Tandem Cells with Novel Transparent Conductive Oxide Interlayer and Poly(3--hexylthiophene): Fullereneactive Layers. *Sol. Energy Mater. Sol. Cells*. 2010, *94*, 376–380.
[59] Larsen-Olsen, T. T.; Bundgaard, E.; Sylvester-Hvid, K. O.; Krebs, F. C. A Solution Process for Inverted Tandem Solar Cells. *Org. Electron.* 2011, *12*, 364–371.
[60] Tung, V. C.; Kim, J.; Cote, L. J.; Huang, J. Sticky Interconnect for Solution-processed Tandem Solar Cells. *J. Am. Chem. Soc.* 2011, *133*, 9262–9265.
[61] Chen, D.; Wang, D.; Yang, Y.; Huang, Q.; Zhu, S.; Zheng, Z. Self Healing Materials for Next-generation Energy Harvesting and Storage Devices. *Adv. Energy Mater.* 2017, *7*, 1700890/1-23.
[62] Banerjee, S.; Tripathy, R.; Cozzens, D.; Nagy, T.; Keki, S.; Zsuga, M.; Faust, R. Photoinduced Smart, Self-healing Polymer Sealant for Photovoltaics. *ACS Appl. Mater. Interfaces*. 2015, *7*, 2064–2072.
[63] Conway, B. E.;. *Electrochemical Supercapacitors-scientific Fundamentals and Technological Applications*; Plenum Press, Springer: New York, USA, 1999.
[64] Salunkhe, R. R.; Lin, J. J.; Malgras, V.; Dou, S. X.; Kim, J. H.; Yamauchi, Y. Large-scale Synthesis of Coaxial Carbon nanotube/$Ni(OH)_2$ Composites for Asymmetric Supercapacitor Application. *Nano Energy*. 2015, *11*, 211–218.
[65] Tang, J.; Salunkhe, R. R.; Liu, J.; Torad, N. L.; Imura, M.; Furukawa, S.; Yamauchi, Y. Thermal Conversion of Core-shell Metal-organic Frameworks: A New Method for Selectively Functionalized Nanoporous Hybrid Carbon. *J. Am. Chem. Soc.* 2015, *137*, 1572–1580.
[66] Zhou, C.; Zhang, Y. W.; Li, Y. Y.; Liu, J. P. Construction of High-Capacitance 3D CoO@Polypyrrole Nanowire Array Electrode for Aqueous Asymmetric Supercapacitor. *Nano Lett.* 2013, *13*, 2078–2085.
[67] Wu, Q.; Xu, Y. X.; Yao, Z. Y.; Liu, A. R.; Shi, G. Q. Supercapacitors Based on Flexible Graphene/Polyaniline Nanofiber Composite Films. *ACS Nano*. 2010, *4*, 1963–1970.
[68] Zhang, K.; Zhang, L. L.; Zhao, X. S.; Wu, J. Graphene/Polyaniline Nanofiber Composites as Supercapacitor Electrodes. *Chem. Mater.* 2010, *22*, 1392–1401.
[69] Lu, Q.; Zhou, Y. K. Synthesis of Mesoporous polythiophene/MnO_2nanocomposite and Its Enhanced Pseudocapacitive Properties. *J. Power Sources*. 2011, *196*, 4088–4094.
[70] Snook, G. A.; Kao, P.; Best, A. S. Conducting-polymer-based Supercapacitor Devices and Electrodes. *J. Power Sources*. 2011, *196*, 1–12.
[71] Cho, S.; Kim, M.; Jang, J. Screen-Printable and flexible RuO_2 Nanoparticle-decorated PEDOT: PSS/graphene Nanocomposite with Enhanced Electrical and Electrochemical Performances for High-capacity Supercapacitor, ACS Appl. *Mater. Interfaces*. 2015, *7*, 10213–10227.
[72] Xia, X.; Chao, D.; Fan, Z.; Guan, C.; Cao, X.; Zhang, H.; Fan, H. J. A New Type of Porous Graphite Foams and Their Integrated Composites with Oxide/polymer Core/shell Nanowires for Supercapacitors: Structural Design, Fabrication, and Full Supercapacitor Demonstrations. *Nano Lett.* 2014, *14*, 1651–1658.
[73] Mia, H.; Zhang, X.; Ye, X.; Yang, S. Preparation and Enhanced Capacitance of Core-shell Polypyrrole/polyaniline Composite Electrode for Supercapacitors. *J. Power Sources*. 2008, *176*, 403–409.
[74] Hashmi, S. A.; Kumar, A.; Tripathi, S. K. Investigations on Electrochemical Supercapacitorsusing Polypyrrole Redox Electrodes and PMMA Basedgel Electrolytes. *Eur. Polym. J.* 2005, *41*, 1373–1379.
[75] Aradilla, D.; Estrany, F.; Armelin, E.; Alemán, C. Ultraporous Poly(3,4-ethylenedioxythiophene) for Nanometric Electrochemical Supercapacitor. *Thin Solid Films*. 2012, *520*, 4402–4409.
[76] Tana, Y. T.; Ran, F.; Kong, L. B.; Liu, J.; Kang, L. Polyaniline Nanoparticles Grown on the Surface of Carbon Microspheres Aggregations for Electrochemical Supercapacitors. *Synth. Met.* 2012, *162*, 114–118.
[77] Wang, Y.; Li, H.; Xia, Y. Ordered Whisker like Polyaniline Grown on the Surface of Mesoporous Carbon and Its Electrochemical Capacitance Performance. *Adv. Mater.* 2006, *18*, 2619–2623.
[78] Li, L.; Song, H.; Zhang, Q.; Yao, J.; Chen, X. Effect of Compounding Process on the Structure and Electrochemical Properties of Ordered Mesoporous Carbon/polyaniline Composites as Electrodes for Supercapacitors. *J. Power Sources*. 2009, *187*, 268–274.
[79] Zheng, L.; Wang, Y.; Wang, X.; Li, N.; An, H.; Chen, H.; Guo, J. The Preparation and Performance of Calcium Carbide-derived Carbon/polyaniline Composite Electrode Material for supercapacitors. *J Power Sources*. 2010, *195*, 1747–1752
[80] Lei, C.; Wilson, P.; Lekakou, C. Effect of Poly (3, 4-ethylenedioxythiophene) (PEDOT) in Carbon-based Composite Electrodes for Electrochemical Supercapacitors. *J. Power Sources*. 2011, *196*, 7823–7827.
[81] Frackowiak, E.; Khomenko, V.; Jurewicz, K.; Lota, K.; Beguin, F. Supercapacitors Based on Conducting Polymers/nanotubes Composites. *J. Power Sources*. 2006, *153*, 413–418.
[82] Gómez, H.; Ram, M. K.; Alvi, F.; Villalba, P.; Stefanakos, E.; Kumar, A. Graphene-conducting Polymer Nanocomposite as Novel Electrode for Supercapacitors. *J. Power Sources*. 2011, *196*, 4102–4108.
[83] Li, X.; Song, Q.; Hao, L.; Zhi, L. Graphenal Polymers for Energy Storage. *Small*. 2014, *10*, 2122–2135.
[84] Chin, S. F.; Pang, S. C. Tetrapropylammonium-manganese Oxide/polypyrrole Hybrid Nanocomposite Thin Films as Novel Electrode Materials for Supercapacitors. *Mater. Chem. Phys.* 2010, *124*, 29–32.
[85] Li, X.; Gan, W.; Zheng, F.; Li, L.; Zhu, N.; Huang, X. Preparation and Electrochemical Properties of RuO_2 /polyaniline Electrodes for Supercapacitors. *Synth. Met.* 2012, *162*, 953–957.

[86] Sawangphruk, M.; Kaewsongpol, T. Direct Electrodeposition and Superior Pseudocapacitive Property of Ultrahigh Porous Silver-incorporated Polyaniline Films. *Mater. Lett.* 2012, *87*, 142–145.
[87] Su, H.; Wang, T.; Zhang, S.; Song, J.; Mao, C.; Niu, H.; Jin, B.; Wu, J.; Tian, Y. Facile Synthesis of polyaniline/TiO_2 /graphene Oxide Composite for High Performance Supercapacitors. *Solid State Sci.* 2012, *14*, 677–681.
[88] Mia, H.; Zhang, X.; Yang, S.; Ye, X.; Luo, J. Polyaniline Nanofibers as the Electrode Material for Supercapacitors. *Mater. Chem. Phys.* 2008, *112*, 127–131.
[89] Zhao, G. Y.; Li, H. L. Preparation of Polyaniline Nanowire Arrayed Electrodes for Electrochemical Supercapacitors. *Microporous Mesoporous Mater.* 2008, *110*, 590–594.
[90] Güllü, M.; Yigit, D. A Novel Asymmetric Pseudocapacitor Based on Poly(5,12-dihydrothieno [3′,4′:2,3][1,4]dioxocino[6,7-b]quinoxaline) Coated Graphite Anode and Poly(ethylenedioxythiophene) Coated Graphite Cathode. *Synth. Met.* 2012, *162*, 1434–1442.
[91] Ghenaatian, H. R.; Mousavi, M. F.; Rahmanifar, M. S. High Performance Hybrid Supercapacitor Based on Two Nanostructuredconducting Polymers: Self-doped Polyaniline and Polypyrrole Nanofibers. *Electrochim. Acta*. 2012, *78*, 212–222.
[92] Kim, B. C.; Too, C. O.; Kwon, J. S.; Ko, J. M.; Wallace, G. G. A Flexible Capacitor Based on Conducting Polymer Electrodes. *Synth. Met.* 2011, *161*, 1130–1132.
[93] Wang, L.; Feng, X.; Ren, L.; Piao, Q.; Zhong, J.; Wang, Y.; Li, H.; Chen, Y.; Wang, B. Flexible Solid-state Supercapacitor Based on a Metal–organic Framework Interwoven by Electrochemically-deposited PANI. *J. Am. Chem. Soc.* 2015, *137*, 4920–4923.
[94] Sivakkumar, S. R.; Oh, J. S.; Kim, D. W. Polyaniline Nanofibres as a Cathode Material for Rechargeable Lithium-polymer Cells Assembled with Gel Polymer Electrolyte. *J. Power Sources*. 2006, *163*, 573–577.
[95] Wang, J.; Too, C. O.; Zhou, D.; Wallace, G. G. Novel Electrode Substrates for Rechargeable Lithium/polypyrrole Batteries. *J. Power Sources*. 2005, *140*, 162–167.
[96] Liu, L.; Tian, F.; Wang, X.; Yang, Z.; Zhou, M.; Wang, X. Porous Polythiophene as a Cathode Material for Lithium Batteries with High Capacity and Good Cycling Stability. *React. Funct. Polym.* 2012, *72*, 45–49. DOI: 10.1016/j.reactfunctpolym.2011.10.006.
[97] Tang, J.; Kong, L.; Zhang, J.; Zhan, L.; Zhan, H.; Zhou, Y.; Zhan, C. Solvent-free, Oxidatively Prepared Polythiophene: High Specific Capacity as a Cathode Active Material for Lithium Batteries. *React. Funct. Polym.* 2008, *68*, 1408–1413. DOI: 10.1016/j.reactfunctpolym.2008.07.001.
[98] Murugan, A. V.;. Novel Organic–inorganic Poly (3, 4-ethylenedioxythiophene)based Nanohybrid Materials for Rechargeable Lithium Batteries and Supercapacitors. *J. Power Sources*. 2006, *159*, 312–318. DOI: 10.1016/j.jpowsour.2006.04.033.
[99] de Azevedo, W. M.; de Mattos, I. L.; Navarro, M.; EF, D. S. Preparation and Characterization of Conducting Polymer/silver Hexacyanoferrate Nanocomposite. *Appl. Surf. Sci.* 2008, *255*, 770–774.
[100] Karami, H.; Mousavi, M. F.; Shamsipur, M. A New Design for Dry Polyaniline Rechargeable batteries.J. *Power Sources*. 2003, *117*, 255–259. DOI: 10.1016/S0378-7753(03)00168-X.
[101] Gurunathan, K.; Amalnerkar, D. P.; Trivedi, D. C. Synthesis and Characterization of Conducting Polymer Composite (Pan/tio_2) for Cathode Material in Rechargeable Battery. *Mater. Lett.* 2003, *57*, 1642–1648. DOI: 10.1016/S0167-577X(02)01045-5.
[102] Cui, L.; Shen, J.; Cheng, F.; Tao, Z.; Chen, J. SnO_2 Nanoparticles@ Polypyrrole Nanowires Composite as Anode Materials for Rechargeable Lithium-ion Batteries. *J. Power Sources*. 2011, *196*, 2195–2201.
[103] Murugan, A. V.; Quintin, M.; Delville, M. H.; Campet, G.; Gopinath, C. S.; Vijayamohanan, K. Exfoliation-induced Nanoribbon Formation of Poly (3,4-ethylenedioxythiophene) PEDOT between MoS2 Layers as Cathode Material for Lithium Batteries. *J. Power Sources*. 2006, *156*, 615–619. DOI: 10.1016/j.jpowsour.2005.06.022.
[104] Tang, J.; Song, Z. P.; Shan, N.; Zhan, L. Z.; Zhang, J. Y.; Zhan, H.; Zhou, Y. H.; Zhan, C. M. Poly[3,4-(ethylene-dithio)thiophene]: High Specific Capacity Cathode Activematerial for Lithium Rechargeable Batteries. *J. Power Sources*. 2008, *185*, 1434–1438. DOI: 10.1016/j.jpowsour.2008.07.079.
[105] Mirmohseni, A.; Solhjo, R. Preparation and Characterization of Aqueous Polyaniline Battery Using a Modified Polyaniline Electrode. *Eur. Polym. J.* 2003, *39*, 219–223.
[106] Rehan, H. H.;. A New Polymer/polymer Rechargeable Battery: Polyaniline/$LiClO_4$ (Mecn)/poly-1-naphthol. *J. Power Sources*. 2003, *113*, 57–61. DOI: 10.1016/S0378-7753(02)00485-8.
[107] Wang, C. Y.; Ballantyne, A.; Hall, S. B.; Too, C. O.; Officer, D.; GG, W. Functionalised Polythiophene Coated Textile: A New Anode Material for A Flexible Battery. *J. Power Sources*. 2006, *156*, 610–614. DOI: 10.1016/j.jpowsour.2005.06.020.
[108] Wang, C. Y.; Tsekouras, G.; Wagner, P.; Gambhir, S.; Too, C. O.; Officer, D.; GG, W. Functionalised Polyterthiophenes as Anode Materials in polymer/polymerBatteries. *Synth. Met.* 2010, *160*, 76–82. DOI: 10.1016/j.synthmet.2009.10.001.
[109] Aydin, M.; Esat, B.; Kilic, C.; Kose, M. E.; Ata, A.; Yılmaz, F. A Polythiophene Derivative Bearing TEMPO as A Cathode Material for Rechargeable Batteries. *Eur. Polym. J.* 2011, *47*, 2283–2294. DOI: 10.1016/j.eurpolymj.2011.09.002.
[110] Nakahara, K.; Iwasa, S.; Satoh, M.; Morioka, Y.; Iriyama, J.; Suguro, M.; Hasegawa.Rechargeable, E. Batteries with Organic Radical cathodes. *Chem Phys Lett*. 2002, *359*, 351–354.
[111] Nystrom, G.; Razaq, A.; Strømme, M.; Nyholm, L.; Mihranyan, A. Ultrafast All-polymer Paper-based Batteries. *Nano Lett.* 2009, *9*, 3635–3639. DOI: 10.1021/nl901852h.
[112] Ji, L.; Yao, Y.; Toprakci, O.; Lin, Z.; Liang, Y.; Shi, Q.; Medford, A. J.; Millns, C. R.; Zhang, X. Fabrication of Carbon Nanofiber-driven Electrodes from Electrospun Polyacrylonitrile/polypyrrole Bicomponents for High-performance Rechargeable Lithium-ion Batteries.

J. Power Sources. 2010, *195*, 2050–2056. DOI: 10.1016/j.jpowsour.2009.10.021.

[113] Baibarac, M.; Lira-Cantu, M.; Sol, J. O.; Baltog, I.; Casan-Pastor, N.; Gomez-Romero, P. Poly(N-vinyl Carbazole) and Carbon Nanotubes Based Composites and Their Application to Rechargeable Lithium Batteries. *Compos. Sci. Technol.* 2007, *67*, 2556–2563. DOI: 10.1016/j.compscitech.2006.12.008.

[114] He, B. L.; Dong, B.; Wang, W.; HL, L. Performance of Polyaniline/multi-walled Carbon Nanotubes Composites as Cathode for Rechargeable Lithium Batteries. *Mater. Chem. Phys.* 2009, *114*, 371–375. DOI: 10.1016/j.matchemphys.2008.09.035.

[115] Wang, J.; Too, C. O.; Wallace, G. G. A Highly Flexible Polymer Fibre Battery. *J. Power Sources*. 2005, *150*, 223–228. DOI: 10.1016/j.jpowsour.2005.01.046.

[116] Wang, J.; Wang, C. Y.; Too, C. O.; Wallace, G. G. Highly-flexible Fibre Battery Incorporating Polypyrrole Cathode and Carbon Nanotubes Anode. *J. Power Sources*. 2006, *161*, 1458–1462. DOI: 10.1016/j.jpowsour.2006.05.038.

[117] Chao, Z. S.; Lan, Z.; Jinhua, Y. Preparation and Electrochemical Properties of Polysulfide polypyrrole.J. *Power Sources*. 2011, *196*, 10263–10266. DOI: 10.1016/j.jpowsour.2011.08.092.

[118] Sun, M. M.; Zhang, S. C.; Jiang, T.; Zhang, L.; Yu, J. H. Nano-wire Networks of Sulfur-polypyrrole Composite Cathode Materials Rechargeable Lithium Batteries. *Electrochem. Commun.* 2008, *10*, 1819–1822.

[119] Liang, X.; Liu, Y.; Wen, Z. Y.; Huang, L. Z.; Wang, X. Y.; Zhang, H. A Nano-structured and Highly Ordered Polypyrrole-sulfur Cathode for Lithium–sulfur Batteries. *J. Power Sources*. 2011, *196*, 6951–6955. DOI: 10.1016/j.jpowsour.2010.11.132.

[120] Zhang, Y.; Bakenov, Z.; Zhao, Y.; Konarov, A.; Doan, T. N. L.; Malik, M.; Paron, T.; Chen, P. One-step Synthesis of Branched Sulfur/polypyrrole Nanocomposite Cathode for Lithium Rechargeable Batteries. *J. Power Sources*. 2012, *208*, 1–8. DOI: 10.1016/j.jpowsour.2012.02.006.

[121] Kong, Y.; Wang, C.; Yang, Y.; Too, C. O.; Wallace, G. G. A Battery Composed of A Polypyrrole Cathode and A Magnesium Alloy anode-Toward A Bioelectric Battery. *Synth. Met.* 2012, *162*, 584–589. DOI: 10.1016/j.synthmet.2012.01.021.

[122] Wang, C.; Zheng, W.; Yue, Z.; Too, C. O.; Wallace, G. G. Buckled, Stretchable Polypyrrole Electrodes for Battery Applications. *Adv. Mater.* 2011, *23*, 3580–3584. DOI: 10.1002/adma.201101067.

[123] Arvay, A.; Yli-Rantala, E.; Liu, C.-H.; Peng, X.-H.; Koski, P.; Cindrella, L.; Kauranen, P.; Wilde, P.M.; Kannan, A. M. Characterization Techniques for Gas Diffusion Layers for Proton Exchange Membrane Fuel Cells E A Review. *J. Power Sources*. 2012, *213*, 317–337.

[124] Garraín, D.; Lechón, Y.; de la Rúa, C. Polymer Electrolyte Membrane Fuel Cells (PEMFC) in Automotive Applications: Environmental Relevance of the Manufacturing Stage. Smart Grid Renewable Energy. 2011, *2*, 68–74. DOI: 10.4236/sgre.2011.22009.

[125] Tang, H.; Peikang, S.; Jiang, S. P.; Wang, F.; Pan, M. A Degradation Study of Nafion Proton Exchange Membrane of PEM Fuel Cells. *J. Power Sources*. 2007, *170*(1), 85–92. DOI: 10.1016/j.jpowsour.2007.03.061.

[126] Kim, D. J.; Jo, M. J.; Nam, S. Y. A Review of Polymer–nanocomposite Electrolyte Membranes for Fuel Cell Application. *J. Ind. Eng. Chem.* 2015, *21*, 36–52. DOI: 10.1016/j.jiec.2014.04.030.

[127] Tohidian, M.; Ghaffarian, S. R.; Nouri, M.; Jaafarnia, E.; Haghighi, A. H. Polyelectrolyte Nanocomposite Membranes Using Imidazole-functionalized Nanosilica for Fuel Cell Applications. *J. Macromol. Sci. Part B Phys.* 2015, *54*(1), 17–31. DOI: 10.1080/00222348.2014.982485.

[128] Luo, Z.; Gong, Y.; Liao, X.; Pan, Y.; Zhang, H. Nanocomposite Membranes Modified by Graphene-based Materials for Anion Exchange Membrane Fuel Cells. *RSC Adv.* 2016, *6*(17), 13618–13625. DOI: 10.1039/C5RA21104B.

[129] Adjemian, K. T.; Dominey, R.; Krishnan, L.; Ota, H.; Majsztrik, P.; Zhang, T.; Mann, J.; Kirby, B.; Gatto, L.; Velo-Simpson, M.;; et al. Function and Characterization of Metal Oxide–Nafion Composite Membranes for Elevated-Temperature H2/O2PEM Fuel Cells. *Chem. Mater.* 2006, *18*(9), 2238–2248.

[130] Jiang, R.; Kunz, H. R.; Fenton, J. M. Composite silica/Nafion® Membranes Prepared by Tetraethylorthosilicate Sol–gel Reaction and Solution Casting for Direct Methanol Fuel Cells. *J. Membr. Sci.* 2006, *272*(1), 116–124. DOI: 10.1016/j.memsci.2005.07.026.

[131] Sood, R.; Iojoiu, C.; Espuche, E.; Gouanve, F.; Gebel, G.; Mendil-Jakani, H.; Lyonnard, S.; Jestin, J. Proton Conducting Ionic Liquid Doped Nafion Membranes: Nano-structuration, Transport Properties and Water Sorption. *J. Phys. Chem. C*. 2012, *116*(46), 24413–24423.

[132] Ramani, V.; Kunz, H. R.; Fenton, J. M. Metal Dioxide Supported heteropolyacid/Nafion® Composite Membranes for Elevated Temperature/low Relative Humidity PEFC Operation. *J. Membr. Sci.* 2006, *279*(1–2), 506–512. DOI: 10.1016/j.memsci.2005.12.044.

[133] Sahu, A. K.; Ketpang, K.; Shanmugam, S.; Kwon, O.; Lee, S.; Kim, H. Sulfonated graphene–Nafion Composite Membranes for Polymer Electrolyte Fuel Cells Operating under Reduced Relative Humidity. *J. Phys. Chem. C*. 2016, *120*(29), 15855–15866. DOI: 10.1021/acs.jpcc.5b11674.

[134] Zarrin, H.; Higgins, D.; Jun, Y.; Chen, Z.; Fowler, M. Functionalized Graphene Oxide Nanocomposite Membrane for Low Humidity and High Temperature Proton Exchange Membrane Fuel Cells. *J. Phys. Chem. C*. 2011, *115*(42), 20774–20781. DOI: 10.1021/jp204610j.

[135] Liu, Y.-L.; Su, Y.-H.; Chang, C.-M.; Suryani, W. D.-M.; J-Y, L. Preparation and Applications of Nafion-functionalized Multiwalled Carbon Nanotubes for Proton Exchange Membrane Fuel Cells. *J. Mater. Chem.* 2010, *20*(21), 4409–4416.

[136] Cele, N.; Ray, S. S. Recent Progress on Nafion-based Nanocomposite Membranes for Fuel Cell Applications. *Macromol. Mater. Eng.* 2009, *294*(11), 719–738. DOI: 10.1002/mame.200900143.

[137] Sharaf, O. Z.; Orhan, M. F. An Overview of Fuel Cell Technology: Fundamentals and Applications. *Renew. Sust. Energ Rev.* 2014, *32*, 810–853. DOI: 10.1016/j.rser.2014.01.012.

AIP Conference Proceedings

The effect of variable electrolyte thickness of LMO vs Li metal foil lithium-ion coin cell using simulation

Cite as: AIP Conference Proceedings **2281**, 020025 (2020); https://doi.org/10.1063/5.0026247
Published Online: 15 October 2020

M. Aditya, Naveen Kumar, Chetan Kumar, and Amrish K. Panwar

View Online Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Multi-body dynamic optimization for upper arm of industrial manipulator
AIP Conference Proceedings **2281**, 020022 (2020); https://doi.org/10.1063/5.0027965

On work of external forces at complex loading
AIP Conference Proceedings **2281**, 020036 (2020); https://doi.org/10.1063/5.0026231

Structural health monitoring using smart sensors
AIP Conference Proceedings **2281**, 020034 (2020); https://doi.org/10.1063/5.0026292

© 2020 Author(s).

The Effect of Variable Electrolyte Thickness of LMO vs Li Metal Foil Lithium-Ion Coin Cell using Simulation

M Aditya [a], Naveen Kumar [b], Chetan Kumar [c], and Amrish K. Panwar [d]

Department of Applied Physics, Delhi Technological University, India.
(Formerly Delhi College of Engineering)

[a] adityam98@gmail.com
[b] naveenkumar_bt2k16@dtu.ac.in
[c] chetan_bt2k16@dtu.ac.in
[d] amrish.phy@dtu.ac.in

Abstract. In current times, Lithium-ion batteries have become an important part of a plethora of devices as a power source with good energy and power densities. The new concern now is to have a battery that has an efficient charge-discharge cycle. In this work simulation of a 2-D cross-section of a coin-cell using a Poly(vinylidene fluoride-co-hexafluoropropylene) p(VdF-HFP) as the polymer soaked in $LiPF_6$ in an in 1:2 EC: DMC as the electrolyte. Lithium manganese oxide, $LiMn_2O_4$ Spinel (LMO) with a thickness of 0.6 mm is used as a cathode, and Lithium metal is used as an anode. For this simulation study, COMSOL Multiphysics has been used. In this study, the thickness of the $LiPF_6$ (1:2 EC: DMC) soaked p(VdF-HFP) polymer electrolyte has been varied from 0.1 mm to 1.0 mm. At these variable thicknesses: battery voltage, lithium-ion concentration on the particle surface, and Electrolyte Potential have been simulated. It is seen that as the thickness of the polymer electrolyte goes on increasing, the discharge lifetime of the cell decreases.

Keywords. Li-ion battery, Polymer Electrolyte, LiPF6, LMO, Li-anode

INTRODUCTION

Energy plays a significant role in all facets of our lives. Energy demand won't wane as long as there's a desire for a better living standard. We rely heavily on conventional fossil fuels like, coal , natural gas and nuclear energy to cater to the increasing demand for energy.[1] Limitations in supply, exploitation of natural resources and long-term harm to the atmosphere caused by a dramatic increase in carbon dioxide levels have brought research into renewable energy sources to the forefront.[1]. The major hindrance for the practicality of renewable energy use involves the development of efficient energy storage devices for industrial systems and consumers. The simplest way to save energy is by electrochemical storage such as batteries. The latest Lithium-ion based rechargeable battery technology has few limitations which need to be balanced against the advantages [1]. The key drawback with the Lithium-ion battery is safety. Cells require additional control systems to prevent such conditions. Short-circuiting a Lithium-ion battery results in fire and overheating accidents and even death. To solve this problem, advances need to be made in Li-ion battery electrolytes, with polymer electrolytes being the prime candidate in current times. Polymer electrolytes refer to polymer-salt complexes, normally formed between the lithium-ion-conducting salts and a polymer that provides structural integrity [2]. The electrolyte plays two roles in polymer-based batteries which are, as a Lithium-ion carrier and an electrode spacer, thus eliminating need for an inert porous separator. They may be formed into a thin film to enhance the battery energy density [1]. Excellent mechanical properties, which is an advantageous behaviour of polymer electrolytes, helps withstand the electrode volume change during the charging-discharging process, but also makes possible to assemble the battery using thin films [3] .While additional properties like resistance to redox shuttling and prevention of lithium dendrites are also desired, providing all of these properties into a polymer electrolyte is a daunting task

Third International Conference on Inventive Material Science Applications
AIP Conf. Proc. 2281, 020025-1–020025-7; https://doi.org/10.1063/5.0026247
Published by AIP Publishing. 978-0-7354-4011-1/$30.00

[3]. Given their relative stability and high energy density, the importance of Polymer electrolytes in the next generation of lithium-ion batteries will be significant. The fundamental properties of different polymer electrolytes such as ion-transport mechanisms and preparation techniques play a major role in determining their feasibility [2]. In this work, Battery Voltage, Lithium concentration on the particle surface, and Electrolyte Potential of Poly(vinylidene fluoride-co-hexafluoropropylene) p(VdF-HFP) + $LiPF_6$ (1:2 EC:DMC) polymer electrolyte Lithium-ion based batteries has been studied for LMO vs Li metal as cathode and anode materials, respectively using COMSOL Multiphysics on varying the thickness of the electrolyte.

METHODOLOGY AND THEORETICAL ANALYSIS

The schematics of the polymer electrolyte lithium-ion battery model and the 2-D cross section of a simulated coin cell used in this study are shown in Fig 1 and Fig 2, respectively. In this work, the structure of the cell is simulated and physical parameters of the anode, cathode and polymer electrolyte are assigned and varied as shown in Fig 1. The anode is Lithium metal and the cathode is constructed with a 0.6 mm film of LMO (LiMn2O4 spinel). The electrolyte is a polymer film Poly(vinylidene fluoride-co-hexafluoropropylene) (p(VdF-HFP)) with 1000 mole/m3 LiPF6 in 1:2 EC: DMC.The electro-chemical reactions at the anode and cathode are represented by [4].

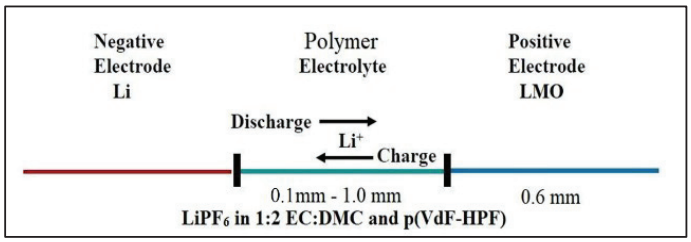


FIGURE 1. Schematic of the polymer electrolyte lithium-ion battery model.

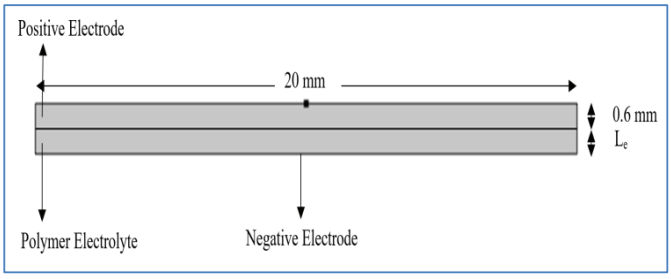


FIGURE 2. 2-D cross-section of the simulated coin cell.

$$Li \rightarrow Li^{+} + e^{-} \quad (1)$$

$$LiMn_2O_4 \rightarrow Li_{1-x}Mn_2O_4 + xLi^{+} + xe^{-} \quad (2)$$

where the reactions shown in eqs. (1) and (2) are described using the Butler-Volmer kinetics shown in eqs (3) to eqs (5) [5][6].

$$i_n = Fk_n\left(\frac{c_{Li+}}{c_{Li+,0}}\right)^{\alpha_n}\left(e^{\frac{\alpha_n F\eta}{RT}} - e^{\frac{-(1-\alpha_n)F\eta}{RT}}\right) \quad (3)$$

$$i_p = i_{0,p}\left(e^{\frac{\alpha_p F\eta}{RT}} - e^{\frac{-(1-\alpha_p)F\eta}{RT}}\right) \quad (4)$$

$$i_{0,p} = Fk_p\left(\left(\frac{(c_{Li,max}-c_{Li})c_{Li+}}{(c_{Li,max}-c_{Li,min})c_{Li+,0}}\right)^{\alpha_p} - \left(\frac{(c_{Li}-c_{Li,max})}{(c_{Li,max}-c_{Li,min})c_{Li+,min}}\right)^{1-\alpha_p}\right) \quad (5)$$

Where F is the charge on 1 mole of e^-, k_n and k_p are the reaction rate-constants at anode and cathode, $c_{Li+,0}$ is the initial concentration of Lithium-ions in the polymer electrolyte, α_n and α_p are the coefficients of charge transfer, $c_{Li,\,max.}$ and $c_{Li,\,min.}$ are the highest and lowest concentrations of lithium ions in the cathode [5][7][8]. During the charging process, the reactions in LMO cathode occur and the produced Lithium ions move to the anode, Lithium metal foil. Likewise, for the discharge process, the counter-reactions at the surface of LMO occur and Lithium-ions obtained from the anode come to the LMO cathode and are consumed at the LMO cathode [9][10]. The cathode is made of cubic LMO Spinel which consists of trivalent and tetravalent Manganese ion species with an average oxidation state of +3.5 with both Mn^{+3} and Mn^{+4} existing in the spinel in equal concentration [4][9]. During charging, more Mn^{+4} is formed but the Lithium stays in the $LiMn_2O_4$ lattice, thus no MnO_2 is formed [4]. Cubic LMO spinel's have good chemical stability, affordability, are safe, nontoxic and eco-friendly. The process of Lithium-ion intercalation and de-intercalation is easily adjusted in the spatial configuration of the LMO and is impermeable to the solvent. During cycling, volumetric change is less than that of layered structures, but LMO does not provide good cyclability and stability at high temperatures due to disproportionation of the Mn^{+3} ion into the Mn^{+2} and Mn^{+4} oxidation states as shown in the reaction (6) [11].

$$2Mn^{3+}(s) \rightarrow Mn^{2+}(l) + Mn^{4+}(s) \qquad (6)$$

The standard diffusion equation is used explain the transport of lithium-ions inside the cathode, which is modelled as a porous electrode. Fick law in equation (7), here explains the movement of lithium species in the cathode [5] [12].

$$N_{Li} = -D_{Li}\nabla C_{Li} \qquad (7)$$

here D_{Li} and c_{Li} are the diffusion and concentration coefficient of lithium ions in the cathode.

RESULTS AND DISCUSSION

The variation of cell voltage with time for variable thicknesses of electrolyte (0.1 mm - 1.0mm) for simulated cells is shown in the Fig 3. It is observed for a simulated cell that the voltage decreases on discharging the cell at the current density of 50 A/m^2 as it was charged with the same current density and it follows the similar discharge characteristic as usually observed for rechargeable batteries, showing polarization effects, ohmic losses and concertation losses. It can be seen that the thinner electrolytes perform better as compared to the thicker electrolytes, this can be attributed to the fact that Li ion transport is much easier through the thinner electrolytes thus leading to a higher concentration of Lithium ions on the surface of the electrolytes creating a higher voltage. It is also observed that the rate of Voltage drop in the battery is also steeper in the thicker electrolytes that can be attributed to the same reason above, as the battery behaves like it is close to discharge. It is also seen from the graph that the successive thinning of the electrolyte leads to a minimal increase in the performance of the battery. The thinning of the electrolyte of the cell also increases the probability of a short circuit exponentially due to dendrites forming on the lithium metal anode. Thus the optimal thickness of the electrolyte is found to be 0.2 mm without much loss in performance.

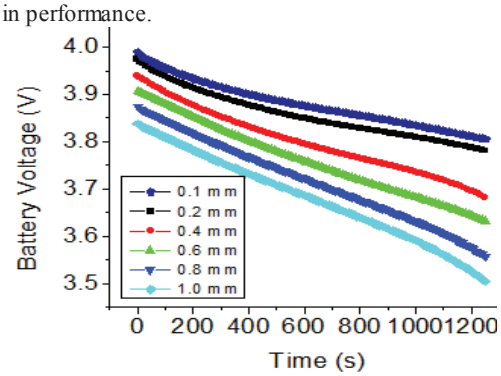


FIGURE 3. Variation of voltage with time during discharging of LMO vs Li ion coin cell of variable thickness of the electrolyte at a constant current density of 50 A/m^2.

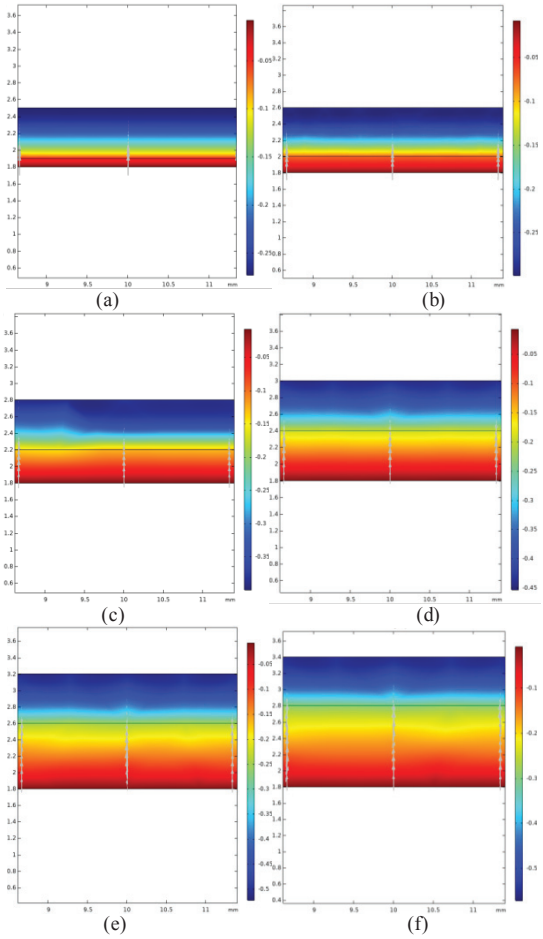


FIGURE 4. Electrolyte Potential in the cell (x-axis and y-axis are the geometric dimensions in mm) at t=1200s (a) 0.1 mm (b) 0.2 mm (c) 0.4 mm (d) 0.6 mm (e)0.8 mm (f) 1.0 mm

The variation of the simulated Electrolyte potential is given in Fig 4. It can be observed that for the same cathode configuration the Electrolyte Potential gradient (Fig 4) inside the Electrolyte is higher for L_b=0.2 mm and 0.1 mm due to the same change in the potential across a thinner electrolyte as compared to the thicker electrolytes. This can be explained as the Battery voltage does decrease with the increase in the thickness of the electrolyte, the thickness of the electrolyte changes more drastically thus causing the effects to nullify. The change in the electrolyte thickness being more than the battery voltage drop causes the net Electrolyte potential gradient to rise thus causing the Electrolyte potential gradient to be higher in the 0.1 mm and 0.2 mm thick electrolyte.Fig 5 shows the simulated Li Ion Concentration on the Surface of the Cathode. The concentration gradient of Lithium ions on the surface of the cathode is observed to be the densest on the 0.2 mm electrolyte film, this is attributed to the fact that the Electrolyte potential gradient is the highest in the 0.2 mm electrolyte film.

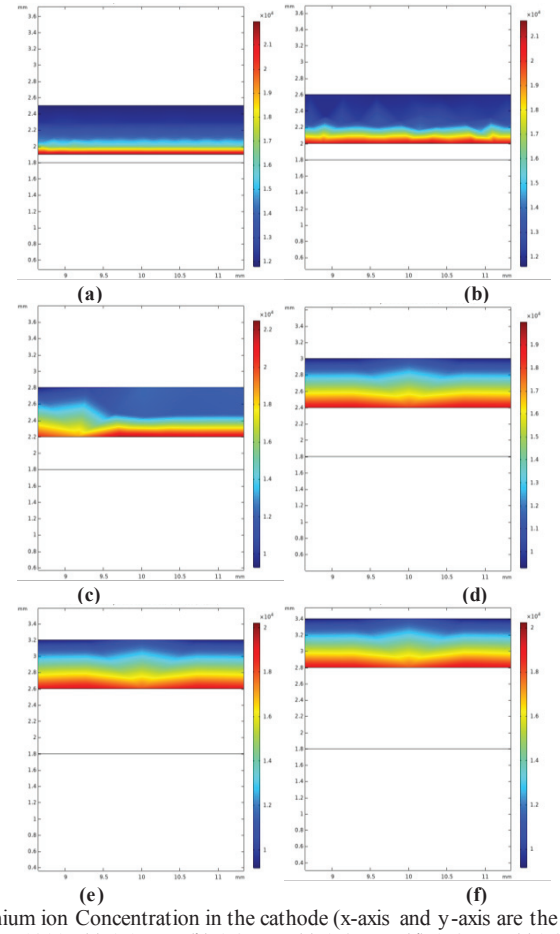


FIGURE 5. Lithium ion Concentration in the cathode (x-axis and y-axis are the geometric dimensions in mm) at t=1200s (a) 0.1 mm (b) 0.2 mm (c) 0.4 mm (d) 0.6 mm (e)0.8 mm (f) 1.0 mm

The higher Electrolyte potential gradient, thus higher potential at the interface compared to the other cells, and the fact t hat the electrode is modelled as a porous electrode acts as a compounding factor in the comparatively higher concentration gradient of the Lithium ions on the cathode surface. The $LiPF_6$ Salt Concentration Gradient in the Electrolyte and the Porous Electrode is shown in Figure 6. This gradient again can be attributed to the Electrolyte Potential Distribution inside the Cell due to the applied charging potential. The gradient in the 0.2 mm polymer film electrolyte is highest as a direct consequence of the Electrolyte Potential gradient that causes the Lithium ions to accumulate in regions toward the negative electrode.

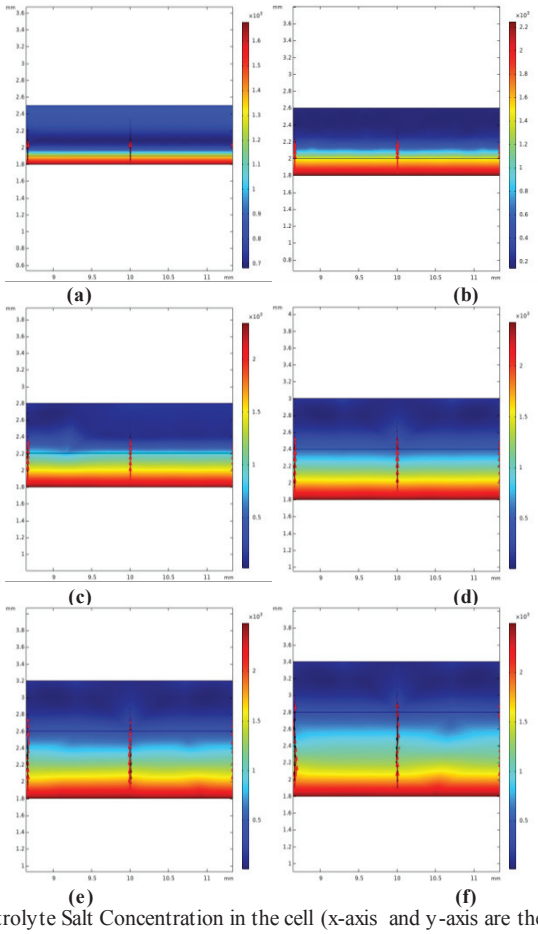


FIGURE 6. Electrolyte Salt Concentration in the cell (x-axis and y-axis are the geometric dimensions in mm) at t=1200s (a) 0.1 mm (b) 0.2 mm (c) 0.4 mm (d) 0.6 mm (e)0.8 mm (f) 1.0 mm

CONCLUSION

This work presents the simulated results of a thin-film polymer electrolyte lithium-ion coin cell using COMSOL Multiphysics. The results show that varying thicknesses of polymer electrolyte cause the different Battery Voltages, Electrolyte Potentials and Lithium Ion Concentrations. The different curves and Figures are obtained based 2D models. It is observed that thinner electrolytes lead to better performance of the cell in terms of discharge lifetime and higher battery potential. While this is of great importance thinner electrolytes also increase the probability of battery failure due to dendrite formation on the anode. Thus from the simulation the optimized thickness was found out to be 0.2 mm which provided the best relative performance. The 0.1 mm electrolyte is not feasible as the performance increase as compared to the 0.2 mm electrolyte is negligible. The presented findings would be useful for the improvement of commercial polymer lithium-ion batteries modelled by micrometre thin film polymers in the near future.

REFERENCES

[1] S. Ramakumar,C. Deviannapoorani, L. Dhivya, L. Shankar, R. Murugan. (2017). Lithium garnets: Synthesis, structure, Li+ conductivity, Li+ dynamics and applications. Progress in Materials Science. vol 88. 10.1016/j.pmatsci.2017.04.007.

[2] D. Zhou, D. Shanmukaraj, A. Tkacheva, M. Armand, G. Wang. (2019). Polymer Electrolytes for Lithium-Based Batteries: Advances and Prospects. Chem. 10.1016/j.chempr.2019.05.009.

[3] M. Liu, B. Jin, X. Zhan,F. Chen. (2018). High-performance solid polymer electrolytes for lithium ion batteries based on sulfobetaine zwitterion and poly (ethylene oxide) modified polysiloxane. Journal of Alloys and Compounds. 742. 10.1016/j.jallcom.2018.01.263.
[4] P. Kurzweil, J. Garche,(2017). Overview of batteries for future automobiles. 10.1016/B978-0-444-63700-0.00002-7.
[5] M. Fukukawa ,L. Tong (2018) 'Study of Electrochemical Impedance Spectroscopy of All-Solid-State Lithium-Ion Batteries' Keisoku Engineering System Co., Ltd.
[6] P.Singh, N. Khare, Chaturvedi. (2014). 'COMSOL Multiphysics Modelling for Li-ion Battery Ageing.' Comsol Multiphysics Conference 2014, At Banglore, Volume: 1,2014
[7] S. Chepyala. (2017). Simulation of Li-ion coin cell using COMSOL Multiphysics.
[8] J. Xie, Jos F. M. Oudenhoven, D. Li, "High power and high capacity 3D-structured TiO2 electrodes for lithium-ion micro batteries", Journal of The Electrochemical Society, 163 (10) A2385-A2389 (2016)
[9] J. P. Carmo, R. P. Rocha, A. F. Silva, L. M. Goncalves, and J. H. Correia, "A thin-film rechargeable battery for integration in standalone microsystems", Procedia Chemistry 1, 453-456 (2009)
[10] L. Tong. (2019). Three-dimensional numerical analysis of a pouch-type lithium-ion battery. The Proceedings of Mechanical Engineering Congress, Japan. 2019. J03102. 10.1299/jsmemecj.2019.J03102.
[11] M. Young, S. Letourneau, R. Warburton, W. Dose, Johnson, Christopher, J. Greeley, J. Elam. (2019). High Rate Spinel $LiMn_2O_4$ (LMO) Following Carbonate Removal and Formation of Li-Rich Interface by ALD Treatment. The Journal of Physical Chemistry C. 123. 10.1021/acs.jpcc.9b04418.
[12] A. Patil, V. Patil, D. W. Shin, J. W. Choi, D. S. Paik, and S. J. Yoon, "Issue and challenges facing rechargeable thin film lithium batteries", Materials Research Bulletin 43, 1913-1942 (2008).
[13] J. B. Bates, N. J. Dudney, B. Neudecker, A. Ueda, and C. D. Evans, "Thin-film lithium and lithium-ion batteries", Solid State Ionics 135, 33- 45 (2000).
[14] M. Kotobuki. (2019). Polymer Electrolytes. 10.1002/9783527805457.ch1.
[15] Y. Choo, D. Halat,I. Villaluenga, K. Timachova,N. Balsara. (2020). Diffusion and Migration in Polymer Electrolytes. Progress in Polymer Science. 103. 101220. 10.1016/j.progpolymsci.2020.101220.
[16] A. Thankappan, M. Stephan, S. Thomas. (2019). Thermal Characterization of Polymer Electrolytes. 10.1002/9783527805457.ch3.
[17] L. Long, S. Wang, M. Xiao, Y. Meng . (2016). Polymer Electrolytes for Lithium Polymer Batteries. J. Mater. Chem. A. 4. 10.1039/C6TA02621D.
[18] Q. Zhao, X. Liu, S. Stalin. et al. Solid-state polymer electrolytes with in-built fast interfacial transport for secondary lithium batteries. Nat Energy 4, 365–373 (2019). https://doi.org/10.1038/s41560-019-0349-7
[19] F. Baskoro, H.Wong, H-J. Yen. (2019). Strategic Structural Design of a Gel Polymer Electrolyte Towards a High Efficiency Li-ion Battery. ACS Applied Energy Materials. 10.1021/acsaem.9b00295.
[20] W.H. Meyer, (1998), Polymer Electrolytes for Lithium- Ion Batteries. Adv. Mater., 10:439-448.doi:10.1002/(SICI)1521-4095(199804)10:6<439::AID ADMA439>3.0.CO;2-I

Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)
IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

Web Security in IoT Networks using Deep Learning Model

Sanjay Patidar
Computer Science and Engineering Department
Delhi Technological University
New Delhi, India
sanjaypatidarcs@gmail.com

Inderpreet Singh Bains
Computer Science and Engineering Department
Delhi Technological University
New Delhi, India
inder.rockstar07@gmail.com

***Abstract*—The vision of IoT is to interface the things utilized in our day-to-day lives (which have the capacity of detecting and activation) via internet platform. This may or might possibly include humans. IoT field is developing and has many open digital issues. Internet of things (IoT) is still remaining in its early stages and has pulled in much enthusiasm for some mechanical parts including clinical fields, tech savvy, urban communities, and automotive. Anyway, as a paradigm, it is defenseless towards a scope of significant intrusion threats. In IoT whenever there is a web attack then it is required to remove the attack by installing software and so by using these models the attack can be removed from the system. This paper presents a threat investigation on IoT and utilizations on artificial neural network (ANN) to battle these threats. In this paper, profound learning method to incorporate digital security and prevention against attacks is also deployed, where a convolution 1d with multiple convolutions is used to increase the accuracy of the user. Profound models of learning are proposed and assessed those utilizing with the most recent CICIDS2017 datasets for DDoS assault recognition that has given most noteworthy precision of about 99.38%. It is essential to create an efficient intrusion identification framework that uses a deep learning mechanism to overcome attack issues in IoT framework. In this paper, a convolutional neural network [CNN] is developed with multiple convolution layers and accuracy of attack detection is also increased.**

Keywords—CNN; attacks; IoT; threats; deep learning.

I. INTRODUCTION

A. IoT architecture

The Internet of Things (IoT) is viewed as a future web that stretches out the internet to all way of genuinely shrewd gadgets in reality. A Cisco study reports the association between these brilliant gadgets and the Internet by around 50 billion constantly 2020. [1] The Internet of Things (IoT) is viewed as a future internet which stretches out the internet to all way of truly shrewd gadgets in reality. [2] IoT will give the wise and self-created cyberphysical frameworks in the fields of clever lattices, savvy urban communities, lodging, wise clinical and clinical administrations, wearable advancements, travel frameworks, and so on., by connecting these billion brilliant gadgets to the Internet. [3] The main work under IoT is for cyber security. Web of Thing is the most recent rising and encouraging innovation that interfaces everything across the globe through web. IoT innovation assures to enhance and support our own, proficient life and culture [4]. IoT comprises of a system of brilliant protests that are spread far and wide through internet with no human obstruction, which is incredible yet it is powerless to digital assaults like some more system. [5]The security of IoT framework is a significant test in modernization. [6] While many existing arrangements depend on human-characterized highlights to create AI (ML) based attack finders against noticeable adventures, such highlights are turning out to be progressively costly and less compelling in the smart network. [7]

The client's unit handles significant security-situated by taking care of frameworks entirely for information security. For example, encoding, unscrambling, and protection order forms. In IoT, these computationally escalated exercises can't be dealt by asset constrained keen machines as a wellbeing focused tasks that support the noisy computing strain. [8]One of the significant difficulties in cybersecurity is the arrangement of a robotized and compelling cyber-threats identification procedure. To overcome this difficulty, deep learning techniques are generally used in cyber security. [9] For this, several intrusion detection systems are implemented, which can check if any attack happens under software layers.

An important technique for detecting cyberattacks in any network is the intrusion detection system (IDS). A considerable lot of the new IDS depend on network-based AI algorithms to train and detect cyberattacks. Fog computing is an improved augmentation of brought together distributed computing wherein disseminated mist hubs are nearer to IoT arrange articles and address adaptability congestion, high data transmission usage nature of administration humbling, and minimal cloud computing high inertness. Fog-to-node registering is suitable for IoT networks being deployed in operation and being efficient. The figure below illustrates the haze to-hub model architecture with appropriated equal calculation giving insight to the disseminated hazes by giving IDS closer to the IoT organize objects computation, control, and storage. In this IoT networks are created which has fog, and has IDS which detects the attacks from users before connecting to the cloud. Compared with the cloud, IDS detect cyber threats easily and rapidly at fog nodes. IoT network consists of links between various types of savvy objects going from supercomputers to minuscule gadgets, which may have low processing capacity, so it is difficult to access these types of networks. Therefore, cybersecurity is a significant shortcoming in IoT network implementation. [10]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

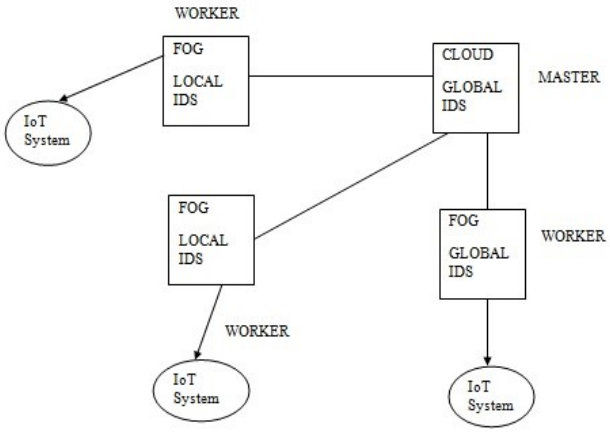


Fig. 1. IoT network architecture

B. CNN Methodology

Using the convolutional operation on the input datasets with pre processing and special layers and filters are applied in CNN method. Fig. 2 shows the base model representation of the CNN based model. The input data sets are fed using their features. Then convolution operation is applied. Maxpool, dropout and fully connected layer parts of CNN methodology is applied for the final output through dense layers. In this methodology convolutional layers are considered; however, will make it deeper by adding more convolutional layers, as well as maxpooling layers. Max pooling layer is included to dispose of highlights with low score and keep just highlights with most elevated score. The outcomes are down tested or pooled include maps that feature the most present component in the fix. This has been found to work preferred by and by over normal pooling for PC vision errands like picture arrangement. Dropout layer is added to spare framework from warming. Yield from the dropout layer is taken care of to completely associated layer which then give contribution to the thick bed alongside sigmoid capacity.

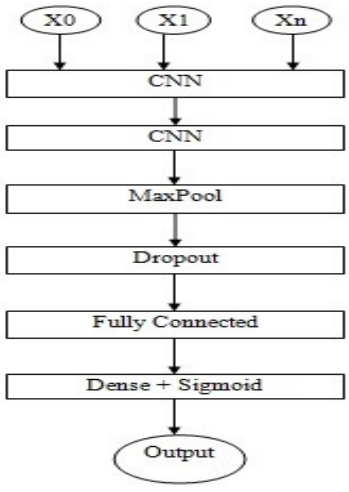


Fig. 2. CNN methodology

This paper is composed as follows: in section II, background work is mentioned. In section III, literature of previous year papers and existing information about IOT in brief is mentioned. In the section proposed IV work is presented. In section V implementation of approach of deep learning in cyber security for IOT is presented. Finally, in the next section results are shown and concluded the paper

II. BACKGROUND

The configuration of the 1D deep CNN model consists of an information layer, a convolutional layer, a max pooling layer, a completely associated layer, and a yield layer as shown in figure 3. We have applied a 1D Convolutional Neural Network on our data. A 1D CNN is exceptionally successful when you would like to get intriguing features from shorter (fixed-length) portions of the general educational record and where the territory of the component inside the area isn't of high significance. This applies well to the examination of time successions of sensor information. It additionally applies to the investigation of any sort of sign information over a fixed-length period, (for example, sound signs). Another application is NLP.

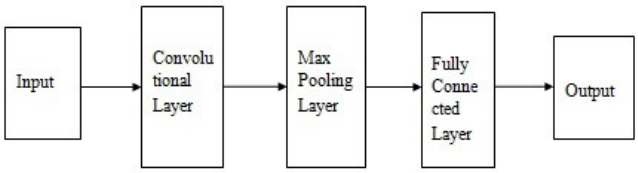


Fig. 3. 1d CNN architecture

III. LITERATURE REVIEW

A. The Internet of Things

The Internet of Things (IoT) as an idea has its roots in the early 1990s was likely one of the first people to recognize the eventual rise of a system where on-demand computing was available to everyone through a combination of hardware and software, connected using wires or radio communication. N.Y. Parotkin et al. [11] in modern times, this idea has materialized as a viable and imminent future technology, characterized by a massively connected system of items or devices which can associate with each another over a network connection.

R. Mahmoud et al. [12] today's widespread, worldwide telecommunications network lays the foundation for massive IoT in the near future. The rise of IoT as a technology platform is also partially attributable to the rapid downward scaling (miniaturization) of transistors; a trend that has been fairly consistent since the late 20th century and continues into the 21st century. Transistors form the bedrock of all silicon-based intelligence in today's time. This observed law of miniaturization is well documented in Dennard's Scaling and Moore's Law; two well-known observations relating to computing in recent times. Current advances in miniaturization allow small sensors and computing modules to operate cheaply, efficiently, and be deployed at a large scale across a wide range of real-world applications.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

The Industrial applications of IoT (IoT) are of particular interest to this paper. IoT devices are primarily used to ingest and analyses data from industrial equipment, operational technology, physical locations, and human resources. Cyber security is an intensive industrial activity in the modern world; given its essential nature, IoT can make significant improvements to it. [13]

B. Deep Learning Methods for Security in IOT

Subsequently, our exploration work advocates improving IoT security by utilizing Deep Learning calculations. W. Abbass et al. [14] Deep Learning is in no way, shape or form an ongoing worldview. It is a subfield pertaining to machine training that has its underlying foundations in Artificial Intelligence. Profound Learning helps essentially perform characterization assignments straightforwardly from writings, pictures and sounds. As of now, Deep Learning is to a great extent engaging the IT scene by tackling different issues. The neurons of the ANN are utilized to form complex hypotheses; the more neurons, the progressively complex the speculations. Assessing the speculations is finished by setting the info hubs in a criticism procedure and the occasion streams are spread through the network to the yield where it is named typical or bargained. J. Lee et al. [15] at this stage the inclination plunges are utilized to push the blunder in the yield hub back through the network by a back-proliferation process so as to gauge the mistake in the concealed hubs. The inclination of the expense – capacity would thus be able to be determined [16]. Neural network framework experiences preparing so as to gain proficiency with the example made in the framework. There are many techniques applied in cyber security with the help deep learning. Some of the techniques are LTSM which long term short memory. Recurrent neural network known as RNN is used in some papers. Also, some researchers have used natural language processing techniques and support vector machine is also used.

Subsequently, different systems like convolutional neural networks are likewise effectively investigated around there, which incorporates input surface, convolution surface, pooling surface, completely associated surface, and yield surface. Konstantinos P. Ferentinos [17] discusses plant disease detection and diagnosis models through a database that contains photographs of healthy and infected plant leaves. The future direction in this is to gather a wide range of training data from various sources from different geographic areas, conditions of cultivation, and other factors.

In SenseBox architecture, H M Sajjad Hossain et al. [18] proposed a DeActive model. This algorithm is executed much more quickly than other algorithms. Dynamic learning can assist us with alleviating the manual effort expected to compile ground-level data about truth and decrease preparing time. With far less marked cases, DeActive can give better precision, which also ensures lower annotation effort.

Chris Xiaoxuan Lu et al. [19] examined security aspects by sniffing a deep learning smartwatch password that is a Snoopy method. Snoopy uses a uniform structure to separate movement information portions, albeit passwords are inserted, and utilizes new profound neurological systems directed toward surmise the real watchwords. This system can effectively spy information on moving out of sight while entering passwords. Without devouring noteworthy force/ computational assets, it can successfully extract password segments of motion data on smartwatches in real-time.

Parisa Pouladzadeh et al. [20] presented an app that utilizes the image of the nourishment, taken by the client's cell phone, to perceive different nourishment things in a similar food to evaluate the calorie and sustenance of the nourishment. In this, the client is challenged to rapidly recognize the broad territory of nourishment by an outline a bouncing ring on the nourishment image by contacting the canopy. The framework, at that point, utilizes picture handling and statistical insight for food item acknowledgment.

Jennifer S. Raj et al. [21] introduced a study of the computational smart strategies as they appear to be viable option for the man-made brainpower by conquering the disappointments and disadvantages in it. Examination of the different computational strategies to locate the ideal one in the identification of false access will be a future heading.

Samuel Manoharan et al. [22] introduced measures to improve the likelihood of the wellbeing for the vehicles with the capacity of self driving worked with man-made reasoning empowered processors. Further the model guaranteed the ability of the security calculation in oneself driving vehicles.

IV. PROPOSED WORK

In this section, detailed methodology and proposed models are explained. The main objectives are to implement deep learning method with higher accuracy in cyber security to compare the accuracy with existing methods.

A. Proposed Methodology

The flow diagram of our framework is shown in figure 4; here the input is the csv files which were gathered with several diverse digital assaults alongside with usual data for five consecutive days. In this, first applied the convolution 1d to our system and then divide the dataset into two categories: BENIGN and Web attack. BENIGN is neural and rest are different web attacks Then in the next step label the data by assigning benign as 0 and web attack as 1. Then divide the information into 70% preparing as well as 30% for examination. Then set the batch size to 32 and epoch which are the number of rounds to 100. Then use for loop for epoch=1, 2, 3 and so on. For this set the count to 0 and repeat until count+batch size< no. of training samples else it goes back to the dataset division. Then train the classifier on benign and web attack to help him learn to classify different web attacks, and then set count equal to batch size and then perform testing on benign and web attack. The last step which is the output is to compute the exactness, accuracy, recall and f1-score from the confusion matrix to know the accuracy of model.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

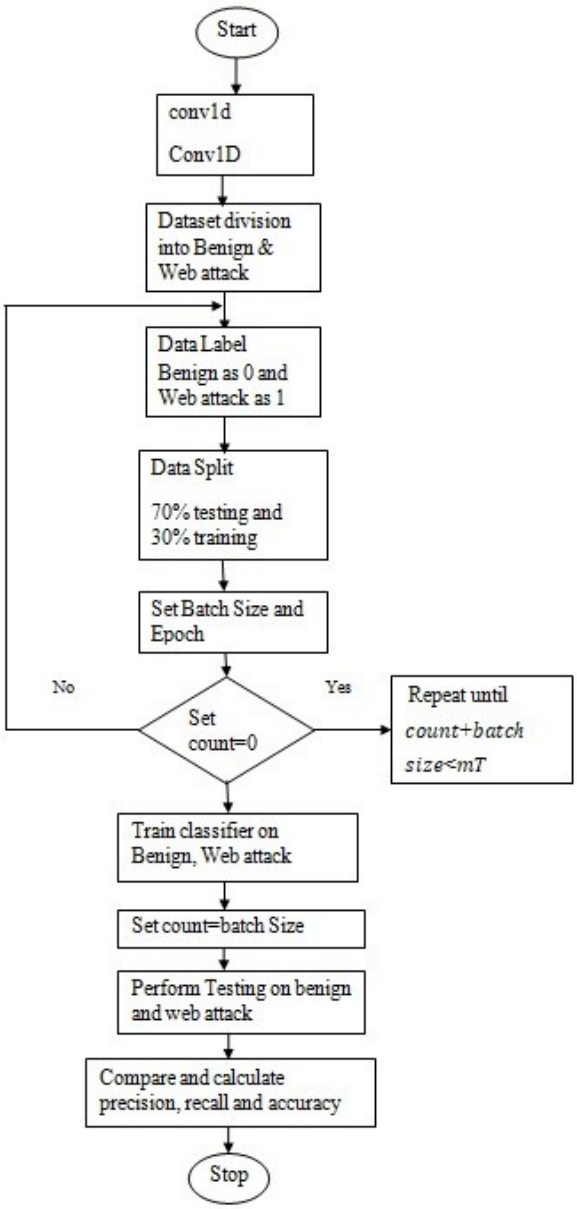


Fig. 4. Proposed Methodology

B. Algorithm

The proposed algorithm is given below:

Input: m- No. of samples
n- No. of features
0- Label for BENIGN data
1- Label for Web attack
mT- No. of training samples
mt- No. of testing samples
$XTmT\times n$- Training samples
$YTmT\times 1$- Training labels
$Xtmt\times n$- Testing samples
$Ytmt\times 1$- Testing labels
$Ypredmt\times 1$- Predicted labels
Output: tp, tn, fp, fn

1. Prepared dataset $Xm\times n$ and label $Ym\times 1$ where $Y\in \{0, 1\}$
2. Split data $Xm\times n$ and label $Yn\times 1$ into 70% training and 30% testing sets
3. ($XTmT\times n$,$\times 1$) is the training set and ($Xtmt\times n$,$Ytmt\times 1$) is the testing set
4. Set $batch\ \ size$=512 $and\ \ epoch\ s$=100
5. for $epoch$ =1,2,3,…….,$epoch\ s$
6. Set $count$=0
7. Repeat until $count$+$batch\ \ size$<mT. Train classifier on (XT$count$+1 $to\ \ count$+$batch\ \ size$, Y$T count$+1 $to\ \ count$+$batch\ \ size$)
6. Calculate accuracy
7. Set $count$=$batch\ \ size$
8. end for
9. Perform testing on $Xtmt\times n$ and find $Ypredmt\times 1$
10. Compare $Ypredmt\times 1$ and $Ytmt\times 1$ and calculate precision, recall, F1-Score, Accuracy

First, divide the information into the preparation and testing part, where 70% of data have utilized for preparing and rest 30% part for testing. Since, the dataset is highly imbalanced, where attacks are in minimum quantity as compared to the BENIGN. Therefore, a different strategy such as multilevel classification can be adopted where the first decision will be whether the data is BENIGN or ATTACKED. If data comes in ATTACKED category then will predict the nature of the attack. Then set the batch size to 512 and epoch which are the number of rounds to 100. Then use for loop for epoch=1, 2, 3 and so on. For this set the count to 0 and repeat until count+batch size< no. of training samples else it goes back to the dataset division. Then train the classifier on benign and web attack and then set count=batch size and then perform testing on benign and web attack. The last step is to compute the exactness, accuracy, recall and f1-score from the confusion matrix to know the accuracy of model.

C. Proposed CNN 3 Layer Model

The proposed model 1 also known as CNN 3 layers is used, which is the modified CNN based deep learning algorithm as shown in the figure 5. In this figure, multiple layers of deep learning-based convolution and max pooling are applied to improve the accuracy. In this there are convolution bed followed by max pooling bed after that dropout layer is applied. The dropout layers are combined to spare the framework from warming. Yield from the dropout layer is taken care of to flatten bed which at that point provide input for the dense bed which then provide input to the subsequent dropout layer. Yield from the dropout bed is taken care of to the second dense layer with sigmoid, relu, and softmax activation function. This CNN model is ideal when requirement is of less computation as there is less parameter required in this model. Here the accuracy of the model is 99.10%.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

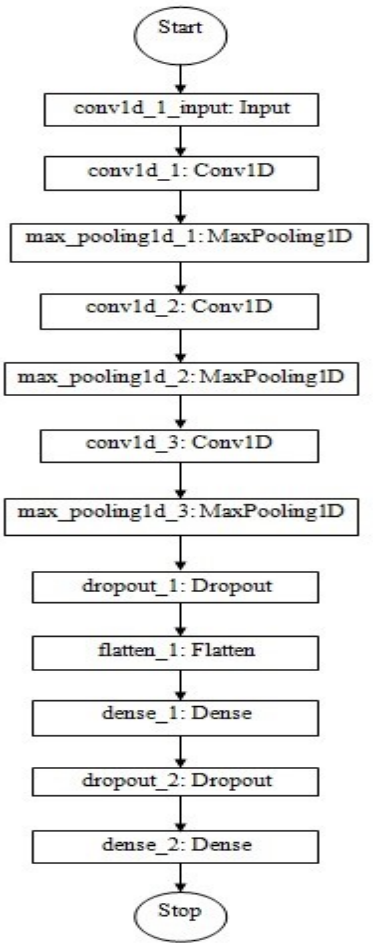


Fig. 5. Proposed CNN 3 Layer Model

D. Proposed Multiheaded CNN Layer Model

The proposed model 2 also known as multiheaded CNN layers is used, which basically concatenate three CNN processes as shown in the figure 6. They are followed by maxpooling layers, dropout layers and flatten layers and then it connects them with concatenate layer. Yield from the concatenate layer is connected to the dense bed that gives contribution to the dropout bed. The dropout bed is added to spare framework from warming. Yield from the dropout bed is connected to the second dense bed with sigmoid, relu, and softmax activation function. In this model the layers are connected parallelly which is why this model is also known as multiheaded CNN layers. It gives significant improvement in the accuracy and other measures as it applies multiple CNN layers to reduce the error to minimum. This CNN model is ideal when requirement is of high computation as it requires more parameters. Here the accuracy of the model is 99.38% which is the highest of all the models.

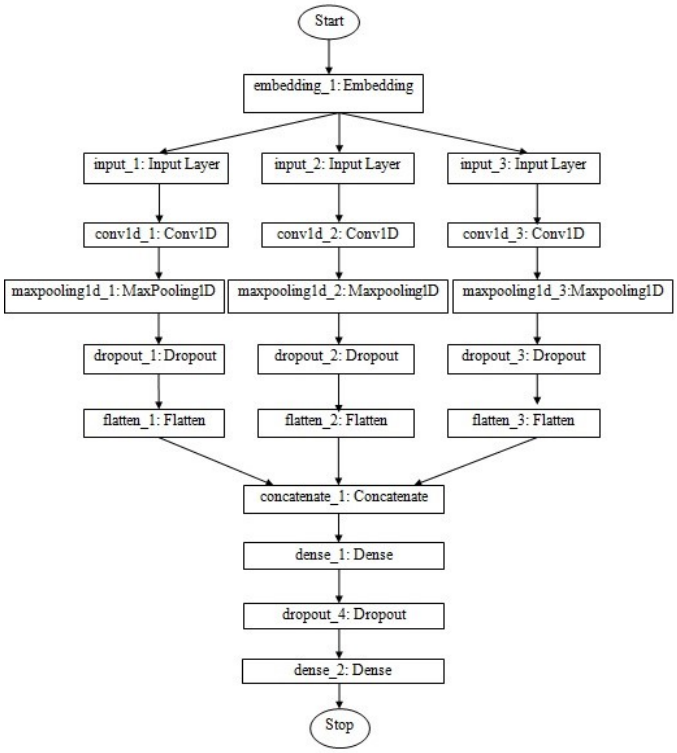


Fig. 6. Proposed Multiheaded CNN Layer Model

V. IMPLEMENTATION

In this segment first we illustrated the experimental setup and then the dataset used in the paper.

A. Experimental Setup

This work is utilized by using Keras [23] on Tensorflow bundle for profound study on 64-bit Intel core-i5 CPU with 30 GB RAM in Windows 7 environment. In this there are local parameters and global variables.

Local parameter

Local parameters are announced within a function and can be utilized distinctly inside that function. Local variables of the same name can be used in different functions. Local parameters in this are: number of features, label for benign, label for web attack, number of training samples, training samples, testing samples, training labels, testing labels, predicted labels, true positive, true negative, false positive, false negative.

Global parameter

Global parameters are announced outside any function, and they can be utilized on any capacity in the program.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

In this, TP is genuine positive, FN is bogus negative, TN is genuine negative and FP is bogus positive. F1 score depends on precision and recall. Since the dataset is imbalanced, therefore the exhibition of the models is assessed as far as Precision, Recall, F1-Score, and Accuracy.

B. Dataset

For leading proposed work we have utilized most recent DDoS assault CICIDS2017 dataset. CICIDS2017 datasets contain exceptional genuine work arrange taking after information. This dataset was accumulated for 5 sequential days with different cyberattacks alongside ordinary information. This dataset contains latest cutting-edge arrange information with and without assault which is near the genuine work organize data.

In this implementation, python language is used. There are 2830743 data samples and 79 features. After preprocessing like removing columns having all zeros and normalization, we have total 76 features.

In this implementation, further we will discuss about the dataset. It contains total 2830743 data samples from 15 different categories having 79 features. The categories and the percentage of data samples in the dataset are as follows: BENIGN (80.3%), Infiltration (0.0013%), DDoS (4.52%), Bot (0.069%), Web assault sql infusion (0.0007%), SSH-Patarator (0.21%), DoS slowloris (0.20%), DoS Hulk (8.16%), PortScan (5.61%), Heartbleed (0.0004%), DoS Slowhttptest (0.19%), DoS GoldenEye (0.36%), FTP-Patarator (0.28%), Web assault Brute Force (0.053%), Web assault XSS (0.023%).

Out of 15 categories BENIGN is neural and rest are different web attacks. As we can see that percentage of individual attacks data in many cases are around 0%. BENIGN is the most used data sample in the dataset with the second being DoS Hulk. Therefore, we have divided data into two parts BENIGN and Web Attacks and then profound learning models are applied to arrange the information into either of the two categories. The dataset contains 79 features, some of highlights are: goal port, stream duration, aggregate forward parcels, complete in reverse bundles, absolute length of forward bundles, all out length of in reverse bundles, forward bundle length max, forward bundle length min, forward parcel length mean, forward parcel length standard, backward parcel length max, backward parcel length min, backward bundle length mean, backward bundle length standard, stream byte/s, stream packet/s, stream IAT mean, stream IAT standard, stream IAT max, stream IAT min and so on.

VI. RESULTS

In this section, results are presented. The testing of results is done in complete week bases with different attacks and methods on each day. The results are basically comparison of the performance metrics explained in previous section. The exhibition of the classifier is presented in Table 1. The performance metrics are then analyzed for each of the models and further explained in results section.

The confusion matrix for all the models is presented in figure 7, 8, and 9. From this we can get the values of genuine positive, genuine negative, bogus positive, bogus negative and ascertain exactness, recall, along with precision.

A. Confusion matrix of CNN Model

The disarray network of the CNN model which is the base model is shown in figure 7. For example Heartbleed attack 100% incorrectly classifies the attack as BENIGN. However the second attack DoS slowloris 97% correctly classifies that the attack is DoS Slowloris but 2 % incorrectly the attack as BENIGN and DoS Slowhttptest. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

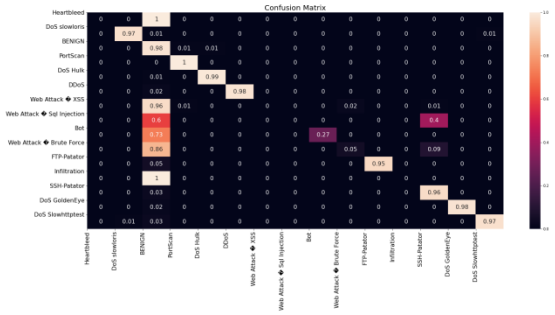


Fig. 7. Confusion matrix of CNN model

B. Confusion matrix of Proposed CNN 3 Layer Model

The disarray network of the proposed CNN 3 layer model is shown in figure 8. For example Web attack-brute force attack does not correctly classifies that it is the same attack. Instead it 92% incorrectly classifies the attack as BENIGN. Similarly the second attack Web attack- XSS 99% incorrectly classifies that the attack is BENIGN. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

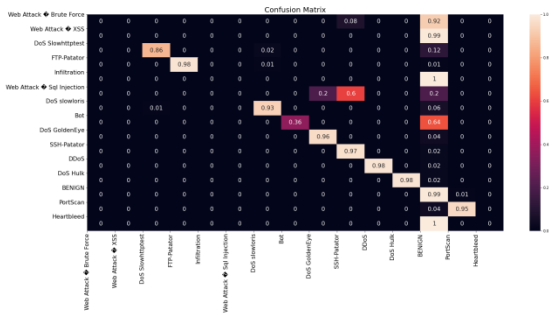


Fig. 8. Confusion matrix of proposed CNN 3 layer model

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

C. *Confusion matrix of Proposed Multiheaded CNN Layer Model*

The disarray network of the proposed multiheaded CNN layer model is shown in figure 9. For example Heartbleed 100% incorrectly classifies the attack as BENIGN. The second attack SSH-patator 97% correctly classifies the attack as SSH-patator but 3% incorrectly classifies the attack as BENIGN. Similarly we calculate the values of all attacks and then calculate the accuracy recall, precision of the model.

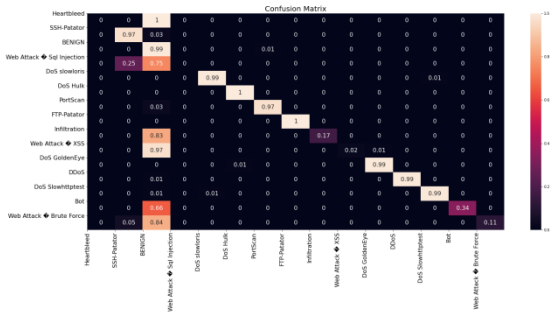


Fig. 9. Confusion matrix of proposed multiheaded CNN layer model

D. *Performance Metrics Evaluation*

The performance metrics evaluation is shown in table 1. In this the CNN model has precision of 94.33%, recall of 97.62%, f1-score of 96.41%, and accuracy of 98.32%. The second model which is the proposed CNN 3 layer model has precision of 96.54%, recall of 98.44%, f1-score of 97.48%, and accuracy of 99.10%. The third model which is the proposed multiheaded CNN layer model has precision of 98.70%, recall of 99.33%, f1-score of 99.01%, and has an accuracy of 99.38% which is the highest of all the models.

Table I. Performance Metrics Evaluation

Model Name	Precision	Recall	F1-Score	Accuracy
CNN Model	94.33 %	97.62 %	96.41 %	98.32 %
Proposed CNN 3 Layer	96.54 %	98.44 %	97.48 %	99.10 %
Proposed Multiheaded CNN Layer	98.70 %	99.33 %	99.01 %	99.38 %

The epochs vs. accuracy and epochs vs. loss chart for all models are presented in Figure 10, 11, and 12.

E. *Epoch vs. Accuracy and Epoch vs. Loss curve of CNN Model*

In this CNN model the testing accuracy versus epoch is significantly high in the below figure 10 for the first model. As we increase the number of epochs the testing accuracy also increases. Here epochs are the number of rounds in which we are running the model. The loss vs. epoch curve shows that loss is decreasing as the number of rounds increases.

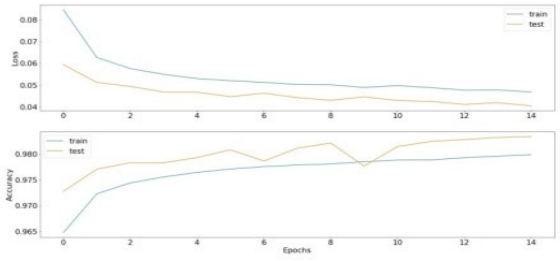


Fig.10. Epochs vs. Accuracy and Epochs vs. Loss curve of CNN Model

F. *Epoch vs. Accuracy and Epoch vs. Loss curve of Proposed CNN 3 Layer Model*

The CNN 3 layer model for the representation results is shown in figure 11; variation in accuracies is seen as per the increase in the epochs. As number of rounds increases there are variations in testing accuracy. The variation in accuracy shows that the accuracy is not constant it keeps on fluctuating. The loss vs. epoch curve shows variation in loss as per the increase in the epochs.

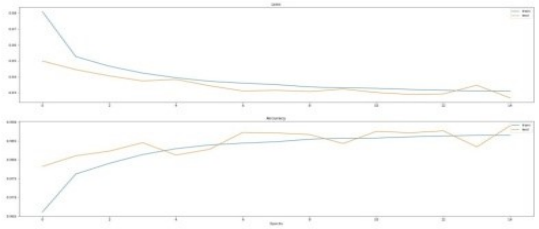


Fig. 11. Epochs vs. Accuracy and Epochs vs. Loss curve of proposed CNN 3 layer model

G. *Epoch vs. Accuracy and Epoch vs. Loss curve of Proposed Multiheaded CNN Layer Model*

The final proposed multiheaded CNN model result for accuracy and loss curve is seen as shown in figure 12, which improved in terms of high accuracy and low loss techniques result. In this the testing accuracy is improved in terms of low loss in the model. The loss vs. epoch curve shows loss is decreasing and as a result testing accuracy of the model increases. This model has the highest accuracy as compared to the rest of the models.

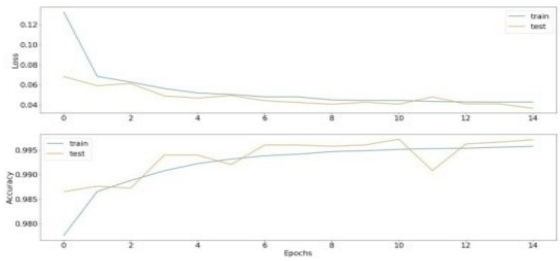


Fig. 12. Epochs vs. Accuracy and Epochs vs. Loss curve of proposed multiheaded CNN layer model

H. *Comparison of Evaluation parameters for all models*

Comparison of all three-model implementation on the basis of four parameters, accuracy, F1 score, recall and precision are

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.

shown in figure 13 for deep learning method in cyber security. Here model 1 is the cnn model which is the base model, model 2 is the proposed cnn 3 layer model, and model 3 is the proposed multiheaded cnn layer model. It is seen that proposed multiheaded cnn layer model is the best amongst the others and has a higher accuracy amongst all that means has the least error.

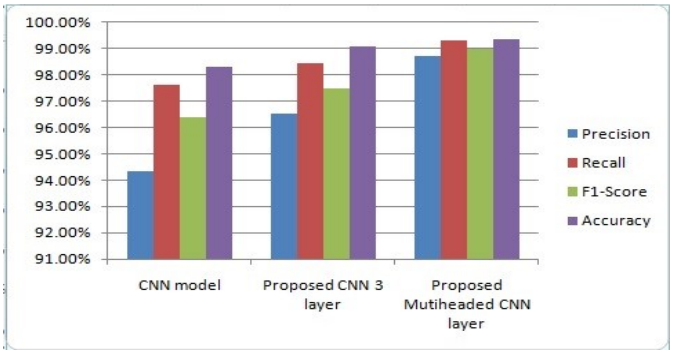


Fig. 13. Comparison of evaluation parameters for all models

VII. CONCLUSION

Web of Thing is the most recent significant development that interfaces all around the globe through web. IoT development helps to improve and bolster the proficient lifestyle and culture. The proposed work is focused on deep learning innovation using CNN and its different model variants. The datasets are used and tested with all possibilities to give better results and innovation in cyber security issues for internet of things. The loss and accuracy are the main parameter for analysis in the ground pertaining to security improvement under the web of things domain. Hence, this paper has provided an improved method to detect security issues in IoT by using a modified deep learning method. The proposed results indicate a higher accuracy in the CNN modified algorithm. In future it can be tested on edge servers and cloud assisted servers. As it is highly unbalanced by duplicating data, the dataset is balanced for this study; this could be enhanced in the future by developing a deep learning model that could run on the unbalanced dataset.

REFERENCES

[1] M. Roopak, G. Yun Tian and J. Chambers, "Deep Learning Models for Cyber Security in IoT Networks," *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 2019, pp. 0452-0457, doi: 10.1109/CCWC.2019.8666588.

[2] Thamilarasu, Geethapriya & Chawla, Shiven. (2019). Towards Deep-Learning-Driven Intrusion Detection for the Internet of Things. Sensors. 19. 1977. 10.3390/s19091977.

[3] K. Gupta and S. Shukla, "Internet of Things: Security challenges for next generation networks," 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH), Noida, 2016, pp. 315-318, doi: 10.1109/ICICCS.2016.7542301.

[4] A. Samandari, M. Ge, J. B. Hong and D. S. Kim, "Evaluating the Security of IoT Networks with Mobile Devices," *2018 IEEE 23rd Pacific Rim International Symposium on Dependable Computing (PRDC)*, Taipei, Taiwan, 2018, pp. 171-180, doi: 10.1109/PRDC.2018.00028

[5] A. Abdullah, R. Hamad, M. Abdulrahman, H. Moala and S. Elkhediri, "CyberSecurity: A Review of Internet of Things (IoT) Security Issues, Challenges and Techniques," *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/CAIS.2019.8769560.

[6] A. Aldaej, "Enhancing Cyber Security in Modern Internet of things (IoT) Using Intrusion Prevention Algorithm for IoT (IPAI)," in *IEEE Access*, doi: 10.1109/ACCESS.2019.2893445.

[7] F. Ullah *et al.*, "Cyber Security Threats Detection in Internet of Things Using Deep Learning Approach," in *IEEE Access*, vol. 7, pp. 124379-124389, 2019, doi: 10.1109/ACCESS.2019.2937347.

[8] F. Rahman, M. Farmani, M. Tehranipoor and Y. Jin, "Hardware-Assisted Cybersecurity for IoT Devices," *2017 18th International Workshop on Microprocessor and SOC Test and Verification (MTV)*, Austin, TX, 2017, pp. 51-56, doi: 10.1109/MTV.2017.16

[9] Y. Lu and L. D. Xu, "Internet of Things (IoT) Cybersecurity Research: A Review of Current Research Topics," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2103-2115, April 2019, doi: 10.1109/JIOT.2018.2869847.

[10] A. Chadd, "DDoS attacks: past, present and future," *Network Security,* vol. 2018, pp. 13-15, 2018.

[11] N. Y. Parotkin and V. V. Zolotarev, "Information Security of IoT Wireless Segment," *2018 Global Smart Industry Conference (GloSIC)*, Chelyabinsk, 2018, pp. 1-7, doi: 10.1109/GloSIC.2018.8570144.

[12] R. Mahmoud, T. Yousuf, F. Aloul and I. Zualkernan, "Internet of things (IoT) security: Current status, challenges and prospective measures," *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, London, 2015, pp. 336-341, doi: 10.1109/ICITST.2015.7412116.

[13] M. R. Schurgot, D. A. Shinberg and L. G. Greenwald, "Experiments with security and privacy in IoT networks," "*2015 IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Boston, MA, 2015, pp. 1-6, doi: 10.1109/WoWMoM.2015.7158207".

[14] W. ABBASS, Z. BAKRAOUY, A. BAINA and M. BELLAFKIH, "Classifying IoT security risks using Deep Learning algorithms," 2018 6th International Conference on Wireless Networks and Mobile Communications (WINCOM), Marrakesh, Morocco, 2018, pp. 1-6, doi: 10.1109/WINCOM.2018.8629709.

[15] J. Lee, J. Kim, I. Kim and K. Han, "Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles," in *IEEE Access*, vol. 7, pp. 165607-165626, 2019, doi: 10.1109/ACCESS.2019.2953095.

[16] D. Wilson, Y. Tang, J. Yan and Z. Lu, "Deep Learning-Aided Cyber-Attack Detection in Power Transmission Systems," *2018 IEEE Power & Energy Society General Meeting (PESGM)*, Portland, OR, 2018, pp. 1-5, doi: 10.1109/PESGM.2018.8586334.

[17] Dan, C., Meier, U., Masci, J., Gambardella, L.M., Schmidhuber, J. "Flexible, high execution convolutional neural systems for image classification," Events of the 22nd International Joint Conference on Artificial Intelligence, vol. 2, pp. 1237–1242, 2011.

[18] Hande Alemdar, TLM van Kasteren, and Cem Ersoy, "Active learning with ambiguity sampling for broad-scale activity identification in a smart dormitory," Journal of Ambient Intelligence and Smart Environments 9, 2, 209-223, 2017.

[19] Mariamn Harbach, Alexander De Luca, and Serge Egelman, "The anatomy of smartphone unlocking: A field study of android lockscreens" In ACM Conference on Human Factors in Computing Systems, CHI, 2016.

[20] Parisa Pouladzadeh, Pallavi Kuhad, Sri Vijay Bharat Peddi, Abdulsalam Yassine, and Shervin shirmohammadi. "Calorie measurement and food classification using deep learning neural network," in Proceedings of the IEEE International Conference on Instrumentation and Measurement Technology, 2016.

[21] Raj, Jennifer S. "A Comprehensive Survey On The Computational Intelligence Techniques And Its Applications." Journal of ISMAC 1, no. 03 (2019): 147-159.

[22] Manoharan, S. (2019). "An Improved Safety Algorithm For Artificial Intelligence Enabled Processors In Self Driving Cars",Journal of Artificial Intelligence, 1(02), 95-104.

[23] Keras deep learning P.W.D. Charles Project Title Available: https://github.com/charlespwd/project-title

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on October 19,2020 at 06:14:15 UTC from IEEE Xplore. Restrictions apply.