SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(87th Edition)

MARCH 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 **Eighty Seventh** 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 **March, 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

10. Grid interfaced solar water pumping system with improved space vector modulated direct torque control, Priyanka Singh and *7.1.Prerna Gaur,* Electrical, DTU

11. High-Performance Dual Cavity-Interferometric Volatile Gas Sensor Utilizing Graphene/PMMA Nanocomposite, Dnyandeo Pawar, Rajesh Kanawade, Ajay Kumar, Ch. N. Rao, Peijiang Cao, Shankar Gaware, Dattatray Late, Sangeeta N. Kale, S. T. Navale, W. J. Liu, D. L. Zhua, Y.M. Lu, *3.Ravindra K. Sinha*, Physics, DTU

12. Machine Learning & Concept Drift based Approach for Malicious Website Detection, *8.Siddharth Singhal*, *8.Utkarsh Chawla* and Rajeev Shorey, Computer Science, DTU

13. Microstates-based resting frontal alpha asymmetry approach for understanding affect and approach/ withdrawal behavior, *6.Ardaman Kaur*, Vijayakumar Chinnadurai and *3.Rishu Chaujar*, Physics, DTU

14. Modeling Involuntary Dynamic Behaviors to Support Intelligent Tutoring (Student Abstract), *8.Mononito Goswami*, Lujie Chen y, Chufan Gao and Artur Dubrawski, Computer, DTU

15. Multi-Criteria decision analysis for optimizing proper selection medium for CSR in automobile industry, *8.Reetik Kaushik*, *8.Roopa Singh*, *8.Sanchit Gupta* and *3.Ranganath M Singari[2]*, [1]Mechanical and [2]Production & Industrial, DTU

16. Narrow-Band UVB-Emitting Gd-Doped SrY2O4 Phosphors, VIJAY SINGH, K. SWAPNA, SUMANDEEP KAUR, *3.A.S. RAO* and J.L. RAO, Physics, DTU

17. Participation of Women in Engineering & Technology Education and Employment, *6.Addissie Melak* and *3.Seema Singh*, Humanities, DTU

18. Path Tracing and Object Avoidance Algorithm for Robotic Manipulators incorporating Constrained Filters, *6.Vipul Garg* and *3.Vikas Rastogi*, Mechanical, DTU

19. Sarcasm Detection in Mash-up Language using Soft-Attention based Bidirectional LSTM and feature-rich CNN, Deepak Jain1, *3.Akshi Kumar* and *3.Geetanjali Garg*, Computer Science, DTU

20. Study of hybrid photovoltaic thermal systems, Ashish Saurabh, Deepali Atheaya and *3.Anil Kumar*, Mechanical, DTU

21. Two Finger Gesture Imitating Animatronic Hand, *6.Y Verma* and *3.V Rastogi*, Mechanical, Production and Industrial and Automobile, 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*

A Multi-task Approach to Open Domain Suggestion Mining (Student Abstract)

Minni Jain[*†]**, Maitree Leekha**[*]**, Mononito Goswami**[*]

Department of Computer Science & Engineering,
Delhi Technological University, New Delhi, India
{minnijain, maitreeleekha_bt2k16, mononito_bt2k16}@dtu.ac.in

Abstract

Consumer reviews online may contain suggestions useful for improving the target products and services. Mining suggestions is challenging because the field lacks large labelled and balanced datasets. Furthermore, most prior studies have only focused on mining suggestions in a single domain. In this work, we introduce a novel up-sampling technique to address the problem of class imbalance, and propose a multi-task deep learning approach for mining suggestions from multiple domains. Experimental results on a publicly available dataset show that our up-sampling technique coupled with the multi-task framework outperforms state-of-the-art open domain suggestion mining models in terms of the F-1 measure and AUC.

1 Introduction

Consumers often express their opinions towards products and services through online reviews and discussion forums. These reviews may include useful suggestions which can help companies better understand consumer needs and improve their products and services. However, manually mining *suggestions* amid vast numbers of *non-suggestions* can be cumbersome, and equated to finding needles in a haystack. Therefore, designing systems which can automatically mine suggestions is essential. The recent *SemEval* challenge on Suggestion Mining saw many researchers using different techniques to tackle the domain specific task. However, *open-domain suggestion mining*, which obviates the need for developing separate suggestion mining systems for different domains, is still an emerging research problem. Building on the work of (Negi 2019), we design a framework to detect suggestions from multiple domains. Specifically, we formulate a multitask classification problem to identify both the domain and nature (suggestion or non-suggestion) of reviews. Furthermore, we also propose a novel language model-based text up-sampling approach to address the class imbalance problem.

---

[*]All the authors contributed equally, and wish that they be regarded as joint First Authors.

[†]Mobile Number: +91 8587898334

2 Methodology

Dataset & Pre-processing

We use the first publicly available and annotated dataset for suggestion mining from multiple domains created by (Negi 2019). It comprises of reviews from four domains namely, `hotel`, `electronics`, `travel` and `software`. During pre-processing, we remove all URLs (eg. *https://...*) and punctuation marks, convert the reviews to lower case and *lemmatize* them. We also pad the text with start `S` and end `E` symbols for up-sampling.

Up-Sampling using Language Model: LMOTE

One of the major challenges in mining suggestions is the imbalanced distribution of classes, *i.e.* the number of non-suggestions greatly outweigh the number of suggestions. To this end, studies frequently utilize *Synthetic Minority Oversampling Technique* (SMOTE) (Chawla et al. 2002) to up-sample the minority class samples using the text embeddings as features. However, SMOTE works in the euclidean space and therefore does not allow an intuitive understanding and representation of the over-sampled data, which is essential for qualitative and error analysis of the classification models.

We introduce a novel over-sampling technique, **L**anguage **M**odel-based **O**ver-sampling **Te**chnique (LMOTE), exclusively for text data and note comparable (and even slightly better sometimes) performance to SMOTE. We use LMOTE to up-sample the number of suggestions before training our classification model. For each domain, LMOTE uses the following procedure to over-sample suggestions:

Find top η n-grams: From all reviews labelled as suggestions (positive samples), sample the top η=100 most frequently occurring n-grams (n=5). For example, the phrase *"nice to be able to"* occurred frequently in many domains.

Train language model on positive samples: Train a BiLSTM language model on the positive samples (suggestions). The BiLSTM model predicts the probability distribution of the next word (w_t) over the whole vocabulary ($V \cup \mathbf{E}$) based on the last n=5 words ($w_{t-5}, ..., w_{t-1}$), *i.e.*, the model learns to predict the probability distribution $P(w_i \mid w_{t-5} \ w_{t-4} \ w_{t-3} \ w_{t-2} \ w_{t-1}) \ \forall i \in (V \cup \mathbf{E})$, such that $w_t = \underset{w_i}{\arg\max} \ P(w_i \mid w_{t-5} \ w_{t-4} \ w_{t-3} \ w_{t-2} \ w_{t-1})$.

Generate synthetic text using language model and frequent n-grams: Using the language model and a randomly chosen frequent 5-gram as the seed, we generate text till the end symbol `E` is predicted.

Mining Suggestion using Multi-task Learning

Multi-task learning (MTL) has been successful in many applications of machine learning since sharing representations between auxiliary tasks allows models to generalize better on the primary task. Figure 1B illustrates 3-dimensional Uniform Manifold Approximation and Projection (UMAP) visualization of *text embeddings* of suggestions, coloured by their domain. These embeddings are outputs of the penultimate layer (dense layer before the final softmax layer) of the *Single task* (STL) ensemble baseline. It can be clearly seen that suggestions from different domains may have varying feature representations. Therefore, we hypothesize that we can identify suggestions better by leveraging domain-specific information using MTL. Therefore, in the MTL setting, given a review r_i in the dataset, D, we aim to identify both the domain of the review, as well as its nature.

Classification Model

We use an ensemble of three architectures namely, CNN to mirror the spatial perspective and preserve the `n-gram` representations; Attention Network to learn the most important features automatically; and a BiLSTM-based text RCNN model to capture the context of a text sequence. In the MTL setting, the ensemble has two output softmax layers, to predict the domain and nature of a review. The STL baselines on the contrary, only have a singe softmax layer to predict the nature of the review. We use ELMo (Peters et al. 2018) word embeddings trained on the dataset, as input to the models.

Domain	Baseline	STL	STL + SMOTE	STL + LMOTE	MTL + LMOTE
Hotel	0.77 (LSTM)	0.79	0.83	0.83	**0.86**
Electronics	0.78 (SVM)	0.80	0.80	0.83	**0.83**
Travel	0.66 (SVM)	0.65	0.68	0.69	**0.71**
Software	0.80 (LSTM)	0.79	0.81	0.84	**0.88**
Pooled AUC		0.876 ±0.014	0.883 ±0.013	0.897 ±0.012	**0.907** **±0.012**

Table 1: Performance evaluation using F-1 score & pooled Area under ROC curve (AUC) with 95% confidence intervals. Multi-task Learning with LMOTE outperforms other alternatives in open-domain suggestion mining.
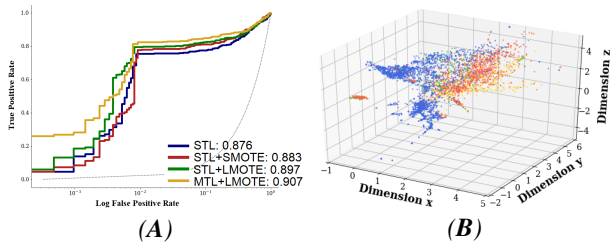

Figure 1: *(A)* ROC (TPR vs. Log FPR) curves pooled across all domains for all models used in this work. *(B)* 3-dimensional UMAP visualization of text embeddings of *suggestions* coloured by domain.

3   Results and Discussion

We conducted experiments to assess the impact of over-sampling, the performance of LMOTE and the multi-task model. We used the same train-test split as provided in the dataset for our experiments. All comparisons have been made in terms of the F-1 score of the suggestion class for a fair comparison with prior work on representational learning for open domain suggestion mining (Negi 2019) (refer *Baseline* in Table 1). For more insightful evaluation, we also find the Pooled Area Under the Receiver Operating Characteristic (ROC) curves for all models used in this work. Table 1 summarizes the results of our experiments, and there are several interesting findings:

Over-sampling improves performance To examine the impact of over-sampling, we compared the performance of our ensemble classifier with and without over-sampling *i.e.* we compared results under the *STL*, *STL + SMOTE* and *STL + LMOTE* columns. Our results confirm that in general, over-sampling suggestions to obtain a balanced dataset improves the performance (F-1 score & AUC) of our classifiers.

LMOTE performs comparably to SMOTE We compared the performance of SMOTE and LMOTE in the single task settings (*STL + SMOTE* and *STL + LMOTE*) and found that LMOTE performs comparably to SMOTE (and even outperforms it in the `electronics` and `software` domains). LMOTE also has the added advantage of resulting in intelligible samples which can be used to qualitatively analyze and troubleshoot deep learning based systems. For instance, consider the following sample *created* by LMOTE: *"It would be good if oversight bixby developed bug feels wide back content zoom should be an option."* While the suggestion may not be grammatically correct, its constituent phrases are nevertheless semantically sensible.

Multi-task learning outperforms single-task learning We compared the performance of our classifier in single and multi-task settings (*STL + LMOTE* and *MTL + LMOTE*) and found that by multi-task learning improves the performance of our classifier. We qualitatively analysed the single and multi task models, and found many instances where by leveraging domain-specific information the multi task model was able to accurately identify suggestions. For instance, consider the following review: *"Bring a Lan cable and charger for your laptop because house-keeping doesn't provide it."* While the review appears to be an assertion (*non-suggestion*), by predicting its domain (`hotel`), the multi-task model was able to accurately classify it as a suggestion.

References

Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; and Kegelmeyer, W. P. 2002. Smote: Synthetic minority over-sampling technique. *Journal of artificial intelligence research* 16:321–357.

Negi, S. 2019. *Suggestion Mining from Text*. Ph.D. Dissertation, National University of Ireland Galway (NUIG).

Peters, M. E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; and Zettlemoyer, L. 2018. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*.

Research Article

Biodegradation studies of polypropylene/natural fiber composites

Priyanka Luthra[1,2] · Kakkarakkal Kottiyath Vimal[2] · Vishal Goel[2] · Ram Singh[1] · Gurpreet Singh Kapur[2]

Abstract

Polypropylene composites were prepared using natural fibers such as pigeon pea stalk fibers and banana peel. Aerobic biodegradation studies in compost have been carried out as per ASTM with composites prepared where cellulose and polypropylene has been taken as positive and negative reference respectively. Various analytical tools like SEM, TGA, XRD, DSC and Color have been used to study the change after biodegradation in composites. Highest fiber loaded composites i.e. 40 wt% depicted highest biodegradation in comparison to 10 and 20 wt% loading of untreated fiber in polypropylene. In comparison to untreated fiber composites, maximum biodegradation was observed in treated fiber composites with same amount of the fiber loading.

1 Introduction

Many commercially available synthetic polymers and polymer-based composites are catering to human's day to day requirement but at the same time their usage is responsible in contaminating/polluting the environment due to their non-biodegradability [1–3]. There has been shift towards the use of biodegradable polymers like poly(caprolactone), poly(lactic acid), poly(ethylene glycol), polyhydroxyalkanoates etc. in recent times [4–6]. But the market of these kinds of polymers is still facing various kinds of problems like relatively high prices and the lack of infrastructure for effective composting, an extremely critical aspect for biodegradable polymers market success [7].

Another area that has gained attention is the production of natural fiber filled polymer composites due to easy and abundant availability as byproducts in the form of peels, stalks and fibers from the plants sector after harvesting of crops [8–10]. These natural materials/fibers replace the synthetic fillers being added to increase the strength of the commercial synthetic polymers and further making these polymers semi-biodegradable. Lot of work has been done by the researchers which have also provided promising results. Verma et al. [11] worked on biodegradability of coir fiber reinforced epoxy composites under the effect of cow dung by soil burial test (ASTM D 5988). Kiprono [12] worked on degradation behaviors of polypropylene and cellulose blends where biodegradation measurements were also done by soil burial method. Vu et al. [13] in their study used polypropylene as a matrix with rice straw fiber used after alkali and subsequent peroxide treatment. The composites prepared showed soil biodegradability with a weight loss of 3.02 to 23.11 wt% after 50 days of burial. Sahi et al. [14] conducted work on low density polyethylene/alkali treated corn flour composites. The biodegradability of the composites was enhanced with increasing corn flour content in the matrix. Obasi [15] in its work

✉ Ram Singh, ramsingh@dtu.ac.in; ✉ Gurpreet Singh Kapur, kapurgs@indianoil.in | [1]Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India. [2]Indian Oil Corporation Ltd., Research & Development Centre, Sector 13, Faridabad, Haryana 121007, India.

studied corncob flour obtained from a waste product of corn threshing which was blended with high density polyethylene for preparation of composites. Soil burial method was used to study the biodegradation where loss in weight (nine weeks) was observed in case of prepared composites with time thereby confirming biodegradation.

The present study focuses on the biodegradation studies of composites prepared from banana peel and Pigeon pea stalk fibers using polypropylene (PP) as matrix. The fibers used for the study are untreated and treated (alkali and *laccase* enzyme) with and without addition of maleic anhydride grafted polypropylene as compatibilizer. Composites prepared have been subjected to aerobic biodegradation test in compost as per ASTM D 5338. The composites before and after biodegradation have been characterized from visual appearance (Color), morphology (SEM), thermal stability (TGA) and percent crystallinity (XRD and DSC) to observe the change in the properties due to microbial attack.

2 Experimental section

The composites for the biodegradation study have been prepared according to our reported methods (details given in supplementary information of the manuscript) [16–18]. The composition and sample designation are given in Table 1.

Table 1 Sample composition of the composites

Sample designation	PP (wt%)	Fiber content (wt%)	Compatibilizer (PP-g-MA) (wt%)
PP	100	0	0
PB1	90	BP 10	0
PB2	80	BP 20	0
PB4	60	BP 40	0
PB4C1	59	BP 40	1
PB4C1A	59	BP 40	1
PB4C1E	59	BP 40	1
PS1	90	PS 10	0
PS2	80	PS 20	0
PS4	60	PS 40	0
PS4C2	58	PS 40	2
PS4C2A	58	PS 40	2
PS4C2E	58	PS 40	2

P PP; *B* Banana peel; *S* Pigeon pea stalk fibers, *C* Compatibilizer, *A* alkali treated, *E* enzyme treated fibers

3 Biodegradability test

ASTM D 5338–98 has been followed for the laboratory-scale biodegradability testing apparatus which determines the rate and extent of aerobic biodegradation for plastic materials under controlled-composting conditions [19]. The apparatus comprises of three different components. The first component is the carbon dioxide-free air supply for controlling the amount of CO_2 free air to be supplied. The second component is temperature controlling chamber i.e. incubator in which sample along with compost is present in different bioreactors. The third component is carbon dioxide trapping assembly with three conical flasks in series containing CO_2 scrubbing solution i.e. barium hydroxide for each sample.

The test materials are exposed to compost inoculums obtained from municipal solid waste. Cellulose has been taken as a positive reference and PP has been taken as a negative reference. Close monitoring of temperature, aeration and humidity is carried out. The rate of biodegradation is determined by the conversion of carbon present in the sample to CO_2. The percentage of biodegradability is obtained by determining the percentage of carbon in the test substance that is converted to CO_2 during the duration of the test. A total of 12 composting vessels were taken and labelled accordingly. The sample and mature compost were weighed for each composting vessel. A continuous stream of pressurized CO_2-free air is set for supplying to the composting vessels with a fixed aeration rate. The composting vessels were incubated in dark inside the biodegradability testing apparatus for a period of 45 days. The test temperature was 58 °C for 45 days. The exit stream of air was directly analyzed continuously titrimetrically after sorption in dilute alkali. The CO_2 concentrations in the outgoing air is measured daily after the first week for the remainder of the test by titration of $Ba(OH)_2$ scrubbing solution present in the CO_2-trapping apparatus against hydrochloric acid (HCl). The mean value (from the three replicates) of the net CO_2 produced by controlled composting of the test substances was determined by subtracting the mean CO_2 production from the compost. The percent biodegradation was calculated by using Eq. 1

$$\text{Biodegradation} = \frac{CO_{2(test)} - CO_{2(blank)}}{(CO)_{2(Th)}} \qquad (1)$$

where $CO_{2(test)}$ is the cumulative amount of carbon dioxide evolved in each composting reactor containing a tested film sample in grams; $CO_{2(blank)}$ is the cumulative amount of carbon dioxide evolved from the blank reactor in grams; and $CO_{2(Th)}$ = theoretical amount of carbon dioxide of the test sample in grams.

4 Results and discussion

4.1 Biodegradation studies

In the process of biodegradation, the microorganisms start utilizing the polymer surface as its food and grow which is confirmed by the formation of a layer of biofilm (deposition of microorganisms and their polysaccharides) on the polymer surface [20]. The different factors affecting the biodegradation are polymer characteristics such as their tacticity, crystallinity, molecular weight, type of functional groups, type of organism and nature of pre-treatment [21]. The test is considered as unacceptable if 70% of the cellulose is not degraded in 45 days. Total organic carbon content in percentage for cellulose as a reference, pure PP as a negative reference and composites as samples have been summarized in Table 2.

Table 2 Total organic carbon content (%) in positive, negative and test samples

S. no	Sample code	TOC (%)	S. no.	Sample code	TOC (%)
1	Cellulose (positive)	42.62	8	PB4C1E	80.20
2	PP (negative)	94.18	9	PS1	90.39
3	PB1	92.00	10	PS2	88.22
4	PB2	90.00	11	PS4	85.34
5	PB4	86.00	12	PS4C2	80.78
6	PB4C1	84.00	13	PS4C2A	78.15
7	PB4C1A	82.38	14	PS4C2E	75.64

The effect of microorganisms on composites degradation of composites prepared with varied natural filler content, compatibilizer and treatment have been studied. Figure 1 depicts the CO_2 generation for cellulose reference, PP neat, PP/Banana peel and PP/Pigeon pea stalk composites. Cellulose, used as a positive reference showed 80% biodegradation when tested for 45 days which conforms the validity of the test conducted as per ASTM refereed for the test. It was observed that with the increase in natural filler content in PP matrix, the generation of carbon dioxide (CO_2) increased. The composites PB1 and PB4 released 0.021 and 0.032 g/g CO_2 respectively whereas PS1 and PS4 released 0.023 and 0.036 g/g respectively. However, no CO_2 generation was observed for neat PP which further confirms that the natural fillers act as a substrate or feed for the growth of microorganisms. This is in agreement with the literature where high natural filler content leads to fastening of biodegradation due to their susceptibility towards water uptake leading to crack generation of the polymer surface that provide easy attack by microorganism [22].

During biodegradation in compost, some factors such as temperature, water and air, accelerate the biodegradation rate [23]. It is an accelerated/simulated test under defined conditions. During the biodegradation cycle three phases are observed. First is the lag phase during which the microbial population adapts to the available carbon test substrate. After that comes the biodegradation phase during which the adapted microbial population begins to use the carbon substrate for its growth which in turn is confirmed by the conversion of the carbon in the test material to CO_2. The final stage is the plateau phase
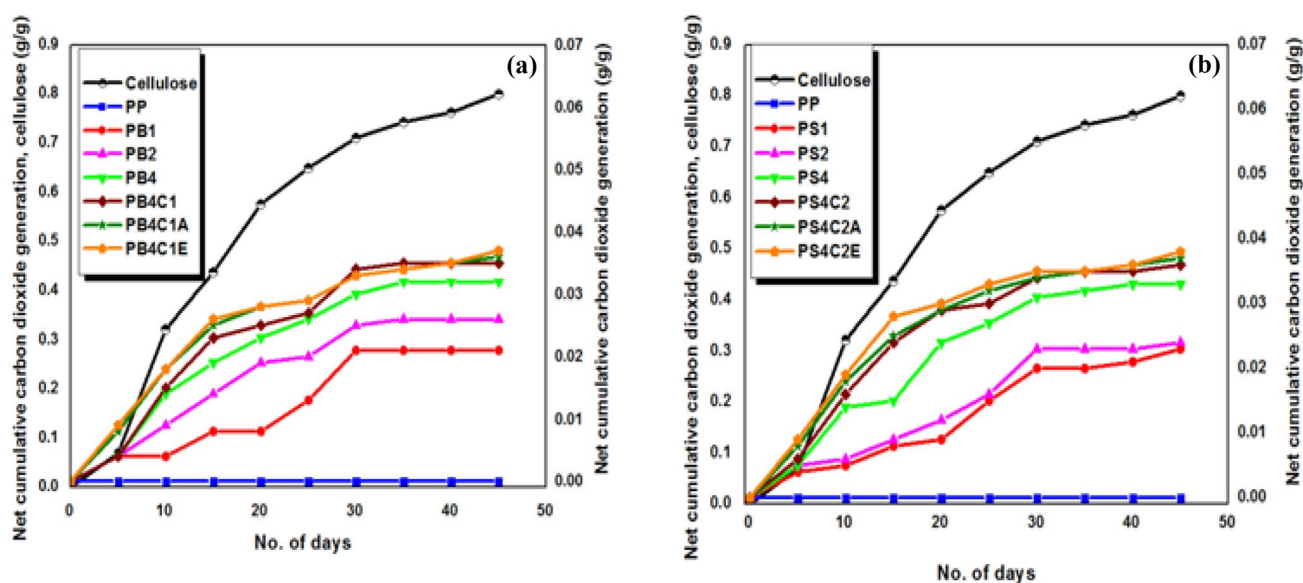

Fig. 1 Net cumulative carbon dioxide per gram of **a** PP/Banana peel and **b** PP/Pigeon pea stalk composites with cellulose

when the carbon test substrate is mostly consumed by microbial population [24]. Biodegradation of PP/Banana peel and PP/Pigeon pea stalk composites in comparison to neat PP ranged from 0 to 3.7 and 0 to 3.8% respectively as can be observed from Fig. 2 against the standard cellulose showing 80 wt% degradation. The observation is in line with the work on biodegradation studies carried out on coffee-based composites [25]. Cellulose have a higher rate of degradation as the hydrophilic hydroxyl group along with polysaccharide chain makes it more susceptible to microbial attack [26]. With the addition of coupling agent and alkali and enzyme treated natural fibers, an increase in biodegradation was observed in composites prepared. This further confirms the reported studies that untreated natural fibers are more resistant to attack of microorganisms due to presence of lignin and hemicelluloses with greater amount than the treated fibers, as lignin act as an adhesion agent in the plant tissue [27].

4.2 Thermal gravimetric analysis (TGA)

TGA of composites has been studied as a function of % weight loss with the increase in temperature. The study has been carried out to analyze the composites thermal stability by TGA under nitrogen atmosphere before and after 45 days of biodegradation. It has been observed that there is no change in the thermal stability of neat PP which confirms that there has been no biodegradation occurred in 45 days, tenure of the test. However, in case of composites, two step degradation has been observed, first step is due to the natural fiber which consist of cellulosic (cellulose and hemicelluloses) and non-cellulosic materials (lignin) and second due to the polymer matrix part in the composite. Degradation in thermal stability of the composites is observed with time (0 day to 45 day)

when exposed to the compost environment. Tables 3 and 4 show the thermal degradation of PP/Banana peel and PP/Pigeon pea stalk composites respectively which conforms 5% (T5) and 10% (T10) degradation occur at a lower temperature for the composites that have been exposed to compost for 45 days vis-a-vis comparison to 0 day of exposure. This further confirms that some microorganism activity has been initiated which attacks on the fiber part of the composites thus degrading it. TGA scan overlay comparison for PP/Banana peel composites (PB4, PB4C1, PB4C1A and PB4C1E) before and after biodegradation have been depicted in Fig. 3 whereas that of PP/Pigeon pea stalk composites (PS4, PS4C2, PS4C2A and PS4C2E) before and after biodegradation have been depicted in Fig. 4.

Table 3 The 5%, 10% and peak degradation temperature of PP/banana peel composites before and after biodegradation

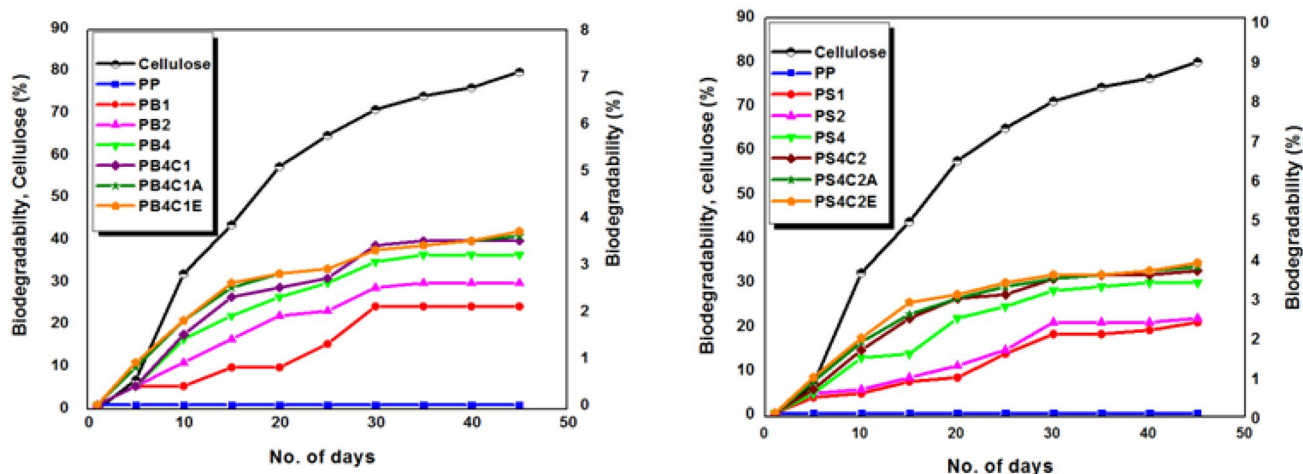| Sample	T5 (°C)	T10 (°C)	T_{max} (°C)
PP-0	393.0	409.9	443.5
PP-45	392.0	409.0	443.6
PB1-0	336.4	410.3	328.6; 453.1
PB1-45	258.6	303.2	240.2; 450.6
PB2-0	309.6	390.1	298.1; 446.5
PB2-45	259.8	312.0	282.3; 449.1
PB4-0	243.1	283.4	341.5; 450.8
PB4-45	235.1	269.8	286.7; 452.8
PB4C1-0	238.9	293.9	329.2; 439.1
PB4C1-45	248.2	281.4	287.4; 421.6
PB4C1A-0	243.7	288.5	325.9; 451.7
PB4C1A-45	240.0	287.7	329.2; 453.4
PB4C1E-0	257.1	299.2	324.7; 452.4
PB4C1E-45	256.0	291.7	320.0; 440.6

Fig. 2   % Biodegradation of PP/Banana peel and PP/Pigeon pea stalk composites with cellulose

Table 4 The 5%, 10% and peak degradation temperature of PP/Pigeon pea stalk composites before and after biodegradation

Sample detail	T5 (°C)	T10 (°C)	T_{max} (°C)
PS1-0	300.0	352.0	351.7; 449.2
PS1-45	295.2	339.4	328.6; 453.1
PS2-0	272.5	324.7	344.5; 449.9
PS2-45	270.0	314.5	343.0; 407.5
PS4-0	250.4	286.5	334.9; 451.9
PS4-45	234.7	282.9	342.2; 449.7
PS4C2-0	255.3	290.5	347.2; 444.9
PS4C2-45	251.7	301.4	332.7; 446.3
PS4C2A-0	258.4	294.2	354.2; 450.2
PS4C2A-45	257.9	308.6	333.3; 450.5
PS4C2E-0	257.4	288.3	344.4; 427.8
PS4C2E-45	253.6	291.2	348.6; 450.6

4.3 Crystallization studies of polypropylene/natural fiber composites (before and after biodegradation)

4.3.1 Differential scanning calorimetry (DSC)

The change in composite properties with respect to change in crystallization temperature (T_c), melting temperature (T_m) and % crystallinity after biodegradation have been studied with DSC. For T_m and % crystallinity, the second heating curves have been taken for the study. For PP/PS composites, not much change in the T_m is observed after biodegradation. An increase in the % crystallinity and crystallization temperature is observed for the composites prepared and exposed for 45 days due to biodegradation test which thereby confirms that not much change has been initiated by the microbial attack on the matrix part of the composites which being synthetic in nature.

The melting point of composites is in the range of 161.8–163.9 °C before and after biodegradation. Therefore,
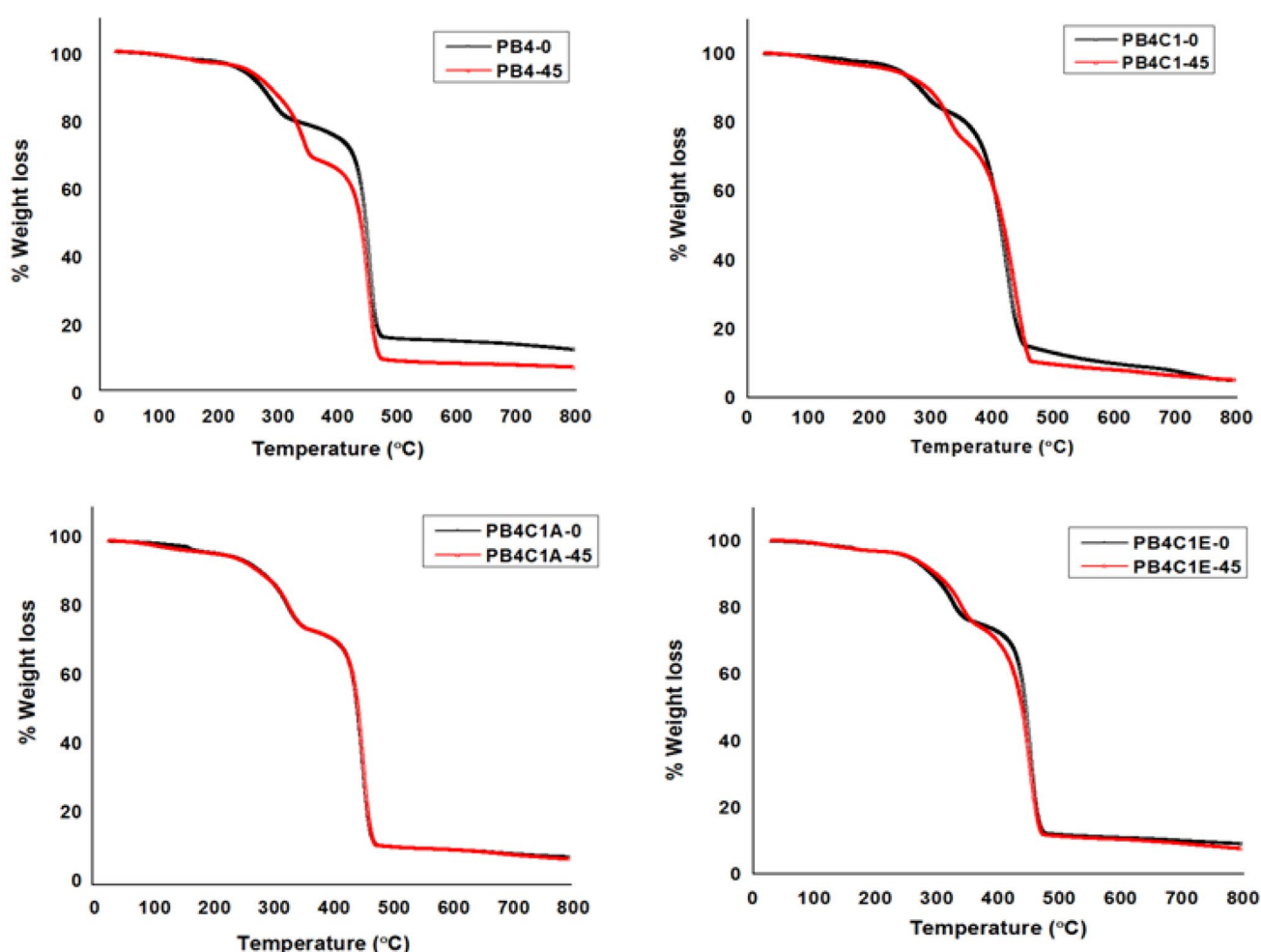
Fig. 3 Overlay comparison of TGA curve of PB4, PB4C1, PB4C1A and PB4C1E before and after biodegradation
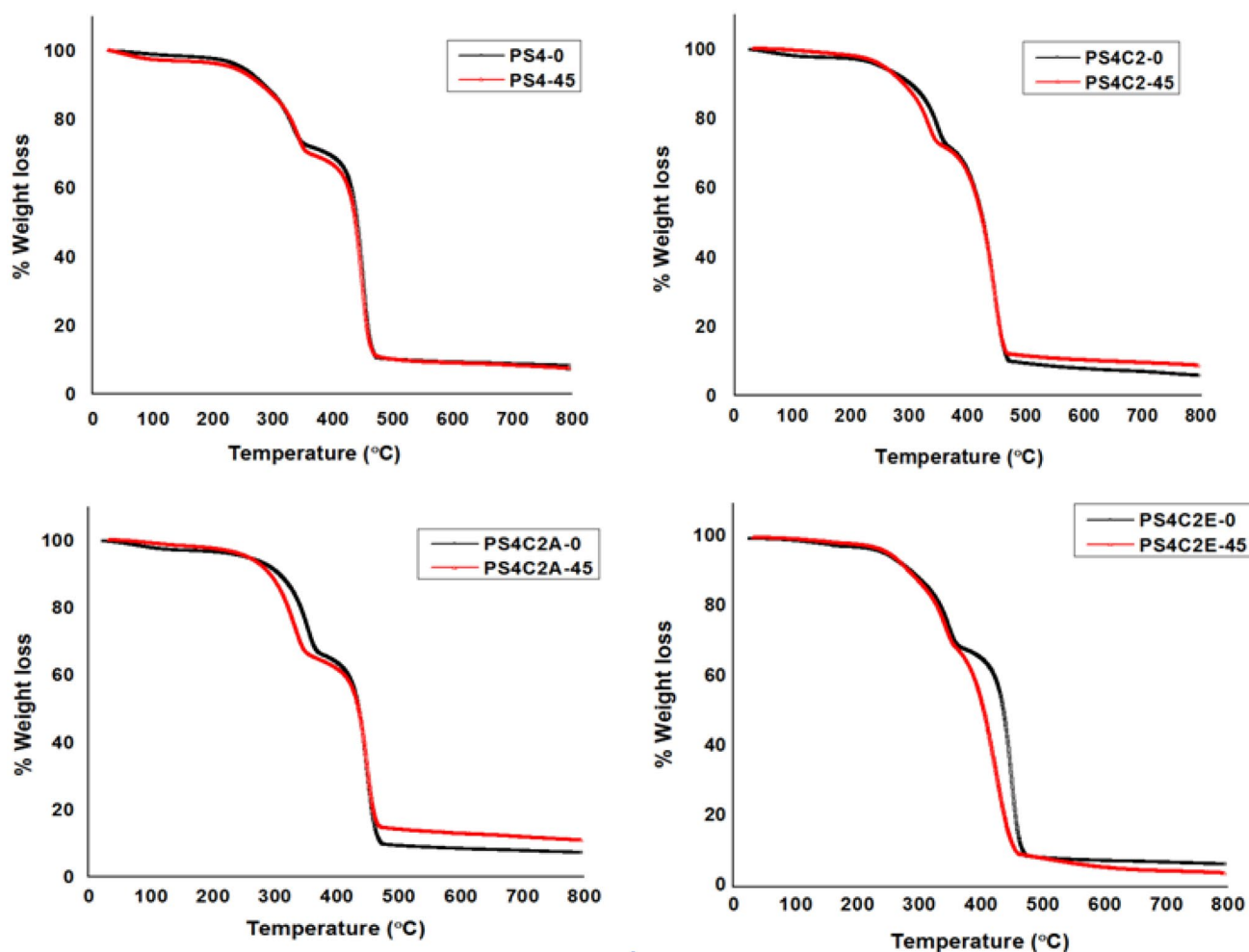
Fig. 4 Overlay comparison of TGA curve of PS4, PS4C2, PS4C2A and PS4C2E before and after biodegradation

it is the natural fiber in the polymer composites, which is the amorphous portion and is easily attacked by the microorganisms, and therefore results in increase in the overall crystallinity of the polymer composites.

The obtained result is in agreement with the literature [28]. The same type of trend is observed for PP/BP composites also. In case of PP/Pigeon pea stalk composites, the percent increase in crystallinity is higher than in case of PP/Banana peel composites as observed from the values tabulated in Table 5. This thereby confirms the slightly higher percent biodegradation of PP/Pigeon pea stalk composites than the PP/Banana peel composites as observed from the biodegradation pattern of the composites studied. In Table 5, after biodegradation of the sample PS4C2, there is an enhancement in % crystallinity by 45%. It could be due to higher fiber length of pigeon pea stalk than the banana peel fiber which leads to increase in biodegradation.

4.3.2 X-ray diffraction (XRD)

The XRD plots of PP/BP and PP/PS composites before and after the biodegradation has been shown in Fig. 5. The plots show the wide-angle X-ray diffraction patterns of untreated and treated composites. Five peaks in the 2θ range of 10–30° at (110), (040), (130), (111) and (041) lattice plane in XRD scan for PP were observed which confirms to the monoclinic α form without any peculiar peak corresponding to beta (β) and gamma (γ) form of PP. A peak that conforms to beta crystalline phase at 2θ = 16.30 (degrees) represents the (300) diffraction plane in the composites prepared. Considering the peak areas obtained by integration in relation to crystallinity; the larger the area of the peaks, the more crystalline the material.

The reason for a slight increase in crystallinity after biodegradation could be due to changes in amorphous part (low density area) which is more prone to biodegradation first than the crystalline part of the composite.

Table 5 DSC results of composites before and after biodegradation test for 45 days

Sample code	T_c (°C)		T_m (°C)		% Crystallinity	
	Before	After	Before	After	Before	After
PP	117.63	117.71	161.87	163.90	48	51
PB1	116.02	119.46	162.11	163.18	47	49
PB2	115.63	118.32	162.11	163.13	46	43
PB4	115.05	116.50	161.92	162.10	43	39
PB4C1	118.53	119.39	163.44	163.62	53	54
PB4C1A	121.44	120.35	161.38	163.89	55	49
PB4C1E	120.48	117.67	163.41	163.71	52	53
PS1	119.27	119.46	163.10	163.18	46	49
PS2	118.70	118.32	161.96	163.13	47	43
PS4	116.24	116.50	161.84	162.10	44	51
PS4C2	118.98	119.39	162.99	163.62	44	64
PS4C2A	113.80	120.35	163.15	163.89	44	49
PS4C2E	111.97	117.67	162.44	163.71	47	51
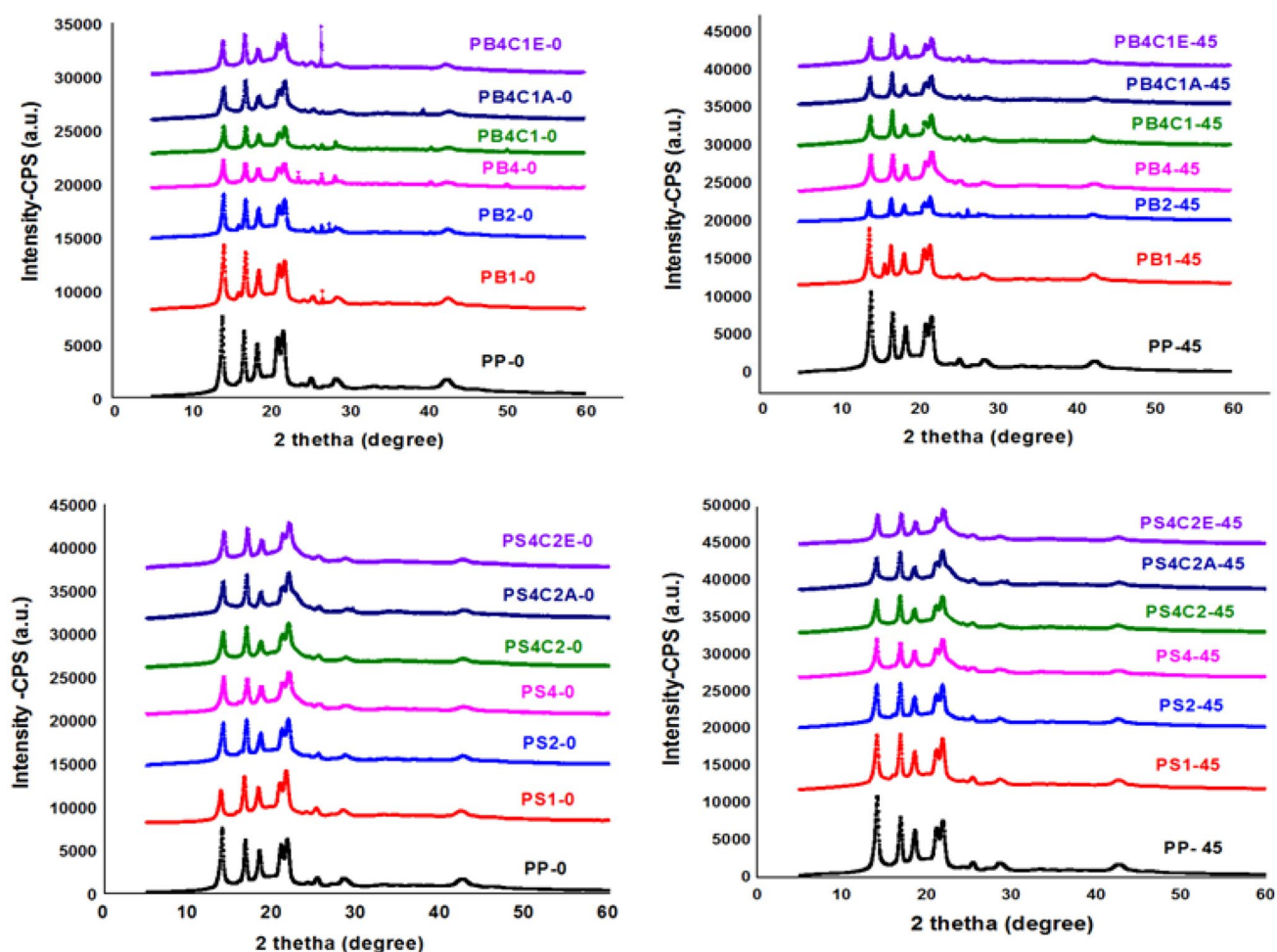
Fig. 5 XRD plots PP/Banana peel and PP/Pigeon pea stalk composites before and after biodegradation

Similar kind of results have been reported by Huanga et al. in their work on study on degradation of composite material polybutylenes succinate/ polycaprolactone [29]. The overlay comparison of XRD scans before and after biodegradation of 45 days for 40 wt% untreated and 40 wt% alkali treated fiber for both PP/Banana peel and PP/Pigeon pea stalk composites have been shown in Fig. 6.

4.4 Scanning Electron Microscope (SEM)

SEM characterization tool is used for the examination of the changes in the morphology before and after biodegradation of the composites. Samples morphology can be observed under high resolution. Cryofracture samples were analyzed microscopically for the changes in the morphology. The figures from SEM (Fig. 7) illustrate the morphology of PP/Banana peel and PP/Pigeon pea stalk composites respectively before and after biodegradation. Formation of cracks, voids have been observed in the

samples subjected to biodegradation in compost environment in comparison to the control samples for the composites prepared. These cracks and voids have initiated easy attack of the microbes and thereby leading to degradation in the material.

4.5 Color measurement

Color is an important parameter used in the assessing the quality of polymer by visual appearance. It is the first physical way to assess the deterioration in the visual appearance which could be the first step of indication of the microbial attack. One of the ways of expressing the color of an object is L, a, b color space which is a 3-dimensional color space. It consists of three parameters L* (lightness) axis, a* (red-green) axis, b* (blue-yellow) axis. It has been observed that after the exposure of the samples to the biodegradation the color parameters have increased in comparison to 0-day data. Increase in L* (Fig. 8) value
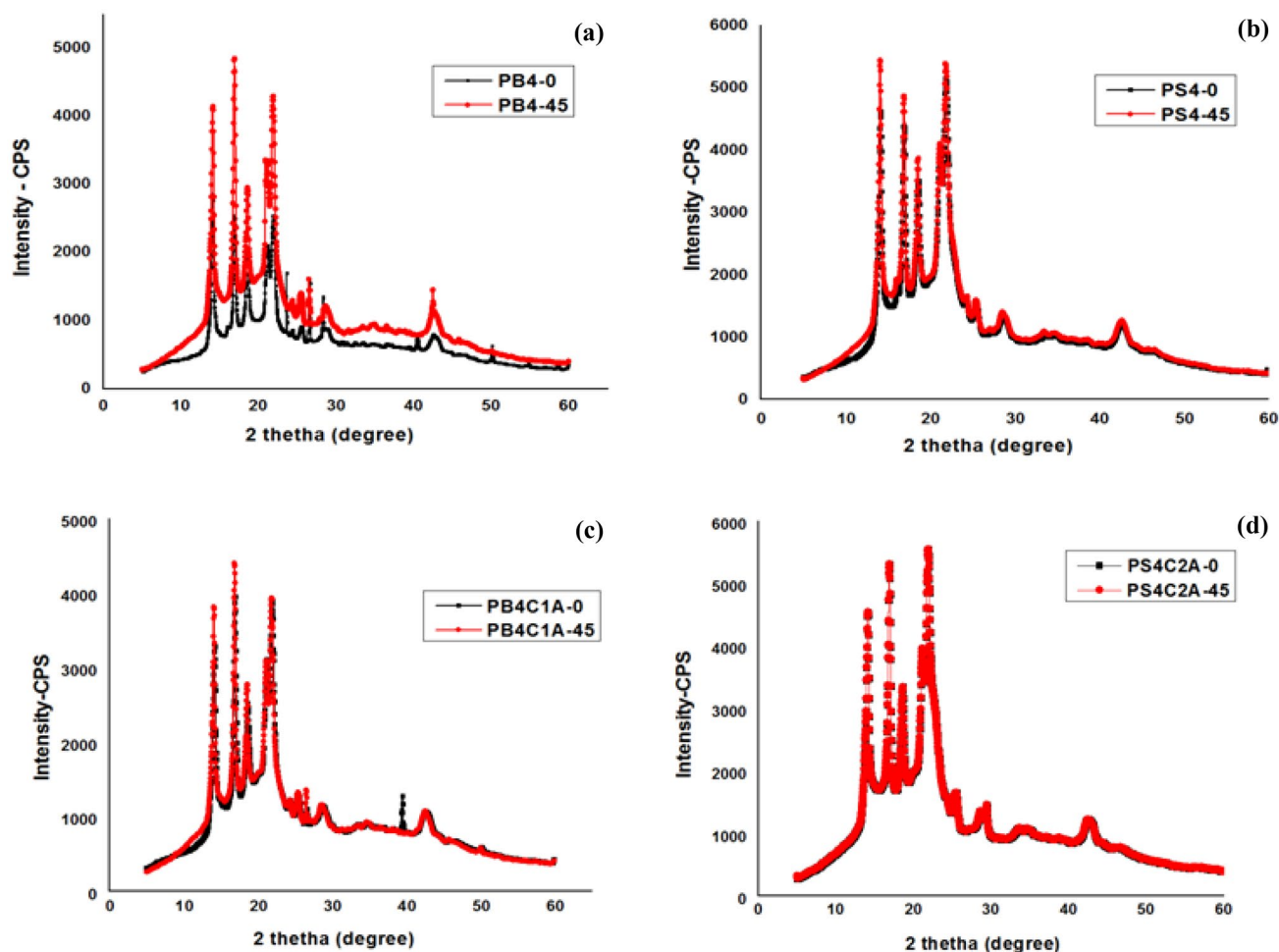

Fig. 6 Comparison of XRD plots of **a** PB4 and **c** PB4C1A (PP/BP composites) and **b** PS4 and **d** PS4C2A (PP/PS composites) before and after biodegradation
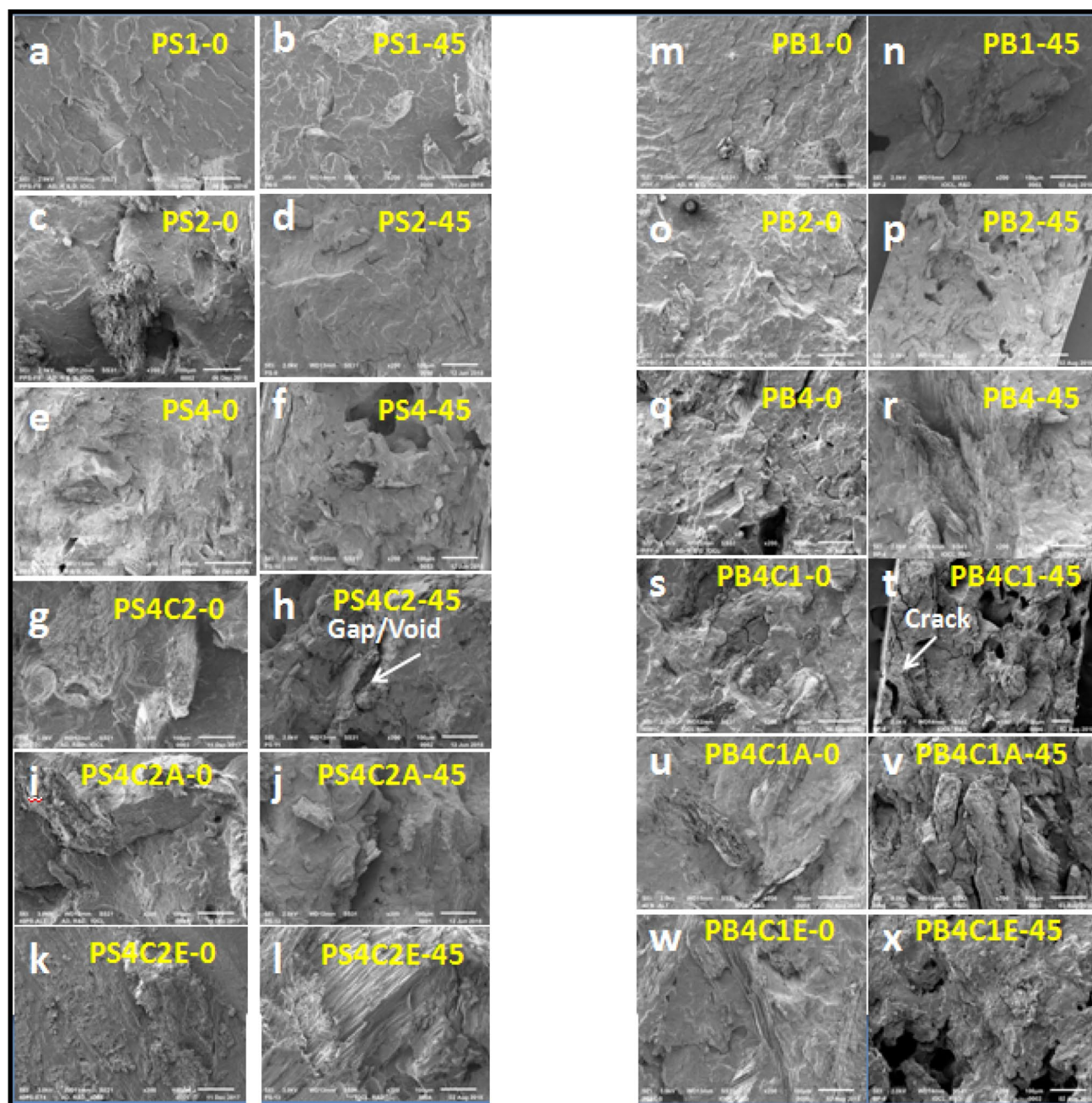
Fig. 7 SEM micrographs PP/Pigeon pea stalk and PP/Banana peel composites before and after biodegradation

after 45 days confirms the sample becoming light in visual appearance. Similar way the a* and b* increased after 45 days thereby confirming the increase in green and yellow tinge in the samples. ΔE which summarizes the change in all the three parameters has also shown an increase after 45 days of exposure of samples to biodegradation. Before biodegradation the samples exhibited dark color, which is expressed on the basis of Lightness /Darkness (L*) of the specimens. Changes in the chemical structure of lignocellulosic complex of the natural fiber of the composites

upon biodegradation has led to change in the color of the composites due to the presence of lignocellulosic material. Same kind of behavior of surface color changes has been observed by Butylina et al. [30] during their work with natural fiber filler and biobased biodegradable polymer. It has also been observed from the results that the color change is more in case of untreated fiber composites (PS1, PS2, PS4, PB1, PB2 and PB4) whereas the treated fiber the change in the color is much less which could be due to removal of lignin by the pretreatment procedure of the

Fig. 8 L* comparison of PP/Banana peel and PP/Pigeon pea stalk composites before and after biodegradation
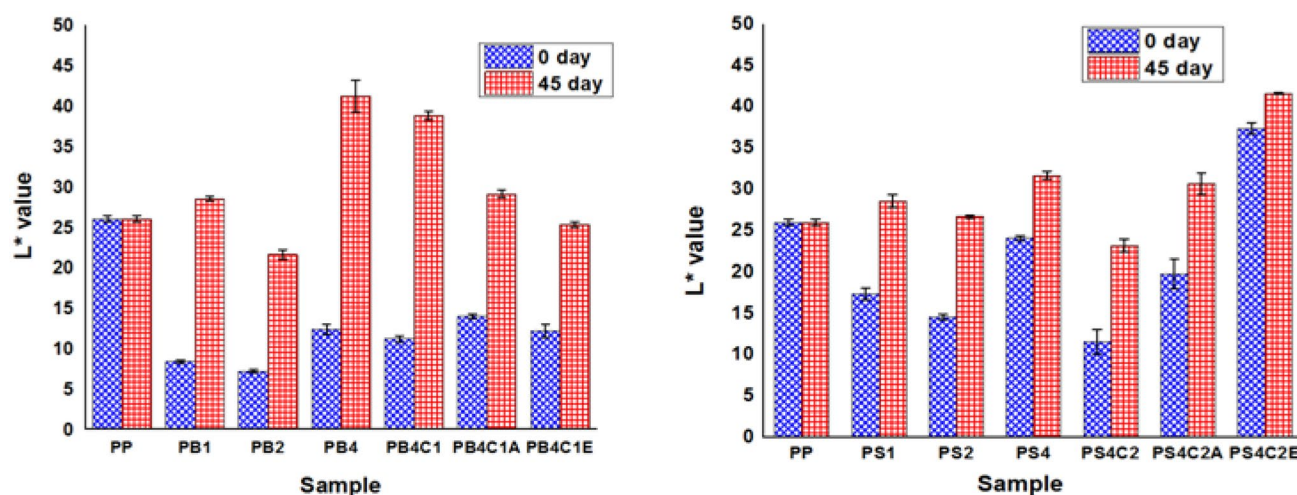
natural fiber carried out before the preparation of the composites. The values of color parameters of the prepared composites before and after biodegradation have been tabulated as Tables 6 and 7 for PP/Banana peel and PP/ Pigeon pea stalk composites respectively. Change in color as observed visually has been complied as Fig. 9 and 10 for PP/BP and PP/PS composites.

5 Conclusion

In the recent work, the biodegradation study of polypropylene/ natural fiber composite has been done where PP is used as a matrix and banana peel and pigeon pea stalk

fibers have been used as reinforcing agent. These composites i.e. PP/Banana peel and PP/Pigeon pea stalk have been compatibilized using maleic anhydride PP-g-MA. The samples were subjected to biodegradation test (compost) and were characterized before and after the test for using various techniques.

- Biodegradation of composites have been observed for PP/Banana peel and PP/Pigeon pea stalk composites prepared in comparison to the standard cellulose which was taken as a reference. Maximum degradation was observed with composites prepared with higher loading i.e. 40 wt% of untreated fiber (Banana peel and Pigeon pea stalk). Treated fiber composites (alkali and enzyme) showed more biodegradation vis a vis comparison to untreated fiber composites which thereby

Table 6 Color measurement for PP/banana peel composites before and after biodegradation

Sample detail	L*	a*	b*	ΔE
PP- 0 day	26.0 ± 0.35	-0.79 ± 0.02	-5.3 ± 0.03	0.01 ± 0.001
PP-45 day	25.5 ± 0.44	0.76 ± 0.04	5.4 ± 0.02	0.03 ± 0.01
PB1-0 day	8.4 ± 0.2	3.9 ± 0.1	4.5 ± 0.06	0.02 ± 0.02
PB1-45 day	28.5 ± 0.3	4.7 ± 0.3	10.0 ± 0.16	0.05 ± 0.01
PB2-0 day	7.2 ± 0.2	2.9 ± 0.2	3.5 ± 0.1	0.2 ± 0.03
PB2-45 day	21.6 ± 0.6	5.1 ± 0.1	9.9 ± 0.12	0.5 ± 0.5
PB4-0 day	12.4 ± 0.6	3.7 ± 0.2	6.3 ± 0.2	0.4 ± 0.2
PB4-45 day	41.2 ± 2.0	8.3 ± 0.4	18.3 ± 0.5	2.6 ± 1.3
PB4C1-0 day	11.2 ± 0.4	2.8 ± 0.1	3.6 ± 0.2	0.3 ± 0.2
PB4C1-45 day	38.8 ± 0.5	5.4 ± 0.1	11.8 ± 0.14	0.4 ± 0.2
PB4C1A-0 day	14.0 ± 0.3	4.5 ± 0.3	8.4 ± 0.8	0.6 ± 0.3
PB4C1A -45 day	29.1 ± 0.5	5.8 ± 0.1	11.0 ± 0.2	0.7 ± 0.14
PB4C1E -0 day	12.2 ± 0.8	3.8 ± 0.2	6.9 ± 0.2	0.5 ± 0.12
PB4C1E -45 day	25.3 ± 0.3	5.4 ± 0.04	11.0 ± 0.1	0.2 ± 0.3

Table 7 Color measurement for PP/Pigeon pea stalk composites before and after biodegradation

Sample detail	L*	a*	b*	ΔE
PS1-0 day	17.4 ± 0.7	9.5 ± 0.5	15.0 ± 0.7	0.9 ± 0.3
PS1-45 day	28.6 ± 0.8	7.2 ± 0.3	13.0 ± 0.8	0.03 ± 0.01
PS2-0 day	14.6 ± 0.3	7.4 ± 0.5	11.1 ± 0.8	0.7 ± 0.1
PS2-45 day	26.7 ± 0.2	7.5 ± 0.2	13.7 ± 0.4	2.5 ± 2.9
PS4-0 day	24.1 ± 0.3	7.5 ± 0.05	13.0 ± 0.06	0.3 ± 0.04
PS4-45 day	31.7 ± 0.5	5.8 ± 0.6	11.4 ± 0.7	0.7 ± 0.5
PS4C2-0 day	11.6 ± 1.5	5.9 ± 0.7	7.3 ± 1.4	1.1 ± 0.9
PS4C2-45 day	23.2 ± 0.8	6.1 ± 0.2	10.6 ± 0.8	0.8 ± 1.0
PS4C2A-0 day	19.8 ± 1.8	11.9 ± 0.1	18.4 ± 0.3	1.6 ± 1.3
PS4C2A-45 day	30.7 ± 1.3	8.4 ± 0.3	16.8 ± 0.13	1.8 ± 0.1
PS4C2E-0 day	37.4 ± 0.7	11.4 ± 0.4	21.3 ± 1.0	1.1 ± 0.1
PS4C2E-45 day	41.6 ± 0.1	10.7 ± 0.7	21.7 ± 1.6	1.4 ± 0.3
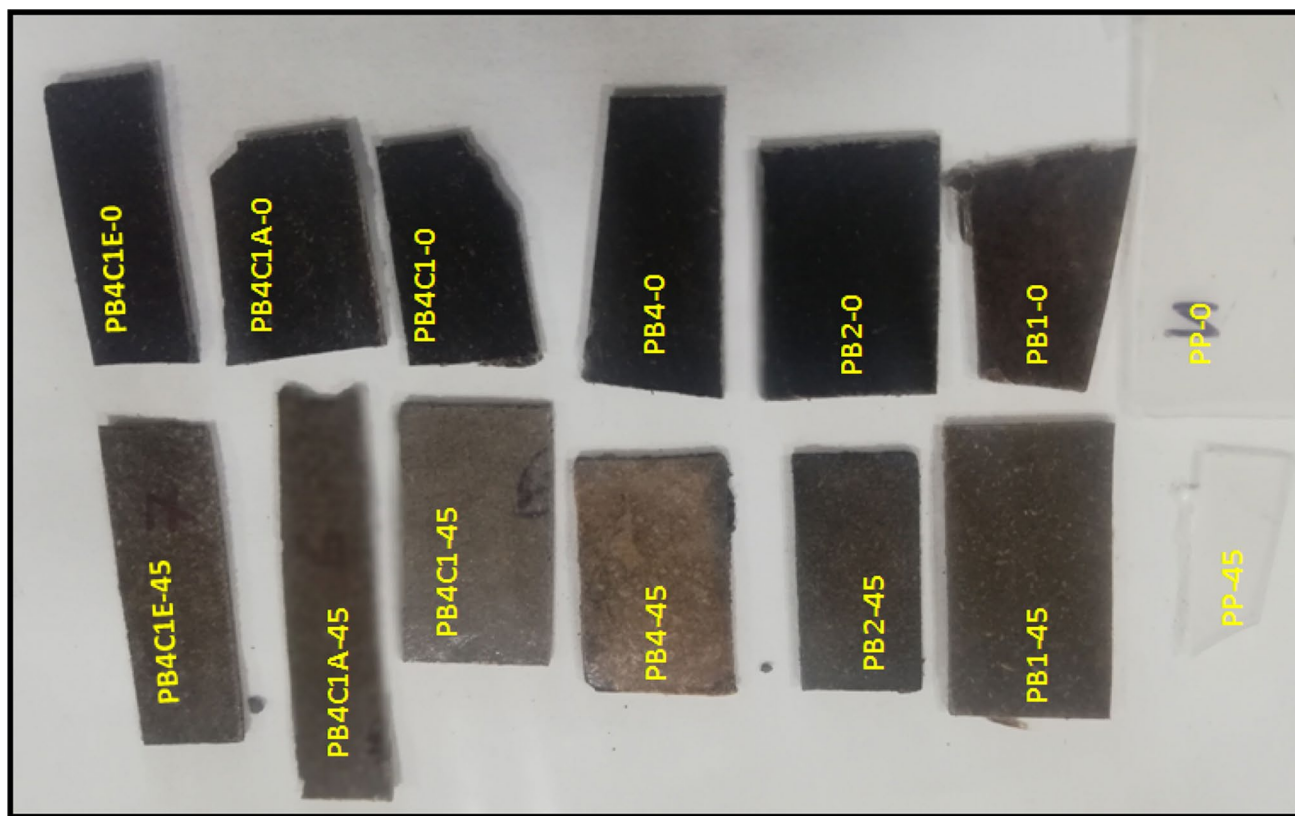
Fig. 9 Color change of PP/Banana peel composites before and after biodegradation
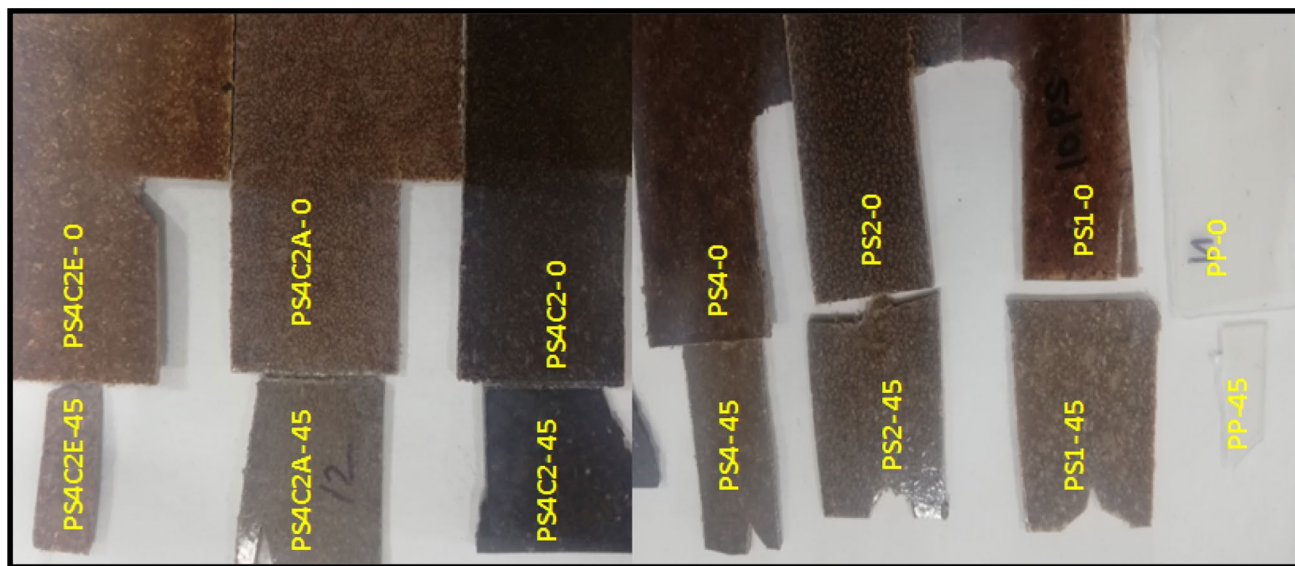
Fig. 10 Color change of PP/Pigeon pea stalk composites before and after biodegradation

confirms that the removal of lignin after treatment which provides a hindrance to the attack of the micro-organisms thereby slowing the rate of biodegradation.

Escalation in percent crystallinity and crystallization temperature is observed by DSC and XRD techniques. However, not much change in the melting point is seen

by DSC for the composites prepared and exposed to 45 days to biodegradation test.

- TGA data concludes that on exposure to biodegradation test for 45 days, PP/Banana peel and PP/Pigeon pea stalk composites weight loss scan, showed a steep/fast thermal degradation which thereby confirms the microbial activity.
- SEM results showed a change in the morphology of the composites prepared. Cracks and voids observed after biodegradation in the composites.
- Visual appearance of the composites also changed which is characterized by the measurement of color which lightened after the samples were subjected to biodegradation in compost.

This thereby confirms that the natural fibers (Banana peel and Pigeon pea stalk) used in the preparation of the composites are prone to microbial attack and can be used for the preparation of polymer composites thereby replacing the synthetic fiber, which being non-biodegradable in nature. These natural fibers filled polymer composites is a suitable alternate to synthetic fiber filled polymer composites in various sectors i.e. automotive, construction etc.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Sidorov OF, Shishov MG, Deryugin AA, Sidelnikov Y (2016) Environmental impact of polymer waste disposal. Coke Chem 59:117–121
2. Rhodes CJ (2018) Plastic pollution and potential solutions. Sci Prog 101:207–260
3. Shastri VP (2003) Non-degradable biocompatible polymers in medicine: past, present and future. Curr Pharmaceut Biotechnol 4:331–337
4. Akos IN, Wyasu G, Ladan Z (2014) Effect of fiber load and compatibilization on biodegradation of poly(ε-caprolactone)/poly(lactic acid) composites. Int Res J Mater Sci Eng 1:2–11
5. Leja K, Lewandowicz G (2010) Polymer biodegradation and biodegradable polymers—a review. Pol J Environ Stud 19:255–266
6. Vroman I, Tighzert L (2009) Biodegradable polymers. Materials (Basel) 2:307–344
7. Rujnić-Sokele M, Pilipović A (2017) Challenges and opportunities of biodegradable plastics: a mini review. Waste Manag Res 35:132–140
8. Dixit S, Goel R, Dubey A, Shivhare PR, Bhalavi T (2017) Natural fibre reinforced composite materials—a review. Polym Renew Res 8:71–78
9. Pandey P, Bajwa S, Bajwa D (2018) Fiber from DDGS and corn grain as alternative fillers in polymer composites with high density polyethylene from bio-based and petroleum sources. J Polym Environ 26:2311–2322
10. Potluri R (2019) Natural fiber-based hybrid bio-composites: processing, characterization, and applications. In: Muthu S (eds) Green composites Textile Science and Clothing Technology. Springer: Singapore.
11. Verma P, Dixit S, Asokan P (2013) Biodegradation of coir/epoxy composites. Appl Polym Compos 1:75–84
12. Kiprono KP (2013) Mechanical, diffusion and degradation behavior of polypropylene and cellulose blends. School of Pure and Applied Sciences of Kenyatta University, Nairobi City
13. Vu ND, Tran HT, Nguyen TD (2018) Characterization of polypropylene green composites reinforced by cellulose fibers extracted from rice straw. Int J Polym Sci article ID 1813847.
14. Sahi S, Djidjelli H, Boukerrou A (2015) Biodegradation study of bio-corn flour filled low density polyethylene composites assessed by natural soil. J Polym Eng 36:245–252
15. Obasi HC (2012) Studies on biodegradability and mechanical properties of high density polyethylene/corncob flour-based composites. Int J Sci Eng Res 3(8):1–14
16. Luthra P, Singh R, Kapur GS (2019) Effect of chemical treatments on pigeon pea stalk fibers. Eur J Sci. https://doi.org/10.29198/ejs1901
17. Luthra P, Singh R, Kapur GS (2019) Development of polypropylene/banana peel (treated and untreated) composites with and without compatibilizer and their studies. Mat Res Exp 6: article number 095313
18. Luthra P, Singh R, Kapur GS (2019) Preparation and studies of pigeon pea stalk/polypropylene composites with and without compatibilizer. Polym Polym Compos 27:337–346
19. ASTM D 5338: Standard Test Method for determining aerobic biodegradation of plastic materials under controlled composting conditions incorporating thermophillic conditions
20. Arkatkar A, Arutchelvi J, Sudhakar M, Bhaduri S, Uppara PV, Doble M (2009) Approaches to enhance the biodegradation of polyolefins. Open Env Eng J 2:68–80
21. Muthukumar A, Veerappapillai S (2015) Biodegradation of plastics—a brief review. Int J Pharm Sci Rev Res 31:204–209
22. Tazi M, Erchiqui F, Kaddami H, Bouazara M, Poaty B (2015) Evaluation of mechanical properties and durability performance of HDPE-wood composites. AIP Conf Proc 1664:150001
23. Siddiquee KM, Helali MdM (2014) Effects of fiber length and fiber ratio on the biodegradability of jute polymer composites. Int J Sci Eng Res 2:64–69
24. Song JH, Murphy RJ, Narayan R, Davies GBH (2009) Biodegradable and compostable alternatives to plastics. Philos Trans R Soc B Biol Sci 364:2127–2139
25. Moustafa H, Guizani C, Dupont C, Martin V, Jeguiri M, Dufresne A (2017) Utilization of torrefied coffee grounds as reinforcing agent to produce high-quality biodegradable PBAT composites for food packaging applications. ACS Sust Chem Eng 5:1906–1916
26. Lynd LR, Weimer PJ, Zyl WHV, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577
27. Costa Junior AE, Barreto ACH, Rosa DS, Maia FJN, Lomonaco D, Mazzetto SE (2015) Thermal and mechanical properties of biocomposites based on a cashew nut shell liquid matrix reinforced with bamboo fibers. J Compos Mat 49:2203–2215
28. Roy SB, Shit SC, Sengupta RA, Shukla PR (2015) Biodegradability studies of bio-composites of polypropylene reinforced by potato starch. Int J Innov Res Sci Eng Technol 4:1120–1130

29. Huanga J, Cuia C, Yana G, Huanaga J, Zhang M (2016) A study on degradation of composite material PBS/PCL. Polym Polym Compos 24:143–148
30. Butylina S, Hyvärinen KT (2012) A study of surface changes of wood-polypropylene composites as the result of exterior weathering. Polym Degrad Stab 97:337–345

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

Clustering Glossary Terms Extracted from Large-Sized Software Requirements using FastText

Kushagra Bhatia
kushagra1198@gmail.com
EE Department
Delhi Technological University
Delhi, India

Siba Mishra
sibam@iiserb.ac.in
EECS Department
IISER Bhopal
Bhopal, Madhya Pradesh, India

Arpit Sharma
arpit@iiserb.ac.in
EECS Department
IISER Bhopal
Bhopal, Madhya Pradesh, India

ABSTRACT

Specialized terms used in the requirements document should be defined in a glossary. We propose a technique for automated extraction and clustering of glossary terms from large-sized requirements documents. We use text chunking combined with WordNet removal to extract candidate glossary terms. Next, we apply a state-of-the art neural word embeddings model for clustering glossary terms based on semantic similarity measures. Word embeddings are capable of capturing the context of a word and compute its semantic similarity relation with other words used in a document. Its use for clustering ensures that terms that are used in similar ways belong to the same cluster. We apply our technique to the CrowdRE dataset, which is a large-sized dataset with around 3000 crowd-generated requirements for smart home applications. To measure the effectiveness of our extraction and clustering technique we manually extract and cluster the glossary terms from CrowdRE dataset and use it for computing precision, recall and coverage. Results indicate that our approach can be very useful for extracting and clustering of glossary terms from a large body of requirements.

CCS CONCEPTS

• **Software → Software Engineering**; *Requirements/Specifications*; • **Computing Methodologies → Artificial Intelligence**; *Natural Language Processing*; • **Information Systems → Information Storage and Retrieval**; *Linguistic Processing*.

KEYWORDS

Requirements Engineering, Natural Language Processing, Word Embeddings, Clustering, Glossary, FastText.

1 INTRODUCTION

Requirements are descriptions of how the system should behave, or of a system property or attribute [10, 24]. They are the basis for every project, defining what the stakeholders in a potential new system need from it, and also what the system must do in order to satisfy that need. A requirements document is used to communicate software requirements to customers, system users, managers and system developers. All subsequent steps in software development are influenced by the quality of requirements document. A major cause of poor quality requirements document is that the stakeholders involved in the development process may have different interpretations of technical terms used in the document. For example, domain experts may use specialized jargon, whose meaning may not be clear to readers with a different technical background and domain expertise. In order to avoid these issues and to improve the understandability of requirements, it is necessary that all stakeholders of the development process share the same understanding of the terminology used. Specialized terms used in the requirements document should therefore be defined in a glossary. A glossary defines specialized or technical terms and abbreviations which are specific to an application domain. For example, if the system is concerned with health care, it would include terms like "hospitalization", "prescription drugs", "physician", "hospital outpatient care", "durable medical equipment", "emergency services" etc. Additionally, requirements glossaries are also useful for text summarization and term-based indexing.

In order to develop a glossary, the terms to be defined and added need to be first extracted from the requirements document. Glossary term extraction for the requirements document is an expensive and time-consuming task. This problem becomes even more challenging for large-sized requirements document, e.g., [20, 21]. Once the glossary terms have been extracted they need to be grouped into clusters of related terms. Clustering of glossary terms helps in improving the overall structure of requirements glossary and writing definitions for the glossary terms. It also helps in detecting multiple terms with the same meaning, i.e., synonyms. Finally, it is also helpful in glossary updation as the definitions of all the glossary terms related to a technical concept can be updated simultaneously.

This paper focuses on automatic extraction and clustering of glossary terms from large-sized requirements documents. A first step in this direction is to extract the candidate glossary terms from a requirements document by applying text chunking. Text chunking consists of dividing a text in syntactically correlated parts of words. Since 99 percent of all the relevant terms are noun phrases (NPs) [2, 12], we only focus on extracting the NPs from a requirements document. Next, we use FastText [3, 11] which is a
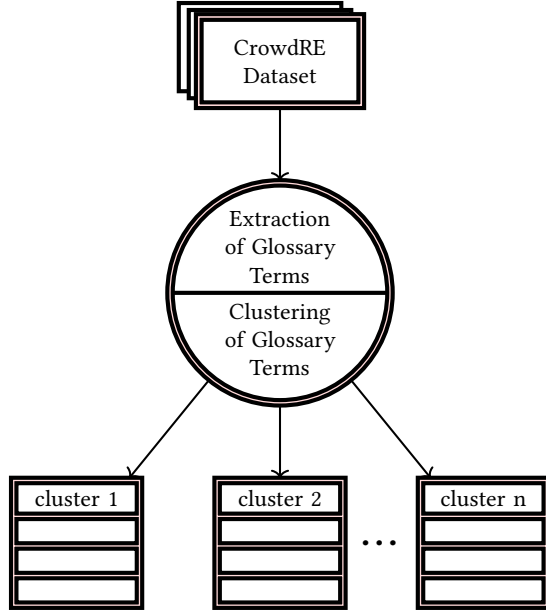
Figure 1: Overview of our approach.

word embeddings model to generate semantic similarity matrix for extracted glossary terms. Word embeddings represent individual words as real-valued vectors in a predefined vector space. Each word is mapped to one vector and the vector values are learnt based on the usage of words. Words that are used in similar ways tend to have similar representations. This means that distance between two words which are semantically very similar is going to be smaller. More formally, the cosine of the angle between such vectors should be close to 1. We have trained our word embeddings model on a domain-specific corpus generated by crawling the home automation (HA) category on Wikipedia. Finally, we give this similarity matrix as an input to the clustering algorithms for generating clusters of glossary terms.

We have applied our approach to the CrowdRE dataset [20, 21], which contains about 3, 000 crowd-generated requirements for smart home applications. To measure the effectiveness of our technique we manually extract the glossary terms from a subset of randomly selected 100 CrowdRE requirements and use this ground truth data for computing the precision and recall of our extraction technique. In addition to these quality metrics we also compute the requirements coverage of these extracted glossary terms. Similarly, we manually cluster the glossary terms extracted from these 100 requirements and use this ground truth data to compute precision and recall for our clustering algorithms. Our detailed experiments show that text chunking and word embedding based glossary term clustering can be very effective for automated extraction and grouping of technical terms in a large-sized requirements document. An overview of our approach is shown in Figure 1. More specifically, our main contributions are as follows :

- We propose a technique for extracting candidate glossary terms by applying text chunking and WordNet removal.

- We propose a word embeddings based technique for clustering glossary terms that are semantically similar.
- We apply our technique to the CrowdRE dataset, which is a large-sized dataset with around 3000 crowd-generated requirements for smart home applications. We extract glossary terms for these requirements. Next, we use FastText model trained on domain-specific corpora to generate semantic similarity matrix which in turn is used by the clustering algorithms for creating clusters.
- To measure the effectiveness of our term extraction technique we manually extract the glossary terms from a subset of randomly selected 100 CrowdRE requirements and use this ground truth data for computing the precision, recall and F-score. Additionally, we also compute the requirements coverage of these extracted glossary terms.
- To measure the effectiveness of our clustering techniques we manually cluster the glossary terms extracted from these 100 requirements and use it for computing precision and recall.

The remainder of the paper is structured as follows: Section 2 discusses the related work. Section 3 provides the required background. Section 4 explains our approach. We present the results of our experiments in Section 5. Section 6 discusses threats to the validity of our experimental results. Finally, Section 7 concludes the paper and provides pointers for future research.

2 RELATED WORK

Word Embeddings for RE. In [6], an approach based on word embeddings and Wikipedia crawling has been proposed to detect domain-specific ambiguities in natural language text. More specifically, in this paper authors investigate the ambiguity potential of typical computer science words using Word2vec algorithm and perform some preliminary experiments. In [7], authors estimate the variation of meaning of dominant shared terms in different domains by comparing the list of most similar words in each domain-specific model. This method was applied to some pilot scenarios which involved different products and stakeholders from different domains. Recently, in [19], we have measured the ambiguity potential of most frequently used computer science (CS) words when they are used in other application areas or subdomains of engineering, e.g., aerospace, civil, petroleum, biomedical and environmental etc. For every ambiguous computer science word in an engineering subdomain, we have reported its most similar words and also provided some example sentences from the corpus highlighting its domain-specific interpretation. Our detailed experiments with several different subdomains show that word embeddings based techniques are very effective in identifying domain specific ambiguities. Our findings also demonstrate that these techniques can be applied to documents of varying sizes. All these applications of word embeddings to requirements engineering are very recent and only focus on detecting ambiguity in requirements documents.

Glossary Extraction and Clustering for RE. In [1, 9], authors have developed tools for finding repeated phrases in natural language requirements. These repeated phrases have been termed as *abstractions*. In [23], authors have described an approach for automatic domain-specific glossary extraction from large document collections using text analysis. In [13], authors described a case study

on application of natural language processing for extracting terms from the text written by domain experts, and build a domain ontology using them. In [27], a term extraction technique has been proposed using parsing and parse relations. A text mining technique using term ranking and term weighing measures for the automatic extraction of the most relevant terms used in Empirical Software Engineering (ESE) documents has been proposed in [28]. In [29], authors have developed a procedure for automatic extraction of single and double-word NPs from existing document collections. Dwarakanath et al. [5] presented a method for automatic extraction of glossary terms from unconstrained natural language requirements using linguistic and statistical techniques. Menard et al. [15] retrieved domain concepts from business documents using a text mining process. Recently, a hybrid approach which uses both linguistic processing and statistical filtering for extracting glossary terms has been proposed in [8]. This technique has been applied to the same CrowdRE dataset which we have used in our paper for experiments. Arora et al. [2] have recently proposed a solution for automatic extraction and clustering of candidate glossary terms from natural language requirements. This technique has been evaluated on three-small sized industrial case studies. Note that in [2] syntactic, e.g., Jaccard, Levenstein, Euclidean and knowledge-based similarity measures, e.g., WUP, LCH, PATH have been used for clustering of glossary terms.

All these term extraction techniques can be broadly classified into linguistic, statistical or hybrid approaches. Linguistic approaches detect glossary terms using syntactic properties. In contrast, statistical approaches select terms based on the frequency of their occurrence. A hybrid approach combines both linguistic and statistical approaches, e.g., [8]. In this paper we take a linguistic approach for glossary term extraction and a semantic approach for glossary clustering. Our paper builds on that investigated in [2] and extends it in the following ways :

- The NLP and filtering pipeline used in our approach is different from the one used in [2], e.g., we only focus on NPs, we use lemmatized chunks and we apply WordNet removal.
- In addition to precision and recall, we also compute the requirements coverage to measure quality of extracted glossary terms.
- Unlike [2] where authors have used syntactic and knowledge-based similarity measures, we use a state-of-the art neural word embeddings model for clustering glossary terms by capturing the natural semantic relationship of their meanings. To compute the semantic relationship of two different terms we have used cosine similarity.
- Unlike [2] where the glossary term extraction and clustering approach has been evaluated on three-small sized case studies (110-380 requirements), we have evaluated our solution on a large-sized requirements document with around 3000 crowd-generated requirements for smart home applications.
- Finally, in [2], authors have concluded that (knowledge-based) semantic measures do not have a significant impact on clustering accuracy and that individual semantic measures should not be applied on their own. In contrast, our experiments have shown that (corpus) word embeddings based

clustering approach can be very effective for grouping semantically similar glossary terms. Our approach falls under the corpus-based semantic similarity category which determines the similarity between terms according to information gained from large corpora.

To the best of our knowledge, this is the first time that a word embeddings based approach has been proposed for grouping semantically similar glossary terms extracted from a large body of requirements.

3 PRELIMINARIES

This section discusses some basic concepts that are needed for the understanding of the rest of this paper.

3.1 Requirements Glossary

Terminological issues are the most common sources of confusion in requirements document. It is possible that a single reader can interpret the requirements in more than one way or that multiple readers come to different interpretations. Additionally, domain experts may use specialized jargon, whose meaning may not be clear to readers with a different technical background and domain expertise. A glossary defines specialized terms used in a requirements document. It contains the following elements [26] :

- Context-specific technical terms
- Abbreviations and acronyms
- Everyday concepts that have a special meaning in the given context
- Synonyms, i.e., different terms with the same meaning
- Homonyms, i.e., identical terms with different meanings

3.2 CrowdRE Dataset

The CrowdRE dataset was created by acquiring requirements from members of the public, i.e., *crowd* [20]. This dataset contains about 3000 requirements for smart home applications. A study on Amazon Mechanical Turk was conducted with 600 workers. This study measured personality traits and creative potential for all workers. A two-phase sequential process was used to create requirements. In the first phase, user stories for smart home applications were collected from 300 workers. In the second phase, an additional 300 workers rated these requirements in terms of clarity and creativity and produced additional requirements.

Each entry in this dataset has 6 attributes, i.e., role, feature, benefit, domain, tags and date-time of creation. Since we are interested in extracting domain-specific terms from this dataset, we only focus on feature and benefit attributes of this dataset. Some example requirements obtained from this dataset after merging feature and benefit attributes are shown in Table 1. For further details, we refer the interested reader to [20, 21].

3.3 FastText

Word embeddings are a powerful approach for analyzing language and have been widely used in information retrieval and text mining. They provide a dense representation of words in the form of numeric vectors which capture the natural semantic relationship

Table 1: Some examples of CrowdRE requirements.

Req Id	Textual Requirements
R1	my smart home to be able to order delivery food by simple voice command, i can prepare dinner easily after a long day at work.
R2	my smart home to turn on certain lights at dusk, i can come home to a well-lit house.
R3	my smart home to sync with my biorhythm app and turn on some music that might suit my mood when i arrive home from work, i can be relaxed.
R4	my smart home to ring when my favorite shows are about to start, i will never miss a minute of my favorite shows.

of their meaning. Word embeddings are considered to be an improvement over the traditional bag-of-words model which results in very large and sparse word vectors. The model "Word2vec" developed by the researchers at Google [16] learns and generates word embeddings from a natural language text corpus. The skip gram negative sampling (SGNS) implementation of Word2vec [18] predicts a collection of words $w \in V_W$ and their contexts $c \in V_C$, where V_W and V_C are the vocabularies of input words and context words, respectively. Context words of a word w_i is a set of words w_{i-wind}, ..., w_{i-1}, w_{i+1}, ..., w_{i+wind} for some fixed window size $wind$. Let D be a multi-set of all word-context pairs observed in the corpus. Let $\vec{w}$, $\vec{c} \in \mathbb{R}^d$ be the d-dimensional word embeddings of a word w and context c. These vectors (both word and context) are created by the Word2vec model from a corpus and are analyzed to check the semantic similarity between them. The main objective of negative sampling (NS) is to learn high-quality word vector representations on a corpus. A logistic loss function is used in NS for minimizing the negative log-likelihood of words in the training set. FastText [3, 11] is an extension to Word2Vec proposed by Facebook's AI research lab in 2016. Unlike Word2vec which feeds individual words into the neural network, FastText breaks words into several n-grams (sub-words). For example, the tri-grams for the word orange is ora, ran, ang and nge. The word embedding vector for orange will be the sum of all these n-grams. Once the neural network has been trained we obtain word embeddings for all the n-grams given the training dataset. FastText properly represents rare words since it is highly likely that some of their n-grams also appears in other words. Note that Word2vec fails to provide any vector representation for words that are not in the corpus.

Word Similarity Computation. The FastText model uses the *cosine similarity* to compute the semantic relationship of two different words in vector space. Let us assume two word embedding vectors $\vec{w}$ and $\vec{w''}$, then the cosine angle between these two word embedding vectors is calculated using Equation (1).

$$cos(\vec{w}, \vec{w''}) = \frac{\vec{w} \bullet \vec{w''}}{|\vec{w}| \, |\vec{w''}|} \tag{1}$$

The range of similarity score is between 0 to 1. If the score is closer to 1 this means that the words are semantically more similar and used in almost the same context. On the other hand, if the score is closer to 0 it means that the words are less related to each other.

3.4 Clustering

Clustering involves grouping together of objects with similar characteristics. It divides the dataset into groups, based on the similar data points. Clustering uses unsupervised mode of learning, where the task of the machine is to group unsorted information according to *similarities*, *patterns* or *differences* without any prior training on data. Clustering helps the businesses to manage their data better by image segmentation, grouping web pages, market segmentation and information retrieval. In this paper we have used two clustering algorithms, namely, *K-means* [14] and *Expectation Maximization* (EM) [4].

K-means. : *K*-means is a hard boundary clustering algorithm. It partitions a set of data points into K clusters. Each cluster is represented by a centroid which is effectively the mean value of all the data points belonging to the cluster. The points are assigned to the clusters in such a way that the intra-cluster similarity is high (the data points are closer) and the inter-cluster similarity is low (the data points are farther). The centroids are initialized randomly and re-evaluated iteratively along with the cluster membership functions until they converge. The algorithm works by iterating through the assignment step where each data point is assigned to a cluster based on its similarity to the centroid and updation step which involves evaluating the centroids until each centroid converges.

Expectation Maximization (EM). : EM assumes apriori that there are K Gaussian and the algorithm tries to fit the data into the K Gaussian by expecting the classes of all data point and then maximizing the maximum likelihood of Gaussian centers. The algorithm iterates through the expectation step which computes the expected classes of all data points for each class and maximization step which computes the maximum likelihood means given our data class membership distribution.

4 APPROACH

This section discusses the approach used for extraction and clustering of glossary terms from large-sized requirements documents.

4.1 Extraction of Glossary Terms

This subsection presents the steps that have been used for extracting the candidate glossary terms. The first step includes the process of *corpus building*. In the second step, we perform *data preprocessing* to extract the candidate glossary terms. The final step is *filtering of candidate glossary terms* which provides us the final set of domain-specific terms. The rest of this section elaborates each of these steps.

4.1.1 Corpus Building.

CrowdRE. For each user story in the CrowdRE dataset, we merge the feature and benefit attributes to obtain a single textual requirement. This is done by using a comma (,) between the text present

in two attributes and a full stop (.) to terminate the requirement. Let C_{CRE} denotes the CrowdRE corpora obtained after applying the above mentioned transformations.

4.1.2 Data Preprocessing. This step involves transforming raw natural language text into an understandable format. All the steps of data preprocessing have been implemented using the Natural Language Toolkit (NLTK)[1] in Jupyter notebook[2], an open-source python IDE. The NLP pipeline used in data preprocessing is shown in Figure 2. In the first step of this NLP pipeline all the words are expanded from its shorter version (expanding contractions). The contractions are shortened version of words. In case of English language, the contractions are generated by removing one of the vowels from the word phrase. For example: "are not" is shortened to *aren't* and would not is written as *wouldn't*. These contractions exist either in written or spoken forms. The textual data (sentences) of each corpus are broken into tokens of words (*tokenization*) followed by the cleaning of all special symbols, i.e., alpha-numeric words. Note that tokenization preserves the syntactic structure of sentences. Next, the tokens of each sentence are tagged according to their syntactical position in the text. The tagged tokens are encoded as 2-tuples, i.e., (PoS, word), where PoS denotes the part of speech. We have used the NLTK (pos_tag)[3] Tagger, which is a perceptron tagger for extracting PoS tags. A perceptron part-of-speech tagger implements part-of-speech tagging using the averaged, structured perceptron algorithm. It uses a pre-trained pickled model by calling the default constructor of the PerceptronTagger class[4]. This tagger has been trained and tested on the Wall Street Journal corpus. Since for technical documents, 99% of all the glossary terms are NPs [12], we develop a regular grammar rule that extracts NPs from the text. The used grammar rule is obtained by a certain combination of POS tags that have been experimentally designed to chunk the most relevant NPs (text chunking). Final step deals with lowering of words (case conversion). This step terminates with producing a set of extracted *candidate glossary terms*.

4.1.3 Filtering of Candidate Glossary Terms. The candidate glossary terms extracted in the previous step serve as input for this step whose output would be the final set of glossary terms. Firstly, we lemmatize[5] the candidate glossary terms (*lemmatization*). Lemmatization removes the inflectional endings and returns the base or dictionary form of a word, i.e., *lemma*. For example, after the lemmatization step, *grows* becomes *grow* and *playing* becomes *play*. In the next step, we clean all the extracted chunks such that all the tokens annotated as determiners, predeterminers and possessive pronouns are removed. Additionally, all the single letter tokens (like 'i' or 'e') and empty strings are also removed from the glossary terms (*removal of determiners, pre-determiners, possessive pronouns, single letter words and empty strings*). Next, all the single word terms that are present in the WordNet English database dictionary[6] are also removed due to the fact that such common words are usually irrelevant for a domain-specific glossary (*WordNet removal*). Out

---

[1] https://www.nltk.org/
[2] https://jupyter.org/
[3] https://www.nltk.org/_modules/nltk/tag.html
[4] https://www.nltk.org/_modules/nltk/tag/perceptron.html#PerceptronTagger
[5] https://www.nltk.org/_modules/nltk/stem/wordnet.html
[6] https://wordnet.princeton.edu/
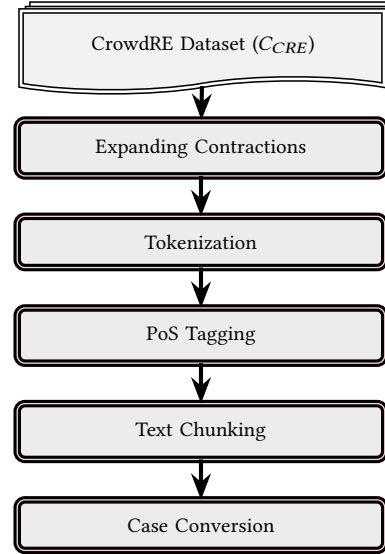
Figure 2: NLP Pipeline.

of the remaining terms, we remove the duplicate terms (*removal of duplicate terms*) and this gives us the final set of glossary terms. All the steps of filtering pipeline are shown in Figure 3. The descriptive
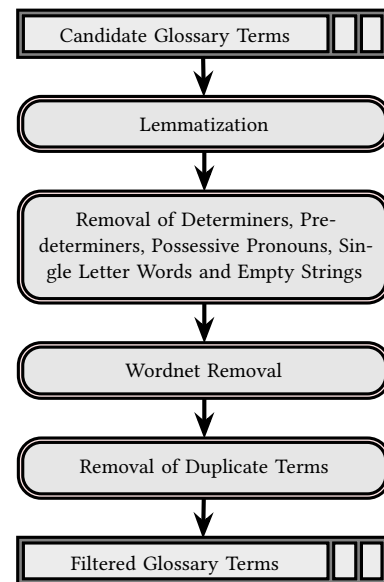
Figure 3: Filtering Pipeline.

statistics of the glossary terms obtained after applying filtering pipeline is shown in Table 2.

4.2 Clustering of Glossary Terms

In this subsection, we present the approach used for clustering the extracted glossary terms. We start with a brief description of

Table 2: Descriptive statistics of the extracted glossary terms.

Process	Total Terms
Lemmatized candidate glossary terms	13,172
Removal of DTs, PDTs, PRPs, single letter words, empty strings	12,957
Removal of WordNet terms	4,525
Removal of duplicate terms (without WordNet removal)	4,084
Removal of duplicate terms (with WordNet removal)	2,890

the steps involved in generating a domain-specific corpora which has been used for training the FastText model. Next, we discuss the training procedure and the steps involved for generating the similarity matrix.

4.2.1 Domain-specific Corpora. We have used some standard web scraping packages available in python[7] to crawl and build the corpora from Wikipedia home automation category[8]. Wikipedia categories have a tree structure and they group together pages on similar subjects. Categories are found at the bottom of an article page. They support auto linking and multi-directional navigation. For our case, the maximum depth used for subcategory traversal is 2. This is primarily because increasing the depth results in extraction of less relevant pages from Wikipedia. For the sake of completeness, we have crosschecked all the results (data extraction for the home automation Wikipedia category) with the help of a widely used Wikipedia category data extraction tool known as PetScan[9]. PetScan (previously CatScan) is an external tool which can be used to find all the pages that belong to a Wikipedia category for some specified criteria. This corpora has been used for training the word embeddings model, i.e., FastText.

For comparative purposes, we have also used a pre-trained Wiki-news FastText word embeddings model [17]. It consists of 1 million word vectors trained with subword information on Wikipedia 2017, UMBC webbase corpus and statmt.org news dataset (16B tokens)[10]. This generic dataset has been used to illustrate the fact that a general purpose embedding does not necessarily adequately represent domain-specific terms.

4.2.2 Training. For FastText training, we have used the Gensim library[11]. For implementing the clustering algorithms, i.e., K-means and EM, we have used SciPy[12] and scikit-learn library[13], respectively. Google Colaboratory[14] has been used for providing GPU support to train the FastText model and run clustering algorithms.

---

[7]https://selenium-python.readthedocs.io/
[8]https://en.wikipedia.org/wiki/Home_automation
[9]https://petscan.wmflabs.org/
[10]https://fasttext.cc/docs/en/english-vectors.html
[11]https://radimrehurek.com/gensim/models/fasttext.html
[12]https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.vq.kmeans.html
[13]https://scikit-learn.org/stable/modules/generated/sklearn.mixture.GaussianMixture.html
[14]https://colab.research.google.com/notebooks/welcome.ipynb

The parameters used for training the FastText model are as follows : dimension=300, window=4, epochs=100, min_count=0, sample=$6e$-5, alpha=0.03, min_alpha = 0.0007 and negative=20.

4.2.3 Similarity Matrix. For generating the similarity matrix, we follow the *optimal matching* approach originally proposed in [22]. Since our glossary terms can be multi-word terms, we need to combine the token-level similarity scores. Recall that token-level similarity scores can be easily obtained by computing the cosine similarity between word vectors generated by FastText model. For a given pair of multi-word glossary terms, the terms are used as bags of tokens. Next, the similarity scores for all pair of tokens are computed using the cosine similarity. To compute the optimal matching of the glossary terms, we select the matching that maximizes the sum of token-level similarity scores. To obtain the final similarity score, we compute the normalized sum for the optimal matching. For further details, we refer the interested reader to [2, 22]. Using the above mentioned procedure we can compute the similarity score between each pair of extracted glossary terms and generate a similarity matrix with dimension $n \times n$ where n is the total number of (multi-word) glossary terms extracted from the requirements document. This similarity matrix is given as an input to the clustering algorithms for grouping semantically similar glossary terms.

5 EXPERIMENTAL RESULTS

This section presents the results of our detailed experiments. Subsection 5.1 evaluates the quality of extracted glossary terms by computing precision, recall, F-score and requirements coverage. Similarly, subsection 5.2 evaluates our clustering techniques by computing precision, recall and F-score.

5.1 Extraction of Glossary Terms

5.1.1 Generation of Ground Truth. Ground truth is used for checking the results of machine learning for accuracy against the real world. For glossary term extraction, generation of ground truth involves manual creation of correct glossary terms by domain experts or by a team of experienced requirements engineers or researchers. Since CrowdRE dataset does not contain a reference list of correct glossary terms and it is not possible to generate the list of glossary terms manually for 3000 requirements, we have manually created the ground truth for a random subset of 100 requirements. This ground truth allows us to assess the performance of our approach by computing precision, recall and F-score. A total of 120 glossary terms have been manually extracted from a random subset of 100 CrowdRE requirements. Some examples of manually extracted glossary terms have been shown in Table 4.

5.1.2 Precision and Recall. To evaluate the quality of our glossary term extraction technique we compute the precision and recall on this random subset of 100 CrowdRE requirements. Precision gives us the fraction of relevant instances among the retrieved instances. On the other hand, recall gives us the fraction of relevant instances that have been retrieved over the total amount of relevant instances. Precision and recall values are computed using True Positives (TP), False Positives (FP) and False Negatives (FN). True Positives are the total number of extracted glossary terms which are also present in

Table 3: Precision, recall and F-score for extracted terms.

Process	Extracted Glossary Terms	Precision	Recall	F-score
With WordNet Removal	143	32.16	51.68	39.65
Without WordNet Removal	292	24.65	69.90	36.45

the ground truth. False Positives are the total number of extracted glossary terms which are not present in the ground truth. Similarly, the ground truth terms that are not extracted by our approach are False Negatives (FN). Precision, recall and F-score are computed as follows:

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \quad F\text{-}score = \frac{2 \times P \times R}{R + P}$$

Precision, recall and F-score values for glossary terms extracted from the random subset of 100 CrowdRE requirements with Word-Net removal are shown in the first row of Table 3. A total of 143 terms have been extracted from the random subset of 100 CrowdRE requirements. Similarly, the values of these quality metrics for terms without WordNet filtering are shown in the second row of Table 3. Here, a total of 292 glossary terms have been extracted from the same subset of 100 requirements. Some example glossary terms extracted by our approach (with WordNet removal) are shown in Table 5. In [2], authors argue that recall is more important than precision for measuring the quality of extraction approach because a low recall means that many of the important domain-specific terms are not present in the extracted glossary. From Table 3 it can be observed that the value of recall is significantly improved if we do not apply WordNet filtering. However, it is important to remember that we are dealing with large-sized requirements documents and for such documents filtering of candidate terms is crucial. For large-sized requirements one needs to strike a balance between

Table 4: Some examples of manually extracted glossary terms.

Req Id	Textual Requirements	Glossary Terms
R1	my <u>smart home</u> to turn on <u>certain lights</u> at dusk, i can come home to a well-lit house.	smart home, certain lights
R2	my <u>smart home</u> to sync with my <u>biorhythm app</u> and turn on some <u>music</u> that might suit my mood when i arrive home from work, i can be relaxed.	smart home, biorhythm app, music

the number of extracted glossary terms and recall. We therefore believe that WordNet removal is important for large-sized documents and our future efforts should focus on improving the recall of our extraction approach. As pointed out in [2], low precision is easy to handle as it only involves removing undesired terms from the list of extracted glossary terms. We believe that statistical filtering [8] combined with semantic filtering could be very helpful in improving the precision of our approach.
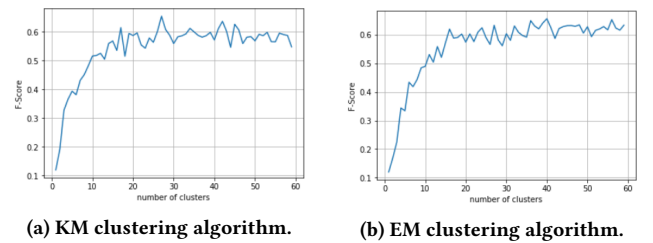
Table 5: Examples of extracted glossary terms using our automated approach.

Glossary Terms	Glossary Terms
automatic fragrance sprinkler	smart lock
automatic garage door	smart touch screen tv
desired temperature	smart watch application
hallway motion detector	wireless charger

5.1.3 Coverage. In [8], requirements coverage has been advocated as another metric for a glossary's quality. Roughly speaking, the definition of coverage is the extent to which something is addressed, covered or included. In the context of glossary term extraction for software requirements, coverage gives us the percentage of requirements that are covered by the terms present in the glossary. For CrowdRE dataset, without WordNet removal we obtain a total of 4084 glossary terms with a coverage rate of 100%, i.e., 2966 of 2966 requirements are covered by these glossary terms. On applying WordNet removal the number of glossary terms reduces to 2, 890 and the corresponding coverage rate is 94.37%, i.e., 2799 of 2966 requirements are covered by this new set of glossary terms. This reduction in coverage rate can be attributed to the following reason. The common nouns or NPs which are not domain-specific but do appear frequently in requirements document would not be part of the final set of glossary terms obtained after applying WordNet filter. Note that even after applying WordNet filtering we get a very high coverage rate which indicates that our extracted glossary is of a reasonably good quality.

(a) KM clustering algorithm.   (b) EM clustering algorithm.

Figure 4: F-score curves for KM and EM algorithms with Fast-Text model trained on Wikipedia home automation corpora.

5.2 Clustering of Glossary Terms

5.2.1 Ground Truth Generation. Since CrowdRE dataset does not contain any information about the ideal clusters (ground truth)
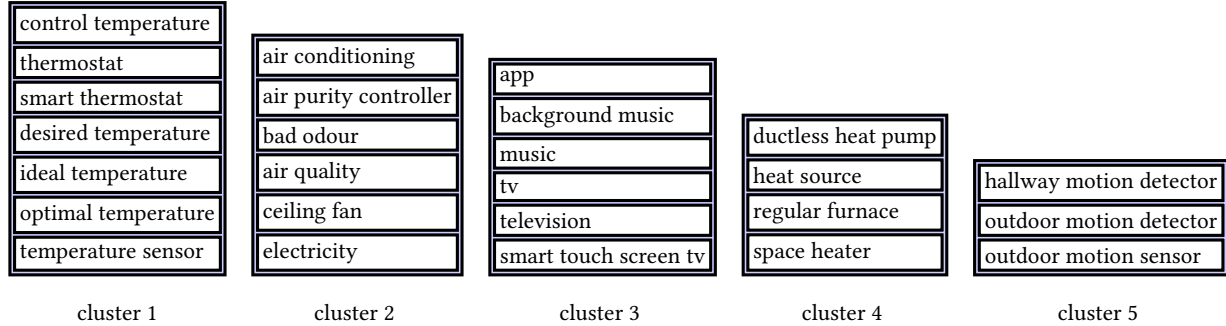
cluster 1 **cluster 2** **cluster 3** **cluster 4** **cluster 5**

Figure 5: Examples of manually clustered glossary terms.

Table 6: Performance evaluation results for automated clustering of glossary terms in ground truth.

Training Approach	Cluster Range	Clustering Algorithm	Precision	Recall	F-score	#Clusters	Normalized AUC
FastText (trained on Wikipedia home automation dataset)	1-20	K-means	56.08	67.91	61.43	16	
	21-40		72.26	59.70	65.38	26	0.5377
	41-60		85.19	50.74	63.60	41	
	1-20	EM	59.21	64.92	61.93	16	
	21-40		84.05	53.73	65.55	40	<u>0.5542</u>
	41-60		93.78	50.00	65.22	56	
FastText (pre-trained Wiki news word embedding model)	1-20	K-means	52.88	58.95	55.75	19	
	21-40		78.57	48.50	59.98	36	0.5063
	41-60		79.87	45.52	57.99	54	
	1-20	EM	53.10	63.43	57.81	16	
	21-40		75.98	50.00	60.31	37	0.5192
	41-60		89.20	48.50	62.84	54	

for glossary terms, we have manually clustered the 120 glossary terms into groups of semantically related terms. More specifically, we have manually grouped the 120 glossary terms into 30 clusters. Note that some of these ideal clusters are overlapping. Fig. 5 shows examples of manually generated clusters for glossary terms.

5.2.2 Precision and Recall. In this section, we compute the precision, recall and F-score for our clustering algorithms, i.e., K-means and EM. As pointed out in [2], if the number of glossary terms is n then it makes little sense to increase the value of K beyond $\frac{n}{2}$ as this would generate very small-sized clusters. To compute precision, recall and F-score for clusters, we have used the same procedure as suggested in [2]. This procedure computes the overall precision and recall by computing the weighted averages of the precisions and recalls of the ideal clusters. We first apply EM and K-means algorithms for clustering 120 glossary terms that have manually extracted (ground truth) from a random subset of 100 CrowdRE requirements. Results of our experiments are shown in

Table 6. Since number of extracted terms is 120, maximum number of clusters is 60. For each clustering algorithm we have two combinations, i.e., word embeddings model trained on domain-specific corpora or trained on Wiki-news. For each combination we have three rows in the table. The first row reports the results where F-score has maximum value for clusters ranging from 1-20. Similarly, second and third rows reports the maximum value of F-score for the clusters ranging from 21-40 and 41-60, respectively. The last column gives the value of normalized Area Under the Curve (AUC), i.e., average F-score for each curve obtained by computing its Area Under the Curve and normalizing the result for each combination. The curves of F-score against the number of clusters for KM and EM algorithms with FastText model trained on Wikipedia home automation (domain-specific) corpora are shown in Figure 4. In Table 6 for K-means clustering with FastText model trained on Wikipedia home automation (domain-specific) dataset,

Table 7: Performance evaluation of term extraction combined with clustering.

WordNet Removal	Training Approach	Cluster Range	Clustering Algorithm	Precision	Recall	F-score	#Clusters	Normalized AUC
Yes		1-23		48.36	40.29	43.96	24	
		24-47	K-means	62.33	39.55	48.39	45	0.4071
		48-71		74.40	39.55	51.64	57	
		1-23		42.15	45.52	43.77	23	
	FastText (trained on Wikipedia home automation dataset)	24-47	EM	61.26	42.53	50.21	37	0.4126
		48-71		68.52	43.28	53.05	48	
		1-48		40.97	45.52	43.13	35	
No		49-97	K-means	67.16	40.29	50.37	97	0.3991
		98-146		71.58	38.05	49.69	113	
		1-48		44.28	41.79	43.00	46	
		49-97	EM	44.28	41.79	43.00	86	0.4125
		98-146		78.23	38.80	51.87	123	
		1-23		44.38	38.80	41.41	24	
Yes		24-47	K-means	61.16	38.80	47.48	32	0.387
		48-71		72.12	35.82	47.86	67	
		1-23		44.67	37.31	40.66	21	
	FastText (pre-trained Wiki news word embedding model)	24-47	EM	65.12	36.56	46.83	47	0.3855
		48-71		76.05	34.32	47.30	64	
		1-48		47.73	38.05	42.35	48	
No		49-97	K-means	49.67	41.79	45.39	74	0.3785
		98-146		81.35	34.32	48.28	142	
		1-48		44.97	38.05	41.22	48	
		49-97	EM	60.03	35.82	44.87	91	0.3848
		98-146		67.30	35.82	46.75	106	

the maximum value of F-score is 65.38 and the corresponding number of clusters is 26. The highest value of F-score among all the four clustering combinations is obtained for EM clustering with FastText model trained on Wikipedia home automation dataset. Here, F-score is 65.55 and the number of clusters is 40. As the Table suggests, the highest value of normalized AUC (0.5542) is obtained for EM clustering with FastText model trained on Wikipedia home automation dataset. Note that for this combination the number of clusters, i.e., 40 corresponding to the highest F-score comes very close to the ideal number of clusters (ground truth), i.e., 30. Therefore, we believe that EM combined with FastText model trained on

domain-specific corpora is the best choice for clustering glossary terms extracted from CrowdRE requirements.

Performance Evaluation of Term Extraction Combined with Clustering. : In the previous section, we have evaluated the quality of our clustering approach by automated clustering of glossary terms present in the ground truth and computing precision, recall and F-score using the ideal clusters. In this section, we first extract the glossary terms from the same subset of 100 CrowdRE requirements using our extraction approach. Next, we group these terms using our clustering algorithms and compute the precision, recall and F-score using our ideal clusters. For term extraction we have two combinations, i.e., with or without WordNet removal. As mentioned
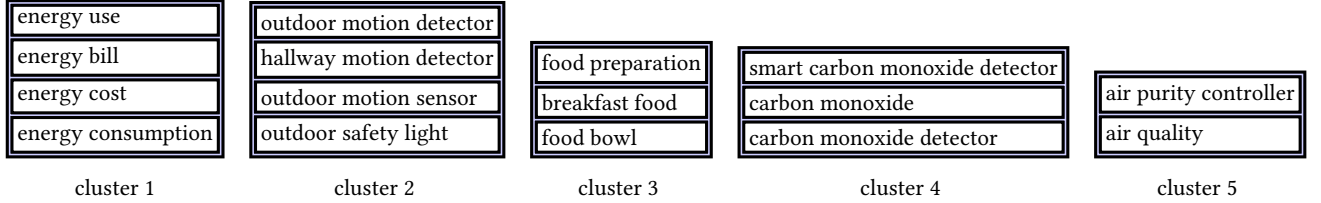
Figure 6: Examples of automated clustering of glossary terms (WordNet + EM + FastText (Wikipedia home automation)).

in the previous section, for clustering we have four combinations. This means that in total we have to perform our experiments on eight combinations. From Table 3 we know that with WordNet removal the total number of extracted terms is 143. This means that the upper bound on number of clusters is 71. Similarly, for term extraction without WordNet removal upper bound on the number of clusters is $\frac{292}{2} = 146$. The results of our experiments are shown in Table 7. The highest normalized AUC, i.e., 0.4126 is obtained for the second combination where WordNet removal is combined with EM algorithm and FastText model trained on domain-specific corpora. Some example clusters for glossary terms extracted with WordNet removal and grouped using EM algorithm and FastText model trained on Wikipedia home automation corpora are shown in Figure 6. From this example it can be observed that all the terms present in a cluster are semantically very similar or related. For example, all types of outdoor motion detectors and safety light have been placed in cluster 2. Similarly, all the terms related to enery usage, e.g., energy bill, energy cost and energy consumption have been placed in cluster 1.

It is important to note that the fourth combination in Table 7 gives us the second highest normalized AUC value, i.e., 0.4125 which comes very close to the highest value. Considering the fact that for large-sized requirements documents filtering of candidate glossary terms, i.e., WordNet removal is crucial and second combination in Table 7 gives us the highest normalized AUC, we believe the second combination is the best choice for glossary term extraction and clustering.

6 THREATS TO VALIDITY

This section discusses threats to the validity of our experimental results.

Internal Validity: One of the main threats to internal validity is the researcher bias, which may have occurred in manual generation of the ground truth data for glossary terms and ideal clusters. To reduce this threat ground truth data was generated individually by all the three authors of this paper and the final ground truth data was generated through consensus after detailed discussions among the authors. Since the word embeddings model has been trained on different datasets, threats related to the bias in model training have been reduced.

External Validity: Concern towards generalizibility may arise because of singleton case study to assess the empirical feasibility of the proposed approach. Having said this we believe that our approach can be used for any application domain by training the FastText model on corresponding domain-specific corpora.

Conclusion Validity: All the conclusions drawn in this paper are shown to have been rooted in section 5 of this paper and thus there is traceability.

7 CONCLUSIONS

This paper proposes an automatic approach for extracting and clustering of glossary terms from large-sized requirements documents. The first step of our solution involves extracting candidate glossary terms by applying text chunking and WordNet removal. In the next step, we apply a state-of-the art word embeddings model, i.e. FastText for generating semantic similarity measures. The model has been trained on a domain-specific corpus created by crawling home automation (HA) category on Wikipedia. Finally, we propose a technique to cluster semantically similar glossary terms by applying K-means and Expectation Maximization clustering algorithms. We have applied our technique to a large-sized requirements documents with around 3000 crowd-generated requirements. We have measured the quality of our extraction approach by computing precision and recall using manually extracted glossary terms from a random subset of 100 CrowdRE requirements. Similarly, the quality of our clustering technique has been measured by computing precision and recall using manually clustered glossary terms extracted from the same subset of 100 requirements. Our experiments indicate that word embeddings based semantic clustering of glossary terms can be very effective for large-sized requirements document. More specifically, we recommend our NLP pipeline combined with filtering for glossary term extraction. For clustering, we suggest using EM algorithm with FastText model trained on domain-specific corpora.

Future Work. This research work can be extended in several interesting directions which are as follows :

- Improve the recall of our term extraction approach.
- Investigate the practical utility of our approach by applying it to several large-sized requirements documents from different application domains.
- Extend this technique to automatically extract and cluster the glossary terms from a large body of natural language requirements for software product lines [25].
- Extend this technique to automatically extract and cluster the glossary terms from documents where natural language text is supplemented with images, tables and figures.
- Compare the effectiveness of FastText semantic clustering with other word embedding techniques, e.g., GloVe.
- Compare the performance of K-means and *EM* algorithms with other well known clustering techniques.

REFERENCES

[1] Christine Aguilera and Daniel M. Berry. 1990. The use of a repeated phrase finder in requirements extraction. *Journal of Systems and Software* 13, 3 (1990), 209–230.

[2] Chetan Arora, Mehrdad Sabetzadeh, Lionel C. Briand, and Frank Zimmer. 2017. Automated Extraction and Clustering of Requirements Glossary Terms. *IEEE Trans. Software Eng.* 43, 10 (2017), 918–945.

[3] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics* 5 (2017), 135–146.

[4] A. P. Dempster, N. M. Laird, and D. B. Rubin. 1977. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, 1 (1977), 1–22.

[5] Anurag Dwarakanath, Roshni R. Ramnani, and Shubhashis Sengupta. 2013. Automatic extraction of glossary terms from natural language requirements. In *21st IEEE International Requirements Engineering Conference (RE)*. 314–319.

[6] Alessio Ferrari, Beatrice Donati, and Stefania Gnesi. 2017. Detecting Domain-Specific Ambiguities: An NLP Approach Based on Wikipedia Crawling and Word Embeddings. In *25th IEEE International Requirements Engineering Conference Workshops (REW)*. 393–399.

[7] Alessio Ferrari, Andrea Esuli, and Stefania Gnesi. 2018. Identification of Cross-Domain Ambiguity with Language Models. In *5th International Workshop on Artificial Intelligence for Requirements Engineering (AIRE)*. 31–38.

[8] Tim Gemkow, Miro Conzelmann, Kerstin Hartig, and Andreas Vogelsang. 2018. Automatic Glossary Term Extraction from Large-Scale Requirements Specifications. In *26th IEEE International Requirements Engineering Conference*. IEEE Computer Society, 412–417.

[9] Leah Goldin and Daniel M. Berry. 1997. AbstFinder, A Prototype Natural Language Text Abstraction Finder for Use in Requirements Elicitation. *Automated Software Engineering* 4, 4 (October 1997), 375–412.

[10] Elizabeth Hull, Ken Jackson, and Jeremy Dick. 2005. *Requirements Engineering* (second ed.). Springer.

[11] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. Bag of Tricks for Efficient Text Classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Association for Computational Linguistics, 427–431.

[12] John S. Justeson and Slava M. Katz. 1995. Technical terminology: some linguistic properties and an algorithm for identification in text. *Natural Language Engineering* 1, 1 (1995), 9–27.

[13] Leonid Kof. 2004. Natural Language Processing For Requirements Engineering: Applicability to Large Requirements Documents. In *Workshop on Automated Software Engineering*.

[14] J. MacQueen. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. University of California Press, 281–297.

[15] Pierre André Ménard and Sylvie Ratté. 2016. Concept extraction from business documents for software engineering projects. *Automated Software Engineering* 23, 4 (December 2016), 649–686.

[16] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* (2013).

[17] Tomas Mikolov, Edouard Grave, Piotr Bojanowski, Christian Puhrsch, and Armand Joulin. 2018. Advances in Pre-Training Distributed Word Representations. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*.

[18] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13)*. 3111–3119.

[19] Siba Mishra and Arpit Sharma. 2019. On the Use of Word Embeddings for Identifying Domain Specific Ambiguities in Requirements. In *27th IEEE International Requirements Engineering Conference Workshops (REW)*. 234–240.

[20] Pradeep K. Murukannaiah, Nirav Ajmeri, and Munindar P. Singh. 2016. Acquiring Creative Requirements from the Crowd: Understanding the Influences of Individual Personality and Creative Potential in Crowd RE. In *24th IEEE International Requirements Engineering Conference (RE)*. 176–185.

[21] Pradeep K. Murukannaiah, Nirav Ajmeri, and Munindar P. Singh. 2017. Toward Automating Crowd RE. In *25th IEEE International Requirements Engineering Conference (RE)*. 512–515.

[22] Shiva Nejati, Mehrdad Sabetzadeh, Marsha Chechik, Steve Easterbrook, and Pamela Zave. 2012. Matching and Merging of Variant Feature Specifications. *IEEE Transactions on Software Engineering* 38, 6 (November 2012), 1355–1375.

[23] Youngja Park, Roy J. Byrd, and Branimir K. Boguraev. 2002. Automatic Glossary Extraction: Beyond Terminology Identification. In *COLING 2002: The 19th International Conference on Computational Linguistics*. https://www.aclweb.org/anthology/C02-1142

[24] Klaus Pohl. 2010. *Requirements Engineering - Fundamentals, Principles, and Techniques* (first ed.). Springer.

[25] Klaus Pohl, Günter Böckle, and Frank J. van der Linden. 2005. *Software Product Line Engineering: Foundations, Principles and Techniques*. Springer-Verlag, Berlin, Heidelberg.

[26] Klaus Pohl and Chris Rupp. 2011. *Requirements Engineering Fundamentals - A Study Guide for the Certified Professional for Requirements Engineering Exam: Foundation Level - IREB compliant*. rockynook. I–XVIII, 1–163 pages.

[27] Daniel Popescu, Spencer Rugaber, Nenad Medvidovic, and Daniel M. Berry. 2008. Reducing Ambiguities in Requirements Specifications Via Automatically Created Object-Oriented Models. In *Innovations for Requirement Analysis. From Stakeholders' Needs to Formal Designs*, Barbara Paech and Craig Martell (Eds.). 103–124.

[28] Francisco P. Romero, José A. Olivas, Marcela Genero, and Mario Piattini. 2008. Automatic Extraction of the Main Terminology Used in Empirical Software Engineering Through Text Mining Techniques. In *Proceedings of the Second ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '08)*. 357–358.

[29] Xuchang Zou, Raffaella Settimi, and Jane Cleland-Huang. 2010. Improving automated requirements trace retrieval: a study of term-based enhancement methods. *Empirical Software Engineering* 15, 2 (April 2010), 119–146.

Comparative Analysis of SVM and ANN Classifiers using Multilevel Fusion of Multi-Sensor Data in Urban Land Classification

Rubeena Vohra[1,2] · K. C. Tiwari[3]

Abstract
Multi-sensor data fusion has recently received remarkably more attraction in urban land classification. The fusion of multi-resolution and multi-sensor remote sensing data can help in comprehending more information about the same land cover features, thereby, enhancing the classification accuracy. In this field of study, a combination of hyperspectral data in a long-wave infrared range and a very high-resolution data in a visible range has been extensively used for exploring the spectral and spatial features for decision level fusion classification. This paper proposes a novel method of integrating the classifier decisions with the additional ancillary information derived from spectral and spatial features for improvement in the classification accuracy of natural and man-made objects in urban land cover. The paper also presents a detailed performance comparative evaluation of two classifiers i.e., support vector machine (SVM) and artificial neural network (ANN) to show the effectiveness of these classifiers. The results obtained from a decision-based multilevel fusion of spectral and spatial information using hyperspectral and visible data have shown improvement in classification accuracy. The results also reveal that the classification accuracy of the SVM classifier is better than ANN in multi-sensor data using decision level fusion of combined feature set analysis.

Keywords Support vector machine · Multisensor data · Decision level fusion · Feature extraction · Artificial neural network

✉ Rubeena Vohra
  rubeena.vohra@gmail.com

1 Delhi Technological University, New Delhi, India

2 Electronics and Communication, Bharatividyapeeth's College of Engineering, New Delhi, India

3 Civil Engineering Department, Delhi Technological University, New Delhi, India

 Springer

1 Introduction

The requirement of greater accuracy in image and data analysis not only fuelled the development of technology in accessing better resolution remote sensing data but also in exploring multiresolution and multi-sensor data [1]. Currently, a variety of multiresolution and multi-sensor remote sensing data is available which primarily includes, multispectral and hyperspectral data. Multispectral remote sensing data comprises of a smaller number of spectral bands with a larger bandwidth. Therefore, the data can be used for studying the spatial characteristics of the ground objects. But the interpretation of multispectral data is difficult particularly, in land-cover applications where intra-class variability is important. Hyperspectral remote sensing data includes large data sets comprising of 100–200 contiguous spectral bands that have ample spectral information of various intra-class ground objects in the scene. But the handling of large amounts of data will, however, dramatically increase the complexity and processing time. Also, the effective and proper selection of relevant bands for the study of any application becomes the primary task while handling the hyperspectral remote sensing data.

Interpretations from multiresolution and multi-sensor data can be enhanced with the knowledge of digital image fusion techniques [2, 3]. Fusion of multiresolution and multi-sensor data not only helps in sharpening the details of low-resolution data but also provides complementary data observed from the same site with superior apprehension which is not possible with the single sensor data [4]. The single-source data is not fully capable of extracting accurate ground objects either due to spectral similarities within different objects or due to spatial adjacencies between the same objects. Consequently, in addition to the spectral response, object characteristics, such as shape and spatial relations should be well interpreted to enhance the accuracy assessment of the data. The use of hyperspectral and multispectral data simultaneously seems to be a promising way for exploring the features related to spectral and spatial domains that can significantly improve the accuracies and enhance the interpretations of remotely sensed data [5, 6]. In the literature it is found that image processing and data fusion of hyperspectral/multispectral multiresolution and multi-sensor data plays a vital role in pan-sharpening, change detection [8–10], image classification [10], image modality fusion [11], geospatial information science with a different applicability. With recent technological advances in remote-sensing systems, fusion of very high-resolution data collected by a digital color camera and a new coarse resolution hyperspectral data in the long-wave infrared (LWIR) range for urban land-cover classification has been extensively enticed much consideration and turned into a research hot spot in image analysis and data fusion research community [13, 14, 31].

The classification of urban land features can be broadly categorized as parametric or non-parametric depending upon whether statistical parameters are used to group pixels or other parameters such as artificial neural network and fuzzy parameters [6, 16, 17]. Parametric techniques can be further divided into two types namely, supervised and unsupervised. Unsupervised classification involves the clustering of the pixels based on the reflectance properties of the pixels. In the case of supervised

classification, however, user-selected reference data is considered for classification. Depending upon whether pixels or objects form the basis of data in classification, these are referred to as a pixel or object-based classification [18, 19]. Another set of classification techniques is biologically inspired by non- parametric techniques. It deals with little or no prior knowledge of the distribution of input or reference data to quantify the complex patterns for classification.

In concern to the discussion above in this paper, a multi-level fusion strategy is proposed for urban land classification on hyperspectral and multispectral multiresolution, multi-sensor data to improve the classification accuracy of objects/classes. In this context, TIR HS (thermal infrared hyperspectral) and VIS RGB (visible band) data are used to extract multiple features about natural and man-made objects/classes. For this purpose, salient features in a spatial, textural and spectral-domain are computed. Further, the multiple features generated through feature level fusion are fed one by one to parametric and non-parametric classifiers i.e., Support Vector Machine (SVM) and Artificial Neural Network (ANN). The comparative classification accuracy assessment of classifiers is obtained through a multi-level fusion strategy using spectral and spatial features. Multi-level fusion strategy is applied at two stages; one at the classifier stage with each feature analysis and other at the combination of all feature analysis. The results have shown that by fusion of all the feature vectors through majority voting the overall classification accuracy of natural and man-made objects has been improved.

The paper is presented in six sections. The next section describes the study area and the available data set. Section 3 presents a detailed description of the theoretical background followed by Sect. 4 that defines the flow graph and the implementation process for the proposed work. In the last section, the experimental results of the classification accuracy are discussed. Finally, the conclusions are presented.

2 Study Area and Data Set

The data consists of an urban area near Thetford Mines, Quebec, Canada [20]. The characteristics of the data sets are described below:

(a) The first airborne data set is an LWIR hyperspectral data that has been acquired using the Telops' Hyper-Cam. The region in the mosaic consists of a variety of natural and man-made objects such as trees, vegetation, soil, roads, and buildings. The airborne LWIR hyperspectral imagery consists of 84 spectral bands in the 868–1280 cm^{-1} region (7.8–11.5 μm) at a spectral resolution of 6 cm^{-1} (full-width-half maximum). The average spatial resolution of LWIR hyperspectral imagery is approximately 1 m. The grey-scale image of the 74th spectral band is shown in Fig. 1a.

(b) The second airborne data has been acquired using a digital color camera (2 Megapixels) mounted on the same platform as that for LWIR Hyperspectral data. The airborne visible imagery is a high spatial resolution digital data with sparse ground coverage over the same area as the LWIR hyperspectral imagery. The
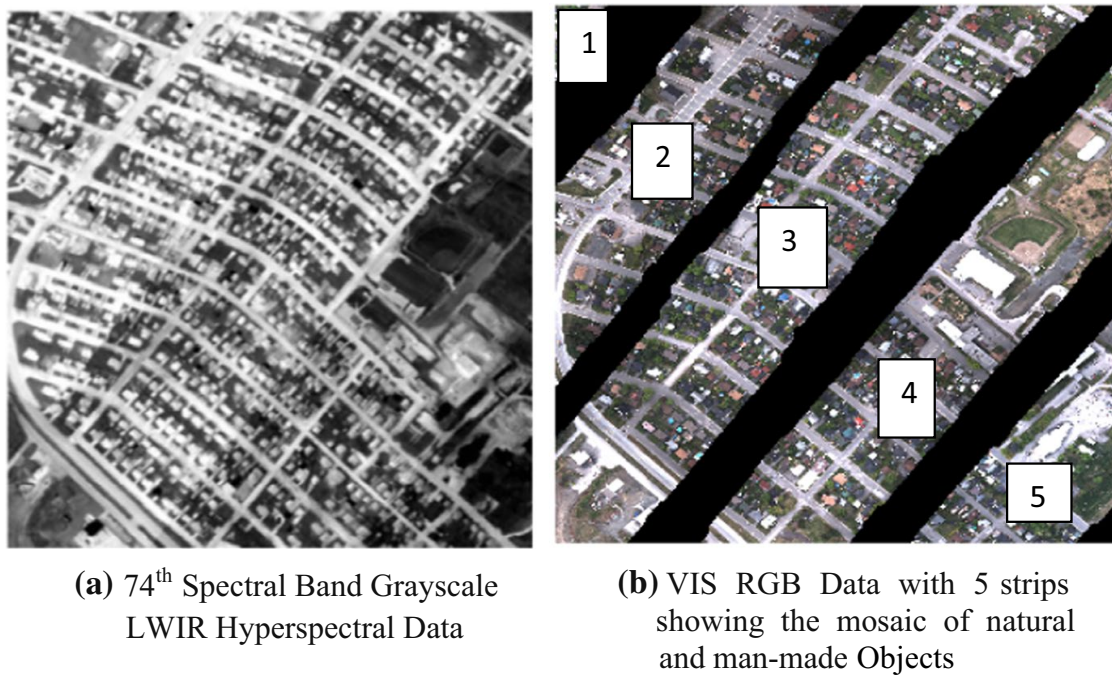
(a) 74th Spectral Band Grayscale LWIR Hyperspectral Data

(b) VIS RGB Data with 5 strips showing the mosaic of natural and man-made Objects

Fig. 1 **a** 74th Spectral Band Grayscale LWIR Hyperspectral Data. **b** VIS RGB Data with 5 strips showing the mosaic of natural and man-made Objects\

average spatial resolution of visible imagery is approximately 0.1 m. Figure 1b shows a VIS RGB image obtained in five strips as shown.

3 Theoretical Background

This section presents a review of the basic concepts of some of the main topics considered in this study.

1. *Feature Extraction* Feature extraction is a way to describe the object in an image in terms of measurable values. The extracted features show the properties of the concerned objects. These features are used with the classifier to assign the class to the object. In this paper, five classes/objects have been identified under the natural and man-made category. These classes/objects are buildings, roads, trees, vegetation, and soil. To extract these objects, the knowledge of mathematical spatial, textural and spectral features is required. In addition to this for the best possible description of the objects, feature level fusion techniques can be employed for extracting the features. Feature level fusion involves the fusion of the features from different sensors to create a feature vector for classifying the objects/classes. In this paper, NDVI (Normalized Difference Vegetation Index) [21] has been used for extracting vegetation, trees and soil components through feature level fusion of multisensory data. NDVI is calculated using hyperspectral data that has bands in an infrared range that extends from the nominal red edge of the visible spectrum at 700 nm to 1 mm. The mean of all the bands is computed to gather the average spectral information of near-infrared range for the healthy vegetation. As the

hyperspectral data has only infrared range information, therefore, the red band information is extracted from VIS RGB data for computing the vegetation index. For a depiction of healthy vegetation, the maximum absorption in red band and maximum reflectance in the infrared band is band required, hence, the fusion of bands from LWIR and VIS RGB data has complemented the extraction of computing the vegetation index in the range from $+1$ to -1. MBI (Morphological building Index) is used for the extraction of buildings by morphological operations [22, 23] through multiresolution segmentation. And textural features have been calculated using occurrence and co-occurrence matrices through GLCM (Gray Level Co-occurrence Matrix) for enhancing the spatial descriptors of the concerned objects/classes [24].

2. *Connected Component Labeling* It is used to detect the connected regions based on pixel connectivity [25] i.e., all pixels having similar pixel intensity values are said to be connected either through 4-pixel or 8-pixel connectivity. Each connected component through 4-pixel or 8-pixel connectivity ensures that the underlying pixels form the same object characteristics. It segments the image into various objects/classes.

3. *SVM Classifier* SVM is a popular method of supervised classification. In this study, SVM with the one-vs-one approach is used because of its effective processing and the ability to handle unbalanced training datasets (G.H. Halldorsson 2004). SVM classifiers (F.Melgani 2004) of form $f(x) = w.\Phi(x) + b$ are made to learn from the data $\{x_i, y_i\}$, where x_i is an n-dimensional input feature vector, $f(x)$ denotes the hyperplane that separates the class labels based on the support vectors on the hyperplane $y_i = \pm 1$ on each side, w and b are the parameters i.e., weights and bias of the hyperplane. The hyperplane calculation can be optimized as follows (G.Camps-Valls 2005):

$$\min_{w,b,\varepsilon} \frac{1}{2}w^2 + C \sum_{i=0}^{N} \varepsilon_i \tag{1}$$

where C is a regularization parameter and ε_i is slack variable. SVM separates the objects/classes based on separating hyper-plane [26]. For a given a set of training data, support vectors for the concerned objects/classes are generated near or far from the hyper-plane. The larger the distance of the support vectors from the hyper-plane the better is the separation of the objects/classes from the other set of objects/classes [27, 28]. Few distinct features of using SVMs in multidimensional data classification are their ability to use the hyperspectral data without considering dimensional reduction algorithms and their insensitivity to the size of training samples and the quality of the input data.

4. *Artificial Neural Network* Many ANN approaches are used in urban land surface classification but the most frequently used neural network is a supervised multi-layered perceptron (MLP) [29]. In the present study of MLP, a back-propagation (BP) training algorithm is used for training the ANN. In the literature, many methods have been developed for training neural networks. The significant techniques widely used are Steep Descent Algorithm, also known as the error

back-propagation algorithm (EBA) and Gauss–Newton Algorithm. Both these algorithms are an important source of least square optimization problems in data fitting but show some problems with convergence and computational speed. Steep Descent algorithm has slow final convergence and fails in finding the minimizer of second-degree polynomial whereas the Gauss–Newton algorithm shows quadratic convergence only if one or more measurement errors are small; otherwise, large error components are difficult to handle and the speed of the final convergence declines. It has been found that for initial stages of neural network training Steep Descent algorithm is preferred and for final stages where quadratic convergence is required Gauss–Newton algorithm is preferred. In the present study, Levenberg–Marquardt (LM) algorithm is used which blends these two minimization curve fitting algorithms. It adapts the fast speed characteristic of Gauss–Newton algorithm and stability factor of Steep Descent algorithm. Hence, Levenberg–Marquardt algorithm [30] provides a numerical solution to non-linear problems as it is fast and has stable convergence is used in the study.

5. *Majority Voting Fusion* Decision level fusion is a higher level of integration of multi-sensor data [31]. In this type of fusion, the classified results from the data of each sensor are merged by various criteria or algorithms to obtain the final classified data. In this paper, the majority voting rule is applied to enhance the classification accuracy of the classified results. The majority voting rule technique is applied on SVM and ANN classified results through various features by setting a set of rules. Here, the class of sample *s* can be determined by:

$$\text{Class (s)} = \text{mode (classified result)} \tag{2}$$

V is the number of votes of class(*s*) in the classified results from the various feature sets. The mode function calculates the maximum number of votes V for a class(s) in a given set of classified results. The class(s) that holds the maximum votes from all the classified results will be considered as the voted class.

The majority voting fusion rule is applied at the end of the classification stage to improve classification accuracy. This rule is generally applied to find out the object in a given set of classes that has more votes than other objects. The class labels assigned to the objects that are obtained from the SVM classifier and ANN are fed as inputs to the general local discrimination strategy. This helps in extracting the object with maximum votes of class labels to fall under the class and thereby increases the classification accuracy. The proposed methodology applies weighted majority voting at two levels. At level 1, binary SVMs using decision level at each feature set are calculated and at level 2, binary SVMs using decision level for combined feature set are calculated. Hence, the decision level is integrated into the classifier's decision at two levels for obtaining the classification accuracy. This leads in understanding the classification accuracy at spatial and spectral domain level as well as considers the combination of spatial and spectral information for enhancing the classification accuracy.

6. *Confusion Matrix* In binary classification techniques, the accuracy is statistically measured to check the performance of the classifier. The results of the classifier

can have four situations [32]; true positive, true negative, false positive and false negative. The classifier result is said to be true positive if the chosen class label n test class matches otherwise false positive. The classifier result is said to be true negative if the chosen class label n test class matches otherwise false negative [33].

$$\text{True Positive} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{3}$$

$$\text{False Positive} = \frac{\text{False Positive}}{\text{False Positive} + \text{True Positive}} \tag{4}$$

4 Implementation

The following section shows the flowchart and the salient aspects of the implementation-

Step I To extract pixels belonging to the classes, first, the features are extracted both in the spectral and spatial domain. These features include normalized difference vegetation index, morphological building index and textural features which can deal with the ambiguities in object recognition that occur due to spectral similarities and spatial adjacencies between various kinds of objects. This is followed by connected component labeling on the ground truth data generated on VIS RGB data by ENVI software (Fig. 2).

Step II Image region property i.e., bounding boxes are calculated and superimposed on all matrixes of all features extracted earlier. As the size of each bounding box applied onto features is different, therefore, the feature data is resized to 25×25 which will make the computations easier. These bounding box features are then divided into training and testing samples

Step III Further, on these training and testing samples SVM and ANN classification algorithms are applied. The study data is classified into three natural and two man-made object classes i.e., vegetation, trees, soil, building, and roads. The classified result is obtained by calculating the true positive values of each class.

Step IV To improve the classification accuracy, the majority voting fusion rule is applied to the classified results. The comparison is done on the results obtained through classified results by applying independent features and by the combination of feature sets to observe the classification accuracy.

Step V Finally, McNemar test is performed to observe the performance of classifiers using spatial and spectral features.
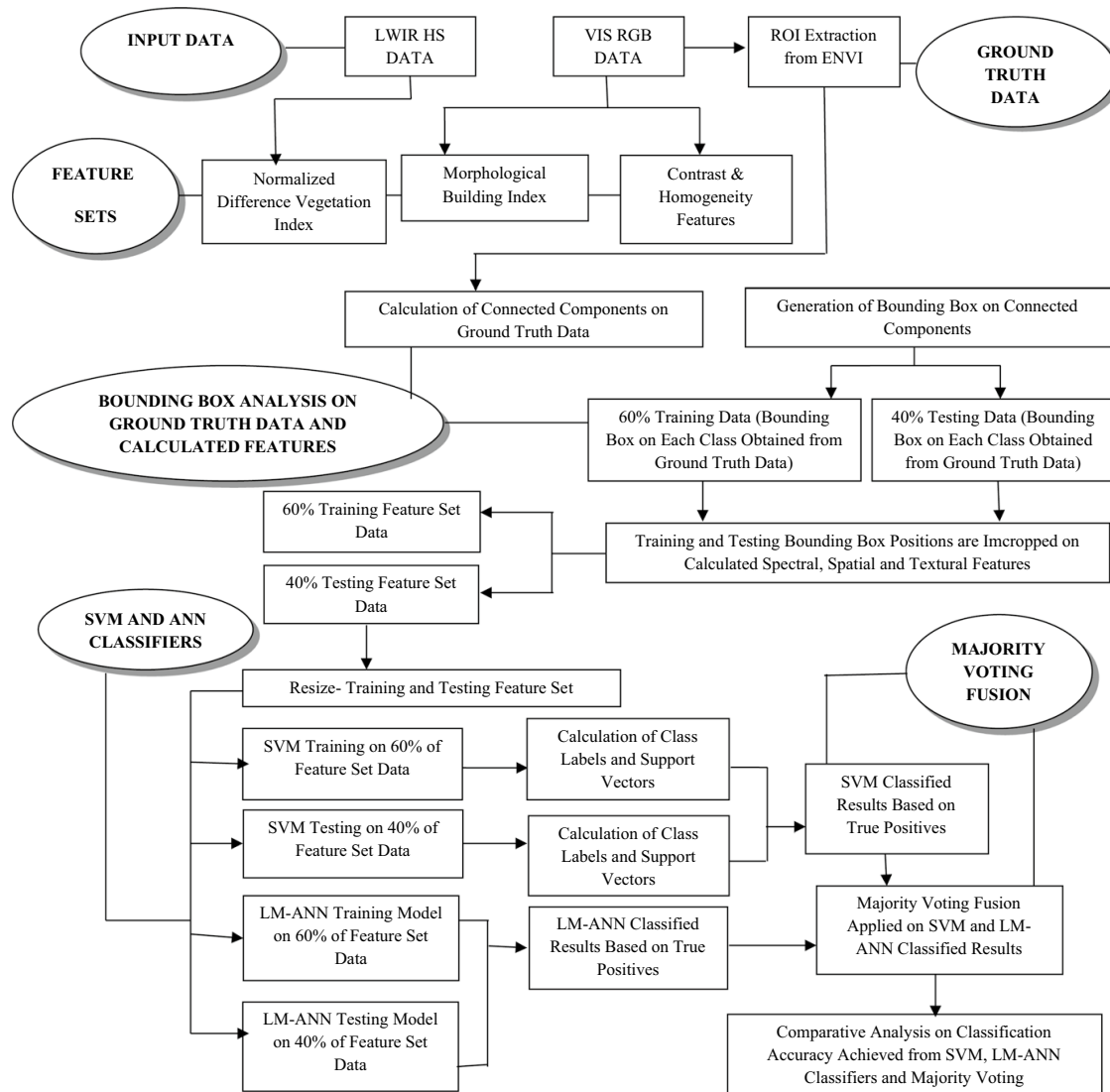
Fig. 2 Flowchart for implementation steps

5 Results and Discussions

This section presents a discussion on the outcome of classification by two algorithms in urban environments using feature level and decision level fusion of multi-resolution and multi-sensor VIS RGB and LWIR data to classify natural and man-made objects. The stepwise discussion is given below:

Step I*: Features and ROI Extraction* Various features such as NDVI, MBI, and textural features have been computed and the results are as given below.

(a) *Vegetation Index* NDVI calculations are done with the available hyperspectral data that has bands in the infra-red range that extends from the nominal red edge of the visible spectrum at 750 nm to 1150 nm. The mean of all the bands is computed to gather the average spectral information of near-infrared range. As the hyperspectral data available has only infrared range information, therefore, the red band information is extracted from VIS RGB data for computing the

Fig. 3 NDVI image
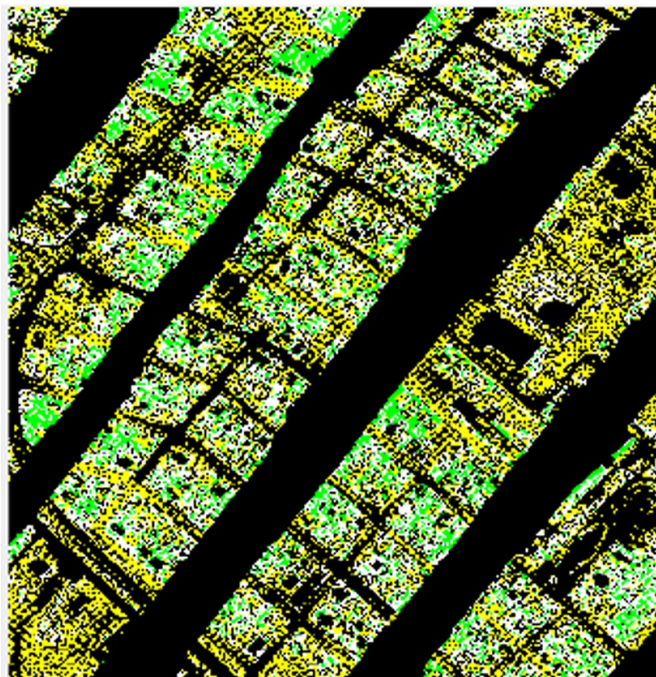
FIG. 4 MBI image
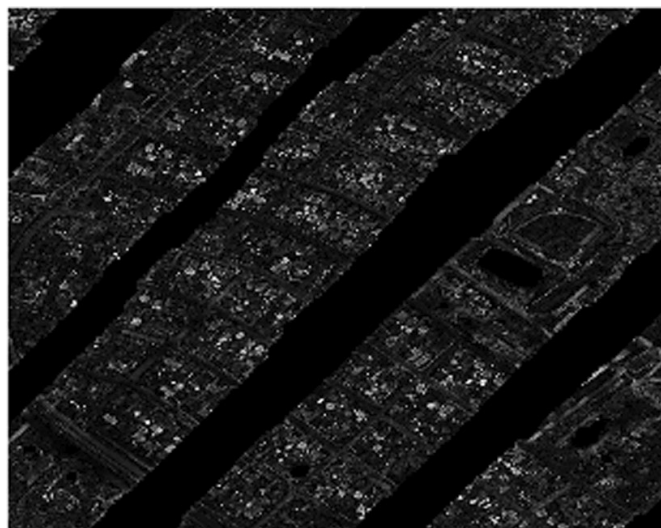
vegetation index. For the depiction of healthy vegetation, the maximum absorption in red band and maximum reflectance in the infra-red band is band required, hence, the fusion of bands from LWIR and VIS RGB data has complemented the extraction of computing the vegetation index in the range from $+1$ to -1. Figure 3 shows the computed NDVI image.

$$NDVI = \frac{\text{mean spectra of NIR Bands of LWIR data } (750 - 11500nm)}{\text{Red Band of VIS RGB data}} \tag{5}$$

(b) *Morphological Building Index* MBI indicates buildings directly and automatically by describing their spectral, spatial characteristics. Figure 4 shows the MBI image obtained after the mathematical computations. MBI index can emphasize

the bright structures with high local contrast, corresponding to the potential buildings. In our study, the values of MBI lie in the range from 0 to 21.159. The drastic changes over the values in the matrix from the low contrast to high contrast are segmented as buildings.

(c) *Textural Features* Another set of features is obtained from the textural analysis by calculating the second-order statistical textural features using GLCM's. It is a way of extracting statistics such as contrast and homogeneity which helps in finding the sharp edges of different objects or to find out the similarity between the connected pixels. These textural features are obtained from VIS RGB data. The total number of statistical values (contrast and homogeneity) obtained is 486. The window size applied to compute the textural features is 3×3. This is applied to VIS RGB data, resized to 25×25 due to lower computational issues. The total number of windows that run across the data horizontally and vertically would be 9 and hence the computed size of the feature data set would be 27×27 for all the bands of VIS RGB data. GLCM is applied to all the bands to find the occurrence and co-occurrence statistics.

(d) *Generation of Region of Interest* The ground truth data that consists of a region of interest of all the five classes, namely, building, road, soil, trees, and vegetation is generated from VIS RGB data through ENVI software. From the region of interest through ENVI software, we get the information about the total number of pixels selected for the object along with their spatial coordinates. This data is converted into text files. These five text files obtained through ENVI are then displayed on MATLAB as black and white images, where the coordinates of the region of interest are highlighted as whites and the rest of the spatial coordinates are considered as a black background.

Step II*: Bounding Box Extraction from Connected Component Labelling* The spatial and spectral features computed from already existing mathematical formulations are then incorporated in the connected regions of the bounding box. Hence, in this manner, all the connected components in the bounding box will have all feature index values for each class. The spatial intensity and spectral reflectance values of each feature are extracted in the bounding box of connected components for all classes under the category of natural and man-made objects. The total connected components for each class are shown below in Table 1. Form the total number of

Table 1 Total connected components into training and testing samples

Classes	Total connected components	Training samples	Testing samples
Building	456	274	182
Road	363	218	145
Soil	140	84	56
Trees	552	331	221
Vegetation	594	356	238
Total	2105	1263	842

connected components (i.e., 2105), each connected component is further divided into two parts i.e., training samples and testing samples. The training samples are 60% of connected components of each class and testing samples are 40% of the connected components of each class. The division of the total number of connected components into training and testing samples is shown in Table 1.

Step III*: SVM Classifier* Since, the data is to be classified into five classes; therefore, in total 5 class indexing is done. The class index 1 is given to class building, 2 to roads, 3 to the soil, 4 to the trees and 5 to the vegetation. The training SVM network in our paper has 20 binary pairings i.e., class 1 paired with class 2, class 3, class 4 and class 5; in similar manner, pairing is done for other rest of the classes. In order to train the network, class labels are generated onto the training bound box feature sets obtained from vegetation features, MBI and textural features. For each binary pair, the network provides support vectors for each class. The support vectors calculated for each class are the data points on the hyperplane for supporting the class. Given below are the Tables 2, 3, 4 showing the support vectors generated on three training feature sets.

The testing bound box feature set is classified based on a support vector training database and the final assignment of the testing sample is given to a class based on the relationship between the support vectors distant from the hyperplane to which the testing class matches with. The following Tables 5, 6, 7 show the confidence measure of each class against the other class from SVM generated network which is based upon each feature set analysis.

From the above analysis shown in Tables 5, 6, 7, the best confidence measure for each class is chosen and listed below in Table 8 indicating the best classification accuracy obtained of each classified class with respect to each feature set.

Step IV*: LM-ANN* The network training function used in MATLAB updates the weight and bias values according to the LM optimization technique. Back-propagation is used to calculate Jacobian parameters with respect to weight and bias

Table 2 SVM training database by vegetation index

Building (Class 1)	SVM training data	Road (1,2)	Soil (1,3)	Tree (1,4)	Vegetation (1,5)
	Support vectors	247	151	21	251
Road (Class 2)	SVM training data	Building (2,1)	Soil (2,3)	Tree (2,4)	Vegetation (2,5)
	Support vectors	247	196	248	213
Soil (Class 3)	SVM training data	Building (3,1)	Road (3,2)	Tree (3,4)	Vegetation (3,5)
	Support vectors	151	196	209	201
Tree (Class 4)	SVM training data	Building (4,1)	Road (4,2)	Soil (4,3)	Vegetation (4,5)
	Support Vectors	359	248	209	394
Vegetation (Class 5)	SVM training data	Building (5,1)	Road (5,2)	Soil (5,3)	Tree (5,4)
	Support vectors	252	213	201	394

Table 3 SVM training database by textural features

Building (Class 1)	SVM training data	Road (1,2)	Soil (1,3)	Tree (1,4)	Vegetation (1,5)
	Support vectors	290	163	413	420
Road (Class 2)	SVM training data	Building (2,1)	Soil (2,3)	Tree (2,4)	Vegetation (2,5)
	Support vectors	392	142	281	456
Soil (Class 3)	SVM training data	Building (3,1)	Road (3,2)	Tree (3,4)	Vegetation (3,5)
	Support vectors	174	173	256	150
Tree (Class 4)	SVM training data	Building (4,1)	Road (4,2)	Soil (4,3)	Vegetation (4,5)
	Support vectors	280	281	256	449
Vegetation (Class 5)	SVM training data	Building (5,1)	Road (5,2)	Soil (5,3)	Tree (5,4)
	Support vectors	420	390	253	346

Table 4 SVM training database by morphological building index

Building (Class 1)	SVM training data	Road (1,2)	Soil (1,3)	Tree (1,4)	Vegetation (1,5)
	Support vectors	92	166	385	137
Road (Class 2)	SVM training data	Building (2,1)	Soil (2,3)	Tree (2,4)	Vegetation (2,5)
	Support vectors	282	203	310	288
Soil (Class 3)	SVM training data	Building (3,1)	Road (3,2)	Tree (3,4)	Vegetation (3,5)
	Support vectors	199	203	253	245
Tree (Class 4)	SVM training data	Building (4,1)	Road (4,2)	Soil (4,3)	Vegetation (4,5)
	Support vectors	382	310	253	357
Vegetation (Class 5)	SVM training data	Building (5,1)	Road (5,2)	Soil (5,3)	Tree (5,4)
	Support vectors	298	288	245	357

variables. The training function support training and testing class labels using data division function in MATLAB. The validation vector is used to stop the training if the performance of the network fails to improve or remains the same for maximum epochs. Test vectors are used as a further check that networking is generating well. Table 9 shows the dimension parameters of LM ANN based on spatial and spectral feature sets.

The training and testing samples are the same as discussed above in step II of this section. From the training data knowledge base, the testing class labels are classified by the LM ANN network. The best confidence measure for each class with respect to each feature set is obtained and listed in Table 10.

Table 5 Confidence measure of one class against another class using NDVI feature

Class	Accuracy			
Building	Road (%)	Soil (%)	Tree (%)	Vegetation (%)
	69.2	63.1	100	59.3
Road	Building	Soil	Tree	Vegetation
	62.0	36.5	67.5	66.9
Soil	Building	Road	Tree	Vegetation
	82.1	64.2	73.2	75.0
Tree	Building	Road	Soil	Vegetation
	73.3	75.5	57.9	57.9
Vegetation	Building	Road	Soil	Tree
	87.3	90.3	74.7	65.1

Table 6 Confidence measure of one class against another class using MBI feature

Class	Accuracy			
Building	Road (%)	Soil (%)	Tree (%)	Vegetation (%)
	100.0	100.0	93.4	100.0
Road	Building	Soil	Tree	Vegetation
	53.7	36.5	46.9	53.7
Soil	Building	Road	Tree	Vegetation
	76.7	66.0	73.2	75.0
Tree	Building	Road	Soil	Vegetation
	69.6	68.7	47.0	60.1
Vegetation	Building	Road	Soil	Tree
	77.7	75.6	70.1	68.9

Table 7 Confidence measure of one class against another class using TEXTURAL feature

Class	Accuracy class			
Building	Road (%)	Soil (%)	Tree (%)	Vegetation (%)
	63.7	73.6	100.0	54.4
Road	Building	Soil	Tree	Vegetation
	0.0	62.7	56.5	75.1
Soil	Building	Road	Tree	Vegetation
	51.7	19.6	39.2	100.0
Tree	Building	Road	Soil	Vegetation
	75.5	73.7	61.5	58.3
Vegetation	Building	Road	Soil	Tree
	72.6	64.2	60.0	0.02

Table 8 Highest confidence measure of each class from spectral and spatial features

Class	Accuracy		
	Vegetation index (%)	Textural features (%)	MBI features (%)
Building	100	100	100
Road	67.5	53.7	75.1
Soil	82.1	76.7	100
Tree	75.5	69.6	75.5
Vegetation	90.3	77.7	72.6

Table 9 Dimensions of LM ANN based on NDVI, MBI and textural features

Dimensions of LM ANN			
Number of inputs	01		
Number of hidden layers	02		
Number of outputs	01		
Number of weight elements	NDVI	MBI	Textural
	1255	1255	977

Table 10 Highest confidence measure of each class from spectral and spatial features

Class	Accuracy		
	Vegetation index (%)	Textural features (%)	MBI features (%)
Building	48	51	76
Road	55	90	100
Soil	100	100	100
Tree	100	100	100
Vegetation	100	100	100

Step V*: Majority Voting Fusion* The voting is given to the most frequent of the class that is resulted in binary SVM classifier and ANN classifier network to improve the classification accuracy. The individual probabilities (confidence measures) from all such pairs of binary SVMs and ANNs are combined to uniquely represent the object feature to one of the classes. After summing up all the probabilities, the class with the highest probability value (confidence measure) represents the object. The benefit of using the weighted majority voting is that to decide the true class through the decision from all possibilities. The confidence majors from classifiers are the weights in weighted majority voting. Therefore, in case of noisy data if one or two pair of SVM and ANN classifiers comes with low probability, still the high probability of all other classifier pairs make it fall into true class. It fails only if all the pair of classifiers comes with low probability and an unstable decision. Finally, weighted majority voting performance is

measured for each feature and as well as for combined features for binary SVMs and ANNs. The producer's accuracy is calculated which is defined in terms of the true positive values i.e., the number of correctly classified samples to the total number of samples of a class. The majority rule is applied through mode function in MATLAB which identifies the most frequent class occurring in the SVM classified result and LM-ANN through three feature sets. Tables 11 and 12 gives the classification accuracy by majority voting rule on each feature set. Table 13 shows the comparative analysis of classification accuracy by two classifiers (SVM and LM ANN) obtained from the majority voting rule on the combination of all features.

Following are the observations drawn from the tabulated results on classification accuracy achieved by SVM and ANN classifier using multilevel majority voting fusion:

(a) From the calculations shown in Tables 11 and 12, it is seen that the classification accuracy is 100% for all the classes under the natural object category. This may be because ANN's often converge on local minima rather than global minima. Hence, ANN outperforms SVM classifier for the classification of natural objects under spectral and spatial feature analysis.

(b) The comparison of the ANN classifier result shown in Tables 12 and 13 shows that the classification performance of building and road has not been improved by using the decision of WMV fusion of classifier networks based on a combination of the spatial and spectral feature set analysis. Hence, ANN's may not require fusion of classifiers' decisions based on the combination of multiple features to enhance the classification accuracy of the objects in urban land classification.

(c) The performance parameters shown in Table 13 which are obtained through majority voting fusion of MBI features for classification of man-made and natural object classification has outperformed the results in comparison to the rest of the features i.e., NDVI and textural feature analysis. This shows that in our study ANN has shown the best performance for the classification of natural and man-made objects using morphological spatial-spectral feature analysis.

(d) The performance parameters shown in Table 11 which are obtained through majority voting fusion of NDVI features for classification of man-made and natural object classification has outperformed the results in comparison to the rest of the features i.e., MBI and textural feature set analysis. This shows that in our study SVM classifier based on NDVI feature set analysis has shown the best performance for classification of natural and man-made objects.

(e) The comparative performance of two classifier networks shown in Table 13 shows that the classification accuracy of man-made objects i.e., building and road (100% and 89.6%) using SVM classifier network that is based on majority voting fusion of all features is more as compared to the classification accuracy results (of building and road is 55.4% and 67.95%) obtained using ANN classifier network based on majority voting fusion of all features. This shows that SVMs outperform ANNs if classification is carried on using the combination of decision level fusion of spatial and spectral feature analysis. Hence, for the

Table 11 Classification accuracy of SVM (Testing) network for natural and man-made objects using majority voting fusion of spectral and spatial features

Objects	Classes	Classifier's Accuracy		
		SVM classifier by weighted majority voting on vegetation index (%)	SVM classifier by weighted majority voting on MBI (%)	SVM classifier by weighted majority voting on textural features (%)
Man-made	Building	100.0	100.0	100.0
	Road	79.3	66.9	66.8
Natural	Soil	91.0	85.7	57.1
	Trees	82.8	75.1	83.7
	Vegetation	91.6	84.8	75.6

Table 12 Classification accuracy of SVM (testing) network for natural and man-made objects using majority voting fusion of spectral and spatial features

Objects	Classes	Classifier's Accuracy		
		ANN classifier by majority voting on vegetation index (%)	ANN classifier by weighted voting on MBI (%)	ANN classifier by majority voting on textural statistics (%)
Man-made	Building	62	81.3	54.9
	Road	28	66.9	53.9
Natural	Soil	100	100	100
	Trees	100	100	100
	Vegetation	100	100	100

Table 13 Comparative assessment of classification accuracy of SVM and ANN classifiers for natural and man-made objects

Objects	Classes	Classifier's Accuracy	
		SVM-based weighted majority voting fusion using all features (%)	ANN-based weighted majority voting fusion using all features (%)
Man-made	Building	100	55.4
	Road	89.6	67.9
Natural	Soil	98.2	100
	Trees	96.8	100
	Vegetation	96.2	100

enhancement in classification accuracy of natural and man-made objects, SVM works well with more exploration of features in the spatial and spectral domain.

(f) The overall accuracy (OA), actual accuracy (AA) and kappa coefficient (K %) has been obtained : (1) simply enlarging the original HS image by cubic interpolation (Cub); (2) PCA component substitution method (PCA); (3) guided filter in PCA domain (GFP); (4) MPs computed on original RGB image (MPs) [34]. The results have been compared with SVM and ANN classifier networks generated using fused features by weighted majority voting. The comparative analysis shows that the fusion of spatial, spectral and spectral-structural features using SVM classifier performs better than other features and fusion schemes (Table 14).

Step VI*: Comparative Assessment of Two Classifiers i.e., SVM and ANN Based on McNemar's Test* The results of classifications from two classifiers are assessed using the confusion matrix in terms of true positives to calculate the overall accuracy. The accuracy assessment also includes the producer's and user's accuracy to evaluate the omission and commission errors for each class which is shown in Tables 15 and 16. To evaluate the variability of the classifications, the same

Table 14 Comparison of classification accuracies of different schemes

Accuracy (%)	Features								
	Cub 1	PCA 2	GFP 3	MPs 4	NDVI 5	Textural 6	MBI 7	SVM-fused features	ANN-fused features
OA	78.1	78.8	91.6	88.7	93.8	76.6	90.3	97.8	91.1
AA	72.4	73.9	84.7	86.4	91.1	73.4	89.8	96.4	89.9
K %	72.8	73.8	89.5	87.8	92.1	72.8	88.8	94.3	88.9

Table 15 Accuracy assessment (confusion matrix in terms of user's accuracy) of natural and man-made classes using SVM classifier

Accuracy assessment	Reference class						
Predicted class	Natural and man-made classes	Building	Road	Soil	Tree	Vegetation	Producer's accuracy (%)
	Building	**1**	0	0	0	0	88.35
	Road	0.0966	**0.8897**	0.0138	0	0	89.22
	Soil	0	0.0357	**0.9643**	0	0	95.09
	Tree	0.0226	0.0633	0.0317	**0.8824**	0	99.05
	Vegetation	0.0126	0.0084	0.0042	0.0084	**0.9664**	100
Accuracy	Overall accuracy = 94.05%						

The bold values signify the maximum value of the classification accuracy attained in comparison to other classes

Table 16 Accuracy assessment (confusion matrix in terms of user's accuracy) of natural and man-made classes using ANN classifier

Accuracy assessment	Reference class						
Predicted class	Natural and man-made classes	Building	Road	Soil	Tree	Vegetation	Producer's accuracy (%)
	Building	**0.5549**	0.4451	0	0	0	63.1
	Road	0.3241	**0.6759**	0	0	0	60.29
	Soil	0	0	**1**	0	0	100
	Tree	0	0	0	**1**	0	100
	Vegetation	0	0	0	0	**1**	100
Accuracy	Overall accuracy = 84.6%						

The bold values signify the maximum value of the classification accuracy attained in comparison to other classes

set of samples are used in the experiment. Since the reference data is not independent, the statistical significance of the difference between two classifications is evaluated using McNemar's test. The McNemar test calculates the z value [35] which is shown in Table 17. If z ≥ 3.2, this demonstrates a significant difference between two classifiers at the 99% confidence level   [36]. Here, a fully rigorous and exhaustive approach is adopted for expressing the statistical significance of classification output differences.

The SVM classifier is now consistently (often significantly) more accurate than the ANN classifier (as shown by the confusion matrix accuracy statistics). For instance, using the majority voting feature fusion, the overall SVM accuracy is 94.05%, compared to overall ANN accuracy of 84.6%.

The magnitude of differences is quantified with accuracy assessment which is shown in Tables 15, 16, and the statistical significance of the differences between the two classifiers is evaluated using McNemar's test. From overall accuracy statistics, it is seen that the SVM classifier has shown better results than the ANN classifier when trained using the combination of spatial and spectral features by object-based classification technique.

The choice of classifier is important in determining the success of classifying complex urban land cover and the optimum choice will vary depending on image data characteristics. The results deduced from McNemar's test show that there lies a significant difference in the statistical analysis of three classes out of five classes i.e., the classified results for building, road and trees have a high statistical difference if two non-parametric classifiers are compared. The class vegetation shows the marginal difference as the value of z appears to be 3 which is borderline. Hence, from the accuracy assessment and McNemar's test, it can be deduced that if a combination of spectral and spatial features are extracted from LWIR and VIS RGB data then SVM shows good performance than ANN classifier.

6 Conclusion

This paper presents a multi-level fusion of multiresolution and multisensory data to classify the urban land cover features. The spatial, textural and spectral characteristics from the VIS RGB and LWIR data sets have enhanced the feature knowledge database that has helped in improving the classification accuracy which can be seen from the quantitative assessments. From the evaluations, in terms of classification

Table 17 Z value of each class using McNemar test

Classes	Z value
Building	9.0
Road	4.0640
Soil	1.0
Tree	2.6458
Vegetation	3.0

accuracy, it emerges that the SVM classifier network outperforms man-made objects classification in comparison to the non-parametric ANN classifier network. ANN classifier network outperforms for natural objects classification in urban land cover. Also, the majority voting fusion of all sets of features affects the performance measure of the SVM classifier network more than ANN for enhancing the classification, whereas, this additional step is not required for the ANN classifier network. SVM classifier shows a phenomenal change in the performance measures due to the majority voting fusion of all the features. Hence, the novelty in characterizing the image objects to a class by multilevel fusion of multiresolution and multi-sensor data has proved that SVM and LM-ANN classifiers show remarkable improvement in the overall accuracy of urban land surface objects. In short, the fusion of LWIR and VIS RGB dataset seems promising and their characteristics can be well explored for classification of urban areas.

References

1. Rodriguez-Galiano, V. F., Ghimire, B., Pardo-Iguzquiza, E., Chica-Olmo, M., & Congalton, R. G. (2012). Incorporating the downscaled Landsat TM thermal band in land-cover classification using random forest. *Photogrammetric Engineering & Remote Sensing, 78*(2), 129–137.
2. Luo, Xiaoqing, Zhang, Zhancheng, & Xiaojun, Wu. (2016). A novel algorithm of remote sensing image fusion based on shift-invariant Shearlet transform and regional selection. *AEU- International Journal of Electronics and Communication., 70*(2), 186–197.
3. Berger, C., Riedel, F., Rosentreter, J., Stein, E., Hese, S., & Schmullius, C. (2015). Fusion of airborne hyperspectral and LiDAR remote sensing data to study the thermal characteristics of urban environments. In Computational Approaches for Urban Environments (pp. 273-292). Springer, Cham.
4. Ma, Jiayi, Ma, Yong, & Li, Chang. (2018). Infrared and visible image fusion methods and applications: a survey. *Information Fusion, 45,* 153–178.
5. Dianat, R., & Kasaei, S. (2010). Dimension reduction of remote sensing images by incorporating spatial and spectral properties. *AEU-International Journal of Electronics and Communication, 64*(8), 729–732.
6. Zhang, Lefei, & Tao, Dacheng. (2012). On combining multiple features for hyperspectral remote sensing image classification. *IEEE Transaction on Geoscience and Remote Sensing., 50*(3), 879–893.
7. Bruzzone, L., & Bovolo, F. (2013). A novel framework for the design of change-detection systems for very-highresolution remote sensing images. *Proceedings of the IEEE, 101*(3), 609–630.
8. Huang, X., Zhang, L., & Zhu, T. (2014). Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing., 7*(1), 105–115.
9. Tian, J., Reinartz, P. (2011). Multitemporal 3D change detection in urban areas using stereo information from different sensors. In International Symposium on Image and Data Fusion (pp. 1–4).
10. Gao, Jianqiang, & Lizhong, Xu. (2015). An efficient method to solve the classification problem for remote sensing. *AEU- International Journal of Electronics and Communication, 69*(1), 198–205.
11. Moser, G., & Serpico, S. B. (2013). Benediktsson. *Land cover mapping by markov modelling of spatial contextual information in very high-resolution remote sensing images, Proceedings of IEEE., 101*(3), 631–651.
12. Eslami, M., & Mohammadzadeh, A. (2015). Developing a spectral-based strategy for urban object detection from airborne hyperspectral TIR and visible data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(5), 1808–1816.
13. Li, J., Zhang, H., Guo, M., Zhang, L., Shen, H., & Du, Q. (2015). Urban classification by the fusion of thermal infrared hyperspectral and visible data. *Photogrammetric Engineering & Remote Sensing, 81*(12), 901–911.
14. Lu, X., Zhang, J., Li, T., & Zhang, G. (2015). Synergetic classification of long-wave infrared hyperspectral and visible images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(7), 3546–3557.

15. Melgani, F., & Serpico, S. B. (2002). A statistical approach to the fusion of spectral and spatio-temporal contextual information for the classification of remote-sensing images. *Pattern Recognition Letters, 23*(9), 1053–1061.

16. Bardossy, A., & Samaniego, L. (2002). Fuzzy rule-based classification of remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing, 40*(2), 362–374.

17. Bruzzone, L., & Cossu, R. (2002). A multiple-cascade-classifier system for a robust and partially unsupervised updating of land-cover maps. *IEEE Transactions on Geoscience and Remote Sensing, 40*(9), 1984–1996.

18. Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification Performance. *International Journal of Remote Sensing, 28*(5), 823–870.

19. Won Hee Lee. (2015). Object-oriented classification of urban areas using lidar and aerial images. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography., 33*(3), 173–179.

20. IEEE GRSS Data Fusion Contest, Presented to Image Analysis and Data Fusion Technical Committee, IEEE Geoscience and Remote Sensing Society (GRSS) (2014).

21. Haboudane, Driss. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment, 90,* 337–352.

22. Huang, Xin, & Zhang, Liangpei. (2012). Morphological building/shadow index for building extraction from high- resolution imagery over urban areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing., 5*(1), 161–172.

23. Liao, Wenzhi, & Bellens, Rik. (2012). Classification of hyperspectral data over urban areas using directional morphological profiles and semi-supervised feature extraction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing., 5*(4), 1177–1190.

24. Marceau, D., Howarth, P., Dubois, J., & Gratton, D. (1990). Evaluation of the grey-level co-occurrence matrix method for land-cover classification using Spot imagery. *IEEE Transaction of Geoscience Remote Sensing., 28*(4), 513–519.

25. Askarpour, S., Anari, M. S., Brumandnia, A., & Javidi, M. M. (2014). The analysis of connected components and clustering in segmentation of persian texts. *Research Journal of Recent Sciences, 3*(4), 71–72.

26. XinHuang, Liangpei Zhang. (2013). An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing, 51*(1), 257–272.

27. Luo, Shezhou, Wang, Cheng, Xi, Xiaohuan, Zeng, Hongcheng, Li, Dong, Xia, Shaobo, et al. (2015). Fusion of airborne discrete-return LiDAR and hyperspectral data for land cover classification. *Remote Sensing, 8*(1), 3–12.

28. Melgani, Farid. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*(8), 1778–1790.

29. Yuan, Hui, & Van Der Wiele, Cynthia F. (2009). An automated artificial neural network system for land use/land cover classification from LANDSAT TM imagery. *Journal of Remote Sensing. ISSN, 2072–4292,* 243–265.

30. Madsen, K., Neilsen, H. B. (2004) Methods for non-linear least squares problems. 2nd Edition.

31. Licciardi, G. (2009). Decision fusion for the classification of hyperspectral data: outcome of the 2008 GRS-S data fusion contest. *IEEE Transactions on Geoscience and Remote Sensing, 47*(11), 3857–3865.

32. Santra, A. K., & Christy, C. J. (2012). Genetic algorithm and confusion matrix for document clustering. *International Journal of Computer Science Issues., 9*(1), 322–328.

33. Singh, Kunwar K., Vogler, John B., Shoemaker, Douglas A., & Meentemeyer, Ross K. (2014). LiDAR-Landsat data fusion for large-area assessment of urban land cover: Balancing spatial resolution, data volume and mapping accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing., 74,* 110–121.

34. Liao, W., et al. (2014). IEEE data fusion best paper contest.

35. Momeni, R., Aplin, P., & Boyd, D. (2016). Mapping complex urban land cover from spaceborne imagery: the influence of spatial resolution, spectral band set and classification approach. *Remote Sensing, 8*(2), 88.

36. Rees, D. G. (1987). *Foundations of statistics*. London: Chapman & Hall.

Constraints for effective distribution network expansion planning: an ample review

Mandhir Kumar Verma[1] · Vivekananda Mukherjee[1] · Vinod Kumar Yadav[2] ·
Santosh Ghosh[3]

Abstract The efficacy and practical relevance of the solution of any technical problem depends not only upon effective formulation of an objective function, but also on the choice of constraints, which limits the solution space, and same is the case with distribution network expansion planning (DNEP) problems as well. The objective functions of DNEP problems need to be bound by appropriate technical, social and economic constraints to arrive at the optimal solution for any site. Constraints are the conditions imposed in DNEP on various objective functions to achieve the most cost-effective, consistent and environment-friendly optimized solution. Hence the choice of appropriate constraints has a significant role in the execution of DNEP problems effectively. The present work aims to summarize the constraints considered in more than eighty selected original articles published over a decade. The constraints considered by various eminent researchers have been discussed by classifying them into technical, non-technical, time & financial and miscellaneous constraints. Special considerations required for deciding constraints for DNEP concerning assimilation of distributed generation (DG) and distributed energy storage system (DESS) have been discussed as well. The summary of constraints, through comprehensive review of the gigantic volume of literature on DNEP published over the past decade will help enduring and budding researchers in this field, pinpointing any breaks in reviewed literature that can be address in impending research for further enhancement and can be employed to arrive at an optimum solutions to DNEP problems.

Keywords Distribution network expansion planning (DNEP) · Constraints · Distributed generation (DG) · Distributed energy storage system (DESS) · Technical and non-technical constraints

List of symbols

✉ Mandhir Kumar Verma
  mandhir.verma@glbajajgroup.org

  Vivekananda Mukherjee
  vivek_agamani@yahoo.com

  Vinod Kumar Yadav
  vinodkumar@dtu.ac.in

  Santosh Ghosh
  santosh.ghosh@kbl.co.in

[1] Department of Electrical Engineering, Indian Institute of Technology, Dhanbad, Jharkhand 826004, India

[2] Department of Electrical Engineering, Delhi Technological University, New Delhi, Delhi 110042, India

[3] Corporate R&D Department, Kirloskar Brothers Limited, Pune 411045, India

Indices

b, c, d	Index of branch
g	Index of alternative
h	Index for hour h in a typical day
i, j	Index of nodes in the distribution network
k, l	Index of nodes to build up DESS
k	Index of bus
L	Index of line
β	Index of binary variables
ss	Index of substation
s	Index of stage
sc	Index of typical daily load scene
u	Index of unit

Sets

ψ_{AN1}, ψ_{AN2}	Set of branches of the additional network and alternatives for each branch d
$\psi_{B1}, \psi_{B2}/N_{Br1}, N_{Br2}$	Set/number of primary network branches (existing and proposed)
ψ_{BFN1}, ψ_{BFN2}	Set of branches of the fixed network and distinct changes for each branch b
$\psi_{CESS,s}$	Set of options of building up a DESS
ψ_{N2}^{DG}	Set of distributed generation nodes
ψ_{N2}^{DESS}	Set of nodes to build up DESS
ψ_{DGA}	Set of substitutes for DG
ψ_{F1}	Set of feeder of system
ψ_k	Set of elements connected to bus k
$\psi_{L1}, \psi_{L2}/N_{L1}, N_{L2}$	Set/number of lines (existing and proposed)
ψ_{N1}, ψ_{N2}	Set of number of bus substation nodes (existing and planned)
ψ_{NW2}	Set of new wirings
ψ_{RN1}, ψ_{RN2}	Set of branches of replacement network and alternatives for each branch c
$\psi_{SS1}, \psi_{SS2}/N_{SS1}, N_{SS2}$	Set/number of substations (existing and planned)
ψ_s	Set of stages
ψ_W	Set of available types of wires for planning and secondary systems project

Constants and variables

ac	Modifications defined for branch c of addition network
C	Charging power of DESS nodes
C_i	Binary variable for capacitor allocation at node i
CRS_z	Customer reliability of service at load point z
$C_{Bud,s}$	Budget allotted for stage s
$C_{Opp,s}$	Maintenance and losses cost at stage s
$C_{Inst,s}$	Installations and reconductoring cost at stage s
D	Discharging power of DESS nodes
E	Energy stored in DESS
fb	Modifications defined for branch b of fixed network
$g_{l,t}^S, g_{max,l,t}^S, g_{max,l}^{SL}$	Generation at node l at stage t and substations' maximum capacity
I_f^{Low}, I_f^{Up}	Lower and upper bound of current magnitudes at feeder section f

$I_{b,s}^F$	Currents of branch b in the fixed network at stage s
$I_{c,s}^R$	Current of branch c in the replacement network at stage s
$I_{d,s}^A$	Current of branch d in the addition network at stage s
$IC_{DESS}, IC_{DESS}^{max}$	Capacity (& maximum capacity) of installed DESSs
M	Big number used in disparities of disjunctive limitations
N_{B1}, N_{B2}	Number of existing buses and new buses
N_{CT}	Number of circuit types
N_{CP2}^{max}	Maximum number of capacitors that can be added on to a system
$N_{CP2,i}^{max}$	Maximum number of capacitors that can be installed in node i of system
N_{DG2}, N_{DG2}^{max}	Number (and Maximum number) of new DG units
N_{G1}	Number of generating units
$N_{H,miss}$	Number of hours during which power not served
N_{L2}, N_{L2}^{max}	Number (and maximum number) of lines rewiring
N_{LP2}, N_{LP2}^{max}	Number (and maximum number) or new connected load points
N_{LB1}	Number of load buses
N_{LS}	Number of load steps refer to load duration curve (LDC)
N_{PY}	Number of years of planning limit
$N_{recon}, N_{recon}^{max}$	Number (and maximum number) of network reconfigurations (switches operations)
N_{SW2}, N_{SW2}^{max}	Number (and maximum number) of new switches
N_{St}	Number of stages
N_{TR2}	Number of transformers that can be added to a substation
N_{VR2}^{max}	Maximum number of VR that can be added to the system
$P_{DG,i}$	Active power generation of DG at node i
$P_{DG,i}^{cur}, Q_{DG,i}^{cur}$	Active and reactive power generation curtailment of DG at node i
$P_{DG,i,s,t}$	Output active power of DG at node i and load step s of year t
$P_{L,i}$	Active load at node i at time t
$P_{L,i}^{cur}, Q_{L,i}^{cur}$	Active and reactive load curtailment on node i
$P_{L,k}, Q_{L,k}$	Active and reactive load at bus k
P_{LD}	Active load demand

$P_{b,f,ss}, Q_{b,f,ss}$	Active and reactive power flow in branch b of the network feeder f at substation ss
$P_{ij,f,ss}, Q_{ij,f,ss}$	Active and reactive power flow between node ij resp
P^1_{Purch}	Active power procured from an upstream network in all states
$P_{G,k}, Q_{G,k}$	Active and reactive generation at bus k
$P_{loss,s,t}$	Active power losses
$P_{loss,s,t}$	Power losses of the distribution system at stage s of year t
P_{SR}	Spinning reserve
$Q_{DG,i,s,t}$	Output reactive power of DG in the bus i and load step s of year t
rd	Modifications defined for branch d of the replacement network
$[S^F], [S^R], [S^A]$	Node-branch incidence matrix for the fixed, replacement and addition network
S^{max}_b	Maximum load flow in branch b
S_{LD}	Load demand
$S_{DG,i}, S^{max}_{DG,i}$	Distributed generation at node i stage s and its maximum capacity
S^1_{Purch}	Power procured from an upstream network in all states
$S_{b,f,ss}, S^{max}_{b,f,ss}$	Power (and maximum power) flow in branch b of the network feeder f at substation ss
S^{Tran}_{SS}	Output power of the substation transformer
$S^{Tran}_{SS,max}$	Maximum output power of the substation transformer
S_{SS}, S^{max}_{SS}	Capacity (and upper bound capacity) of substation ss
$S_{LD,s,t}$	Total load demand of distribution system at stage s of year t
$S_{G,u}, S^{max}_{G,u}$	Generation capacity (and maximum generation capacity) of unit u
$S_{NLS,s}$	Vector of nodal-load shedding at stage s
$S_{L,s}$	Vector of nodal load at stage s
T_{ij}	Limitation of the absolute value of active power flow between node ij
T_P	Time horizon of planning
T_{OH}	Outage hours of load node
$T_{on,u}, T_{off,u}$	On/off time of unit u
$T_{MUT,u}, T_{MDT,u}$	Minimum-up/-downtime of unit u
UR_u, DR_u	Up and down rate of unit u
$v_{VR,c1}$	Binary variable allied with VR allocation in circuit c1
V^{Low}_i, V^{Up}_i	Lower and upper bound of nodal voltage magnitudes at node i
$V_{j,d}, \tilde{V}_{j,d}$	Non-regulated and regulated voltage magnitude
$V_{NVI,s}$	Column vectors of the nodal voltages injection at stage s
x	Available area of substations for expansion
$x^{SL}_{i,s}$	Alternative SL for of substation node i expansion at stage s
$x[0]$	Initial network
$x[1]$	Target network
y	Validation flag of lines
$y^{fb}_{b,s}$	Alternative fb for the deployment of branch b at stage s
$y^{ac}_{c,s}$	Alternative ac for the deployment of additional branch c at stage s
$y^{rd}_{d,s}$	Alternatives rd for the deployment of replacement branch d at stage s
$y_{b,sc}$	Decision variable to replace the cable in branch b of the secondary circuit sc
Z_{ij}	Impedance between node ij
Z^{F1}_b	Impedance of alternative related to branch b of fixed network
Z^{R0}_b, Z^{R1}_b	Initial and alternative impedances related to branch b of the replacement network
Z^{A1}_b	Impedance of alternative AK related to branch b of addition network
η	Charging/ discharging efficiency of DESS
$\beta^{cb}_{i,s}$	Binary variable allied with capacitor bank allocation at node i at stage s
$\beta^{DG}_{i,g,s}$	Binary variable allied with DG allocation at node considering the substitute at stage s
σ	Fabricated coefficient to measure a reduction in outage hours
θ_i	Voltage angle at node i
$\emptyset^L_i, \emptyset^{DG}_i$	Power factor angle of load and DG on node i

1 Introduction

To cope up with the ever-increasing population and expanding industrialization, the electric power systems need to be expanded in all of its three dimensions viz. generation, transmission, and distribution. To reduce the infrastructure cost and high losses incurred during power transmission, policymakers all over the world are endeavoring to bring the point of power generation closer to the consumption center. Hence the distribution systems are becoming a complete power system (micro-grids) in

themselves. Hence its planning process needs to be more robust and to keep a check at all aspects; the constraints need to be chosen expansively and carefully.

In the present work, a comprehensive review of selected original research works (more than eighty) spread over a decade, on DNEP constraints, have been presented. After deciding the objective functions for optimal DNEP, the significant and foremost task is to bind those outputs of objective functions with suitable constraints appropriate for the particular site being planned. To fulfill the objectives of DNEP, different constraints have been considered by researchers and categorized based on different criteria. For example, Murty and Kumar (2015) and Humayd and Bhattacharya (2017) have differentiated constraints according to their degree of fulfillment. Those constraints which have to be fulfilled 100% are termed as stiff constraints, and those who have satisfaction probability up to some low degree are termed as probabilistic constraints. Xing et al. (2015) have distinguished constraints as technical and non-technical types. The authors Mansor and Levi (2015) and Tabares et al. (2016) classified constraints as hard and soft ones, out of which satisfying the former one is obligatory and latter one can be diluted with the provision of violation minimization. Further, Labbi and Attous (2014) and Bansal et al. (2017) has separated DNEP constraints as equality and non-equality type.

Depending upon the output of objectives, in this article constraints have been classified into five categories. First one is technical constraints (Mori and Yamada 2007; Carrano et al. 2008; Cossi et al. 2009; Ouyang et al. 2010; Aghaei et al. 2014; Xing et al. 2015; Mokryani et al. 2017; Bai et al. 2018; Ugranli 2019; Nahman and Peri 2020) and second one are non-technical constraints (Mori and Yamada 2007; Cossi et al. 2009; Al-Muhaini et al. 2010; Wang et al. 2011; Borges and Martins 2012a; Gitizadeh et al. 2013; Ravadanegh and Roshanagh 2014; Heidari et al. 2015; Tabares et al. 2016; Shen et al. 2018, Quevedo et al. 2019; Xie et al. 2020). The third category is based upon special considerations required for DG and DESS incorporation into DNEP used by Borges and Martins (2012b), Gitizadeh et al. (2013), Huang et al. (2014), Delgado et al. (2015), Tabares et al. (2016), Ahmadigorji et al. (2017), Bai et al. (2018), Mehrjerdi(2019) etc. With more dissemination of DG in the DNEP, the emergence of microgrids has increased. DG has become an inevitable part of modern power systems to mitigate prevailing energy crises. It has attracted attention worldwide due to its ecological as well as economic benefits by its integration with the electricity grid. The diffusion of the distributed energy storing systems (DESSs) with DNEP is taking the lead in providing more effective and feasible solutions. DESS planning and scheduling are becoming not only alternative to DG technically, but it also justifies the higher investment.

Some researchers have also consider time and money related constraints (classified as fourth category) such as

Carrano et al. (2008), Cossi et al. (2009), Al-Muhaini et al. (2010), Cossi et al. (2012), Zonkoly (2013), Sadeghi and Kalantar (2014), Delgado et al. (2015), Gholami et al. (2017), Ugranli (2019), Xie et al. (2020). Rest of all the types have been discussed under the classification of miscellaneous constraints to be considered for steady-state operation of distribution networks (Haffner et al. 2008; Borges and Martins 2012a; Camargo et al. 2013; Saldarriaga et al. 2013; Ravadanegh and Roshanagh 2014; Heidari et al. 2015; Gholami et al. 2017). These are related to type of conductors, available area for expansion, spinning reserve, ramp rate, salvage of the line, etc. which have been considered by very few researchers but are of prominent values.

In view of the above discussion, the main contributions of this work are the followings:

- Nearly eighty acclaimed articles study have been carried out.
- Authors have summarized the DNEP constraints considered in more than seventy selected articles published over a decade and categorized them into five heads viz. *technical constraints, non-technical constraints, DG and DESS constraints, time and money constraints* and *miscellaneous constraints.*
- Constraints can't be classified merely as equality and non-equality type but have to be more diversified so that all-round check can be implemented for higher optimization degree.
- Each stage of DNEP viz—modeling, calculation, multi-scenario simulation, all-inclusive scheduling, and rectification—contains diverse blends of components, elements, strategy, & optimization techniques that influence the ultimate planning conclusion. So the constraints for above mention stages must be precisely defined separately.

Ensuing this preliminary section, the rest of the article is structured as follow. Section 2 is dedicated to a brief overview of the significant classification of DNEP constraints. The first classification, i.e., technical constraints, is described in detail in Sect. 3, followed by a summary of non-technical constraints presented in Sect. 4. Constraints related to DG and DESS are précised in Sect. 5. A few investigators have also considered time and financial constraints, which are discussed in Sect. 6, followed by all types of miscellaneous constraints in Sect. 7. Finally, Sect. 8 eloborates all findings and authors' views.

2 DNEP constraints

While executing DNEP, the planner has to find a solution within certain prescribed natural and human-made limits. The same restrictions are termed as constraints when take the shape of mathematical equations and incorporated

within the objective functions of DNEP. These constraints can also be of equality type or inequality type. By putting a cap on various DNEP objectives, the most optimal and appropriate solution can be obtained. While fulfilling the objective of DNEP, many constraints have been considered by various researchers and are categorized according to their implementations. A vast number of constraints considered by various researchers can be broadly divided into the following five categories: *technical constraints, non-technical constraints, DG & DESS constraints, time & money constraints,* and *miscellaneous constraints.*

It can be observed from Table 1 that, survey regarding DNEP constraints such as technical constraints are considered in 94.9% of the reviewed literature, whereas the non-technical constraints were used in 62.03% of reviewed literature. The DG and DESS related constraints, financial and time constraints, and miscellaneous constraints were considered in 50.6%, 24.1%, and 19% of reviewed literature respectively.

3 Technical constraints

To ensure that after taking steps for enlargement of the distribution network and henceforth the system performance meets the technical performance parameters, technical constraints are introduced in the objective function of the DNEP (Table 2). The main technical constraints (which are aslo termed as operation constraints in many articles) incorporated in objective functions of DNEP problem are:

(A1) voltage magnitude limit for nodes and busses (taken with 88.2% of reviewed work), the limit of voltage angle at each bus and is quantified as (Mori and Yamada 2007):

$$V_i^{Low} \leq V_i \leq V_i^{Up} \tag{1}$$

(A2) Power balance Eqs. (2–4) (considered in 62.4% of reviewed work) which comprises of nodal power balance equation constraints (2) (Pudaruth and Li 2007) Kirchhoff's laws constraints (Haffner et al. 2008) derived by KCL (3) and KVL (4):

$$P_{DG,i} + P_{DG,i}^{cur} - \sum_{i=1}^{N_B} \left[\frac{\left(|V_i| - |V_j| \right)^2}{|Z_{ij}|} \right] . pf - P_{L,i}$$
$$- \sum_{i=1}^{N_B} P_{b,f,ss}(V_i, \theta_i)$$
$$= 0 \tag{2}$$

$$[S]^F I_{b,s}^F + [S]^R I_{c,s}^R + [S]^A I_{d,s}^A + V_{NVI,s} + S_{NLS,s} = S_{L,s} \forall t$$
$$= 1, 2, \ldots, T \tag{3}$$

$$\left. \begin{array}{l} \left| Z_b^{F1} I_{b,s}^F + [S^F]_{rowi}^T V_{NVI,s} \right| \leq M \left(1 - y_{b,s}^{fb} \right) \\ \{ \forall i \in \psi_{BFN1}, I \in \psi_{BFN2} \} \\ \left| Z_b^{R0} I_{c,s}^R + [S^R]_{rowj}^T V_{NVI,s} \right| \leq M \left(1 - y_{d,s}^{ro} \right) \\ \{ \forall j \in \psi_{RN1} \} \\ \left| Z_b^{R1} I_{c,s}^R + [S^R]_{rowj}^T V_{NVI,s} \right| \leq M \left(1 - y_{d,s}^{rd} \right) \\ \{ \forall j \in \Psi^R, J \in \psi_{RN2} \} \\ \left| Z_b^{A1} I_{d,s}^A + [S^A]_{rowk}^T V_{NVI,s} \right| \leq M \left(1 - y_{c,s}^{ac} \right) \\ \{ \forall k \in \psi_{AN1}, K \in \psi_{AN2} \} \end{array} \right\} \tag{4}$$

(A3) Branch current flow constraints (Nerves and Julian 2007), (Considered in 49.4% of reviewed work) and are expressed as (5),

$$I_f^{Low} \leq I_f \leq I_f^{Up} \tag{5}$$

(A4) Radiality of the distribution network (considered in 47.1% of reviewed work) is given by (6) based upon graph theory (Xing et al. 2016). Only (6) cannot assure a radial topology arrangement. The network must be connected, and all new loads also have to fulfill power balance constraints.

$$\sum_{i \in \psi_{NW2}} N_{Br2} = N_{LP2} - \sum_{s2 \in \psi_{SS2}} N_{SS2} \tag{6}$$

(A5) Branch thermal limits (Pudaruth and Li 2007) (Considered in 20% of reviewed work) is quantified as:

$$\sqrt{\left[P_{b,f,ss}(V_{ij}, \theta) \right]^2 + \left[Q_{b,f,ss}(V_{ij}, \theta) \right]^2} \leq S_b^{max} \tag{7}$$

(A6) Power flow capacity constraints (Considered in 15.3% of reviewed work) can be quantified by:

$$\left| S_{b,f,ss} \right| \leq S_{b,f,ss}^{max} b \in \psi_{B1}, f \in \psi_{F1}, ss \in \psi_{N1}, \psi_{N2} \tag{8}$$

(A7) The power factor has been considered as a technical constraint in 10.6% of reviewed work, which can be expressed by (9) and (10), where (9) represents power factor equality (Xing et al. 2015) and (10) retains the power factor of DG units persistent for every load step of load duration curve (LDC) during the planning prospects (Ahmadigorji et al. 2017).

$$\left. \begin{array}{l} Q_{L,i}^{cur} = P_{L,i}^{cur} . tan \left(\emptyset_i^L \right) \\ Q_{DG,i}^{cur} = P_{DG,i}^{cur} . tan \left(\emptyset_i^{DG} \right) \end{array} \right\} \tag{9}$$

$$\left. \begin{array}{l} pf_{i,t,s} = \frac{P_{DG,i,s,t}}{\sqrt{\left(P_{DG,i,s,t} \right)^2 + \left(Q_{DG,i,s,t} \right)^2}} = const. \\ \forall i \in N_{LB1}, \forall s \in N_{LS}, \forall t \in N_{PY} \end{array} \right\} \tag{10}$$

(A8) Uni-directional flow (Sadeghi and Kalantar 2014) (Considered in 5.9% of reviewed work) and is expressed as:

$$S_{Purch}^1 \geq 0 \tag{11}$$

Table 1 summary of various constraints from reviewed articles

S.	Constraints	Reviewed articles	x*	x %**
A	Technical	Mori and Yamada (2007), Nerves and Julian (2007), Carrano et al. (2007), Pudaruth and Li (2007), Carrano et al. (2008), Haffner et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Sayed and Arram (2009), Najafi et al. (2009), Ouyang et al. (2010), Popovic and Popovic (2010), Wang et al. (2011), Sahoo et al. (2011), Falaghi et al. (2011), Martins and Borges (2011), Borges and Martins (2012a), Zou et al. (2012), Salehi and Haghifam (2012), Saberi et al. (2012), Mendoza et al. (2012), Cossi et al. (2012), Taroco et al. (2012), Millar et al. (2012), Junjie et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a, b), Ganguly et al. (2013a), Samper and Vargas (2013a, b), Camargo et al. (2013), Morillo et al. (2013), Saldarriaga et al. (2013), Zonkoly (2013), Aghaei et al. (2014), Ravadanegh and Roshanagh (2014), Muttaqi et al. (2014), Franco et al. (2014), Huang et al. (2014), Sadeghi and Kalantar (2014), Labbi and Attous (2014), Heidari et al. (2015), Murty and Kumar (2015), Delgado et al. (2015), Miloca et al. (2015), Bagheri et al. (2015), Gonçalves et al. (2015), Xing et al. (2015), Hemmati et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Mansor and Levi (2015), Saboori et al. (2015), Mazhari et al. (2016), Tabares et al. (2016), Xing et al. (2016), Ahmadigorji and Amjady (2016), Tarôco et al. (2016), Verma et al. (2016), Asija et al. (2017), Ahmadigorji et al. (2017), Shen et al. (2018), Hasanvand et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017)Gholami et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Mokryani et al. (2017), Hemmati et al. (2017), Bai et al. (2018), Raviprabakaran and Subramanian (2018), Li et al. (2018), Lin et al. (2019), Mehrjerdi (2019), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Bhadoria et al. (2019), Nahman and Peri (2020), Xie et al. (2020)	76	89.4
B	Non-technical	Mori and Yamada (2007), Cossi et al. (2009), Sayed and Arram (2009), Najafi et al. (2009), Al-Muhaini et al. (2010), Ouyang et al. (2010), Wang et al. (2011), Falaghi et al. (2011), Borges and Martins (2012b), Salehi and Haghifam (2012), Saberi et al. (2012), Taroco et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a, b), Samper and Vargas (2013a, b), Morillo et al. (2013), Zonkoly (2013), Aghaei et al. (2014), Ravadanegh and Roshanagh (2014), Muttaqi et al. (2014), Sadeghi and Kalantar (2014), Heidari et al. (2015), Miloca et al. (2015), Bagheri et al. (2015), Gonçalves et al. (2015), Xing et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Saboori et al. (2015), Tabares et al. (2016), Ahmadigorji and Amjady (2016), Tarôco et al. (2016), Verma et al. (2016), Shen et al. (2018), Hasanvand et al. (2017), Gholami et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Pombo et al. (2017), Mokryani et al. (2017), Bansal et al. (2017), Hemmati et al. (2017), Bai et al. (2018), Li et al. (2018), Lin et al. (2019), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Xie et al. (2020)	50	58.8
C	DG & DESS	Pudaruth and Li (2007), Haffner et al. (2008), Sayed and Arram (2009), Al-Muhaini et al. (2010), Falaghi et al. (2011), Borges and Martins (2012a, b), Zou et al. (2012), Saberi et al. (2012), Junjie et al. (2012), Gitizadeh et al. (2013), Morillo et al. (2013), Saldarriaga et al. (2013), Zonkoly (2013), Aghaei et al. (2014), Huang et al. (2014), Labbi and Attous (2014), Sadeghi and Kalantar (2014), Delgado et al. (2015), Bagheri et al. (2015), Xing et al. (2015), Hemmati et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Saboori et al. (2015), Tabares et al. (2016), Xing et al. (2016), Ahmadigorji and Amjady (2016), Ahmadigorji et al. (2017), Asija et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017), Shen et al. (2018), Hasanvand et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Mokryani et al. (2017), Hemmati et al. (2017), Bai et al. (2018), Raviprabakaran and Subramanian (2018), Li et al. (2018), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Bhadoria et al. (2019), Xie et al. (2020)	46	54.1
D	Time & money	Carrano et al. (2008), Haffner et al. (2008), Cossi et al. (2009), Najafi et al. (2009), Al-Muhaini et al. (2010), Cossi et al. (2012), Junjie et al. (2012), Zonkoly (2013), Sadeghi and Kalantar (2014), Delgado et al. (2015), Shivaie et al. (2015), Gholami et al. (2017), Humayd and Bhattacharya (2017), Pombo et al. (2017), Hemmati et al. (2017), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Xie et al. (2020)	19	22.4
E	Miscellaneous	Haffner et al. (2008), Borges and Martins (2012a), Cossi et al. (2012), Taroco et al. (2012), Camargo et al. (2013), Saldarriaga et al. (2013), Zonkoly (2013), Ravadanegh and Roshanagh (2014), Franco et al. (2014), Huang et al. (2014), Heidari et al. (2015), Gonçalves et al. (2015), Mansor and Levi (2015), Gholami et al. (2017), Pombo et al. (2017)	15	17.7

*x Number of research work reviewed in which the concern constraint has been considered

*x% Percentage of research work reviewed

(A9) Reliability constraint (Zonkoly 2013) has been considered in 3.5% of reviewed work and is expressed by (12) and (13), where (12) ensures that service reliability remains in upper and lower allowable values and power shortage rate reliability is expressed by (13).

$$CRS_z^{min} \leq CRS_z \leq CRS_z^{max} \tag{12}$$

$$CRS_z = 1 - \frac{T_{OH}}{T_P} \tag{13}$$

In addition to technical constraints discussed above, load demand considered by Borges and Martins (2012b),

Table 2 Summary of technical constraints of DNEP

S.	Technical Constraints	Reviewed articles	x^*	$x\,\%^{**}$
A1	Voltage limits	Mori and Yamada (2007), Nerves and Julian (2007), Carrano et al. (2007), Pudaruth and Li (2007), Carrano et al. (2008), Haffner et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Sayed and Arram (2009), Najafi et al. (2009), Ouyang et al. (2010), Popovic and Popovic (2010), Wang et al. (2011), Sahoo et al. (2011), Falaghi et al. (2011), Martins and Borges (2011), Borges and Martins (2012a), Zou et al. (2012), Salehi and Haghifam (2012), Saberi et al. (2012), Mendoza et al. (2012), Cossi et al. (2012), Taroco et al. (2012), Millar et al. (2012), Junjie et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a, b), Ganguly et al. (2013a), Camargo et al. (2013), Morillo et al. (2013), Saldarriaga et al. (2013), Zonkoly (2013), Aghaei et al. (2014), Ravadanegh and Roshanagh (2014), Muttaqi et al. (2014), Franco et al. (2014), Huang et al. (2014), Sadeghi and Kalantar (2014), Labbi and Attous (2014), Heidari et al. (2015), Murty and Kumar (2015), Delgado et al. (2015), Miloca et al. (2015), Bagheri et al. (2015), Gonçalves et al. (2015), Xing et al. (2015), Hemmati et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Mansor and Levi (2015), Saboori et al. (2015), Mazhari et al. (2016), Tabares et al. (2016), Xing et al. (2016), Ahmadigorji and Amjady (2016), Tarôco et al. (2016), Verma et al. (2016), Ahmadigorji et al. (2017), Asija et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017)Shen et al. (2018), Hasanvand et al. (2017), Gholami et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Mokryani et al. (2017), Lin et al. (2019), Mehrjerdi (2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Bhadoria et al. (2019), Nahman and Peri (2020), Xie et al. (2020)	75	88.2
A2	Power balance equations	Mori and Yamada (2007), Pudaruth and Li (2007), Carrano et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Sayed and Arram (2009), Ouyang et al. (2010), Popovic and Popovic (2010), Martins and Borges (2011), Borges and Martins (2012a), Taroco et al. (2012), Junjie et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a), Samper and Vargas (2013a, b), Camargo et al. (2013), Zonkoly (2013), Ravadanegh and Roshanagh (2014), Franco et al. (2014), Sadeghi and Kalantar (2014), Labbi and Attous (2014), Heidari et al. (2015), Miloca et al. (2015), Bagheri et al. (2015), Gonçalves et al. (2015), Xing et al. (2015), Ahmadigorji and Amjady (2015), Mansor and Levi (2015), Saboori et al. (2015), Mazhari et al. (2016), Tabares et al. (2016), Xing et al. (2016), Ahmadigorji and Amjady (2016), Verma et al. (2016), Ahmadigorji et al. (2017), Shen et al. (2018), Hasanvand et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Mokryani et al. (2017), Asija et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017), Bai et al. (2018), Raviprabakaran and Subramanian (2018), Mehrjerdi (2019), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Nahman and Peri (2020), Xie et al. (2020)	53	62.4
A3	Current limits	Mori and Yamada (2007), Nerves and Julian (2007), Carrano et al. (2007, 2008), Haffner et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Najafi et al. (2009), Ouyang et al. (2010), Popovic and Popovic (2010), Martins and Borges (2011), Zou et al. (2012), Salehi and Haghifam (2012), Taroco et al. (2012), Junjie et al. (2012), Ganguly et al. (2013a, b), Morillo et al. (2013), Saldarriaga et al. (2013), Ravadanegh and Roshanagh (2014), Huang et al. (2014), Delgado et al. (2015), Miloca et al. (2015), Gonçalves et al. (2015), Xing et al. (2015), Ahmadigorji and Amjady (2015), Mansor and Levi (2015), Saboori et al. (2015), Tabares et al. (2016), Xing et al. (2016), Tarôco et al. (2016), Verma et al. (2016), Ahmadigorji et al. (2017), Shen et al. (2018), Hasanvand et al. (2017), Mokryani et al. (2017), Bai et al. (2018), Lin et al. (2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Nahman and Peri (2020), Xie et al. (2020)	42	49.4
A4	Radiality	Mori and Yamada (2007), Nerves and Julian (2007), Carrano et al. (2007, 2008), Ouyang et al. (2010), Popovic and Popovic (2010), Falaghi et al. (2011), Martins and Borges (2011), Borges and Martins (2012b), Taroco et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a), Morillo et al. (2013), Aghaei et al. (2014), Ravadanegh and Roshanagh (2014), Franco et al. (2014), Huang et al. (2014), Heidari et al. (2015), Delgado et al. (2015), Miloca et al. (2015), Bagheri et al. (2015), Xing et al. (2015), Hemmati et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Mansor and Levi (2015), Mazhari et al. (2016), Tabares et al. (2016), Xing et al. (2016), Tarôco et al. (2016), Verma et al. (2016), Ahmadigorji et al. (2017), Shen et al. (2018), Lin et al. (2019), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Verma et al. (2019), Nahman and Peri (2020), Xie et al. (2020)	40	47.1
A5	Branch thermal limits	Pudaruth and Li (2007), Ouyang et al. (2010), Popovic and Popovic (2010), Wang et al. (2011), Mendoza et al. (2012), Millar et al. (2012), Huang et al. (2014), Murty and Kumar (2015), Bagheri et al. (2015), Xing et al. (2015), Mansor and Levi (2015), Mazhari et al. (2016), Xing et al. (2016), Verma et al. (2016), Hasanvand et al. (2017), Asija et al. (2017), Verma et al. (2019)	17	20.0
A6	Line capacity	Borges and Martins (2012a), Cossi et al. (2012), Samper and Vargas (2013a, b), Zonkoly (2013), Ravadanegh and Roshanagh (2014), Hemmati et al. (2015), Saboori et al. (2015), Mazhari et al. (2016), Shen et al. (2018), Saboori and Hemmati (2017), Bai et al. (2018), Li et al. (2018)	13	15.3
A7	Power factor	Zou et al. (2012), Sadeghi and Kalantar (2014), Delgado et al. (2015), Xing et al. (2015, 2016), Ahmadigorji and Amjady (2016), Verma et al. (2016), Ahmadigorji et al. (2017), Bai et al. (2018)	09	10.6

S.	Technical Constraints	Reviewed articles	x*	x %**
A8	Uni-directional flow	Franco et al. (2014), Delgado et al. (2015), Miloca et al. (2015), Pombo et al. (2017), Quevedo et al. (2019)	5	05.9
A9	Reliability	Huang et al. (2014), Humayd and Bhattacharya (2017), Nahman and Peri (2020)	3	03.5

*x Number of research work reviewed in which the concern constraint has been considered

*x% Percentage of research work reviewed

Zou et al. (2012), Cossi et al. (2012), Ganguly et al. (2013a), Xing et al. (2015), Mansor and Levi (2015), Xing et al. (2016), Verma et al. (2016), Li et al. (2018), short circuit ratio by Ouyang et al. (2010), Millar et al. (2012), Saboori et al. (2015), Hemmati et al. (2017), number of nodes by Morillo et al. (2013), Mansor and Levi (2015), Verma et al. (2019), phase angle taken by Mansor and Levi (2015), Saboori et al. (2015), Mehrjerdi (2019), Verma et al. (2019) etc. have also been considered during various DNEP studies as well.

4 Non-technical constraints

Non-technical constraints are also termed as utilization or environmental constraints in many reviewed articles. Mainly these are related to an optimal number of substation, electrical instruments, and power equipment for expansion planning such as number of transformers, feeders, new switches, buses, substations, standard capacitors, circuit breakers, and voltage regulator (Table 3). The non-technical constraint also includes the capacity of instruments such as that of the transformers, feeders, and substations. Other than quantity and capacity, non-technical constraints are related to voltage regulators.

*x: Number of research work reviewed in which the concern constraint has been considered; *x%: Percentage of research work reviewed (B1) Non-technical constraints related to the quantity of electrical apparatus are expressed by (14)–(19). And these were considered by 7.1% of articles reviewed. (14) express the number of substations (Gonçalves et al. 2015), which are to be set up for optimal DNEP.

$$N_{SS2} \in \{0, 1\} \forall i \in \psi_{N1}, \psi_{N2} \tag{14}$$

The standard capacitors units mounted in one node (Zonkoly 2013) and total installed in the system is limited by (15) and (16), respectively.

$$0 \le N_{CP2,i}^{max} \le \overline{N_{CP2,i}^{max}} C_i \forall i \in \psi_{N1} \tag{15}$$

$$\sum_{i \in \psi_{N1}} C_i \le \overline{N_{CP2}^{max}} C_i \quad \forall\, i \in \psi_{N1} \tag{16}$$

(17) restricts the use of only one circuit breaker at each node (Tabares et al. 2016),

$$\sum_{u \in \psi_{N1}} \beta_{i,s}^{cb} \le 1 \quad \forall\, i \in \psi_{N1} \tag{17}$$

while the maximum number of circuit breakers installed during DNEP is limited by (18).

$$\sum_{u \in \psi_s} \sum_{g \in \psi_{DGA}} \beta_{i,g,s}^{DG} \le 1 \quad \forall\, i \in \psi_{N1} \tag{18}$$

Parameter N_{max}^{VR} expresses the max. number of voltage regulators (VR), that can be installed during DNEP, is established as:

$$\sum_{ij \in \psi_{B1}} v_{VR,c1} \le N_{VR2}^{max} \tag{19}$$

(B2) Non-technical constraints related to electrical apparatus capacity limits are expressed by (20)–(23), comprises of 33% of total reviewed articles, where (20) express capacity of transformers constraint (Gitizadeh et al. 2013), (21) and (22) provides feeder capacity constraints (Sayed and Arram 2009) and (23) expresses substation capacity constraints (Mori and Yamada 2007).

$$0 \le S_{SS}^{Tran} \le S_{SS,max}^{Tran} i = 1, 2, 3, \ldots, N_{SS1} \tag{20}$$

$$S_{b,f,ss} \le S_{b,f,ss}^{max} \quad \forall\, ij \in feeder\, links \tag{21}$$

$$\left.\begin{aligned} N_{L1} &= N_{B1} - 1 \\ N_{recon} &\le N_{recon}^{max} \\ N_{SW2} &\le N_{SW2}^{max} \\ N_{DG2} &\le N_{DG2}^{max} \\ N_{L2} &\le N_{L2}^{max} \\ N_{LP2} &\le N_{LP2}^{max} \end{aligned}\right\} \tag{22}$$

Table 3 Summary of non-technical constraints of DNEP

S.	Non-technical constraints	Reviewed articles	x*	x %**
B1	Number of Substation and Power equipment	Camargo et al. (2013), Saldarriaga et al. (2013), Franco et al. (2014), Gonçalves et al. (2015), Tabares et al. (2016), Humayd and Bhattacharya (2017)	6	7.1
	Capacitor	Franco et al. (2014), Gonçalves et al. (2015), Humayd and Bhattacharya (2017)	3	3.3
	Circuit breakers	Gonçalves et al. (2015), Tabares et al. (2016)	2	2.3
	Substation	Camargo et al. (2013), Saldarriaga et al. (2013)	2	2.3
	Voltage regulator	Tabares et al. (2016)	1	1.2
B2	Capacity of Substation and Power equipment	Mori and Yamada (2007), Haffner et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Najafi et al. (2009), Sahoo et al. (2011), Falaghi et al. (2011), Saberi et al. (2012), Mendoza et al. (2012), Cossi et al. (2012), Gitizadeh et al. (2013), Ganguly et al. (2013a), Samper and Vargas (2013a, b), Camargo et al. (2013),Morillo et al. (2013), Saldarriaga et al. (2013), Heidari et al. (2015), Delgado et al. (2015), Miloca et al. (2015), Shivaie et al. (2015), Mansor and Levi (2015), Tabares et al. (2016), Xing et al. (2016), Humayd and Bhattacharya (2017), Mokryani et al. (2017), Bansal et al. (2017), Bai et al. (2018)	28	33.0
	Substation	Mori and Yamada (2007), Haffner et al. (2008), Mori and Yoshida (2009), Najafi et al. (2009), Sahoo et al. (2011), Falaghi et al. (2011), Cossi et al. (2012), Ganguly et al. (2013a), Samper and Vargas (2013a, b), Morillo et al. (2013), Saldarriaga et al. (2013), Heidari et al. (2015), Miloca et al. (2015), Shivaie et al. (2015), Mansor and Levi (2015), Humayd and Bhattacharya (2017), Mokryani et al. (2017), Bai et al. (2018)	19	22.4
	Transformers	Cossi et al. (2009), Saberi et al. (2012), Mendoza et al. (2012), Gitizadeh et al. (2013), Delgado et al. (2015), Tabares et al. (2016), Bansal et al. (2017), Verma et al. (2019)	8	9.4
	Feeders	Camargo et al. (2013)	1	1.2
B3	Voltage regulator	Tabares et al. (2016), Xing et al. (2016), Bansal et al. (2017)	3	3.5

$$0 \leq g^S_{l,t} \leq g^S_{max,l,t} + \sum_{L \in \Psi^S_l} \left[\sum_{\tau=1}^{t} x^{SL}_{l,\tau} \right] g^{SL}_{max,l} \right\}$$
$$\left\{ \forall l \in \Psi^S \right\} \tag{23}$$

(B3) With the assumption that tapping of VRs is an uninterrupted variable and at least one VR is assigned in each circuit, (24), (25) model the operation of a VR (Gonçalves et al. 2015) (comprises of 3.5% of total reviewed articles).

$$[1 - R\%]^2 \tilde{V}^{sqr}_{j,d} \leq V^{sqr}_{j,d} \leq [1 + R\%]^2 \tilde{V}^{sqr}_{j,d}, \\ \forall ij \in \Omega_l, \\ \forall d \in \Omega_d \right\} \tag{24}$$

$$\left| V^{sqr}_{j,d} - \tilde{V}^{sqr}_{j,d} \right| \leq [V^2_{max} - V^2_{min}] v_{ij}, \\ \forall ij \in \Omega_l, \\ \forall d \in \Omega_d \right\} \tag{25}$$

In addition to the non-technical constraints discussed above, number of feeders taken by Camargo et al. (2013), number of transformers by Heidari et al. (2015), Hemmati et al. (2015), Verma et al. (2016), number buses by Borges and Martins (2012a), new paths by Verma et al. (2016, 2019) normally open points & switches by Borges and Martins (2012a), Cossi et al. (2012), Pombo et al. (2017), Verma et al. (2019), etc. were considered by few researchers as well.

5 DG & DESS constraints

Various DG related constraints considered in DNEP are its capacity, penetration level, active power, and reactive power generated (Table 4), which are expressed mathematically and a few of which are considered as non-mathematical constraints viz. wind speed, solar radiation intensity, number and type of DG.*x: Number of research work reviewed in which the concern constraint has been considered; *x%: Percentage of research work reviewed

(C1) DG capacity constraints are expressed by (26) (Haffner et al. 2008), which comprises 36.5% of the total reviewed articles. **(C2)** & **(C3)** Constraints of active and reactive power generation are expressed by (27) and (28) respectively, which have been considered in 20% and 22.4% of total reviewed articles, respectively. **(C4)** DG penetration level constraint (Borges and Martins 2012b) (see Eq. (29)) is considered in 11.4% of the total reviewed articles.

$$0 \leq S_{DG,i} \leq S^{max}_{DG,i} \quad \forall n \in \psi^{DG}_{N2} \tag{26}$$

$$P_{G,k} - P_{L,k} - \sum_{j \in \psi_k} P_{ij,f,ss} = 0 \tag{27}$$

$$Q_{G,k} - Q_{L,k} - \sum_{j \in \psi_k} Q_{ij,f,ss} = 0 \tag{28}$$

$$S_{DG,i} \leq S^{max}_{DG,i} \tag{29}$$

Table 4 Summary of DG & DESS constraints of DNEP

S.	DG & DESS related constraints	Reviewed articles	x*	x %**
C1	DG capacity	Haffner et al. (2008), Sayed and Arram (2009), Al-Muhaini et al. (2010), Falaghi et al. (2011), Borges and Martins (2012b), Zou et al. (2012), Saberi et al. (2012), Junjie et al. (2012), Gitizadeh et al. (2013), Morillo et al. (2013), Saldarriaga et al. (2013), Zonkoly (2013), Aghaei et al. (2014), Huang et al. (2014), Sadeghi and Kalantar (2014), Delgado et al. (2015), Bagheri et al. (2015), Xing et al. (2015), Ahmadigorji and Amjady (2015), Shivaie et al. (2015), Tabares et al. (2016), Xing et al. (2016), Ahmadigorji and Amjady (2016), Ahmadigorji et al. (2017), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Hemmati et al. (2017), Bai et al. (2018), Li et al. (2018), Mehrjerdi(2019), Ugranli (2019)	31	36.5
C2	Active power generated	Pudaruth and Li (2007), Labbi and Attous (2014), Xing et al. (2015), Hemmati et al. (2015), Tabares et al. (2016), Xing et al. (2016), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Asija et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017), Mokryani et al. (2017), Raviprabakaran and Subramanian (2018), Quevedo et al. (2019), Ugranli (2019), Bhadoria et al. (2019), Xie et al. (2020)	17	20.0
C3	Reactive power generated	Pudaruth and Li (2007), Labbi and Attous (2014), Xing et al. (2015), Hemmati et al. (2015), Zou et al. (2012), Xing et al. (2015), Tabares et al. (2016), Xing et al. (2016), Humayd and Bhattacharya (2017), Saboori and Hemmati (2017), Asija et al. (2017), Gacem and Benattous (2017), Bansal et al. (2017), Mokryani et al. (2017), Raviprabakaran and Subramanian (2018), Quevedo et al. (2019), Ugranli (2019), Bhadoria et al. (2019), Xie et al. (2020),	19	22.4
C4	Penetration of DG	Sadeghi and Kalantar (2014), Delgado et al. (2015), Bagheri et al. (2015), Xing et al. (2016), Ahmadigorji and Amjady (2016), Ahmadigorji et al. (2017), Hasanvand et al. (2017), Humayd and Bhattacharya (2017), Li et al. (2018), Ugranli (2019)	10	11.8
C5	Number of DG	Saldarriaga et al. (2013), Huang et al. (2014), Xing et al. (2015), Tabares et al. (2016), Mehrjerdi(2019)	05	05.9
C6	DESS Stored energy limits	Shen et al. (2015), Saboori et al. (2015), Xing et al. (2016), Shen et al. (2018), Saboori and Hemmati (2017), Bai et al. (2018), Li et al. (2018), Quevedo et al. (2019), Xie et al. (2020)	09	10.6
C7	Initial value prerequisite of daily dispatch	Shen et al. (2015), Saboori et al. (2015), Xing et al. (2016), Shen et al. (2018), Saboori and Hemmati (2017), Bai et al. (2018), Li et al. (2018), Quevedo et al. (2019)	08	09.4
C8	Charging & discharging power limits	Shen et al. (2015), Saboori et al. (2015), Xing et al. (2016), Shen et al. (2018), Saboori and Hemmati (2017), Bai et al. (2018), Li et al. (2018), Xie et al. (2020)	08	09.4
C9	Outage time constraints	Shen et al. (2015), Xing et al. (2016), Shen et al. (2018), Saboori and Hemmati (2017)	4	04.7
C10	Capacity of DESS	Pombo et al. (2017), Quevedo et al. (2019), Xie et al. (2020)	3	03.5

(C5) Constraints applied on sitting of number of DG have been taken by 5.9% articles reviewed and are incorporated non-mathematically in the planning.

$$\forall k, l \in \psi_{N2}^{DESS}, S \in \psi_{CESS,s}, h, sc, t \tag{30}$$

$$E_{k,h+1,sc,t}^{S} = E_{k,h,sc,t}^{S} + \left[\eta_k^C C_{k,h,sc,t}^{S} - \frac{1}{\eta_k^D} D_{k,h,sc,t}^{S} \right] \tag{31}$$

$$\sum_{k \in ESSN} y_{k,t}^{DESS,S} E_{k,min}^{S} \leq E_{k,h,sc,t}^{S} \leq \sum_{k \in ESSN} y_{k,t}^{DESS,S} E_{k,max}^{S} \tag{32}$$

$$E_{k,h,sc,t}^{S} = E_{k,0}^{S} h = 1, 24 \tag{33}$$

$$D_{k,h,sc,t}^{S} \leq \sum_{k \in ESSN} y_{k,t}^{DESS,S} D_{k,max}^{S} \tag{34}$$

$$C_{k,h,sc,t}^{S} \leq \sum_{k \in ESSN} y_{k,t}^{DESS,S} C_{k,max}^{S} \tag{35}$$

$$\left. \begin{array}{l} T_{OH,n,t} = T_{OH,n,max,t} - \sigma \times IC_{DESS} \forall k \in \psi_{N2}^{DESS} \\ T_{OH,n,max,t} \forall k \notin \psi_{N2}^{DESS} \end{array} \right\} \tag{36}$$

$$IC_{DESS} = \sum_{sc} \sum_{h} \frac{E_{k,h,sc,t}}{d_{h,sc,t}} \tag{37}$$

$$IC_{DESS} \leq IC_{DESS}^{max} \tag{38}$$

The inclusion of distributed energy storing systems (DESS) is the latest trend in DNEP. The constraints imposed for optimal operation of DESS are power balance Eq. (31) based, which comprises of **(C6)** stored energy limits (Shen et al. 2015) represented by (32) (included in 10.6% of total DNEP studies reviewed); **(C7)** initial value required for daily dispatch for DESS represented by (33) (considered in 9.4% of total reviewed articles) and **(C8)** charging and discharging power limits of DESS represented by (34) and (35) respectively (considered in 9.4% of total reviewed articles). **(C9)** Outage time constraints are

Table 5 Summary of time & financial constraints of DNEP

S.	Time & financial constraints	Reviewed articles	x*	x %**
D1	Budget limitations	Carrano et al. (2008), Haffner et al. (2008), Cossi et al. (2009), Mori and Yoshida (2009), Najafi et al. (2009), Cossi et al. (2012), Delgado et al. (2015), Shivaie et al. (2015), Gholami et al. (2017), Humayd and Bhattacharya (2017), Pombo et al. (2017), Hemmati et al. (2017), Mehrjerdi(2019), Quevedo et al. (2019), Ugranli (2019), Xie et al. (2020)	16	18.8
D2	Purchased power	Junjie et al. (2012)	1	01.2
D3	Operational cost	Carrano et al. (2008)	1	01.2
D4	Minimum-up & -down time	Zonkoly (2013)	1	01.2

given by (36) wherein (37), IC_{DESS} signifies the capability of DESS to backing a load of each node when the outage happens (Shen et al. 2018). The impact of DESS capability to provide backup during an outage is measured by coefficient **σ**. **(C10)** DESS capacity is restricted by (38) where IC_{DESS}^{max} indicates installed DESS maximum capacity.

6 Time and financial constraints

When the most utilized objective function is found out to be minimizing the cost of expansion planning, then it is evident that cost minimization can be done by putting constraints on financial and economic aspects as well. Those are (Table 5): budget limitations, purchased power constraints and fuel cost.

$$C_{Inst,s} + C_{Opp,s} \leq C_{Bud,s} \forall s = 1, 2, \ldots, N_{St} \quad (39)$$

$$\left. \begin{array}{c} P_{Purch}^1 = S_{LD,s,t} + P_{loss,s,t} - \sum_{i=1}^{N_{LB1}} P_{DG,i,s,t} \\ \forall s \in N_{LS}, \forall t \in N_{PY} \end{array} \right\} \quad (40)$$

$$\left. \begin{array}{c} C_{Opp,s} = C_{Opp,s-1} + C_{Inst,s} \forall s = 1, 2, \ldots, N \\ x[0] + \sum_{s=1}^{N} C_{Inst,s} = x[1] \end{array} \right\} \quad (41)$$

(D1) Every stage expansion has budget allocation (Carrano et al. 2008), which is governed by (39). **(D2)** Purchased power constraints (Junjie et al. 2012) and **(D3)** fuel cost (Carrano et al. 2008) have been described with (40) and (41).

Keeping uncertainties in mind while doing the DNEP, time constraints also play an essential role (Table 5). **(D4)** (42) and (43) provide a limit for the minimum time period for which any unit must be on or off before it can be shut down or carried operational (Zonkoly 2013).

$$T_{on,u} \geq T_{MUT,u} \quad (42)$$

$$T_{off,u} \geq T_{MDT,u} \quad (43)$$

Besides up and downtime constraints, computational time, planning period (Gholami et al. 2017), meantime failure (Al-Muhaini et al. 2010), mean time to repair (Al-Muhaini et al. 2010) etc., are some other time related constraints which have been included in various research articles without any mathematical expressions.

Table 6 Miscellaneous constraints of DNEP

S.	Miscellaneous constraints	Reviewed articles	x*	x %**
E1	Conductor type	Cossi et al. (2012), Saldarriaga et al. (2013), Franco et al. (2014), Heidari et al. (2015), Gonçalves et al. (2015)	5	05.9
E2	Available area	Heidari et al. (2015)	2	02.4
E3	Spinning reserve	Zonkoly (2013)	1	01.2
E4	Ramp rate	Zonkoly (2013)	1	01.2
E5	Salvage of a line	Heidari et al. (2015)	1	01.2

7 Miscellaneous constraints

In addition to technical and non-technical constraints, there are various miscellaneous constraints found in several articles which bring more confinement of the obtained results. These are: types of cable/conductor, spinning reserve, ramp rate, salvage of a line, and available area constraints (Table 6).

(E1) While installation or replacement of any wire gage in a circuit (44) will ensure that only a single type of conductor is selected (Cossi et al. 2009). **(E2)** During DNEP, constraints on the substation area (Heidari et al. 2015) is given by (45), which also ensures that at least one transformer (term N_{TR2}) can be added to an upcoming substation.

(E3) Spinning reserve constraint (Zonkoly 2013) is expressed by (46). **(E4)** (47) and (48), which provides limits of up and down ramp rate (Zonkoly 2013) for a dedicated dispatchable unit. **(E5)** Finally, the last mathematically expressed DNEP constraint for salvage (Heidari et al. 2015) of an existing line is expressed by (49)–(51).

$$\sum_{k \in C} y_{b,sc} \leq 1, \quad \forall ij \in \psi_{B1} \tag{44}$$

$$\sum_{t=1}^{T} \sum_{ss=1}^{N_{SS1}} x_{j,t,ss} \leq N_{TR2} \quad \forall j \in \psi_{SS1}, \psi_{SS2} \tag{45}$$

$$\sum_{u=1}^{N_{G1}} S_{G,u}^{max} \geq P_{LD} + P_{loss,s,t} + P_{SR} \tag{46}$$

$$S_{G,u}(t) - S_{G,u}(t-1) \leq UR_u \tag{47}$$

$$S_{G,u}(t-1) - S_{G,u}(t) \leq DR_u \tag{48}$$

$$\left. \begin{array}{c} -1 \leq \sum_{\tau=1}^{t} x_{b,\tau,a} \leq 0 \\ \forall j \in \psi_{L1}, \\ \forall a = \text{Type}_L, \\ \forall t \in \{1, 2, \ldots, T\} \end{array} \right\} \tag{49}$$

$$\left. \begin{array}{c} 0 \leq \sum_{\tau=1}^{t} x_{L,\tau,a} \leq 1 \\ \forall L \in \psi_{L1}, \\ \forall a \in \{1, 2 \ldots \psi_{L1}\} \\ a \neq \text{Typ}_L, \\ \forall t \in \{1, 2 \ldots, T\} \end{array} \right\} \tag{50}$$

$$\left. \begin{array}{c} 0 \leq \sum_{\tau=1}^{t} x_{L,\tau,a} \leq 1 \\ \forall L \in \psi_{L2} \\ \forall a \in \{1, 2 \ldots \psi_{L1}\} \\ \forall t \in \{1, 2 \ldots, T\} \end{array} \right\} \tag{51}$$

In addition to above mathematical expressions of miscellaneous constraints, planning alternatives (Borges and Martins 2012b), geographical limitations (Ravadanegh and Roshanagh 2014), length of feeders (Wang et al. 2011; Mendoza et al. 2012; Camargo et al. 2013; Mansor and Levi 2015; Verma et al. 2016; Gholami et al. 2017), effective age and failure rate of feeders (Gholami et al. 2017), restrictions of choice of network sections (Pombo et al. 2017) etc. have also been included in few DNEP studies as well.

8 Conclusions

With continuous change in demographics and consumption patterns of energy, traditional distribution systems are getting converted into microgrids having more and more penetration of DG. A DNEP problem is no longer a single-objective one. It has now become a multi-objective, complex polynomial non-convex problem. The continuous changes and advancements in DNEP are also leading researchers towards providing greener solutions (by having more and more participation of renewable-based DG and DESS), which requires more dedicated checks and boundaries on technical and non-technical objective outputs. In present work, authors have summarized the DNEP constraints considered in nearly eighty selected articles published over a decade and categorized them into five heads viz. *technical constraints, non-technical constraints, DG and DESS constraints, time and money constraints* and *miscellaneous constraints*. Major conclusions which can be drawn from the review of the colossal literature available on DNEP constraints are:

- 89.4% of total reviewed articles considered technical constraints as the most imperative one in for DNEP problems, followed by non-technical constraints (58.8%). With the changing scenarios of distribution networks, few new constraints have been added to the DNEP library.
- The foremost aim and decision involved in DNEP carried out is to minimize the investment cost while complying with various technical and physical constraints.
- Constraints can't be classified merely as equality and non-equality type but have to be more diversified so that all-round check can be implemented for higher optimization degree.
- Very few appropriate conditions for the long-term accomplishment of a commercial microgrid have been talked about. New constraints have to be defined, as the conventional distribution system is transferring from passive to active network and are becoming renewable-based DG integrated microgrid. Enhanced conditions for a microgrid formation will be defined by addressing additional real-market constraints and uncertainties.

- E-vehicles are a new type of large loads with limited research work and which can significantly influence the stability and reliability of distribution systems. Complete appraisal of the impacts of e-vehicles on DNEP, considering them with appropriate constraints is prescribed.
- The governing parameters of metaheuristic algorithms should be regulated for optimization as they expressively affect their computational working. Conversely, in DNEP literature, practically no research work has put constraints to control metaheuristics parameters.
- Most of the prevailing research articles have not considered all types of constraints related to system uncertainty, such as in the case of non-dispatchable DG units and power charges. Considering these constraints related to uncertainties into account make available more accurate modeling of distribution systems leading to more apposite elucidations for real-world DNEP.
- Due to environmental and economic issues, constraints must also be designed with flexibility so that alternative options can be incorporated in DNEP. For example, in case the planer has to meet the fast load growth and construction of new substations is not feasible, then DG and capacitor bank units can also be taken as an alternative for the situation.
- Many energy storage elucidations based on DESS, fuel cells, renewable, etc. are offered, but they are diverse in terms of provisions and features, making it hard to handpick a sole technology for all the applications. It is suggested that appropriate constraints must be put to recognize the concrete level of infiltration of technologies as well as on the physical properties of elements for a safe environment.
- Most researchers have separated the DNEP issue into restricted parts alongside forced constraints and attempted to recommend their specific solutions. Some examination works have not considered all related constraints for DNEP, which should be considered simultaneously. Overlooking such limitations prompts the aftereffects of no practical value and noteworthiness.
- Each stage of DNEP viz—modeling, calculation, multi-scenario simulation, all-inclusive scheduling, and rectification—contains diverse blends of components, elements, strategy, & optimization techniques that influence the ultimate planning conclusion. So the constraints for above mention stages must be precisely defined separately.

Not many research articles are presently available on the solution of DNEP in which all the constraints (to be applicable) revealed in this article are fully and adequately considered. The above discussions and conclusions show the swing of modern DNEP towards more cohesive, multi-objective, and DESS and fuel cells assimilated ones with control functionalities, in contrast to the traditional DNEP prototypes. This article presents a thorough review of colossal literature on DNEP constraints. This work is initiated and purposed with a conclusion to expand information base on DNEP and to pinpointing any breaks in the reviewed literature that can be addressed in impending research for further enhancement. It is foreseen that this survey will be valuable to the engineers, planners, and researchers required with DNEP for growing new methodologies.

References

Aghaei J, Muttaqi KM, Azizivahed A, Gitizadeh M (2014) Distribution expansion planning considering reliability and security of energy using modified PSO (particle swarm optimization) algorithm. Energy 65:398–411. https://doi.org/10.1016/j.energy.2013.10.082

Ahmadigorji M, Amjady N (2015) Optimal dynamic expansion planning of distribution systems considering non-renewable distributed generation using a new heuristic double-stage optimization solution approach. Appl Energy 156:655–665. https://doi.org/10.1016/j.apenergy.2015.07.042

Ahmadigorji M, Amjady N (2016) A multiyear DG-incorporated framework for expansion planning of distribution networks using binary chaotic shark smell optimization algorithm. Energy 102:199–215. https://doi.org/10.1016/j.energy.2016.02.088

Ahmadigorji M, Amjady N, Dehghan S (2017) A novel two-stage evolutionary optimization method for multiyear expansion planning of distribution systems in presence of distributed generation. Appl Soft Comput 52:1098–1115. https://doi.org/10.1016/j.asoc.2016.09.020

Al-Muhaini M, Heydt GT, Huynh A (2010) The reliability of power distribution systems as calculated using system theoretic concepts. In: IEEE PES general meeting. https://doi.org/10.1109/pes.2010.5588116

Asija D, Soni KM, Sinha SK et al (2017) Multi-objective optimization and network security enhancement for congestion management in wholesale electricity market. Int J Syst Assur Eng Manag 8:1775–1782. https://doi.org/10.1007/s13198-017-0668-7

Bagheri A, Monsef H, Lesani H (2015) Integrated distribution network expansion planning incorporating distributed generation considering uncertainties, reliability, and operational conditions. Int J Electr Power Energy Syst 73:56–70. https://doi.org/10.1016/j.ijepes.2015.03.010

Bai L, Jiang T, Li F, Chen H, Li X (2018) Distributed energy storage planning in soft open point based active distribution networks incorporating network reconfiguration and DG reactive power capability. Appl Energy 210:1082–1091. https://doi.org/10.1016/j.apenergy.2017.07.004

Bansal JC, Jadon SS, Tiwari R et al (2017) Optimal power flow using artificial bee colony algorithm with global and local neighborhoods. Int J Syst Assur Eng Manag 8:2158–2169. https://doi.org/10.1007/s13198-014-0321-7

Bhadoria VS, Pal NS, Shrivastava V (2019) Artificial immune system based approach for size and location optimization of distributed

generation in distribution system. Int J Syst Assur Eng Manag 10:339. https://doi.org/10.1007/s13198-019-00779-9

Borges C, Martins V (2012a) Active distribution network integrated planning incorporating distributed generation and load response uncertainties. In: IEEE power and energy society general meeting. https://doi.org/10.1109/pesgm.2012.6343936

Borges C, Martins V (2012b) Multistage expansion planning for active distribution networks under demand and Distributed Generation uncertainties. Int J Electr Power Energy Syst 36(1):107–116. https://doi.org/10.1016/j.ijepes.2011.10.031

Camargo V, Lavorato M, Romero R (2013) Specialized genetic algorithm to solve the electrical distribution system expansion planning. In: IEEE power & energy society general meeting. https://doi.org/10.1109/pesmg.2013.6672615

Carrano EG, Guimaraes FG, Takahashi RH, Neto OM, Campelo F (2007) Electric distribution network expansion under load-evolution uncertainty using an immune system inspired algorithm. IEEE Trans Power Syst 22(2):851–861. https://doi.org/10.1109/tpwrs.2007.894847

Carrano EG, Cardoso R, Takahashi R, Fonseca C, Neto O (2008) Power distribution network expansion scheduling using dynamic programming genetic algorithm. IET Gener Transm Distrib 2(3):444. https://doi.org/10.1049/iet-gtd:20070174

Cossi A, Romero R, Mantovani J (2009) Planning and projects of secondary electric power distribution systems. IEEE Trans Power Syst 24(3):1599–1608. https://doi.org/10.1109/tpwrs.2009.2021208

Cossi A, Silva LD, Lázaro R, Mantovani J (2012) Primary power distribution systems planning taking into account reliability, operation and expansion costs. IET Gener Transm Distrib 6(3):274. https://doi.org/10.1049/iet-gtd.2010.0666

Delgado GM, Contreras J, Arroyo JM (2015) Joint expansion planning of distributed generation and distribution networks. In: IEEE power & energy society general meeting. https://doi.org/10.1109/pesgm.2015.7285602

Falaghi H, Singh C, Haghifam M, Ramezani M (2011) DG integrated multistage distribution system expansion planning. Int J Electr Power Energy Syst 33(8):1489–1497. https://doi.org/10.1016/j.ijepes.2011.06.031

Franco JF, Rider MJ, Romero R (2014) A mixed-integer quadratically-constrained programming model for the distribution system expansion planning. Int J Electr Power Energy Syst 62:265–272. https://doi.org/10.1016/j.ijepes.2014.04.048

Gacem A, Benattous D (2017) Hybrid GA-PSO for optimal placement of static VAR compensators in power system. Int J Syst Assur Eng Manag 8:247–254. https://doi.org/10.1007/s13198-015-0347-5

Ganguly S, Sahoo NC, Das D (2013a) Multi-objective particle swarm optimization based on fuzzy-Pareto-dominance for possibilistic planning of electrical distribution systems incorporating distributed generation. Fuzzy Sets Syst 213:47–73. https://doi.org/10.1016/j.fss.2012.07.005

Ganguly S, Sahoo NC, Das D (2013b) Multi-objective planning of electrical distribution systems using dynamic programming. Int J Electr Power Energy Syst 46:65–78. https://doi.org/10.1016/j.ijepes.2012.10.030

Gholami M, Moeini-Aghtaie M, Safdarian A, Fard AA (2017) Maintenance based distribution network expansion planning. In: Iranian conference on electrical engineering (ICEE). https://doi.org/10.1109/iraniancee.2017.7985185

Gitizadeh M, Vahed AA, Aghaei J (2013) Multistage distribution system expansion planning considering distributed generation using hybrid evolutionary algorithms. Appl Energy 101:655–666. https://doi.org/10.1016/j.apenergy.2012.07.010

Gonçalves RR, Franco JF, Rider MJ (2015) Short-term expansion planning of radial electrical distribution systems using mixed-integer linear programming. IET Gener Transm Distrib 9(3):256–266. https://doi.org/10.1049/iet-gtd.2014.0231

Haffner S, Pereira L, Pereira L, Barreto L (2008) Multistage model for distribution expansion planning with distributed generation—part I: problem formulation. IEEE Trans Power Deliv 23(2):915–923. https://doi.org/10.1109/tpwrd.2008.917916

Hasanvand S, Nayeripour M, Waffenschmidt E, Fallahzadeh-Abarghouei H (2017) A new approach to transform an existing distribution network into a set of micro-grids for enhancing reliability and sustainability. Appl Soft Comput 52:120–134. https://doi.org/10.1016/j.asoc.2016.12.013

Heidari S, Fotuhi-Firuzabad M, Kazemi S (2015) Power distribution network expansion planning considering distribution automation. IEEE Trans Power Syst 30(3):1261–1269. https://doi.org/10.1109/tpwrs.2014.2339301

Hemmati R, Hooshmand R, Taheri N (2015) Distribution network expansion planning and DG placement in the presence of uncertainties. Int J Electr Power Energy Syst 73:665–673. https://doi.org/10.1016/j.ijepes.2015.05.024

Hemmati R, Saboori H, Siano P (2017) Coordinated short-term scheduling and long-term expansion planning in microgrids incorporating renewable energy resources and energy storage systems. Energy 134:699–708. https://doi.org/10.1016/j.energy.2017.06.081

Huang Y, Alvehag K, Soder L (2014) Distribution network expansion planning considering distributed generation using probabilistic voltage constraints. In: International conference on probabilistic methods applied to power systems (PMAPS). https://doi.org/10.1109/pmaps.2014.6960581

Humayd AS, Bhattacharya K (2017) Distribution system planning to accommodate distributed energy resources and PEVs. Electr Power Syst Res 145:1–11. https://doi.org/10.1016/j.epsr.2016.12.016

Junjie M, Yulong W, Yang L (2012) Size and location of distributed generation in distribution system based on immune algorithm. Syst Eng Procedia 4:124–132. https://doi.org/10.1016/j.sepro.2011.11.057

Labbi Y, Attous DB (2014) A hybrid particle swarm optimization and pattern search method to solve the economic load dispatch problem. Int J Syst Assur Eng Manag 5:435. https://doi.org/10.1007/s13198-013-0186-1

Li Y, Feng B, Li G, Qi J, Zhao D, Mu Y (2018) Optimal distributed generation planning in active distribution networks considering integration of energy storage. Appl Energy 210:1073–1081. https://doi.org/10.1016/j.apenergy.2017.08.008

Lin Z, Hu Z, Song Y (2019) Distribution network expansion planning considering $N-1$ criterion. Trans Power Syst 34(3):2476–2478. https://doi.org/10.1109/TPWRS.2019.2896841

Mansor NN, Levi V (2015) Expansion planning of medium voltage distribution networks. In: 50th International universities power engineering conference (UPEC). https://doi.org/10.1109/upec.2015.7339776

Martins VF, Borges CLT (2011) Active distribution network integrated planning incorporating distributed generation and load response uncertainties. IEEE Trans Power Syst 26(4):2164–2172. https://doi.org/10.1109/TPWRS.2011.2122347

Mazhari SM, Monsef H, Romero R (2016) A multi-objective distribution system expansion planning incorporating customer choices on reliability. IEEE Trans Power Syst 31(2):1330–1340. https://doi.org/10.1109/tpwrs.2015.2430278

Mehrjerdi H (2019) Simultaneous load leveling and voltage profile improvement in distribution networks by optimal battery storage planning. Energy 181:916–926. https://doi.org/10.1016/j.energy.2019.06.021

Mendoza JE, López ME, Peña HE, Labra DA (2012) Low voltage distribution optimization: site, quantity and size of distribution transformers. Electr Power Syst Res 91:52–60. https://doi.org/10.1016/j.epsr.2012.05.004

Millar RJ, Kazemi S, Lehtonen M, Saarijarvi E (2012) Impact of MV connected microgrids on MV distribution planning. IEEE Trans Smart Grid 3(4):2100–2108. https://doi.org/10.1109/tsg.2012.2212922

Miloca S, Volpi N, Yuan J, Pinto C (2015) Expansion planning problem in distribution systems with reliability evaluation: an application in real network using georeferenced database. Int J Electr Power Energy Syst 70:9–16. https://doi.org/10.1016/j.ijepes.2015.01.004

Mokryani G, Hu YF, Papadopoulos P, Niknam T, Aghaei J (2017) Deterministic approach for active distribution networks planning with high penetration of wind and solar power. Renew Energy 113:942–951. https://doi.org/10.1016/j.renene.2017.06.074

Mori H, Yamada Y (2007) An efficient multi-objective meta-heuristic method for distribution network expansion planning. In: IEEE Lausanne power tech. https://doi.org/10.1109/pct.2007.4538346

Mori H, Yoshida T (2009) An efficient multi-objective memetic algorithm for uncertainties in distribution network expansion planning. In: IEEE power & energy society general meeting. https://doi.org/10.1109/pes.2009.5275469

Morillo JL, Perez JF, Cadena AI (2013) Dynamic multi-objective planning for distribution systems with distributed generation. In: IEEE PES ISGT Europe 2013. https://doi.org/10.1109/isgteurope.2013.6695399

Murty V, Kumar A (2015) Optimal placement of DG in radial distribution systems based on new voltage stability index under load growth. Int J Electr Power Energy Syst 69:246–256. https://doi.org/10.1016/j.ijepes.2014.12.080

Muttaqi KM, Le AD, Negnevitsky M, Ledwich G (2014) An algebraic approach for determination of DG parameters to support voltage profiles in radial distribution networks. IEEE Trans Smart Grid 5(3):1351–1360. https://doi.org/10.1109/tsg.2014.2303194

Nahman JM, Peri DM (2020) Radial distribution network planning under uncertainty by applying different reliability cost models. Electr Power Energy Syst 117:105655. https://doi.org/10.1016/j.ijepes.2019.105655

Najafi S, Hosseinian S, Abedi M, Vahidnia A, Abachezadeh S (2009) A framework for optimal planning in large distribution networks. IEEE Trans Power Syst 24(2):1019–1028. https://doi.org/10.1109/tpwrs.2009.2016052

Nerves AC, Julian GB (2007) Optimal feeder configuration in expansion planning using simulated annealing. In: TENCON 2007 - IEEE region 10 conference. https://doi.org/10.1109/tencon.2007.4428981

Ouyang W, Cheng H, Zhang X, Yao L, Bazargan M (2010) Distribution network planning considering distributed generation by genetic algorithm combined with graph theory. Electr Power Compon Syst 38(3):325–339. https://doi.org/10.1080/15325000903273429

Pombo AV, Murta-Pina J, Pires VF (2017) Multiobjective formulation of the integration of storage systems within distribution networks for improving reliability. Electr Power Syst Res 148:87–96. https://doi.org/10.1016/j.epsr.2017.03.012

Popovic Z, Popovic D (2010) Graph theory based formulation of multi-period distribution expansion problems. Electr Power Syst Res 80(10):1256–1266. https://doi.org/10.1016/j.epsr.2010.04.009

Pudaruth GR, Li F (2007) Costs and benefits assessment considering deferral of assets expansion in distribution systems. In: 42nd International universities power engineering conference. https://doi.org/10.1109/upec.2007.4469064

Quevedo PMD, Muñoz-Delgado G, Contreras J (2019) Impact of electric vehicles on the expansion planning of distribution systems considering renewable energy, storage and charging stations. Trans Smart Grid 10(1):794–804. https://doi.org/10.1109/TSG.2017.2752303

Ravadanegh SN, Roshanagh RG (2014) On optimal multistage electric power distribution networks expansion planning. Int J Electr Power Energy Syst 54:487–497. https://doi.org/10.1016/j.ijepes.2013.07.008

Raviprabakaran V, Subramanian RC (2018) Enhanced ant colony optimization to solve the optimal power flow with ecological emission. Int J Syst Assur Eng Manag 9:58–65. https://doi.org/10.1007/s13198-016-0471-x

Saberi R, Falaghi H, Ramezani M, Khosravi M, Vahidi T, Dadgar M (2012) Optimal planning of medium voltage distribution networks in the presence of wind power generations. In: CIRED workshop: integration of renewables into the distribution grid. https://doi.org/10.1049/cp.2012.0735

Saboori H, Hemmati R (2017) Maximizing DISCO profit in active distribution networks by optimal planning of energy storage systems and distributed generators. Renew Sustain Energy Rev 71:365–372. https://doi.org/10.1016/j.rser.2016.12.066

Saboori H, Hemmati R, Jirdehi MA (2015) Reliability improvement in radial electrical distribution network by optimal planning of energy storage systems. Energy 93:2299–2312. https://doi.org/10.1016/j.energy.2015.10.125

Sadeghi M, Kalantar M (2014) Multi types DG expansion dynamic planning in distribution system under stochastic conditions using covariance matrix adaptation evolutionary strategy and monte-carlo simulation. Energy Convers Manag 87:455–471. https://doi.org/10.1016/j.enconman.2014.07.010

Sahoo N, Ganguly S, Das D (2011) Simple heuristics-based selection of guides for multi-objective PSO with an application to electrical distribution system planning. Eng Appl Artif Intell 24(4):567–585. https://doi.org/10.1016/j.engappai.2011.02.007

Saldarriaga CA, Hincapie RA, Salazar H (2013) A holistic approach for planning natural gas and electricity distribution networks. IEEE Trans Power Syst 28(4):4052–4063. https://doi.org/10.1109/tpwrs.2013.2268859

Salehi J, Haghifam M (2012) Long term distribution network planning considering urbanity uncertainties. Int J Electr Power Energy Syst 42(1):321–333. https://doi.org/10.1016/j.ijepes.2012.04.005

Samper ME, Vargas A (2013a) Investment decisions in distribution networks under uncertainty with distributed generation—part I: model formulation. IEEE Trans Power Syst 28(3):2331–2340. https://doi.org/10.1109/tpwrs.2013.2239666

Samper ME, Vargas A (2013b) Investment decisions in distribution networks under uncertainty with distributed generation—part II: implementation and results. IEEE Trans Power Syst 28(3):2341–2351. https://doi.org/10.1109/tpwrs.2013.2239667

Sayed MAE, Arram AA (2009) Dispersed generation impact on distribution network expansion planning. In: Power systems conference. https://doi.org/10.1109/psamp.2009.5262329

Shen X, Zhu S, Zheng J, Han Y, Li Q, Nong J, Shahidehpour M (2015) Active distribution network expansion planning integrated with centralized and distributed energy storage system. In: IEEE power & energy society general meeting. https://doi.org/10.1109/pesgm.2015.7286069

Shen X, Shahidehpour M, Han Y, Zhu S, Zheng J (2018) Expansion planning of active distribution networks with centralized and distributed energy storage systems. In: IEEE power & energy society general meeting (PESGM). https://doi.org/10.1109/pesgm.2018.8585964

Shivaie M, Ameli MT, Sepasian MS, Weinsier PD, Vahidinasab V (2015) A multistage framework for reliability-based distribution

expansion planning considering distributed generations by a self-adaptive global-based harmony search algorithm. Reliab Eng Syst Saf 139:68–81. https://doi.org/10.1016/j.ress.2015.03.001

Tabares A, Franco JF, Lavorato M, Rider MJ (2016) Multistage long-term expansion planning of electrical distribution systems considering multiple alternatives. IEEE Trans Power Syst 31(3):1900–1914. https://doi.org/10.1109/tpwrs.2015.2448942

Taroco CG, Carrano EG, Takahashi RH, Neto OM (2012) A faster genetic algorithm for substation location and network design of power distribution systems. In: IEEE congress on evolutionary computation. https://doi.org/10.1109/cec.2012.6256532

Tarôco CG, Takahashi RH, Carrano EG (2016) Multiobjective planning of power distribution networks with facility location for distributed generation. Electr Power Syst Res 141:562–571. https://doi.org/10.1016/j.epsr.2016.08.020

Ugranli F (2019) Analysis of renewable generation's integration using multi-objective fashion for multistage distribution network expansion planning. Electr Power Energy Syst 106:301–310. https://doi.org/10.1016/j.ijepes.2018.10.002

Verma MK, Mukherjee V, Yadav VK (2016) Greenfield distribution network expansion strategy with hierarchical GA and MCDEA under uncertainty. Int J Electr Power Energy Syst 79:245–252. https://doi.org/10.1016/j.ijepes.2016.01.004

Verma MK, Yadav VK, Mukherjee V, Ghosh S (2019) A multi-criteria approach for distribution network expansion through pooled MCDEA and Shannon entropy. Int J Emerg Electr Power Syst. https://doi.org/10.1515/ijeeps-2019-0043

Wang D, Ochoa L, Harrison G (2011) Modified GA and data envelopment analysis for multistage distribution network expansion planning under uncertainty. In: IEEE power and energy society general meeting. https://doi.org/10.1109/pes.2011.6039008

Xie S, Hu Z, Yang L, Wang J (2020) Expansion planning of active distribution system considering multiple active network managements and the optimal load-shedding direction. Electr Power Energy Syst 115:105451. https://doi.org/10.1016/j.ijepes.2019.105451

Xing H, Cheng H, Zhang L, Zhang S, Zhang Y (2015) Second-order cone model for active distribution network expansion planning. In: IEEE power & energy society general meeting. https://doi.org/10.1109/pesgm.2015.7286204

Xing H, Cheng H, Zeng P, Zhang Y (2016) Active distribution network expansion planning integrating dispersed energy storage systems. IET Gener Transm Distrib 10(3):638–644. https://doi.org/10.1049/iet-gtd.2015.0411

Zonkoly AME (2013) Multistage expansion planning for distribution networks including unit commitment. IET Gener Transm Distrib 7(7):766–778. https://doi.org/10.1049/iet-gtd.2012.0289

Zou K, Agalgaonkar AP, Muttaqi KM, Perera S (2012) Distribution system planning with incorporating DG reactive capability and system uncertainties. IEEE Trans Sustain Energy 3(1):112–123. https://doi.org/10.1109/tste.2011.2166281

ORIGINAL ARTICLE

Ectopic overexpression of cytosolic ascorbate peroxidase gene (*Apx1*) improves salinity stress tolerance in *Brassica juncea* by strengthening antioxidative defense mechanism

Saurabh C. Saxena[1] · Prafull Salvi[2] · Nitin Uttam Kamble[2] · Pankaj K. Joshi[3] · Manoj Majee[2] · Sandeep Arora[4]

Abstract

Salinity stress is considered to be a key constrain that reduces the crop productivity by impairing plant growth and development. During salt stress condition, an underlying mechanism for reduction in crop yield is increase in ROS level that can potentially harm cellular macromolecules, leading to disruption of essential physiological and biochemical processes. Plants possess a complex antioxidative defense machinery for scavenging these ROS. Ascorbate peroxidase (APX, E.C. 1.11.1.11), is a crucial antioxidant enzyme involved in Ascorbate–Glutathione pathway that primarily detoxifies the negative impact of H_2O_2 in cell. The efficient scavenging of H_2O_2 is a prerequisite for enhanced tolerance to salinity stress. Here, we have inspected whether over-expression of APX could provide protection against salinity stress. Cytosolic ascorbate peroxidase (*Apx1*) gene, isolated from *Arabidopsis thaliana*, was chosen as the candidate gene for strengthening the antioxidative defense system of *Brassica juncea*. Physiological parameters were employed to analyze the growth status of transgenic plants. Leaf disc assay was done to evaluate the salinity stress tolerance potential of transgenic plants, using several physiological and biochemical parameters. Under salinity stress, the transgenic plants performed well as compared to their non-transgenic counterparts; as revealed through greater proline accumulation, increased chlorophyll stability index, lower chlorophyll a/b ratio, and higher antioxidative enzyme activities. Further, the lower H_2O_2 levels were well correlated with lesser membrane damage as measured through MDA content. Collectively, our results clearly depicted that ectopic overexpression of *AtApx1* gene was able to confer salinity stress tolerance by strengthening the antioxidative defense system in *B. juncea*.

Keywords Salinity stress · *Brassica juncea* · Reactive oxygen species · Transgenic plants · Antioxidant system

✉ Saurabh C. Saxena
  saurabhbiochem@gmail.com

[1] Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India

[2] National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110067, India

[3] Sigma Aldrich Pvt. Ltd., Plot No. 12, Bommasandra-Jigani Link Road, Bengaluru 560 100, India

[4] Department of Molecular Biology and Genetic Engineering, CBSH, G. B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand 263145, India

Introduction

Plant's encounter to salt stress adversely affects its growth and productivity worldwide (Wang et al. 2003). Among several abiotic stresses, salinity stress is not only an environmental restrain that reduces the yield potential of a crop plant, but it also imposes a huge negative impact on growth and survival of crop plant. Salinity stress covers a wide range of alteration at physiological, biological and molecular

levels. It disturbs the cellular homeostasis by disrupting the ionic equilibrium and consequently induces the physiological drought as well as other secondary stresses (Apel and Hirt 2004). In response to these secondary stresses, plants induce the accumulation of osmolytes, anti-stress proteins and trigger the scavenging of reactive oxygen species (ROS) to overcome the devastating milieu. Primarily, accelerating the ROS detoxification system is one of the preferred approaches to ameliorate such adverse effects of salinity, albeit other approaches are also significantly important. The ROS detoxification is sustained by enzymatic and non-enzymatic antioxidant systems (Gill and Tuteja 2010; Miller et al. 2010). Plants have developed several antioxidant enzymes such as catalase (CAT), superoxide dismutase (SOD), ascorbate peroxidase (APX) and peroxidase (POD) that effectively restrict the ROS and hence protect the cell from oxidative injury imposed by ROS accumulation (Mittler and Zilinskas 1992; Noctor and Foyer 1998).

APX enzyme catalyzes the reaction for converting the H_2O_2 into H_2O and O_2, the APX enzymes are further categorized as cAPX (cytosolic), mitAPX (mitochondrial), sAPX (chloroplastic stroma), and tAPX (chloroplast thylakoids) on the basis of their subcellular localization (Shigeoka et al. 2002). Among the different APX isoforms, the cytosolic form is studied extensively and reported to be the most responsive form to environmental constrain (Davletova et al. 2005; Fourcroy et al. 2004; Mittler and Zilinskas 1994). cAPXs are mostly encoded by more than one gene and characterized from several plants including *Arabidopsis thaliana, Oryza sativa, Pisum sativum* and *Lycopersicum esculentum* to name a few (Gadea et al. 1999; Kim and Chung 1998; Kubo et al. 1992; Mittler and Zilinskas 1992). APX activity is known to be enhanced during various environmental stresses, including dehydration (Tanaka et al. 1990), high temperature (Sato et al. 2001), deficiency in microelements and excessive light (Cakmak and Marschner 1992), UV stress (Saxena et al. 2011), and salinity (Hernandez et al. 2000). Cytosolic APX1 provides cross-compartment protection to chloroplastic APX under high light intensity (Davletova et al. 2005). Rice showed the increase in transcript of *cApx* gene by heat stress; interestingly seedlings were having prior exposure to heat stress displayed improved tolerance to cold stress (Sato et al. 2001). Arabidopsis possesses eight *Apx* isoforms including two cytosolic isoforms (*AtApx1* and *AtApx2*) which are reported to be differentially induced under heat and high light stresses (Davletova et al. 2005). *AtApx1* promoter possesses ABA response element (ABRE), heat-shock element (HSE) while *AtApx2* promoter has two HSEs. The presence of different *cis*-regulatory elements *AtApx1* and *AtApx2* is attributed to their differential gene expression under stress (Mittler and Zilinskas 1992).

Engineering crop plants with improved antioxidative enzyme machinery pave the way for the better abiotic stress potential. Here we hypothesize that enhancing the activity of APX enzyme through genetic engineering will strengthen the antioxidative defense system and will enable it to limit the increased production of hydrogen peroxide under salinity stress.

For the present investigations, we have chosen *Brassica juncea* as the target crop, an important oilseed cash crop globally. Brassica is affected by various stresses which translate to the significant loss of quantity and quality produced. *B. juncea* and Arabidopsis belongs to the same family with high similarity in functional genes present between them allows us to study these functionally valuable genes (Anjum et al. 2012). Here, we ectopically expressed *AtApx1* in *B. juncea* and the growth of transgenic plants was analyzed for several physiological parameters under controlled conditions. We have compared the physiological and biochemical markers in transgenic and wild type (WT) plants under salt stress treatment. Transgenic plants were further examined for the ROS scavenging potential by determining the antioxidative enzyme activity of APX, POD, GPX and CAT. Taken together, our data indicated that the potential of *AtApx1* gene, improves the salinity stress tolerance by limiting the ROS-induced damage in *B. juncea*. These findings present a way forward to engineer the salt tolerance potential in agriculturally important crops to minimize the yield loss by salinity stresses.

Materials and methods

Plant materials and growth conditions

Seeds for *B. juncea* variety pusajaikisan were obtained from National Research Centre for Rapeseed and Mustard (NRCRM), Bharatpur (Rajasthan), India and were used in the present study. Brassica were grown in a growth chamber at 22 °C for 16 h light and 15 °C for 8 h dark cycle with 70% relative humidity under white fluorescent light.

Plant transformation

Agrobacterium tumefaciens strain C58C1rifR (pGV2260) harboring *AtApx1* (TAIR: AT1G07890) was used for plant transformation. Construct was generated by cloning the *Arabidopsis* cytosolic ascorbate peroxidase cDNA (*AtApx1*) with EcoRV and XhoI sites (Primers are listed in Table S3) under the control of a constitutive promoter (CaMv35S) and was transferred to *B. juncea* by *Agrobacterium*-mediated transformation. For plant transformation, hypocotyls from 6 day old seedlings were co cultivated with *Agrobacterium* culture for 24 h and further washed and selected from MS medium supplemented with kanamycin (30 mgL^{-1}). Selected explants were regenerated on shoot regenerating

medium (MS medium supplemented with combination of growth hormones such as 2.5 mgL^{-1} N^6-benzylaminopurine and 1.0 mgL^{-1} 1-naphthaleneacetic acid). Well-developed shoots were shifted to root regeneration medium (MS medium supplemented with 0.5 mgL^{-1} indole butyric acid). Plantlets were grown at 22 °C under white fluorescent light and 16-h photoperiod. Completely developed plants were hardened and were grown to maturity. The progeny of the primary transformants were multiplied by selfing and T_2 generation seeds were used for further experiments.

Confirmation of transgenic plants

Genomic DNA (gDNA) was isolated from one-month old transgenic plants along with WT plants using modified CTAB method (Augustine et al. 2013). For PCR amplification, genomic DNA was used as the template and amplified using gene specific primers and *nptII* primers (Primers are listed in Table S3). For southern blot, 10 μg of gDNA was digested with EcoRV and the samples were electrophoresed on agarose gel. After electrophoresis, the gel was treated with depurination, denaturation and neutralization solutions followed by hybridization. After hybridization and washing, the blot was exposed to X-ray film and the film was developed. We have used three independent transgenic lines for further experiments which are designated in figure and/or graphs as follows: L1 represents *AtApx1* transgenic line 3; L2 represents *AtApx1* transgenic line 4; L3 represents *AtApx1* transgenic line 5.

Quantitative real-time PCR

For isolation of total RNA from transgenic and WT leaf sample, TRI reagent (Sigma) was used. RNA (2 μg) was reverse transcribed, used for cDNA synthesis (using ABI cDNA synthesis kit). qPCR was performed with *AtApx1* primers (listed in Table S3) and endogenous control (*ACT*) along with negative control (lacking cDNA sample) (Petla et al. 2016). Three biological replicates were used with triplicate reaction in each assay. Primers used here are listed in Table S3.

Measurement of physiological growth parameters

To know the physiological and growth status of transgenic plants, rate of photosynthesis was assessed using Infrared gas analyzer (IRGAs) following the manufacturer's instruction. The flow rate of air through the sample chamber was (11 cm^2) set at 0.4 LPM and leaf temperature was maintained at 24 ± 0.8 °C and measured at a time interval of 15/20 s that was kept constant for all samples. Fully expanded leaves in triplicates (transgenic and WT) were used for analysis and five recurrent readings were obtained each time.

Chlorophyll fluorescence measurement

To estimate photosynthetic activity in WT and transgenic plants, we measured chlorophyll fluorescence (Fv/Fm) with a handy chlorophyll fluorescence measuring device (Handy PEA, Hansatech, UK) following the manufacturer's instruction. Measurements for Fv/Fm were recorded in the forenoon hours and for fluorescence level (F_0) in modulated light. The maximum fluorescence level (Fm) was obtained by 1.5 s saturating pulse at 300 mol m^{-2} s^{-1} on dark adopted leaves (20 min).

Soil plant analysis development (SPAD) measurements

SPAD values were recorded in sunlight following the manufacturer's instruction (SPAD meter, Opti Science, CMM-200, USA). The instrument calculates the transmission of red light and infrared light at 650 nm and 940 nm, respectively. Leaves were randomly sampled from matured transgenic and WT plants. Five readings obtained by portable SPAD meter from individual leaves were averaged and calculated.

Leaf disc assay

To examine the morphological difference under salt stress, leaf disc assay was performed on transgenic along with WT plants using varied saline concentrations (NaCl-100, 150 and 200 mM) to the half MS medium (Singla-Pareek et al. 2008). Leaf discs (1 cm) were cut from independent transgenic lines of *B. juncea* and WT plants and floated on 4 ml half MS medium supplemented with different concentrations of NaCl and kept for 48 h at 22 °C under a 16-h photoperiod. Finally, 200 mM concentration of NaCl was chosen for performing further assays on the basis of morphological assessment of leaf discs. The chlorophyll content of the leaf discs was determined for getting information regarding salt stress-induced chlorophyll loss in the transgenic and WT plants. Leaf discs floated on half MS medium without salt served as the experimental control. Leaf samples from control and salt-treated matured plant (both WT and transgenic) were used for further physiological analysis.

Seed germination study and salinity stress treatment

For seed germination study and stress experiments, seeds from both WT and transgenic *B. juncea* were germinated on filter paper (Whatman no. 2) water-logged with 4 mL of the ½ MS medium supplemented with and without 200 mM of NaCl. Seed germination was evaluated after every 12 h. For stress experiments, seeds were germinated on germination

paper with 4 mL of ½ MS medium supplemented with and without 200 mM of NaCl and allowed to grow for 10 days. Ten day old seedlings were chosen for measuring stress markers and antioxidant enzyme activity.

Chlorophyll and carotenoid content

For determining chlorophyll content, the method given by Hiscox and Israelstam (1979) is followed. Fresh leaves (200 mg) were extracted in 5 ml dimethylsulfoxide (DMSO) and incubated at 65 °C for 40 min. After cooling, a 2 ml aliquot of the chlorophyll content was taken into a cuvette and measurement was done spectrophotometrically at 470, 645, 663 nm using DMSO as blank. Respective absorption was used to measure the Chlorophyll a, b, and total chlorophyll (Chl $a + b$), based on the equations reported by (Joshi et al., 2013). Carotenoid content was measured using the method of (Terzi and Kadioglu 2006). Chlorophyll stability index was determined using ratio of chlorophyll content of stressed leaf to control leaf sample (Terzi and Kadioglu 2006). All quantifications were performed in triplicate.

Proline content

Free proline content was measured using the method reported by Bates et al. (1973). Proline concentration was calculated with the help of calibration curve made using L-proline. Further, the final values were expressed as $\mu g \ gm^{-1} \ FW^{-1}$. All quantifications were performed in triplicate.

Malondialdehyde and hydrogen peroxide content

MDA content was estimated using thiobarbituric acid (TBA) method reported by Heath and Packer (1968). The MDA concentration was calculated using an extinction coefficient of 155 $mM^{-1} \ cm^{-1}$. Values are presented from the measurements of three replicates. Hydrogen peroxide (H_2O_2) level was determined following potassium iodide (KI) method spectrophotometrically as described by Salvi et al.(2016). All quantifications were performed in triplicate.

Antioxidative assay ofAPX, GPX, POD and CAT

For enzyme assay, protein was extracted in protein extraction buffer containing 100 mM sodium phosphate buffer (pH 7.0), 5 mM ascorbate, 10% glycerol and 1 mM EDTA (Ghosh et al. 2020). Bradford reagent was used to determine the protein concentration (Bradford 1976).

Specific APX activity was measured following the method reported by Nakano and Asada (1981). Briefly, ascorbate oxidation by H_2O_2 was measured in 1.5 ml assay mix containing 50 mM phosphate buffer (pH 6.0), 0.1 μM

EDTA, 0.5 mM ascorbate, and 1.0 mM H_2O_2 with the addition of 50 μl of leaf extract at 290 nm. APX activity was calculated using extinction coefficient of 2.8 $mM^{-1} \ cm^{-1}$. GPX assay was performed by the method reported in Urbanek et al. (1991). The tetraguaiacol formed was calculated with molar extinction coefficient (26.6 $mM^{-1} \ cm^{-1}$). The enzyme activity was expressed in terms of $\mu mol \ min^{-1} \ mg^{-1}$ protein. For POD activity, the method described in Siegel and Galston (1967) was followed. For CAT activity, we used the method reported by Beers and Sizer (1952) with few modifications. In this method, the reaction mixture of 1.5 ml contains 50 μl of the enzyme extract, 100 mM phosphate buffer (pH 7.0), 0.1 μM EDTA, and 20 mM H_2O_2. The decline in H_2O_2 content was observed at 240 nm, immediately and 1 min after H_2O_2 addition. The difference in absorbance (ΔA_{240}) was divided by the H_2O_2 molar extinction coefficient (36 $M^{-1} \ cm^{-1}$) and the enzyme activity was expressed as μmol of $H_2O_2 \ min^{-1} \ mg^{-1}$ protein.

Statistical analysis

Data presented in this study are mean $\pm$ standard deviation (SD). For statistical analysis (one-way analysis of variance) we used Duncan's multiple range test (DMRT, $\alpha = 0.05$) using SPSS program (SPSS, Chicago, IL, USA). Letters mentioned on the graph designate the similarities or differences between mean values.

Results

Molecular analysis of transgenic lines expressing *AtApx1*

The C58C1Rif^R (pGV2260) strain harboring *AtApx1* construct (described in material and methods Sect. 2.2) (Fig. S1) was transformed into the *B. juncea* using *Agrobacterium*-mediated genetic transformation (Fig. 1a). To optimize the regeneration frequency, different combinations of growth hormones were used, and the appropriate concentration was selected (Table S1). Among different tissues, the segment of hypocotyl has been used as a source of explants to achieve the successful transformation. After transformation several transgenic lines were obtained, later on five independent transgenic lines (L1–L5) were analyzed. To validate the AtApx1 integration, we carried out PCR using *AtApx1* (898 bp) and *nptII* (Kan) (734 bp) gene specific primers in the T_0 generation plants indicating the integration of the *AtApx1* gene (Fig. S2A). For confirming the single transgene integration, southern hybridization was performed. Transgenic lines L3, L4 and L5 were shown to have single copy insertion while WT did not show hybridization signal (Fig. S2B). Further, T_2 plants were raised from the seeds of T_1
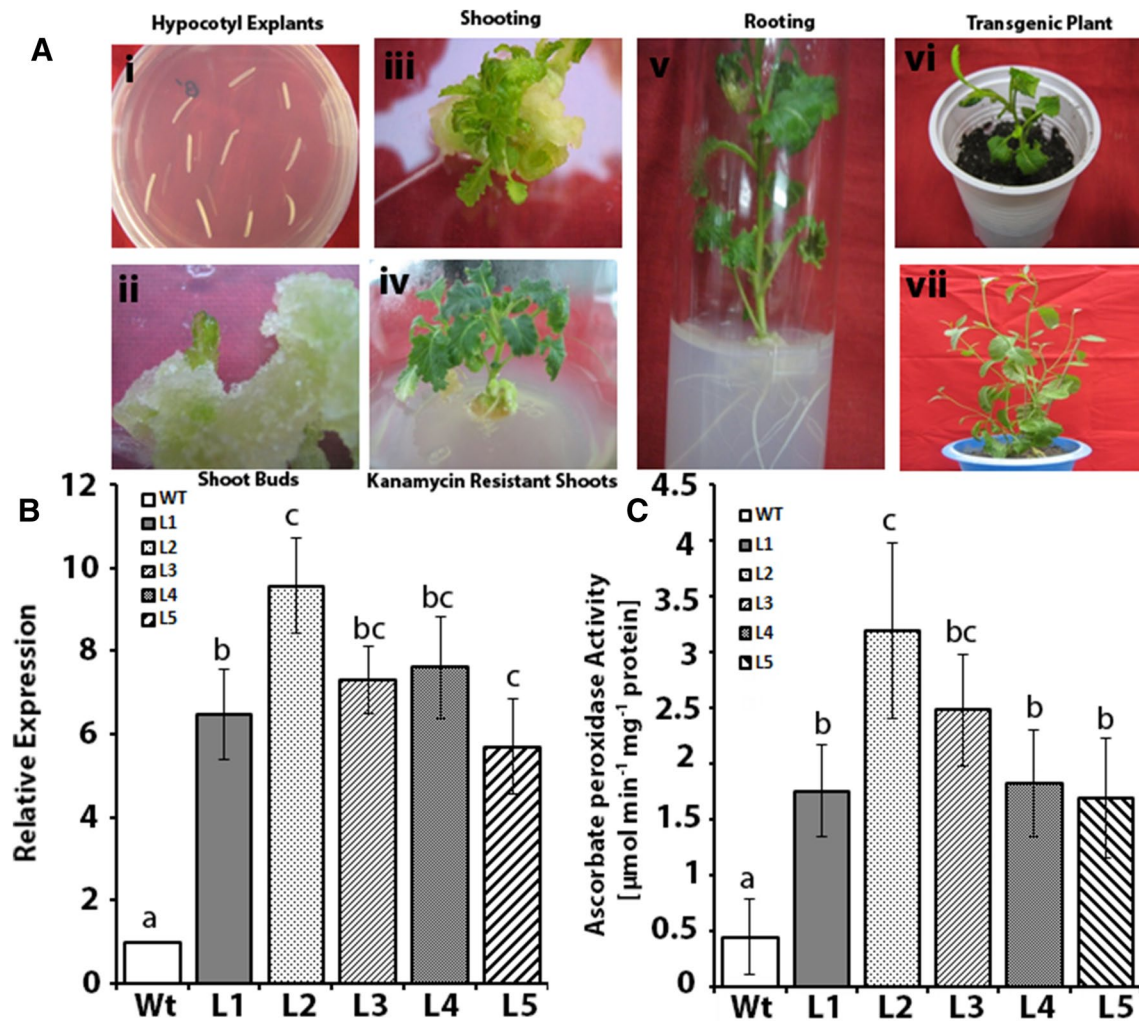
Fig. 1 Generation of transgenic *Brassica juncea* and their analysis. **a** Different stages of *Agrobacterium*-mediated genetic transformation of *Brassica juncea* (I–VI). **I** Hypocotyl explants, **II** shoot buds, **III** shooting. **IV** kanamycin-resistant shoots, **V** rooting, **VI**, **VII** transgenic plant. **b** qRT-PCR of transgenic lines. The relative expression values of each gene were normalized to an endogenous control *ACTIN2* and calculated using $\Delta\Delta$CT method. Values are means of three biological replicates each with three technical replicates. **c** APX activity in transgenic lines. Protein extract (50 µl) was assayed in the reaction mix as described in materials and methods. Error bars indicate the standard deviation. Significant differences among means ($\alpha = 0.05$) are denoted by the different letters

generation (via self-pollination). In control condition, transgenic plants (at T_1 and T_2) appeared similar to WT in general morphology. We measured *AtApx1* transcript as well as APX activity under control conditions in the selected transgenic lines and WT plants. Each transgenic line displayed significant elevation of *AtApx1* transcript and activity (Fig. 1b, c).

AtApx1 overexpression does not alter normal physiological parameters and displays comparable growth with WT plants

Physiological status and growth were compared to rule out the occurrence of any alteration in the physiology of the transgenic plant because of genomic alteration (overexpression of *AtApx1* gene). Gas exchange was measured at the vegetative stage with the young fully expanded leaves of WT as well as transgenic plants under normal growth conditions. During gas exchange measurement, we did not observe significant variation in net photosynthetic rate and stomatal conductance of transgenic leaves than WT leaves (Fig. 2a, b) that revealed their normal photosynthetic efficiency. To assess the health integrity of photosynthetic apparatus, we further determine the cholorophyll flurescence (Fv/Fm). Photosynthetic efficiency of transgenic plants was comparable to the WT plants under control conditions (Fig. 2c) and the architecture of the photosynthetic apparatus of transgenic plants was found to be normal. In addition to the above parameters, SPAD value was also examined which gives an indication of the nitrogen status and relative greenness of leaves. The results suggested that the relative greenness

Fig. 2 Physiological attributes of wild type (WT) and *AtAPX1* transgenic plants under control conditions. **a** Photosynthetic rate, **b** stomatal conductance, **c** chlorophyll fluorescence, **d** soil plant analysis development (SPAD) value. Measurements were taken in one-month old pot grown plant under normal growth conditions. (*WT* wild Type, L1 represents *AtApx1* transgenic line 3; L2 represents *AtApx1* transgenic line 4; L3 represents *AtApx1* transgenic line 5). Error bars indicate the standard deviation. Significant differences among means ($\alpha = 0.05$) are denoted by the different letters
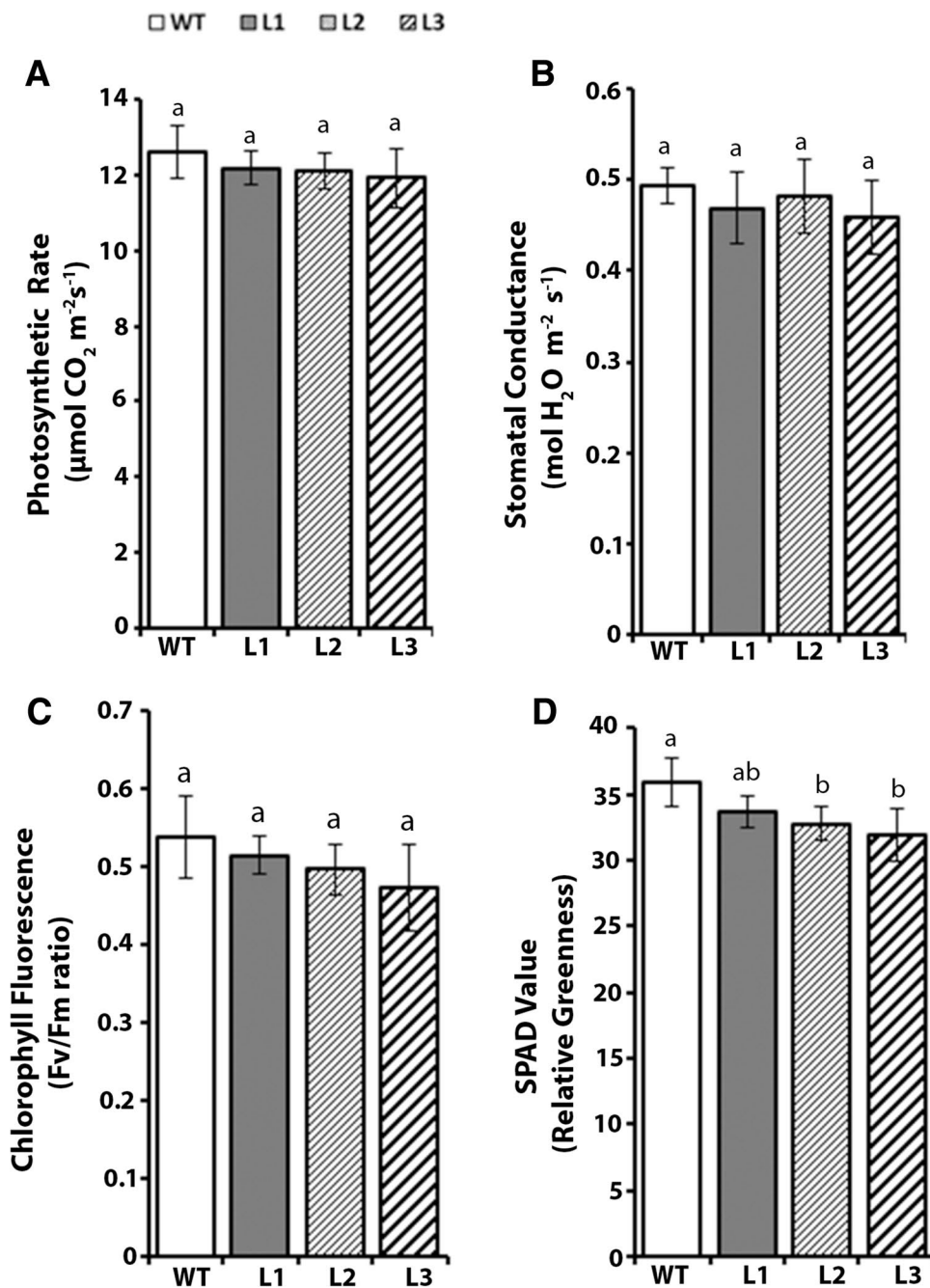
was slightly better in WT than the transgenic under normal growth conditions (Fig. 2d).

AtApx1 overexpression leads to higher chlorophyll retention and enhanced salt tolerance

As shown in Fig. 1b, c the transgenic *B. juncea* plant over-expressing *AtApx1* gene had a higher transcript and APX activity than WT plants. The homozygous lines were used to test, if the increased APX activity would render the plants an advantage in tolerating the salinity stress. As the bleaching

in the leaves is considered as a good indicator of salinity-induced damage, the pot grown transgenic *B. juncea* plants were first evaluated for the leaf disc assay (Fig. 3a).

Under control condition, we did not observe evident difference between WT and transgenic plants. Whereas under leaf disc assay, WT leaf disc treated with salt stress (200 mM for 24 h) showed substantial damage as compared to transgenic leaves. Further during chlorophyll and carotenoid content measurement, we found that chlorophyll *a*, chlorophyll *b* and total chlorophyll concentrations were significantly declined in all the plants exposed to
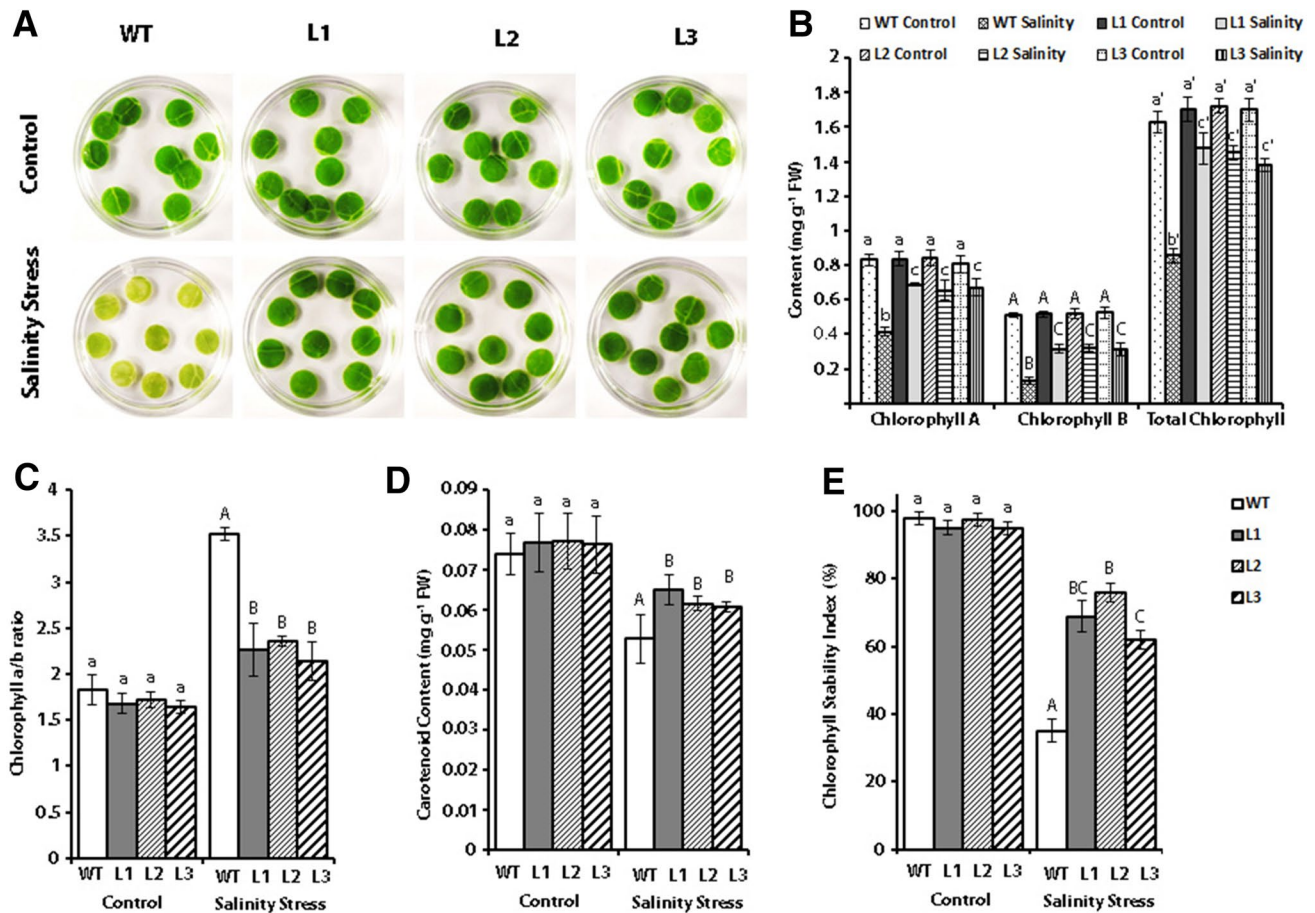
Fig. 3 Salt stress tolerance assay of *AtApx1 Brassica juncea* transgenic plants. Retardation of salt stress promoted senescence in one-month old transgenic *Brassica* plant overexpressing *AtApx1*, indicating the tolerance at cellular level towards toxic levels of 200 mM salt after 48 h of treatment. Leaf disc floating in MS medium without salt served as the experimental control. **a** Comparative phenotype analysis in WT and transgenic *Brassica* leaf disc under control (upper panel) and salt treated (lower panel) conditions, **b** chlorophyll content, **c** chlorophyll *a/b* ratio, **d** carotenoid content, **e** chlorophyll stability index (%). (*WT* wild Type, L1 represents *AtApx1* transgenic line 3; L2 represents *AtApx1* transgenic line 4; L3 represents *AtApx1* transgenic line 5). Error bars indicate the standard deviation. Significant differences among means ($\alpha = 0.05$) are denoted by the different letters. Values are means of three biological replicates each with three technical replicates

salinity stress; however, such a decrease was more profound in WT (Fig. 3b). The total chlorophyll content decreased by 48% in WT plants under 200 mM NaCl stress as compared to the transgenic leaf discs which recorded about 15% decrease under salinity stress, over that of their respective control. (Fig. 3b). Chlorophyll *a/b* ratio was enhanced in both WT and transgenic after salt stress treatment, however, the increase in WT was 2-fold while it was 1.3-fold in transgenic over their respective control (Fig. 3c).

Total carotenoid content decreased in leaves of WT and transgenic plants after salinity stress; however, the decrease was lesser in transgenic plants as compared to WT (Fig. 3d). These results were also supported by the chlorophyll stability index which is a measure of stress tolerance potential of plants (Fig. 3e). Chlorophyll stability index was 1.9-fold

higher in the leaves of transgenic plant than the WT (Wanichananan 2003).

AtApx1 overexpression lines accumulate higher proline, lower H_2O_2 and MDA content under salinity stress

Plants exposed to salt stress preferentially accumulate higher concentrations of proline that acts as a rescue molecule against oxidative injury (Jain et al. 2001; Jebara et al. 2005; Koca et al. 2007). Therefore, WT and transgenic plants were evaluated for the accumulation of proline content after salinity stress. As anticipated, after stress exposure proline content accumulated in both WT and transgenics; however, net proline accumulation was 1.8-fold higher in the transgenic as compared to WT (Fig. 4a).

The cellular H_2O_2 level in WT plant was about 2.8-fold higher in contrast to transgenic plant under salinity stress (Fig. 4b). To further extend this analysis, we evaluated malondialdehyde (MDA) content to evaluate H_2O_2-mediated damage. MDA content was found to increase drastically by almost 2.5-fold in WT plants when challenged to salt stress, whereas its level in the transgenic plants increased only by 1.5-fold, indicating that WT plants experienced higher loss in membrane integrity (Fig. 4c). These results clearly depicted that under similar stress condition transgenic plants hold better capacity to detoxify ROS in comparison to WT.

AtApx1 overexpression resulted in improved germination rate after salinity stress

Under salinity stress treatment, transgenic lines showed better % germination rate (Fig. 5a and b). After 48 h of stress treatment % germination was recorded to be just 50% in WT. However, the germination rate was similar in both WT and transgenic seeds under control condition (without NaCl). Seeds were counted as germinated just after the radicle production from the seeds.

AtApx1 overexpression confers salinity stress tolerance and improves antioxidant enzyme activities in *B. juncea*

We analyzed the activity profile of different enzymes (APX, GPX, CAT and POD) associated with ROS scavenging pathway (Fig. 6a–d). The activity profile of all plants was similar under control condition while an evident upregulation of APX, GPX and POD was observed in transgenic plants (Fig. 6a–c).

The APX specific activity was found to be 1.9-fold higher in transgenic than the WT after salt stress treatment (Fig. 6a). It is clear that APX activity is significantly induced under salinity stress, although a distinctly higher APX activity in the transgenics is attributed to the over-expression of *AtApx1* gene. Additionally, we also estimated the GPX and POD activity to explore the antioxidative machinery of transgenic plants under salinity stress. The results revealed that under salinity stress transgenic plant possess 1.5-fold higher GPX
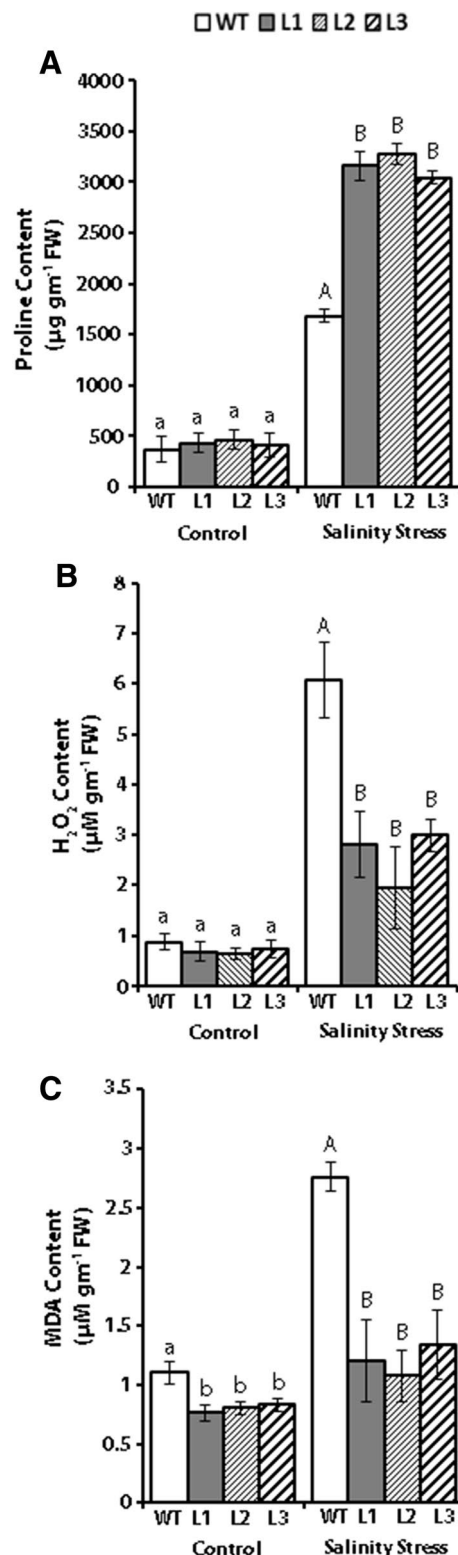
activity than that of WT (Fig. 6b). Similarly, total peroxidase activity also showed a marked increase in transgenic plant (twofold) than WT (Fig. 6c). To our surprise, catalase activity was found to decrease in WT as well as the transgenic

Fig. 5 *Brassica juncea AtApx1* transgenic lines show better germination rate under salt stress treatment. **a** Comparative seed germination study in WT and transgenic *Brassica* under control conditions (upper panel) and salt treated (lower panel). **b** Graph showing seed germination rate in WT and transgenic lines after 48 h under control conditions and salt treatment. **c** Stress treatment to WT and transgenic *Brassica* seedlings (10 days old) (*WT* wild Type; L1 represents *AtApx1* transgenic line 3; L2 represents *AtApx1* transgenic line 4; L3 represents *AtApx1* transgenic line 5). Error bars indicate the standard deviation. Significant differences among means ($\alpha = 0.05$) are denoted by the different letters. Values are means of three biological replicates each with three technical replicates
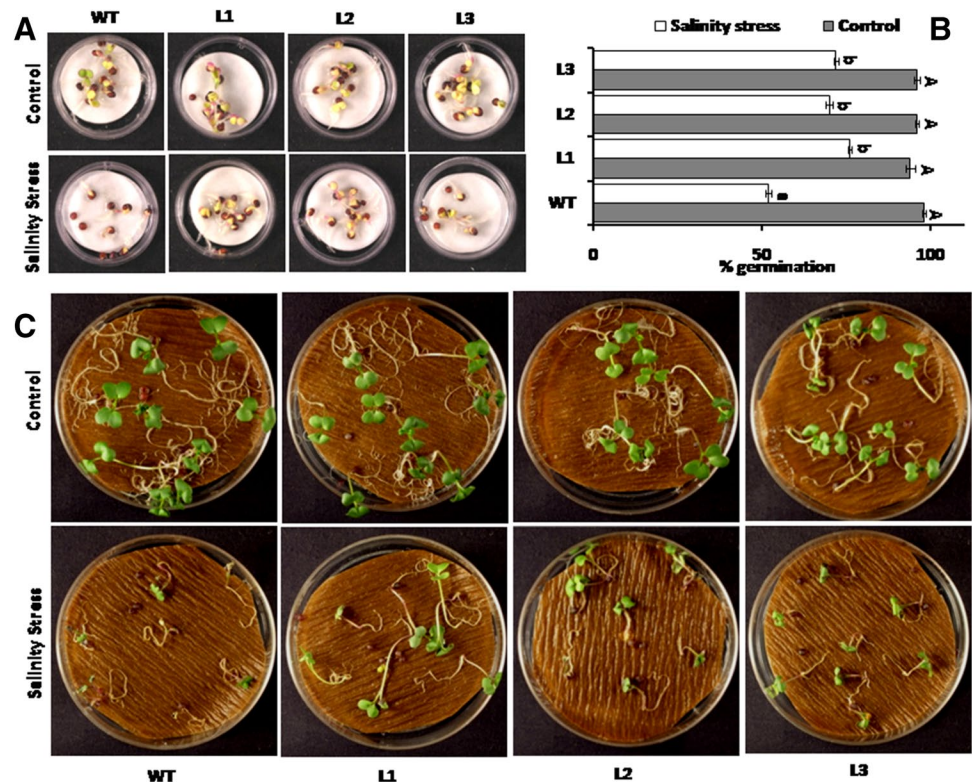
after salt stress (Fig. 6d). These results lead us to speculate that the ectopic overexpression of *AtApx1* activated several other enzymes associated with the ascorbate–glutathione cycle in the transgenic plants under salinity stress.

Discussion

The salinity stress is inevitably associated with enhanced ROS accumulation and oxidative stress which drastically reduces the crop yield. Among several adaptive strategies, activating the antioxidative system to limit the detrimental effects of ROS is a preferred choice for plants (Foyer and Noctor 2003; Mittler and Zilinskas 1994). A positive correlation has been reported between the antioxidant enzyme activity of enzymes such as APX, POD and SOD and detoxification of ROS under salinity stress (Cakmak and Marschner 1992; Jebara et al. 2005). APX is a key antioxidative enzyme which detoxifies H_2O_2 to H_2O and O_2 (Davletova et al. 2005; Kubo et al. 1992; Nakano and Asada 1981).

Since, H_2O_2 is easily diffusible through biological membranes, it accumulates excessively in the cytosol. Incidentally, cytosolic isoform of APX is the one which responds maximally to externally imposed stress (Davletova et al. 2005). Therefore, we attempted to genetically engineer *B. juncea* (var. pusajaikisan) by overexpressing a cytosolic APX so as to strengthen the antioxidative defense system of the plants.

In this study, we have demonstrated that ectopic expression of ascorbate peroxidase gene (*AtApx1*) in the economically important oilseed crop *B. juncea* could improve salinity stress tolerance. These findings are well correlated with other reports which suggested that increased APX activity plays a promising role under salinity stress tolerance (Gossett et al. 1994; Jebara et al. 2005).

To analyze the growth status of transgenic *B. juncea* several physiological parameters were examined viz photosynthetic rate, stomatal conductance, chlorophyll fluorescence and SPAD values. Stomatal conductance is well correlated with the gas exchange and CO_2 assimilation (Lincoln and Zeiger 2006; Salisbury and Ross 1992) which ultimately affect photosynthetic rate and in turn is linked with the productivity of plant. Chlorophyll fluorescence is an important criterion to determine the photosynthetic efficiency. The Fv/Fm ratio is used to measure the quantum efficiency of the electron transport in photosystem II (PSII) (Maxwell and Johnson 2000). Higher the variable fluorescence, higher is the photosynthetic capacity of the leaf. Thus, the parameters, maximal fluorescence (Fm), variable fluorescence (Fv) and the Fv/Fm ratio characterize the functional state of PSII (Moradi and Ismail 2007). Similarly, chlorophyll content in terms of relative greenness measured by SPAD meter indicates the growth status of plants. The lower SPAD values for transgenic plants might result through the channeling of carbon and metabolic energy towards the continuous synthesis of APX and
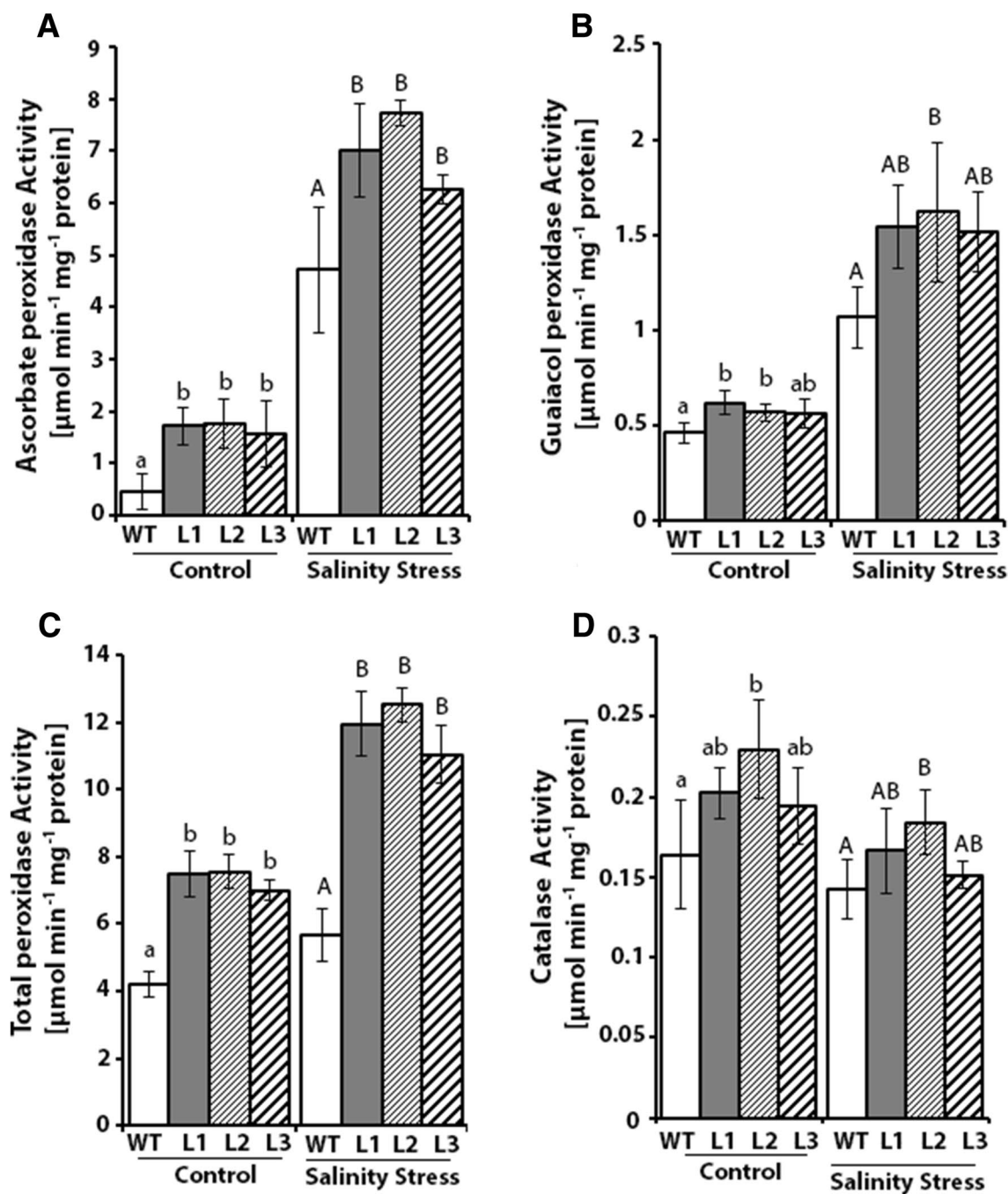
Fig. 6 *Brassica juncea AtApx1* transgenic lines maintain higher anti-oxidative potential after salt stresses. Quantitative measurement of antioxidative activity of **a** ascorbate peroxidase, **b** guaiacol peroxidase, **c** total peroxidase, **d** catalase activity. Quantifications were done in control and after salt stress treatments in one-month old plant as described in the materials and methods section. (*WT* wild Type; L1 represents *AtApx1* transgenic line 3; L2 represents *AtApx1* transgenic line 4; L3 represents *AtApx1* transgenic line 5). Error bars indicate the standard deviation. Significant differences among means ($\alpha = 0.05$) are denoted by the different letters. Values are means of three biological replicates each with three technical replicates

possible changes in the flavonoid and carotene contents of plants. On comparative analysis, these parameters which inferred the constitutive expression of *AtApx1* did not obtrude the photosynthetic rate, stomatal conductance and chlorophyll fluorescence of transgenic plants under normal growth conditions (Fig. 2). This avails us to ascertain about the uniformity of further experimental analysis.

Salinity stress is known to impair the biosynthesis of enzymes involved in the maintenance of photosynthetic apparatus and may also trigger chlorophyll degradation

which precisely reduced the chlorophyll content in salt-stressed plant (El-Tayeb 2005; Fang et al. 1998). Chlorophyll content can be precisely correlated with the cellular metabolic status; thus, its reduction indicates oxidative damage in tissues due to accumulation of ions and other active oxygen species. Loss of chlorophyll pigment under salt stress was observed for some salt-sensitive plant species (Fang et al. 1998; Saxena et al. 2013). In our investigation, leaf disc senescence assay displayed higher degree of bleaching in WT and turned yellow in the presence of 200 mM NaCl as compared to the transgenic. The reduced chlorosis in the leaf discs of transgenics were well correlated with the quantitative measurement of chlorophyll content in salt-treated matured plant (Fig. 3). Likewise, the lower chlorophyll content is possibly associated with its impaired biosynthesis due to salinity-induced ROS attack. The higher activity of APX in transgenics could protect chlorophyll biosynthesis or curtail its degradation (Maruta et al. 2010; Stenbaek and Jensen 2010). Reduced sensitivity to salinity stress in *AtApx1* over-expressing transgenic leaf is similar to the earlier observation (Badawi et al. 2004; Jebara et al. 2005; Moradi and Ismail 2007). In addition, the chlorophyll stability index and carotenoid content (Fig. 3d, e) were higher in the transgenic plants compared to WT plants, indicating a more efficiently functioning photosynthetic apparatus. The higher chlorophyll stability index (CSI) is considered as a potential indicator of stress tolerance capacity of plants. Higher CSI is associated with the better availability of chlorophyll, thus generating a higher photosynthetic rate (Moradi and Ismail 2007). Also, the chlorophyll *b* which showed higher reduction than chlorophyll *a* indicates that chlorophyll b is more sensitive than chlorophyll a under stress conditions (Netondo et al. 2004). Further, the higher chlorophyll *a/b* ratio has been associated with the alteration of pigment composition of photosynthetic system which possesses lower level of light harvesting chlorophyll proteins (LHCPs) following abiotic stress treatment (Loggini et al. 1999; Nakano and Asada 1987).

Due to excess ROS, the proper functioning of macromolecule present in cytosol as well as sub cellular organelles such as chloroplasts and mitochondria are severely hampered. Enhanced salt tolerance accompanies the accumulation of a high level of proline content during stress conditions. Being a cytosolic osmolytica, proline not only offers osmotic balance, rather it may also stabilize the structure and function of other cellular macromolecules such as regulatory proteins and nucleic acid (Jain et al. 2001; Kubo et al. 1992; Saxena et al. 2011; Strizhov et al. 1997). Higher proline content in transgenic leaves under control conditions could be attributed to the changes in the normal biosynthetic reactions in the cells, which are being forced to continuously express *AtApx1* gene introduced under the control of a constitutive promoter. However, under salt stress treatment a significantly

higher proline accumulation in transgenic seedlings could be attributed to the protective role of proline as an osmoprotectant as well as a ROS scavenger (Matysik et al. 2002). Higher accumulation of free proline in salinity stress tolerance has been well documented (Jebara et al. 2005). It has been reported that overexpression of proline biosynthesis enzymes endows reduced oxidative damage and improved photosynthesis (Strizhov et al. 1997). H_2O_2 is a detrimental ROS, which imposes severe effects on cellular machinery (Salvi et al. 2017, 2020). Salinity treatments caused significant increase in H_2O_2 content and thus lipid peroxidation. MDA content reflects on the degree of lipid peroxidation and thus indicates ROS-triggered cellular damage (Salvi et al. 2017). Relatively higher H_2O_2 accumulation and lipid peroxidation has been documented in salt stress sensitive varieties as compared to resistant variety (Gossett et al. 1994). The apparently lower level of MDA, a lipid peroxidation product and consequently reduced membrane damage in transgenic plants can be explained by lower H_2O_2 level in respective plants. Thus, the enhanced tolerance observed is likely to be a result of lower H_2O_2 level that is attributed to significant protection against lipid peroxidation and eventually protects the cellular integrity under oxidative injury. H_2O_2 imposes toxic effects on cell metabolism which eventually impair the growth and development of plant, thus cells require a fast and efficient way of minimizing H_2O_2 toxicity and the formation of the highly toxic hydroxyl radicals (Tanaka et al. 1985).

It is well known fact that salt stress can cause excessive accumulation of ROS; including hydrogen peroxide (H_2O_2) resulting in oxidative damage of cells (Miller et al. 2010). The H_2O_2 formed as a product of superoxide dismutase activity is a potential damaging agent under different abiotic stresses (Cakmak and Marschner 1992).The lower accumulation of H_2O_2 content in transgenic plants with salinity treatment can be directly correlated with ectopic overexpression of *AtApx1*. Induction in enzyme activity of antioxidative pathways is positively correlated with reduced ROS and enhanced stress tolerance capacity in plants. In *Mesembryanthemum crystallinum* cytosolic *Apx* transcripts and its enzyme activities are found to be induced under salt stress treatment. Being an intricate system, different enzymes of antioxidative pathways are interrelated, thus co-regulated. So, the over-expression of one enzyme can also influence the activity of other enzymes (Gossett et al. 1994; Loggini et al. 1999). Similarly, the induction in the enzyme activity of peroxidase and guaiacol peroxidase was observed in this study which may be due to the overexpression of APX activity. Though guaiacol peroxidase and total peroxidase activity induced in both WT and transgenic, however net GPX activity was higher in the transgenic than the WT in salinity stress. Kim et al. (2005) documented the induction in GPX activity, along with APX and catalase activity under

salinity stress in barley shoots and roots. A consequently increased activity of total peroxidase and guaiacol peroxidase in transgenics, under salt treatment, indicate the existence of an intricate cross-talk mechanism among different peroxidases in the cell. In contrary, catalase activity did not exhibit significant difference, under control and salt-stressed conditions. The decrease in catalase activity under stress conditions might be due to inhibition of biosynthesis or inactivation of catalase activity due to stress (Shim et al. 2003). A similar decline in CAT activity has been reported in rice subjected to salt stress (Lee et al, 2001). Shi et al. (2001) also found a decline in CAT activity under heat stress and salt stress in *Arabidopsis* plants overexpressing peroxisomal APX. Protective roles of APX and related antioxidant enzymes including CAT, GPX and POD in mitigating salt stress have been reported (Gossett et al. 1994; Mittler and Zilinskas 1992).Taken together our results indicated that *AtApx1* overexpressing plants are more tolerant to salinity than WT plants, with transgenic plants having higher chlorophyll retention, antioxidant enzyme activities, proline content and better net photosynthesis than WT plants.

Conclusion

Collectively, in the present studies a successful attempt has been made to strengthen the antioxidative defense potential of *B. juncea* (*var*. pusajaikisan), through overexpression of cytosolic ascorbate peroxidase gene. The over expressed enzyme could substantially improve the salinity stress tolerance profile of the host plant, as assessed through various physiological and biochemical parameters. Thus, *AtApx1* in *B. juncea* plays a pivotal role in preventing the excess accumulation of ROS and helps in enhancing the stress tolerance potential of host plant. APX along with the coordination of other antioxidative enzymes like GPX, catalase and POD maintains the ROS homeostasis and provides tolerance to the cell.

Compliance with ethical standards

References

Anjum NA, Gill SS, Ahmad I, Pacheco M, Duarte AC, Umar S, Khan NA, Pereira ME (2012) The plant family Brassicaceae: an introduction the plant family Brassicaceae. Springer, Berlin, pp 1–33

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Ann Rev Plant Biol 55:373–399

Augustine R, Mukhopadhyay A, Bisht NC (2013) Targeted silencing of BjMYB28 transcription factor gene directs development of low glucosinolate lines in oilseed *Brassica juncea*. Plant Biotechnol J 11(7):855–866

Badawi GH, Kawano N, Yamauchi Y, Shimada E, Sasaki R, Kubo A, Tanaka K (2004) Overexpression of ascorbate peroxidase in tobacco chloroplasts enhances the tolerance to salt stress and water deficit. Physiologia Plant 121(2):231–238

Bates L, Waldren R, Teare I (1973) Rapid determination of free proline for water-stress studies. Plant Soil 39(1):205–207

Beers RF, Sizer IW (1952) A spectrophotometric method for measuring the breakdown of hydrogen peroxide by catalase. J Biol Chem 195(1):133–140

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72(1):248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Cakmak I, Marschner H (1992) Magnesium deficiency and high light intensity enhance activities of superoxide dismutase, ascorbate peroxidase, and glutathione reductase in bean leaves. Plant physiol 98(4):1222–1227

Davletova S, Rizhsky L, Liang H, Shengqiang Z, Oliver DJ, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of Arabidopsis. Plant Cell 17(1):268–281

El-Tayeb M (2005) Response of barley grains to the interactive e. ect of salinity and salicylic acid. Plant Growth Regul 45(3):215–224

Fang Z, Bouwkamp JC, Solomos T (1998) Chlorophyllase activities and chlorophyll degradation during leaf senescence in non-yellowing mutant and wild type of *Phaseolus vulgaris* L. J Exp Bot 49(320):503–510

Fourcroy P, Vansuyt G, Kushnir S, Inzé D, Briat J-F (2004) Iron-regulated expression of a cytosolic ascorbate peroxidase encoded by the APX1 gene in Arabidopsis seedlings. Plant physiol 134(2):605–613

Foyer CH, Noctor G (2003) Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. Physiol Plant 119(3):355–364

Gadea J, Conejero V, Vera P (1999) Developmental regulation of a cytosolic ascorbate peroxidase gene from tomato plants. MGG 262(2):212–219

Ghosh S, Kamble NU, Verma P, Salvi P, Petla BP, Roy S, Rao V, Hazra H, Varshney V, Kaur H, Majee M (2020) *Arabidopsis* protein L-ISOASPARTYL METHYLTRANSFERASE repairs isoaspartyl damage to antioxidant enzymes and increases heat and oxidative stress tolerance. J Biol Chem 295(3):783–799. https://doi.org/10.1074/jbc.RA119.010779

Gill SS, Tuteja N (2010) Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiol Biochem 48(12):909–930

Gossett DR, Millhollon EP, Lucas M (1994) Antioxidant response to NaCl stress in salt-tolerant and salt sensitive cultivars of cotton. Crop Sci 34(3):706–714

Heath RL, Packer L (1968) Photoperoxidation in isolated chloroplasts: I. Kinetics and stoichiometry of fatty acid peroxidation. Arch Biochem Biophys 125(1):189–198

Hernandez J, Jimenez A, Mullineaux P, Sevilia F (2000) Tolerance of pea (*Pisum sativum* L.) to long term salt stress is associated with induction of antioxidant defences. Plant Cell Environ 23(8):853–862

Hiscox JT, Israelstam G (1979) A method for the extraction of chlorophyll from leaf tissue without maceration. Can J Bot 57(12):1332–1334

Jain M, Mathur G, Koul S, Sarin N (2001) Ameliorative effects of proline on salt stress-induced lipid peroxidation in cell lines of groundnut (*Arachis hypogaea* L.). Plant Cell Rep 20(5):463–468

Jebara S, Jebara M, Limam F, Aouani ME (2005) Changes in ascorbate peroxidase, catalase, guaiacol peroxidase and superoxide dismutase activities in common bean (*Phaseolus vulgaris*) nodules under salt stress. J Plant Physiol 162(8):929–936

Joshi R, Ramanarao MV, Baisakh N (2013) Arabidopsis plants constitutively overexpressing a myoinositol 1-phosphate synthase gene (SaINO1) from the halophyte smooth cordgrass exhibits enhanced level of tolerance to salt stress. Plant Physiol Biochem 65:61–66

Kim I-J, Chung W-I (1998) Molecular characterization of a cytosolic ascorbate peroxidase in strawberry fruit1. Plant Sci 133(1):69–77

Kim S-Y, Lim J-H, Park M-R, Kim Y-J, Park T-I, Seo Y-W, Choi K-G, Yun S-J (2005) Enhanced antioxidant enzymes are associated with reduced hydrogen peroxide in barley roots under saline stress. BMB Rep 38(2):218–224

Koca H, Bor M, Özdemir F, Türkan I (2007) The effect of salt stress on lipid peroxidation, antioxidative enzymes and proline content of sesame cultivars. Environ Exper Bot 60(3):344–351

Kubo A, Saji H, Tanaka K, Tanaka K, Kondo N (1992) Cloning and sequencing of a cDNA encoding ascorbate peroxidase from *Arabidopsis thaliana*. Plant Mol Biol 18(4):691–701

Lee DH, Kim YS, Lee CB (2001) The inductive responses of the antioxidant enzymes by salt stress in the rice (*Oryza sativa* L.). J Plant Physiol 158(6):737–745. https://doi.org/10.1078/0176-1617-00174

Lincoln T, Zeiger E (2006) Secondary metabolites and plant defense. Plant Physiol 13:125 **[Fourth Edition. Sinauer Associates, Inc. Capítulo]**

Loggini B, Scartazza A, Brugnoli E, Navari-Izzo F (1999) Antioxidative defense system, pigment composition, and photosynthetic efficiency in two wheat cultivars subjected to drought. Plant physiol 119(3):1091–1100

Maruta T, Tanouchi A, Tamoi M, Yabuta Y, Yoshimura K, Ishikawa T, Shigeoka S (2010) Arabidopsis Chloroplastic Ascorbate Peroxidase Isoenzymes play a dual role in Photoprotection and gene regulation under Photooxidative stress. Plant Cell Physiol 51(2):190–200. https://doi.org/10.1093/pcp/pcp177

Matysik J, Alia, Bhalu B, Mohanty P (2002) Molecular mechanisms of quenching of reactive oxygen species by proline under stress in plants. Curr Sci 82(5):525–532

Maxwell K, Johnson GN (2000) Chlorophyll fluorescence—a practical guide. J Exp Bot 51(345):659–668

Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R (2010) Reactive oxygen species homeostasis and signalling during drought and salinity stresses. Plant Cell Environ 33(4):453–467

Mittler R, Zilinskas B (1992) Molecular cloning and characterization of a gene encoding pea cytosolic ascorbate peroxidase. J Biol Chem 267(30):21802–21807

Mittler R, Zilinskas BA (1994) Regulation of pea cytosolic ascorbate peroxidase and other antioxidant enzymes during the progression of drought stress and following recovery from drought. Plant J 5(3):397–405

Moradi F, Ismail AM (2007) Responses of photosynthesis, chlorophyll fluorescence and ROS-scavenging systems to salt stress during seedling and reproductive stages in rice. Ann Bot 99(6):1161–1173

Nakano Y, Asada K (1981) Hydrogen peroxide is scavenged by ascorbate-specific peroxidase in spinach chloroplasts. Plant Cell Physiol 22(5):867–880

Nakano Y, Asada K (1987) Purification of ascorbate peroxidase in spinach chloroplasts; its inactivation in ascorbate-depleted medium and reactivation by monodehydroascorbate radical. Plant Cell Physiol 28(1):131–140

Netondo GW, Onyango JC, Beck E (2004) Sorghum and salinity: I. Response of growth, water relations, and ion accumulation to NaCl salinity. Crop Sci 44(3):797

Noctor G, Foyer CH (1998) Ascorbate and glutathione: keeping active oxygen under control. Annu Rev Plant Biol 49(1):249–279

Petla BP, Kamble NU, Kumar M, Verma P, Ghosh S, Singh A, Rao V, Salvi P, Kaur H, Saxena SC, Majee M (2016) Rice protein l-isoaspartyl methyltransferase isoforms differentially accumulate during seed maturation to restrict deleterious isoAsp and reactive oxygen species accumulation and are implicated in seed vigor and longevity. New Phytol 211(2):627–645. https://doi.org/10.1111/nph.13923

Salisbury F, Ross C (1992) Mineral nutrition. Plant Physiol 4:116–135

Salvi P, Saxena SC, Petla BP, Kamble NU, Kaur H, Verma P, Rao V, Ghosh S, Majee M (2016) Differentially expressed galactinol synthase(s) in chickpea are implicated in seed vigor and longevity by limiting the age induced ROS accumulation. Sci Rep 6:35088. https://doi.org/10.1038/srep35088

Salvi P, Kamble NU, Majee M (2017) Stress inducible Galactinol Synthase of Chickpea (*CaGolS*) implicates in heat and oxidative stress tolerance through reducing stress induced excessive reactive oxygen species accumulation. Plant Cell Physiol. https://doi.org/10.1093/pcp/pcx170

Salvi P, Kamble NU, Majee M (2020) Ectopic over-expression of ABA-responsive Chickpea galactinol synthase (CaGolS) gene results in improved tolerance to dehydration stress by modulating ROS scavenging. Environ Exp Bot 171:103957. https://doi.org/10.1016/j.envexpbot.2019.103957

Sato Y, Murakami T, Funatsuki H, Matsuba S, Saruyama H, Tanida M (2001) Heat shock-mediated APX gene expression and protection against chilling injury in rice seedlings. J Exp Bot 52(354):145–151

Saxena S, Joshi P, Grimm B, Arora S (2011) Alleviation of ultraviolet-*C*-induced oxidative damage through overexpression of cytosolic ascorbate peroxidase. Biologia 66(6):1052–1059

Saxena SC, Salvi P, Kaur H, Verma P, Petla BP, Rao V, Kamble N, Majee M (2013) Differentially expressed myo-inositol monophosphatase gene (CaIMP) in chickpea (*Cicer arietinum* L.) encodes a lithium-sensitive phosphatase enzyme with broad substrate specificity and improves seed germination and seedling growth under abiotic stresses. J Exp Bot 64(18):5623–5639. https://doi.org/10.1093/jxb/ert336

Shi W, Muramoto Y, Ueda A, Takabe T (2001) Cloning of peroxisomal ascorbate peroxidase gene from barley and enhanced thermotolerance by overexpressing in *Arabidopsis thaliana*. Gene 273(1):23–27

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y, Yoshimura K (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot 53(372):1305–1319

Shim I-S, Momose Y, Yamamoto A, Kim D-W, Usui K (2003) Inhibition of catalase activity by oxidative stress and its relationship to salicylic acid accumulation in plants. Plant Growth Regul 39(3):285–292

Siegel BZ, Galston A (1967) The isoperoxidases of *Pisum sativum*. Plant Physiol 42(2):221–226

Singla-Pareek SL, Yadav SK, Pareek A, Reddy MK, Sopory SK (2008) Enhancing salt tolerance in a crop plant by overexpression of glyoxalase II. Transgenic Res 17(2):171–180. https://doi.org/10.1007/s11248-007-9082-2

Stenbaek A, Jensen PE (2010) Redox regulation of chlorophyll biosynthesis. Phytochemistry 71(8–9):853–859

Strizhov N, Ábrahám E, Ökrész L, Blickling S, Zilberstein A, Schell J, Koncz C, Szabados L (1997) Differential expression of two *P5CS* genes controlling proline accumulation during salt stress requires ABA and is regulated by ABA1, ABI1 and AXR2 in Arabidopsis. Plant J 12(3):557–569

Tanaka K, Suda Y, Kondo N, Sugahara K (1985) O3 tolerance and the ascorbate-dependent H_2O_2 decomposing system in chloroplasts. Plant Cell Physiol 26(7):1425–1431

Tanaka K, Masuda R, Sugimoto T, Omasa K, Sakaki T (1990) Water deficiency-induced changes in the contents of defensive substances against active oxygen in spinach leaves. Agric Biol Chem 54(10):2629–2634

Terzi R, Kadioglu A (2006) Drought stress tolerance and the antioxidant enzyme system. Acta Biol Cracov Ser Bot 48:89–96

Urbanek H, Kuzniak-Gebarowska E, Herka K (1991) Elicitation of defence responses in bean leaves by *Botrytis cinerea* polygalacturonase. Acta Physiol Plant (Poland) 13:43–50

Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218(1):1–14

Wanichananan P, Kirdmanee C, Vutiyano C (2003) Effect of salinity on biochemical and physiological characteristics in correlation to selection of salt-tolerance in aromatic rice (*Oryza sativa* L.). Sci Asia 29(4):333–339

Enhancing Accuracy for Human Trajectory Forecasting in Crowded Scenes

Aditya Singhal
Delhi Technological University
Delhi, India
aditya_bt2k16@dtu.ac.in

S. Indu
Delhi Technological University
Delhi, India
s.indu@dtu.ac.in

ABSTRACT

We address the problem of accurately predicting the trajectories of pedestrians in crowded scenes. Mapping and plotting of human trajectories are well-known and key problems in autonomous systems, the goal of which is to enable self-driving cars to explore a new environment autonomously. Extending the classical RNN (recurrent neural networks) and LSTM (long short-term memory) based models, we propose in this work a learning-based approach for error reduction and accuracy improvement by utilizing visualizations of the crowded scenes. Our method works effectively with state-of-the-art models on ETH and UNIV-UCY datasets. It encompasses several pooling methods to join social, navigation, and semantic tensors. Our proposed model is tested in unstructured environments and demonstrates how awareness of social interactions helps to correctly predict future positions. We validate our approach in simulated real-world environments and demonstrate error reductions over competitive baseline approaches.

CCS Concepts

• **Computing methodologies** → **Artificial intelligence** → **Computer vision** → **Computer vision tasks** → **Vision for robotics**

Keywords

Computer Vision; Trajectory Forecasting; Autonomous Systems; LSTM model; Deep Learning

1. INTRODUCTION

Human trajectory forecasting is an apposite topic in computer vision and has numerous real-life applications. Autonomous systems such as socially aware robots and self-driving cars need to anticipate human path in order to optimize their paths and to better comply with human motion. Delivery robots could arrive at their destinations by avoiding people and obstacles and consequently reduce energy consumption. Moreover, human anomaly detection is possible using fixed cameras not only in urban open spaces (e.g., parks, streets, etc.), but also in crowded scenes (e.g., subway stations, airports, shopping malls, etc.). This process of trajectory prediction by a mobile robot has been a key problem in computer vision for the past several decades.

Even though recurrent neural networks for time-series prediction have produced significant results, many problems still remain. Most of the previous work has focused on data-driven approaches, which are usually unaware of surrounding elements which constitute one of the main reasons of direction changes in an urban scenario. As a result, these systems are less reliable in real-world environments.

When moving through a crowded space, people are inclined to rely on previous experience to conform to observed patterns and visual stimuli to select shortest route or avoid threats. In addition, while walking in an urban environment, they take into account which kind of *objects* they encounter in the neighborhood. Velocity direction changes are also possible in several situations. For example, people tend to move around a roundabout rather than across it. In some cases, they use different pace according to weather conditions. For instance, a sudden snowfall can cause people to move indoors quickly to take cover. In this context, the LSTM networks, which have the ability to learn, remember and forget through gates, and are one of the most suitable solution for solving sequence-to-sequence problems have been extensively used over the past few years [1, 2]. Despite their success, they mainly focus on modelling human-human interactions and human-space interactions [3, 4, and 5]. The early works on human-space interactions have demonstrated the importance of scene context information in improving prediction of future positions by adding different physical constraints and more realistic paths [6, 7, and 8].

In this work, we propose an approach to utilize the visualizations of the crowded scenes in tackling the problem of increased error. Our goal is to maximize the accuracy of predicted trajectories by learning the entire pipeline from a visual environment and calculating values for numerical stability.

In particular, we aim to exploit the local-awareness of surrounding space by combining social and semantic elements. Our work is built upon the Social-LSTM model proposed by Alahi *et al.* [3], and Social and Scene Aware-LSTM model proposed by Lisotto *et al.* [9], by implanting new factors which encode human-space interactions in order to attain more accurate predictions. More specifically, we use a data-driven approach to extract social conventions from observed trajectories and augment it with semantic information from neighborhood in an LSTM based architecture.

Figure 1. The ground-truth positions of pedestrians in Frame 1830 of ETH and Frame 360 of UNIV-UCY.

2. RELATED WORK

There is a large body of literature on human trajectory prediction. We present a review of human path prediction involving human-human and human-space interactions. While the former only models mutual interaction between pedestrians; the latter, takes into account interactions with surrounding elements: obstacles (e.g. benches, buildings, trees), and walkable area (e.g. roads, sidewalks).

2.1 Human – Human Interactions

The Social Force model introduced by Helbing and Molnar [10] uses hand crafted functions to form coupled *Langevin Equations* which describe social interaction among pedestrians. Several other methods employed hand crafted features [11, 12, and 13], however the prediction becomes much more difficult due to the loss of information. Recent works based on LSTM models rely on social pooling mechanism proposed in [3], which allows pedestrians to share their hidden representations. It merges hidden states of pedestrians in immediate neighborhood of each other to make each trajectory *aware* of its neighborhood. Even so, pedestrians are still unaware of elements such as trees or benches, which might influence them to change directions in the absence of other humans. The inherent multi-modal nature of trajectory forecasting task is used in [5] to differentiate between multiple plausible paths using GAN (Generative Adversarial Network). The model fails to discriminate between structured and unstructured environments, as it only considers relative positions between two pedestrians. In [14], the author proposes a model to detect groups of people moving in the same direction. Vemula *et al.* [15] use spatio-temporal graph which models position evolution and mutual interaction to handle prediction. Hasan *et al.* [4] embodies *vislet* information within the social-pooling mechanism also relying on mutual faces orientation to augment space perception.

2.2 Human - Space Interactions

Sadeghian *et al.* [16] use GAN to take into account both past crossed areas and semantic context to predict social and context-aware positions. Ballan *et al.* [6] propose a Bayesian framework based on previously observed motions to infer unobserved paths. They use it for transferring learned motions to unobserved scenes. Likewise, circular distributions model dynamics and semantics for long-term trajectory predictions are introduced by Coscia *et al.* in [17]. The work presented by Sadeghian *et al.* [16] is based on based on a two-level attention mechanism. It uses bird's eye view images together with past observations. The work mainly focuses on scene cues partially addressing agents' interactions. Bartoli *et al.* [18] proposed attractions towards static objects (e.g., portraits, paintings and other artworks) which deflect straight paths in various contexts (e.g., art galleries, museums) but the approach is severely restricted in application.

There are other approaches in addition to the ones discussed above. [19] uses Inverse Reinforcement Learning (IRL) to focus on transfer learning for pedestrian motion at intersections. [20] attains best performance on the challenging Stanford Drone Dataset (SDD). Lisotto *et al.* [9] propose three pooling mechanisms, namely Social, Navigation and Semantic. Social pooling mechanism takes into account the neighborhood in terms of other people, merging their hidden state. Navigation pooling mechanism exploits past observations to discriminate between equally likely predicted positions using previous information about the scene. Finally, Semantic pooling uses semantic scene segmentation to recognize not crossable areas.

3. PROPOSED MODEL

Different factors are associated in guiding pedestrians towards their destinations in urban scenarios. These factors influence pedestrian dynamics. For example, sidewalks and marked zebra crossings are more likely to be used by pedestrians than grass or roads. Similarly, fast walkers tend to swiftly turn around benches. Moreover, people generally don't enter a building through window. They use an authorized gate or door. Therefore, while forecasting paths, it is crucial to consider human dynamics with neighboring pedestrians, neighborhood semantics and past observations from experience. To this aim, we extend the Social-LSTM model [3] and Social and Scene Aware-LSTM model [9], as schematized in Figure 2. More specifically, for every i^{th} pedestrian, trajectory in a 2D sequence $T_i =$

$(x_1^i, y_1^i), (x_2^i, y_2^i) \dots (x_8^i, y_8^i)$ = observed time steps

and

$(x_9^i, y_9^i), (x_{10}^i, y_{10}^i) \dots (x_{20}^i, y_{20}^i)$ = predicted time steps

The social, navigation and semantic tensors are embedded into three vectors, $a_i t$, $n_i t$, $s_i t$. In addition, the spatial coordinates are represented by $e_i t$. In LSTM cell,

$e_i^t = \Phi(x_i^t, y_i^t; W_e)$

$a_i^t = \Phi(H_i^t; W_a)$

$n_i^t = \Phi(N_i^t; W_n)$

$s_i^t = \Phi(S_i^t; W_s)$

$g_i^t = \Phi(concat(a_i^t, n_i^t, s_i^t); W_g)$

$h_i^t = LSTM(h_i^{t-1}, concat(e_i^t, g_i^t); W_h)$

where Φ represents the ReLU activation function and W_h are the LSTM weights.

Finally, by minimizing the parameters of the network, we obtain negative log-Likelihood loss L_i for the ith pedestrian.

4. EXPERIMENTS

4.1 Implementation Details

The model is based on TensorFlow library [21], and is trained using a single GPU. The number of hidden units for each LSTM cell are 128, while 64 embedding dimensions of spatial coordinates are used. The learning rate is set to 0.003 and RMS-prop is used as optimizer with a decay of 0.95. The model is trained for 50 epochs, and N_o; N_n; and N_s are set to 8, 32 and 20, respectively. A leave-one-out-cross-validation approach has been used, i.e., the models are trained on N-1 scenes and tested on the remaining one. The results are averaged over the two datasets.

4.2 Datasets

We train and test our model on two datasets: ETH [22] and UNIV-UCY [23]. They are captured using drone from a bird's-eye view and involve numerous challenging situations, such as interacting pedestrians, and highly nonlinear trajectories.
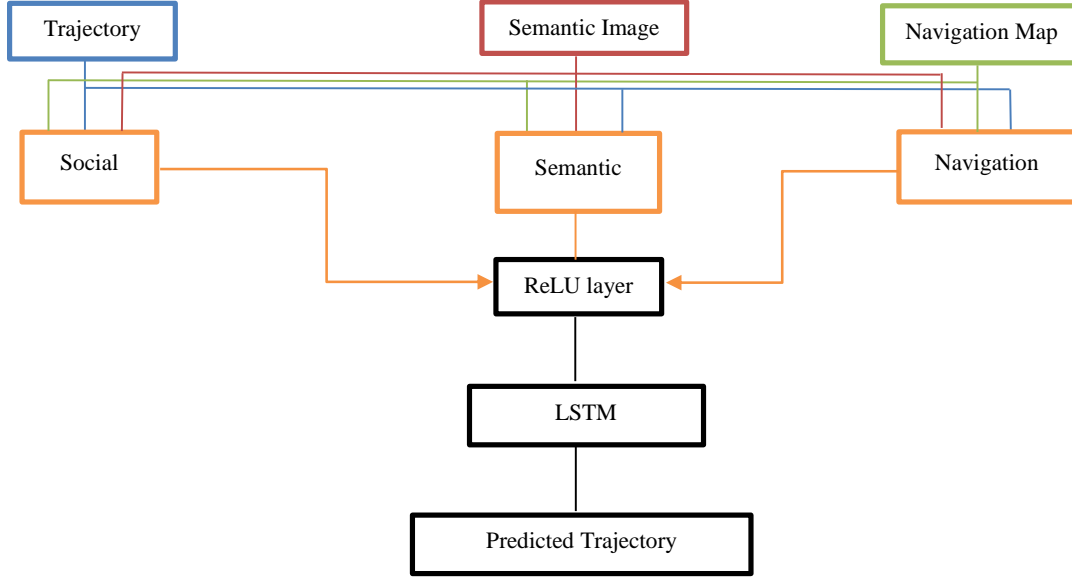

Figure 2. Overview of the model. Trajectories, semantic image and navigation map are combined using social, semantic and navigation pooling mechanisms.

Table 1. Quantitative results of the architecture. The errors are reported in meters.

Metric	Scene	Vanilla [3]	S-LSTM [3]	SS-LSTM [9, 24]	SN-LSTM [9]	SNS-LSTM [9]	SNS-LSTM [Proposed]
Average Displacement Error	ETH	0.51	0.52	0.48	0.47	0.58	0.38
	UNIV-UCY	0.55	0.52	0.43	0.39	0.37	0.32
	Average	0.53±0.13	0.52 ± 0.11	0.45 ± 0.10	0.43 ± 0.09	0.47	**0.35 ± 0.09**
Final Displacement Error	ETH	2.82	2.84	2.57	1.8	2.43	2.40
	UNIV-UCY	3.04	2.92	2.54	2.1	2.08	2.05
	Average	2.93 ± 0.71	2.88 ± 0.6	2.55 ± 0.54	**1.95 ± 0.3**	2.25	2.22 ± 0.43

4.3 Ablation Study

We also studied the ablation versions of this method. The effect of navigation and semantic factors were tested individually, as well as in combination with social interactions, on two different models, namely SN-LSTM (Social-Navigation-LSTM) and SS-LSTM (Social-Semantic-LSTM), respectively. Thereafter, they were combined in SNS-LSTM (Social-Navigation-Semantic) model. As proposed in [22] and [9], a pedestrian trajectory is observed for 3.2s in order to predict the next 4.8s. At frame level, the network is trained on first 8 frames and predicted for the next 12 frames.

Finally, to reduce the error and increase accuracy of prediction, we consider the epsilon value as 1e-25 for numerical stability.

4.4 Evaluations

As error metrics, we report the Average Displacement Error (ADE) and the Final Displacement Error (FDE). ADE represents the average Euclidean distance between the predicted and ground-truth positions, while FDE consists in the average Euclidean distance between the final predicted and ground-truth position.

5. RESULTS

5.1 Quantitative Results

The results for ETH and UNIV/UCY are shown in Table 1.

Comparison with existing models:

- S-LSTM [3]: We extract every 10th frame consisting of one or multiple pedestrians per time step.
- Vanilla LSTM [3]: Realistic paths are not predicted due to absence of any kind of interactions.
- SS-LSTM [24, 9]: Uses neighborhood interactions and scene information.
- SN-LSTM [9]: Uses past observations and neighborhood interactions.

We report on these methods to predict human behaviors and movements. The worst model that takes no internal/external factor into account, is namely Vanilla-LSTM. It is observed that the presence of discriminative regions improves the performance of SN-LSTM model considerably, in comparison to S-LSTM model when two or more path are plausible. In addition, introduction of semantics with SS-LSTM improves the model performance for ETH scene. Compared to other datasets, navigation map enables better predicted trajectories on the ETH dataset.

For the UNIV-UCY, the effect of social-navigation pooling mechanism reduces the error metrics by 45.6%. Therefore, to make more accurate predictions, the combination of proposed factor seems to reinforce the value of integrating navigation and semantic factors. Figure 3 shows the ADE as a function of ranging the amount of training data from 0% to 50%. We observe that the ADEs are significantly reduced as the amount of training data for both datasets increases. In comparison with [25], our model of SNS-LSTM performs significantly better.
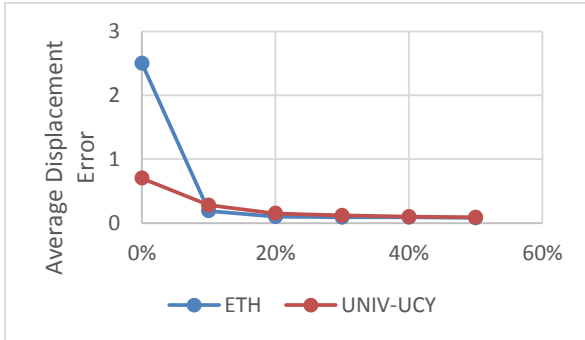
Figure 3. The impact of the amount of training data on Average Displacement Error (ADE) for ETH and UNIV-UCY video sequences using all-grid model in SNS-LSTM.

5.2 Qualitative Results

In this section, we first present the static comparisons between different models and baselines, and then visualize multiple predicted trajectories of both datasets over time at consecutive time steps.

5.2.1 Static Visualizations

Figures 4 and 5 show examples of predicted trajectories from ETH and UNIV-UCY datasets respectively. The figures illustrates two different scenarios: single pedestrian, and multiple pedestrians. For visualization clarity, only ground-truth, S-LSTM and SNS-LSTM models are shown for multiple pedestrians. Of course, this method is not perfect, there are some bad cases in which the predicted trajectories deviate substantially from the ground truth, as shown in the figure.

5.2.2 Dynamic Visualizations

We use temporal evaluations on our dataset. The first 8 frames are visualized as the observed path, followed by the next 12 frames based on prediction models. Figure 6 shows such trajectories for a single pedestrian drawn from S-LSTM, SN-LSTM and SNS-LSTM (proposed) models. The effectiveness of pooling mechanism to avoid deviations and collisions is demonstrated for ETH dataset. The semantic pooling avoids the model's collision with obstacles such as snow.

The results for multiple pedestrians in same time steps are presented in Figure 7.

The temporal visualizations illustrate the efficacy of the elements implemented to better capture complex human behaviors in crowded environments where static objects often contribute to the process of path generation.
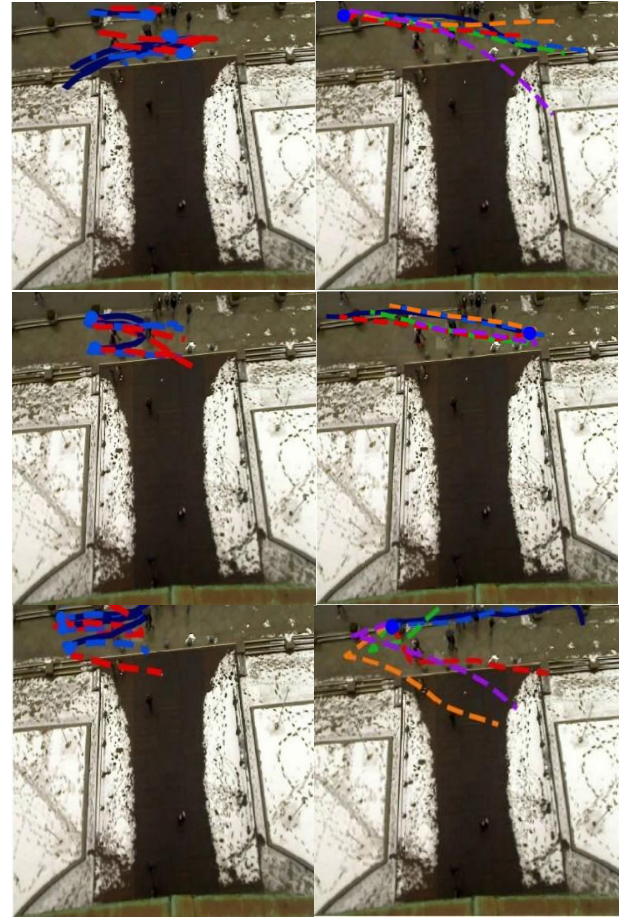

Figure 4. The solid blue lines represent ground-truths, while the predicted paths are shown by dashed-lines in ETH. The first column shows predictions for multiple pedestrians. The second column shows predicted paths for single pedestrian using all models: Vanilla (Violet), S-LSTM (Red), SN-LSTM (Green), SS-LSTM (Orange), and SNS-LSTM (Light blue).
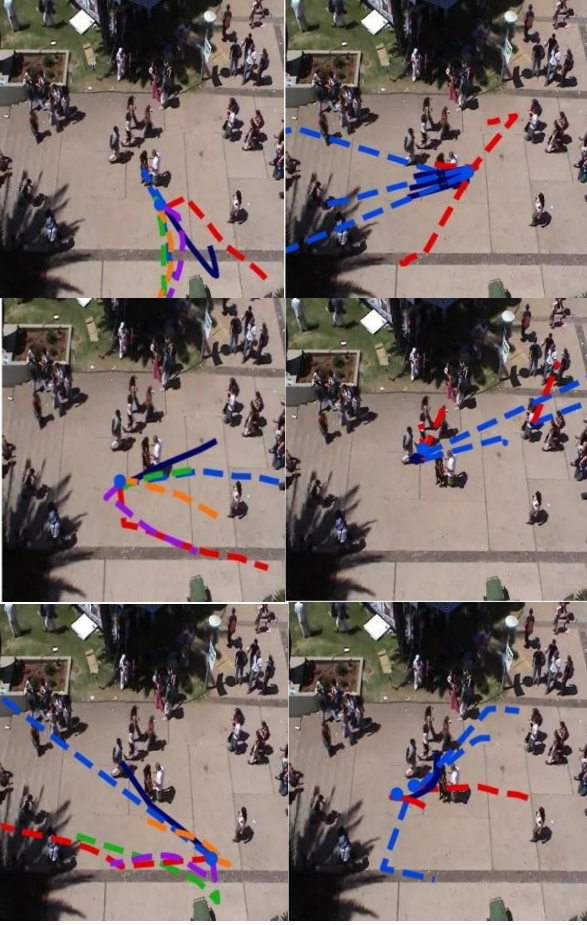
Figure 5. The above figures show predicted paths for UNIV-UCY dataset. The first column shows outputs for all models: Ground-truth (Dark blue), Vanilla (Violet), S-LSTM (Red), SN-LSTM (Green), SS-LSTM (Orange), and SNS-LSTM (Light blue), for a single pedestrian. The second column shows output cases for S-LSTM, SNS-LSTM, and ground-truths for multiple pedestrians.

6. CONCLUSION

We evaluate human-human and human-space interactions for trajectory forecasting in challenging scenarios of ETH and UNIV-UCY datasets. The Navigation and Semantic pooling mechanisms use past observations about the scene and semantics of crossed areas to form SNS-LSTM. Such an approach favors more reliable predicted paths when multiple paths are simultaneously possible to reach desired destinations. Results from this study show better performance than previous state-of-the-art methods.

In addition, our proposed method for error calculation significantly reduces error for common metrics, thereby increasing the accuracy of these models. Our future work will explore various datasets where more nuanced dynamics are captured as well as different entities such as vehicles, skateboarders, or cyclists, normally share the same area. We plan to explore in-depth the effect of semantic segmentation of neighborhood areas in trajectory prediction.
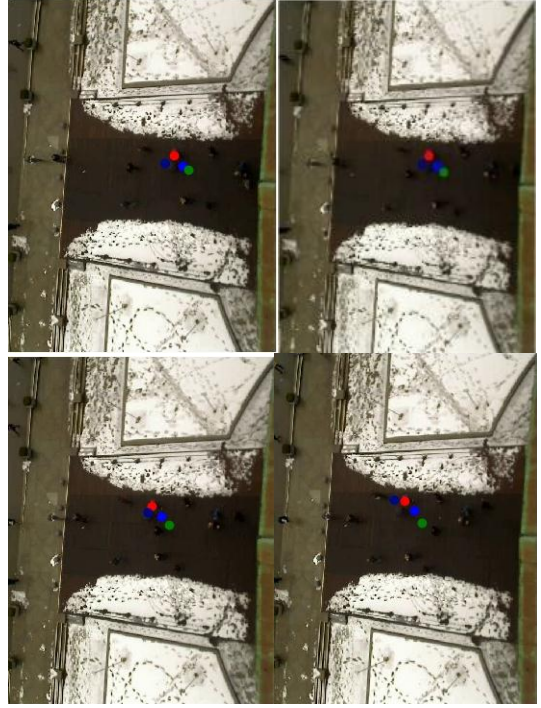
Figure 6. Temporal sequences visualizations drawn from ETH dataset for single pedestrian. The circles represent ground-truth (light blue), S-LSTM (red), SN-LSTM (green) and SNS-LSTM (dark blue). The images correspond to Frame IDs 8980, 8990, 9000, and 9010 in left-to-right manner.
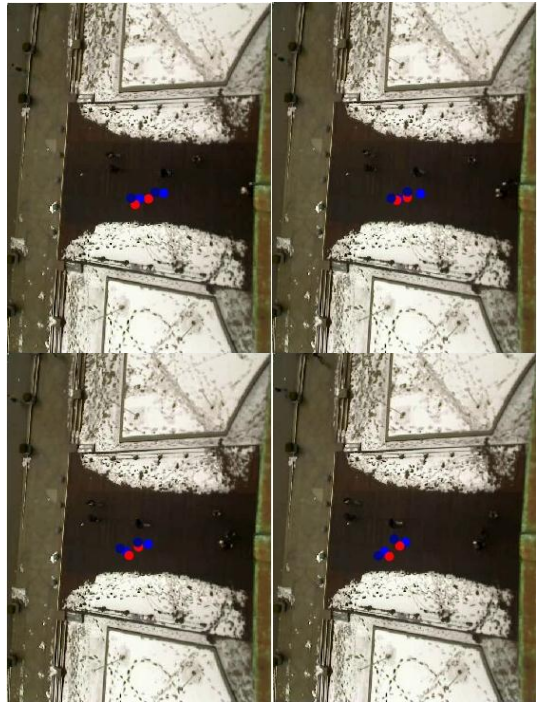
Figure 7. Temporal sequences visualizations drawn from ETH dataset for multiple pedestrians. The circles represent ground-truth (light blue), S-LSTM (red) and SNS-LSTM

(dark blue). The images correspond to Frame IDs 1530, 1540, 1550, and 1560 in left-to-right manner.

7. REFERENCES

[1] Fernando, T., Denman, S., Sridharan, S. and Fookes, C. 2018. Soft Hardwired attention: An LSTM framework for human trajectory prediction and abnormal event detection. *Neural Networks*. 108, (2018), 466–478.

[2] Karpathy, A., Johnson, J. and Fei-Fei, L. 2020. Visualizing and Understanding Recurrent Networks. *arXiv.org*. https://arxiv.org/abs/1506.02078.

[3] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L. and Savarese, S. 2016. Social LSTM: Human Trajectory Prediction in Crowded Spaces. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. (2016).

[4] Hasan, I., Setti, F., Tsesmelis, T., Bue, A.D., Galasso, F. and Cristani, M. 2018. MX-LSTM: Mixing Tracklets and Vislets to Jointly Forecast Trajectories and Head Poses. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2018).

[5] Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S. and Alahi, A. 2018. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2018).

[6] Ballan, L., Castaldo, F., Alahi, A., Palmieri, F. and Savarese, S. 2016. Knowledge Transfer for Scene-Specific Motion Prediction. *Computer Vision – ECCV 2016 Lecture Notes in Computer Science*. (2016), 697–713

[7] Kitani, K., Ziebart, B., Bagnell, J. and Herbert, M. 2020. *Ri.cmu.edu*. http://www.ri.cmu.edu/pub_files/2012/10/Kitani-ECCV2012.pdf.

[8] Kooij, J.F.P., Schneider, N., Flohr, F. and Gavrila, D.M. 2014. Context-Based Pedestrian Path Prediction. *Computer Vision – ECCV 2014 Lecture Notes in Computer Science*. (2014), 618–633.

[9] Lisotto, M., Coscia, P. and Ballan, L. 2020. Social and Scene-Aware Trajectory Prediction in Crowded Spaces. *arXiv.org*. https://arxiv.org/abs/1909.08840.

[10] Helbing, D. and Molnár, P. 1995. Social force model for pedestrian dynamics. *Physical Review E*. 51, 5 (Jan. 1995), 4282–4286.

[11] Liz Murphy and Paul Newman. 2008. Using incomplete online metric maps for topological exploration with the gap navigation tree. *In Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2792–2797.

[12] P. Newman, M. Bosse, and J. Leonard. 2003. *Autonomous feature-based exploration. In 2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, Vol. 1. 1234–1240 vol.1. https://doi.org/10.1109/ROBOT.2003. 1241761

[13] F. H. Wullschleger, K. O. Arras, and S. J. Vestli. 1999. A flexible exploration framework for map building. In (*Eurobot '99) 1999 Third European Workshop on Advanced Mobile Robots, 1999*. 49–56. https://doi.org/10.1109/EURBOT.1999. 827621

[14] Bisagno, N., Zhang, B. and Conci, N. 2019. Group LSTM: Group Trajectory Prediction in Crowded Scenarios. *Lecture Notes in Computer Science Computer Vision – ECCV 2018 Workshops*. (2019), 213–225.

[15] Vemula, A., Muelling, K. and Oh, J. 2018. Social Attention: Modeling Attention in Human Crowds. *2018 IEEE International Conference on Robotics and Automation (ICRA)*. (2018).

[16] Sadeghian, A., Legros, F., Voisin, M., Vesel, R., Alahi, A. and Savarese, S. 2018. CAR-Net: Clairvoyant Attentive Recurrent Network. *Computer Vision – ECCV 2018 Lecture Notes in Computer Science*. (2018), 162–180.

[17] Coscia, P., Castaldo, F., Palmieri, F.A., Alahi, A., Savarese, S. and Ballan, L. 2018. Long-term path prediction in urban scenarios using circular distributions. *Image and Vision Computing*. 69, (2018), 81–91.

[18] Bartoli, F., Lisanti, G., Ballan, L. and Bimbo, A.D. 2018. Context-Aware Trajectory Prediction. *2018 24th International Conference on Pattern Recognition (ICPR)*. (2018).

[19] Sadeghian, A., Kosaraju, V., Sadeghian, A., Hirose, N., Rezatofighi, S. and Savarese, S. 2020. SoPhie: An Attentive GAN for Predicting Paths Compliant to Social and Physical Constraints. *arXiv.org*. https://arxiv.org/abs/1806.01482.

[20] Becker, S., Hug, R., Hübner, W. and Arens, M. 2019. RED: A Simple but Effective Baseline Predictor for the TrajNet Benchmark. *Lecture Notes in Computer Science Computer Vision – ECCV 2018 Workshops*. (2019), 138–153.

[21] TensorFlow: A System for Large-Scale Machine Learning: *https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf*. Accessed: 2019-10-30.

[22] Pellegrini, S., Ess, A., Schindler, K. and Gool, L.V. 2009. Youll never walk alone: Modeling social behavior for multi-target tracking. *2009 IEEE 12th International Conference on Computer Vision*. (2009).

[23] Lerner, A., Chrysanthou, Y. and Lischinski, D. 2020. Crowds by Example.

[24] SS-LSTM: A Hierarchical LSTM Model for Pedestrian Trajectory Prediction, In IEEE Winter Conference on Applications of Computer Vision (WACV), pages 1186–1194, 2018.

[25] Manh, H. and Alaghband, G. 2020. Scene-LSTM: A Model for Human Trajectory Prediction. *arXiv.org*. https://arxiv.org/abs/1808.04018.

[26] Robicquet, A., Sadeghian, A., Alahi, A. and Savarese, S. 2016. Learning Social Etiquette: Human Trajectory Understanding In Crowded Scenes. *Computer Vision – ECCV 2016 Lecture Notes in Computer Science*. (2016), 549–565.

[27] Shen, M., Habibi, G. and How, J.P. 2018. Transferable Pedestrian Motion Prediction Models at Intersections. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. (2018).

[28] Zhi, X., He, X. and Schwertfeger, S. 2019. Learning Autonomous Exploration and Mapping with Semantic Vision. *Proceedings of the 2019 International Conference*

on Image, Video and Signal Processing - IVSP 2019. (2019).

[29] Sadeghian, A., Kosaraju, V., Gupta, A., Savarese, S. and Alahi, A. 2020. Trajectory Forecasting. *Trajnet.stanford.edu*. http://trajnet.stanford.edu/.

[30] Xie, D., Shu, T., Todorovic, S. and Zhu, S.-C. 2018. Learning and Inferring "Dark Matter" and Predicting Human Intents and Trajectories in Videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 40, 7 (Jan. 2018), 1639–1652.

A Comparative Study of 6T and 8T SRAM Cell With Improved Read and Write Margins in 130 nm CMOS Technology

2 authors, including:

Ram Murti Rawat
Delhi Technological University

4 PUBLICATIONS   **2** CITATIONS

Some of the authors of this publication are also working on these related projects:

A Novel Low power and Swing Restoration SRAM Logic Circuit Technique View project

A Comparative Study of 6T and 8T SRAM Cell With Improved Read and Write Margins in 130 nm CMOS Technology

RAM MURTI RAWAT, VINOD KUMAR
Department of Computer Science and Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Bawana Road, Delhi-110042
INDIA
rammurtirawat@dtu.ac.in, vinod_k@dtu.ac.in

Abstract: - This paper examines the factors that affect the Static Noise Margin (SNM) of a Static Random Access memories which focus on optimizing Read and Write operation of 8T SRAM cell which is better than 6T SRAM cell Using Swing Restoration for Dual Node Voltage. The read and Write time and improve Stability. New 8T SRAM technique on the circuit or architecture level is required. In this paper Comparative Analysis of 6T and 8T SRAM Cells with Improved Read and Write Margin is done for 130 nm Technology with Cadence Virtuoso schematics Tool.

Key-Words: - SRAM, Swing Restoration Logic, Dual Node Voltage, Low Power, Read and Write margin

1 Introduction

Low power Static Random Access Memories have become a critical component of many VLSI chips. This is especial consideration for microprocessors where the on chip cache sizes are growing with each generation to bridge the increasing divergence in the speeds of the processors and the main memory [P. Barnes 2010, S. Hesley, et.al, 2009]. one of the major issues in the design of an SRAM cell in stability. The cell stability determines the sensitivity of the memory to process tolerances and operating conditions.

The stability of Static Random access memory cell in the presence of DC noise is measured by the static noise margin (SNM). Static Noise Margin is the amount of voltage noise required at the output nodes to flip the state of the cell. This can be obtained using the voltage transfer characteristic (VTC) of the two cross coupled inverters of the SRAM cell [6].

Figure 1 illustrates the schematic of a 6 transistor SRAM cell for simulating the static noise margin. The sources Vn are the noise sources at the state nodes of the cell [6].
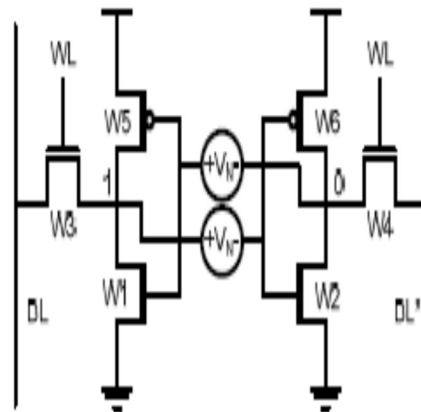
Fig.1 Schematic of a 6T SRAM bit cell with noise voltage sources for measuring SNM [6].

The cross-coupled inverters maintain a bi-stable state and their output nodes retain the data stored in the cell. However, as the noise Vn increases, the stability of the cell degrades because of the fluctuations at the node voltages. The Static Noise Margin quantifies the allowed levels of these noise voltages and thus the ability of these inverters to retain their state in the presence of noise.

The goal of this paper is to determine the effect of several circuit parameters on the SNM of the 6T SRAM cell designed in 180 nm CMOS process technology and compare it with the model derived in [6].

The SNM of the SRAM Cell When in standby or retain mode, read operation, and write operation. The SNM of the SRAM cell is obtained by plotting the VTCs of the Two cross-coupled inverters. The VTC of one of the inverters is flipped with respect to the line y = x in order to form a "butterfly curve". The SNM is the side of the smaller square that can be fitted inside the "eye" of the graph as shown in Figure1 [6].
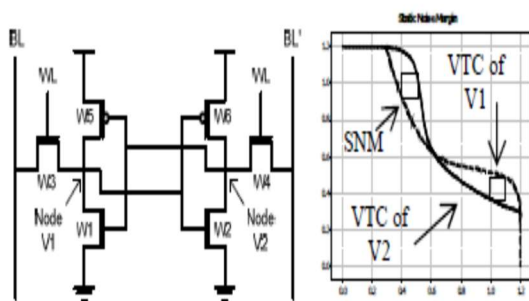
Fig.2 Schematic of a 6T SRAM bit cell and sample SNM-the side of the largest square fitted inside the graph [6].

This Paper is organized as follows: the characteristics of 6T SRAM cell are described are represented in section 2. In section 3, Proposed 8T SRAM cell is described. In section 4, Standard 8T SRAM cell is described. Section 5 includes the simulation results which give comparison of different parameters of 6T and 8T SRAM cells. section 6, DC analysis and section 7 conclusion the work.

2 Six Transistor SRAM Cell

In a conventional 6T SRAM cell, the data storage nodes are directly accessed through the bit-line access transistors during read operations, as shown in Fig.3. While reading, the storage node voltages are disturbed between cross-coupled inverter pair and bit lines. The BL and BLB are the bit lines and WL is the word line. The access transistors are controlled by WL (word line) to perform the operation of read and write operation. Bit lines act as input and output nodes. During a read operation,

bit lines transfer the data from SRAM cells to a sense amplifier. Based on the technology the minimum length of the transistors is 180nm [1].
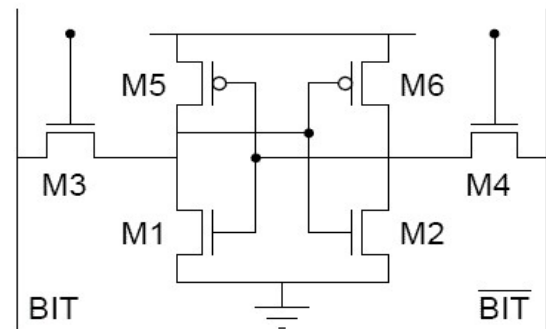
Fig.3 6T SRAM Cell [1].

3 Proposed 8T SRAM Cell

This is proposed SRAM, Dual Voltage with Swing Restoration Logic Perform node voltage in Hold, read and write operation and other parameters like Delay, Stability, are used in Fig.4. Comparison between Low power 6T SRAM and proposed 8T SRAM Cell designs is done. The comparison results revels that read, write and hold mode operation for 8T SRAM cell is better than 6T SRAM cell. This is because higher noise margin are obtained which ensure good write ability for the bit-cell [5].
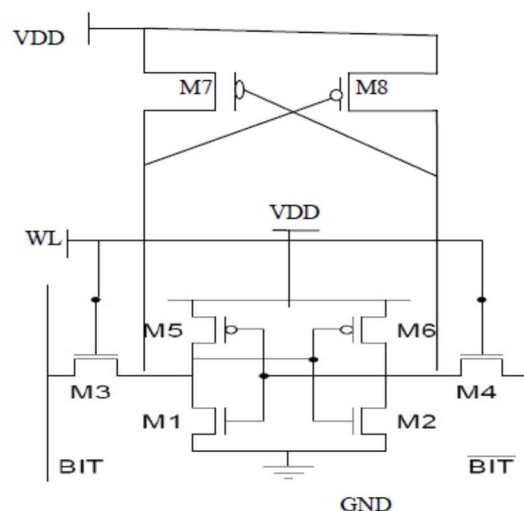

Fig.4 Proposed 8T SRAM Cell [5].

4 Standard 8T SRAM Cell

The standard 8T-SRAM Cell in shown in Fig.5. As it is Seen, read and write cycles use different wordlines and bitlines. Note that the standard 8T SRAM cell uses a single-ended read scheme which reduces the swing of bitlines. The 8T-SRAM cell

provides significantly improved RSNM (similar to Hold Static Noise Margin (HSNM) of the standard 6T-SRAM cell) with similar access time, write time, and write margin [7] [8].
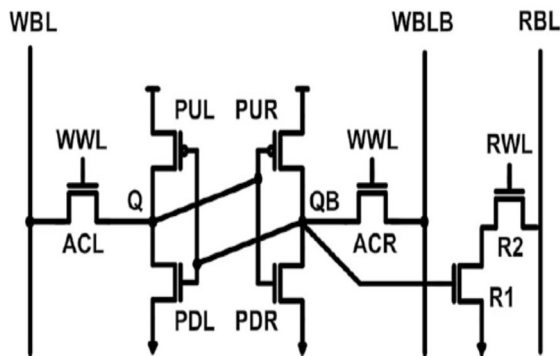

Fig.5 STANDARD 8T-SRAM CELL [7][8].

5 Simulation and Results

Analysis of proposed 8T SRAM cell in terms of write ability, read stability and hold static noise has been carried out in this section. These results are compared with standard 6T and 8T SRAM cell. The circuit is characterized by using the 130 nm Technology with the supply voltage of 1.2 volt.

5.1 Hold Stability

Static noise margin (SNM) is the most common approach to measure hold stability and read stability of the cell. Hold stability is calculated when the SRAM cell is in hold state.

In hold state the word lines are off, so the cell is totally disconnected from the bit lines. SNM defines the largest noise that can be imposed to the storage nodes before flipping the content of the cell. Fig.6, 7 and 8 shows the hold static noise margin of 6T, Proposed 8T and standard 8T SRAM cells respectively.

5.2 Read Stability

The Read stability is measured by read static noise margin (RSNM) in SRAM Cell. In the proposed 8T SRAM cell due to storing nodes isolation we get better RSNM comparable to conventional 6T SRAM and standard 8T SRAM Cells. Fig. 9, 10 and 11 represents the read stability of 6T SRAM and proposed 8T and standard 8T SRAM cells respectively.

5.3 Write Stability

The Write stability is measured by write static noise margin (WSNM). In the proposed 8T SRAM cell

due to storing nodes isolation we get better WSNM comparable to conventional 6T SRAM and standard 8T SRAM Cells. Fig. 12, 13 and 14 represents the write stability of 6T SRAM and proposed 8T and standard 8T SRAM cells respectively.
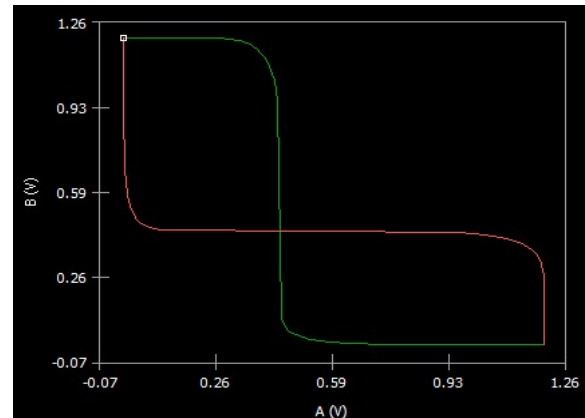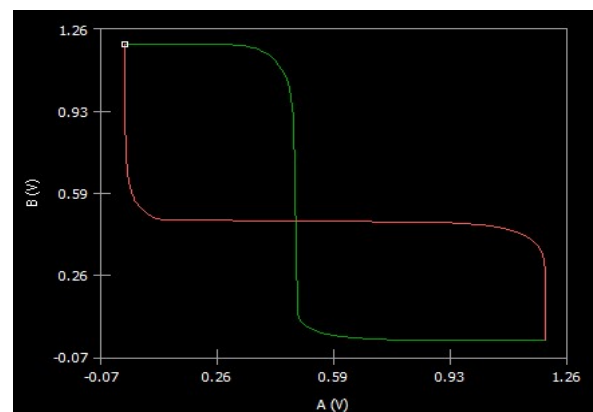

Fig.6 HSNM 6T SRAM CELL

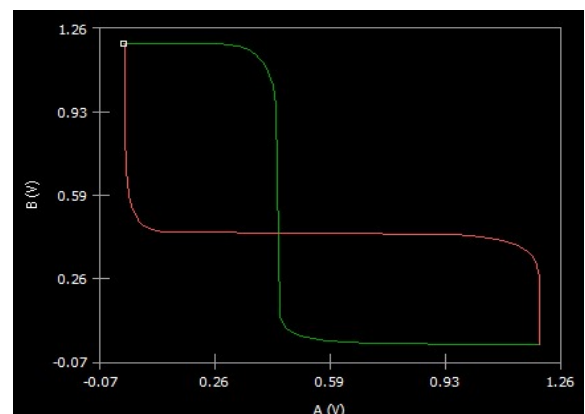Fig.7 HSNM PROPOSED 8T SRAM CELL
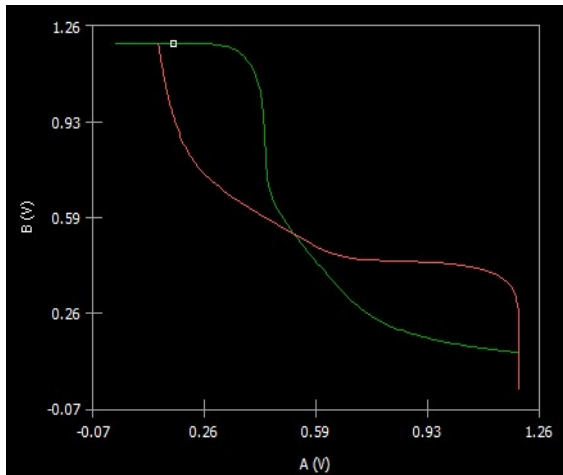

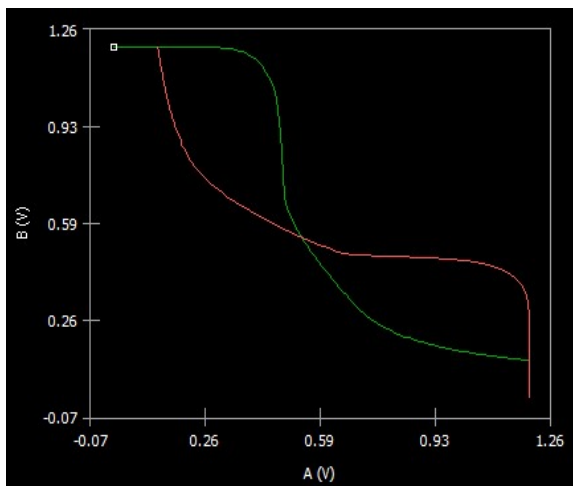Fig.8 HSNM STANDARD 8T SRAM CELL

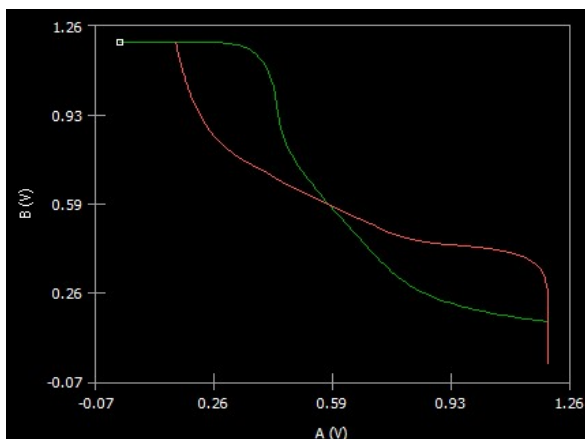Fig.9 RSNM 6T SRAM CELL

Fig.10 RSNM PROPOSED 8T SRAM CELL

Fig.11 RSNM STANDARD 8T SRAM CELL

5.3.1 Write Trip Point

Write trip point is the measure of write ability of the cell. It shows how difficult it is to the storing nodes of the cell.

 The bit-line voltage is swept from 0 to Vdd, and the flipping of the cell, when Q and Q bar flip their their content is captured. The value of bit-line voltage at the crossing point of internal storage nodes Q and QB bar represents write trip point.
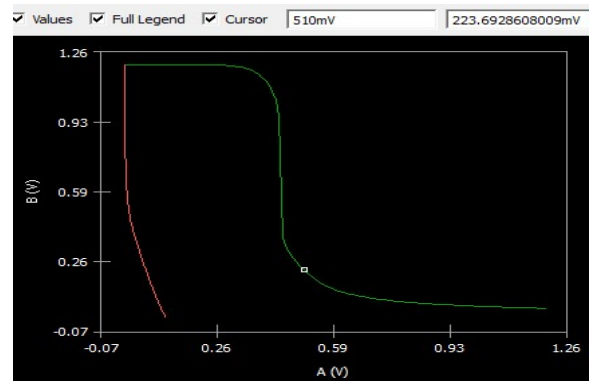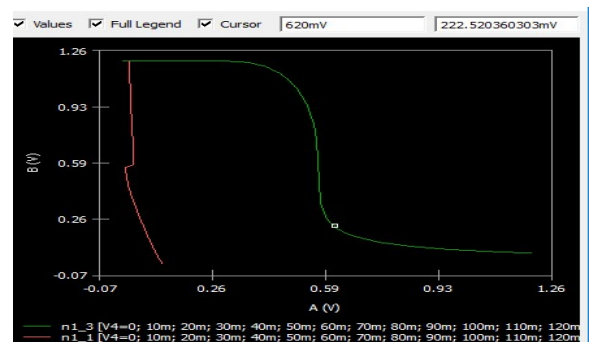

Fig.12 6T SRAM Write Trip Point
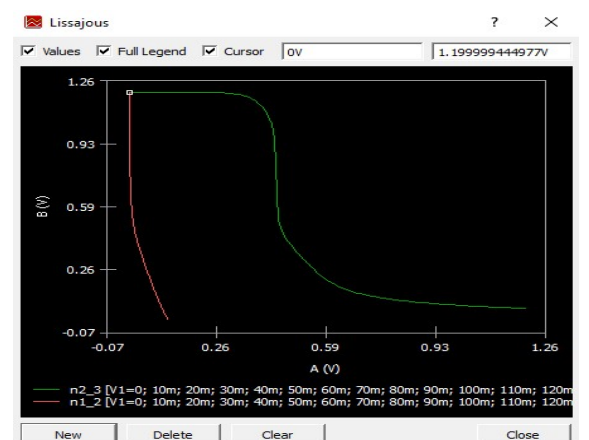

Fig.13 PROPOSED 8T SRAM Write Trip Point

Fig.14 STANDARD 8T SRAM Write Trip Point

6 D.C. Analysis

6.1 Six Transistor SRAM Cell
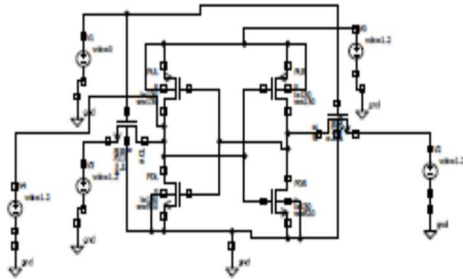
6.1.1 Six Transistor SRAM Cell Output

Fig.15 6T SRAM Cell

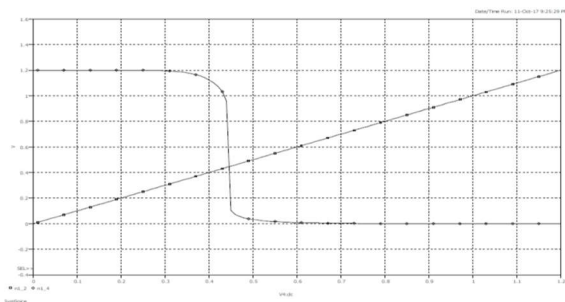### 6.1.2 Six Transistor SRAM Cell DC Analysis

Fig.16 6T SRAM Cell DC Analysis

6.2 Proposed Eight Transistor SRAM Cell

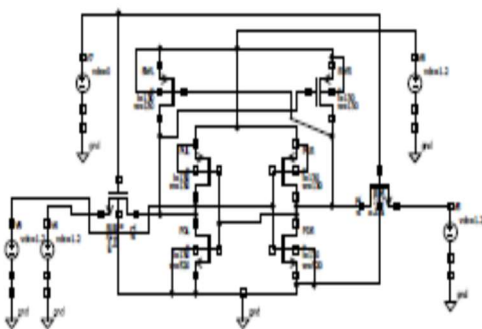### 6.2.1 Proposed Eight Transistor SRAM Cell Output

Fig.17 PROPOSED 8T SRAM Cell

6.2.2 Proposed Eight Transistor SRAM Cell DC Analysis
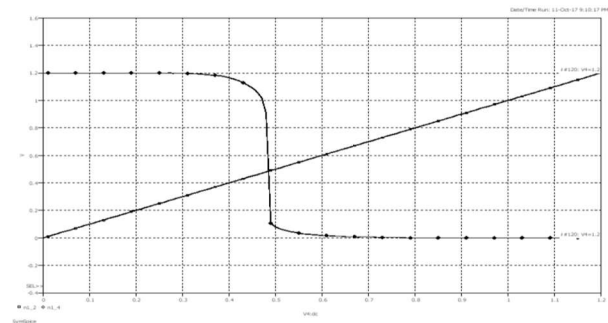

Fig. 18 PROPOSED 8T SRAM CELL DC Analysis

6.3 Standard Eight Transistor SRAM Cell

6.3.1 Standard Eight Transistor SRAM Cell Output
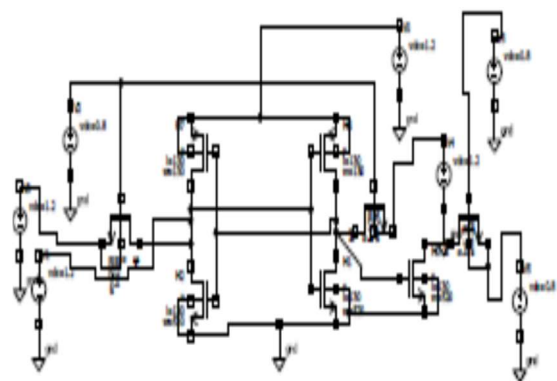

Fig.19 STANDARD 8T SRAM CELL

6.3.2 Standard Eight Transistor SRAM Cell DC Analysis
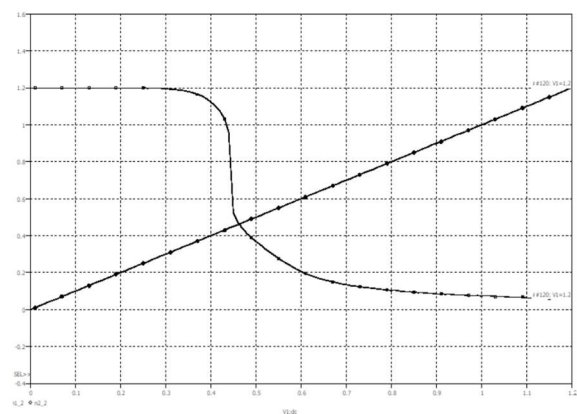

Fig.20 STANDARD 8T SRAM CELL DC Analysis

7 Conclusion

A low power and swing node restoration Static Random Access memory logic circuit technique is presented in the paper. In this paper comparative analysis of 6T and 8T SRAM cells in 130 nm Technology is also presented.

This is the proposed SRAM cell and Dual node voltage with Swing Restoration logic perform D.C. analysis Hold mode operation. D.C. analysis HOLD operation good noise margin proposed 8T SRAM cell is better than 6T SRAM and STANDARD 8T SRAM Cells.

This conclusion is good for power consumption is low. Then PROPOSED 8T SRAM Cell Write mode is power analysis is better Then 6T SRAM and STANDARD 8T SRAM Cells. Then speed is higher for proposed 8T SRAM Cell.

References:

[1] Seevinck, "Static-Noise Margin Analysis of MOS SRAM Cells,"IEEE Journal of Solid-State Circuits, SC-22,5 (Oct. 1987),748-754.

[2] H.L. yeger et al, "Domino circuit Topology", U.S. Patent 6784695, Aug. 31,2004.

[3] A.S. Sedra and K.C. Smith Microelectronic circuits, Fourth Edition, new York: oxford,1988.

[4] S.M. Kang and Leblebici, CMOS Digital Integrated circuits, Analysis and Design, Second Edition, Mc Graw-Hill,1999.

[5] R. M. Rawat, A Novel Low power and Swing Restoration SRAM Logic Circuit Technique, March 2018, International Journal of Engineering and Technical Research V7(02).

[6] A. Pavlov and Manoj sachdev, CMOS SRAM Circuit Design and parametric Test in Nano-scaled Technologies. Springer 2008.

[7] Calhoun, B.H., Chandrakasan, A.P.: Static noise margin variation for sub-threshold SRAM in 65 nm CMOS. IEEE J. Solid-state Circuits 41, 1673-1679 (2006).

[8] Farshad , 8T-SRAM Cell with improved read and write margins in 65 nm CMOS Technology,IFIP AICT 464,PP.95-109,2015.

[9] Sedra/ Smith, Microelectronic Circuit, chapter 15, 6th edition.

[10] Wai-Kai-Chen, Memory, Microprocessor an ASIC, 2003.

[11] Tegze P. Haraszti, CMOS Memory Circuits, Kluwer Publication, 2002.

[12] David Money Harris & Sarah L. Harris, Digital Design and Computer Architecture.

[13] Mohammed Sharifkhani, Design and Analysis of low power Sram, thesis report, Waterloo University, Canada.

[14] Prof. Etienna Sicard and Asst Prof. Sonia Delmas Bendhia, INSA Electronic Engineering School of Toulouse, France, Advanced CMOS Cell Design.

[15] Randall L. Geiger, Phillip E. Allen and Noel R. Strider, VLSI Design Techniques for Analog and Digital Circuits.

[16] R. Jacob Baker, Harry W. Li and David E. Boyce, CMOS Circuit Design, Layout and Simulation.

[17] Neil Weste, CMOS VLSI Design: A Circuit and System Perspective, International Edition.
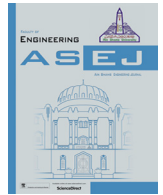
Electrical Engineering

Grid interfaced solar water pumping system with improved space vector modulated direct torque control

Priyanka Singh *, Prerna Gaur

Netaji Subhas Institute of Technology, New Delhi, New Delhi 110078, India

ARTICLE INFO

ABSTRACT

In this paper, a new simplified space vector modulated direct torque control (SVM-DTC) method for induction motor drive is used in solar water pumping system for agricultural and irrigational loads that have smart sharing of power between grid and solar photovoltaic array is proposed. A boost converter with variable step size MPPT that automatically tracks the next step size to maximize the power output is used. Four distinct types of modes with grid operative system and SPV array are proposed such that the unceasing performance of pump is guaranteed and surfeit power is fed back to the grid. The proposed system is modeled in Matlab/Simulink, and the plausibility and novelty of the system in amalgamating the SPV source into the grid can be verified by the results.

© 2020 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Ain Shams University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

As India is a developing country, the provision of water and its accessibility for different locations are different due to many rural areas and uncertain climate conditions. For such cases, solar water pumping system for irrigation and other purposes is a most feasible approach [1–3], as it is more reliable with less maintenance and ease to install.

As the electrical power generated by PV system depends upon weather conditions such as temperature, irradiance, and changes according to them, a maximum power point tracking (MPPT) control is essential for the maximum power extraction from the PV arrays to handle such problems. Many different types of techniques such as fixed duty cycle, perturb & observe (P&O) and incremental conductance (INC) for MPPT have been proposed [4–10,22], and these techniques vary from each other inaccessibility, simple in implementation, cost, time-response, accuracy, easiness. Among all MPPT techniques, incremental conductance (INC-MPPT)

method is considered to be the best for the water pumping system. Therefore, in this paper, a modified INC-MPPT with the variable step size is proposed. Fixed step size with a fixed iteration in INC-MPPT is basically required regulation by precision with continuous speed tracking [10]. Thus, the trade-off in between the dynamics and steady-state oscillations should be satisfactorily inscribed by the relative design. To overcome these problems, INC-MPPT with variable step size is proposed and based on the intrinsic PV array characteristics the step size is automatically tuned. If the operating point is away from the maximum power point, it increases the step size which in turn empowers a rapid tracking ability and if the operating point is near to the maximum power point it minimizes the step size which thus, leads to the reduction in oscillation and contributing to higher efficiency.

An efficient standalone solar water pumping system can be achieved by the use of highly advanced power electronics converters and motor drives [11–18] and [30]. Different types of motors are used for this system, and among all these motors induction motor is low in cost and weight, better reliable with less maintenance required are used widely for irrigation purposes as because of brushes and commutator in DC motors frequent breakdown occurs which leads to poor efficiency. Photovoltaic pumping system with different configuration of induction motor is briefly explained in [11] and [17] with open end winding for dual and multilevel inverter.

Water pumping system requires speed control, there are various schemes of speed control for induction motor drive such as

* Corresponding author.
   E-mail address: siradhna.p@gmail.com (P. Singh).

scalar control (v/f), vector control and direct torque control. [20] and [27] explains the photovoltaic water pumping system with scalar and vector control for induction motor drive. Scalar control is considered to be the simplest among all methods and is widely used but as flux and torque are not directly controlled in this method, the efficiency and precision is not as good as other. Whereas, vector control method provides good efficiency and better precision control it needs high computational power and is very complex due to co-ordinate system and also costly. In 1984 Direct torque control is introduced. In direct torque control (DTC) the torque and speed of three-phase AC electric motors are directly controlled. From the measured values of motor current and voltage the magnetic flux and torque are estimated [25]. In past recent years DTC is also taken in used for brushless DC motor drive and permanent magnet synchronous motor drive for better performance and efficiency [24] and [23]. Performance of DTC method is also improved with advancement of many artificial optimization techniques, space vector modulation and multilevel inverters. In [21] and [18] DTC with fuzzy and sliding mode controllers gives better results compared to conventional DTC, but again it reduces the simplicity of the method.

For photovoltaic water pumping system with induction motor control requires converter for maintaining DC bus voltage constant at inverter side. Various types of converters are introduced and used in recent years for this application such as push pull converter, flyback converter and boost converter [12].

Over the period of time grid interfaced solar PV system is also widely used for many household applications and industrial purposes [14–21] with different control strategies, as the excess power from SPV can be fed to the grid that can further be utilized by other users.

In this paper an efficient hybrid grid integrated solar water pumping system, operated from the grid integrated SPV array system is proposed with modified SVM based DTC drive control, where look up table, identifying angle and sector are not required to reduce the burden on the processor. The ripples are reduced by a substantial amount with the switching patterns to make the system more efficient.

2. Design of the proposed system

Fig. 1 is the basic structure of the smart water pumping system proposed which consists of three phase induction motor drive with improved DTC control in which PWM ac voltage is provided by voltage source inverter. A boost converter with MPPT control for SPV system, a single phase rectifier at grid side and LC filters for reducing the high switching ripples in AC mains are also the part of the pumping system. Kw is $a * \sqrt[3]{P_{pv}}$ where, a is a constant value for generating reference speed from PV array.

In this paper 2.7 kw rated solar PV array is designed to feed 2.2 kw (3 HP) induction motor drive. Tables 1 and 2 given below show the design parameters and specifications of the Solar PV array and also the estimation of the parameters required for the proposed system.

3. Control mechanism for the proposed system

As the pump used in the system is centrifugal in nature, the water discharge is directly proportional to the induction motor drive's speed, and the motor's speed is controlled by the proposed improved SVM-DTC technique. The block diagram of the system in Fig. 1 shows that the DC link voltage is maintained by PI controller that gives the reference current i_s^*. This current is then compared with the system sensed supply current and the differences in currents fed into the hysteresis current controller [28] for generating

the switching pulses for the converter. The proposed MPPT technique for the SPV system and improved SVM-DTC technique for the induction motor drive is discussed in detail in Sections 3.1 and 3.3.

3.1. MPPT control

The INC-MPPT with variable step size is introduced in this paper, in this control according to the intrinsic PV array characteristics the step size is automatically tuned. If the operating point is far from the maximum power point, it increases the step size which in turn enables a fast tracking ability and if the operating point is away from the maximum power point it decreases the step size to very small which thus, leads to the reduction in oscillation and contributing to a higher efficiency giving better output power results as compared to fixed step size INC-MPPT. Fig. 2 shows the flowchart of the modified INC-MPPT with variable step size where P is photovoltaic power, V(n) and I(n) are taken as the PV array output current and voltage at time n and k is the scaling factor [10].

3.2. Different types of modes of smart power sharing

The proposed water pumping system is powered by SPV array and from the grid when the power from the SPV is not sufficient to run the motor. Also, power from SPV array is fed to grid when it is more than the required power to drive the pump. Four different modes of power transfer from both the sources as per need are explained in detail below.

Mode 1: In this mode power generated by SPV array is considered to be sufficient enough to run the pump at rated speed and gives the desired discharge. And if the available insolation is higher and can generate the power higher than the maximum power required for the rated speed of pump, the MPPT is not really made to operate in such case, and also if the generated power is less than the required rated power for the IMD, then in that case from the maximum power the reference frequency is estimated.

Mode 2: During night time when solar power is not available, and also for low radiation case, mode 2 is conducted. In this mode the pump is operated at rated speed and discharge by the single phase grid supply, and voltage source converter (VSC) is used to improve the power quality at AC mains, as VSC helps in mitigating the issues such as unbalanced loads, burden on reactive power, flow of high current, and harmonics in waveforms [29].

Mode 3: As radiation from the sun depends on many factors such as radiation intensity, panel position, and weather conditions that are different for each day, so maximum power available from SPV varies throughout the day in such case pump is operated by both the sources that are from the solar to supply the required discharge for the system and remaining power from the grid supply.

Mode 4: During the time when pump is not used in access or in other case when the power generated from the SPV is much more than the required power for the pump operation then the excess power is fed to the grid.

3.3. Induction motor drive control

There are many techniques used for control of induction motor drive [13–15,23–27]. V/F is the simplest among all but does not give good accuracy in both torque and speed responses. Whereas in Vector Oriented Control (VOC) accuracy is good with fast responses as compared with other techniques but it deals with the co-ordinate system as motor equations are transformed which increases its complexity [27]. Another widely used method is Direct Torque Control method (DTC), in which the torque and speed of three-phase AC electric motors are directly controlled,
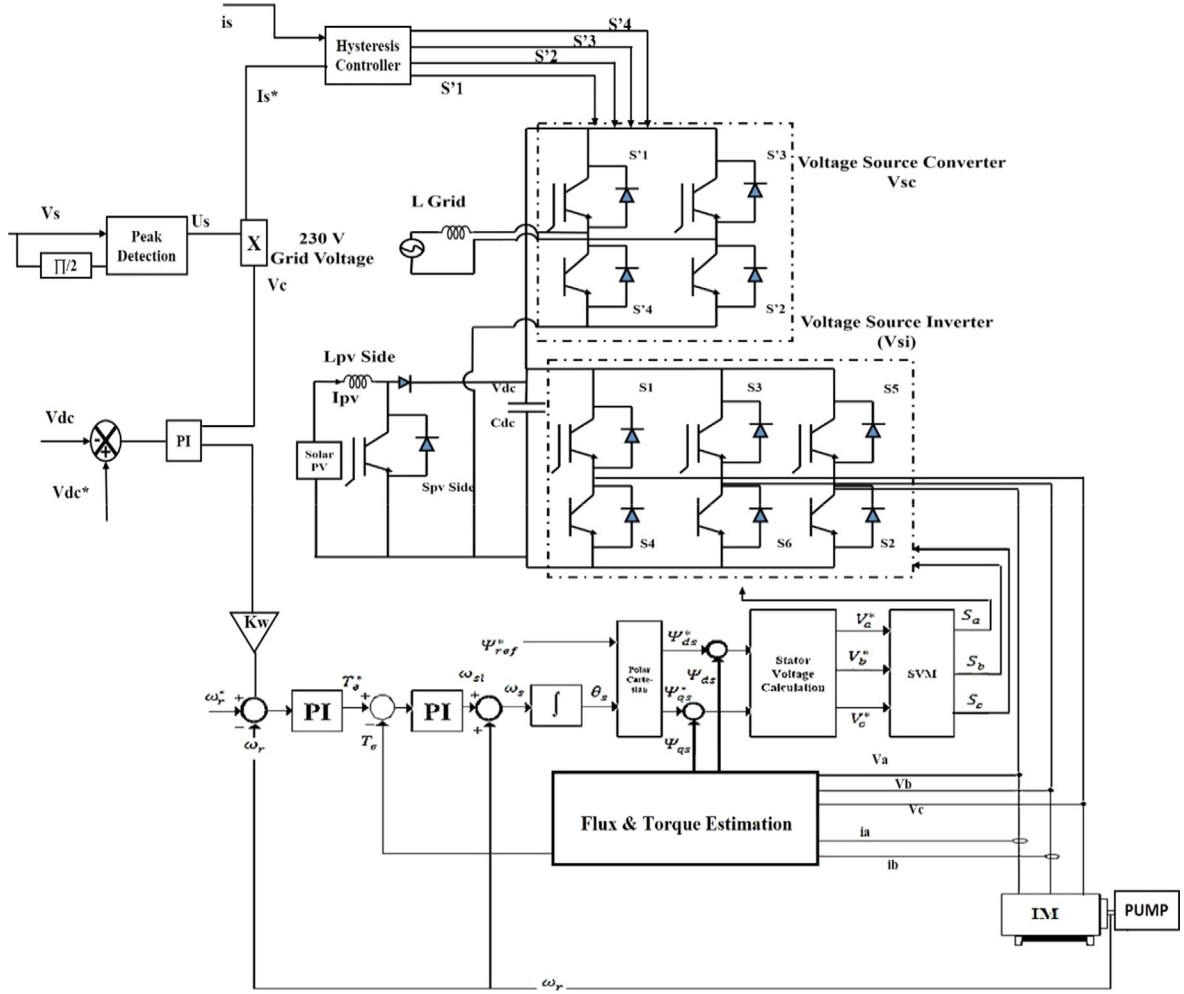
Fig. 1. Block diagram of proposed solar water pumping system.

Table 1
Parameter specification of SPV array.

Parameter of SPV array	Values calculated for SPV array
Open circuit voltage of module V_{OC}	49 V
Short circuit current of module I_{sh}	9.56 A
Maximum power point voltage of module	42 V
Maximum power point current of module	8.13 A
Maximum power of SPV array	2.7 kw
Maximum power point voltage of SPV array V_{MPP}	344 V
Maximum power point current of SPV array I_{MPP}	8.13 A
Number of modules connected in series	8
Number of modules connected in parallel	1

and from the measured values of motor current and voltage, the magnetic flux and torque are estimated.

Fig. 3 shows the block diagram of conventional DTC also known as classical DTC which depends upon the hysteresis controllers for flux and torque control and also needs a voltage vector selection table [21,23]. The torque and flux control of classical DTC can be explained through the following Eqs. (3.1)–(3.5).

$$T_e = \left(\frac{3}{2}\right)\left(\frac{P}{2}\right)\frac{L_m}{\sigma L_s L_r}|\Psi_s||\Psi_r|(\sin(\delta)) \tag{3.1}$$

where, $\sigma = 1 - \frac{L_m^2}{L_s L_r}$ is leakage factor, L_m, L_r, L_s are mutual, rotor and stator inductances, Ψ_s, Ψ_r are stator and rotor flux, δ is the angle between rotor and stator flux vectors. Eq. (3.1) shows that any variation in angle δ leads to change in torque and also as they are directly proportional to each other considering magnitudes of rotor and stator flux vectors constant. And for the flux control the stator flux linkage of the induction motor in the stationary reference frame can be written as:

$$\frac{d\Psi_s}{dt} = V_s - [r_s]\,i_s \tag{3.2}$$

$$\Psi_s = \int V_s - [r_s]\,i_s \tag{3.3}$$

$$\Psi_s = V_s.\Delta t + \Psi_s/_{t=0} \tag{3.4}$$

where, i_s and V_s are the estimated stator current and voltage, $\Psi_{s/t=0}$, is the inceptive stator flux linkage at the instant of switching and r_s, is the measured stator resistance. Ψ_s, tends to in the direction of applied voltage vectors if the stator term is removed from the stator flux estimation.

$$\Delta\Psi_s = V_s\Delta t \tag{3.5}$$

Eq. (3.5) shows that the amplitude of stator flux linkage can be easily controlled by applying a essential voltage vector. Classical DTC is simple in implementing but high ripples in current, flux and torque are its foremost drawback.

Table 2
Specifications of the parameters used in the system.

Components used	Expression	Parameter description	Data Input	Calculated value	Selected value
DC link voltage V_{DC}^*	$\frac{2\sqrt{2}}{\sqrt{3}}V_{LL}$	V_{LL} = line voltage of motor	$\frac{2\sqrt{2}}{\sqrt{3}}*230$	375 V	400 V
DC link voltage C_{DC}	$\frac{6\beta VIT}{[V_{DC}^2-V_{DC'}^2]}$	V_{DC}^* = reference DC bus voltage of voltage source inverter β = overloading factor $V_{DC'}$ = minimum DC link voltage t = time taken to reach allowable DC link voltage	$\frac{6*1.2*133*8.3*0.005}{400^2-375^2}$	2026 μf	2200 μf
Boost converter duty ratio D_{pv}	$\frac{V_{DC}-V_{mp}}{V_{DC}}$	V_{mp} = Maximum power point voltage of PV array V_{DC} = DC link voltage	$\frac{400-344}{400}$	0.14	0.14
Boost converter inductor L_{pv}	$\frac{V_{mp}D}{\Delta I*f_s}$	V_{mp} = Maximum power point voltage of PV array D = duty ratio f_s = switching frequency	$\frac{344*0.14}{0.2*8.13*10.000}$	2.96 mH	2.96 mH
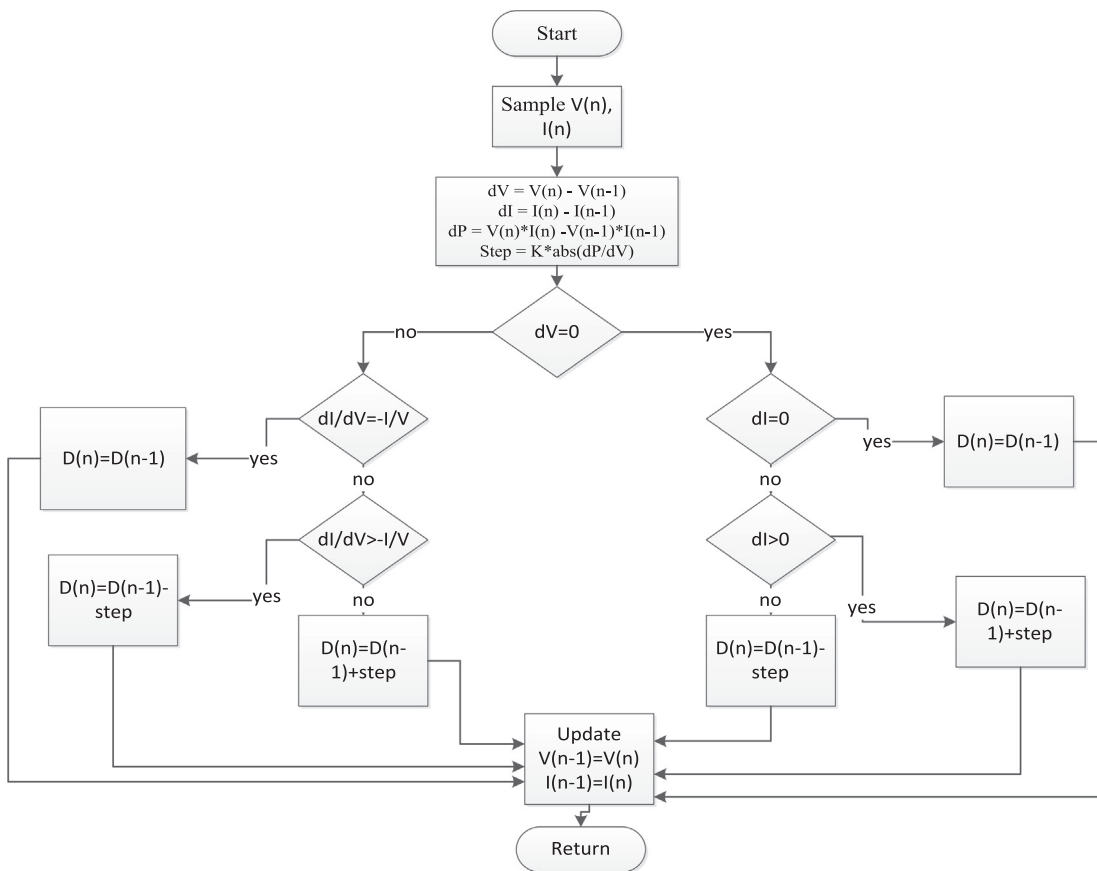
Fig. 2. Flowchart of modified variable step size INC-MPPT.

The main objective of SVM technique is to control the inverter space vectors in order to generate harmonically optimum PWM voltages waves at output. Fig. 4 shows that the modulating command voltages of a 3-ϕ inverter form a rotating space vector V* as it is always sinusoidal in nature. V* can be resolved as follows:

$$V^* \sin(\frac{\pi}{3}-\alpha) = V_a.\sin\frac{\pi}{3} \tag{3.6}$$

$$V^* \sin\alpha = V_b.\sin\frac{\pi}{3} \tag{3.7}$$

$$V_a = \frac{2}{\sqrt{3}}V^* \sin(\frac{\pi}{3}-\alpha) \tag{3.8}$$

$$V_b = \frac{2}{\sqrt{3}}V^* \sin\alpha \tag{3.9}$$

where V_a and V_b are the components of V* oriented in the direction of V_1 and V_2. Considering the sampling period T_{ss} is the considered sampling period for which the command should match average output, it can be depicted as

$$V^* = V_a + V_b = V_1.\frac{T_1}{T_{ss}} + V_2.\frac{T_2}{T_{ss}} + (V_0 or V_7)\frac{T_0}{T_{ss}} \tag{3.10}$$

$$V^*T_{ss} = V_1T_1 + V_2T_2 + (V_0 or V_7).T_0 \tag{3.11}$$

where,

$$T_1 = T_{ss}.\alpha\frac{\sin(\frac{\pi}{2}-\alpha)}{\sin60°} \tag{3.12}$$

$$T_2 = T_{ss}.\alpha\frac{\sin\alpha}{\sin60^0} \tag{3.13}$$
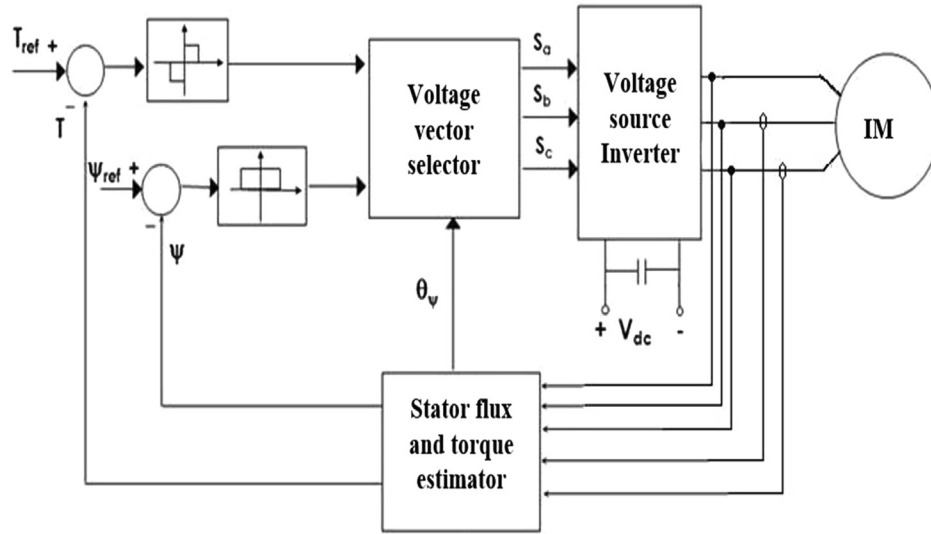
Fig. 3. Classical DTC scheme.
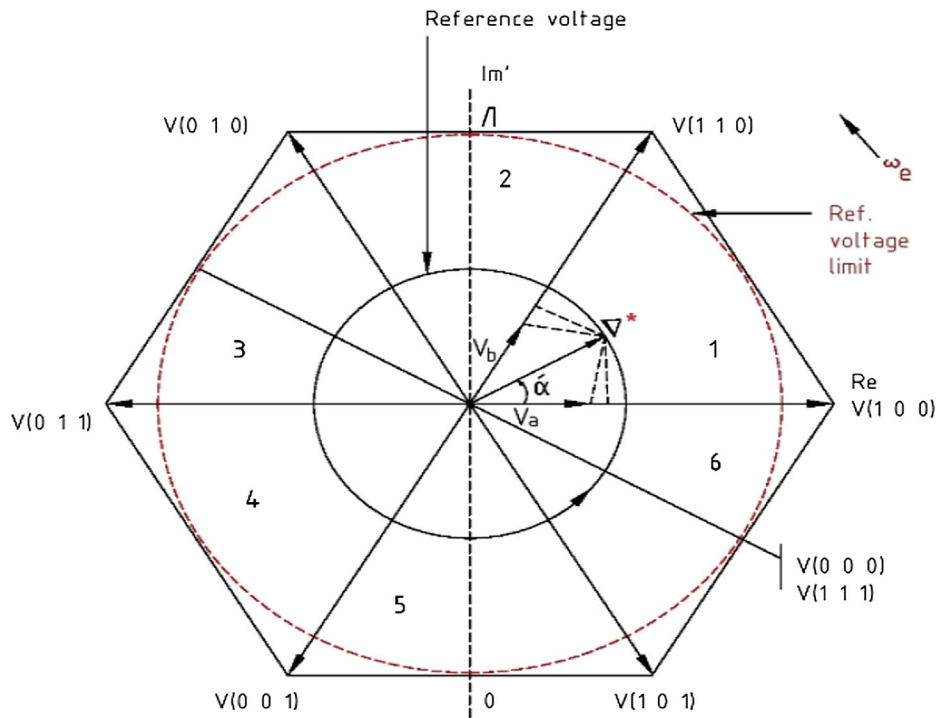
Fig. 4. Space vector of 3-φ inverter.

Here,$\alpha = \frac{V^*}{V_1} = \frac{V^*}{V_2}$, T_{ss} = sampling time,T_1, T_2 and T_0 are the time period for which V_1, V_2 and V_0orV_7 are applied, V_0andV_7 are null vectors, α is d- axisangle of V^* and $V_1 = V_2 = \frac{2}{3}.V_{DC}$

From the above it can be concluded that the conventional implementation of the SVPWM involves identification of sector, transferal of the active vector switching time periods into the inverter leg switching timings, calculation of the active vector switching time period, generation of the gating signals for each devices using the inverter leg switching timing. Therefore, all the required parameters and algorithm for implementing the classical SVM technique is complex and makes the overall process laborious.

A simple scheme is proposed in this paper where look up table, identifying angle and sector is not required by which the burden on the processor is reduced. Also, the ripples are reduced by a substantial amount with the switching patterns and the results are alike to SVM DTC. Simulated results for the same are compared with the classical DTC as shown in Figs. 7–10 and from the results, it is seen that this technique posse much superiority than the classical one.

Fig. 5 shows the SVM based DTC scheme for the induction motor drive for the solar water pumping system. For the improved SVM DTC, difference of V_{DC}^* and DC bus voltage V_{DC} processed in controller gives the reference speed. Electromagnetic torque and
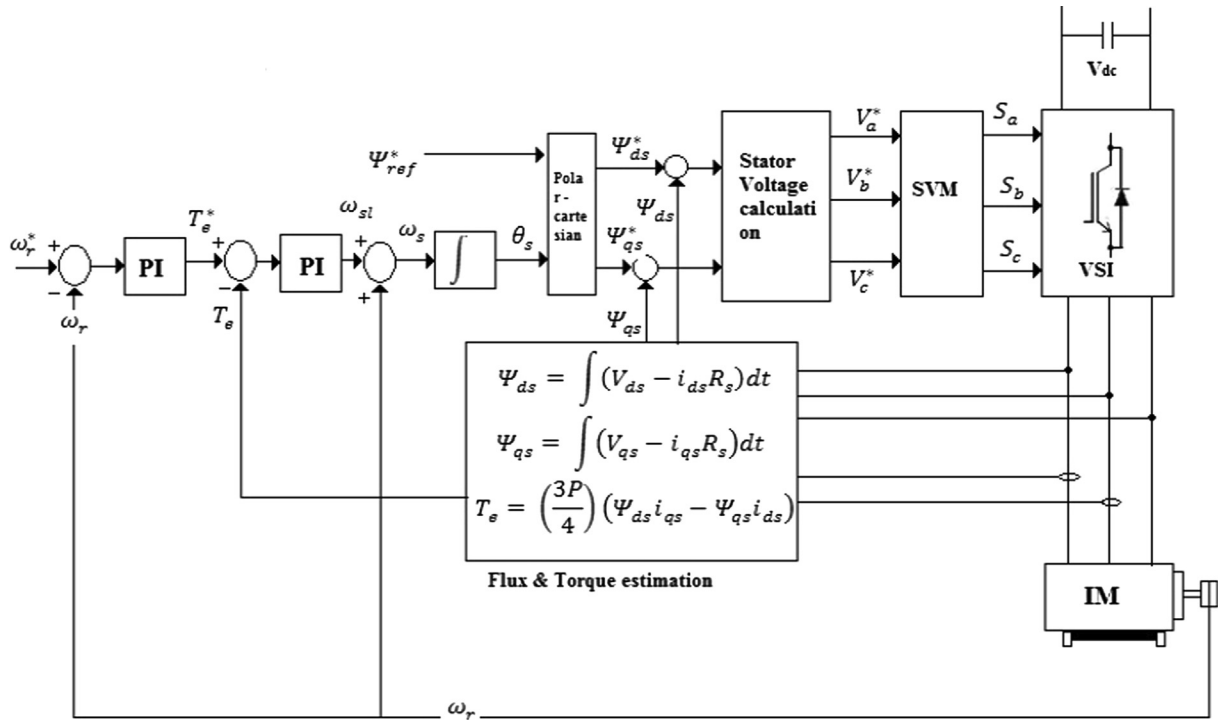
Fig. 5. SVM-DTC scheme.
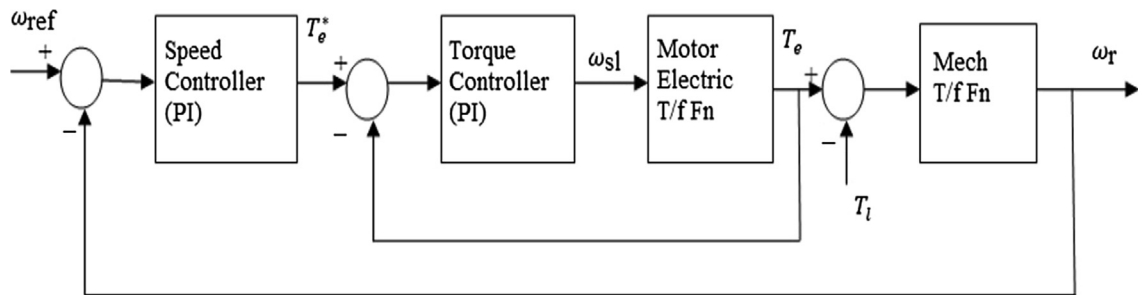
Fig. 6. Speed Loop Block Diagram.
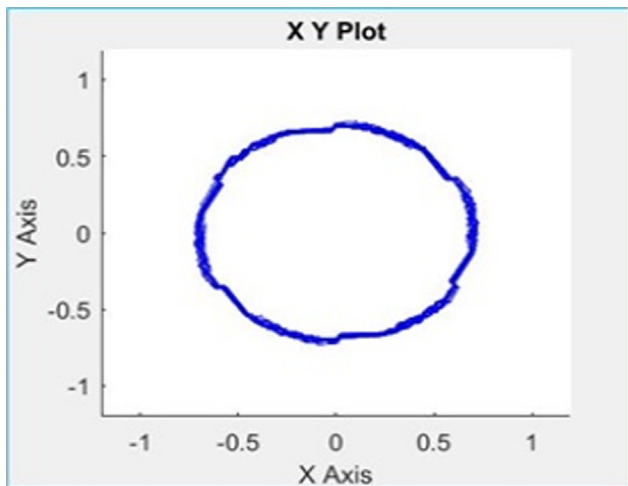
Fig. 7. Flux locus of classical DTC.
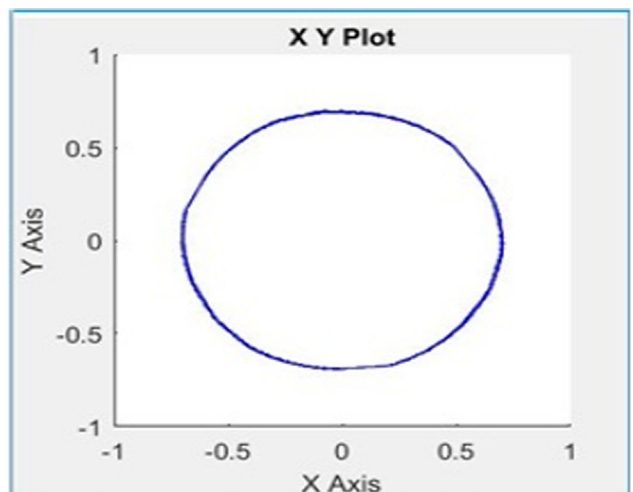
Fig. 8. Flux locus of improved SVM-DTC.
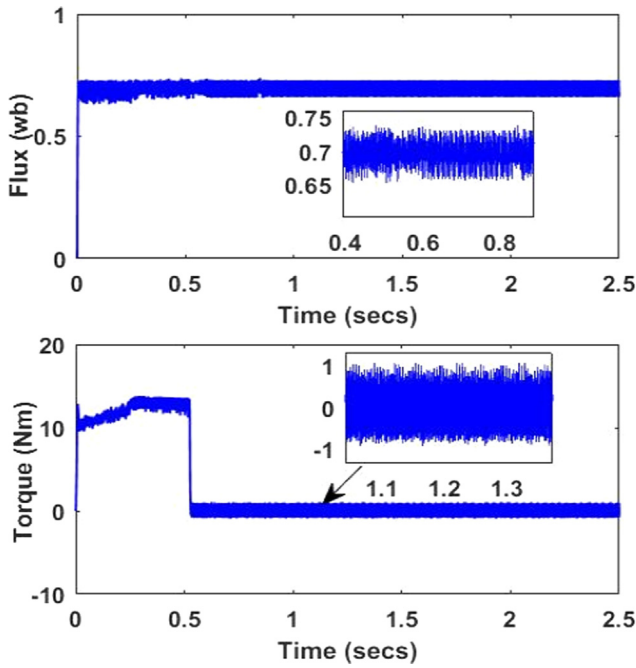
Fig. 9. Estimated torque and flux for classical DTC.

Fig. 10. Estimated torque and flux for SVM-DTC.

Fig. 11. Step change in speed in classical DTC.
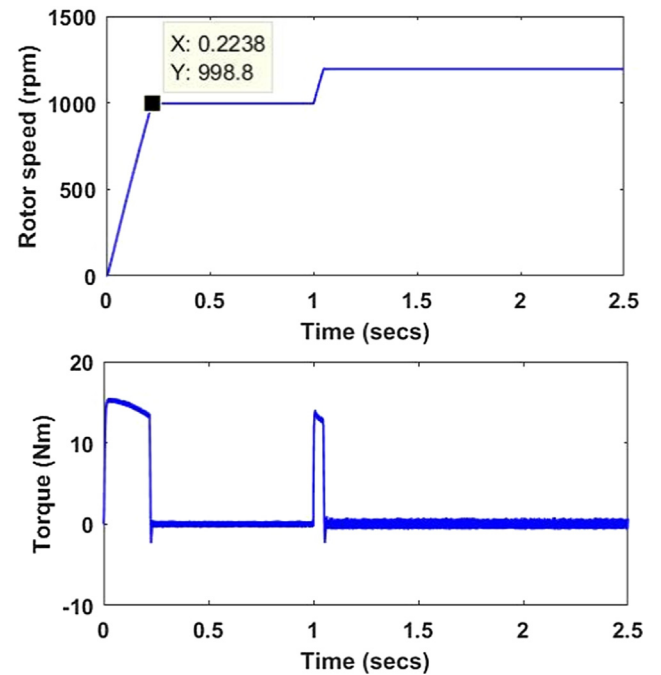
Fig. 12. Step change in speed in SVM-DTC.

flux vectors in d-q axis components are derived from any two phase currents and voltages of the motor and the phase angle between the flux vectors is obtained by integrating the synchronous speed which is sum of the slip speed and the motor speed [25]. Following are the equations for the real switching times of the inverter.

Stator voltage equation in stationary reference frame in d-q is as given

$$V_s = R_s i_s + \frac{d\psi}{dt} \tag{3.14}$$

After neglecting the stator resistance, the equation is simplified as:

$$\Delta\psi_s = V_s.\Delta t \tag{3.15}$$

$$\Delta\psi_{sd} + j.\Delta\psi_{sq} = (V_{sd} + j.\Delta V_{sq})\Delta t \tag{3.16}$$

Taking analogous of the imaginary & real components of Eq. (3.16) results:
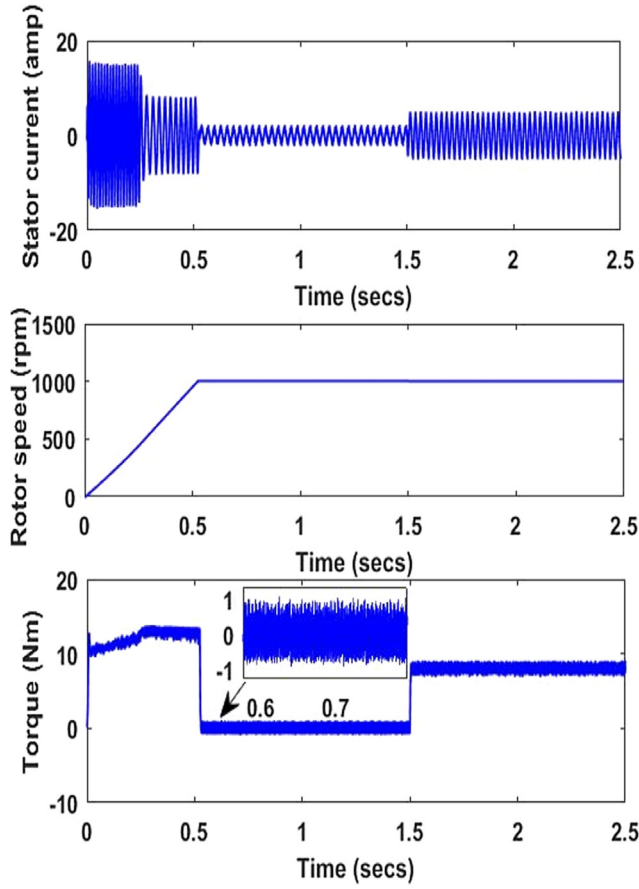
Fig. 13. Torque, speed and current Response of classical-DTC.
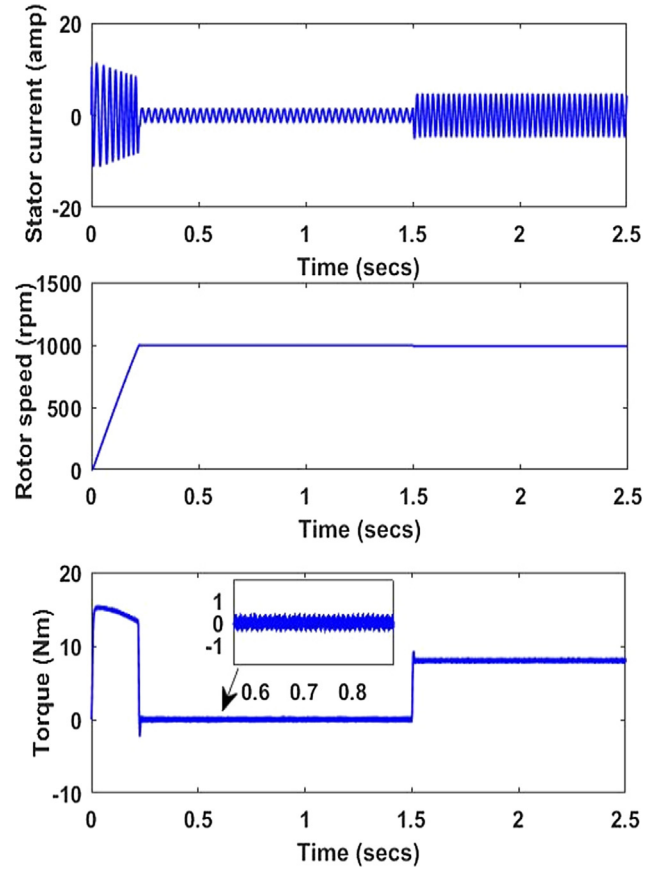
Fig. 14. Torque, speed and current Response of SVM-DTC.

$$V_{sd} = \frac{\Delta \psi_{sd}}{\Delta t} = \frac{\psi_{sd}^* - \psi_{sd}}{\Delta t} \qquad (3.17)$$

$$V_{sq} = \frac{\Delta \psi_{sq}}{\Delta t} = \frac{\psi_{sq}^* - \psi_{sq}}{\Delta t} \qquad (3.18)$$

where, Δt is sampling time T_{s*}.

Thus, the imaginary switching instants in d-q stationary reference frame are derived as following:

$$T_{sd} = \frac{V_{sd}}{V_{DC}}.T_{s*} = \frac{\Delta \psi_{sd}}{\Delta t (= T_{s*}) * V_{DC}} = \frac{\psi_{sd}^* - \psi_{sd}}{V_{DC}} \qquad (3.19)$$

$$T_{sq} = \frac{V_{sq}}{V_{DC}}.T_{s*} = \frac{\Delta \psi_{sq}}{\Delta t (= T_{s*}) * V_{DC}} = \frac{\psi_{sq}^* - \psi_{sq}}{V_{DC}} \qquad (3.20)$$

Also,

$$T_s = T_{sd} + j.T_{sq} = \frac{V_{sd}}{V_{DC}}.T_{s*} + j.\frac{V_{sq}}{V_{DC}}.T_{s*} \qquad (3.21)$$

Hence,

$$T_s = \frac{V_s}{V_{DC}}.T_{s*} = \frac{\psi_s^* - \psi_s}{V_{DC}} \qquad (3.22)$$

From Eq. (3.22) the real switching times of the inverter are calculated from the imaginary switching time vectors. The imaginary switching times T_{sa}, T_{sb} and T_{sc} are obtained from two to three phase conversion of imaginary time vector components [26].

By finding the relationship between the slip speed and the electromagnetic torque generated from Fig. 6, gains of PI controllers can be designed for the proposed SVM-DTC technique.
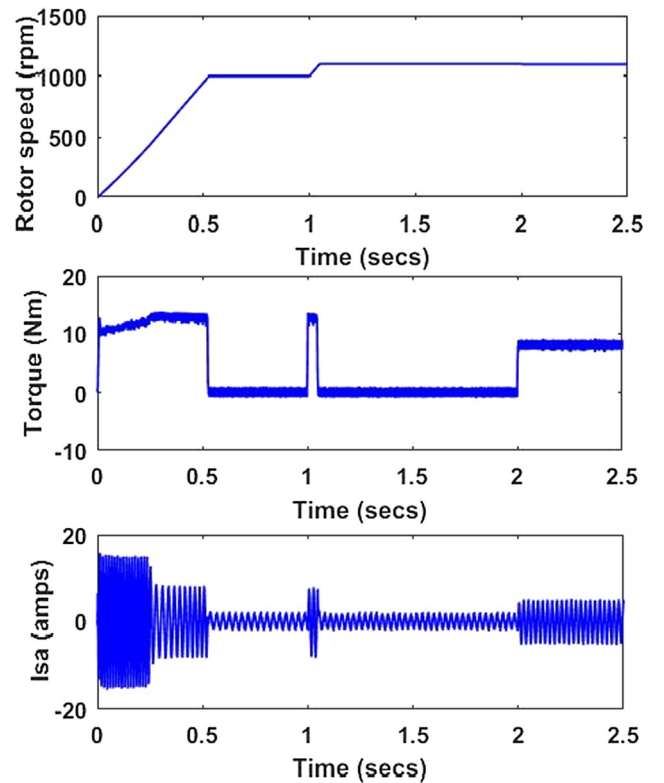
Fig. 15. Step change in speed & torque in classical DTC.
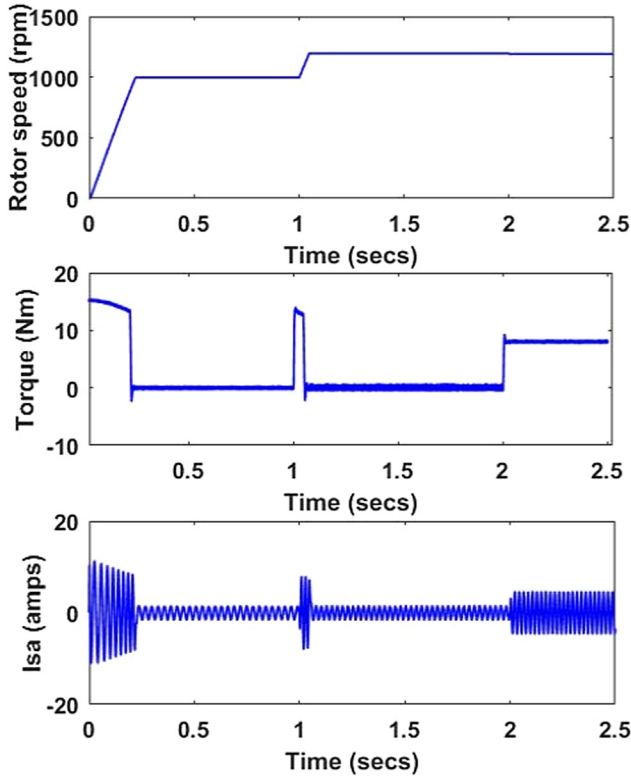
Fig. 16. Step change in speed & torque inSVM-DTC.

Table 3
Comparison table.

Method	Parameter	Settling time(seconds)	Ripple
Classical DTC	Torque	0.5232	1.8Nm
	Flux	0.5236	0.08wb
	Speed	0.5244	
SVM- DTC	Torque	0.2250	0.8Nm
	Flux	0.2247	0.01wb
	Speed	0.2238	

From the stator and rotor flux equations of induction motor [15] the rotor flux can be expressed in terms of the stator flux

$$\overrightarrow{\Psi_r}(s) = \frac{\frac{L_m}{L_s}}{s\sigma\frac{L_r}{R_r} + \left(1 - j\omega_m\sigma\frac{L_r}{R_r}\right)}\overrightarrow{\Psi_s}(s) \tag{3.23}$$

Assuming that $\overrightarrow{\Psi_s} = |\Psi_s|e^{j\theta_s} = |\Psi_s|e^{j\omega_s t}$ and keeping amplitude of the stator flux $\overrightarrow{\Psi_s}$ constant and that $\overrightarrow{\Psi_s}$ rotates at an angular speed ω_s, then the Laplace transform is given by

$$\overrightarrow{\Psi_s}(s) = \frac{|\Psi_s|}{s - j\omega_s} \tag{3.24}$$

By substituting it in the above Eq. (3.23) and taking inverse laplace transform we get the rotor flux as

$$\overrightarrow{\Psi_r}(t) = L^{-1}\left\{\frac{\frac{L_m}{L_s}}{s\sigma\frac{L_r}{R_r} + \left(1 - j\omega_m\sigma\frac{L_r}{R_r}\right)}\frac{|\Psi_s|}{s - j\omega_s}\right\} \tag{3.25}$$

Thus

$$\overrightarrow{\Psi_r}(t) = \frac{L_m}{L_s}\frac{\sqrt{1 + \left(e^{-\frac{t}{\tau}}\right)^2 - 2e^{-\frac{t}{\tau}}\cos\left((\omega_s - \omega_m)t\right)}}{\sqrt{1 + ((\omega_s - \omega_m)\tau)^2}}$$
$$* |\Psi_s|e^{j\left(\tan^{-1}\left(\frac{y}{x}\right) - \tan^{-1}\left((\omega_s - \omega_m)\tau\right)\right)} \tag{3.26}$$

where, $\tau = \sigma\frac{L_r}{R}, x = \cos(\omega_s t) - e^{-\frac{t}{\tau}}\cos(\omega_m t)$

$$x = \sin(\omega_s t) - e^{-\frac{t}{\tau}}\sin(\omega_m t) \tag{3.27}$$

The torque can be expressed as
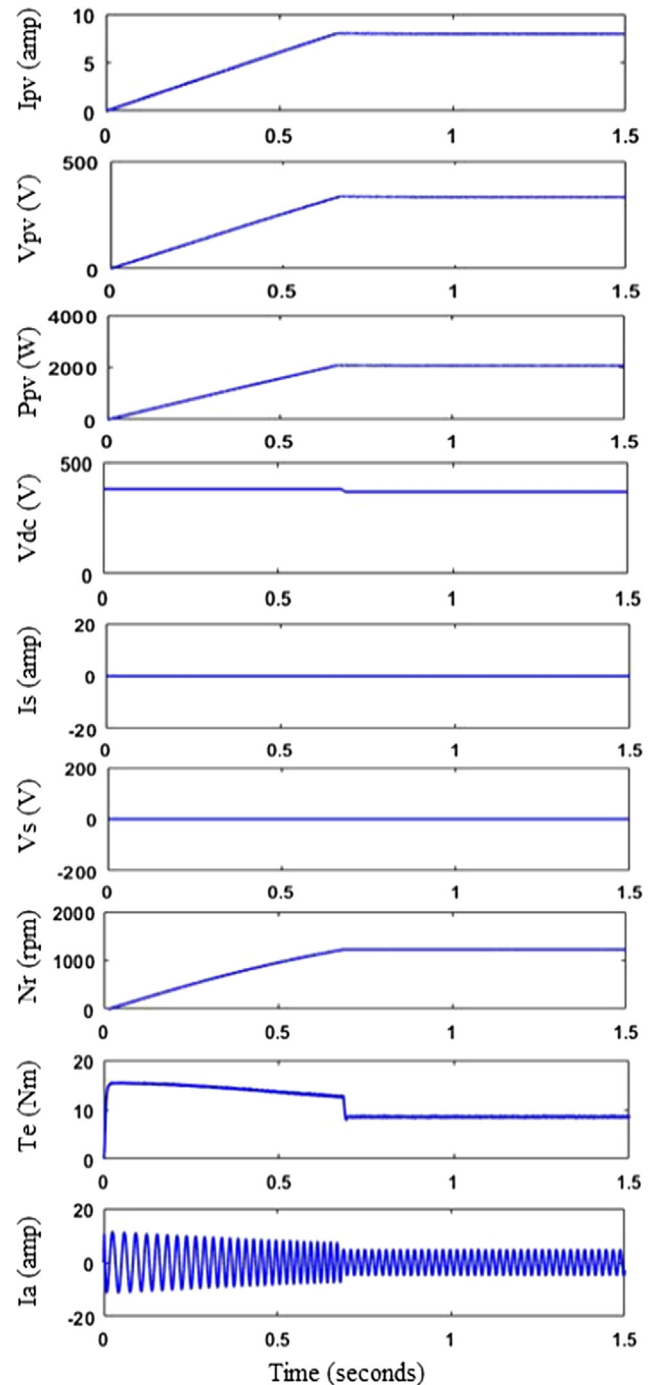
Fig. 17. Mode I: starting performance.

$$T_e(t) = \frac{3}{2} P \frac{L_m}{\sigma L_s L_r} \overrightarrow{\Psi_r}(t) * |\Psi_s| e^{j\omega_s t} \qquad (3.28)$$

By substituting Eq. (3.13) in (3.15)

$$T_e(t) = \frac{3}{2} P \frac{L_m}{\sigma L_s L_r} \frac{L_m}{L_s}$$

$$\times \frac{\sqrt{1 + \left(e^{-\frac{t}{\tau}}\right)^2 - 2e^{-\frac{t}{\tau}} \cos\left((\omega_s - \omega_m)t\right)}}{\sqrt{1 + ((\omega_s - \omega_m)\tau)^2}} |\Psi_s|^2$$

$$* \sin\left\{\omega_s t - \tan^{-1}\left(\frac{y}{x}\right) - \tan^{-1}\left((\omega_s - \omega_m)\tau\right)\right\} \qquad (3.29)$$

The non-linear relationship of Eq. (3.29) helps in determining the dynamic response of torque by the amplitude and rotating speed of the stator flux vector. If the amplitude of stator flux is kept constant the torque can be easily regulated by controlling rotating speed of the stator flux vector $\overrightarrow{\Psi_s}$. Therefore, the expression given above for torque can be simplified to Eq. (3.30) in which the slip speed is small.

$$T_e(t) = \left\{\frac{3}{2} \frac{P}{2} \frac{L_m^2}{R_r L_s^2} |\Psi^*_s|^2\right\} \left(1 - e^{-\frac{t}{\tau}}\right)(\omega_s - \omega_m) \qquad (3.30)$$
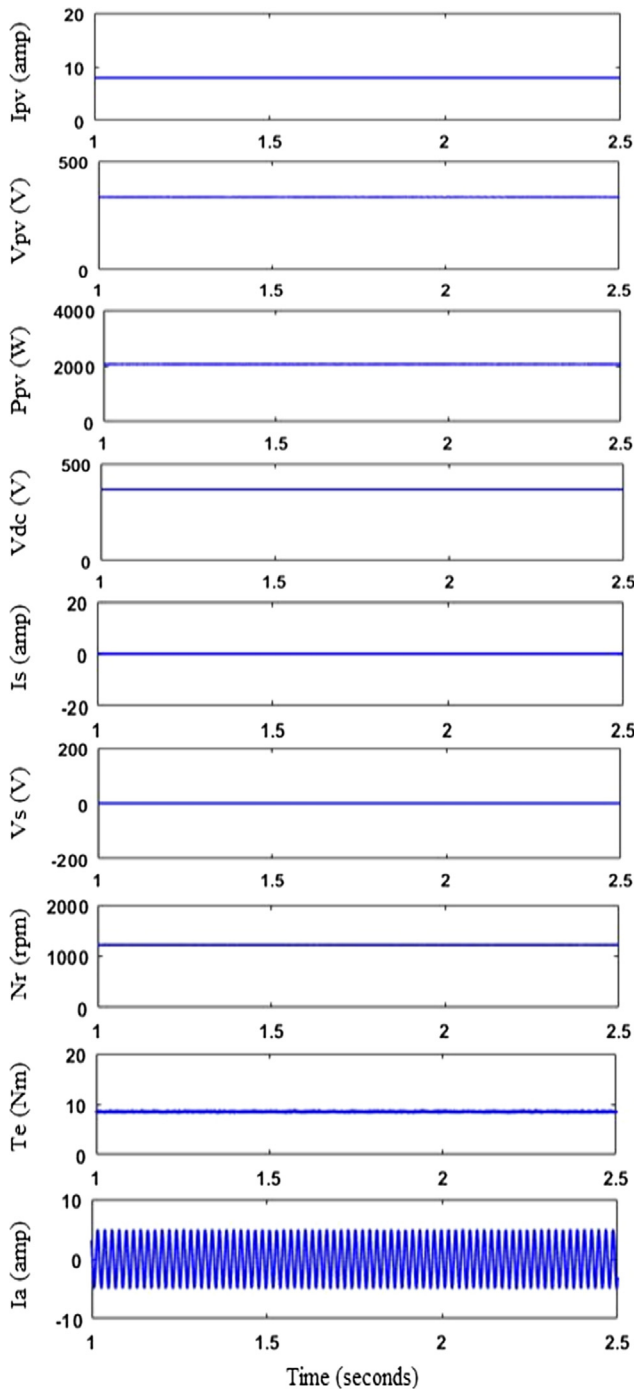
Fig. 18. Mode I: Steady state performance.
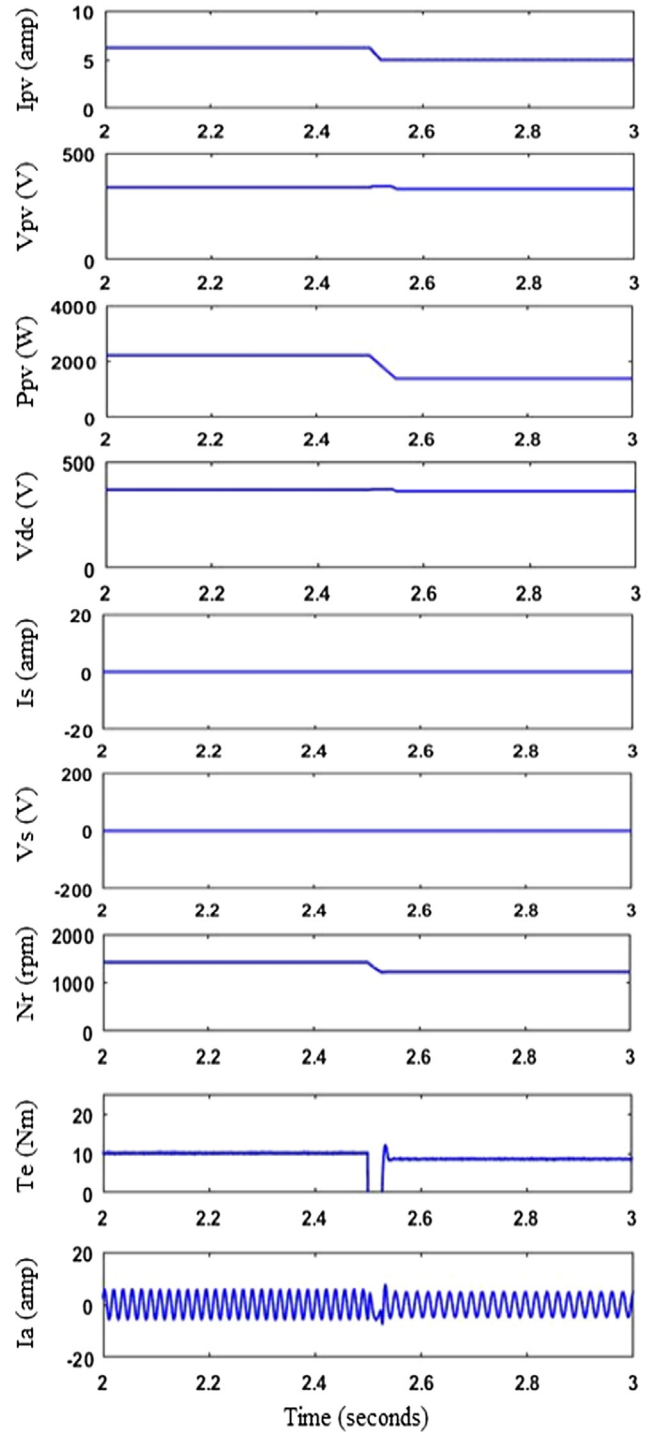
Fig. 19. Mode I: Decrease in insolation performance 800 W/m² – 500 W/m².
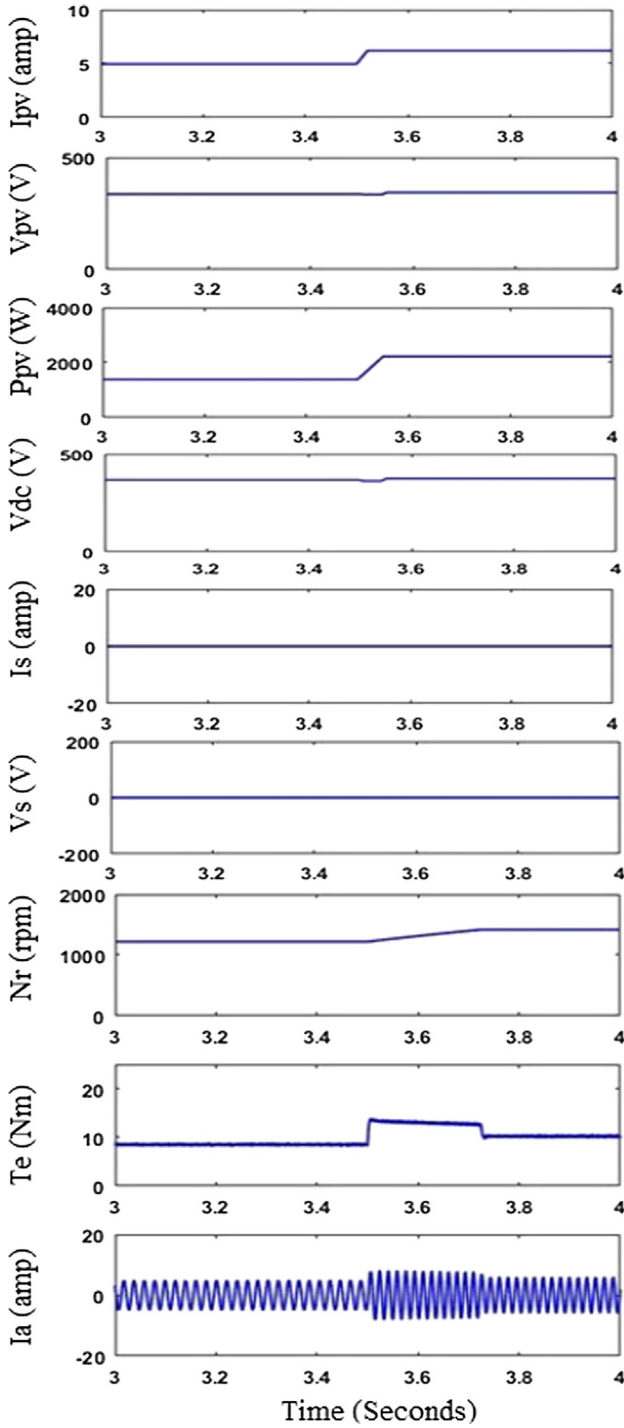
Fig. 20. Mode I: Increase in insolationperformance 500 W/m² – 800 W/m².

$$T_e(t) = K\left(1 - e^{-\frac{t}{\tau}}\right)(\omega_s - \omega_m) \tag{3.31}$$

From this relation we can see that the torque is proportional to slip speed. And the above equation can be simplified as

$$\frac{T_e}{\omega_{sl}^*} = \frac{K_m}{1 + sT_m} \tag{3.32}$$

where $K_m = \frac{3}{2}\frac{P}{2}\frac{L_m^2}{R_r L_s^2}|\Psi^*{}_s|^2$ and $T_m = \frac{\sigma L_r}{R_r}$

The mechanical load balance equation of the machine is given by
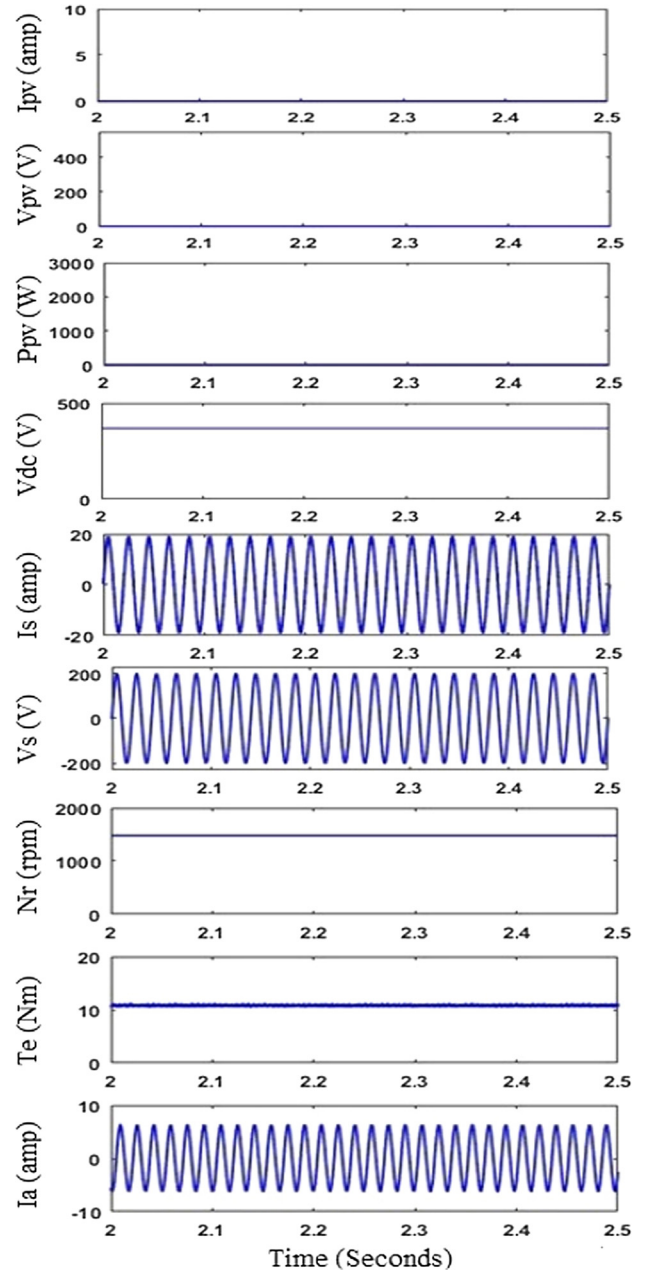
Fig. 21. Mode II: Steady state performance.

$$T_e - T_l = J\frac{d\omega_m}{dt} + B\omega_m \tag{3.33}$$

So now the mechanical transfer function of the motor can be obtained from Eq. (3.33) as

$$\frac{\omega_r}{T_e - T_l} = \frac{1}{Js + B}$$

From the above analysis, the gains of PI controller for the inner torque loop can be calculated first and then the outer speed loop PI gains can be calculated in order to get a desired response.

4. Simulation results and discussion

Figs. 7–10 show the flux locus, torque and flux waveforms of both classical and improved SVM DTC, and by comparison it is clearly observed that flux locus of improved SVM DTC is pure circle
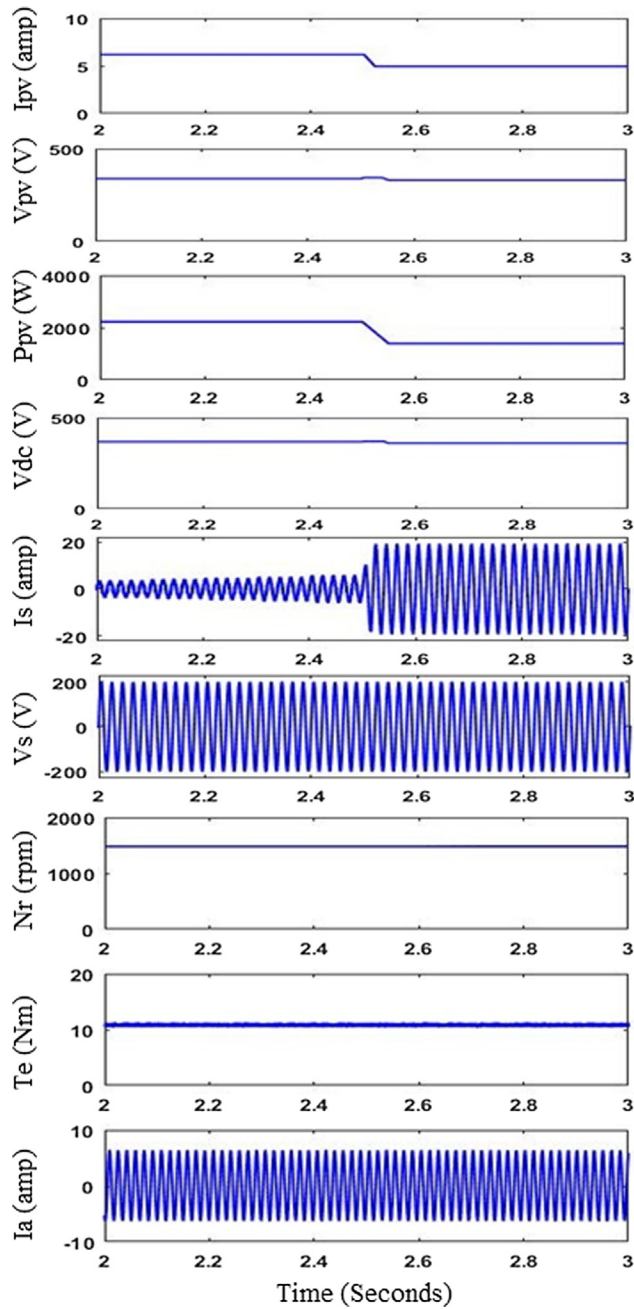
Fig. 22. Mode III: Decrease in insolation performance 800 W/m^2 – 500 W/m^2.
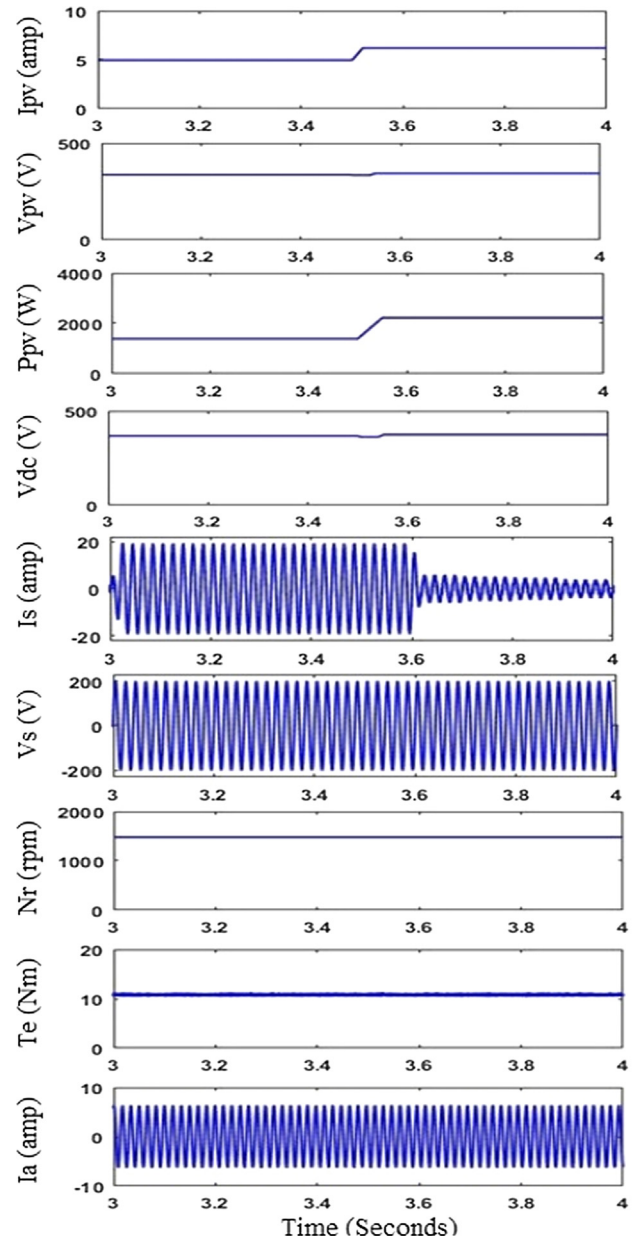
Fig. 23. Mode III: Increase in insolation performance 500 W/m^2 – 800 W/m^2.

without any ripples, and torque response is also taking less settling time with less ripples in SVM DTC, so the improved SVM DTC is preferred here for the proposed system.

A step change in the reference speed from 1000 rpm to 1200 rpm at 1 sec and the settling time of speed for both the methods are shown in Figs. 11 and 12 below.

The torque response, current and speed response of the SVM-DTC & classical DTC for motor is as shown below in Figs. 13 and 14. The response is plotted for a step change in the load torque from 0 to 8 N-m at 1.5 s.

Overall behavior of torque, stator current and rotor speed of motor for both the schemes are as given in Figs. 15 and 16 for a step change in the load torque from 0 to 8 N-m at 2 secs and step change in the reference speed from 1000 rpm to 1200 rpm at 1 sec.

Comparison between the two schemes is shown in the Table 3 below which clearly shows that the settling time and the ripples in the SVM-DTC is much less as compared to the classical DTC.

Figs. 17–20 gives the results of the system for mode I of section 3.2, where pump is driven by only solar PV system. The starting and steady state performance of the system for mode I is shown in Figs. 17 and 18 for radiation 800 W/m^2, here the DC link voltageV_{DC}, V_{mpp} and I_{mpp} are maintained at their respective values. For the starting mode torque is observed to be little high for some instant and then settled, and for steady state torque and speed are constant at nearly 10Nm and 1500 rpm.

Effect of change in radiation is seen in current and power of PV array and motor parameters i.e. when the radiation is decreased from 800 W/m^2 to 500 W/m^2 the Ipv and Ppv are also decreased which also lead to decrease in torque, speed and current of motor drive and when the radiation is increased from 500 W/m^2 to 800 W/m^2, current Ipv and power P_{pv} of PV array is increased and same is seen in torque, speed and current of motor drive
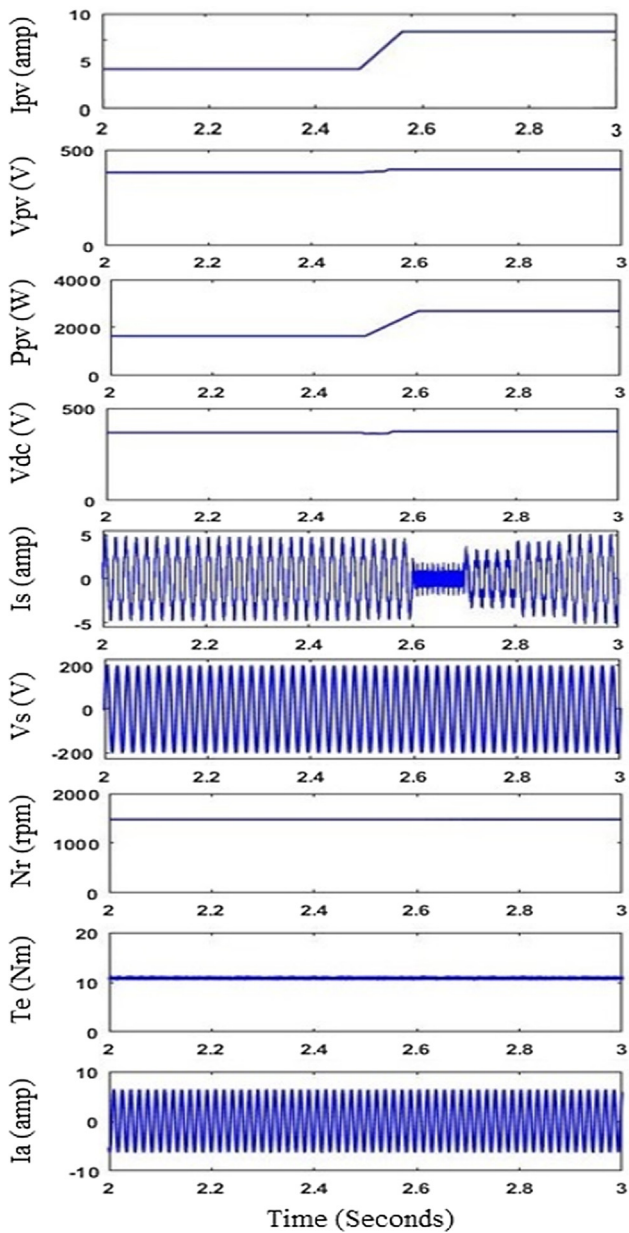
Fig. 24. Mode IV: Increase in insolation performance 500 W/m² – 1000 W/m².

parameters Figs. 19 and 20, whereas change in V_{pv} and DC link voltage V_{DC} is almost negligible in both the cases.

In mode II of section 3.2, in which motoris driven by grid supply only so the steady state performance is observed in Fig. 21, where it is clearly seen that the I_s is in phase with V_s and rated discharge is obtained from the pump in this mode as the motor is run at its rated speed and torque.

Figs. 22 and 23 give the results for the mode III in which the system is operated by both the sources i.e. from SPV array and grid supply. When the power from the SPV array is not sufficient the requirement is fulfilled by the utility grid supply, from the results it is clearly seen that when the radiation is decreased grid current increases and similarly when radiation is increased grid current decreases with almost negligible change in torque and speed of the motor.

For mode IV 1000 W/m² radiation is considered. The maximum power at 1000 W/m² is 2.7Kw which is more than the drive's required rated capacity, so the excess power is delivered to the grid as seen in Fig. 24.

5. Conclusion

A smart water pumping system with simplified SVM-DTC with new switching scheme that reduces the complexity of identifying the sector and the angle of the voltage vector for switching the inverter, also interfaced with grid and operates in different modes is modeled and verified in this paper.

A smart power sharing between the SPV array and utility grid, undisputable power transfer between the sources, reduced harmonics distortion that helps in power saving and improves the efficiency further with less complexity and less burden on processor, the faster settling time and less torque ripples are the main features of the proposed system. The system is beneficially and practically suited for irrigation and household purposes as it managed to cut down the electricity bill to a immense amount as the designed system draws the maximum power from the SPV array which in result reduces the burden on the utility grid.

References

[1] Mapurunga Caracas JV, de Carvalho Farias G, Moreira Teixeira LF, de LA, Souza Ribeiro. Implementation of a high-efficiency, high-lifetime, and low-cost converter for an autonomous photovoltaic water pumping system. IEEE Trans Pwr Electron 2014;50(1):631.
[2] Thampi N, Nisha R. A solar powered water pumping system with efficient storage and energy management. Int J Adv Res Electr Commun Engg 2016;338.
[3] Chen Y-M, Wu H-C, Chen Y-C. DC bus regulation strategy for grid-connected PV power generation system. Proc IEEE ICSET 2018;437.
[4] Esram T, Chapman PL. Comparison of photovoltaic array maximum power point tracking techniques. IEEE Trans Energy Conver 2007;439.
[5] Xiao W, Dunford WG. A modified adaptive hill climbing MPPT method for photovoltaic power systems. Proc 35th annual conference of IEEE Power Electronics specialists conference, Aachen, Germany, 2004.
[6] Pandey A, Dasgupta N, Mukherjee AK. Design issues in implementing MPPT for improved tracking and dynamic performance. IEEE IECON, Paris, France, 2008.
[7] Hua C, Lin J. A modified tracking algorithm for maximum power tracking of solar array. Energy Convers Manage 2004;911.
[8] Dzung PQ, Khoa LD, Lee HH, Phuong LM, Vu NTD. The new MPPT algorithm using ANN based PV. Proc Int Forum Strategic Technol 2010;401:401.
[9] Zakzouk N, Elsaharty M, Abdelsalam A, Helal A, Williams B. Improved performance low-cost incremental conductance PV MPPT technique. IET Renew Power Gener 2016;10(4):561.
[10] Mei Qiang, Shan Mingwei, Liu Liying, Guerrero Josep M. A novel improved variable step-size incremental-resistance MPPT method for PV systems. IEEE Trans Ind Electron 2011;58(6):2427.
[11] Jain S, Thopukara AK, Karampuri R, Somasekhar V. A single-stage photovoltaic system for a dual inverter-fed open end winding induction motor drive for pumping applications. IEEE Trans Pwr Electron 2015;30(9):4809.
[12] Razzak MA, Chowdhury ASK, Salam K. Induction motor drive system using push-pull converter and three-phase spwm inverter fed from solar photovoltaic panel. Power Energy Systems Conf: Towards Sustain Energy 2014;1.
[13] Vas P. Vector control of AC machines. USA: Oxford University Press; 1990.
[14] Peter V. Sensorless vector and direct torque control (1998).
[15] Bose BK. Power electronics and motor drives: advances and trends. Academic Press; 2010.
[16] Mohan N. Advanced electric drives: analysis, control, and modeling using MATLAB/Simulink. John Wiley & Sons; 2014.
[17] Jain S, Karampuri R, Somasekhar V. An integrated control algorithm for a single-stage PV pumping system using an open-end winding induction motor. IEEE Trans Ind Elec 2016;63(2):956.
[18] Achour A, Rekioua D, Mohammedi A, Mokrani Z, Rekioua T, Bacha S. Application of direct torque control to a photovoltaic pumping system with sliding-mode control optimization. Electric Power Comp Syst 2016;44(2).
[19] Muljadi E. PV water pumping with a peak-power tracker using a simple six-step square-wave inverter. IEEE Trans Ind Appl 1997;33(3):714.
[20] Sharma U, Kumar S, Singh B. Solar array fed water pumping system using induction motor drive. 1st IEEE Intern Conf on Power Electronics, Intelligent Control and Energy Systems. Delhi 2016.
[21] Franklin T, Cerqueira J, de Santana E. Fuzzy and PI controllers in pumping water system using photovoltaic electric generation. IEEE Trans Latin America 2014;12(6):1049.
[22] de Brito MAG, Galotto L, Sampaio LP, e Melo GDA, Canesin CA. Evaluation of the main MPPT techniques for photovoltaic applications. IEEE Trans Ind Electron 2013;60(3):1156.
[23] Foo GHB, Zhang X. Constant switching frequency based direct torque control of interior permanent magnet synchronous motors with reduced ripples and fast torque dynamics. IEEE Trans Power Electron 2016;31(9):6485.
[24] Premalatha D, Sakthi Rubini A. Direct torque control of brushless DC motor using PI and fuzzy controller. Int J Sci, Eng Technol Res 2015;4(4):922.

[25] Zhang Jun, Rahman MF. Analysis and design of a novel Direct flux control scheme for induction machine. IEEE international electric machines and drives conference, San Antonio, USA; 2005. p. 426.

[26] Chung Dae Woong, Kim Joohn Shoek, Sul Seung Ki. Unified voltage modulation technique for real time three phase power conversion. IEEE Trans Ind Appl 1998;34(2):374.

[27] Sharma U, Dwivedi S, Jain C. Single stage solar PV array fed field oriented controlled induction motor drive for water pump. National Power Electronics Conf (NPEC), IIT Bombay 2015.

[28] Chaudhari MA, Suryawanshi HM, Renge MM. A three-phase unity power factor front-end rectifier for AC motor drive. IET Power Electron 2012;5(1):1.

[29] Zeng Z, Yang H, Zhao R, Cheng C. Topologies and control strategies of multifunctional grid-connected inverters for power quality enhancement: a comprehensive review. J Renew Sustain Energy Rev 2013;24:223–70.

[30] Hasanien Hany M. Performance improvement of photovoltaic power systems using an optimal control strategy based on whale optimization algorithm. Electr Power Syst Res 2018;157(168).

Priyanka Singh received the B. Tech. degree in Electrical Engineering fromRajasthan Technical University, India in 2010, and M. Tech. degree in Power Electronics from Birla Institute of Technology, India. She has two years of teaching experience and two years industry experience. She is currently pursuing the Ph.D degree at Netaji Subhas Institute of Technology, India. Her current research interest areas includes power electronics converters, electrical drives and renewable energy systems.

Prerna Gaur received the B. Tech. degree in Electrical Engineering from GB Pant University of technology and agriculture, India in 1988, M. Tech. degree from Delhi college of Engineering, India in 1996, and Ph.D degree from Delhi University, India in 2008. She has six years of industry experience and twenty two years of teaching experience in Delhi college of Engineering and Netaji Subhas institute of technology india. she is also head of ICEdepartment. Besides teaching she is actively associated with IEEE (Senior Member), ISTE (Life Member) and Fellow IE. She has important positions in IEEE and has received many honors and awards. Her area of interest includes renewable energy, power electronics, power quality, artificial intelligence based control and electric drives.

Journal Pre-proof

High-Performance Dual Cavity-Interferometric Volatile Gas Sensor Utilizing Graphene/PMMA Nanocomposite

Dnyandeo Pawar, Rajesh Kanawade, Ajay Kumar, Ch.N. Rao, Peijiang Cao, Shankar Gaware, Dattatray Late, Sangeeta N. Kale, S.T. Navale, W.J. Liu, D.L. Zhu, Y.M. Lu, Ravindra K. Sinha

High-Performance Dual Cavity-Interferometric Volatile Gas Sensor Utilizing Graphene/PMMA Nanocomposite

Dnyandeo Pawar[a,b*], Rajesh Kanawade[c], Ajay Kumar[c], Ch. N. Rao[a], Peijiang Cao[a*], Shankar Gaware[d], Dattatray Late[e], Sangeeta N. Kale[d], S. T. Navale[a], W. J. Liu[a], D. L. Zhu[a], Y.M. Lu[a], Ravindra K. Sinha[c,f]

[a]*College of Materials Science and Engineering, Guangdong Research Center for Interfacial Engineering of Functional Materials, Shenzhen University, Shenzhen 518055, PR China*

[b]*Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, PR China*

[c]*Advanced Materials and Sensors Division (V4), CSIR-Central Scientific Instruments Organisation, Sector 30-C, Chandigarh 160030, India*

[d]*Department of Applied Physics, Defence Institute of Advanced Technology, Pune 411025, India*

[e]*Centre for Nanoscience and Nanotechnology, Amity University Mumbai, Maharashtra 410206, India*

[f]*TIFAC-CORE, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India*

[*]*Corresponding authors:*

Email addresses: pjcao@szu.edu.cn (P. Cao), pawar.dnyandeo@gmail.com (D. Pawar)

Graphical Abstract:
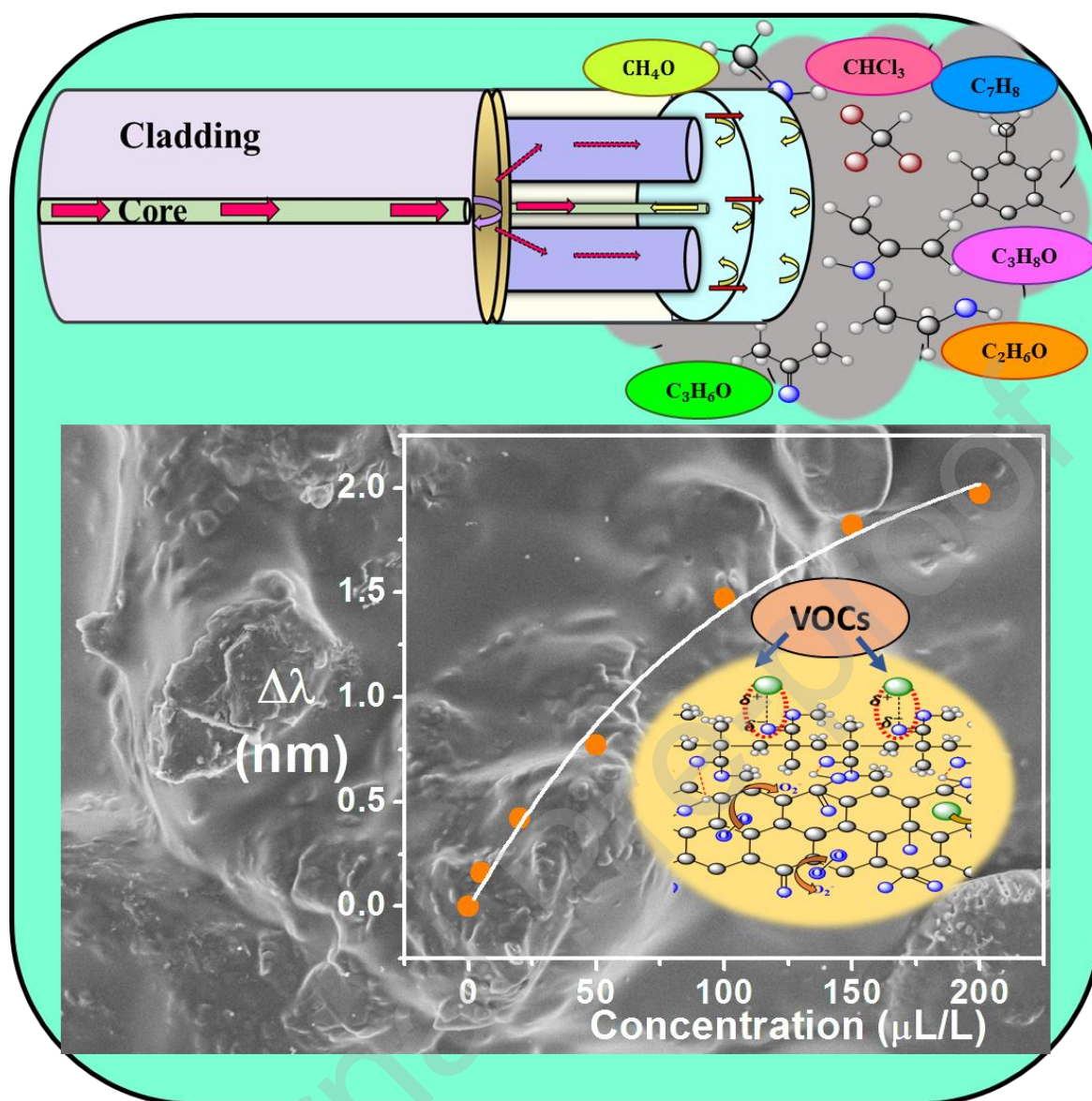
Highlights:

- **A new design based on Dual-cavity interferometer coated with graphene, PMMA and graphene-PMMA is demonstrated for various volatile gases detection**
- **The sensing mechanism is based on change in refractive index of graphene, effective swelling of PMMA and simultaneous change in their refractive index/swelling under the exposure of different vapors**
- **The experimental data is well fitted with the models based on Langmuir adsorption isotherm and Hansen solubility parameters**
- **A simple, flexible design makes it attractive for many industrial applications and environmental monitoring.**

2

ABSTRACT

Volatile organic compounds (VOCs) monitoring is become a vital need for many industrial and environmental air quality monitoring applications. We proposed and demonstrated a polarization maintaining fiber (PMF) dual-cavity Fabry-Perot interferometric (DFPI) sensor by immobilizing graphene, PMMA and graphene/PMMA composite for detection of various volatile gases. The optical properties of these materials changes when it is exposed to different VOCs, which proportionally leads to wavelength and power shift in an interference spectrum due to effective change in refractive index of graphene and swelling of PMMA respectively. The experimental results follow Langmuir dependence and are analyzed through DFPI theory and follows Hansen solubility parameters (HSP) model. This study focuses on insight into the physical/chemical interaction of volatile gases with versatile sensing materials along with new optical fiber configuration approach, which could be explored for remote detection and air quality monitoring.

KEYWORDS: Chemical sensors; Refractive index; Swelling; Fiber-optic, Langmuir isotherm; Hansen solubility.

1. Introduction

   Clean air is an essential need for well human health and maintaining sustainable environment. Large number of volatile organic compounds (VOCs) are releasing day by day and it is creating serious impact on human well-being and environment [1]. Many VOCs are toxic in nature and considered to be carcinogenic [2]. VOCs are considered as a major class of pollutant especially in indoor places (almost >10 times than outdoors) where the human activity is too much involved [3]. Furthermore, most of the VOCs are not sensitive at room temperature and therefore it's difficult to monitor their concentration. Therefore, it is a great importance to develop a sensor to monitor such trace level of VOCs. Since last decade, consistent effort has been made by the researchers in the development of a reliable, cost-effective and sensitive VOCs sensor based on electronic as well as on optical way. VOCs sensors based on mass spectrometry [4] and gas chromatography [5] have achieved greater limit of detection (LOD) and wide measuring range. However, high instrumental cost and bulky set up limits its use for personal and mobility. Alongside, colorimetric-based sensors are also being developed for detection of various vapors with simple, cost-effective and real-time monitoring approach [6]. The demerit of this technique is to obtain the actual concentration of VOCs and lack in selectivity in case of VOCs mixture. Metal oxides (MOXs) based chemiresistor sensors are proved to be very effective for gases detection. Many MOXs such as In_2O_3 [7], ZnO [8],[9], TiO_2 [10],[11], SnO_2 [12],[13] and others [14] have been widely explored for gas and vapor sensing. However, despite its higher sensitivity towards chemical species, the MOXs are not suitable choice for the development of gas or vapor detection. The high working temperature ($>200\,^0C$) and large power consumption limits its use at room temperature-based applications. However, certain room temperature MOXs based sensors have been fabricated but still the issues are not properly resolved. Also, at high temperature, detection of flammable and explosive vapors like ethanol and acetone is extremely dangerous by

considering the human safety and infrastructure damage. So, it requires special precautions to avoid accidental harm.

Therefore, considering the above facts, the devices based on optical techniques especially a fiber-optics based sensors have been attracted a considerable attention due to their many amazing features include capability of real time monitoring, multiplexing, not producing any electrical spark and ability to work in complex environment [15]. Along with these very miniaturized and low-cost sensory platforms can be fabricated and deployed anywhere. Fiber-optics based sensors are proved to be very effective over conventional based sensors in terms of reliability, selectivity, and stability. With advancements of nanotechnology and many fabrication techniques, the nanomaterials combined with fiber-optics based sensors have bright future in the field of chemical sensing. Since last decade, different carbon-based nanomaterials have been emerged as an excellent kind of materials for various gases and vapors monitoring. Carbon nanotube (CNTs) [16] and graphene based derivatives including reduced graphene oxide [17] and graphene oxide [18],[19] are well explored in gas sensing field. These materials have been attracted due to their high surface area, ability of fast electron transfer and low signal noise and therefore could be considered as a promising candidate for portable sensors [20],[21]. Lot of efforts have been made to enhance the sensor performance by combining graphene and MOXs for gas sensing [22],[23] The heterojunction formed at graphene and MOXs boost the reaction kinetics multifold times and enhances the sensor performance. Utilization of a sensor at room temperature is very crucial task. Therefore, most of the graphene and polymer-based composites have been effectively deployed for gas sensing application at room temperature. Polymer is a class of material in which either their conductivity changes or it swells in presence of gas/vapors and also efficient at room temperature and therefore, extensively studied for detection of various gases and vapors [24],[25],[26],[27],[28]. Combining their properties by modulating the carrier concentration of

5

graphene and ability of swelling the polymer under the exposure of different vapors could become a potential competitor for monitoring various gases and vapors at room temperature. Due to several advantages of optical fiber, lot of fiber-optics sensors have been fabricated for the detection of various kind of vapors. Table 1 shows the advancement of nanomaterials including graphene, MOXs and polymer modified optical fiber sensors for VOCs monitoring. These approaches consisting certain geometries including coated on side polished fibers [29], coating on unclad part of fiber surface [30],[31],[32],[33],[34], coating on tapered fiber surface [35], material coated at the tip of the fiber [36], polarization maintaining fiber coated [18] and Fiber gratings (FGs) coated [37],[38],[39] etc. From table, it observed that most of them possesses certain drawbacks such as some shows high response and recovery time, exposure of very high concentration, lack in sensitivity, fragility due to tapering fiber structure and not efficient to detect the mixture of VOCs.

We observed that most of the common geometries such as Fabry-Perot, SPR and fiber grating based sensors have been implemented for VOCs detection. All these configurations have certain limitations as discussed above and therefore, a different kind of microfiber based dual-cavity Fabry-Perot interferometric (DFPI) structure has been aimed to fabricate due to its flexibility, simplicity and highly sensitive nature to chemical moieties. Compared with normal optical fibers, the microfibers have certain advantages such as low loss transmission, ability to confinement of light field, tunability of sensing region and ease to fabricate and therefore the multimodal interference spectrum can be obtained which could be modulated easily by controlling the parameters and properties of microfiber. By using microfibers [40],[41],[42] and specialty fibers [43],[44] numerous versatile geometries have been fabricated and successfully employed for gas and another parameters detection.

In this work, we have proposed and demonstrated performance of DFPI-based sensors for monitoring individual vapor moieties as well as multi-vapors in the concentration range from 5 μL/L to 200 μL/L. The DFPI-based structure is fabricated by using a PMF microfiber tip utilized with sensitive materials like graphene, PMMA and graphene/PMMA composite. The interaction of various volatile gases with sensing materials based on experimental analysis and theoretical base are well analyzed and elaborately presented.

2. Experimental details

2.1. Synthesis and characterization of sensing materials

The sensing materials such as synthesized graphene, PMMA (Alfa Aesar, India) and graphene/PMMA were used for sensing various VOCs. To obtain graphene sample, 100 mg exfoliated graphite (Nacional de Grafite Co., Brazil) taken into 100 ml of isopropanol and treated for 2 hours in an ultrasonic bath followed by drying in an incubator to obtain graphene powder [45]. To obtain graphene/PMMA nanocomposite; firstly, 75 mg of PMMA in 10 ml of toluene was taken in a mortar and grinded well by using a pestle for 10 minutes. Later, 25 mg of as-prepared graphene powder was added into it and again grinded it for 10 minutes. Due to uniform grinding, the solution was homogeneously mixed and when it turned into viscous form, the PMF (Nufern, PM1550-HP, beat length 4 mm at 1550 nm) tip was coated by using dip coating technique.

The sensing materials were characterized and analysed through numerous characterization techniques viz Field Emission Scanning Electron Microscopy (FESEM; Carl Zeiss's Sigma), Raman ((LABRAM HR-800) at wavelength of 532 nm, Thermo-Gravimetric Analysis (TGA; Perkin-Elmer STA 6009) and Fourier Transform Infrared Spectroscopy (FTIR; Bruker, Germany-3000).

2.2. Fabrication of dual-cavity Fabry-Perot interferometric (DFPI) probe

The sensing mechanism of this developed sensor is based on DFPI principle. To fabricate the DFPI, a PMF of length 5 mm was spliced at the distal end of single mode fiber (SMF) by using a fusion splicer (T35, Fujikura, Japan). The sensing material is coated at the distal end of PMF by using dip coating method. The three sensing probes namely graphene coated dual-cavity Fabry-Perot interferometric (G-DFPI), PMMA coated dual-cavity Fabry-Perot interferometric (P-DFPI), and graphene/PMMA coated dual-cavity Fabry-Perot interferometric (GP-DFPI) were prepared by dip coating technique. After that the coated probes were dried at room temperature for 24 hours.

The schematic of nanomaterial/polymer-modified DFPI is shown in Figure 1(a). It consists of three reflecting surfaces in which the light reflects and get coupled into the coupler and generates an interference spectrum. The corresponding reflection coefficients from these three surfaces denoted as R_1, R_2 and R_3 from the interfaces between SMF-PMF, PMF-sensing film surface, and sensing film-air respectively. The majority of light couple into the solid core of PMF (PMF structure is shown in Figure 1(b)) and remaining light coupled into the cladding part of PMF (splicing image of SMF-PMF is shown in Figure 1(c)). At R_2, certain proportion of light reflects and majority of light travel into an FP cavity. Thus, this geometry consists of two cavities created from combination of PMF and sensing material. When light couple into the PMF through collapsing region and reflects from PMF-film interface forms a first cavity and second cavity formed due to the sensing material where light get multiply internally reflected. The combined geometry of both cavities contributes to generate a very fine interference spectrum.

2.3. Experimental

The experimental set up consists of a 3-dB coupler whose one end is connected to the C-band laser source (wavelength range: 1525-1565 nm, Opto-link Corporation Ltd., Hong-Kong) and another end to the optical spectrum analyzer (OSA) (Wavelength range 600-1700 nm, wavelength resolution: 20 pm, AQ6370C, Yokogawa, Japan) (Figure S1). The three-different cavities have been prepared and studied their response under the exposure of different volatile compounds such as acetone, methanol, chloroform, toluene, ethanol and isopropanol in concentration ranging from 5 µL/L to 200 µL/L. The sensing probe was inserted into the glass gas chamber (volume: 100 ml, Borosil) and the known amount of VOC liquid (varied from 0.5 µL - 20 µL) was injected into it by using a micropipette (range 0.5-100 µL, Fisher Scientific, USA). The concentration of liquid (ranging from 5 µL/L, 20 µL/L, 50 µL/L, 100 µL/L, 150 µL/L and 200 µL/L) was calculated from the liquid taken in micropipette per total volume of the gas chamber. When sensor exhibits saturation, the probe allowed to interact with outside air by opening the gas-outlet of the gas chamber. The recovery of the sensor was achieved in air at room temperature. The air is blown as a precaution to remove any contamination to the measured VOCs. This experimental procedure was followed for each measurement for all three sensing probes. In case of detection of VOCs mixture, same sensing probes were utilized. In this measurement, the mixture of VOCs of specific concentration (mixing ratio varied from 0.5 µL - 20 µL) was prepared by mixing an appropriate concentration of each VOC and then injected inside the gas chamber by using a micropipette. The concentration of mixture of VOCs is varied from 5 µL/L, 20 µL/L, 50 µL/L, 100 µL/L and 200 µL/L respectively. Furthermore, the sensors responses are also measured with binary mixtures of acetone and other VOCs in concentration ranging from 5 µL/L to 200 µL/L. The four type of binary mixtures such as acetone + ethanol, acetone + methanol, acetone + chloroform and acetone + toluene was prepared in equal

proportion (mixing ratio varied from 0.5 μL -20 μL). The gas concentration of binary mixture varied from 5 μL/L, 50 μL/L, 100 μL/L and 200 μL/L, which are obtained from the appropriate mixing of VOC liquid solution by using micropipette. For all measurement, the recovery of the sensor was measured and confirmed from the wavelength dip of an interference pattern. In the experimental process, the temperature and relative humidity was around 23 ^{0}C and 65 %RH respectively which measured by a digital thermo/hygrometer instrument (288-ATH, HTC).

In this study, the response of the sensor is measured in terms of wavelength shift ($\Delta\lambda$) upon exposure of analyte vapor concentration (C). The sensitivity of the sensor is defined as: $S = \Delta\lambda\ (nm)/C(\mu L/L)$ where, $\Delta\lambda$ is the total wavelength shift per analyte concentration (C). The sensor response is the period in which the sensor reaches 90% response value upon exposure of analyte vapor and recovery time is defined as period in which the sensor response value changes to its 10% of the response value after the analyte vapor is removed. The limit of detection (LOD) of sensor is defined as the minimum concentration of analyte vapor can be detected under given conditions. Along with sensor sensitivity, the selectivity is also an important parameter in gas sensing. The sensor selectivity is defined as whether a sensor can respond selectively to group of analytes or specifically to a single analyte. All these parameters are analyzed and discussed in following sections.

3. Theoretical analysis:

According to multi-beam theory, interference will take place when three reflection beams coupled back into the SMF and interfere with one another. The total reflected intensity is expressed as follows [42,46].

$$I = I_1 + I_2 + I_3 - 2\sqrt{I_1 I_2}\cos\left(\frac{4\pi B_{PMF}L_1}{\lambda} + \varphi_{10}\right) - 2\sqrt{I_2 I_3}\cos\left(\frac{4\pi n_3 L_2}{\lambda} + \varphi_{20}\right) + 2\sqrt{I_3 I_1}\cos\left(\frac{4\pi(B_{PMF}L_1 + n_3 L_2)}{\lambda} + \varphi_{30}\right) \tag{1}$$

Where, I_1, I_2 and I_3 are the light beam intensities reflected from three reflection surfaces; λ is the dip wavelength; n_3 is the refractive index of sensing material; φ_{10}, φ_{20} and φ_{30} are the initial phases; L_1 and L_2 are the length of the PMF and sensing cavity (material coated at the tip of PMF) and $B_{PMF} = n_2 = n_x - n_y$ is the birefringence of PMF (n_x and n_y are the refractive indices along the slow and fast axis of the PMF).

Here, PMF plays very important role in the fabrication of dual cavity. In SMF-PMF structure, along with core light, most of the cladding light is also propagates into the sensing film where it gets multiply internally reflected and strongly coupled, which further increase the fringe visibility and fineness of the FP cavity. Therefore, the fringe pattern is modulated by an interference of PMF, which gives a higher extinction ratio in fiber structure [47–49].

Therefore, the total optical phase difference (φ) and path difference (δ) can be given by [46,50]:

$$\Delta\varphi = \frac{2\pi}{\lambda}.\delta = \frac{4\pi(n_3L_2 + n_2L_1)}{\lambda} = (2m+1) \; ; \; \delta = 2(n_3L_2 + n_2L_1) \tag{2}$$

Where m is an integral order of interference fringe (m = 0, 1, 2, 3…). It is observed that the total phase difference is depending upon both refractive index and length of the PMF and FP cavity respectively.

When any vapor adsorbs on the sensing film, the length and refractive index of the sensing film will become as $L_2 + \Delta L_2$ and $n_3 + \Delta n_3$ respectively. Here, PMF refractive index n_2 and length L_1 will remain constant as this region is not exposed to vapor. Therefore, the corresponding wavelength shift for m^{th}-order interference is given as [46]:

$$\Delta\lambda = \frac{4\left[\Delta n_3L_2 + (n_3 + \Delta n_3)\Delta L_2\right]}{2m+1} \tag{3}$$

Generally, Langmuir isotherm is ideal to describe the saturation behavior of sensing material with external gas or vapor at low concentration. When sensing surface is exposed to any vapor,

11

initially it shows linear response with respect to vapor concentration and then reaches to an asymptotic value when adsorption sites approach saturation [51,52]. Therefore, from FP theory the resonance wavelength relation with Langmuir adsorption isotherm parameters is given by [52]:

$$\lambda_{res} = \frac{4\gamma\left[\Delta n_3 L_2 + (n_3 + \Delta n_3)\Delta L_2\right]}{(2m+1)}\frac{\left[1-\phi_v(C)\right]}{\phi_v(C)}$$

(4)

Where $\phi_v(c)$ is the maximum adsorption sites, and $\gamma = 3$ or 1 for isotropic volume and linear expansion, respectively. Therefore, the resonance dip position is depending upon the refractive index n_3, length of the FP cavity (due to swelling) of sensing film, maximum adsorption capacity of the sensing film and gas concentration, respectively.

The Figure 2(a, b, c) shows the electric filed distribution (designed in COMSOL Multiphysics 5.2) in different DFPIs such as P-DFPI, G-DFPI and GP-DFPI respectively. The corresponding reflection coefficients from three surfaces are calculated by using Fresnel's equations. In case of P-DFPI, R_1 is 0.010%, R_2 is 0.0029% and R_3 is 3.87%. For G-DFPI, R_1 is 0.010%, R_2 is 7.63% and R_3 is 19.75%. For GP-DFPI case, R_1 is 0.010%, R_2 is 0.014% and R_3 is 4.12%. At R_1, the reflection coefficient is not that much strong comparatively with R_2 and R_3 due to the minimum refractive index difference between optical fiber itself (core of SMF and PMF). Figure 2(d) shows the spatial frequency spectrum for each configuration obtained from their reflected interference spectrum. It shows that not only core light passes into the sensing film but also light from cladding part goes into the film and therefore, it increases the multiple internal reflection inside the sensing film and thus the fringe pattern is modulated by an interference of PMF.

12

4. Results and discussion

4.1. Characterization results

Figure 3 shows the Raman and TGA of Graphene, PMMA and graphene/PMMA composite. The structures of graphene, PMMA, and graphene/PMMA are well confirmed through Raman study as shown in Figure 3(a). Raman spectra gives the characteristic peaks for graphene. Generally, in graphene D band is order of defects states and G band originates from C-C stretching. In case of graphene, the position of disordered state D band and C-C stretching G band overlapped with D originates from the Stokes' scattering by a longitudinal optical phonon are present at 1325.12 cm^{-1}, 1576.50 cm^{-1} and 1620.12 cm^{-1} respectively. The characteristic second-order two-phonon 2D band appeared at 2682 cm^{-1} which strongly depends upon the frequency of laser energy. The G and D modes are due to the localized vibrational modes of the impurities, which interacts with extended phonon modes of graphene. Generally, for bilayer graphene the typical values of I_D/I_G and I_{2D}/I_G ratios are in the range of 0.1-0.3 and 0.4-1 respectively [53,54]. In case of graphene, the I_D/I_G and I_{2D}/I_G ratios are observed around 0.17 and 0.4 confirm the bilayer nature of graphene (n~2). Similarly, in case of graphene/PMMA, the D band, G band and 2D band positions are at 1335.81 cm^{-1}, 1583.10 cm^{-1}, 2632.83 cm^{-1} respectively. Generally, defects states intensity is measured by I_D/I_G ratio and it increases with increase in disorder in the sample. From Figure 3(a), in case of graphene/PMMA it is observed that the intensity of the D state is significantly increased compared with D state of graphene. The corresponding I_D/I_G ratio of around $I_D/I_G = 0.17$ for graphene and $I_D/I_G = 1.4$ for graphene/PMMA composite respectively. It has been reported that the significant increase in I_D/I_G ratio may be due to incorporation of graphene in PMMA matrix in which the defects present on the surface of graphene could promote reaction sites between graphene and polymer chains [55,56]. This indicates that the increase in disorder in graphene/PMMA is attributed due

to the high edge content [45]. The PMMA peak observed at 2950 cm^{-1} in the Raman spectrum of PMMA (inset shows zoomed peak) which is also exist in Raman spectra of graphene/PMMA [57]. In graphene/PMMA, the broadening in G and 2D band is observed may be due to the presence of PMMA which can generate some residual strain in graphene [57,58]. The increased in defect states are also confirmed from the FWHM values of 2D band of graphene (FWHM = 71) and graphene/PMMA (FWHM = 112) [59,60]. The defects states play an important role in deciding the sensor performance.

To find the amount of graphene in graphene/PMMA nanocomposite, the TGA study was conducted as shown in Figure 3(b). It is confirmed that ≈ 25% of graphene is present in PMMA matrix. The possibility of any interaction is estimated through FTIR study (Figure S2). The PMMA shows its signature peaks of C=O and C-H stretching bands which are appeared at 1722 cm^{-1} and 2957 cm^{-1} respectively [61]. The graphene shows its characteristic peaks at 1585 cm^{-1} corresponds to C=C stretching band. The other functional groups such as C=O band and C-O band are present at 1720 cm^{-1} and 1130 cm^{-1} respectively. In case of graphene/PMMA composite no peak shift is observed. The characteristic peaks which are originally present in PMMA are also present in graphene/PMMA composite.

The morphology and fiber-tip coated section of sensing material is studied by taking FE-SEM images of PMMA, Graphene and graphene/PMMA composite fiber-coated samples as shown in Figure 4. The FE-SEM images of graphene, PMMA and graphene/PMMA directly obtained from fiber-tip coated samples are shown in Figure 4(a-i). From Figure 4(a, b), it is observed that the graphene possesses more than one layer. The Figure 4(d, e) depicts the polymeric nature of PMMA. The FE-SEM image of graphene/PMMA (shown in Figure 4(g, h) shows the homogeneous mixing of Graphene in PMMA matrix. The uniform layer of PMMA on graphene is observed. The FE-SEM images of fiber-optic coated probes for graphene, PMMA and

graphene/PMMA are shown in Figure 4(c, f, i). The inset of Fig. 4(c, f, i) shows an enlargement of coated optical fiber probes with measured film thicknesses of ~65 μm, ~75 μm, and ~85 μm for graphene, PMMA, and graphene/PMMA respectively. These characterization results support in describing the experimental results shown by these sensing probes under the exposure of various volatile gases.

4.2. Response of G-DFPI sensor under the exposure of various vapors

The response of the sensor is demonstrated in Figure 5 at different vapor concentration in terms of wavelength shift and power shift respectively. The G-DFPI shows better sensing response towards methanol vapor and therefore, here only the methanol spectrum is depicted (Figure S3, spectrum response for other vapors). Figure 5(a) shows the interference spectrum response of G-DFPI for methanol vapor in concentration from 5 μL/L to 200 μL/L. It shows that as the vapor concentration is increased from 5 μL/L to 200 μL/L, the corresponding resonance dip wavelength position is also shifted (as shown in zoomed graph of Figure 5(b)). The highlighted resonance dip position of the sensor is 1549.15 nm at 0 μL/L concentration and the equivalent spectral shift is measured from this spectral position under exposure of varied concentration of methanol. The dip wavelength changes from 1549.15 nm 1548.37 nm for methanol concentration varied from 5 μL/L to 200 μL/L. Similarly, for other dip wavelength region, the wavelength position is shifted from 1534.90 nm to 1534.12 nm and from 1563.80 nm to 1563.03 nm respectively. It has been confirmed that the total wavelength shift ($\approx$ 0.78 nm) observed at specific dip wavelength is nearly same at all dip wavelength in the interference spectrum. Figure 5(c) shows the wavelength shift response of G-DFPI for different vapors concentration ranging from 5 μL/L to 200 μL/L. It is detected that the wavelength shift follows Langmuir dependence with respect to change in vapors concentration. At 200 μL/L, the

maximum wavelength shift is observed for methanol followed by acetone, chloroform, ethanol, toluene and isopropanol. The sensitivity for these vapors is calculated from the linear fitting. It appears that the wavelength shift follows linear response up to 50 µL/L of vapor concentration and then trails saturation. Therefore, their corresponding wavelength sensitivity values with standard error calculated by linear fitting are of the order of 3.57 ± 0.001 pm/(µL/L), 2.81 ± 0.0004 pm/(µL/L), 3.62 ± 0.0008 pm/(µL/L), 6.14 ± 0.0004 pm/(µL/L), 1.86 ± 0.0007 pm/(µL/L) and 1.76 ± 0.00008 pm/(µL/L) for methanol, acetone, chloroform, ethanol, toluene and isopropanol respectively. The corresponding LOD of G-DFPI for methanol is observed around 5.6 µL/L. The response of G-DFPI is also measured in terms of power shift as well. The response of power shift Vs vapor concentration is shown in Figure S3(a). The G-DFPI sensor shows maximum power shift for chloroform followed by toluene, ethanol, isopropanol, acetone and methanol respectively. The G-DFPI sensor exhibits maximum power shift linear response (up to 50 µL/L) towards chloroform with obtained sensitivity of 0.0013 ± 0.0002 dB/(µL/L).

At 200 µL/L, G-DFPI shows high response towards methanol, this might be due to highly polar nature of methanol. The readily donated electron to the graphene surface changes the carrier electron density of graphene and modifies its electrical conductivity and thus the refractive index [18,23].

4.2.1. Sensing mechanism of Graphene with VOCs

Figure 6 shows sensing mechanism of G-DFPI with different vapors. It is well-known that the defect sites play an important role in gas sensing. Generally, as-prepared graphene contains many oxidation defect sites (C=O), (C-O), (O-C=O) and (O-H)) during synthesis or transfer process, which interacts with vapors through weak hydrogen bonding or van der walls interaction [45]. The interaction of VOCs molecules with graphene is strongly depending upon these

16

oxidation defects present on the surface of graphene. It is expected that the adsorbed oxygen atoms on graphene surface take negative charge by capturing an electron and acts as donors or acceptors depending upon the nature of the VOCs. Therefore, this will change the electron carrier density of graphene which results in change in its refractive index. Thus, it further alters the dip wavelength in an interference spectrum [23,59].

Experimentally, we observed that the graphene offers highest sensitivity in terms of wavelength shift towards methanol. Generally, the polar molecule possesses large dipole moment and hence, induces strong molecular dipole-induced charge density modulation on the surface of graphene, which strongly modifies the refractive index of graphene [59]. Therefore, the VOCs can be arranged as per polarity order viz: methanol > ethanol > acetone > isopropanol > chloroform > toluene. Methanol is highly polar in nature and therefore, can hold the charge moving along graphene layer, resulting in a large increase in electrical resistance and therefore, shows high sensing response compared with other vapors. The electronic properties of graphene get altered by charge transfer due to polar molecules adsorption. The polar molecule such as methanol can form hydrogen bond with the oxygen functionalities present on the graphene surface and hence shows maximum response [59,60,62]. Toluene is less polar and hence obtain very low response. Thus, in case of graphene the polarity plays crucial role in deciding the sensor response. However, there might be certain factors such as vapor pressure and boiling point of the VOCs are also very decisive in vapor sensing [23,63]. The graphene also showed good wavelength sensitivity toward acetone and chloroform. The reason might be the high vapor pressure and low boiling point of acetone and chloroform. The vapors having low boiling point and high vapor pressure tend to physically adsorb more on the graphene surface which further modulates the refractive index of the graphene. The reducing vapors like acetone readily reacts with oxygen vacancies of graphene and modifies the dielectric constant by changing electron

17

carrier density on the surface and therefore, can induce the wavelength shift in the interference spectrum [23,64].

4.3. Response of P-DFPI sensor under the exposure of various vapors

Here, the P-DFPI spectral response towards acetone vapor in concentration ranging from 5 µL/L to 200 µL/L as shown in Figure 7(a) (Figure S4, spectrum response for other vapors). The resonance dip shift with respect to volatile gas concentration is displayed in Figure 7(b). The response of sensor is measured at dip wavelength position at 1553.96 nm (in air) toward various acetone concentration and observed that position of dip wavelength changes as the concentration varies from 5 µL/L to 200 µL/L. In the P-DFPI interference pattern, the position of dip wavelength is varied from 1541.16 nm to 1540.17 nm for the first interference dip and from 1553.96 nm to 1555.06 nm for second interference dip respectively. The total wavelength shift (around 1.0 nm) is nearly same for both dips in the fringe pattern. The sensor shows highest sensitivity in terms of wavelength shift for acetone followed by chloroform, ethanol, toluene, methanol and isopropanol. For P-DFPI, the sensor response is also well fitted with Langmuir model (as shown in Figure 7(c)). It observed that the P-DFPI shows linear response for acetone vapor up to 100 µL/L and then follows saturation. But for other vapors the linear response is up to 50 µL/L and then it follows saturation. Therefore, the wavelength shift sensitivity values with standard error (calculated up to linear fitting) is obtained as 5.96 ± 0.0008 pm/(µL/L) for acetone up to 100 µL/L, and 7.1 ± 0.0035 pm/(µL/L), 1.48 ± 0.0004 pm/(µL/L), 1.52 ± 0.0001 pm/(µL/L), 1.29 ± 0.0002 pm/(µL/L), 0.85 ± 0.0001 pm/(µL/L) at 50 µL/L for chloroform, ethanol, toluene, methanol and isopropanol respectively. The corresponding LOD of P-DFPI for acetone is observed around 3.35 µL/L. The Langmuir R-square fitting value of wavelength shift for all vapors between 5 µL/L to 200 µL/L were observed as 0.98, 0.96, 0.93, 0.99, 0.98, and

0.97 respectively. Similarly, the response of P-DFPI sensor is measured in terms of power shift with different vapors concentration as shown in Figure S4(a). The maximum power shift is obtained for chloroform and least for isopropanol. The corresponding sensitivity of P-DFPI sensor towards chloroform (linear fit up to 50 μL/L) is obtained as 0.00711 ± 0.00183 dB/(μL/L) and minimum of 0.0011 ± 0.0001 dB/(μL/L) for isopropanol.

From results, it is observed that the maximum response of P-DFPI is observed towards acetone and chloroform in terms of wavelength shift and power respectively. The PMMA is well-known for its swelling property which changes the FP cavity length significantly and modulates the dip wavelength in an interference pattern [65]. Therefore, theoretically as per the Hansen solubility parameters (HSP) model, the PMMA should exhibits high response towards acetone and chloroform. The results obtained from P-DFPI are in agreement with theoretical values predicted by Hansen model. The possible trend of swelling of PMMA under the exposure of various vapors is reported in Table 2.

4.3.1. Sensing mechanism of PMMA with VOCs

The PMMA is very attractive material for sensing due to its swelling ability when it interacts with certain vapors. The swelling of PMMA is strongly depending upon the nature of VOCs. As shown in Fig. 8, the PMMA possesses repetitive units of ester pendant group containing =O which actively participate in the interaction which results into the swelling; which further changes the cavity length and induces the wavelength shift in an interference spectrum [62,65,66]. However, this distortion is reversible when vapor encounters with air.

The interaction of PMMA with various vapors could be clearly understood with the help of a model proposed by Hansen. This model gives the preliminary idea about the impact of swelling and its role in sensor responsivity. According to the HSP model, the swelling of any type of

polymer is dependent upon the Hansen solubility parameters. The swelling of polymer depends upon the relative energy difference (RED) between polymer and vapor. When the value of RED <1, which shows that the polymer is more miscible. So, more swelling occurs when the value of swelling ratio is below one and vice versa [59,60,67]. In this research work, we tabulated the values of different HSP parameters of polymer and VOCs. It shows that in case of acetone the RED value is around 0.46 which is below one and close to zero and hence, it will more interacts with PMMA. The other VOCs such as chloroform and toluene are showing RED value below one hence, it may cause significant swelling. However, the response is less than acetone due to high RED value. Other VOCs such as isopropanol, ethanol and methanol, the RED is greater than one and therefore, it is less miscible to PMMA. These vapors may cause some swelling which could be comparatively much lesser than acetone, chloroform and toluene. Therefore, it is confirmed that the solvents which are good solvents to PMMA are more miscible and therefore, give high response compared with less soluble solvents. The more swelling tends to increase in volume and hence, the cavity length and therefore, shows high wavelength shift in the interference spectrum [59,60,66,68].

From Table 2, it is confirmed that the PMMA is more sensitive to acetone followed by chloroform, toluene, isopropanol, ethanol and methanol. However, experimentally, we observed that P-DFPI shows highest response in terms of wavelength shift towards acetone followed by chloroform, ethanol, toluene, methanol and isopropanol. Thus, observed trend is found nearly in agreement with trend predicted by HSP model.

4.4. Response of GP-DFPI sensor under the exposure of various vapors

Figure 9(a, b, c, d) shows the GP-DFPI sensor response in terms of interference spectrum for acetone, close view of resonance dip response towards acetone vapor, response in terms of

20

wavelength shift and power shift towards various vapors (Figure S5, spectrum response for other vapors). For GP-DFPI sensor, we observed overall increased in sensitivity towards all vapors in terms of wavelength shift and power shift respectively compared with individual responses of G-DFPI and P-DFPI. As shown in Fig. 9(a), the GP-DFPI exhibits good response towards acetone in 5 µL/L and 200 µL/L. The close view of change in resonance dip shift towards acetone vapor in concentration range from 5 µL/L and 200 µL/L is shown in Fig. 9(b). Initially, the dip wavelength position is at 1552.83 nm in air and the corresponding change in spectrum dip position is recorded in presence of acetone vapor concentration varied from 5 µL/L to 200 µL/L. As shown in the interference pattern (shown in Fig. 9(a)), the dip wavelength position is changed from 1529.89 nm to 1531.78 nm for first dip, 1537.43 nm to 1539.37 nm for second dip, 1545.08 nm to 1547.04 nm for third dip, 1552.83 nm to 1554.80 nm for fourth dip, and 1560.57 nm to 1562.51 nm for fifth dip respectively. Therefore, the total wavelength shift is nearly identical of around 1.94 nm at all dips in the interference pattern. For other vapors, the total wavelength shift for a specific vapor is also nearly same for all dips in their respective interference pattern. The highest response of GP-DFPI in terms of wavelength shift is observed towards acetone (shown in Figure 9(c)) and followed by toluene, methanol, chloroform, ethanol and isopropanol respectively. In case of GP-DFPI, we detected that the sensor shown linear dependence to acetone vapor at low pressure region and follows saturation at higher concentration suggesting Langmuir type behavior. Interestingly, GP-DFPI shown much increased in linear response towards these vapors compared with linear response shown by G-DFPI and P-DFPI individually. The GP-DFPI shows linear response concentration up to 150 µL/L for acetone and up to 100 µL/L for other vapors. For GP-DFPI, the wavelength sensitivity values with standard error calculated from the linear fitting of the response curves are of 11.62 ± 0.0008 pm/(µL/L) toward acetone (linear fitting up to 150 µL/L), and 3.71 ± 0.0005 pm/(µL/L), 3.20 ± 0.0006 pm/(µL/L),

4.14 ± 0.0002 pm/(μL/L), 2.47 ± 0.0009 pm/(μL/L), 1.88 ± 0.0002 pm/(μL/L) up to 100 μL/L for toluene, methanol, chloroform, ethanol and isopropanol respectively. The calculated LOD of GP-DFPI for each VOCs is 1.72 μL/L, 5.39 μL/L, 6.25 μL/L, 4.83 μL/L, 8.09 μL/L and 10.63 μL/L respectively. The Langmuir curve R-square fitting for wavelength shift with concentration of all VOCs were observed around 0.99, 0.91, 0.97, 0.95, 0.99, 0.92 respectively. The error bar shown in sensitivity values and response curves are the standard deviation in the wavelength dip obtained while performing the experiments multiple times. This shows that these sensing probes are well repeatable.

The GP-DFPI sensor also exhibits power shift under the exposure of different VOCs (shown in Figure S5(a)). The highest response is observed for acetone followed by methanol, toluene, ethanol, chloroform and isopropanol respectively. The sensor shows linear response up to 100 μL/L with corresponding power sensitivity of the order of 0.0107 ± 0.0013 dB/(μL/L), 0.0051 ± 0.0010 dB/(μL/L), 0.0027 ± 0.00009 dB/(μL/L), 0.0054 ± 0.0023 dB/(μL/L), 0.0045 ± 0.0015 dB/(μL/L) and 0.0024 ± 0.0010 dB/(μL/L) respectively. The corresponding Langmuir curve R-square fitting of power shift with concentration of all VOCs were observed around 0.94, 0.96, 0.96, 0.99, 0.95, 0.98 respectively.

The calibration curve shows that at higher gas concentration the curve becomes saturated. This is due to the limited number of active sites and finite surface area of sensing film which is exposed to volatile gas. At higher gas concentration, the decreasing number of active sites which leads to less change in the effective refractive index of sensing layer and hence shows less change in the wavelength shift. This is well agreement with theory discussed by equations (1-4). All the curves are well fitted by simple Langmuir adsorption equation. Similar Langmuir-type behavior have also been observed in various gas sensors studies based on ZnO-PDMS [69], SnO$_2$ [70], Pd-Si [71] and CNTs-PMMA-PS [72].

Thus, from experimental results it is confirmed that the GP-DFPI exhibits highest sensitivity comparatively than G-DFPI and P-DFPI towards various vapors concentration ranging from 5 µL/L to 200 µL/L respectively. In case of GP-DFPI, the change in wavelength shift is attributed to the change in the refractive index and swelling of PMMA polymer resulting into the enhanced in the sensitivity [18,23,59,60,65,66]. The sensing mechanism behind this enhanced in sensitivity is elaborately explained in below section.

4.4.1.   Sensing mechanism of graphene/PMMA with VOCs

The graphene/PMMA composite shows highest sensitivity in terms of wavelength shift and power shift under the exposure of different VOCs. Therefore, to understand the sensing mechanism of graphene/PMMA towards VOCs is very important in application point of view.

It has been reported in the literature that graphene and PMMA composite is formed in such way that the functional groups present on graphene surface can physically interacts with PMMA through hydrogen bonding which provide the well adhesion in PMMA matrix [59,62,68,73]. In case of graphene/PMMA composite, the sensing mechanism can be explained based on polymer swelling and HSP model. When VOC adsorbed on the surface of graphene/PMMA, the PMMA polymer coating becomes swollen which resulting into increase in polymer volume. This phenomenon leads to bending of graphene sheets and hence, increase in contact resistance [59,60,62,68,74]. The amount of swelling is depending upon the solubility of polymer in the solvent. Thus, greater the swelling, more change in the volume and hence, greater the response of the sensor [59,60,62,68]. Therefore, the significant swelling of PMMA (cavity length changes), and vapors interaction with the graphene surface (refractive index change) which results in increase in distance between graphene sheets and hence, the contact resistance could significantly modulates the dip wavelength in an interference pattern, which is also observed in

23

some nanocomposite based gas sensors [60,62,65,74]. However, swelling of PMMA polymer is more predominant than change in dielectric constant of graphene under the exposure of VOCs. Also, according to HSP model, when the RED value is small it will more interactive with the polymer and hence, there will be more response. Therefore, as shown in table 2, the acetone has small RED and hence, interacts more compared with other vapors.

Experimentally, we observed that the graphene/PMMA composite system shows highest sensitivity response in terms of wavelength shift towards acetone, methanol and chloroform. Therefore, the observed results are well co-related with proposed theory. Graphene shows significant shift towards methanol and acetone. The PMMA shows major shift towards acetone. The graphene/PMMA shows enhanced sensitivity towards acetone and methanol. In graphene/PMMA composite, the PMMA wt% is more and therefore, it is obvious that the PMMA will contribute more in the sensing. Thus, as per the mathematical theory (equation 1, 4) and Hansen model, it is expected that the graphene/PMMA could show significant shift towards acetone and methanol. Depending upon the type of vapor either polar/non-polar, or soluble/insoluble to PMMA; once it adsorbs on the surface on the graphene/PMMA, there might be the substantial swelling of PMMA (cavity length) of the sensing material, and hence, more change in the dielectric constant of graphene/PMMA composite which significantly modulates the sensor characteristics.

As per the Hansen parameters shown in Table 2, the order of REDs for graphene/PMMA is observed as: acetone < chloroform < toluene < isopropanol < ethanol < methanol. Therefore, according to HSP model, the graphene/PMMA composite shows high response towards acetone. Here, those vapors are good solvents to PMMA have shown better response compared with those are insoluble or less soluble. The acetone, chloroform, and toluene are good solvents for PMMA and therefore, shown high response. The solvents like methanol, isopropanol, and ethanol in

which the PMMA is less soluble shown less response. Therefore, the response is occurred irrespective to the polar nature of the solvents and therefore, PMMA swelling is the predominant factor in enhancing the sensing response for graphene/PMMA composite [59,60,68]. Along with these, we need to consider certain other factors such as molecular size, dielectric constant and vapor pressure which might be contributing to enhanced in sensitivity for graphene/PMMA case [59,60,75]. The mechanism based on vapor pressure can also explain these experimental results. As per the vapor pressure values, the order of these vapors is arranged as: acetone > chloroform > methanol > ethanol > isopropanol > toluene. It is well-known that higher the vapor pressure, higher is the diffusion and hence more response is expected. Therefore, in case of graphene/PMMA, high vapor pressure and low molecular size of vapors may have greater effects than other factors and thus shows good response towards acetone [59,60,68]. The response trend is much like that as explanation provided above. However, one should note that the exact trend may not be achieved due to close values of vapor pressure, RED values, and dielectric constant of these vapors.

4.5. Sensing response of G-DFPI, P-DFPI and GP-DFPI towards mixture of VOCs

The study of detection of individual sensor towards the mixture of VOCs is also performed. In case of VOCs mixture, it's difficult to calculate the actual concentration of mixture in ppm due to their varied vapor pressure, solubility parameters, ebullition temperature and therefore, the concentration of VOCs are considered in units of μL/L. The GP-DFPI sensor exhibits maximum response towards methanol and therefore, its spectral response is only shown in Figure 11(a) along with its zoomed resonance dip as shown in Figure 11(b) (Figure S6(a) for G-DFPI, Figure S7(a) for P-DFPI, Figure S8(a) for GP-DFPI spectrum response). Figure 11(c, d) shows the sensing response of all sensors in terms of wavelength shift and power shift under the exposure

of different VOCs mixture concentration ranging from 5 µL/L to 200 µL/L. The GP-DFPI shows significant wavelength shift response towards the mixture of VOCs in concentration ranging from 5 µL/L to 200 µL/L (Figure 11(c)). All the characteristic responses (wavelength shift, power shift) with vapor mixture concentration are fitted by Langmuir isotherm curve.

The GP-DFPI shows wavelength sensitivity (calculated from the linear fitting up to 150 µL/L) of the order of 2.4 ± 0.00008 pm/(µL/L) followed by 1.59 ± 0.0002 pm/(µL/L) and 1.4 ± 0.0002 pm/(µL/L) shown by P-DFPI and G-DFPI towards the mixture of VOCs varied from 5 µL/L to 200 µL/L. Similarly, P-DFPI, G-DFPI and GP-DFPI shows power shift linear response up to 50 µL/L in Fig. 11(d). Therefore, the corresponding power sensitivity for these sensing probes towards the mixture of VOCs is around 0.0015 ± 0.0001 dB/(µL/L), 0.00061 ± 0.0002 dB/(µL/L) and 0.00085 ± 0.0001 dB/(µL/L) respectively.

4.6. Sensing response of G-DFPI, P-DFPI and GP-DFPI towards binary mixtures

Additionally, the response (in terms of wavelength shift) of all sensors towards the binary mixture (shown in Figure 12) also measured. (Figure S6(b-e) for G-DFPI, Figure S7(b-e) for P-DFPI, Figure S8(b-e) for GP-DFPI spectrum response). Generally, G-DFPI, P-DFPI and GP-DFPI sensors shows significant response under the exposure of acetone and therefore, sensors response is measured with binary mixture of acetone and other VOCs in concentration ranging from 5 µL/L to 200 µL/L. The four types of binary mixtures such as acetone + ethanol, acetone + methanol, acetone + chloroform and acetone + toluene was prepared as described in experimental section. Figure 12 shows the response fitted by Langmuir isotherm of G-DFPI, P-DFPI and GP-DFPI in terms of wavelength shift towards these mixtures. It is observed that the GP-DFPI (shown in Figure 12(a)) sensor probe shows better sensing response compared with G-DFPI (shown in Figure 12(b)) and P-DFPI (shown in Figure 12(c)) under the binary mixture

concentration varied from 5 µL/L to 200 µL/L. At concentration of 200 µL/L, maximum wavelength shift shown by each sensing probe is around 1.07 nm for acetone + toluene by GP-DFPI as that of 0.33 nm for acetone + methanol, and 0.4 nm for acetone + ethanol by P-DFPI respectively.

However, detection of specific vapor from VOCs mixture is very challenging and has not yet resolved. It is very complex phenomenon and depends upon many factors such as rate of diffusion, vapor pressure, Hansen solubility parameters, boiling point, density of mixtures, thickness of the film, operating temperature and charge carrier concentration [75–77].

4.7. Selectivity response of individual G-DFPI, P-DFPI and GP-DFPI sensors towards VOCs

In Figure 13, the selectivity study of individual sensing probe at 200 µL/L for all vapors is depicted. As three type of sensing probes are studied for different exposure of VOCs gas. Therefore, their responses are also different based on the interaction of material property and nature of gas. Thus, we need to consider selectivity of individual sensor towards the exposure of VOCs. The GP-DFPI shown high selectivity towards acetone and least towards isopropanol. The P-DFPI shows good response towards acetone and very small response to isopropanol gas. For G-DFPI, the sensor exhibits well selectivity to methanol and least towards isopropanol. This study shows that the GP-DFPI sensing probe gives good response (in terms of wavelength shift) towards different vapors compared with individual responses shown by P-DFPI and G-DFPI.

4.8. Sensor's time response characteristics

All sensors responses were measured on OSA and their sensitivity is calculated in terms of wavelength shift respectively. Under different exposure of VOCs, all sensors show immediate

response. Here, we just plotted each sensor response towards acetone only. As shown in Figure 14(a), the response of each sensor is plotted for acetone at concentration range from 5 µL/L to 200 µL/L. It is observed that the estimated response time is few tens of seconds and recovery time is in few seconds. The response of all sensors is presented for acetone at concentration of 5 µL/L to 200 µL/L as shown in Figure 14(b, c). The recovery of each sensor was done in open air at room temperature. The recovery time of sensor was measured in between its saturation value to its initial position. However, it is observed that for all sensors the response time of the order of 48-52 seconds compared with recovery time of just 8-12 seconds. Here, we confirmed that for all sensors, their responses are repeatable over the concentration range of all vapors. The repeatability response is only shown for GP-DFPI due to its higher response towards acetone gas. From Figure 14(d), it is confirmed that for GP-DFPI sensor response in terms of wavelength shift is well repeatable over three cycles of acetone vapor at 200 µL/L. The measured average error in the repetition measurement was observed around 0.1 nm. Therefore, it also confirms the reusability for the detection of VOCs gas. In this study, the sensing probes were prepared by using very simple and cost-effective dip coating technique. Thus, it is very easy to reproduce the sensing probe in short duration of time by monitoring its spectral characteristics on OSA. The GP-DFPI sensor long-term stability is shown in Figure S9. We observed that even after several days, the graphene/PMMA coated sensor interference pattern remains same signifying its good stability (as shown in Figure S9(a)). The sensor response toward 200 µL/L acetone vapors for 10 days is shown in Figure S9(b). It shows that the sensor possesses good long-term stability.

The average time of all sensors to acetone gas is shown in Figure 14(e). To understand the response and recovery time of the sensors in more detail, we plotted their time response individually towards acetone at concentration varied from 5 µL/L to 200 µL/L for G-DFPI, P-DFPI and GP-DFPI as shown in Figure 14(f). From Figure 14(f), it is noticed that the response

time is little longer than the recovery time. Generally, the response/recovery time of the sensor is determined by the thickness of the sensing film and diffusion coefficient of vapor into the sensing film. Therefore, for much larger thickness the recovery time will be much faster than the response time [76,77]. This may because of most of the vapors possess high vapor pressure at room temperature [3]. When it encounters air, the diffused vapor inside the sensing film quickly come outside and hence sensor shows fast recovery.

5. Sensor Algorithm in terms of wavelength shift and power shift with temperature variation:

The sensor response is expressed in terms of wavelength shift and power shift respectively. All sensing probes were reusable and their responses were reproducible. Therefore, it is necessary to express their correlation in more simplified form. As the vapors purged inside the gas chamber, it adsorbs on the sensing material and exchange the electrons with sensing material. The refractive index or the cavity length changes due to interaction and hence, it reflects in terms of wavelength shift in the interference spectrum. The temperature response of the sensor is also conducted and found that all the sensing probes possessed linear dependence with temperature (shown in Figure S10).

Thus, the correlation of wavelength shift and power shift with respect to change in vapor concentration and temperature is expressed in terms of second order of matrix algorithm and is expressed as:

$$\begin{bmatrix} \Delta C \\ \Delta T \end{bmatrix} = M^{-1} \begin{bmatrix} \Delta P \\ \Delta \lambda \end{bmatrix} \tag{5}$$

Where, $M^{-1} = \frac{1}{[A_{22}A_{11} - A_{21}A_{12}]} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$ and A_{11}, A_{12}, A_{21} and A_{22} are the matrix coefficients.

The response of G-DFPI, P-DFPI and GP-DFPI with temperature change from 30 ^{0}C to 70 ^{0}C is measured (Figure S10). We observed that the G-DFPI exhibits high temperature sensitivity of around 0.21 nm/ ^{0}C. The P-DFPI and GP-DFPI shows temperature sensitivity of the order of 0.068 nm/ ^{0}C and 0.12 nm/ ^{0}C respectively. PMMA owns high thermal expansion coefficient and therefore, shows response towards temperature and widely used for temperature based applications [78,79]. Graphene possess good thermal conducting property and hence, explored for temperature measurement as reported elsewhere [80,81]. Therefore, these sensing probes could be suitable for simultaneous measurement of VOCs concentration and temperature.

As we observed maximum sensitivity for GP-DFPI towards acetone and therefore, the sensitivity response of GP-DFPI is expressed in following matrix algorithm. Thus, the sensitivity coefficients of GP-DFPI in terms of power shift and wavelength shift under exposure of acetone concentration and temperature change are given as A_{11}= 0.010 dB/(μL/L), A_{12}= 0.011 dB/^{0}C for power shift and A_{21}=11.62 pm/(μL/L) and A_{22}= 120 pm/ ^{0}C for wavelength shift respectively. Therefore, equation (5) can be modified as follows:

$$\begin{bmatrix} \Delta C \\ \Delta T \end{bmatrix} = \begin{bmatrix} 111.92 & -0.0102 \\ -10.837 & 0.0093 \end{bmatrix} \begin{bmatrix} \Delta P \\ \Delta \lambda \end{bmatrix} \qquad (6)$$

Equation 6 could be used as a performance equation of the developed sensor for the acetone gas under different variables such as temperature and concentration of vapor.

In summary, we proposed a sensor based on DFPI configuration for monitoring different type of vapors. This sensor has significant features include a utilization of micro-fiber properties which modulate the interference spectrum, use of versatile nanocomposite and fast air-recovery obtained at room temperature. The obtained results concluded that the GP-DFPI shown much enhanced sensitivity as compared with G-DFPI and P-DFPI. However, an optimization of graphene in PMMA matrix is very important in device point of view. Therefore, further experiments on different concentration ratio of graphene/PMMA composite need to be

performed to find out optimized ratio to enhance the sensors performance and this will be our upcoming study. PMF is temperature sensitive due to its varied refractive index profile and therefore, this type of sensor could be used for simultaneous monitoring of gas concentration and temperature. Although, there is much scope to enhance this sensor performance further. Additional work could be done in the field of excitation of higher order modes by etching or coating the nanomaterial onto the PMF surface with varied PMF length. The response time of the sensor could be decreased by decreasing the film thickness of the coated material. The detection limit of the sensor can be increased by using high resolution OSA or spectrometer.

6. Conclusions

We have demonstrated a new and sensitive DFPI approach for VOCs monitoring in the concentration varied from 5 µL/L to 200 µL/L at room temperature. The sensor head comprises SMF-PMF, which is tip coated with graphene, PMMA, and graphene/PMMA composite for various VOCs monitoring. Owing to their superior properties such as variation in refractive index of graphene, swelling of PMMA polymer and simultaneous change in their refractive index and swelling of graphene/PMMA composite makes it more attractive for sensing various kinds of VOCs such as acetone, chloroform, toluene, methanol, ethanol and isopropanol respectively. All sensors response follows Langmuir-type dependence and be dependent upon the properties of sensing material and analyte. For all sensors, the average response and recovery time was around 48-52 sec and 8-12 sec which is much better than some reported sensors in Table 1. The details regarding sensing mechanism for each sensing material with different vapors along with mathematical formulation and HSP model is elaborately studied and presented. This nanomaterial-modified DFPI structure is very compact, reproducible and sensitive to VOCs. The

VOCs sensor by using this proposed configuration will be fabricated as a more accurate and efficient real-time gas monitoring.

Supporting Information. Experimental set up (**Figure S1**), FTIR spectra of materials (**Figure S2**), Power shift and spectral response of G-DFPI for various vapors (**Figure S3**), Power shift and spectrum response of P-DFPI sensor towards various vapors (**Figure S4**), Power shift and spectral response of GP-DFPI for various vapors (**Figure S5**), Response of G-DFPI for binary mixture of VOCs (**Figure S6**), Response of P-DFPI towards binary mixture of VOCs (**Figure S7**), Response of GP-DFPI towards binary mixture of VOCs (**Figure S8**), Long-term stability study of GP-DFPI, (**Figure S9**), Temperature study of G-DFPI, P-DFPI and GP-DFPI (**Figure S10**).

Notes: The authors declare no competing financial interest.

Acknowledgements:

References

[1] V. Soni, P. Singh, V. Shree, V. Goel, Effects of VOCs on Human Health, in: N. Sharma, A.K. Agarwal, P. Eastwood, T. Gupta, A.P. Singh (Eds.), Air Pollut. Control. Energy, Environ. Sustain., Springer Singapore, Singapore, 2018: pp. 119–142. https://doi.org/10.1007/978-981-10-7185-0_8.

[2] I. Oakley-Girvan, S.W. Davis, Breath based volatile organic compounds in the detection of breast, lung, and colorectal cancers: A systematic review., Cancer Biomark. 21 (2017) 29–39. https://doi.org/10.3233/CBM-170177.

[3] T. Salthammer, Very volatile organic compounds: An understudied class of indoor air pollutants, Indoor Air. 26 (2016) 25–38. https://doi.org/10.1111/ina.12173.

[4] M.J. Wilde, R.L. Cordell, D. Salman, B. Zhao, W. Ibrahim, L. Bryant, D. Ruszkiewicz, A. Singapuri, R.C. Free, E.A. Gaillard, C. Beardsmore, C.L.P. Thomas, C.E. Brightling, S. Siddiqui, P.S. Monks, Breath analysis by two-dimensional gas chromatography with dual flame ionisation and mass spectrometric detection – Method optimisation and integration within a large-scale clinical study, J. Chromatogr. A. 1594 (2019) 160–172. https://doi.org/https://doi.org/10.1016/j.chroma.2019.02.001.

[5] J.G. Vallarino, A. Erban, I. Fehrle, A.R. Fernie, J. Kopka, S. Osorio, Acquisition of Volatile Compounds by Gas Chromatography–Mass Spectrometry (GC-MS) BT - Plant Metabolomics: Methods and Protocols, in: C. António (Ed.), Springer New York, New York, NY, 2018: pp. 225–239. https://doi.org/10.1007/978-1-4939-7819-9_16.

[6] and R.J. Susmita Dolai,† Susanta Kumar Bhunia,† Stella S. Beglaryan,§ Sofiya Kolusheva,‡ Leila Zeiri, Colorimetric Polydiacetylene − Aerogel Detector for Volatile Organic Compounds (VOCs), ACS Appl. Mater. Interfaces. 9 (2017) 2891−2898.

https://doi.org/10.1021/acsami.6b14469.

[7]   S. Park, S. Kim, G.J. Sun, C. Lee, Synthesis, Structure, and Ethanol Gas Sensing Properties of In2O3 Nanorods Decorated with Bi2O3 Nanoparticles, ACS Appl. Mater. Interfaces. 7 (2015) 8138–8146. https://doi.org/10.1021/acsami.5b00972.

[8]   M. Kumar, V.S. Bhati, S. Ranwa, J. Singh, M. Kumar, Pd/ZnO nanorods based sensor for highly selective detection of extremely low concentration hydrogen, Sci. Rep. 7 (2017) 1–9. https://doi.org/10.1038/s41598-017-00362-x.

[9]   B. P. Caoa, Z. Yanga, S.T. Navalea, D.Z. , S. Hana, X. Liua, W. Liua, Y. Lua, F.J. Stadlerb, Ethanol sensing behavior of Pd-nanoparticles decorated ZnO-nanorod based chemiresistive gas sensors, (n.d.).

[10] Z. Lou, F. Li, J. Deng, L. Wang, T. Zhang, Branch-like hierarchical heterostructure (α-Fe2O 3/TiO2): A novel sensing material for trimethylamine gas sensor, ACS Appl. Mater. Interfaces. 5 (2013) 12310–12316. https://doi.org/10.1021/am402532v.

[11] R.S.M. S.T. Navalea, b, Z.B. Yanga, Chenshitao Liua, P.J. Caoa, V.B. Patil c, N.S. Ramgird, F.J. Stadler, Enhanced acetone sensing properties of titanium dioxide nanoparticles with a sub-ppm detection limit, (n.d.).

[12] Y. Li, H. Ban, M. Yang, Highly sensitive NH3 gas sensors based on novel polypyrrole-coated SnO2 nanosheet nanocomposites, Sensors Actuators B. Chem. (2015). https://doi.org/10.1016/j.snb.2015.10.078.

[13] K. Haddad, A. Abokifa, S. Kavadiya, B. Lee, S. Banerjee, B. Raman, P. Banerjee, C. Lo, J. Fortner, P. Biswas, SnO 2 Nanostructured Thin Films for Room-Temperature Gas Sensing of Volatile Organic Compounds, ACS Appl. Mater. Interfaces. 10 (2018) 29972–29981. https://doi.org/10.1021/acsami.8b08397.

[14] X. Zhou, S. Lee, Z. Xu, J. Yoon, Recent Progress on the Development of Chemosensors

for Gases, (2015). https://doi.org/10.1021/cr500567r.

[15]   X. Wang, O.S. Wolfbeis, Fiber-Optic chemical sensors and biosensors (2013 − 2015),
        Anal. Chem. 88 (2016) 203–227. https://doi.org/10.1021/acs.analchem.5b04298.

[16]   M. Khalaf AL. Arasu PT, Lim HN, Paiman S, Yusof NA, Mahdi MA, Yaacob, Modified
        plastic optical fiber with CNT and graphene oxide nanostructured coatings for ethanol
        liquid   sensing,   Opt.   Express.   25   (2017)   5509–5520.
        https://doi.org/10.1364/OE.25.005509.

[17]   S. Sridevi., K.S. Vasu, N. Bhat, S. Asokan, A.K. Sood, Ultra sensitive NO2 gas detection
        using the reduced graphene oxide coated etched fiber Bragg gratings, Sensors Actuators,
        B Chem. 223 (2016) 481–486. https://doi.org/10.1016/j.snb.2015.09.128.

[18]   Y. Zhang J, Fu H, Ding J, Zhang M, Zhu, Graphene-oxide-coated interferometric optical
        microfiber   ethanol   vapor   sensor,   Appl.   Opt.   56   (2017)   8828–8831.
        https://doi.org/10.1364/AO.56.008828.

[19]   B. Xu, J. Huang, X. Xu, A. Zhou, L. Ding, Ultra-sensitive NO gas sensor based on the
        graphene oxide coated long period fiber grating, ACS Appl. Mater. Interfaces. (2019).
        https://doi.org/10.1021/acsami.9b14212.

[20]   V. Schroeder, S. Savagatrup, M. He, S. Lin, T.M. Swager, Carbon Nanotube Chemical
        Sensors, Chem. Rev. 119 (2018) 599–663. https://doi.org/10.1021/acs.chemrev.8b00340.

[21]   Y. Zhao, X. Li, X. Zhou, Y. Zhang, Review on the graphene based optical fiber chemical
        and   biological   sensors,   Sensors   Actuators   B.   Chem.   231   (2016)   324–340.
        https://doi.org/10.1016/j.snb.2016.03.026.

[22]   H. Fu H, Jiang Y, Ding J, Zhang J, Zhang M, Zhu Y, Li, Zinc oxide nanoparticle
        incorporated graphene oxide as sensing coating for interferometric optical microfiber for
        ammonia   gas   detection,   Sensors   Actuators   B.   Chem.   254   (2017)   239–247.

https://doi.org/10.1016/j.snb.2017.06.067.

[23] D. Pawar, B.V.B. Rao, S.N. Kale, Fe3O4-decorated graphene assembled porous carbon nanocomposite for ammonia sensing: study using an optical fiber Fabry–Perot interferometer, Analyst. 143 (2018) 1890–1898. https://doi.org/10.1039/c7an01891f.

[24] M.T.K. Tushar Ramesh Pavase, Hong Lin, Qurat-ul-ain Shaikh, Hussain Sameer, Zhenxing Li, Ishfaq Ahmed, Liangtao Lv, Lirui Sun, Syed Babar Hussain Shah, Recent advances of conjugated polymer (CP) nanocomposite-based chemical sensors and their applications in food spoilage detection: A comprehensive review, Elsevier B.V., 2018. https://doi.org/10.1016/j.snb.2018.06.118.

[25] W.S.P. Carvalho, M. Wei, N. Ikpo, Y. Gao, M.J. Serpe, Polymer-Based Technologies for Sensing Applications, Anal. Chem. 90 (2018) 459−479. https://doi.org/10.1021/acs.analchem.7b04751.

[26] P.J. Rivero;, J. Goicoechea;, F.J. Arregui, Optical Fiber Sensors Based on Polymeric Sensitive Coatings, Polymers (Basel). 10 (2018) 1–26. https://doi.org/10.3390/polym10030280.

[27] M. Xu B, Li P, Wang DN, Zhao C, Dai J, Yang, Hydrogen sensor based on polymer-filled hollow core fiber with Pt-loaded WO3/SiO2 coating, Sensors Actuators B. Chem. 245 (2017) 516–523. https://doi.org/10.1016/j.snb.2017.01.206.

[28] X. Ma W, Wang R, Rong Q, Shao Z, Zhang W, Guo T, Wang J, Qiao, CO2 gas sensing using optical fiber Fabry-Perot interferometer based on Polyethyleneimine/Poly(Vinyl Alcohol) coating, IEEE Photonics J. 9 (2017). https://doi.org/10.1109/JPHOT.2017.2700053.

[29] Z. Xiao Y, Yu J, Shun L, Tan S, Cai X, Luo Y, Zhang J, Dong H, Lu H, Guan H, Zhong Y, Tang J, Chen, Reduced graphene oxide for fiber-optic toluene gas sensing, Opt.

Express. 24 (2016) 28290–28302. https://doi.org/10.1364/OE.24.028290.

[30] T. Kavinkumar, D. Sastikumar, S. Manivannan, Eﬀect of functional groups on dielectric, optical gas sensing properties of graphene oxide and reduced graphene oxide at room temperature, RSC Adv. 5 (2015) 10816–10825. https://doi.org/10.1039/C4RA12766H.

[31] S. Manivannan, A.M. Saranya, B. Renganathan, D. Sastikumar, G. Gobi, K. Chang, Single-walled carbon nanotubes wrapped poly-methyl methacrylate fiber optic sensor for ammonia , ethanol and methanol vapors at room temperature, Sensors Actuators B. Chem. 171–172 (2012) 634–638. https://doi.org/10.1016/j.snb.2012.05.045.

[32] T. Gong B, Shi T, Zhu W, Liao G, Li X, Huang J, Zhou, UV irradiation-assisted ethanol detection operated by the gas sensor based on ZnO nanowires / optical fiber hybrid structure, Sensors Actuators B. Chem. 245 (2017) 821–827. https://doi.org/10.1016/j.snb.2017.01.187.

[33] S. Devendiran, D. Sastikumar, Gas sensing based on detection of light radiation from a region of modified cladding (nanocrystalline ZnO) of an optical fiber, Opt. Laser Technol. 89 (2017) 186–191. https://doi.org/10.1016/j.optlastec.2016.10.013.

[34] B. Renganathan, D. Sastikumar, S.G. Raj, A.R. Ganesan, Fiber optic gas sensors with vanadium oxide and tungsten oxide nanoparticle coated claddings, Opt. Commun. 315 (2014) 74–78. https://doi.org/10.1016/j.optcom.2013.10.072.

[35] M. Shabaneh A, Girei S, Arasu P, Mahdi M, Rashid S, Paiman S, Yaacob, Dynamic response of tapered optical multimode fiber coated with carbon nanotubes for ethanol sensing application, Sensors. 15 (2015) 10452–10464. https://doi.org/10.3390/s150510452.

[36] S. Tabassum, R. Kumar, L. Dong, Nanopatterned optical fiber tip for guided mode resonance and application to gas sensing, IEEE Sens. J. 17 (2017) 7262–7272.

https://doi.org/10.1109/JSEN.2017.2748593.

[37]   M. Arasu PT, Noor ASM, Shabanesh AA, Yaacob MH, Lim HN, Mahdi, Fiber Bragg
       grating assisted surface plasmon resonance sensor with graphene oxide sensing layer, Opt.
       Commun. 380 (2016) 260–266. https://doi.org/10.1016/j.optcom.2016.05.081.

[38]   R. Hromadka J, Korposh S, Partridge MC, James SW, Davis F, Crump D, Tatam, Multi-
       parameter measurements using optical fibre long period gratings for indoor air quality
       monitoring,   Sensors   Actuators   B.   Chem.   244   (2017)   217–225.
       https://doi.org/10.1016/j.snb.2016.12.050.

[39]   C. Park, Y. Han, K. Joo, Y.W. Lee, S. Kang, H. Kim, Optical detection of volatile organic
       compounds using selective tensile effects of a polymer-coated fiber Bragg grating, Opt.
       Express. 18 (2010) 24753–24761.

[40]   B.X.U. En, C.L.Z. Hao, F.A.N.Y. Ang, H.U.G. Ong, D.N.W. Ang, J.I.X.I.D. Ai, M.I.Y.
       Ang, Sagnac interferometer hydrogen sensor based on panda fiber with Pt-loaded WO 3 /
       SiO 2 coating, 41 (2016) 1594–1597.

[41]   K.S. Park, Y.H. Kim, J.B. Eom, S.J. Park, M.-S. Park, J.-H. Jang, B.H. Lee, Compact and
       multiplexible hydrogen gas sensor assisted by self-referencing technique, Opt. Express. 19
       (2011) 18190–18198. https://doi.org/10.1364/OE.19.018190.

[42]   C. Wang, G. Yan, Z. Lian, X. Chen, S. Wu, S. He, Hybrid-cavity fabry-perot
       interferometer for multi-point relative humidity and temperature sensing, Sensors
       Actuators   B   Chem.   255   (2018)   1937–1944.
       https://doi.org/https://doi.org/10.1016/j.snb.2017.08.224.

[43]   A.M. Shrivastav, G. Sharma, A.S. Rathore, R. Jha, Hypersensitive and Selective
       Interferometric Nose for Ultratrace Ammonia Detection with Fast Response Utilizing
       PANI@SnO2   Nanocomposite,   ACS   Photonics.   5   (2018)   4402–4412.

https://doi.org/10.1021/acsphotonics.8b00828.

[44] S. Hanf, T. Bögözi, R. Keiner, T. Frosch, J. Popp, Fast and highly sensitive fiber-enhanced Raman spectroscopic monitoring of molecular H2 and CH4 for point-of-care diagnosis of malabsorption disorders in exhaled human breath, Anal. Chem. 87 (2015) 982–988. https://doi.org/10.1021/ac503450y.

[45] U. V Patil, A.S. Pawbake, L.G.B. Machuno, R. V Gelamo, S.R. Jadkar, C.S. Rout, D.J. Late, Effect of plasma treatment on multilayer graphene: X-ray photoelectron spectroscopy, surface morphology investigations and work function measurements, RSC Adv. 6 (2016) 48843–48850. https://doi.org/10.1039/C6RA03046G.

[46] J. Peng, W. Feng, X. Yang, G. Huang, S. Liu, Dual Fabry – Pérot Interferometric Carbon Monoxide Sensor Based on the PANI / Co 3 O 4 Sensitive Membrane-Coated Fibre Tip, Z. Naturforsch. (2018) 1–7.

[47] H. Lu, Y. Yue, J. Du, L. Shao, T. Wu, J. Pan, J. Hu, Temperature and liquid refractive index sensor using P-D fiber structure-based Sagnac loop, Opt. Express. 26 (2018) 18920–18927. https://doi.org/10.1364/OE.26.018920.

[48] S. Xiao, Y. Wu, Y. Dong, H. Xiao, Y. Jiang, W. Jin, H. Li, S. Jian, Simultaneous measurement of refractive index and temperature using SMP in Sagnac loop, Opt. Laser Technol. 96 (2017) 254–258. https://doi.org/https://doi.org/10.1016/j.optlastec.2017.05.016.

[49] C. Zhong, C. Shen, Y. You, J. Chu, X. Zou, X. Dong, Y. Jin, J. Wang, A polarization-maintaining fiber loop mirror based sensor for liquid refractive index absolute measurement, Sensors Actuators B Chem. 168 (2012) 360–364. https://doi.org/https://doi.org/10.1016/j.snb.2012.04.035.

[50] H. Sun, X. Zhang, L. Yuan, L. Zhou, X. Qiao, M. Hu, An Optical Fiber Fabry–Perot

Interferometer Sensor for Simultaneous Measurement of Relative Humidity and Temperature, IEEE Sens. J. 15 (2015) 2891–2897. https://doi.org/10.1109/JSEN.2014.2383387.

[51] N. Can, B.C. Ömür, A. Altındal, Modeling of heavy metal ion adsorption isotherms onto metallophthalocyanine film, Sensors Actuators B. Chem. 237 (2016) 953–961. https://doi.org/10.1016/j.snb.2016.07.026.

[52] Y. Zhao, Y. Yuan, W. Gan, M. Yang, Optical fiber Fabry – Perot humidity sensor based on polyimide membrane : Sensitivity and adsorption kinetics, Sensors Actuators A. Phys. 281 (2018) 48–54. https://doi.org/10.1016/j.sna.2018.08.044.

[53] M.P. Lavin-Lopez, J.L. Valverde, M.C. Cuevas, A. Garrido, L. Sanchez-Silva, P. Martinez, A. Romero-Izquierdo, Synthesis and characterization of graphene: influence of synthesis variables, Phys. Chem. Chem. Phys. 16 (2014) 2962–2970. https://doi.org/10.1039/C3CP54832E.

[54] S. Lee, K. Lee, Z. Zhong, Wafer Scale Homogeneous Bilayer Graphene Films by Chemical Vapor Deposition, Nano Lett. 10 (2010) 4702–4707. https://doi.org/10.1021/nl1029978.

[55] S.N. Tripathi, P. Saini, D. Gupta, V. Choudhary, Electrical and mechanical properties of PMMA/reduced graphene oxide nanocomposites prepared via in situ polymerization, J. Mater. Sci. 48 (2013) 6223–6232. https://doi.org/10.1007/s10853-013-7420-8.

[56] R. Li, Z. Li, E. Pambou, P. Gutfreund, T.A. Waigh, J.R.P. Webster, J.R. Lu, Determination of PMMA Residues on a Chemical-Vapor-Deposited Monolayer of Graphene by Neutron Reflection and Atomic Force Microscopy, Langmuir. 34 (2018) 1827–1833. https://doi.org/10.1021/acs.langmuir.7b03117.

[57] M. Xia, Z. Su, S. Zhang, Raman spectra of bilayer graphene covered with Poly(methyl

methacrylate) thin film, AIP Adv. 2 (2012) 32122. https://doi.org/10.1063/1.4739785.

[58] E. Koo, S.-Y. Ju, Role of residual polymer on chemical vapor grown graphene by Raman spectroscopy, Carbon N. Y. 86 (2015) 318–324. https://doi.org/https://doi.org/10.1016/j.carbon.2015.01.055.

[59] C. Rattanabut, W. Wongwiriyapan, W. Muangrat, W. Bunjongpru, M. Phonyiem, Y.J. Song, Graphene and poly(methyl methacrylate) composite laminates on flexible substrates for volatile organic compound detection, Jpn. J. Appl. Phys. 57 (2018) 04FP10. https://doi.org/10.7567/jjap.57.04fp10.

[60] C. Rattanabut, W. Muangrat, M. Phonyiem, W. Bungjongpru, W. Wongwiriyapan, Y. Jae Song, Hybrid graphene and poly(methyl methacrylate) for gas sensor application, Mater. Today Proc. 4 (2017) 6397–6403. https://doi.org/https://doi.org/10.1016/j.matpr.2017.06.144.

[61] V.K. Thakur, D. Vennerberg, S.A. Madbouly, M.R. Kessler, Bio-inspired green surface functionalization of PMMA for multifunctional capacitors, RSC Adv. 4 (2014) 6677–6684. https://doi.org/10.1039/C3RA46592F.

[62] S.K. Mishra, S.N. Tripathi, V. Choudhary, B.D. Gupta, SPR based fibre optic ammonia gas sensor utilizing nanocomposite film of PMMA/reduced graphene oxide prepared by in situ polymerization, Sensors Actuators B Chem. 199 (2014) 190–200. https://doi.org/https://doi.org/10.1016/j.snb.2014.03.109.

[63] D. Pawar, R. Kitture, S.N. Kale, ZnO coated Fabry-Perot interferometric optical fiber for detection of gasoline blend vapors: Refractive index and fringe visibility manipulation studies, Opt. Laser Technol. 89 (2017) 46–53. https://doi.org/https://doi.org/10.1016/j.optlastec.2016.09.038.

[64] C.M. Yang, T.C. Chen, Y.C. Yang, M. Meyyappan, C.S. Lai, Enhanced acetone sensing

properties of monolayer graphene at room temperature by electrode spacing effect and UV illumination, Sensors Actuators, B Chem. 253 (2017) 77–84. https://doi.org/10.1016/j.snb.2017.06.116.

[65] R. Kitture, D. Pawar, C.N. Rao, R.K. Choubey, S.N. Kale, Nanocomposite modified optical fiber: A room temperature, selective H2S gas sensor: Studies using ZnO-PMMA, J. Alloys Compd. 695 (2017) 2091–2096. https://doi.org/https://doi.org/10.1016/j.jallcom.2016.11.048.

[66] C.-B. Yu, Y. Wu, C. Li, F. Wu, J.-H. Zhou, Y. Gong, Y.-J. Rao, Y.-F. Chen, Highly sensitive and selective fiber-optic Fabry-Perot volatile organic compounds sensor based on a PMMA film, Opt. Mater. Express. 7 (2017) 2111–2116. https://doi.org/10.1364/OME.7.002111.

[67] C.M. Hansen, The three-dimensional solubility parameter and solvent diffusion coefficient: Their importance in surface coating formulation, (n.d.).

[68] J.K. Abraham, B. Philip, A. Witchurch, V.K. Varadan, C.C. Reddy, A compact wireless gas sensor using a carbon nanotube/PMMA thin film chemiresistor, Smart Mater. Struct. 13 (2004) 1045–1049. https://doi.org/10.1088/0964-1726/13/5/010.

[69] A. Klini, S. Pissadakis, R.N. Das, E.P. Giannelis, S.H. Anastasiadis, D. Anglos, ZnO − PDMS Nanohybrids : A Novel Optical Sensing Platform for Ethanol Vapor Detection at Room Temperature, J. Phys. Chem. C. 119 (2015) 623–631. https://doi.org/10.1021/jp506632d.

[70] S. Lettieri, A. Setaro, C. Baratto, E. Comini, G. Faglia, G. Sberveglieri, P. Maddalena, On the mechanism of photoluminescence quenching in tin dioxide nanowires by NO2 adsorption, New J. Phys. 10 (2008). https://doi.org/10.1088/1367-2630/10/4/043013.

[71] and I.P. Jeonghoon Yun, Jae-Hyuk Ahn, Dong-Il Moon, Yang-Kyu Choi, Joule heated

and suspended silicon nanowire based sensor for low-power and stable hydrogen detection, ACS Appl. Mater. Interfaces. (2019). https://doi.org/10.1021/acsami.9b15111.

[72] V.C. 2 and J.-F.F. Abhishek Sachan 1, 2 , Mickael Castro 1, vQRS Based on Hybrids of CNT with PMMA-POSS and PS-POSS Copolymers to Reach the Sub-PPM Detection of Ammonia and Formaldehyde at Room Temperature Despite Moisture, Chemosensors. (2017). https://doi.org/10.3390/chemosensors5030022.

[73] G. Wang, Z. Dai, L. Liu, H. Hu, Q. Dai, Z. Zhang, Tuning the Interfacial Mechanical Behaviors of Monolayer Graphene/PMMA Nanocomposites, ACS Appl. Mater. Interfaces. 8 (2016) 22554–22562. https://doi.org/10.1021/acsami.6b03069.

[74] Z. Zabihi, H. Araghi, P.E.D.S. Rodriguez, A. Boujakhrout, R. Villalonga, Vapor sensing and interface properties of reduced graphene oxide–poly(methyl methacrylate) nanocomposite, J. Mater. Sci. Mater. Electron. 30 (2019) 2908–2919. https://doi.org/10.1007/s10854-018-00567-4.

[75] W. Liu, L. Xu, K. Sheng, X. Zhou, B. Dong, G. Lu, H. Song, A highly sensitive and moisture-resistant gas sensor for diabetes diagnosis with Pt@In2O3 nanowires and a molecular sieve for protection, NPG Asia Mater. 10 (2018) 293–308. https://doi.org/10.1038/s41427-018-0029-2.

[76] L. Yi, C. You-ping, S. Han, Z. Gang, L. Yi, C. You-ping, S. Han, Z. Gang, Hydrogen gas sensor based on palladium and yttrium alloy ultrathin film, Rev. Sci. Instrum. 83 (2012) 125003. https://doi.org/10.1063/1.4770329.

[77] Y. Liu, Y. Li, P. Huang, H. Song, G. Zhang, Modeling of hydrogen atom diffusion and response behavior of hydrogen sensors in Pd–Y alloy nanofilm, Sci. Rep. 6 (2016) 37043. https://doi.org/10.1038/srep37043.

[78] U. Sampath, D.G. Kim, H. Kim, M. Song, Cryogenic Temperature Sensor Based on

Fresnel Reflection from a Polymer-Coated Facet of Optical Fiber, IEEE Sens. J. 18 (2018) 3640–3644. https://doi.org/10.1109/JSEN.2018.2813303.

[79] M. Esposito, S. Buontempo, A. Petriccione, M. Zarrelli, G. Breglio, A. Saccomanno, Z. Szillasi, A. Makovec, A. Cusano, A. Chiuchiolo, M. Bajko, M. Giordano, Fiber Bragg Grating sensors to measure the coefficient of thermal expansion of polymers at cryogenic temperatures, Sensors Actuators, A Phys. 189 (2013) 195–203. https://doi.org/10.1016/j.sna.2012.09.016.

[80] A.A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, C.N. Lau, Superior Thermal Conductivity of Single-Layer Graphene 2008, Nano Lett. 8 (2008) 902–907.

[81] M. Wang, D. Li, R. Wang, J. Zhu, Z. Ren, PDMS-assisted graphene microfiber ring resonator for temperature sensor, Opt. Quantum Electron. 50 (2018) 1–8. https://doi.org/10.1007/s11082-018-1395-2.

Figure captions:

Figure 1. (a) Schematic of DFPI sensor, (b) Microscopic image of PMF, and (c) image shows splicing of SMF with PMF.

Figure 2. The electric field distribution inside the DFPI-based configurations for (a) PMMA, (b) graphene, (c) graphene/PMMA composite, and (d) Spatial frequency study for different DFPI structures.

Figure 3. (a) Raman spectrum of graphene and graphene/PMMA composite, and (b) TGA spectrum of graphene, PMMA and graphene/PMMA composite.

Figure 4. FE-SEM images of optical fiber tip coated samples and their corresponding sensing probe images. (a, b, c) for graphene, (d, e, f) for PMMA, and (g, h, i) for graphene/PMMA composite.

Figure 5. Graphene coated DFPI sensor response in terms of its (a) Interference spectrum recorded with methanol, (b) Close view of its resonance dip, and (c) Wavelength shift for different vapors in the concentration ranging from 5 μL/L to 200 μL/L.

Figure 6. Sensing mechanism of graphene with different VOCs gas.

Figure 7. PMMA coated DFPI sensor response in terms of (a) Interference spectrum response for acetone, (b) Zoomed graph of its resonance dip, and (c) Wavelength shift for different VOCs in concentration ranging from 5 μL/L to 200 μL/L.

Figure 8. Sensing mechanism of PMMA under the exposure of different VOCs.

45

Figure 9. The graphene/PMMA coated DFPI sensor response in terms of (a) Spectral response for acetone gas, (b) Zoomed graph of its resonance dip, (c) Wavelength shift for different VOCs in concentration varied from 5 µL/L to 200 µL/L.

Figure 10. Sensing mechanism of graphene/PMMA interaction with different VOCs.

Figure 11. (a) GP-DFPI spectral response towards mixture of VOCs, (b) Zoomed graph of its resonance dip; The Sensor response of G-DFPI, P-DFPI and GP-DFPI in terms of (c) Wavelength shift vs concentration and (d) Power shift under the exposure of mixture of VOCs in concentration from 5 µL/L to 200 µL/L.

Figure 12. Response of G-DFPI, P-DFPI and GP-DFPI in terms of wavelength shift towards various binary mixtures of concentration varies from 5 µL/L to 200 µL/L.

Figure 13. Selectivity response of individual G-DFPI, P-DFPI and GP-DFPI sensors under the exposure of different volatile gases at 200 µL/L.

Figure 14. (a) Time response of all sensors to acetone gas in concentration ranging from 5 µL/L to 200 µL/L, (b) Time response of all sensors at 5 µL/L and (c) at 200 µL/L of acetone vapor, (d) Repeatability of GP-DFPI sensor towards acetone gas at 200 µL/L, (e) Average response and recovery time of all the sensors at 200 µL/L towards acetone (f) Response and recovery time of all sensors under the exposure of acetone in concentration from 5 µL/L to 200 µL/L.

47

Fig. 1.

48

Fig. 2.

Fig. 3.

50

Fig. 4.

Fig. 5.

52

Fig. 6.

Fig. 7.

Fig. 8.

55

Fig. 9.

Fig. 10.

Fig. 11.

58
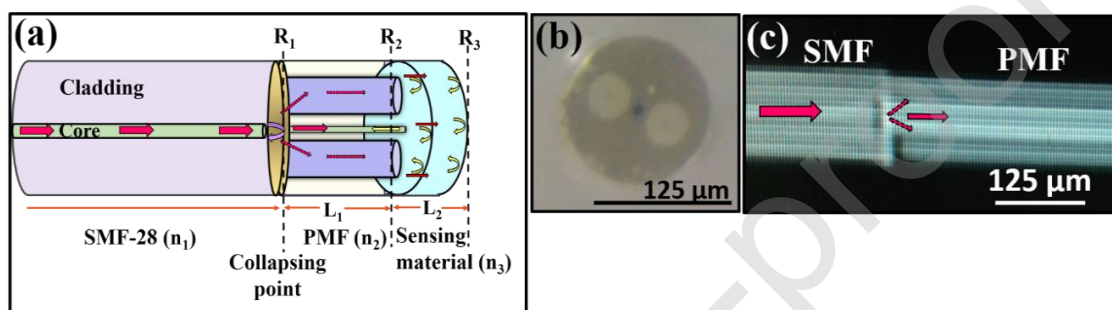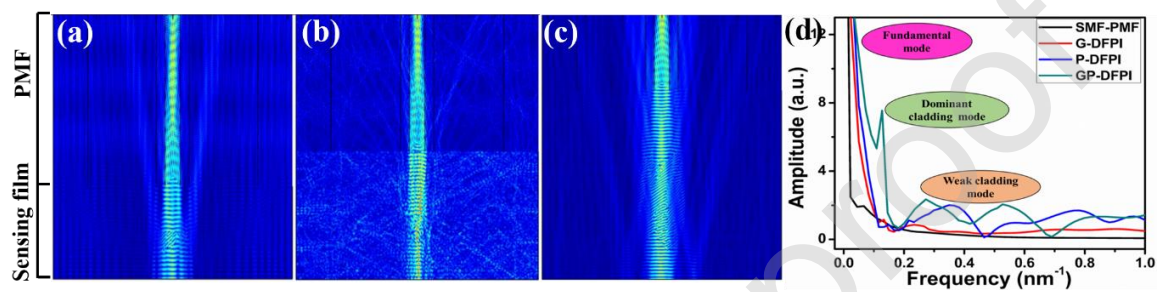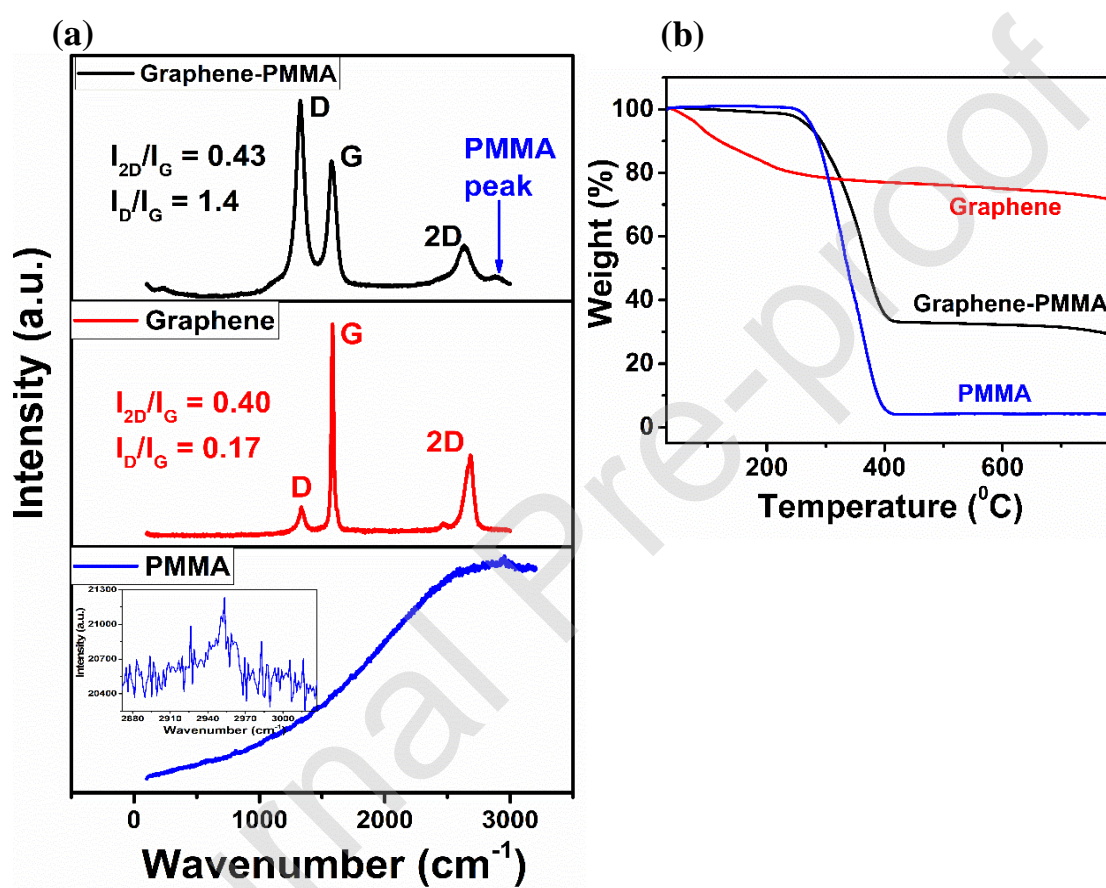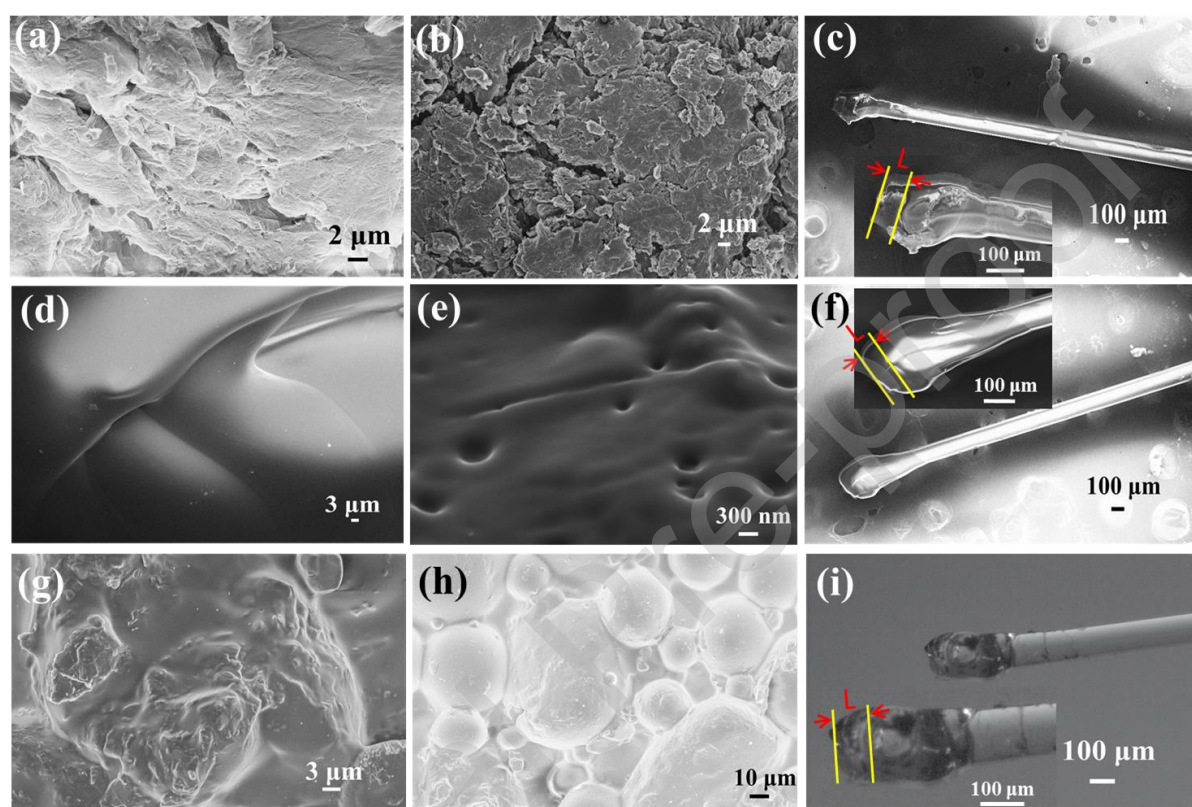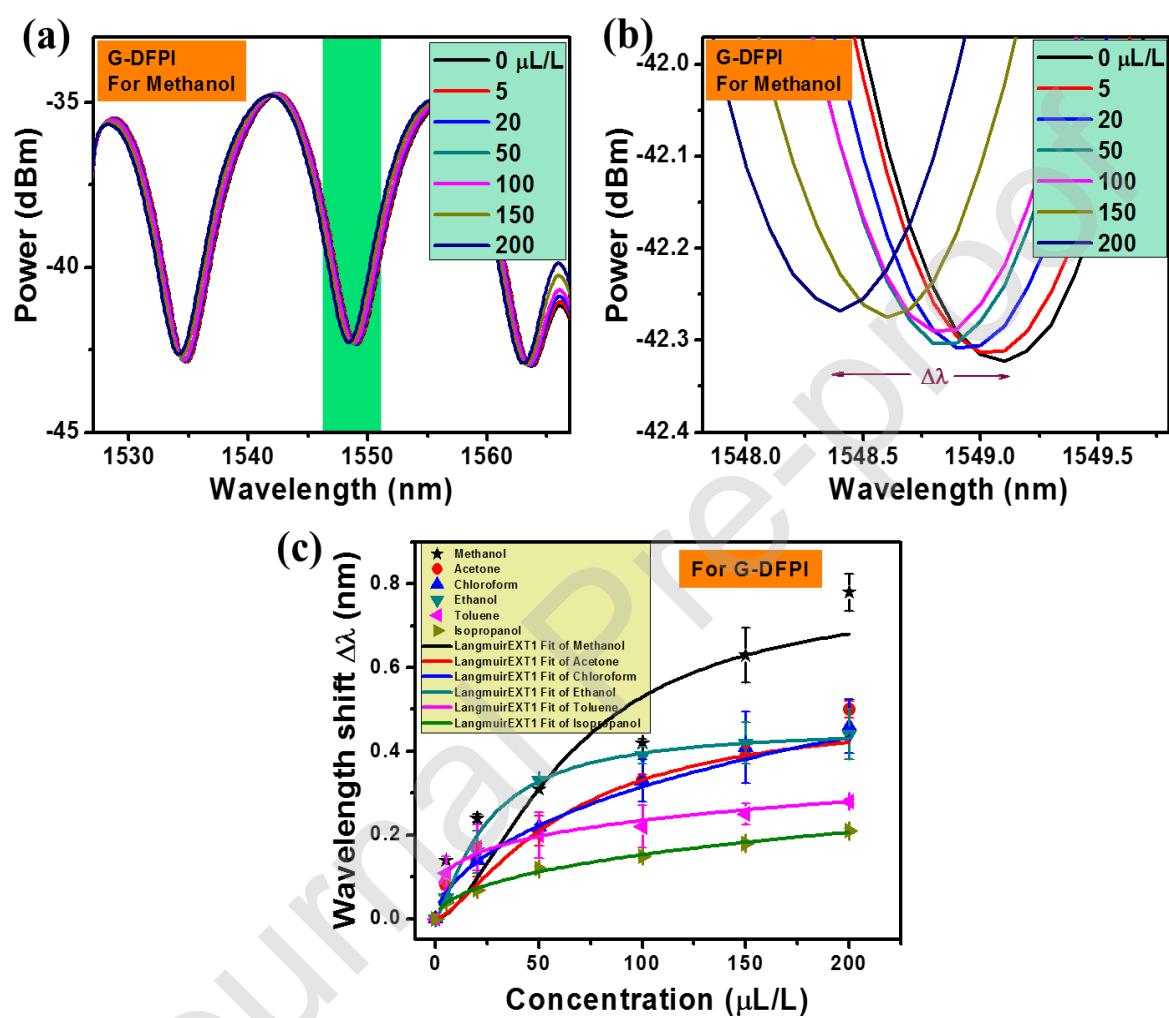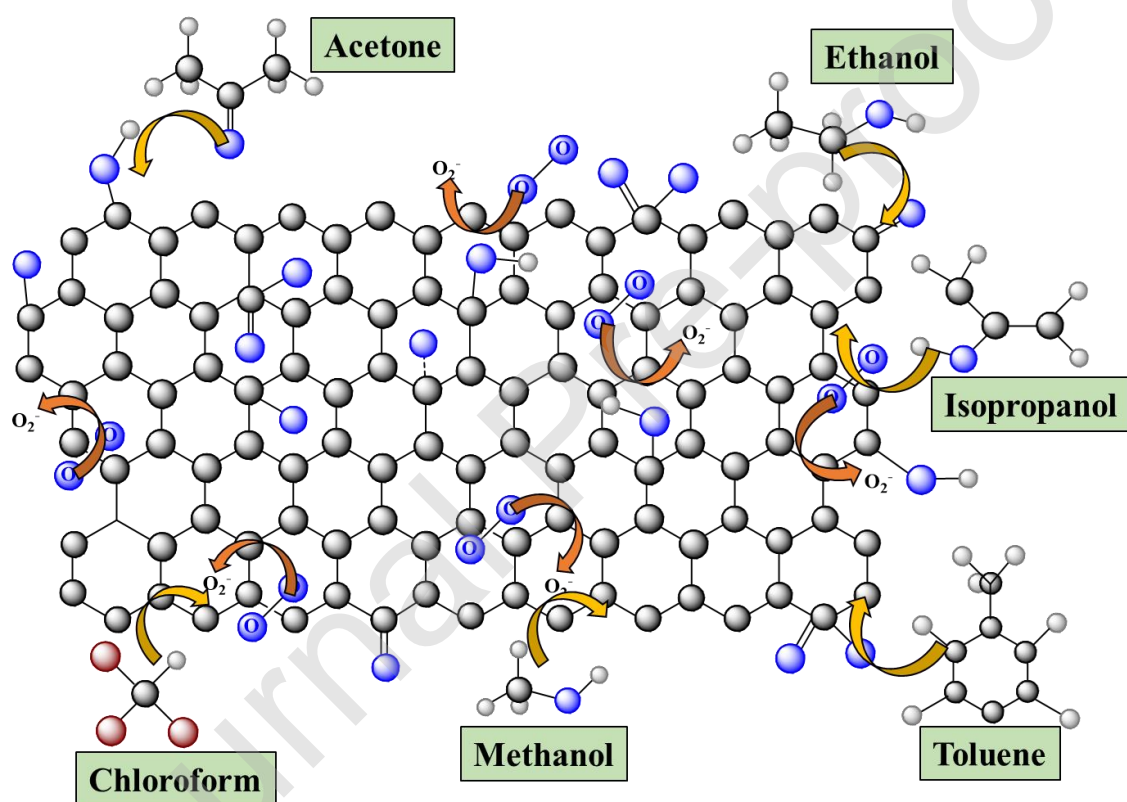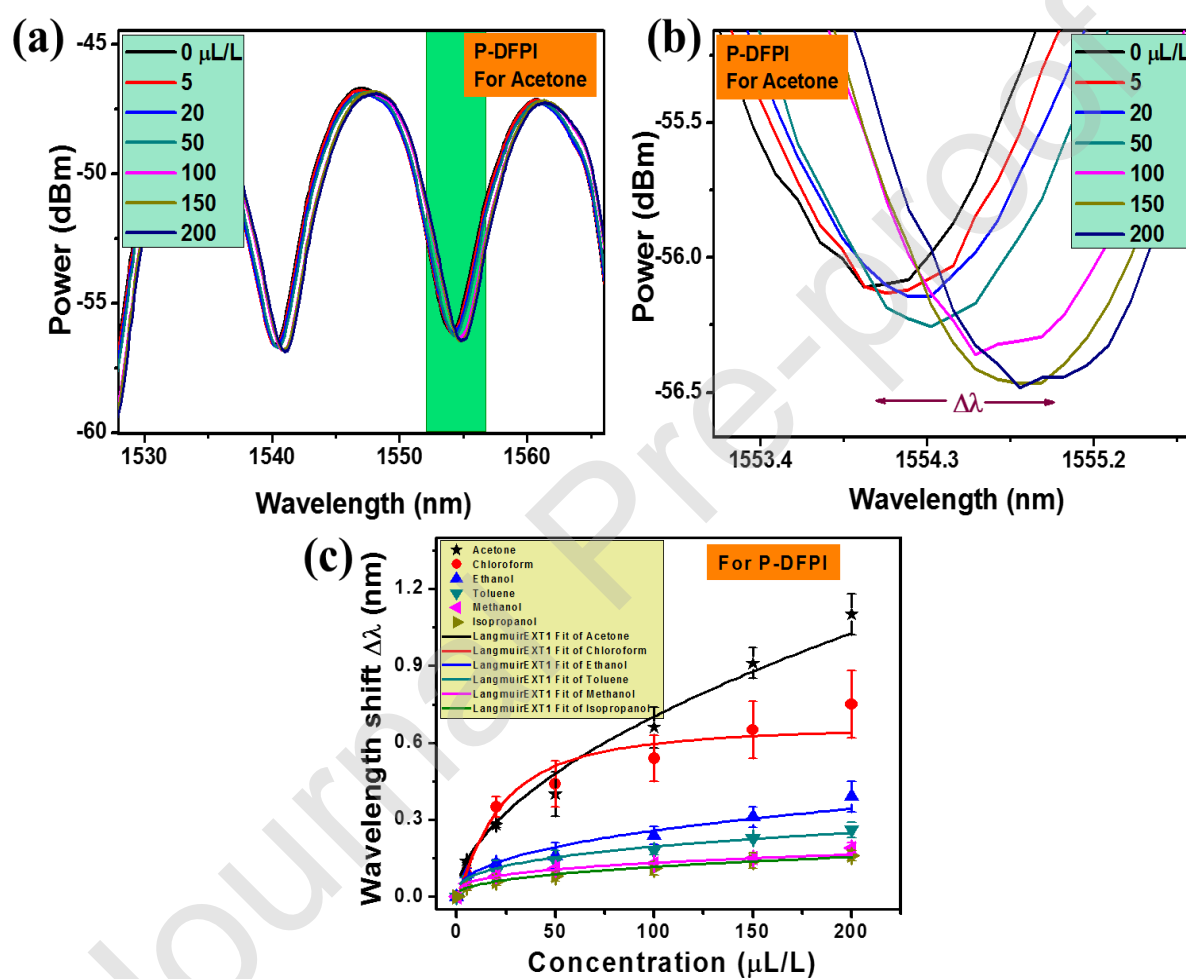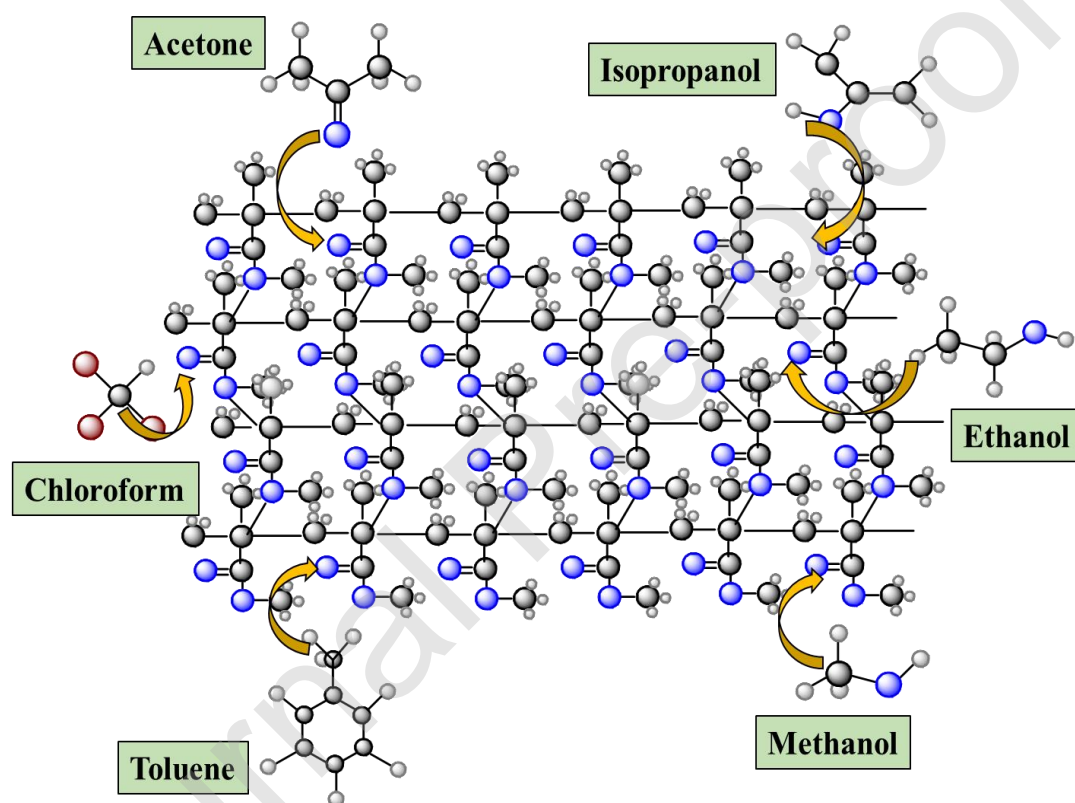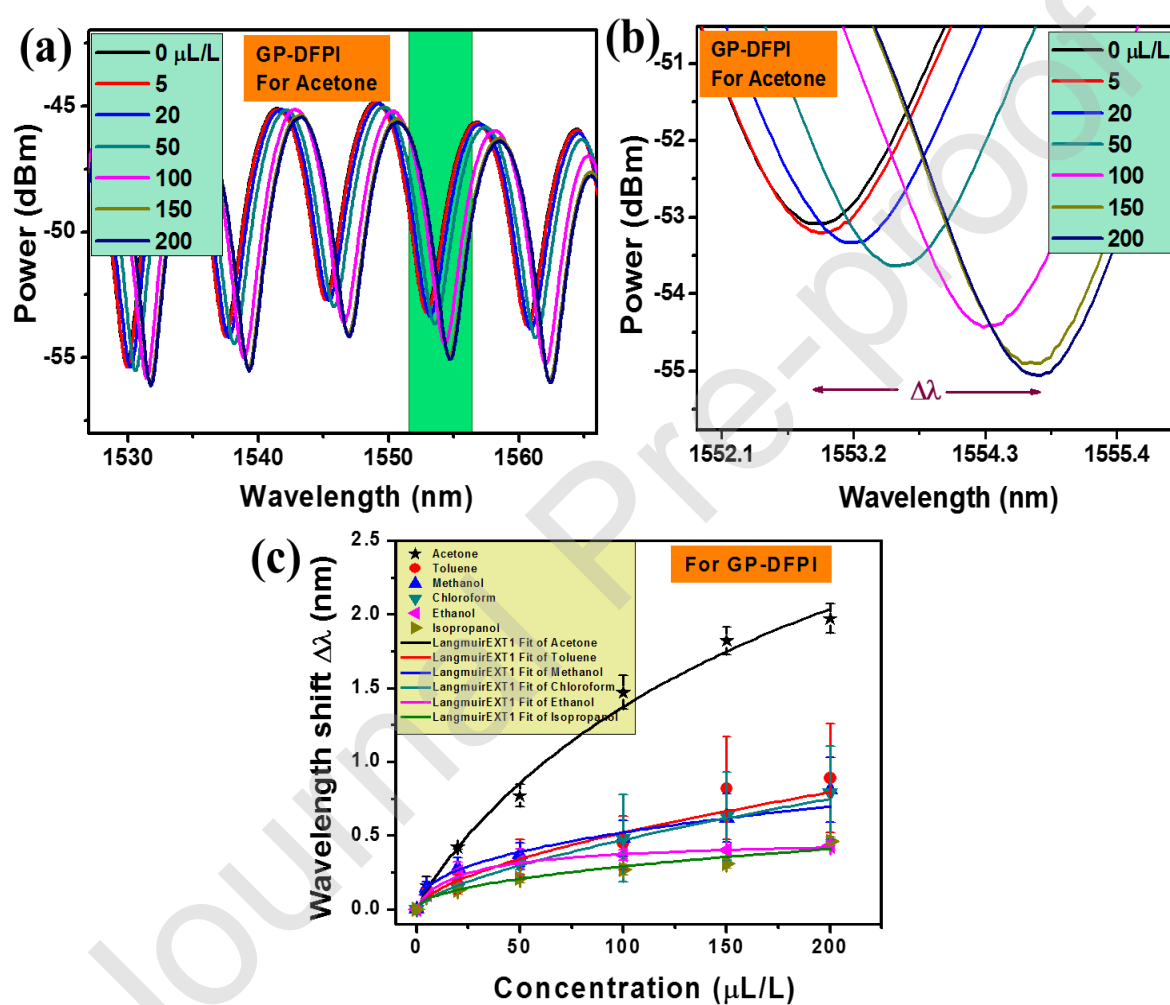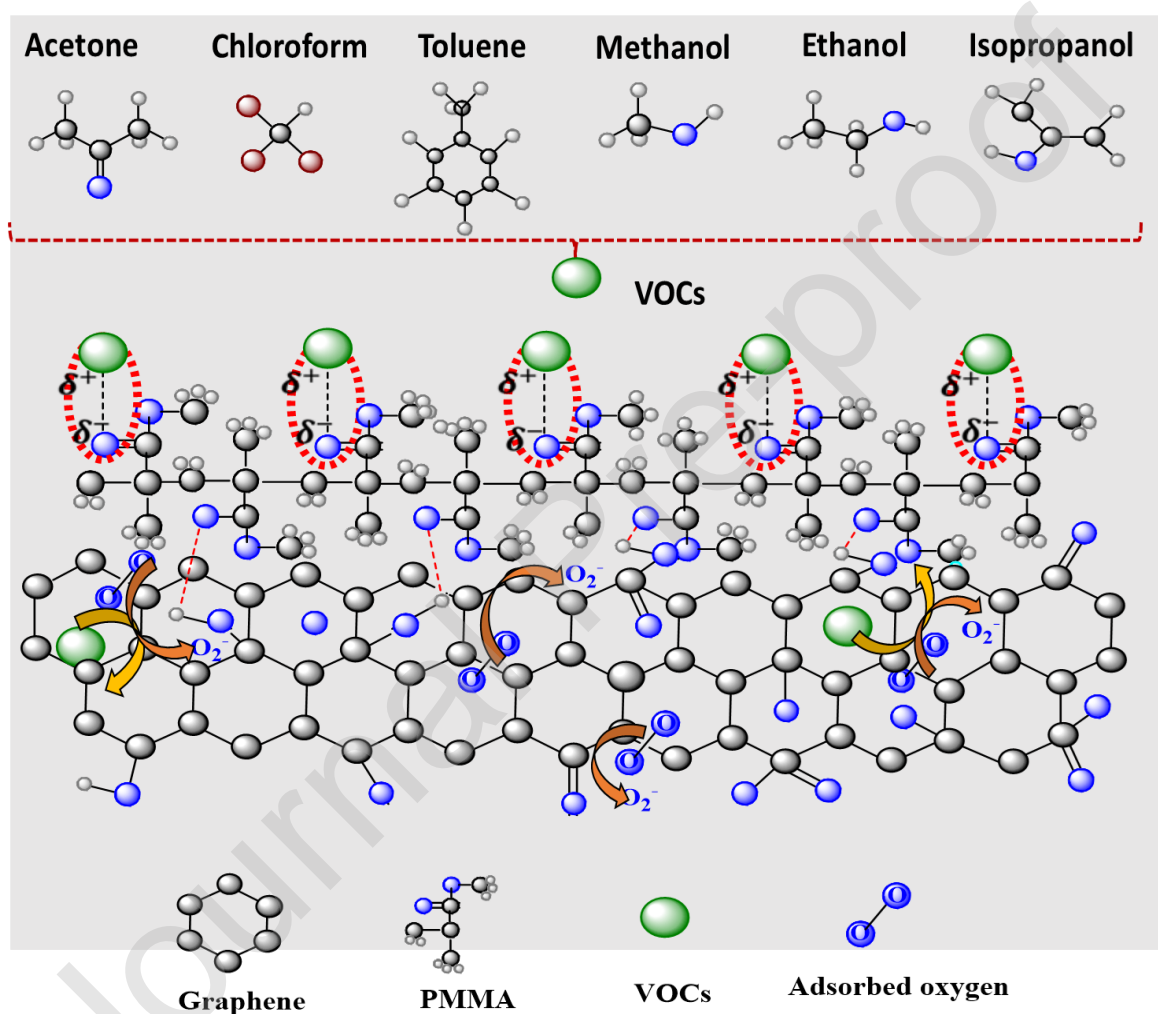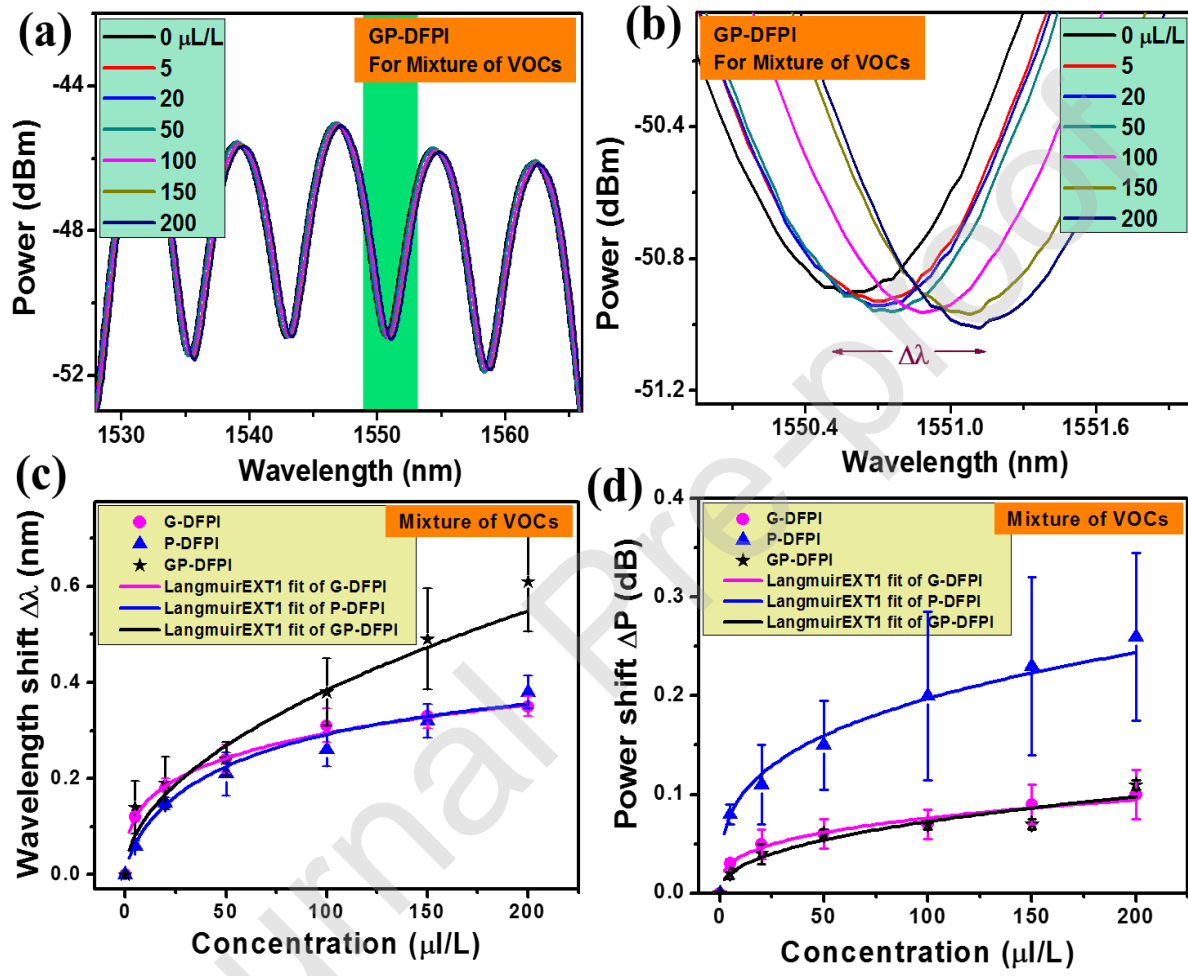
Fig. 12.
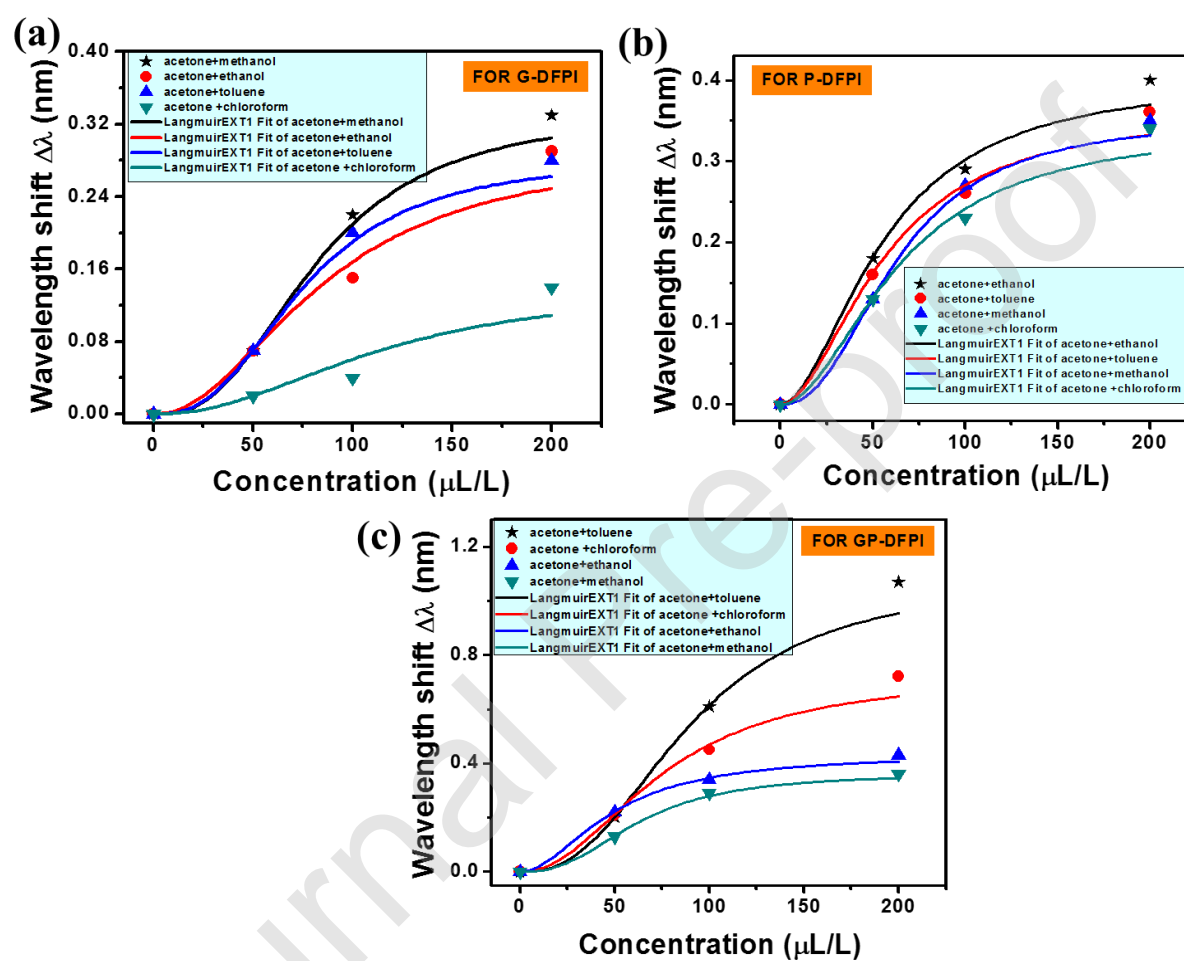
Fig. 13.
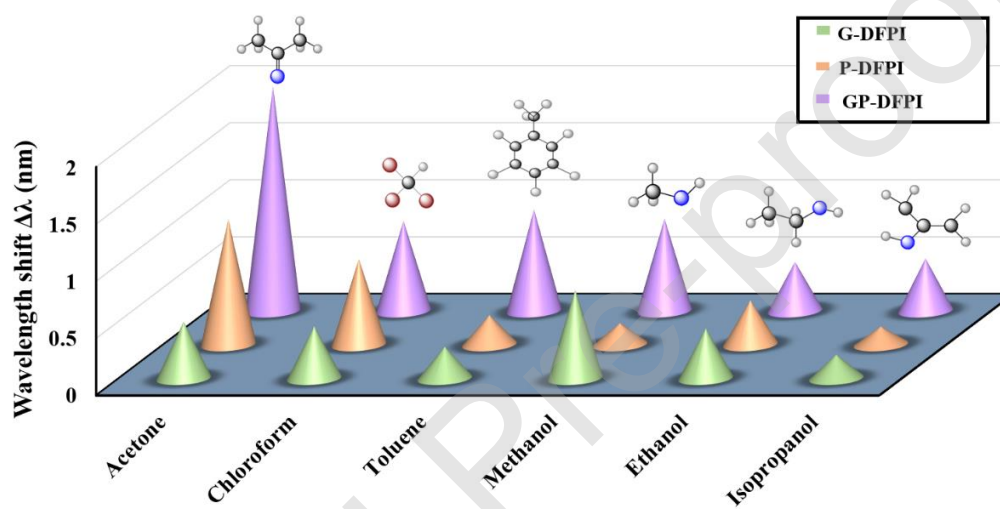
Fig. 14.
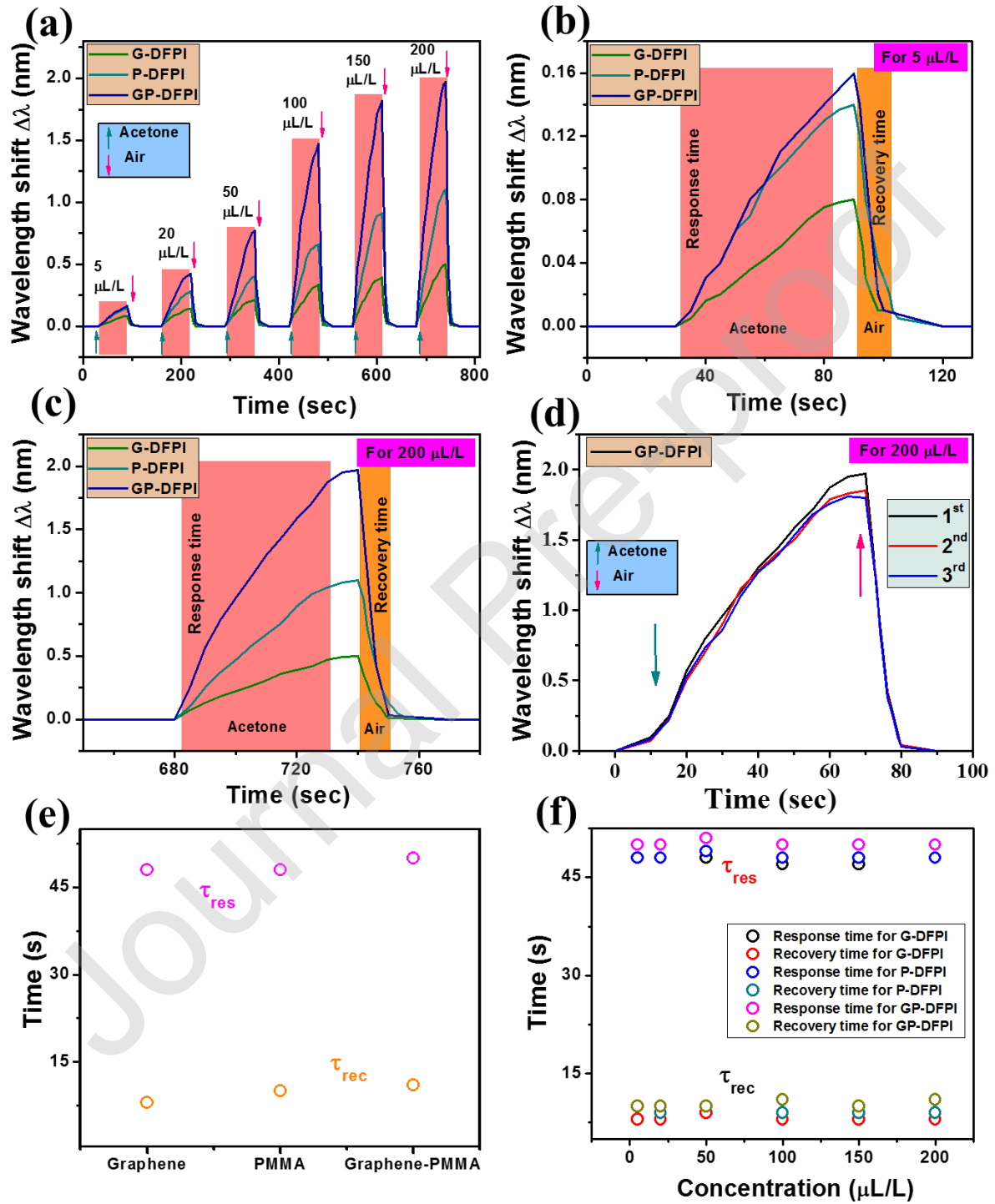
Table 1. Reported nanomaterials (graphene, MOXs, polymer) coated optical fiber sensors for detection of various VOCs.

Materials and type of fiber	Sensing configuration	VOCs type	Film deposition technique	Sensing range	Sensitivity (S)	Response time / Recovery time
rGO coated on side polished SMF [[29]]	Transmission mode	Toluene	Evaporation deposition	40-196 ppm	0.003667 dB/ppm	4.26 min. / 4.26 min.
GO coated PM microfiber [[18]]	Evanescent (Transmission mode)	Ethanol	Drop casting	0-80 ppm	0.138 nm/ppm	6.1 min./-
GO coated on unclad PMMA fiber [[30]]	Transmission mode	Methanol	Dip coating	0-500 ppm	-0.20 counts/ppm	NA
GO coated onto FBG/ SMF [[37]]	Transmission mode	Ethanol	Drop casting	0-100 %	0.45 nm/%C	NA
SWCNTs coated on unclad PMMA fiber [[31]]	Transmission mode	Methanol	Dip coating	0-500 ppm	~40 counts @500 ppm	110 min. / 50 min.
CNT coated on tapered MMF [[35]]	FPI (Reflection mode)	Ethanol	Drop casting	5-20%	0.1441 /vol%	50 s / 53 s
ZnO coated on unclad MMF [[33]]	Evanescent (Radiation mode)	Ethanol	Dip coating	0-500 ppm	56×10^{-3} /kPa	~11 min. / ~8 min.
ZnO nanowires coated on unclad MMF [[32]]	Transmission mode	Ethanol	Magnetron sputtering	50-500 ppb	~ $R_a/R_g = 1.4$	66.33 min. / 4.53 min.
V_2O_5 and WO_3 coated on unclad PMMA fiber [[34]]	Transmission mode	Methanol	Dip coating	0-500 ppm	56×10^{-3} /kPa 59×10^{-3} /kPa	40 min. / 39 min.
LPGs modified with (PAH/SiO2)8 [[38]]	Transmission mode	Acetone, Chloroform, Toluene	Electrostatic self-assembly	298,000 ppmv, 210,000 ppmv, 28,000 ppmv	Wavelength shift = 0.35 nm, 0.58 nm, 0.24 nm	NA

TiO$_2$/GO coated tip of the MMF [[36]]	GMR-based (Transmission mode)	Ethylene	e-beam evaporation	0-909 ppm	0.92 pm/ppm	NA
PDMS coated FBGs [[39]]	Evanescent (Transmission mode)	Acetone	NA	NA	1.63 x 10^{-3} nm/%	12 min. / 23 min.

Table 2: The HSP parameters of PMMA and various VOCs.

Polymer and VOCs	δ_d	δ_p	δ_h	R_a	R_0	RED = R_a/R_0
PMMA	17.9	10.1	5.4	-	11	-
Acetone	15.5	10.4	7.0	5.06		0.46
Chloroform	17.8	3.1	5.7	7.00		0.63
Toluene	18.0	1.4	2.0	9.34		0.84
Isopropanol	15.8	6.1	16.4	12.43		1.13
Ethanol	15.8	8.8	19.4	15.23		1.38
Methanol	15.1	12.3	22.3	17.93		1.63

Machine Learning & Concept Drift based Approach for Malicious Website Detection

Siddharth Singhal, Utkarsh Chawla
Dept. of Computer Science & Engineering
Delhi Technological University
New Delhi, India
{siddharthsinghal_bt2k16, utkarshchawla_bt2k16}@dtu.ac.in

Rajeev Shorey
TCS Research and Innovation &
Dept. of Computer Science & Engineering
Indian Institute of Technology
New Delhi, India
rajeev.shorey@tcs.com

Abstract—The rampant increase in the number of available cyber attack vectors and the frequency of cyber attacks necessitates the implementation of robust cybersecurity systems. Malicious websites are a significant threat to cybersecurity. Miscreants and hackers use malicious websites for illegal activities such as disrupting the functioning of the systems by implanting malware, gaining unauthorized access to systems, or illegally collecting personal information. We propose and implement an approach for classifying malicious and benign websites given their Uniform Resource Locator (URL) as input. Using the URL provided by the user, we collect Lexical, Host-Based, and Content-Based features for the website. These features are fed into a supervised Machine Learning algorithm as input that classifies the URL as malicious or benign. The models are trained on a dataset consisting of multiple malicious and benign URLs. We have evaluated the accuracy of classification for Random forests, Gradient Boosted Decision Trees and Deep Neural Network classifiers. One loophole in the use of Machine learning for detection is the availability of the same training data to the attackers. This data is exploited by the miscreants to alter the features associated with the Malicious URLs, which will be classified as benign by the supervised learning algorithms. Further, owing to the dynamic nature of the malicious websites, we also propose a paradigm for detecting and countering these manually induced concept drifts.

Index Terms—URL Feature Extraction, Malicious Website Detection, Concept Drifts, Feature Vectors, Gradient Boosted Trees, Random Forest, Feedforward Neural Networks

I. INTRODUCTION

We have witnessed an exponential growth in the number of users browsing the web. This is due to several reasons such as, for example, majority of businesses shifting online, an increase in usage of highly reliable cloud storage and computation devices, and availability of high-speed connectivity to the internet across the globe including remote areas at low prices. As systems get connected to the internet, they become vulnerable to a wide range of cyber threats such as spam-advertising, financial fraud, malware implanting, and information theft. Malicious websites are the primary attack vector used by hackers for propagating these attacks. Miscreants provide these URLs as baits to naive users who are unable to distinguish between malicious and benign websites, which necessitates the development of robust solutions for validating an unfamiliar URL or suspicious website.

A variety of methodologies have been proposed and deployed for detecting websites as well as mitigating the cyber attacks. One of the most widely used approaches is to maintain a centralized blacklist of suspicious as well as malicious sites, IP addresses, and domains [5]. The solution, though elegant and straightforward, is not scalable as it is tedious or impossible to validate and maintain a list of all these websites.

To overcome the scalability issues, we propose a system for automatic feature extraction and real-time classification using the input URL. Singh et al. [13] analyse the effectiveness of various machine learning attributes for detecting malicious websites. However, their work is limited to the use of Host-Based and Content-Based features. In our approach, we extract the Lexical features along with Host-Based, and Content-Based features and store these attributes as feature vectors for each URL. We feed these feature vectors as input to the supervised learning algorithm, which classifies these URLs as malicious or benign. The supervised learning algorithms used in our study are Random Forests, Gradient Boosted Trees, and Feed Forward Neural Networks. Our models are trained on a list of malicious and benign URLs curated from a variety of sources. While the proposed approach is agnostic to a variety of attacks and is scalable, it does not conform to the dynamic nature of websites. Miscreants who have access to the same training data and are capable of identifying patterns in detection algorithms attempt to alter certain features of malicious sites with the motive of bypassing the security. The scenario where the relation between the input data and the target variable changes over time is referred to as Concept Drift [8]. Tan et al. [7] propose machine learning techniques such as Linear Support Vector Machines, Logistic Regression, Random Forests, and Naive Bayes for detecting malicious websites along with statistical method for detecting Concept Drifts in websites. In this paper, we define a novel approach for detecting concept drifts by computing the distance between old and new feature space. The novelty of our proposed approach lies in the use of Neural Networks for detecting Malicious websites.

II. DATASET

In an effort to validate our approach, we collect a set of malicious URLs and a set of benign URLs. For each URL in both the datasets, we extract the features listed in Table I.

We then pre-process these features to be suitable for our supervised learning models.

- **Phase 1: Data Collection:** We used the following public blacklists to collect malicious URLs: PhishTank [12], a collaborative platform operated by OpenDNS to distribute and verify phishing websites and Malware Domain List (MDL) [1], which maintains an archive of malware-infected websites. Benign URLs are retrieved from majestic.com [9], which lists the top one million URLs with the most referring subnets. Our final dataset has a total of eighty thousand unique URLs, out of which forty thousand are malicious, and the remaining forty thousand are benign.

- **Phase 2: Feature Extraction:** We categorize the URL features into three groups: lexical, host-based, and content-based. For host-based and content-based, we use the features which are considered to be important by Singh et al. [13]. We also include lexical based features, which have been known to improve the performance of the classifier [2] [3].

  - **Lexical Features:** Lexical features are extracted from the URL (Uniform Resource Locator) string. The idea is to classify a malicious website from a benign one by identifying the differences in various aspects of their URL strings.
    * *URL Length:* Length of the Uniform Resource Locator.
    * *Host Length:* Length of the hostname.
    * *Host token count:* Number of tokens in the hostname delimited by (".").
    * *Path Length:* Length of the pathname.
    * *Number of symbols:* Number of special characters such as '&', '%', '$', '#', '⁷', etc.

  - **Host-Based Features**
    * *Location:* The Geographical location of the website is determined from the IP address through the GeoIP database.
    * *Autonomous System Number:* This is a unique number that is available globally to identify an autonomous system and which enables the system to exchange exterior routing information with other neighbouring autonomous systems.

  - **Content-Based Features:** Content-Based features are obtained by analysing the HTML and the JavaScript code of the website.
    * *HTTPS Enabled:* The presence of the Secure Socket Layer (SSL) protocol ensures that the website has a basic level of privacy and integrity [4].
    * *Applet Count:* The number of java applet tags. Java applets are used to perform malicious exploits [15].
    * *Presence of Eval() function:* Eval() function is used to generate malicious code at runtime, which prevents detection [6].

Feature	Description
URL Length	Length of the Uniform Resourse Locator
Host Length	Length of the hostname.
Host Token Count	Number of tokens in the hostname (delimited by '.')
Path Length	Length of the pathname.
Number of symbols	Number of special characters like '', '%', '$', '', '=', etc.
Location	Geographical Location of the website
ASN	Autonomous System Number of the website
HTTPS Enabled	Presence of Secure Socket Layer
Applet Count	Number of Java Applet tags
Eval() function	Presence of eval() function
XHR	Number of XHR tags
Popups	Presence of windows.open() function
Redirection	Check if the website redirects
unescape() function	Presence of unescape() function

TABLE I: Summary of features extracted from the URL

  * *Count of XML HTTP Requests (XHR) tag:* XHR, although being at the core of AJAX, can be used to inject malicious exploits. The number of XHR tags is stored as a numerical value.
  * *Presence of popups:* The windows.open() function is generally used for advertisements, but can also be used to inject malicious code.
  * *Redirection:* Redirection in a website can be used as a means to redirect from a benign URL to a malicious one [10]. Redirection is checked through the Requests library in python, and the result is stored as a boolean variable.
  * *Presence of unescape() function:* The unescape() function is usually used for decoding encoded malicious code [11].

Table I provides a summary of all the aforementioned features.

III. PROPOSED FRAMEWORK

In this section, we define the design of our proposed framework which classifies the website as malicious or benign. Further, the system is equipped with the ability to retrain the supervised machine learning model in the presence of concept drifts. Our system consists of three modules, namely feature extraction, supervised learning and concept drift detection. The high-level architecture of the system is shown in Figure 1.

- **Feature extraction:** This module makes use of open-source python libraries such as Requests, WHOIS, BeautifulSoup and open-source GeoIP database to extract features from a given URL. All the features given in Table I are obtained using this module.

- **Supervised Learning model:** This module is responsible for training the supervised learning model followed by predicting the degree of maliciousness for unseen URLs. We analysed three algorithms, namely Gradient Boosted Trees, Random Forest and Feedforward Neural Networks as candidates for our learning model.

  - **Gradient Boosted Trees:** Gradient boosting is a machine learning technique which uses an ensemble of decision trees to form the prediction model. It
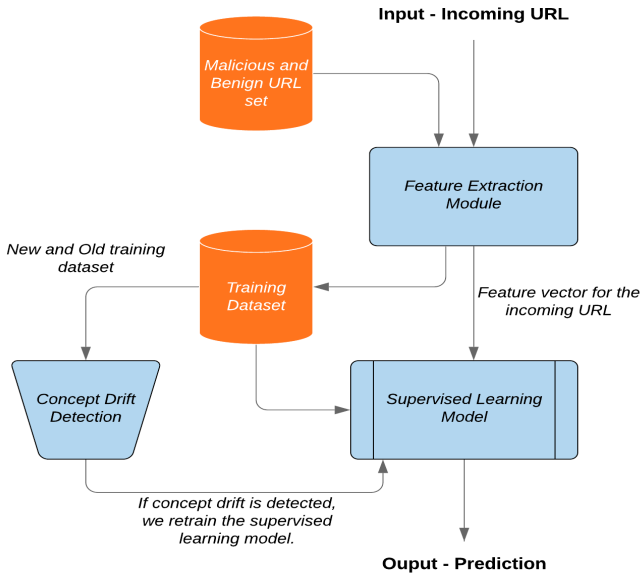
Fig. 1: System Architecture

sequentially builds the model such that the succeeding predictors can learn from the mistakes of the preceding predictors

- **Random Forest:** Random Forest is a machine learning technique used for classification and regression, which is an ensemble of decision trees. It uses bagging and feature randomness for creating a forest of uncorrelated trees having accuracy higher than the individual trees.
- **Feedforward Deep Neural Networks:** This is an artificial neural network where links between the different layers do not form a cycle. Thus, in this network, data only moves in the forward direction.

- **Concept Drift Detection:** One shortcoming of our model is that hackers and miscreants are aware of the standard features that are obtained from the URL. Hence, this provides them with the opportunity to modify the properties of malicious URLs to escape detection. This phenomenon, where the relationship between the input data and the target variable changes over time, is referred to as concept drift [8].

We propose an algorithm to detect concept drifts in real-time efficiently. During concept drift detection, we again collect the feature vectors for all the malicious and benign URLs in our training dataset. The algorithm focuses on finding the difference in data distribution between the old training data and the newly collected feature vectors data. For every malicious URL in the new training dataset, we find the least distant malicious URL feature vector in the old dataset. We repeat the same procedure for all the benign URLs in the new dataset. If the average of all the distances calculated for this time period is more than a pre-set threshold value, then we claim that concept

drift has occurred. We use the Heterogeneous Euclidean-Overlap Metric (HEOM) [14] for calculating the distance between two feature vectors. HEOM has been known to perform well with vectors consisting of both numerical and categorical attributes, hence it fits perfectly with the dataset discussed in the paper. Once concept drift is detected, we retrain our supervised learning model on the new data to improve the model's ability to detect URLs with manipulated features. Occurrence of a concept drift can also be validated by a steep fall in the accuracy of our classification model on our validation set.

Algorithm 1 Concept Drift Detection

1: **Inputs:**
2: **Mal_old** ← Array of feature vectors for all malicious URLs (in training data) created before concept drift
3: **Ben_old** ← Array of feature vectors for all benign URLs (in training data) created before concept drift
4: **Mal_new** ← Array of feature vectors for all malicious URLs (in training data) created after concept drift
5: **Ben_new** ← Array of feature vectors for all benign URLs (in training data) created after concept drift
6: **Ouput:**
7: **1** : Concept Drift detected
8: **0** : No Concept Drift detected
9: **Algorithm:**
10: total_distance ← 0
11: total_count ← 0
12: **for** each item e **in** Mal_new **do**
13:     distance ← infinite
14:     **for** each item f **in** Mal_old **do**
15:         distance ← min(distance, **HEOM(e,f)**)
16:     **end for**
17:     **if** $distance$!= 0 **then**
18:         total_distance ← total_distance + distance
19:         total_count ← total_count + 1
20:     **end if**
21: **end for**
22: **for** each item e **in** Ben_new **do**
23:     distance ← infinite
24:     **for** each item f **in** Ben_old **do**
25:         distance ← min(distance, **HEOM(e,f)**)
26:     **end for**
27:     **if** $distance$!= 0 **then**
28:         total_distance ← total_distance + distance
29:         total_count ← total_count + 1
30:     **end if**
31: **end for**
32: mean_distance ← total_distance / total_count
33: **if** $mean\_distance > threshold$ **then**
34:     **return** 1
35: **end if**
36: **return** 0

IV. Evaluation

We first evaluate the performance of our supervised classification algorithms in detecting malicious websites. Table II provides an overview of the classification performance obtained on our test set for three algorithms. To evaluate the performance, we used the following metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN}.$$

Where TP, TN, FP and FN denotes the number of True Positives, True Negatives, False Positives and False Negatives respectively. We observe that Gradient Boosting Algorithm is the most effective in classifying malicious and benign URLs with an accuracy of 96.4% for a maximum depth (longest path from root to leaf for any estimator tree in the ensemble) of 4. Figure 2 depicts the trends in accuracy achieved by Gradient Boosting for different hyperparameters. As illustrated in the figure, we observe that for fewer estimators in the ensemble, the accuracy is low due to underfitting. The accuracy increases with an increase in the number of estimators reaching a maximum value of 96.4% for 80 estimators. Thereafter, the accuracy falls steeply due to overfitting on the training dataset. We observe that the effectiveness of these algorithms in detecting malicious website begins to decrease when a concept drift occurs. In order to assess the performance of the concept drift detection algorithm, we create an artificial concept drift in our dataset. We start interchanging the target labels of our data such that a fraction of the malicious URLs become benign, and the respective benign URLs become malicious. We interchange the labels of 10% of the URLs in the database and then apply our algorithm to the artificial dataset. Our detection algorithm returns a value of 0.32, which is higher than our pre-set threshold value of 0.20. Thus, we can claim that our algorithm is able to successfully recognize concept drifts.
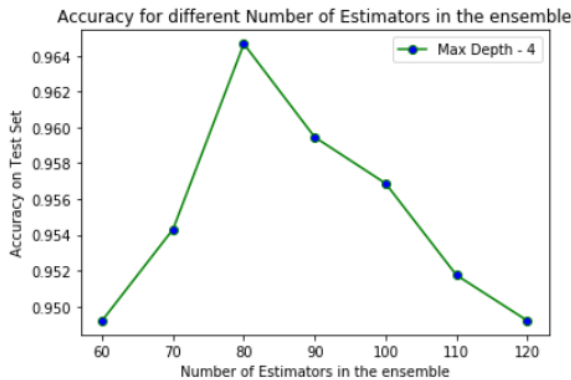
Fig. 2: Accuracy curve for Gradient Boosted Decision Trees at Maximum Depth 4

Model	Accuracy(%)	Precision(%)	Recall(%)
Random Forest	93	92.4	91.8
Gradient Boosting	**96.4**	**96.2**	**96.3**
Neural Network	95.4	94.5	93.8

TABLE II: Classification performance for different supervised learning models

V. Conclusion

In this paper, we have proposed a robust and novel approach for assisting users in detecting malicious and harmful websites and URLs. The dataset used for training and testing has been curated from a variety of sources. Our algorithm achieves a superior accuracy of 96.4% on the test dataset. We have proposed a novel approach for detecting and mitigating attempts by attackers for circumventing malicious website detection algorithms. Results show that the proposed algorithm is able to successfully detect concept drifts.

References

[1] "Malware domain list," https://www.malwaredomainlist.com/.
[2] G. Chakraborty and T. T. Lin, "A url address aware classification of malicious websites for online security during web-surfing," *2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*.
[3] H. Choi, B. B. Zhu, and H. Lee, "Detecting malicious web links and identifying their attack types," *WebApps'11, Proceedings of the 2nd USENIX conference on Web application development, pp.11-11, Berkeley, USA (2011)*.
[4] A. P. Felt, R. Barnes, A. King, C. Palmer, C. Bentzel, and p. Tabriz, "Measuring https adoption on the web," *26th USENIX Security Symposium, 2017, pp. 1323–1338*.
[5] Y. Fukushima, Y. Hori, and K. Sakurai, "Proactive blacklisting for malicious web sites by reputation evaluation based on domain and ip address registration," *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*.
[6] A. Gorji and M. Abadi, "Detecting obfuscated javascript malware using sequences of internal function calls," *Proceedings of the 2014 ACM Southeast Regional Conference on - ACM SE '14, 2014*.
[7] T. Guolin, Z. Peng, L. Qingyun, L. Xinran, Z. Chunge, and D. Fenghu, "Adaptive malicious url detection: Learning in the presence of concept drifts," *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications*.
[8] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning under concept drift: A review," *IEEE Transactions on Knowledge and Data Engineering, 2018*.
[9] Majestic, "Majestic million," https://majestic.com/reports/majestic-million.
[10] H. Mekky, R. Torres, Z.-L. Zhang, S. Saha, and A. Nucci, "Detecting malicious http redirections using trees of user browsing activity," *INFOCOM, 2014 Proceedings IEEE. IEEE, 2014, pp. 1159–1167*.
[11] S. Morishige, S. Haruta, H. Asahina, and I. Sasase, "Obfuscated malicious javascript detection scheme using the feature based on divided url," *Communications (APCC), 2017 23rd Asia-Pacific Conference on. IEEE, 2017, pp. 1–6*.
[12] OpenDNS, "Phishtank," https://www.phishtank.com.
[13] A. K. Singh and N. Goyal, "A comparison of machine learning attributes for detecting malicious websites," *2019 11th International Conference on Communication Systems Networks (COMSNETS)*.
[14] D. R. Wilson and T. R. Martinez, "Improved heterogeneous distance functions," *Journal of Artificial Intelligence Volume 6 Issue 1, January 1997 Pages 1-34*.
[15] C. Wressnegger, F. Yamaguchi, D. Arp, and K. Rieck, "Technical report - analyzing and detecting flash-based malware using lightweight multi-path exploration," *University of Gottingen, Germany, no. December, 2015. [Online]*.

OPEN

Microstates-based resting frontal alpha asymmetry approach for understanding affect and approach/withdrawal behavior

Ardaman Kaur[1,2], Vijayakumar Chinnadurai[1]* & Rishu Chaujar[2]

The role of resting frontal alpha-asymmetry in explaining neural-mechanisms of affect and approach/withdrawal behavior is still debatable. The present study explores the ability of the quasi-stable resting EEG asymmetry information and the associated neurovascular synchronization/desynchronization in bringing more insight into the understanding of neural-mechanisms of affect and approach/withdrawal behavior. For this purpose, a novel frontal alpha-asymmetry based on microstates, that assess quasi-stable EEG scalp topography information, is proposed and compared against standard frontal-asymmetry. Both proposed and standard frontal alpha-asymmetries were estimated from thirty-nine healthy volunteers resting-EEG simultaneously acquired with resting-fMRI. Further, neurovascular mechanisms of these asymmetry measures were estimated through EEG-informed fMRI. Subsequently, the Hemodynamic Lateralization Index (HLI) of the neural-underpinnings of both asymmetry measures was assessed. Finally, the robust correlation of both asymmetry-measures and their HLI's with PANAS, BIS/BAS was carried out. The standard resting frontal-asymmetry and its HLI yielded no significant correlation with any psychological-measures. However, the microstate resting frontal-asymmetry correlated significantly with negative affect and its neural underpinning's HLI significantly correlated with Positive/Negative affect and BIS/BAS measures. Finally, alpha-BOLD desynchronization was observed in neural-underpinning whose HLI correlated significantly with negative affect and BIS. Hence, the proposed resting microstate-frontal asymmetry better assesses the neural-mechanisms of affect, approach/withdrawal behavior.

Understanding the neural mechanisms associated with functional hemispheric asymmetry of affect, approach/withdrawal measures is one of the core focuses in neuroscience. Numerous studies revealed an association of functional hemispheric asymmetry with positive/negative affect and approach/withdrawal dichotomy. This linkage was initially observed in many studies where left hemispheric lesion affected the perception of positive emotions whilst damage to the right hemisphere impaired the perception of negative emotions[1–3]. Subsequently, there was a surge in elucidating the role of frontal hemispheric asymmetry based on the alpha signature of electroencephalography (EEG) in manifesting the individual differences in affect and approach/withdrawal measures[4–6]. Davidson *et al.*[7–9], in their studies, suggested the lateralization of the prefrontal cortex (PFC) with respect to positive/motivational valence. Thus, the right PFC was observed to be linked with avoidance/negative emotion and left PFC with approach/positive emotion. Nevertheless, Carver and Harmon-Jones[10] showed the association of left hemisphere with negative emotion anger and thus proposed to eliminate the differentiation of positive and negative valence from the affective model. Subsequently, a larger number of studies concentrated on EEG frontal asymmetry through the induction of emotional/motivational states or tasks to understand the neural mechanisms associated with the evoked approach/withdrawal behavior[11–18] and other specific tasks[19]. This has led to ample literature which examined alterations in frontal EEG asymmetry in clinical and healthy populations[20–28].

Although the aforementioned studies have proved EEG based frontal asymmetry assessment as a reliable indicator of affect, approach/withdrawal behavior during emotional tasks, it's validity in healthy individuals during resting still remains ambiguous. In one large resting EEG study, Tomarken *et al.*[29,30] revealed a significant

[1]NMR Research Centre, Institute of Nuclear Medicine and Allied Sciences, Lucknow Road, Timarpur, Delhi, 110054, India. [2]Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi, 110042, India. *email: vijayakumar@inmas.drdo.in

negative correlation of resting Frontal asymmetry (FA; channel pair: F4, F3) with negative affect and positive correlation of resting Anterior Temporal Asymmetry (ATA; channel pair: T4, T3) with positive affect for female subjects. Jacobs and Snyder[31], in their study, revealed the negative correlation of resting Frontal Temporal Asymmetry (FTA; channel pair: F8, F7) with negative affect in men, further Hall and Petruzzello[32] showed that resting FA positively predicted the positive affect of both sexes. Pertaining to approach and withdrawal measures, studies by Harmon-Jones and Allen[33] and De Pascalis *et al.*[34] reported a significant positive correlation of approach measure, Behavioral Activation System (BAS) with resting FA. The aforementioned studies are in sync with the hypothesis that positive affect correlates positively with alpha asymmetry $(\ln(\alpha lpha^{Right}) - \ln(\alpha lpha^{Left}))$ and links to the left hemisphere, howbeit negative affect correlates negatively with the same and associates with the right hemisphere. Conversely, in another study[35], absolutely no significant relationship was observed between resting FA and measures of positive and negative valence for both sexes. Similarly, Schneider *et al.*[36] observed an absence of correlation between resting alpha FA and measures of approach/withdrawal behavior. In contradiction to the above hypothesis, Hagemann *et al.*[37] showed that subjects exhibiting greater relative left-hemispheric resting cortical activation at the anterior temporal site reported more intense NA in response to negative stimuli. Further, in the same line of research[38], it was found that subjects scoring high on NA, demonstrated greater relative left-sided resting cortical activation at the anterior temporal region than subjects scoring low on NA.

Most findings of the aforementioned literature are based on two fundamental assumptions. Firstly, the above studies assume the acquired EEG to possess only stable cognitive information. Hence, these studies correlate the single session EEG information directly with affect and approach/ withdrawal measures. However, many studies[29,39] revealed that the stable EEG patterns across previous sessions showed the interrelation of affect and approach/ withdrawal measures with frontal alpha asymmetry. This brings the importance of assessing the stable EEG patterns and information from single session recordings as unstable EEG information may be influenced by interference from many cognitive factors. Recent EEG studies of wakeful rest have shown that global electrical brain activity on scalp remains semi-stable for transient periods[40,41]. In specifics, there exists a finite number of scalp potential topographies in spontaneous resting EEG activity that remains stable for a definite time before rapidly shifting to a different topography that once again attains a stable state. These distinct epochs of topographic stability have been referred to as 'EEG microstates'. Lehman *et al.*[42] substantiated that EEG microstates represent blocks of consciousness, and these microstates are modulated by the content of the thoughts. Additionally, Milz *et al.*[43] postulated the role of intracranial sources in the alpha band in predominantly determining these EEG microstate topographies. Further, Shafi *et al.*[44], in their study, highlighted the role of microstates in individual variability of human fluid intelligence and in response to cognitive training. Howbeit, there is no study to date that has explored the quasi-stable state as assessed by EEG microstates for understanding frontal hemispheric asymmetry. Also, their ability over standard EEG frontal asymmetry in explaining affect and approach/withdrawal dichotomy is still unmapped.

Further, the second important assumption is that EEG alpha power is inversely[45–47] related to neural activation. Hence, an increase in neural activation in the left hemisphere is associated with the increase observed in frontal asymmetry scores. This enables us in concluding that the positive correlation of affect and approach/ withdrawal measures with frontal asymmetry score $(\ln(\alpha lpha^{Right}) - \ln(\alpha lpha^{Left}))$ is the resultant of left hemispherical neuronal activity and vice versa. However, recently, many neuro-vascular studies[48–51] have observed alpha-BOLD synchronization wherein the alpha power correlates positively with neural activation during task engagement. Hence, there is a need to fully understand the neurovascular coupling and neural underpinning associated with frontal EEG asymmetry[5] and how alpha-BOLD synchronization or desynchronization during resting-state associates with affect and approach/withdrawal behavior. Few researchers brought better understanding by studying the role of hemispheric asymmetry in affect, approach/avoidance behavior through functional MR imaging. Rohr *et al.*[52] concluded that the affective elements in the underlying organization of emotion are predominantly associated with the network of right-hemispheric regions. Lindquist *et al.*[53] proposed that the implementation of valence depends on a set of valence-general limbic and paralimbic brain regions. Though the above studies gave significant insights, the congruence between resting-EEG frontal alpha asymmetry and resting-fMRI is still uncharted. This is vital for a better understanding of neuro-vascular aspects of resting frontal asymmetry and their association with affect and approach/withdrawal behavior.

Hence, the present study proposes an EEG microstate based approach for assessment of quasi-stable frontal hemispherical asymmetry measures of resting-state affect and approach/withdrawal behavior. It further aims to compare the performance of microstate based frontal hemispheric asymmetry against the standard resting EEG frontal asymmetry measures. For this purpose, resting EEG was acquired from a sample of 39 healthy male subjects. This multichannel resting-EEG signal from all subjects was parsed into a limited number of distinct quasi-stable microstates. These microstates were back-fitted to each subject's EEG data to obtain microstate time-series data specific to each subject. The microstate time-series was further filtered at alpha frequency band and EEG microstate based frontal asymmetry measures were derived from channel pairs F4/F3 (FA) and F8/F7 (FTA). Further, the robust correlation of both standard and EEG microstate based frontal hemispheric asymmetry with positive/negative affect (PANAS) and approach (BAS)/withdrawal (BIS) behavior was carried out.

Moreover, the study focuses on bringing a better understanding of neural mechanisms associated with functional hemispheric asymmetry of affect and approach/ withdrawal behavior during resting-state. For this purpose, standard and microstates based resting EEG frontal asymmetries were subjected to the EEG informed fMRI approach and the associated neural underpinning of both EEG frontal asymmetries were independently estimated. Thereafter, the hemodynamic lateralization index (HLI) based on the amplitude of hemodynamic response function (HRF) of regions part of the neural underpinning of both EEG frontal asymmetries were assessed. Further, the estimated HLI was subjected to a robust correlation with resting-state affect and approach/ withdrawal psychological scores. Finally, the results were analyzed to understand the ability of proposed EEG
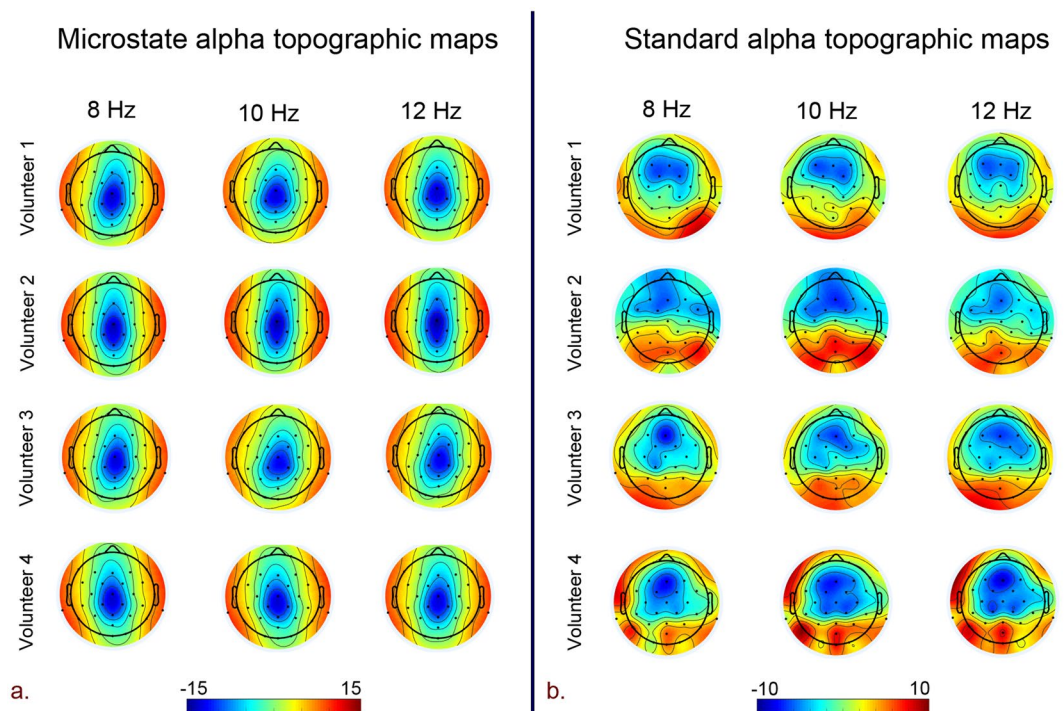
Figure 1. Topographic EEG maps of spectral power density for the alpha band for. (**a**) Proposed microstate based analysis and. (**b**) Standard analysis (CSD referenced). The color bar represents the log-transformed spectral power density ($10*\log 10$ (μv^2/Hz)) where red represents the maximum and blue represents the minimum values.

microstate estimates in revealing neural-vascular insights of association of functional hemispherical asymmetry with resting-state affect and approach/ withdrawal behavior.

Results

Our study focused on exploring the ability of quasi-stable EEG microstate based frontal alpha hemispherical asymmetry measures against standard EEG frontal alpha asymmetry measures in explaining the resting state affect and approach/ withdrawal behavior for healthy young male volunteers during 1-time measurement. The standard alpha topographic maps (CSD referenced) and microstate alpha topographic maps are shown in Fig. 1. Evidently, the maps of standard alpha topography (CSD referenced) in Fig. 1b reveal the typical parietal-occipital alpha activity for eyes-closed resting-state condition[54,55]. However, the parietal-alpha activity is typical of standard alpha topographic maps and has not been observed and reported by any researchers in microstate alpha topographic maps so far. For assessing the association of EEG microstate based frontal hemispheric asymmetry with affect and approach/withdrawal behavior, robust correlation of PANAS and BAS, BIS measures with standard and EEG microstate FA and FTA was estimated. Subsequently, to better understand the neural mechanisms underlying the proposed microstate and standard hemispherical asymmetry measures, they were subjected to the EEG informed fMRI, and their neural underpinnings were estimated. Further, to gain insights into the hemodynamic lateralization associated with the neural underpinnings and its linkage with affect and approach/withdrawal measures, HLI of both asymmetry measures neural underpinnings' was calculated and subsequently subjected to the robust correlation with PANAS and BAS, BIS measures.

Robust correlation of frontal hemispherical asymmetry measures with psychological measures. The robust correlation (Pearson, bend, spearman, and skipped) of proposed microstate and standard frontal hemispheric asymmetry measures with PANAS, BIS/BAS psychological scores are tabulated in Table 1.

Standard FA and FTA revealed no statistically significant correlation with PANAS as well as BIS/BAS measures. Similarly, proposed microstate based FA and FTA yielded insignificant low correlation with positive affect score.

Howbeit, negative affect scores revealed a strong and significant correlation with proposed microstate based FA and FTA. Specifically, microstates based FA yielded high pearson, bend and spearman correlations (Fig. 2a: pearson $r = 0.35$, 95% CI $= [0.07; 0.58]$, pcorr $= 0.04$; Fig. 2b: bend $r = 0.33$, 95% CI $= [-0.02; 0.61]$, pcorr $= 0.05$; Fig. 2c: spearman $r = 0.36$, 95% CI $= [0.04; 0.62]$, pcorr $= 0.03$). Similarly, skipped pearson and spearman robust correlations of microstates based FA with negative affect scores has also yielded stronger correlations (Fig. 2d: pearson skipped $= 0.35$, 95% CI $= [0.04; 0.57]$; spearman skipped $= 0.36$, 95% CI $= [0.005; 0.62]$). In addition, a strong robust pearson, bend and spearman correlation of microstates based FTA with negative affect scores was observed (Fig. 3a: pearson $r = 0.42$, 95% CI $= [0.13; 0.67]$, pcorr $= 0.01$; Fig. 3b: Bend $r = 0.42$, 95% CI $= [0.05;$

3

EEG alpha frontal asymmetry	Channel pair	Behavioral measure	Pearson Correlation		Bend correlation		Spearman correlation		Skipped correlation			
									Pearson		Spearman	
			r	p	r	p	r	p	r	t	r	t
Standard	F4/F3 (FA)	Positive affect	0.22	0.21	0.2	0.23	0.09	0.54	0.22	1.27	0.09	0.54
		Negative affect	−0.1	0.54	−0.04	0.8	−0.05	0.75	−0.1	−0.6	−0.05	−0.31
		BAS	−0.25	0.37	−0.16	0.56	−0.17	0.56	−0.25	−0.92	−0.17	−0.59
		BIS	−0.03	0.9	0.09	0.75	0.09	0.73	−0.03	−0.12	0.09	0.34
	F8/F7 (FTA)	Positive affect	0.03	0.83	−0.11	0.52	−0.12	0.47	0.03	0.21	−0.12	−0.73
		Negative affect	−0.05	0.75	0.01	0.92	−0.004	0.97	−0.05	−0.31	−0.004	−0.02
		BAS	0.18	0.52	0.17	0.55	0.13	0.65	0.18	0.65	0.13	0.46
		BIS	−0.14	0.62	−0.14	0.61	−0.3	0.28	−0.14	−0.5	−0.3	−1.12
Microstates	F4/F3 (FA)	Positive affect	0.03	0.84	0.08	0.61	0.12	0.46	0.03	0.2	0.12	0.73
		Negative affect	0.35	0.04	0.33	0.05	0.36	0.03	0.35	2.13	0.36	2.2
		BAS	−0.09	0.74	−0.04	0.86	0	1	−0.09	−0.32	0	0
		BIS	−0.3	0.29	−0.41	0.14	−0.28	0.32	−0.3	−1.09	−0.28	−1.01
	F8/F7 (FTA)	Positive affect	0.0003	0.99	−0.01	0.91	−0.01	0.92	0.0003	0.0018	−0.01	−0.09
		Negative affect	0.42	0.01	0.42	0.01	0.38	0.02	0.42	2.64	0.38	2.34
		BAS	−0.17	0.54	−0.18	0.52	−0.18	0.53	−0.17	−0.62	−0.18	−0.64
		BIS	−0.32	0.25	−0.45	−1.7	−0.33	−1.22	−0.32	−1.19	−0.33	−1.22

Table 1. Robust correlation (Pearson, bend, spearman and skipped) of standard and proposed microstate based frontal hemispheric asymmetry measures with psychological scores.
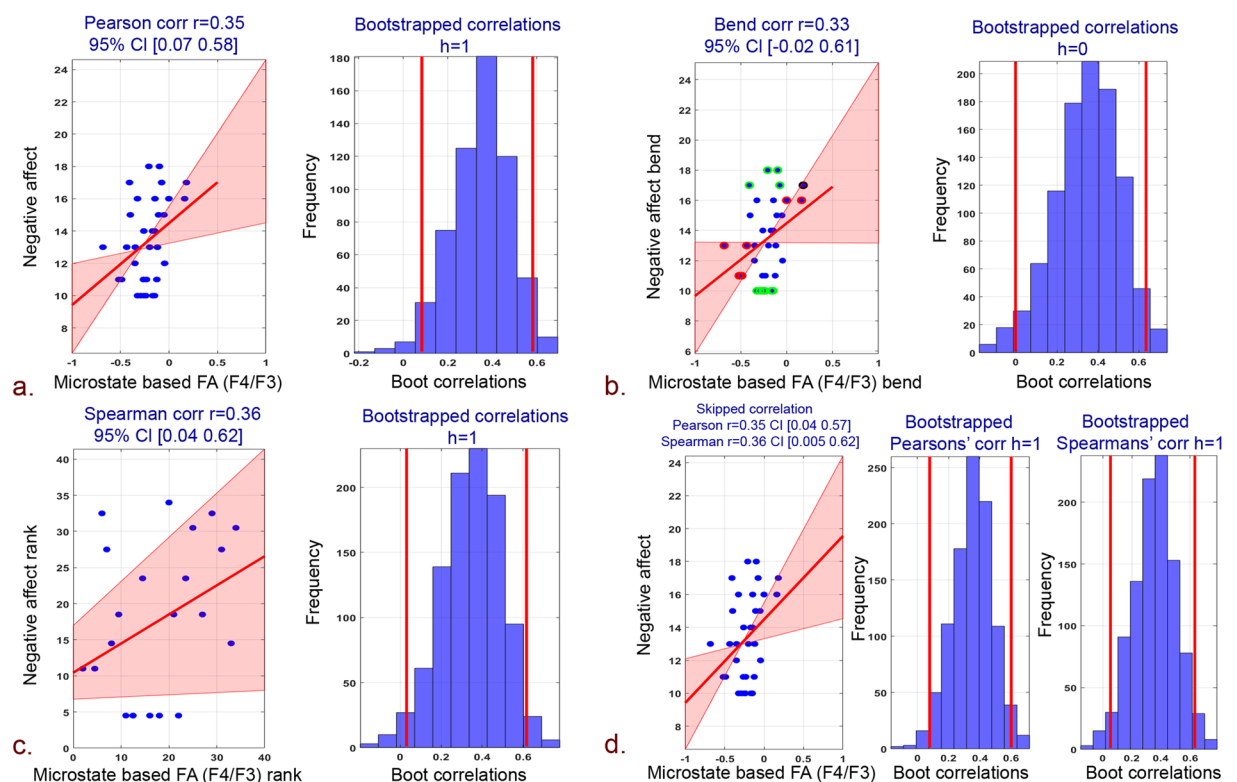
Figure 2. Correlation plots between negative affect scores and microstate based FA (F4/F3) and associated histograms of correlations for bootstrapped data. (**a**) Pearson correlation. (**b**) 20% Bend correlation. (**c**) Spearman correlation. (**d**) Skipped (Pearson and Spearman) correlations.

0.70], pcorr = 0.01; Fig. 3c: spearman r = 0.38, 95% CI = [0.02; 0.68], pcorr = 0.02). Skipped (pearson and spearman) correlations among microstates-derived FTA and negative affect scores has also yielded stronger correlations (Fig. 3d: Pearson skipped = 0.42, 95% CI = [0.14; 0.67]; Spearman skipped = 0.38, 95% CI = [0.04; 0.68]).
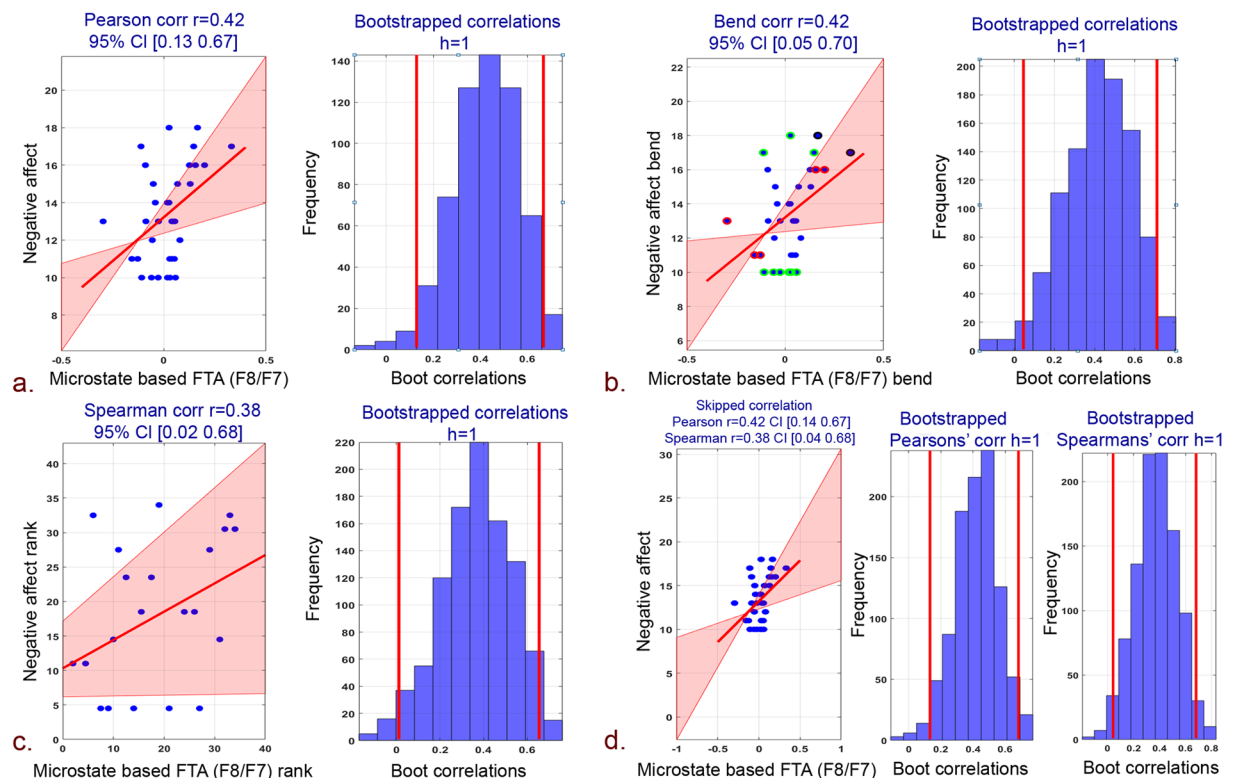
Figure 3. Correlation plots between negative affect scores and microstate based FTA (F8/F7) and associated histograms of correlations for bootstrapped data. (**a**) Pearson correlation. (**b**) 20% Bend correlation. (**c**) Spearman correlation. (**d**) Skipped (Pearson and Spearman) correlations.

However, BAS measures yielded a statistically insignificant low correlation with proposed microstate asymmetry. The analysis with BIS measures for both FA and FTA revealed high correlation, but the p-values remained insignificant.

EEG informed fMRI analysis. The proposed microstate and standard hemispherical asymmetry measures were subjected to the EEG informed fMRI analysis to assess their neural underpinnings, respectively. The observed neural underpinnings were inferred with FDR corrected p-values less than 0.05, and a cluster size of more than 20 voxels were considered for analysis.

Neural underpinnings of standard hemispheric asymmetry. Neural underpinnings of standard FA encompassed right as well as left-hemispheric regions (Fig. 4a). Table 2 comprises of these areas, their peak coordinates, and cluster size. Specifically, in the right hemisphere, EEG frontal asymmetry negatively correlated with BOLD activity in occipital cortex with major clusters in lateral occipital cortex and occipital pole. Additionally, BOLD activity in temporal cortex also correlated negatively with standard FA. However, BOLD of parietal cortex regions, particularly postcentral gyrus, correlated positively with standard FA. Withal, in the left hemisphere, standard FA correlated positively with BOLD activity in the postcentral gyrus. However, activity in the occipital fusiform gyrus and temporal lobe regions correlated negatively with this alpha asymmetry measure. Majority of frontal lobe regions correlated negatively. However, superior frontal gyrus correlated positively (high t-value as compared to the right hemisphere) with standard FA.

Figure 4b shows the neural underpinnings of standard FTA. Both right and left hemispheres revealed positive as well as negative correlations between BOLD activity and standard FTA (Table 3). In the right hemisphere, BOLD activity in occipital lobe regions (cuneal cortex, lingual gyrus, and superior division of lateral occipital cortex) correlated negatively with standard FTA. Major clusters in the frontal lobe, specifically frontal pole, and activity in precuneus cortex also found a negative correlation with this frontal asymmetry index. In the left hemisphere, standard FTA correlated negatively with BOLD activity in Inferior frontal gyrus. Few clusters in parietal, occipital and temporal pole also correlated negatively with standard FTA. The neural underpinnings of standard FA showed left-hemispheric dominance whilst FTA revealed right-hemispheric dominance.

Neural underpinnings of microstate based EEG asymmetry. Right and left-lateralized neural underpinnings of microstate based FA are shown in Fig. 5a. A complete list of activation clusters is provided in Table 4. In the right hemisphere, microstate based FA correlated negatively with BOLD activity in the frontal medial cortex and frontal pole regions of the frontal lobe. Similarly, BOLD activity in the posterior division of cingulate gyrus has also correlated negatively. However, few clusters in the frontal lobe, occipital lobe, and temporal
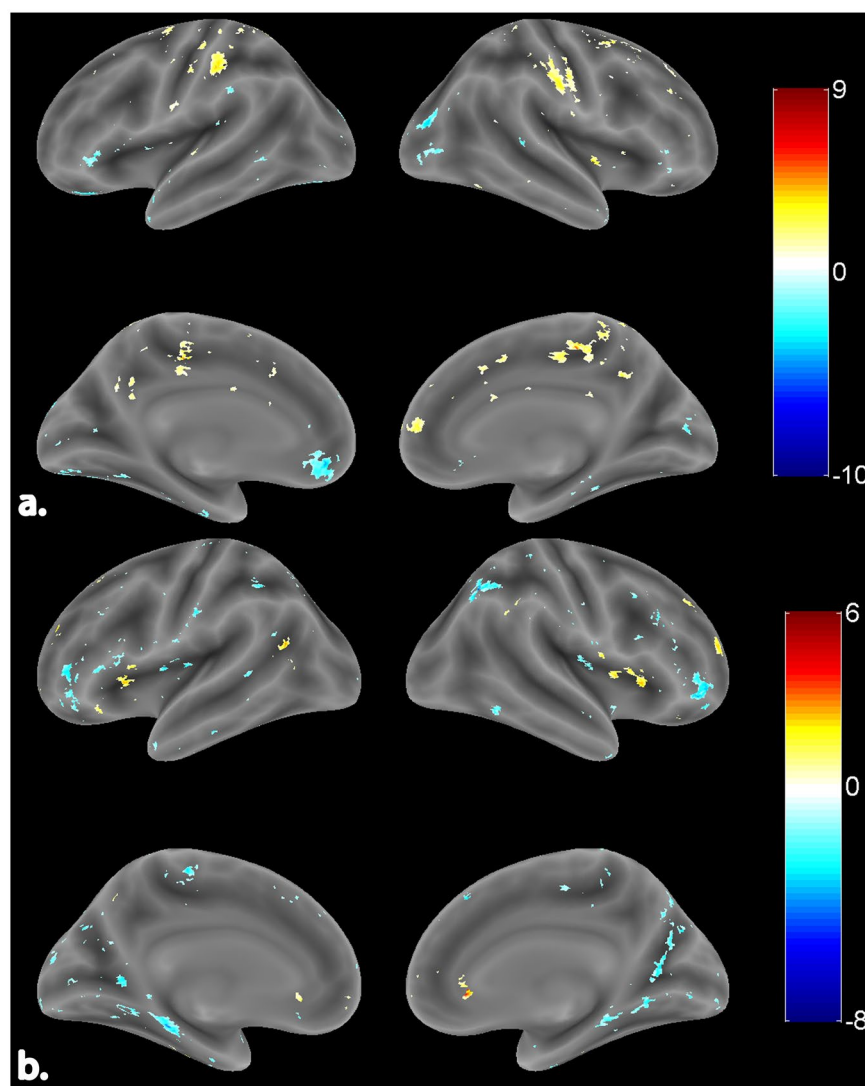
Figure 4. Surface rendered view of neural underpinnings of standard. (**a**) FA (channel pair F4/F3). (**b**) FTA (channel pair F8/F7). The color bar indicates the t-values with blue being the least and red being the highest. The activations are represented at FDR corrected $p < 0.05$.

pole revealed a positive correlation with microstate FA. In the left hemisphere, resting-state microstate based FA correlated positively with major clusters in all lobes with frontal lobe having the maximum cluster extent. This is evident as microstates are known to represent the global brain activity.

Figure 5b shows the neural underpinnings in both right and left hemispheres for resting-state microstate based FTA. Table 5 comprises of these areas, their peak coordinates, and cluster size. In the right hemisphere, activity in the frontal lobe and limbic lobe regions correlated negatively with this EEG alpha asymmetry. BOLD of specific regions of the parietal lobe (Angular gyrus) and temporal lobe (Planum temporale) correlated negatively with microstate based FTA. Pertaining to the left hemisphere, activations in the frontal lobe and limbic lobe (a posterior division of cingulate gyrus) correlated negatively with microstate based FTA. Negative correlation also emanated from BOLD activity in specific regions of the parietal lobe (Angular gyrus, Superior parietal lobule) and lateral occipital cortex of occipital lobe. The neural underpinnings for microstate based FA and FTA showed left-hemispheric dominance.

Robust correlation of HLI with PANAS, BIS/BAS measures. The correlation and p-values for all the significant results obtained for this analysis are tabulated in Table 6. The robust correlation between negative affect measure and HLI of neural underpinnings of microstate frontal alpha asymmetry yielded a significantly strong negative correlation in the anterior division of the middle temporal gyrus. Further, superior frontal gyrus emerged as the positive correlate for correlation among positive affect scores and HLI pertaining to neural underpinnings of microstate alpha asymmetry. Moreover, the correlation of BIS measure with HLI pertaining to neural underpinnings of microstate frontal alpha asymmetry yielded a significantly strong positive correlation in inferior frontal gyrus (pars triangularis) and frontal medial cortex. Further, the HLI of occipital fusiform gyrus correlated negatively with BAS measure.

Cluster Center Region	Cluster No.	Voxels	MNI Coordinates			T-Stats
Right Hemispheric Activations						
Frontal Lobe						
Superior frontal gyrus	1	43	22	4	48	4.271
Parietal Lobe						
Postcentral gyrus	1	246	50	−20	38	6.414
	2	97	16	−28	44	8.087
	3	33	14	−44	60	3.886
	4	23	48	−26	64	−3.349
Superior parietal lobule	1	48	18	−46	64	4.345
Occipital Lobe						
Lateral occipital cortex, superior division	1	96	38	−86	14	−4.214
Lateral occipital cortex, inferior division	1	31	36	−72	−30	−2.84
	2	20	46	−80	2	−3.317
Occipital pole	1	96	22	−90	26	−4.04
Intracalcarine cortex	1	27	10	−80	10	−2.747
Temporal Lobe						
Temporal pole	1	20	42	14	−32	−2.74
Limbic Lobe						
Paracingulate gyrus	1	34	12	50	10	3.121
Insular cortex	1	38	34	−6	−2	9.396
Left Hemispheric Activations						
Frontal Lobe						
Superior frontal gyrus	1	118	−24	−4	62	7.828
	2	26	−6	56	30	−3.36
Frontal medial cortex	1	117	−12	42	−10	−5.587
Frontal orbital cortex	1	60	−22	26	−18	−6.658
Precentral gyrus	1	60	−16	−26	40	6.636
Inferior frontal gyrus	1	27	−46	30	−2	−3.36
Parietal Lobe						
Postcentral gyrus	1	147	−46	−26	38	5.015
	2	36	−36	−28	70	3.992
	3	21	−62	−8	22	2.715
Superior parietal lobule	1	132	−30	−46	64	7.511
Occipital Lobe						
Occipital fusiform gyrus	1	159	−22	−84	−10	−3.179
Occipital pole	1	29	−16	−90	30	−3.58
	2	20	−2	−98	0	−2.736
Temporal Lobe						
Temporal Occipital Fusiform Cortex	1	159	−28	−66	−22	−2.946
Temporal Fusiform cortex, posterior division	1	39	−36	−14	−26	−3.441
Limbic Lobe						
Cingulate gyrus, posterior division	1	34	−8	−54	28	4.73

Table 2. Neural underpinnings of standard FA (channel pair F4/F3). The activations after correction for multiple comparisons are represented at $p < 0.05$ (FDR corrected). The coordinates reported are in Montreal Neurological Institute (MNI) space.

However, the robust correlation between negative affect measure and HLI of neural underpinnings of standard frontal alpha asymmetry yielded low and insignificant correlation with all cortical regions. Whilst correlation of positive affect scores with HLI pertaining to standard alpha asymmetry revealed a significant positive correlation with the insular cortex. Further, the correlation of BAS and BIS measures with HLI revealed a low and insignificant correlation with all cortical regions pertaining to standard alpha asymmetry.

Robust correlation among frontal hemispherical asymmetry measures. Figure 6 shows the Pearson robust correlation of proposed microstate frontal hemispheric asymmetry with standard frontal hemispheric asymmetry measures. Proposed microstate based FA and FTA yielded insignificant low correlation with

Cluster Center Region	Cluster No.	Voxels	MNI Coordinates			T-Stats
Right Hemispheric Activations						
Frontal Lobe						
Frontal pole	1	67	26	54	22	3.822
	2	385	46	38	10	−5.757
	3	385	30	48	−12	−5.113
	4	385	50	44	−10	−2.962
Subcallosal cortex	1	25	6	30	−4	6.522
	2	20	6	14	−4	−3.574
Middle frontal gyrus	1	94	50	14	36	−8.063
Precentral gyrus	1	27	50	6	40	−3.411
Parietal Lobe						
Precuneous cortex	1	392	28	−52	10	−4.401
	2	392	22	−66	26	−2.890
Supramarginal gyrus, posterior gyrus	1	36	64	−46	32	3.054
Occipital Lobe						
Cuneal cortex	1	392	8	−78	38	−5.546
Lateral occipital cortex, superior division	1	174	34	−62	46	−6.328
Occipital fusiform cortex	1	149	26	−68	−26	−4.764
Lingual gyrus	1	210	14	−58	−4	−2.931
	2	48	2	−76	0	−2.582
Occipital pole	1	20	8	−96	2	−2.553
Temporal Lobe						
Inferior temporal gyrus, temporooccipital part		39	56	−54	−14	−3.061
Central operculum cortex		28	36	−12	22	−2.751
Limbic Lobe						
Insular cortex		56	30	20	8	3.558
Left Hemispheric Activations						
Frontal Lobe						
Inferior frontal gyrus	1	305	−50	32	16	−3.912
	2	20	−46	16	26	−2.703
Middle frontal gyrus	1	42	−52	22	30	−3.196
	2	21	−50	14	36	−3.502
Frontal operculum cortex	1	46	−34	18	12	3.808
Precentral gyrus	1	36	−6	−26	52	−2.984
	2	22	−32	−20	72	−3.16
Parietal Lobe						
Supramarginal gyrus, posterior division	1	32	−36	−44	36	−2.772
Postcentral gyrus	1	24	−62	−14	36	−2.86
Occipital Lobe						
Occipital pole	1	60	−4	−94	22	−2.882
Lateral occipital cortex, superior division	1	26	−8	−86	38	−2.893
Lingual gyrus	1	42	−24	−54	2	−2.896
Temporal Lobe						
Temporal pole	1	40	−50	10	−28	−3.633
Limbic Lobe						
Parahippocampal gyrus, posterior division	1	175	−10	−38	−22	−5.375
Parahippocampal gyrus, anterior division	1	23	−30	−10	−30	−3.412
Cingulate gyrus, posterior division	1	22	−10	−40	2	−2.716

Table 3. Neural underpinnings of standard FTA (channel pair F8/F7). The activations after correction for multiple comparisons are represented at p < 0.05 (FDR corrected). The coordinates reported are in Montreal Neurological Institute (MNI) space.
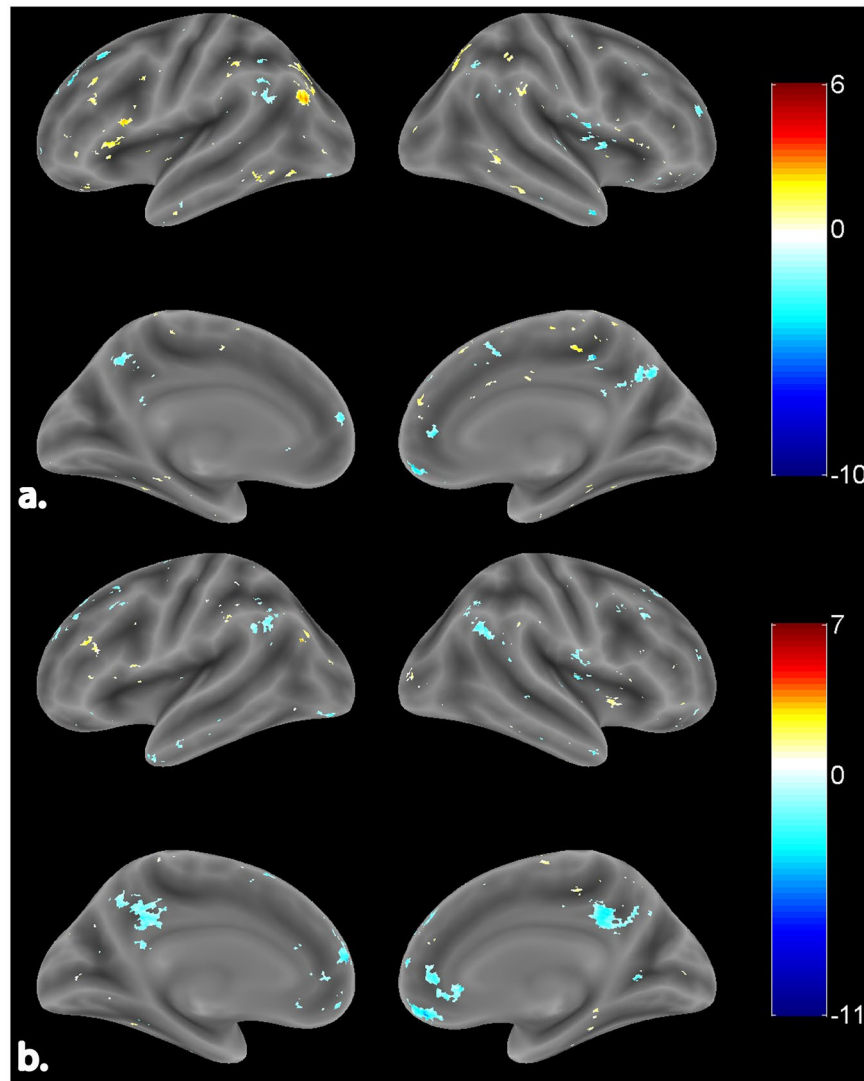
Figure 5. Surface rendered view of neural underpinnings of proposed microstate based. (**a**) FA (channel pair F4/F3). (**b**) FTA (channel pair F8/F7). The color bar indicates the t-values with blue being the least and red being the highest. The activations are represented at FDR corrected p < 0.05.

standard FA and FTA. Pearson correlation among standard and microstate based FA and FTA revealed correlation coefficients and p-values as Pearson r = −0.14, 0.013; pcorr = 0.37, 0.93 respectively.

Discussion

Valence[56–59] and motivation hypothesis[60] propose that higher values of positive affect/approach behavior and negative affect/withdrawal behavior are associated with the greater relative left and right cortical activation, respectively. These hypotheses are established in task-based EEG alpha asymmetry studies where the implications are based on alpha inhibition (desynchronization w.r.t BOLD) in event-specific regions[45,47]. Thus, following this abstraction, the above-mentioned hypothesis holds when standard frontal hemispheric asymmetry $(\ln(\alpha^{Right}) - \ln(\alpha^{Left}))$ correlates positively with positive affect/ approach behavior and negatively with negative affect/withdrawal behavior. Howbeit, the validity of these hypotheses in resting-state recordings which involves sole perception and not induction of valence/ behavior still remains vacillating. The inconsistent results of the relationship between the standard resting frontal asymmetry and affect and approach/withdrawal behavior are tabulated in Table 7. The line of studies by Tomarken et al.[29,30] and Jacob and Snyder[31] supported the above hypothesis. Similarly, for approach/withdrawal dichotomy, Harmon-Jones and Allen[33], Shackman et al.[61], and De Pascalis et al.[34] supported the above-mentioned hypotheses. Nonetheless, Sutton and Davidson[35] and Schneider et al.[36] observed no association of affect, approach/withdrawal dichotomy with frontal asymmetry, respectively. Conversely, the study by Hagemann et al.[38] proposed that subjects with high negative affect exhibited high left cortical activation. Further, Hewig et al.[62] propounded a higher approach measure to be associated with the bilateral frontal cortical activity. Hence, in order to bring more clarity, the present study aims to assess the capability of quasi-stable microstates based frontal hemispheric asymmetry in explaining the affect and approach/ withdrawal dichotomy as against standard frontal hemispheric asymmetry.

Cluster Center Region	Cluster No.	Voxels	MNI Coordinates			T-Stats
Right Hemispheric Activations						
Frontal Lobe						
Middle frontal gyrus	1	32	42	34	40	5.93
	2	26	44	4	58	3.035
Superior frontal gyrus	1	24	4	14	60	3.421
Frontal operculum cortex	1	20	40	22	4	3.036
Frontal medial cortex	1	59	4	44	−14	−6.743
Frontal pole	1	20	22	48	18	−3.366
Parietal Lobe						
Precuneous cortex	1	23	6	−50	66	4.279
Occipital Lobe						
Lateral occipital cortex, superior division	1	64	22	−58	48	4.138
Temporal Lobe						
Temporal occipital fusiform cortex	1	60	32	−40	−28	3.691
Middle temporal gyrus, temporoccipital part	1	42	62	−50	−8	4.211
Inferior temporal gyrus,temporoccipital part	1	22	54	−38	−18	3.538
Central operculum cortex	1	22	50	−8	10	−4.769
Middle temporal gyrus, anterior division	1	23	52	0	−36	−5.849
Limbic Lobe						
Insular cortex	1	48	40	14	−4	6.748
	2	29	36	2	4	−3.17
Parahippocampal gyrus, posterior division	1	27	36	−28	−10	5.738
Cingulate gyrus, posterior division	1	106	8	−52	28	−3.381
Left Hemispheric Activations						
Frontal Lobe						
Inferior frontal gyrus	1	82	−54	10	14	5.142
Frontal operculum cortex	1	37	−40	24	6	3.055
Middle frontal gyrus	1	29	−42	30	42	2.89
	2	42	−24	24	36	−7.953
Frontal pole	1	42	−22	40	32	−4.002
Precentral gyrus	1	41	−36	−10	66	−3.66
Parietal Lobe						
Supramarginal gyrus, anterior division	1	57	−60	−30	46	3.13
Supramarginal gyrus, anterior division	2	23	−44	−36	44	3.112
Postcentral gyrus	1	23	−14	−38	76	3.168
Supramarginal gyrus, posterior division	1	22	−54	−42	54	3.485
	2	35	−38	−48	36	−3.261
Precuneous cortex	1	144	−4	−58	42	−4.005
Occipital Lobe						
Lateral occipital cortex, superior division	1	74	−42	−74	28	4.041
	2	50	−28	−62	30	2.992
	3	41	−30	−78	36	3.427
Lateral occipital cortex, inferior division	1	24	−30	−82	−28	−3.625
Occipital fusiform gyrus	1	21	−34	−86	−20	−3.382
Temporal Lobe						
Temporal fusiform cortex, posterior division	1	66	−38	−48	−32	3.132
Inferior temporal gyrus, temporoocipital part	1	23	−58	−54	−14	2.967
Limbic Lobe						
Parahippocampal Gyrus, posterior division	1	66	−22	−36	−20	3.035

Table 4. Neural underpinnings of proposed microstate based FA (channel pair F4/F3). The activations after correction for multiple comparisons are represented at $p < 0.05$ (FDR corrected). The coordinates reported are in Montreal Neurological Institute (MNI) space.

Prelude to the present research study.   This study primarily focuses on exploring the ability of EEG microstates based frontal hemispherical asymmetry measure against standard Davidson's approach in explaining mechanisms of the resting state affect and approach/ withdrawal behavior. The rationale for examining EEG microstates-derived frontal asymmetry was based on the specific observation that affect and approach/

Cluster Center Region	Cluster No.	Voxels	MNI Coordinates			T-Stats
Right Hemispheric Activations						
Frontal Lobe						
Frontal medial cortex	1	192	2	42	−12	−11.711
Frontal pole	1	81	8	50	42	−3.819
	2	25	8	58	14	−2.958
Subcallosal cortex	1	68	6	28	−4	−4.043
Superior frontal gyrus	1	23	16	32	56	−3.107
Parietal Lobe						
Angular gyrus	1	104	50	−56	28	−3.685
Occipital Lobe						
Lingual gyrus	1	20	26	−56	2	−3.676
Temporal Lobe						
Planum Temporale	1	124	62	−12	6	−3.416
Limbic Lobe						
Cingulate gyrus, posterior division	1	203	4	−44	38	−5.841
Insular cortex	1	124	36	−12	14	−5.241
Left Hemispheric Activations						
Frontal Lobe						
Middle frontal gyrus	1	25	−42	34	24	3.014
	2	51	−26	20	38	−3.049
	3	25	−42	18	48	−3.022
Superior Frontal Gyrus	1	138	−4	52	36	−3.021
	2	30	−6	40	50	−4.585
	3	21	−2	14	66	−3.298
Frontal pole	1	138	−8	58	14	−6.323
	2	75	−20	52	30	−3.548
	3	21	−20	44	38	−2.748
Precentral gyrus	1	21	−36	−12	68	−2.877
Parietal Lobe						
Supramarginal gyrus, anterior division	1	87	−62	−28	40	7.244
	2	41	−44	−36	46	2.927
Angular gyrus	1	181	−46	−56	54	−3.641
	2	181	−58	−54	36	−3.615
Superior Parietal Lobule	1	181	−34	−52	38	−2.478
Occipital Lobe						
Lateral occipital cortex, inferior division	1	91	−30	−88	−18	−3.47
Temporal Lobe						
Temporal pole	1	47	−44	10	−36	−4.092
Limbic Lobe						
Cingulate gyrus, posterior division	1	386	−6	−48	36	−5.096
	2	386	−4	−44	14	−2.506

Table 5. Neural underpinnings of proposed microstate based FTA (channel pair F8/F7). The activations after correction for multiple comparisons are represented at $p < 0.05$ (FDR corrected). The coordinates reported are in Montreal Neurological Institute (MNI) space.

withdrawal measures associated significantly with stable EEG signatures. Microstate analysis estimates the global pattern of coherence across entire EEG channels from temporal EEG data and thus assesses patterns of quasi-stable activities. The interaction within a large scale brain network involves a rapid change in the dynamics of these quasi-stable activity patterns. Further, the neural mechanism associated with any cognitive process generally involves the coordinated activity of many neural assemblies located at different cortexes. Correspondingly, the neural mechanisms of resting-state affect and approach/ withdrawal behavior are also the result of one such coordinated activity of the large scale brain networks.

Thus, in this study, a novel approach is explored, which assesses the frontal hemispherical asymmetry of quasi-stable activity patterns (microstates) from large scale brain interactions of the resting state affect and approach/ withdrawal behavior. These EEG microstates based frontal hemispherical asymmetry measures are further subjected to the EEG informed fMRI analysis to estimate their neural underpinnings. Subsequently, the lateralization index, which measures the hemispherical asymmetry of these large scale brain networks based on their hemodynamic information, is measured and correlated with affect and approach/ withdrawal psychological

Hemodynamic Lateralization Index (HLI)	Behavioral measure	Cortical regions	Pearson Correlation		Bend correlation		Spearman correlation		Skipped correlation			
									Pearson		Spearman	
			r	p	r	p	r	p	r	t	r	t
Standard neural underpinnings	Negative affect	No region survived	—	—	—	—	—	—	—	—	—	—
	BIS	No region survived	—	—	—	—	—	—	—	—	—	—
	Positive affect	Insular cortex	0.44	0.004	0.44	0.005	0.4	0.01	0.53	3.82	0.41	2.75
	BAS	No region survived	—	—	—	—	—	—	—	—	—	—
Microstates Neural underpinnings	Negative affect	Middle temporal gyrus, anterior division	−0.4	0.01	−0.38	0.01	−0.43	0.006	−0.4	−2.67	−0.43	−2.91
	BIS	Inferior frontal gyrus	0.69	0.005	0.63	0.01	0.69	0.005	0.69	3.36	0.69	3.39
		Frontal medial cortex	0.71	0.004	0.76	0.001	0.75	0.001	0.71	3.53	0.75	4.04
	Positive affect	Superior frontal gyrus	0.36	0.02	0.3	0.05	0.31	0.05	0.36	2.37	0.31	2.01
	BAS	Occipital fusiform gyrus	−0.58	0.02	−0.57	0.03	−0.55	0.03	−0.58	−2.51	−0.55	−2.32

Table 6. Robust correlation (Pearson, bend, spearman and skipped) of HLI based on standard and proposed microstate based frontal hemispheric asymmetry measures with psychological scores.

measures. Lastly, the insights brought by the proposed EEG microstates based approach is compared with the standard EEG asymmetry measures to understand the effectiveness of microstate derived asymmetry measures in explaining resting-state affect and approach/ withdrawal behavior. The insights of the present study are summarized in the following subsections.

Standard alpha asymmetry and its HLI reveal no correlation with PANAS and BIS/BAS measures. The current study is in line with the observation of Davidson and colleagues[35] and other earlier studies[4,36,63], where no correlation was observed for affect and BIS/BAS measures with standard hemispheric asymmetry. However, these previous studies never explored the neurovascular underpinnings and associated hemodynamic asymmetry of these underpinnings. In the present study, the absence of linkage of standard hemispheric asymmetry with affect and BIS/BAS measures is further strengthened by the lack of correlation of HLI of neural underpinnings of standard alpha asymmetry with PANAS, BIS, and BAS measures. This supports the understanding that neural mechanisms that are measured as standard EEG frontal alpha asymmetry may not be the marker to explain the affect and/or approach-withdrawal measures during resting state. It might possibly are influenced by the neural activity associated with other ongoing resting-state neural mechanisms, which limit its sensitivity towards the neural mechanisms associated with affect and approach-withdrawal measures during resting state. Thus, our finding strengthens the understanding that the standard EEG alpha asymmetry model, especially in the male population, is effectual in explaining affect or approach-withdrawal measures only when arousing situations such as those relying on mood induction procedures are present.

Microstates based asymmetry correlates with and delineates the neural mechanisms of Negative affect. In contradistinction to the standard hemispheric asymmetry, the proposed microstates based measure brings better insights into the global coordinated activity of large scale brain networks pertaining to negative affect. In this study, the robust correlational analysis revealed a positive correlation of negative affect with microstates based frontal hemispheric asymmetry. This implies that negative affect increases with an increase in right hemispheric alpha activity or a decrease in left-hemispheric alpha activity. Further, the most common neurovascular hypotheses state that when engaged in the task, the brain region exhibits suppression in alpha power with an increase in BOLD signal[47]. This causes a negative correlation between alpha power and BOLD signal and is termed as alpha-BOLD desynchronization. Figure 7 depicts these underlying dynamics for the association between alpha asymmetry measures and the BOLD signal during alpha-BOLD synchronization/ desynchronization. Following this, the positive correlation of negative affect with microstates based frontal hemispheric asymmetry implies left-hemispheric interaction with negative affect. These observations do not support the valence hypothesis explained in the earlier section but goes in line with the observations by Hagemann *et al.*[38], wherein negative affect has been linked to the left-hemisphere. Our results were also in line with a mood induction study by Gale *et al.*[64], where negative mood increased with an increase in left frontal activation. Further, recently Farahi *et al.*[65] showed the associativity of fear positively with the left hemisphere.

Additionally, neural underpinnings of microstate derived asymmetry revealed the involvement of temporal lobe regions. In this study, HLI ($HRF\_Amp_n^R - HRF\_Amp_n^L$), which was estimated by utilizing the amplitude of the HRF of each neural underpinning of microstate based frontal asymmetry linked negatively the HLI of the anterior division of middle temporal gyrus neural underpinning to negative affect. This implies that relatively
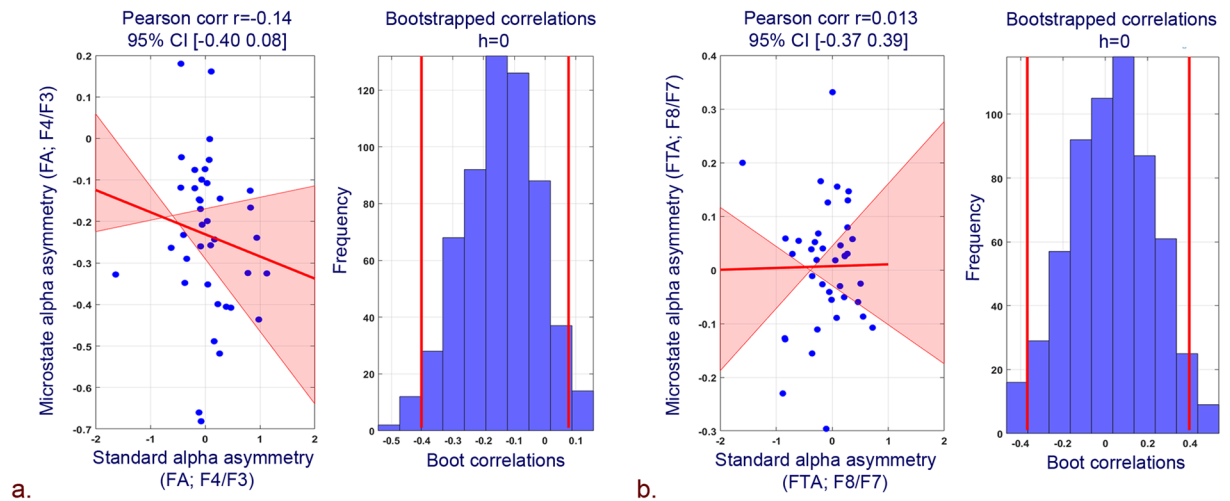
Figure 6. Pearson correlation plots and associated histograms for bootstrapped data for correlation between. (**a**) Standard and microstate based FA (F4/F3). (**b**) Standard and microstate based FTA (F8/F7).

left-lateralized HRF amplitude of temporal underpinning pertaining to microstate based frontal asymmetry is associated with negative affect. Studies in the past have connected left anterior temporal cortical activation as well as temporal lobe per se to the negative affect;[66–69] this proves the efficacy of microstate based frontal asymmetry in explaining the neurovascular mechanism of negative affect which remains absent in the previous literature. Batut et al.[70] signaled the involvement of mesial temporal regions in emotional processes; further, Yun et al.[71] showed that the angst for social communication in social anxiety disorder could be resultant of the imbalanced functional connectivity of left middle temporal gyrus. The association of anterior division of middle temporal gyrus with negative affect is plausible as studies[72–74] have indicated the interaction between middle temporal gyrus and amygdala for better prediction of memory for emotional events. Hence, the middle temporal gyrus may be more tightly functionally coupled with affect specific regions for the memory of negative events. The significant correlation of negative affect with temporal region's HLI, which is independently measured from resting fMRI data for neural underpinnings of microstate frontal asymmetry and its relative left lateralization, also strengthens the finding of positive correlation of negative affect with microstate based frontal asymmetry measures (FA and FTA).

Microstate based asymmetry reveals no correlation with BIS, Positive affect, and BAS. Microstate based asymmetry showed a high but insignificant correlation with BIS measure. It also showed no correlation with positive affect and BAS measure. One possible explanation for these findings is the fact that the positive affect scale is a diverse measure with components of joy, interest, and activation. Each one of these components might involve distinct and sometimes even opposite whole-brain activations[75]. Similarly, BAS is also composed of varied components (reward, drive, and fun)[76]. These varied brain activation patterns might not be producing definite patterns at cortical levels to be picked by the alpha power.

HLI of microstates neural underpinnings reveals significant association with BIS, positive affect, and BAS measures. The hemodynamic lateralization measure of neural underpinnings of the proposed technique revealed a high and positive correlation of BIS in frontal cortical regions. Frontal cortical regions play a very important role in inhibition systems, and it has been one of the cornerstones of neuroscience research[13,77,78]. Further, Fuentes et al.[79] also emphasized the association of individual differences in the behavioral inhibition system with the orbitofrontal cortex. Hence, our results suggest that the HLI, which constitutes the voxel-level hemispheric differences in HRF amplitude of microstates based asymmetry better manifests BIS measure. Further, though microstate based alpha asymmetry found no significant correlation with positive affect and BAS measure, the HLI of occipital fusiform gyrus was found to strongly correlate with BAS measure. This is consonant with the hypothesis where the BAS system is proposed to be modulated by occipital cortices[80]. The nature of correlation was strong but negative and thus was inverse of the BIS system. Moreover, positive affect was correlated positively with hemodynamic lateralization measures in superior frontal gyrus. The link of the prefrontal cortex with positive affect is consistent with recent studies. Wager et al.[81] showed the association of the prefrontal cortex with positive affect as compared to negative affect. Similarly, Roy et al.[82] observed more frequent activity was found in the prefrontal cortex during positive as compared with negative feelings. Hence, hemodynamic lateralization measures of microstates neural underpinnings bring better insight into the positive affect and BAS as compared to the standard EEG based hemispherical asymmetry measures.

Interestingly, the neural underpinnings (middle temporal gyrus (anterior division), inferior frontal gyrus, frontal medial cortex) whose HLI revealed significant correlation (r-value) with negative affect and BIS scores have been observed to be undergoing only alpha-BOLD desynchronization process. They were found to be either correlating positively in the left hemisphere or negatively in the right hemisphere. On the other hand, the neural underpinnings whose HLI correlated with positive affect and BAS scores have

Study	Alpha EEG Asymmetry (R-L)	Mood Measures	Subjects	Main Results
Tomarken et al.[136]	FA (F4/F3);	Acquisition of resting EEG followed by the presentation of affective clips to obtain subjective ratings for emotional reactions	32 females, Cohort A: 17 to 41 years Cohort B: 20 to 54 years	Resting FA significantly predicted self-reported global NA
Tomarken et al.[29]	FA (F4/F3); ATA (T4/T3)	Resting EEG on two occasions; 3 weeks apart; PANAS	90 females, 17–21 years	FA:↓NA ATA:↑PA
Tomarken et al.[30]	Same as in Tokarman et al., 1992a	Same as in Tomarken et al., 1992a	85 females, 17–21 years	Same as in Tomarken et al., 1992a
Jacobs and Snyder,1996[31]	FA (F4/F3); FTA (F8/F7)	Resting EEG on 1-time measurement; PANAS	40 males, 18–53 years	FTA:↓NA
Sutton and Davidson,1997[35]	FA (F4/F3)	Resting EEG on two occasions 6 weeks apart PANAS first session; BIS/BAS scales the second session	46 (23 females) 18–22 years	No correlation between FA and PA, NA, BAS, BIS
Hagemann et al.[37]	FA (F4/F3); ATA (T4/T3)	Acquisition of resting EEG followed by the presentation of affective slides to obtain subjective ratings for emotional reactions	37 (22 females: 15 males: Mean age 24.5)	Subjects with greater relative left-sided anterior temporal cortical activation reported more intense NA in response to negative stimuli
Hagemann et al.[38]	FA (F4/F3); ATA (T4/T3)	Resting EEG; PANAS	36 (24 females) Mean age 24.7	Subjects with high NA exhibited high left cortical activation at the anterior temporal site
Hall and Petruzzello, 1999[32]	FA (F4/F3)	Resting EEG and measures of physical activity; PANAS	41 (26 females) Mean age 68.7	FA positively predicted PA
Harmon-Jones and Allen, 1997[33]	FA (F4/F3);	Resting EEG from females who scored in the upper or lower third of the distribution of social anxiety scores; BAS,BIS	37 females	FA:↑BAS
Hewig et al.[62]	FA (F4/F3); FTA (F8/F7); ATA (T4/T3)	Resting EEG on four occasions; four weeks apart; BAS, BIS	59 (30 females: Mean age 23; 29 males: Mean age 25)	Higher BAS associated with bilateral frontal cortical activity
Shackman et al.[61]	FA (F4/F3); FTA (F8/F7)	Resting EEG on two occasions; several weeks apart; BAS, BIS	51 females Mean age 19.5	A significant relationship between BIS and FA. Higher BIS associated with right posterior DLPFC
De Pascalis et al.[34]	FA (F4/F3); FTA (F8/F7)	Resting EEG; BAS,BIS	51 females Mean age 24.1	FA:↑BAS, Higher BAS associated with left-sided activation in MFG
Schneider et al.[36]	FA (F4/F3);	Two assessments of resting EEG; BAS, BIS	99 (50 females; 49 males aged 10–12 years)	No correlation of BAS, BIS measures with FA

Table 7. List of studies for positive/negative affect and approach/withdrawal dichotomy. EEG, Electroencephelography;↑Positive correlation;↓Negative correlation; FA, Frontal Asymmetry (F4/F3); FTA, Frontal Temporal Asymmetry (F8/F7); ATA, Anterior Temporal Asymmetry (T4/T3); BAS, Behavioral Activation System; BIS, Behavioral Inhibition System; PA, Positive Affect; NA, Negative Affect; DLPFC, Dorsolateral Prefrontal Cortex; MFG, Middle Frontal Gyrus.

revealed both alpha-BOLD synchronization and desynchronization. Particularly, superior frontal gyrus, which correlated with positive affect, underwent both alpha-BOLD synchronization and desynchronization. However, the occipital fusiform gyrus correlated negatively in the left hemisphere, which thus undergoes alpha-BOLD synchronization. Thus, the neural mechanisms involved in negative affect/withdrawal in the resting state exhibited only alpha-BOLD desynchronization. On the contrary, the positive affect and an approach relevant region involved both alpha-BOLD synchronization and desynchronization. However, the underlying innate cause of these mechanisms remains elusive and needs to be explored in the future. Thus, our finding implicates that microstates based frontal alpha asymmetry may provide newer insights into the association of alpha asymmetry with mood and personality measures in both healthy and clinical populations. The plausible explanation is that different cognitive states, including affect and approach/withdrawal behavior, generally involve coordinated activity of many neural assemblies located at the different cortex, and the microstate prototypes could represent these cognitive states.

Absence of correlation among proposed microstate and standard frontal hemispheric asymmetry measures. The proposed microstate based FA and FTA yielded an insignificant low correlation with standard FA and FTA. The proposed microstate based FA and FTA measure the quasi-stable coordinated brain activity and, in the present study, brings better insights into the large scale brain networks of negative affect. Previous works of literature[29,39] have also emphasized the prominence of stability in the standard EEG patterns in bringing forth the linkage among standard frontal alpha asymmetry and affect and approach/withdrawal measures. Hence, the lack of correlation among proposed microstate and standard frontal hemispheric asymmetry measures might be caused by the unstable nature of standard EEG and its frontal alpha asymmetry indices, which

Dynamics of Alpha asymmetry index and BOLD signal for alpha-BOLD desynchronization		
Alpha asymmetry index [Right (alpha)-Left (alpha)]	Dominant BOLD activity trend	Cerebral Lateralization
⬆	⬆ Left Hemisphere	Positively correlating Left hemispheric activation
⬇	⬆ Right Hemisphere	Negatively correlating Right hemispheric activation
Dynamics of Alpha asymmetry index and BOLD signal for alpha-BOLD synchronization		
Alpha asymmetry index [Right (alpha)-Left (alpha)]	Dominant BOLD activity trend	Cerebral Lateralization
⬆	⬆ Right Hemisphere	Positively correlating Right hemispheric activation
⬇	⬆ Left Hemisphere	Negatively correlating Left hemispheric activation

Figure 7. Underlying dynamics associated with alpha asymmetry index and BOLD signal.

is caused by a substantial interference from many other cognitive factors. As this interference is different at different time points across volunteers, the standard EEG and its frontal alpha asymmetry are likely to correlate less with the quasi-stable patterns assessed by the proposed microstate frontal alpha asymmetry indices.

Limitation of the study. The present study utilizes 39 volunteers' data to validate the role of microstate based resting frontal alpha asymmetry in understanding the neural mechanisms of affect and approach/withdrawal behavior. However, affect and approach/withdrawal behavior is known to be elicited by mood induction tasks. Hence, it is necessary to carry out future studies to validate the proposed microstate based frontal alpha asymmetry during such task engagements. Further, the current research involves healthy volunteers from the Indian urban population. Many studies[83–85] in the past have revealed the association of affect and approach/withdrawal behavior with the cultural, ethnic, and social background of the individuals. Thus, it is required to examine the proposed microstate based frontal alpha asymmetry approach in a larger population dataset, which includes individuals from various cultural, ethnic, and social backgrounds.

Also, the topographies of average-referenced, preprocessed standard EEG are known to represent the posterior alpha than frontal alpha, and these topographies have also been studied in comparison with other referencing schemes[54,86]. However, the microstate analysis employed in the current study uses an average referencing scheme for frontal alpha asymmetry estimation. The present study follows average referencing for microstate analysis as various studies[40,87] adequately understand the cognitive phenomena through average-referenced microstate estimations. Further, the effect of different EEG referencing schemes on microstate estimations is still not clearly understood. Extensive, systematic work needs to be undertaken to properly understand the role of varying EEG reference montages based microstate analysis in explaining frontal, posterior, and temporal EEG frequency signatures and topographies.

Conclusion

The above study validates the effectiveness of resting quasi-stable microstate based asymmetry in explaining the neural mechanisms of affect and approach/withdrawal behavior for healthy young male volunteers during 1-time measurement. The novelty of our work emanates from the fact that we estimated the frontal asymmetry of the alpha power from the average GFP amplitude of the quasi-stable microstates topographies, which might reflect the degree of coordination of the neurons underlying alpha-neural underpinnings. Microstate frontal alpha asymmetry correlated positively with negative affect scores, which are defended by the negative correlation of HLI based on microstates' temporal neural underpinning with negative affect. Further, a significant association of HLI based on microstate neural underpinnings with positive affect, BAS and BIS measures concludes that the neural mechanisms of affect and approach/withdrawal dichotomy are better explained by the synchronized global firing of neurons and on-going activity of entire brain networks as assessed by quasi-stable microstates frontal alpha asymmetry. This study also stands unique in exploring the underlying neurovascular synchronization/desynchronization mechanisms of microstate based frontal asymmetry measures. The analysis revealed that neural underpinnings involved both positively and negatively correlating brain regions, thus satisfying alpha-BOLD desynchronization and synchronization criteria. However, specifically the microstates neural underpinnings

whose HLI correlated with negative affect and inhibition involved alpha-BOLD desynchronization, however the positive affect and approach relevant regions involved alpha-BOLD synchronization as well as desynchronization.

Methods

Figure 8 depicts the schema of the methodology adopted in this study.

Sample and procedure.   Thirty-nine healthy participants (all males; age range 18–24 M = 19.57; SD = 1.28) took part in this study after providing a written and informed consent to the protocol. The experiment has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki), and all measurements were also approved by the Institute of Nuclear Medicine and Allied Sciences (INMAS) institutional ethical committee (Number: ECR/824/Inst/DL/2016). All subjects were volunteers recruited among university students and were right-handed. Subjects completed a personality questionnaire for positive affect and negative affect and Behavioral inhibition system (BIS)/Behavioral approach system (BAS). The questionnaires were in the English language, and all the volunteers were fluent in the English language. The resting-state fMRI and EEG data analyzed in this paper were collected after the subject completed the psychological questionnaires. The simultaneous EEG-fMRI resting-state recording lasted for 6 minutes with eyes closed condition.

Behavioral measures.   To assess the dispositional affect and approach/withdrawal parameters in resting state, PANAS scores, and BIS/BAS measures were evaluated for each individual. We also estimated the Profile of mood states using POMS scores for prior exclusion criteria. Table 8 presents descriptive characteristics for the study participants with the mean and standard deviation values.

Positive and negative affect.   Positive and negative affect scores were evaluated for each volunteer. Positive and Negative Affect Schedule (PANAS) consists of mood scales designed to assess affect at the present moment[88]. These scales are highly uncorrelated, stable over time, and consistent, and both scales demonstrate good convergent and discriminant validity[89,90]. Positive and negative affect scores showed good internal consistency in our study (Cronbach's alphas = 0.89; 0.91).

Behavioral approach system (BAS)/behavioral inhibition system (BIS).   BIS and BAS scores were calculated for each subject[91] and evaluation included 24 items (20 score-items and four fillers, each measured on four-point Likert scale), and two total scores for BIS (range = 7–28; 7 items) and BAS (range = 13–52; 13 items). In our study, BIS and BAS scales showed good internal consistency (Cronbach's alphas = 0.93; 0.92).

POMS (Profile of mood states).   Volunteers were also asked to fill in forms for the POMS[92]. It measures six different dimensions of mood swings, namely Tension or Anxiety, Anger or Hostility, Vigor or Activity, Fatigue or Inertia, Depression or Dejection, Confusion, or Bewilderment. These scores formed the basis for exclusion criteria. All selected volunteer returned self-report scores for all the modes within a relatively normal range.

Simultaneous EEG-fMRI data acquisition and preprocessing.   MRI data was acquired in a Siemens 3 T scanner. After acquiring a high-resolution T1-weighted anatomic rapid gradient-echo image (T1 MPRAGE sequence TR 1900ms, TE 2.49 ms, FA 9°, 160 slices with slice thickness 0.9 mm and distance factor of 50%, FoV 240 mm with voxel size $0.9 \times 0.9 \times 0.9$ mm), we acquired 205 T2*-weighted EPI images for resting-state eyes-closed condition (T2* EPI sequence: TR 2000ms, TE 30 ms, FA 90°, 30 slices with thickness 5 mm and distance factor 0%, FoV 240 mm with voxel size $3.8 \times 3.8 \times 5.0$ mm). Continuous EEG data were acquired simultaneously during resting state T2* acquisition using a 32-channel MR-compatible brain vision cap. The electrodes were placed according to the international 10–20 system with a separate electrode called the Reference electrode, placed between Fz and Cz electrodes, that provided the reference for recording the data. Electrocardiogram (ECG) was also recorded. The impedance level for each electrode was kept less than 5 KΩ. The recorded EEG signal was digitized and transmitted with a sampling frequency of 5000 Hz. The acquisition of EEG signals was accomplished using Brain vision analyzer software.

The fMRI data preprocessing for 205 resting-state volumes was done using the default preprocessing pipeline for volume-based analysis in CONN software. The pre-processing procedure included the realignment and unwarping of T2*-weighted image with the mean functional image for motion correction followed by the translation of center to (0, 0, 0) coordinates and slice time correction of functional data. Functional outlier detection (ART- based identification of outlier scans for scrubbing) was performed, followed by segmentation and direct normalization to MNI space. Next, functional smoothening with a Gaussian Kernel with FWHM of 6 mm was carried out. Further, translation of structural center to (0, 0, 0) and simultaneous structural segmentation and normalization were performed.

EEG data were corrected for gradient artifact using the Brain vision analyzer's[93,94] average artifact subtraction algorithm (AAS)[95,96]. A template from MR scanner artifacts was created by averaging the MR scanner artifacts over fixed intervals which were accurately specified by utilizing the fMRI volume markers (labeled as 'TR'). Subsequently, this average was subtracted from the EEG data. Further, the gradient artifact removed data accommodated six seconds of data prior to the start of the first fMRI block acquisition (identified by the first TR marker). These six seconds is the time the fMRI pulse sequence prepares itself before acquiring the first fMRI block. This prior time interval accommodated gradient-contaminated ECG; hence we truncated these 6 seconds prior data and subjected only the data pertaining to the fMRI volumes to the subsequent cardio ballistic (CB) artifact removal. The CB artifact removal was performed in the FMRIB plugin. The method detects the QRS peaks in the ECG data using combined adaptive thresholding[97] and Teager energy operator[98], followed by a correction algorithm. Further, the removal of the CB artifact is performed based on the Optimal Basis Set (OBS) method[99].
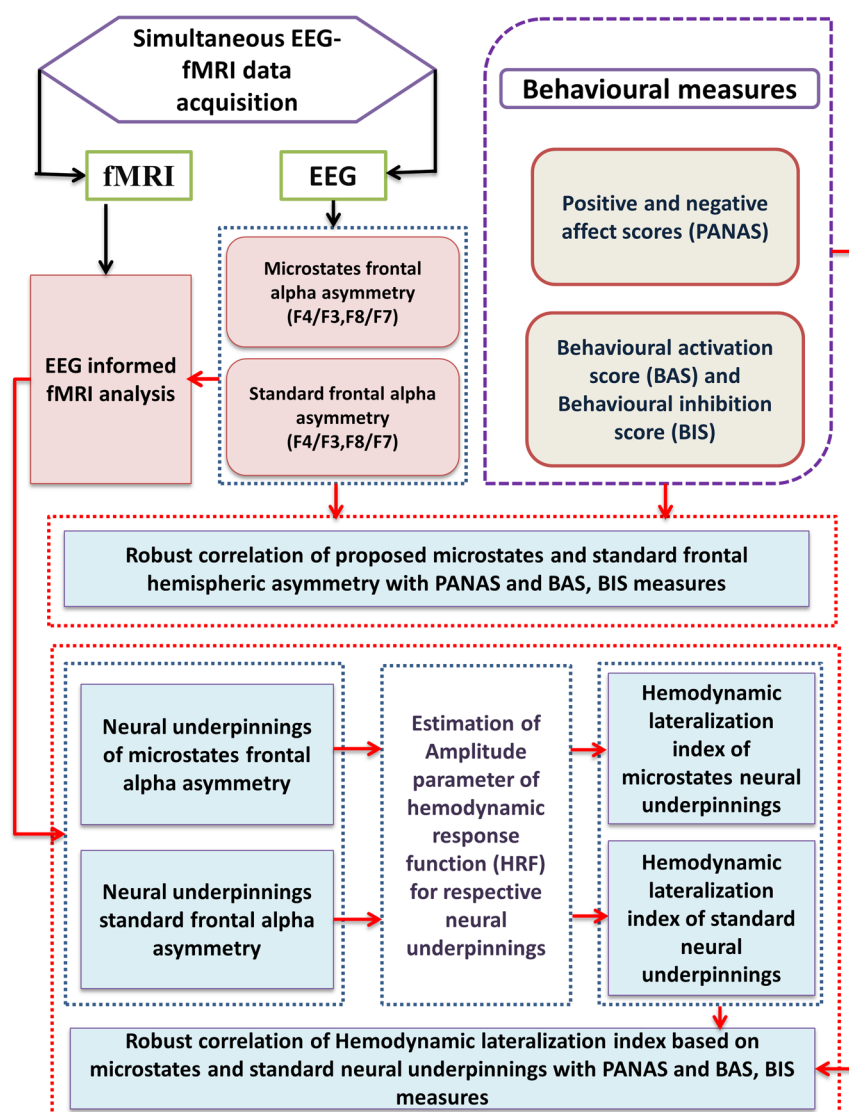
Figure 8. Schema of the methodology adopted in this study.

Variable	Mean (M)	Std. Dev (SD)
Age	19.57	1.28
Positive Affect scores	39.66	5.66
Negative Affect scores	14.64	4.29
BAS scores	23.42	3.5
BIS scores	15.28	2.7

Table 8. Demographic and behavioral characteristics of study participants (N = 39).

In addition, we also employed the HAPPE toolbox[100] for further ensuring the quality of conventional EEG artifact removal from the scanner and CB artifact corrected datasets. The following steps utilizing the HAPPE toolbox were adopted. First, the scanner and CB artifact removed data were subjected to the filtering process with 0.1 Hz high pass and 70 Hz low pass filtering, and all the EEG channels were selected for further analysis. This was followed by removal of the electrical (line) noise using the Cleanline plugin[101] of EEGLAB. The functionality of HAPPE was utilized next to identify and remove the contaminated channels. HAPPE identifies the contaminated channels by evaluating the normed joint probability of average log power across all the channels and rejecting the channels whose joint probability is more than three standard deviations. Wavelet enhanced ICA (W-ICA) approach was implemented subsequently to correct for EEG artifact while retaining the entire length of the data file. The W-ICA approach removes ocular and muscle-related artifacts and also improves the decomposition of later performed ICA, which eventually rejects artifact components. Next, independent components (ICs) with

the extended infomax independent component analysis (ICA) were computed, and the MARA plugin[102,103] of EEGLAB was employed for automatic component rejection. MARA evaluates each component on six features and eventually assigns a probability of artifact contamination to that component. Further, HAPPE's pipeline automatically rejected any components with artifact probabilities higher than 0.5. Subsequently, segmentation of data based on the markers, rejection of segments, and interpolation of removed channels were carried out. Finally, the processing report about the quality of data was generated. The EEG preprocessing procedures in this study have been explained in detail in Supplementary methods and discussion section. Further the processing report about the quality of data for all volunteers has been tabulated in Supplementary Table 1.

To ensure the quality of preprocessing, we also subjected both raw and final artifact removed EEG data (CSD referenced) to the estimation of the power spectrum between 0.2 Hz to 50 Hz frequency range. The median power spectrum plots of both raw and final artifact removed EEG data (CSD referenced) for channels F3, F4, F7, F8, Pz, Oz, and POz are shown in Supplementary Figs. 1 and 2 respectively. The median spectral power of artifact removed EEG data clearly reveals parietal and occipital alpha and beta bands. Data was down-sampled to 250 Hz for further analysis.

Assessment of frontal hemispherical asymmetry measures. The main objective of the study was to understand the neural mechanisms associated with the affect, approach/withdrawal behavior, as explained by the hemispherical asymmetry measures. For this purpose, the present study proposes an EEG microstate based frontal hemispheric assessment approach and aims to compare its advantage over the standard EEG frontal asymmetry approach. The following subsections explain the methods for estimating the proposed EEG microstate based frontal hemispheric asymmetry as well as the standard frontal EEG asymmetry.

EEG microstates based estimation of hemispheric asymmetry. Many recent studies[40,104,105] have clearly indicated that individual brain functions involve massive parallel processing in distributed brain networks. These distributed brain networks are observed as the scalp field potential in EEG, and the state of global neural activity is measured as a topographical map at that moment of time. The changes in this topography reflect changes in the global coordination of neural activity over time. EEG microstates were proposed to represent changes in behavior, thoughts, and emotions and can be classified into few topographies, which have explained 90% of the variance of continuous EEG. Microstate analysis considers millisecond time range signal from all electrodes to create a global picture of a functional state during that interval.

The schema of the methodology adopted for microstate estimation is explained in Fig. 9. The aim of a microstate analysis is first to segment EEG maps into microstate prototypes and second to re-express the spatial-temporal characteristics of the time series of EEG through these microstate prototypes.

In this study, let X be the time series EEG information that was acquired from the volunteers. At first, the EEG data X has been pre-processed for removing the artifacts and was referenced to the average referencing. Then, it was subjected to the estimation of Global field power (GFP). GFP is the measure of global brain response to an event and is represented as:

$$GFP = \sqrt{\left(\sum_{i=1}^{C}(X_i(t) - X_{mean}(t))^2\right)/C}$$

(1)

where X_i is the measured potential at the ith electrode at a given time-point t, X_{mean} is the mean value of all X_i's and C represents the total number of channels. GFP, therefore, represents the standard deviation of the electrode values and indicates, on average, how strong potential is being recorded across the electrode montage[106]. For each volunteer, a selection of data points for the further processing has been carried out by filtering estimated GFPs based on minimum peak distance of 20 milliseconds, and the threshold amplitude of one standard deviation of estimated GFP. Then, the filtered EEG data points of every individual are concatenated to form the GFP datasets for further clustering process as follows:

$$\chi = \{x_{GFP}^1, x_{GFP}^2, \ldots\ldots\ldots x_{GFP}^S\}$$

(2)

where χ is of the concatenated GFP dataset and x_{GFP}^i are selected data points based on the GFP criteria of the ith volunteer, and S is the total number of volunteers. In this study, thirty-nine volunteers dataset has been subjected to analysis.

Further, concatenated GFP dataset χ was subjected to the clustering process through the modified K-means clustering algorithm[107]. Modified K-means clustering algorithm requires the initialization of both number (K) of microstate prototype vectors and their components values[108]. Thus, the clustering algorithm was randomly initialized with a set of microstate prototype vectors as the center of initial clusters as follows:

$$Z = \{z_i | i = 1 \; to \; K\}$$

(3)

where K is the total number of microstate prototype vectors (cluster center). In this study, the K is initialized with 8. The clustering algorithm was allowed to iterate and minimize the orthogonal euclidean distance between the data points in χ as given below.

$$\tau_n = arg\min_k\{D_{kn}^2\}$$

(4)
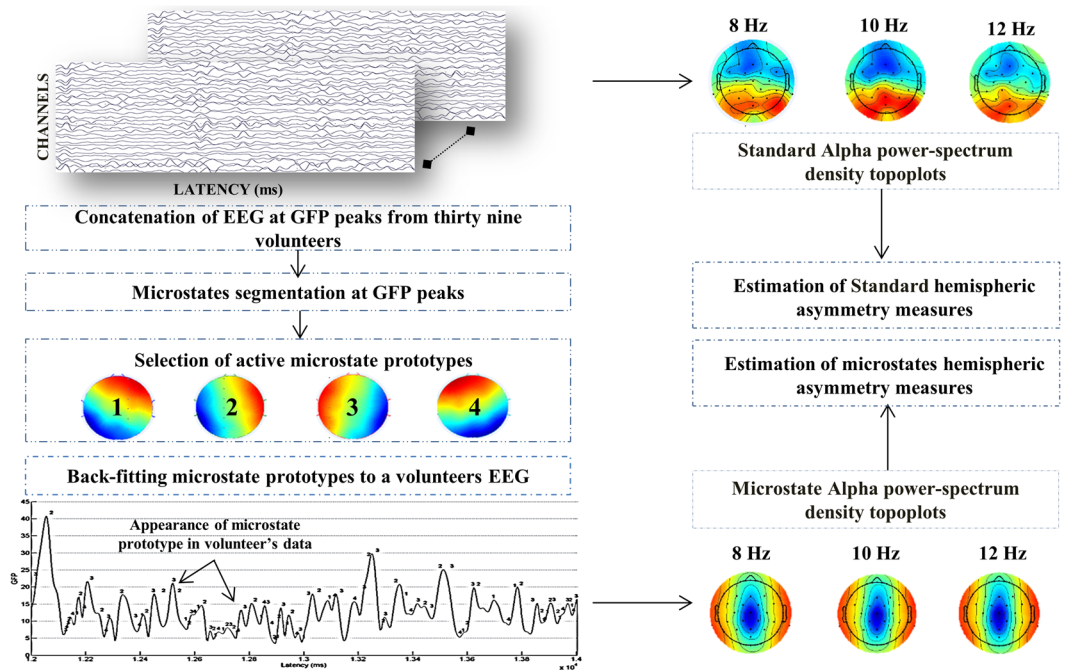
Figure 9. Schema of the methodology adopted for proposed microstate estimation and assessment of standard and microstates based frontal alpha hemispheric asymmetry measures.

$$D_{kn}^2 = \chi_n^T \cdot \chi_n - (\chi_n^T \cdot z_k)^2 \tag{5}$$

where τ_n represents the microstate label for nth sample, χ_n represents the nth time sample of the concatenated dataset, z_k represents the prototypical map for the kth microstate cluster and D_{kn} represents the distance between χ_n and microstate k for the nth time sample. Thus, this clustering algorithm allocates each EEG sample to the cluster whose prototype it is most similar to and then re-estimates microstate prototypes by averaging newly assigned samples[107]. The maximum number of iterations was set to 1000, and the threshold for convergence was set at 1e^{-6} for analysis in this study.

Subsequently, a review of goodness of fit and selection of active microstates is carried out based on global explained variance (GEV) and cross-validation (CV) criterion. It basically evaluates how well microstate segmentation explains the EEG data, which has been used to estimate the prototypes. Therefore, GEV measures how similar the EEG sample and the microstate prototype are; and is calculated as follows.

$$GEV_n = \frac{(Corr(\chi_n, z_{\tau_n}) \cdot x_{GFP_n})^2}{\sum_{n'}^N x_{GFP_{n'}}^2} \tag{6}$$

where χ_n represents the nth time sample of the concatenated dataset, z_{τ_n} ($\tau_n = k$) is the prototypical map for the kth microstate cluster and x_{GFP_n} represents the nth time sample of the GFP data, and N represents the total number of time samples in concatenated dataset χ. GEV is thus the correlation between the EEG dataset and associated microstate prototype weighted by the EEG dataset's fraction of the total squared GFP[107]. Thereafter to calculate the GEV for a given cluster, the GEV of its members is summed. Subsequently, CV which is a measure related to the residual noise $\in$ is estimated as,

$$CV = \sigma^2 \cdot \left(\frac{C-1}{C-K-1} \right) \tag{7}$$

$$\sigma^2 = \frac{\sum_n^N \chi_n^T \cdot \chi_n - (\chi_n^T \cdot z_k)^2}{N(C-1)} \tag{8}$$

where σ^2 is the variance of the residual noise, C is the number of EEG channels, N represents the total number of time samples in concatenated dataset χ, and K is the number of clusters. The aim is to obtain a low value of CV. The active microstate prototypes obtained in this study are consistent with the normative EEG microstate classes identified by many studies[40,41,87,109–111].

Following the selection of an active number of microstate prototypes, the EEG of each volunteer is re-expressed as a sequence of microstate classes by back-fitting these active microstate prototypes on each volunteer's EEG

data. Back fitting implies assigning microstate labels to the EEG dataset based on the dataset's topographic similarity with the microstate prototype. The estimated re-expressed back fitted dataset is represented as follows

$$X_{re-expressed} = \{\mu_n | where \ \mu_n \ \in \ Z_{k'}\} \tag{9}$$

$$where \ \mu_n \ = \ arg \ \min(GMD_n)$$

The global map dissimilarity (GMD) index measures the topographical similarity between each microstate prototype vector with the EEG sample vector. The GMD is calculated as,

$$GMD_n = \frac{\left\| \frac{X_n}{X_{GFP_n}} - \frac{z_{k'}}{z_{GFP_{k'}}} \right\|}{\sqrt{C}} \tag{10}$$

where X_n represents the nth time sample of the preprocessed dataset, $z_{k'}$ represents the prototypical map for the kth microstate cluster. In an ideal condition, if the microstate prototype vector and the EEG sample vector of interest are having the same topographic distribution, then the GMD index will be zero. In case, if both the vectors are topographically opposite, then GMD index would be positively higher. Hence, in this study, instead of the thresholding the GMD index, the microstate prototype vector, which yields a very less GMD index, is chosen as the label for that particular EEG sample vector. Finally, microstates statistics using labels obtained from back-fitted prototypes were calculated.

Subsequently, the amplitude of the microstate prototype vector associated with each label in microstate re-expressed EEG data of every individual is subjected to the alpha band power (8–12 Hz) estimation. The estimated alpha power map of the microstate re-expressed EEG data was used to estimate EEG microstate based frontal hemispheric asymmetry as follows:

$$Asymmetry_{MS} = \ln(\alpha(X_{re-expressed}{}^{Right}) - \ln(\alpha(X_{re-expressed}{}^{Left}) \tag{11}$$

$\alpha(X_{re-expressed}{}^{Right})$ and $\alpha(X_{re-expressed}{}^{Left})$ are the alpha powers measured at the right and left hemispheric channel of microstate re-expressed EEG data, respectively.

Standard EEG estimation of hemispheric asymmetry. In order to estimate standard frontal asymmetry, the preprocessed EEG data is first re-referenced to CSD reference using the CSD toolbox[112,113]. Recent work suggests that the CSD transformation reduces the influence of non-frontal sources to frontal asymmetry and may provide a better index of individual differences in frontal asymmetry[114]. Subsequently, the power spectral density (PSD) of alpha frequency (8–12 Hz) was extracted. The estimated alpha power map EEG data was used to calculate standard EEG frontal hemispheric asymmetry as follows:

$$Asymmetry_{Standard} = \ln(\alpha(X)^{Right}) - \ln(\alpha(X)^{Left}) \tag{12}$$

$\alpha(X)^{Right}$ and $\alpha(X)^{Left}$ are the standard alpha powers measured at the right and left hemispheric channels of individual EEG data, respectively.

Table 9 presents the median and median absolute deviation values for EEG asymmetries for mid-frontal and lateral-frontal sites.

Robust correlation of frontal hemispherical asymmetry measures with psychological measures. Further, estimated EEG microstate and standard frontal hemispherical asymmetries are correlated with PANAS and BAS, BIS measures. These robust correlations were carried out for hemispherical measures that are estimated for both channel pairs F4/F3 i.e. Frontal Asymmetry (FA) and F8/F7 i.e. Frontal Temporal Asymmetry (FTA) independently. The rationale for choosing these channels was based on the linkage of hemispheric asymmetry to mid-frontal (F3, F4) and lateral frontal (F7, F8) sites[39,60,115]. Robust correlations were implemented in Robust correlation Matlab toolbox[116]. This method detects and protects against any bivariate or univariate outliers. Pearson, Bend, and Spearman correlation coefficients, as well as bootstrapped confidence intervals, were computed to evaluate each correlation. Both p-values and confidence intervals were Bonferroni corrected for multiple comparisons.

Assessment of neural mechanisms associated with functional hemispheric asymmetry measures. One of the focuses of the present study is to understand the neural mechanisms associated with proposed and standard functional hemispheric asymmetry measures in explaining the affect and approach/ withdrawal behavior during resting state. For this purpose, both proposed and standard hemispheric asymmetry measures were subjected to the EEG informed fMRI, and their neural underpinnings were estimated. Subsequently, the lateralization index based on differences in the amplitude of hemodynamic response of neural underpinnings of both hemispheric asymmetry measures was assessed. Finally, the estimated lateralization index was correlated with PANAS and BAS, BIS psychological measures to understand the ability of both hemispheric asymmetry measures in explaining affect and approach/ withdrawal behavior during resting state. The following sub-sections explain these operations in detail.

EEG informed fMRI analysis. Estimation of neural underpinnings of proposed microstate based EEG asymmetry and standard asymmetry was carried out as follows. At first, the estimated alpha powers for frontal

Variable	Channel pair F4/F3 (FA)		Channel pair F8/F7 (FTA)	
	Median	Median Absolute Deviation	Median	Median Absolute Deviation
Standard hemispheric asymmetry[a]	0.0347	0.3509	−0.052	0.3655
Microstates based hemispheric asymmetry[a]	−0.2324	0.1427	0.0256	0.0896

Table 9. Median and median absolute deviation of the standard and proposed microstates based frontal hemispheric asymmetry measures. [a]The difference between log-transformed alpha values from one right-hemispheric electrode to the corresponding electrodes on the left.

channels F3, F4 F7, and F8 were down sampled to match the acquisition blocks of fMRI (TR: 2 seconds). This was carried out by taking the median of the alpha powers for these specific channels corresponding to each fMRI scan time, which is 2 seconds. The onset time of EEG and fMRI acquisition were also matched. This yielded one EEG alpha power corresponding to each fMRI scan, respectively. Thereafter, microstate based and standard FA and FTA were estimated. The first-level analysis in the present study was performed in SPM12. Different design matrices were obtained each for microstate based and standard asymmetry respectively for each subject wherein microstate based and standard FA and FTA parametrically modulated the fMRI regressors in EEG informed fMRI analysis[117–121].

The first-level analysis in our study was performed in SPM12, and the time series of fMRI regressors and parametric modulators were convolved with canonical HRF and with its time and dispersion derivatives. Further, at first-level, an F-contrast was defined for parametric modulators subsuming both non-derivative (canonical HRF) and derivative terms (time and dispersion derivatives) for microstate based FA, standard FA, microstate based FTA and standard FTA models.

Subsequently, for the second level of analysis, the first-level contrast images, along with the dispersion and temporal derivatives, were subjected to extraction of amplitude measures from the basis sets[122–126]. The robust regression toolbox[127] was used to conduct group-level random-effects analysis. The robust regression toolbox uses iteratively re-weighted least squares (IRLS), which detects influential extreme outliers. Thus, the IRLS analysis reduces the likelihood of false-positive and negative findings with no reduction in power and minimizes the effect of extreme outliers[128]. The IRLS has proved beneficial with small samples (n = 10), and the benefits tend to increase with larger sample sizes (n = 40). Further, IRLS controls false-positive rates at an appropriate level when no true effects are present. The contrast image for amplitude summary measure was then subjected for the whole brain analysis corrected with voxel-wise False Discovery Rate (FDR) thresholded at q < 0.05. This yielded the underpinning of both microstate based FA and FTA and standard FA and FTA.

Estimation of Hemodynamic lateralization index and its robust correlation with psychological measures. The lateralization index measures the hemispherical dominance within the large scale brain network that integrates the neural underpinnings associated with resting affect and approach/withdrawal behavior. The neural activity associated with the neural underpinnings of each hemisphere causes differential electrical potential on the cortical surface of the respective hemisphere. This is measured as the EEG asymmetry index, as explained in the earlier sections. In the mean-time, these differential neural activities of each hemisphere generate a feed-forward signal, which results in differential hemodynamic response at the location of neural activity. Measurement of these hemodynamic hemispherical differences facilitates a better understanding of hemispherical dominance within the large scale brain interactions. Diverse methods have been proposed to calculate the hemodynamic lateralization index on the basis of fMRI BOLD information. As most of these studies involved task engagement, the hemispherical difference of cluster size and BOLD signal strength[129–133] were normally used to estimate the HLI.

The main motivation behind this estimation is to understand whether hemodynamic asymmetry reveals more insight into understanding the neurovascular mechanisms of the affect and approach /withdrawal behavior. For this purpose, initially, we estimated the hemodynamic response function metric that is hemodynamic response function amplitude (HRF_Amp) at every voxel by independently subjecting the preprocessed resting fMRI data to blind deconvolution method as proposed by Wu et al.[134,135]. The estimation of HRF was carried out independently by assuming acquired fMRI BOLD signal $y(t)$ as the convolution of neural states $n(t)$ with $HRF(t)$. This is represented as,

$$y(t) = conv(n(t), HRF(t)) + \in (t) \tag{13}$$

where $\in(t)$ is the noise in the measurement. Further, $n(t)$ is substituted by a hypothetical neural activation model:

$$\hat{n}(t) = \sum_{\tau=0}^{\infty} \delta(t - \tau) \tag{14}$$

where $\delta(t - \tau)$ is the delta function. This allows fitting $HRF(t)$ according to $\hat{n}(t)$ using a canonical HRF and two derivatives (temporal and dispersion derivatives). This model is subjected to blind deconvolution approach for retrieving the hemodynamic response function ($HRF(t)$) of every voxel. Once $HRF(t)$ is obtained, an approximation of $\tilde{n}(t)$ can be calculated using the inverse Fourier transform (deconvolution). Then, $HRF(t)$ was utilized to estimate the HLI for the neural underpinnings of both microstate based FA and FTA and standard FA and FTA,

all considered together. Hence, the cluster results of EEG informed fMRI were used only for the selection of regions for estimating HLI as follows,

$$HLI(n) = HRF\_Amp_{p_n}^{R} - HRF\_Amp_{p_n}^{L} \qquad (15)$$

where $HRF\_Amp_{p_n}^{R}$ and $HRF\_Amp_{p_n}^{L}$ are the median amplitude of hemodynamic response function of the nth neural underpinnings in the right and left hemispheres, respectively. The median of estimated HLI of neural underpinnings of proposed microstate based EEG asymmetry and the standard EEG asymmetry measures were finally subjected to the robust correlations with PANAS and BIS/BAS measures.

Data availability

The data for this study is available from the corresponding author on a reasonable request.

References

1. Killgore, W. D. S. & Yurgelun-Todd, D. A. The right-hemisphere and valence hypotheses: Could they both be right (and sometimes left)? *Soc. Cogn. Affect. Neurosci.* **2**, 240–250 (2007).
2. Nijboer, T. C. W. & Jellema, T. Unequal impairment in the recognition of positive and negative emotions after right hemisphere lesions: A left hemisphere bias for happy faces. *J. Neuropsychol.* **6**, 79–93 (2012).
3. Aben, H. P. *et al.* Impaired Emotion Recognition after Left Hemispheric. *Stroke: A Case Report and Brief Review of the Literature. Case Rep. Neurol. Med.* **2017**, 1–6 (2017).
4. Palmiero, M. & Piccardi, L. Frontal EEG Asymmetry of Mood: A Mini-Review. *Front. Behav. Neurosci.* **11**, 1–8 (2017).
5. Allen, J. J. B., Keune, P. M., Sch, M. & Nusslock, R. Frontal EEG alpha asymmetry and emotion: From neural underpinnings and methodological considerations to psychopathology and social cognition. 1–6, https://doi.org/10.1111/psyp.13028 (2018).
6. Hewig, J. Intentionality in frontal asymmetry research. *Psychophysiology* **55**, 1–18 (2018).
7. Davidson, R. J., Saron, C. D., Senulis, J. A., Ekman, P. & Friesen, W. V. Approach/withdrawal and cerebral asymmetry: Emotional\ rexpression and brain physiology: I. *J. Pers. Soc. Psychol.* **58**, 330–341 (1990).
8. Davidson, R. J., Kalin, N. H. & Shelton, S. E. Lateralized response to diazepam predicts temperamental style in rhesus monkeys. *Behav. Neurosci.* **107**, 1106–1110 (1993).
9. Davidson, R. J. What does the prefrontal cortex 'do' in affect: Perspectives on frontal EEG asymmetry research. *Biol. Psychol.* **67**, 219–234 (2004).
10. Carver, C. S. & Harmon-Jones, E. Anger Is an Approach-Related Affect: Evidence and Implications. *Psychol. Bull.* **135**, 183–204 (2009).
11. Jacob, T. J. C. *et al.* Light and smell stimulus protocol reduced negative frontal EEG asymmetry and improved mood. *Open Life Sci.* **12**, 51–61 (2017).
12. Rey, B., Clemente, M., Wrzesien, M., Rodríguez, A. & Alcañiz, M. Assessing brain activations associated with emotional regulation during virtual reality mood induction procedures. *Expert Syst. Appl.* **42**, 1699–1709 (2014).
13. Andreas, M. & Hewig, J. Mind the movement: Frontal asymmetry stands for behavioral motivation, bilateral frontal activation for behavior. 1–19, https://doi.org/10.1111/psyp.12908 (2017).
14. Balconi, M., Vanutelli, M. E. & Grippa, E. Resting state and personality component (BIS/BAS) predict the brain activity (EEG and fNIRS measure) in response to emotional cues. *Brain Behav.* **7**, 1–15 (2017).
15. Studer, B., Pedroni, A. & Rieskamp, J. Predicting Risk-Taking Behavior from Prefrontal Resting-State Activity and Personality. *Plos One* **8**, 1–8 (2013).
16. Eftekhari, E., Tran, A. & McGregor, I. Decentering increases approach motivation among distressed individuals. *Pers. Individ. Dif.* **119**, 236–241 (2017).
17. Gollan, J. K. *et al.* Frontal alpha EEG asymmetry before and after behavioral activation treatment for depression. *Biol. Psychol.* **99**, 198–208 (2014).
18. Fedorowicz, G. The Effects of Mindfulness Meditation on Mental Health. (2012).
19. Kelley, N. J. & Schmeichel, B. J. The effects of negative emotions on sensory perception: fear but not anger decreases tactile sensitivity. *Front. Psychol.* **5**, 1–8 (2014).
20. Brzezicka, A., Kamiński, J., Kamińska, O. K., Wołyńczyk-Gmaj, D. & Sedek, G. Frontal EEG alpha band asymmetry as a predictor of reasoning deficiency in depressed people. *Cogn. Emot.* **31**, 868–878 (2017).
21. Greimel, E. *et al.* Resting frontal EEG asymmetry patterns in adolescents with and without major depression. *Biol. Psychol.* **132**, 212–216 (2018).
22. Quaedflieg, C. W. E. M. *et al.* The role of frontal EEG asymmetry in post-traumatic stress disorder. *Biol. Psychol.* **108**, 62–77 (2015).
23. Frenkel, T. I. *et al.* ADHD Symptoms in Post-Institutionalized Children Are Partially Mediated by Altered Frontal EEG Asymmetry. *J. Abnorm. Child Psychol.* **45**, 857–869 (2017).
24. Stewart, J. L. & Allen, J. J. B. Resting frontal brain asymmetry is linked to future depressive symptoms in women. *Biol. Psychol.* **136**, 161–167 (2018).
25. van der Vinne, N., Vollebregt, M. A., van Putten, M. J. A. M. & Arns, M. Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? A meta-analysis. *NeuroImage Clin.* **16**, 79–87 (2017).
26. Lachman, M. E. *et al.* Frontal brain asymmetry, childhood maltreatment, and low-grade inflammation at midlife. *Psychoneuroendocrinology* **75**, 152–163 (2016).
27. Flasbeck, V., Popkirov, S. & Brüne, M. Frontal EEG asymmetry in borderline personality disorder is associated with alexithymia. *Borderline Personal. Disord. Emot. Dysregulation* **4**, 4–9 (2017).
28. Adolph, D. & Margraf, J. The differential relationship between trait anxiety, depression, and resting frontal α-asymmetry. *J. Neural Transm.* **124**, 379–386 (2017).
29. Tomarken, A. J., Davidson, R. J., Wheeler, R. E. & Doss, R. C. Individual-Differences in Anterior Brain Asymmetry and Fundamental Dimensions of Emotion. *J. Pers. Soc. Psychol.* **62**, 676–687 (1992).
30. Tomarken, A. J., Davidson, R. J., Wheeler, R. E. & Kinney, L. Psychometric Properties of Resting Anterior EEG Asymmetry: Temporal Stability and Internal Consistency. *Psychophysiology* **29**, 576–592 (1992).
31. Jacobs, G. D. & Snyder, D. Frontal brain asymmetry predicts affective style in men. *Behav. Neurosci.* **110**, 3–6 (1996).
32. Hall, E. E. & Petruzello, S. J. Frontal Asymmetry, Dispositional Affect and Physical Activity in Older Adults. *J. Aging Phys. Act.* **7**, 76–90 (1999).
33. Harmon-Jones, E. & Allen, J. J. B. Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *J. Abnorm. Psychol.* **106**, 159–163 (1997).

34. De Pascalis, V., Cozzuto, G., Caprara, G. V. & Alessandri, G. Relations among EEG-alpha asymmetry, BIS/BAS, and dispositional optimism. *Biol. Psychol.* **94**, 198–209 (2013).
35. Sutton, S. K. & Davidson, R. J. Prefrontal brain asymmetry: A Biological Substrate of the Behavioral Approach and Inhibition Systems. *Psychol. Sci.* **8**, 204–210 (1997).
36. Schneider, M. *et al.* EEG asymmetry and BIS/BAS among healthy adolescents. *Biol. Psychol.* **120**, 142–148 (2016).
37. Hagemann, D., Naumann, E., Becker, G., Maier, S. & Bartussek, D. Frontal brain asymmetry and affective style: A conceptual replication. *Psychophysiology* **35**, 372–388 (1998).
38. Hagemann, D. *et al.* EEG asymmetry, dispositional mood and personality. *Pers. Individ. Dif.* **27**, 541–568 (1999).
39. Wheeler, R. E., Davidson, R. J. & Tomarken, A. J. Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. *Psychophysiology* **30**, 82–89 (1993).
40. Michel, C. M. & Koenig, T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *Neuroimage* **180**, 577–593 (2018).
41. Khanna, A. Microstates in Resting-State EEG: Current Status and Future Directions. 105–113, https://doi.org/10.1016/j.neubiorev.2014.12.010.Microstates (2016).
42. Lehmann, D., Strik, W. K., Henggeler, B., Koenig, T. & Koukkou, M. Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *Int. J. Psychophysiol.* **29**, 1–11 (1998).
43. Milz, P., Pascual-Marqui, R. D., Achermann, P., Kochi, K. & Faber, P. L. The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. *Neuroimage* **162**, 353–361 (2017).
44. Shafi, M. M. *et al.* EEG Microstate Correlates of Fluid Intelligence and Response to Cognitive Training. *Brain Topogr.* **30**, 502–520 (2017).
45. Wright, D., Makin, A. D. J. & Bertamini, M. Right-lateralized alpha desynchronization during regularity discrimination: Hemispheric specialization or directed spatial attention? *Psychophysiology* **52**, 638–647 (2015).
46. Lenartowicz, A. *et al.* Alpha desynchronization and fronto-parietal connectivity during spatial working memory encoding deficits in ADHD: A simultaneous EEG-fMRI study. *NeuroImage Clin.* **11**, 210–223 (2016).
47. Fink, A., Grabner, R. H., Neuper, C. & Neubauer, A. C. EEG alpha band dissociation with increasing task demands. *Cogn. Brain Res.* **24**, 252–259 (2005).
48. Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P. & Winkler, T. 'Paradoxical' alpha synchronization in a memory task. *Cogn. Brain Res.* **7**, 493–501 (1999).
49. Arakaki, X. *et al.* Alpha desynchronization/synchronization during working memory testing is compromised in acute mild traumatic brain injury (mTBI). *Plos One* **13**, 1–19 (2018).
50. Palva, S. & Palva, J. M. Functional roles of alpha-band phase synchronization in local and large-scale cortical networks. *Front. Psychol.* **2**, 1–15 (2011).
51. Benedek, M., Bergner, S., Könen, T., Fink, A. & Neubauer, A. C. EEG alpha synchronization is related to top-down processing in convergent and divergent thinking. *Neuropsychologia* **49**, 3505–3511 (2011).
52. Rohr, C. S., Okon-Singer, H., Craddock, R. C., Villringer, A. & Margulies, D. S. Affect and the Brain's Functional Organization: A Resting-State Connectivity Approach. *Plos One* **8** (2013).
53. Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J. & Barrett, L. F. The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. *Cereb. Cortex* **26**, 1910–1922 (2016).
54. Tenke, C. E., Kayser, J., Abraham, K., Alvarenga, J. E. & Bruder, G. E. Posterior EEG alpha at rest and during task performance: Comparison of current source density and field potential measures. *Int. J. Psychophysiol.* **97**, 299–309 (2015).
55. Rashed-Al-Mahfuz, M., Islam, M. R., Hirose, K. & Molla, M. K. I. Artifact suppression and analysis of brain activities with electroencephalography signals. *Neural Regen. Res.* **8**, 1500–1513 (2013).
56. Davidson, R. J., Mednick, D., Moss, E., Saron, C. & Schaffer, C. E. Ratings of emotion in faces are influenced by the visual field to which stimuli are presented. *Brain Cogn.* **6**, 403–411 (1987).
57. Davidson, J. Anterior cerebral asymmetry and the nature of emotion. **151**, 125–151 (1992).
58. Baijal, S. & Srinivasan, N. Emotional and hemispheric asymmetries in shifts of attention: An ERP study. *Cogn. Emot.* **25**, 280–294 (2011).
59. Wyczesany, M., Capotosto, P., Zappasodi, F. & Prete, G. Hemispheric asymmetries and emotions: Evidence from effective connectivity. *Neuropsychologia* **121**, 98–105 (2018).
60. Harmon-Jones, E., Gable, P. A. & Peterson, C. K. The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biol. Psychol.* **84**, 451–462 (2010).
61. Shackman, A. J., McMenamin, B. W., Maxwell, J. S., Greischar, L. L. & Davidson, R. J. Right dorsolateral prefrontal cortical activity and behavioral inhibition. *Psychol. Sci.* **20**, 1500–1506 (2009).
62. Hewig, J., Hagemann, D., Seifert, J., Naumann, E. & Bartussek, D. The relation of cortical activity and BIS/BAS on the trait level. *Biol. Psychol.* **71**, 42–53 (2006).
63. Quaedflieg, C. W. E. M., Meyer, T., Smulders, F. T. Y. & Smeets, T. The functional role of individual-alpha based frontal asymmetry in stress responding. *Biol. Psychol.* **104**, 75–81 (2015).
64. Gale, A., Edwards, J., Morris, P., Moore, R. & Forrester, D. Extraversion-introversion, neuroticism-stability, and EEG indicators of positive and negative empathic mood. *Pers. Individ. Dif.* **30**, 449–461 (2001).
65. Farahi, S. M. M., Ebrahimabad, M. J. A., Gorji, A., Bigdeli, I. & Farahi, S. M. M. M. Neuroticism and frontal EEG asymmetry correlated with dynamic facial emotional processing in adolescents. *Front.* **10**, 1–9 (2019).
66. Meador, K. J., Kapur, R., Loring, D. W., Kanner, A. M. & Morrell, M. J. Quality of life and mood in patients with medically intractable epilepsy treated with targeted responsive neurostimulation. *Epilepsy Behav.* **45**, 242–247 (2015).
67. Hennion, S. *et al.* Experiences of self-conscious emotions in temporal lobe epilepsy. *Epilepsy Behav.* **90**, 1–6 (2019).
68. Ritchey, M., Wang, S. F., Yonelinas, A. P. & Ranganath, C. Dissociable medial temporal pathways for encoding emotional item and context information. *Neuropsychologia* **124**, 66–78 (2019).
69. Ives-Deliperi, V. L. & Jokeit, H. Impaired Social Cognition in Epilepsy: A Review of What We Have Learnt From Neuroimaging Studies. *Front. Neurol.* **10**, (2019).
70. Batut, A. C. *et al.* Neural responses associated with positive and negative emotion processing in patients with left versus right temporal lobe epilepsy. *Epilepsy Behav.* **9**, 415–423 (2006).
71. Yun, J. Y. *et al.* The left middle temporal gyrus in the middle of an impaired social-affective communication network in social anxiety disorder. *J. Affect. Disord.* **214**, 53–59 (2017).
72. Dolcos, F., Iordan, A. D. & Dolcos, S. Neural correlates of emotion Á cognition interactions: A review of evidence from brain imaging investigations. **23**, 669–694 (2011).
73. Dolcos, F., Labar, K. S. & Cabeza, R. Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proc. Natl. Acad. Sci. USA* **102**, 2626–2631 (2005).
74. Buchanan, T. W., Tranel, D. & Adolphs, R. Memories for emotional autobiographical events following unilateral damage to medial temporal lobe. *Brain* **129**, 115–127 (2006).
75. Egloff, B., Schmukle, S. C., Burns, L. R., Kohlmann, C. W. & Hock, M. Facets of Dynamic Positive Affect: Differentiating Joy, Interest, and Activation in the Positive and Negative Affect Schedule (PANAS). *J. Pers. Soc. Psychol.* **85**, 528–540 (2003).

76. Taubitz, L. E., Pedersen, W. S. & Larson, C. L. BAS Reward Responsiveness: A unique predictor of positive psychological functioning. *Pers. Individ. Dif.* **80**, 107–112 (2015).

77. Coan, J. A. & Allen, J. J. B. Frontal EEG asymmetry as a moderator and mediator of emotion. *Biol. Psychol.* **67**, 7–49 (2004).

78. Harmon-Jones, E. & Gable, P. A. On the role of asymmetric frontal cortical activity in approach and withdrawal motivation: An updated review of the evidence. *Psychophysiology* **55** (2018).

79. Fuentes, P. *et al.* Individual differences in the Behavioral Inhibition System are associated with orbitofrontal cortex and precuneus gray matter volume. *Cogn. Affect. Behav. Neurosci.* **12**, 491–498 (2012).

80. Barrós-Loscertales, A. *et al.* Behavioral activation system modulation on brain activation during appetitive and aversive stimulus processing. *Soc. Cogn. Affect. Neurosci.* **5**, 18–28 (2010).

81. Wager, T. D., Barrett, L. F., Weber, J., Lindquist, K. A. & Satpute, A. B. The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. *Cereb. Cortex* **26**, 1910–1922 (2015).

82. Roy, M., Shohamy, D. & Wager, T. D. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn. Sci.* **16**, 147–156 (2012).

83. Thayer, J. F. & Koenig, J. Resting Cerebral Blood Flow and Ethnic Differences in Heart Rate Variability: Links to Self-Reports of Affect and Affect Regulation. *Neuroimage* **202**, 116154 (2019).

84. Consedine, N. S., Magai, C., Cohen, C. I. & Gillespie, M. Ethnic variation in the impact of negative affect and emotion inhibition on the health of older adults. *Journals Gerontol. - Ser. B Psychol. Sci. Soc. Sci.* **57**, P396–P408 (2002).

85. Deer, L. K., Shields, G. S., Ivory, S. L., Hostinar, C. E. & Telzer, E. H. Racial/ethnic disparities in cortisol diurnal patterns and affect in adolescence. *Dev. Psychopathol.* **30**, 1977–1993 (2018).

86. Barzegaran, E., Vildavski, V. Y. & Knyazeva, M. G. Fine Structure of Posterior Alpha Rhythm in Human EEG: Frequency Components, Their Cortical Sources, and Temporal Behavior. *Sci. Rep.* **7**, 1–12 (2017).

87. Al Zoubi, O. *et al.* EEG Microstates Temporal Dynamics Differentiate Individuals with Mood and Anxiety Disorders From Healthy Subjects. *Front. Hum. Neurosci.* **13**, 1–10 (2019).

88. Watson, D. & Clark, L. A. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. **54**, 1063–1070 (1988).

89. John, R. & Julie, D. The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *Br. J. Clin. Psychol.* **43**, 245–65 (2004).

90. Tuccitto, D. E., Giacobbi, P. R. & Leite, W. L. The Internal Structure of Positive and Negative Affect: A Confirmatory Factor Analysis of the PANAS. *Educ. Psychol. Meas.* **70**, 125–141 (2010).

91. Cooper, A., Gomez, R. & Aucote, H. The Behavioural Inhibition System and Behavioural Approach System (BIS/BAS) Scales: Measurement and structural invariance across adults and adolescents. *Pers. Individ. Dif.* **43**, 295–305 (2007).

92. Renger, R. A review of the profile of mood states (POMS) in the prediction of athletic success. *J. Appl. Sport Psychol.* **5**, 78–84 (1993).

93. Ritter, P., Becker, R., Freyer, F. & Villringer, A. EEG quality: The image acquisition artefact. *EEG - fMRI Physiol. Basis, Tech. Appl.* 153–171, https://doi.org/10.1007/978-3-540-87919-0_9 (2010).

94. Ritter, P., Becker, R., Graefe, C. & Villringer, A. Evaluating gradient artifact correction of EEG data acquired simultaneously with fMRI. *Magn. Reson. Imaging* **25**, 923–932 (2007).

95. Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R. & Lemieux, L. Identification of EEG events in the MR scanner: The problem of pulse artifact and a method for its subtraction. *Neuroimage* **8**, 229–239 (1998).

96. Allen, P. J., Josephs, O. & Turner, R. A method for removing imaging artifact from continuous EEG recorded during functional MRI. *NeuroImage* **12**, 230–239 (2000).

97. Niazy, R. K., Beckmann, C. F., Iannetti, G. D., Brady, J. M. & Smith, S. M. Removal of FMRI environment artifacts from EEG data using optimal basis sets. *Neuroimage* **28**, 720–737 (2005).

98. Christov, I. I. Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomed. Eng. Online* **3**, 1–9 (2004).

99. Kim, K. H., Yoon, H. W. & Park, H. W. Improved ballistocardiac artifact removal from the electroencephalogram recorded in fMRI. *J. Neurosci. Methods* **135**, 193–203 (2004).

100. Gabard-Durnam, L. J., Leal, A. S. M., Wilkinson, C. L. & Levin, A. R. The harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Front. Neurosci.* **12**, 1–24 (2018).

101. Mullen, T. CleanLine EEGLAB Plugin. San Diego, CA: Neuroimaging Informatics Toolsand Resources Clearinghouse (NITRC). (2012).

102. Winkler, I., Haufe, S. & Tangermann, M. Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. *Behav. Brain Funct.* **7**, 1–15 (2011).

103. Winkler, I. *et al.* Robust artifactual independent component classification for BCI practitioners. *J. Neural Eng.* **11** (2014).

104. Pessoa, L. Understanding brain networks and brain organization. *Phys. Life Rev.* **11**, 400–435 (2014).

105. Sigman, M. & Dehaene, S. Brain Mechanisms of Serial and Parallel Processing during Dual-Task Performance. *J. Neurosci.* **28**, 7585–7598 (2008).

106. Murray, M. M., Brunet, D. & Michel, C. M. Topographic ERP analyses: A step-by-step tutorial review. *Brain Topogr.* **20**, 249–264 (2008).

107. Poulsen, A. T., Pedroni, A., Langer, N. & Hansen, L. K. Microstate EEGlab toolbox: An introductory guide. *bioRxiv* 289850, https://doi.org/10.1101/289850 (2018).

108. Rokach, L. & Maimon, O. *Data Mining with Decision Trees.* (2008).

109. Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T. & Winterer, G. Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. *Neuroimage* **52**, 1149–1161 (2010).

110. Van De Ville, D., Britz, J. & Michel, C. M. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc. Natl. Acad. Sci. USA* **107**, 18179–18184 (2010).

111. Brodbeck, V. *et al.* EEG microstates of wakefulness and NREM sleep. *Neuroimage* **62**, 2129–2139 (2012).

112. Kayser, J. & Tenke, C. E. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clin. Neurophysiol.* **117**, 348–368 (2006).

113. Kayser, J. & Tenke, C. E. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. *Clin. Neurophysiol.* **117**, 369–380 (2006).

114. Smith, E. E., Reznik, S. J., Stewart, J. L. & Allen, J. J. B. Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *Int. J. Psychophysiol.* **111**, 98–114 (2017).

115. Coan, J., Coan, J. A. & Allen, J. J. B. Frontal EEG asymmetry and behavioral activation and inhibition systems. *Frontal EEG asymmetry and the behavioral activation and inhibition systems.* **40**, 106–114 (2003).

116. Pernet, C. R., Wilcox, R. & Rousselet, G. A. Robust correlation analyses: false positive and power validation using a new open source Matlab toolbox. **3**, 1–18 (2013).

117. Abreu, R., Leal, A. & Figueiredo, P. EEG-Informed fMRI: A Review of Data Analysis. *Methods.* **12**, 1–23 (2018).

118. Pisauro, M. A., Fouragnan, E., Retzler, C. & Philiastides, M. G. Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nat. Commun.* **8**, 15808 (2017).

119. Murta, T., Leite, M. & Carmichael, D. W. Electrophysiological Correlates of the BOLD Signal for EEG-Informed fMRI. **00** (2014).

120. Sclocco, R. *et al*. EEG-informed fMRI analysis during a hand grip task: Estimating the relationship between EEG rhythms and the BOLD signal. *Front. Hum. Neurosci.* **8**, 1–13 (2014).
121. Laufs, H. *et al*. EEG-correlated fMRI of human alpha activity. *Neuroimage* **19**, 1463–1476 (2003).
122. Calhoun, V. D., Stevens, M. C., Pearlson, G. D. & Kiehl, K. A. fMRI analysis with the general linear model: Removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. *Neuroimage* **22**, 252–257 (2004).
123. Lindquist, M. A. & Wager, T. D. Validity and power in hemodynamic response modeling: A comparison study and a new approach. *Hum. Brain Mapp.* **28**, 764–784 (2007).
124. Gawlowska, M., Domagalik, A., Beldzik, E., Marek, T. & Mojsa-Kaja, J. Dynamics of error-related activity in deterministic learning - an EEG and fMRI study. *Sci. Rep.* **8**, 2–11 (2018).
125. Kok, P., Van Lieshout, L. L. F. & De Lange, F. P. Local expectation violations result in global activity gain in primary visual cortex. *Sci. Rep.* **6**, 1–10 (2016).
126. Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A. & Grafton, S. T. Differential Recruitment of the Sensorimotor Putamen and Frontoparietal Cortex during Motor Chunking in Humans. *Neuron* **74**, 936–946 (2012).
127. Wager, T. D., Keller, M. C., Lacey, S. C. & Jonides, J. Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage* **26**, 99–113 (2005).
128. Fritsch, V. *et al*. Robust regression for large-scale neuroimaging studies. *Neuroimage* **111**, 431–441 (2015).
129. Seghier, M. L. Laterality index in functional MRI: methodological issues. *Magn. Reson. Imaging* **26**, 594–601 (2008).
130. Bradshaw, A. R., Bishop, D. V. M. & Woodhead, Z. V. J. Methodological considerations in assessment of language lateralisation with fMRI: a systematic review. *PeerJ* **5**, e3557 (2017).
131. Jansen, A. *et al*. The assessment of hemispheric lateralization in functional MRI-Robustness and reproducibility. *Neuroimage* **33**, 204–217 (2006).
132. Branco, D. M. *et al*. Functional MRI of memory in the hippocampus: Laterality indices may be more meaningful if calculated from whole voxel distributions. *Neuroimage* **32**, 592–602 (2006).
133. Jones, S. E., Mahmoud, S. Y. & Phillips, M. D. A practical clinical method to quantify language lateralization in fMRI using whole-brain analysis. *Neuroimage* **54**, 2937–2949 (2011).
134. Wu, G. R., Deshpande, G., Laureys, S. & Marinazzo, D. Retrieving the Hemodynamic Response Function in resting state fMRI: Methodology and application. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2015-Novem**, 6050–6053 (2015).
135. Ding, J.-R. *et al*. A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med. Image Anal.* **17**, 365–374 (2013).
136. Tomarken, A. J., Davidson, R. J. & Henriques, J. B. Resting frontal brain asymmetry predicts affective responses to films. *J. Pers. Soc. Psychol.* **59**, 791–801 (1990).

Acknowledgements

Author contributions

The study was designed by Ardaman Kaur and Dr. Vijayakumar Chinnadurai. Data was acquired and processed by Ardaman Kaur. Data was interpreted and analyzed by Dr. Vijayakumar Chinnadurai, Ardaman Kaur and Dr. Rishu Chaujar. The manuscript was written by Ardaman Kaur with the help of Dr. Vijayakumar Chinnadurai. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-61119-7.

Correspondence and requests for materials should be addressed to V.C.

Reprints and permissions information is available at www.nature.com/reprints.

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

Modeling Involuntary Dynamic Behaviors to Support Intelligent Tutoring (Student Abstract)

Mononito Goswami[1,2*†], Lujie Chen[2†], Chufan Gao[2] and Artur Dubrawski[2]

[1] Delhi Technological University, New Delhi, India
[2] Auton Lab, Carnegie Mellon University, Pittsburgh, USA
mononito_bt2k16@dtu.ac.in, lujiec@andrew.cmu.edu, chufang@andrew.cmu.edu, awd@cs.cmu.edu

Abstract

Problem solving is one of the most important 21st century skills. However, effectively coaching young students in problem solving is challenging because teachers must continuously monitor their cognitive and affective states and make real-time pedagogical interventions to maximize students' learning outcomes. It is an even more challenging task in social environments with limited human coaching resources. To lessen the cognitive load on a teacher and enable affect-sensitive intelligent tutoring, many researchers have investigated automated cognitive and affective detection methods. However, most of the studies use culturally-sensitive indices of affect that are prone to social editing such as facial expressions, and only few studies have explored involuntary dynamic behavioral signals such as gross body movements. In addition, most current methods rely on expensive labelled data from trained annotators for supervised learning. In this paper, we explore a semi-supervised learning framework that can learn low-dimensional representations of involuntary dynamic behavioral signals (mainly gross-body movements) from a modest number of short time series segments. Experiments on a real-world dataset reveal a significant utility of these representations in discriminating cognitive disequilibrium and flow and demonstrate their potential in transferring learned models to previously unseen subjects.

1 Introduction

For young children, solving challenging non-routine math problems emulates real life challenges they will encounter later in their lives, and may often invite them to ride an "emotional roller-coaster" as the child advances through various stages of problem solving (Chen et al. 2016). Thus, effectively coaching young students for problem solving requires teachers to continuously monitor their cognitive and affective states and make real time pedagogical decisions such as when and how to best intervene. Moreover, teachers have to effectively handle the high cognitive load of monitoring diverse cohorts of students. Intelligent Tutoring Systems that attempt to teach problem-solving also face similar challenges. To lessen the cognitive load of teachers and improve the effectiveness of intelligent tutoring, we envision a decision support system which can monitor the cognitive and affective states of multiple students simultaneously in real time. The focus of this paper is on the state detection capability of such a system, specifically needed to discriminate between *cognitive disequilibrium* (CD) and *flow* states, which are the critical inputs to inform appropriate subsequent interventions. In this work, we investigate a method designed to discriminate between CD and flow using involuntary behavioral signals that are less prone to social editing, including head and eye movement, which can be non-invasively collected using inexpensive sensors such as cameras. To overcome limited supply of labeled data, while taking advantage of the large supply of unlabeled data, we explore a semi-supervised approach where deep embedding features are derived from unlabeled time series segments, which are then fed into a supervised learning algorithm.

2 Methodology

Our study is based on a dataset collected in one-to-one coaching scenarios for math problem solving. We extracted a number of features from the dataset along the visual (*Facial Action Units (FAUs), head & eye gaze orientations*) and writing channels (*writing speed*). We computed the first and second order derivatives of all visual features with the exception of FAUs.

The field of affective and cognitive computing relies on supervised learning algorithms, and is therefore heavily dependent on training data from expert annotators or self-reports by participants of a study. Since most advanced and powerful supervised learning algorithms require substantial amounts of training data to learn reliable decision functions, application of affective computing is severely limited by short supply of trained expert annotators or potentially biased self-reports. To this end, we investigated the utility of an unsupervised representation learning model proposed by (Franceschi, Dieuleveut, and Jaggi 2019), which can be trained on a large amount of unlabeled data to learn potentially useful feature representations. By automatically learning useful features for classifying raw data, representation learning algorithms replace manual feature engineering and

---

*Mobile Number: +91 8800592994

†The authors wish that it be known that, in their opinion, the first two authors should be regarded as joint First Authors.

allow systems to identify potential discriminators and use them to support a specific predictive task. Their model comprises of a deep neural network with dilated causal convolutions to handle time series and minimizes an unsupervised triplet loss function which assigns similar time series proximate embeddings based on the assumption that they occur in temporal proximity while a randomly chosen subseries is likely to be dissimilar.

The unsupervised representation learning model was trained on 248 time series segments each having 27 features over 300 time steps and returned 64-dimensional embeddings. Using these output embeddings as feature vectors and manually annotated labels, we trained a random forest classifier to predict the cognitive state (Flow or Cognitive Disequilibrium) of a time series segment. We chose random forests because they are able to learn non-linear and complex decision boundaries, work well with high-dimensional data and can be robust to outliers.

The unsupervised representation learning model coupled with a random forest classifier can function as a semi-supervised model, where the former learns embeddings (features) from a large number of time series segments in a completely unsupervised fashion, and the latter uses these features and a limited number of annotations to learn a decision function. Such a semi-supervised paradigm can be extremely useful in practice of affective computing, where obtaining vast amounts of unlabeled data is extremely easy, but its annotation can be expensive.

3   Results and Discussion

Predictive Utility of Deep Features

We evaluated the predictive utility of deep embedding features by feeding them into a random forest classifier. We conducted two types of experiments: (1) "Random" experiments which make a *random split* between train and test sets. These experiments could yield inflated algorithm performance as the information from the same session and subject may appear in both the training and testing sets, allowing the model to succeed by hooking-onto personal characteristics of some distinct subjects. (2) Leave-one-person(subject)-out experiments which represent a "cold start" scenario where the model is trying to predict for a completely unseen subject. Due to varying degrees of information sharing between training and test set, we expect the performance will degrade from the upper bound case of random split to the conservative (but of high practical utility) LOPO experiments. We also compared the performance of our semi-supervised model with *ResNet*. (Fawaz et al. 2019) in a recent and comprehensive survey found that ResNet can significantly

outperform other deep learning approaches in classifying time series on the UCR/UEA and MTS archives. In addition, they found encouraging results (comparable predictive performance and significantly less training & testing time) while comparing ResNet to other state-of-the-art time series classification algorithms such as HIVE-COTE.

Table 1 compares ResNet and the deep semi-supervised model introduced above on several performance metrics. The results reveal that while ResNet achieved higher accuracy (0.78%) in the LOPO experiments, there was no significant difference (within 95% confidence interval) between the models in terms of AUC in both evaluation scenarios.

How Much Supervision is Necessary?

We conducted sensitivity analysis to demonstrate the utility of unsupervised embeddings in the prediction task. In these set of experiments, we fixed the held-out test set, varied the size of the training set and reported the performance of our semi-supervised approach accordingly. Our experiments (refer Supplementary Material for details) revealed that with deep embeddings features, the model is able to learn effective discrimination with even a small number of labeled data points, and the resulting performance is comparable with a potent fully supervised deep learning alternative which often requires large extents of supervision.

Discussion

In this paper, we explored a semi-supervised framework to model the dynamics of involuntary behavioral signals in order to discriminate between cognitive disequilibrium and flow. Experimental results with a modestly sized multimodal multi-sensor dataset, collected from young children practicing problem solving in a naturalistic environment, reveal several insights. We found that our semi supervised approach was able to effectively generalize from training subjects to previously unseen subjects, as demonstrated by its robust performance with leave-one-person-out experiments. When further validated with a more diverse set of subjects, the proposed approach has the promise to scale up practicality of the task of cognitive and affective state detection that is often bottle-necked by high costs of label acquisition even with abundant unlabeled data. Practically relevant capability of generalization to unseen subjects is also encouraging as the proposed approach would often be expected to work well with out-of-sample subjects in the real world use cases.

References

Chen, L.; Li, X.; Xia, Z.; Song, Z.; Morency, L.-P.; and Dubrawski, A. 2016. Riding an emotional roller-coaster: A multimodal study of young child's math problem solving activities. *International Educational Data Mining Society.*

Fawaz, H. I.; Forestier, G.; Weber, J.; Idoumghar, L.; and Muller, P.-A. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* 33(4):917–963.

Franceschi, J.-Y.; Dieuleveut, A.; and Jaggi, M. 2019. Unsupervised scalable representation learning for multivariate time series. *arXiv preprint arXiv:1901.10738.*

Experiments	Semi-supervised		Supervised (ResNet)	
	Random	LOPO	Random	LOPO
Precision	0.83 (0.037)	0.81 (0.063)	0.81 (0.016)	0.77 (0.074)
Recall	0.82 (0.04)	0.7 (0.111)	0.81 (0.023)	0.78 (0.071)
F1	0.82 (0.04)	0.71 (0.092)	0.81 (0.024)	0.77 (0.069)
Accuracy	0.82 (0.04)	0.7 (0.111)	0.81 (0.022)	0.78 (0.072)
AUC	0.83 (0.052)	0.79 (0.058)	0.8 (0.016)	0.73 (0.081)

Table 1: Performance of semi-supervised model vs. ResNet.

Multi-Criteria decision analysis for optimizing proper selection medium for CSR in automobile industry

Article · January 2018

4 authors, including:

Reetik Kaushik
Delhi Technological University
2 PUBLICATIONS   **0** CITATIONS

SEE PROFILE

Ranganath Muttanna Singari
Delhi Technological University, Formerly Delhi college of Engineering, Delhi, India
101 PUBLICATIONS   **251** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Dates Extended for Full manuscript Submission at ICAPIE 2016, Delhi, India View project

Publications by Prof. Ranganath M Singari View project

Multi-Criteria decision analysis for optimizing proper selection medium for CSR in automobile industry

Reetik Kaushik[1*], Roopa Singh[2], Sanchit Gupta[3], Ranganath M Singari[4]
[1-3] Department of Mechanical Engineering, Delhi Technological University, Delhi, India
[4] Department Production & Industrial Engineering, Delhi Technological University, Delhi, India

Abstract
Industrial sector is comprised of various set of financial and non-financial target. Corporate Social Responsibility (CSR) is a concept that underlies a multi-pillar strategy where many quantitative and qualitative objectives of multiple stakeholders lever a company's success. CSR refers to the totality of a corporation's financial, social, and environmental performance in conducting its business. Making sure CSR initiatives tie back to company's mission and align with their brand identity. It can be tempting to support a wide variety of non-profits, but it is more powerful to pick a few core focus areas in which to make a significant impact. Within the field of operations research, MCDA has evolved over the last three decades as one major discipline, it provides financial decision makers (DMs) and analysts a wide range of methodologies, which are well suited to the complexity of selecting optimum area for CSR in automobile Industry. The aim of this study is to provide an in-depth presentation of the contributions of MCDA in the field of finance, focusing on the methods used and their real-world applications to form a framework for selecting most suitable and profitable CSR initiatives through a survey of companies in automobile Industry.

Keywords: corporate social responsibility, multi-criteria decision aid, finance, multi-objective optimization

1. Introduction

Many believe that the sole responsibility of a business is to make a profit. The maximization of financial efficiency is considered to be the primary goal of an enterprise within the theory of corporate finance In this case, financial goals constitute a common denominator for the evaluation of a firm's outcomes; by this, the integrity and purposefulness of a firm's management are safeguarded. Financial-goal orientation determines the related decision-making. However, as to how the financial goal is expressed and organized is of no importance from the perspective of a firm's financial and non-financial outcomes. Consequently, the financial and non-financial goals of enterprises are deliberated with respect to how goal-setting affects management practices.

Corporate Social Responsibility (CSR) is a concept that suggests that it is the responsibility of the corporations operating within society to contribute towards economic, social and environmental development that creates positive impact on society at large. Although there is no fixed definition, however the concept revolves around that fact the corporations needs to focus beyond earning just profits. The term became popular in the 1960s and now is formidable part of business operations [1]. A CSR strategy begins with a long-term vision and commitment from the top of the executive food chain. Optimizing proper Selection medium for CSR in Automobile Industry can be considered a multi criteria decision making problem as there are many factors (criterion) which are to be considered for selecting the best area of CSR. Analytical Hierarchy Process is a multi-criteria decision making tool which can be used to find an optimum solution for selecting most beneficial area in terms of financial as well non-financial growth of industry AHP is a structured technique for organising and analysing complex decisions rather than focusing on the correct decision, the AHP helps decision makes find most optimum solution for achieving their target. It provides a comprehensive and rational framework for structuring and decision problem, for representing and quantifying its elements for relating those elements with overall goals and for evaluating alternative solutions.

2. Statement Problem

In any company the main objective is allocating its resources both financial and non-financial in cost effect, time efficient and productive manner. Although allocating companies funds in Corporate Social Responsibility (CSR) activities is beneficial to company but when it comes to deciding which areas of CSR to invest in, it becomes a problematic situation. If they make most appropriate choice of selection CSR area company can exhibit following benefits improved financial performance, Lower operating costs, Enhanced brand image and reputation, Increased sales and customer loyalty, bigger productivity and quality, additional ability to draw in and retain workers, Reduced regulative oversight, Access to capital, manpower diversity, Product safety and small liability. The idea of company social responsibility is currently firmly stock-still on the world business agenda. However so as to maneuver from theory to concrete action, several obstacles got to be overcome [2]. This method must consider the various characteristics benefits and risks before implementing solutions for selecting optimum area for CSR in automobile Industry. The aim of this study is to provide an in-depth presentation of the contributions of MCDA in the field of finance, focusing on the methods used and their real-world applications to form a framework for selecting most suitable and profitable CSR initiatives through a survey of companies in automobile Industry. The branch of MCDA that is considered for providing solution is Analytic Hierarchy Process (AHP). It is very powerful and powerful tool because the scores, and therefore the final ranking, are obtained on the basis of the pairwise relative evaluations of both the criteria

and the options provided by the user.

3. Literature Review

For the process of selecting the best alternatives for determining the key areas for CSR activities is very crucial step. Realizing the fact many researchers have done a lot of work in this area. Constantin Zopounidis [3] explained how the complexity of financial decisions has rapidly increased over the decade, thus highlighting the importance of developing and implementing sophisticated and efficient quantitative analysis techniques for supporting and aiding financial decision making. Multi-criteria decision aid (MCDA), an advanced field of operations research, provides financial decision makers (DMs) and analysts a wide range of methodologies, which are well suited to the complexity of financial decision problems.

Further a descriptive study done by Anna Doś [4] explained about how despite the fact that some studies focused on both CSR and finance, little is still known about how MCDMs can be used to integrate financial management and CSR. The research also concluded how maximization of financial efficiency constitutes a common denominator for the evaluation of a firm's outcomes, the integrity and purposefulness of the firm's management are safeguarded.

The key benefits company exhibits while involving itself with CSR activities are summed up as follows:

1. **Improves Brand Image:** Companies that demonstrate their commitment to various causes are perceived as more philanthropic than companies whose corporate social responsibility endeavours are non-existent.
2. **Brand Awareness by Media Coverage:** Media visibility is important as it sheds a positive light on your organization and shows that company is transparent.
3. **Boosts Employee Engagement:** Employees like working for a company that has a good public image and is constantly in the media for positive reasons.
4. **Networking and Maintaining Relationships:** Interaction with client outside the scope of the project to overcome the problem of relationship discontinuity, regular communication with the clients.
5. **Generating Sales leads:** connecting with various employees of target companies as potential clients.
6. **To increase Web Traffic:** with availability of company's profile on various social platforms it becomes easy to redirect to company's website thereby increasing web traffic.

CSR activities are focused on six broad themes with the goal of improving the overall socio- economic status in its areas of operation

1. Promoting healthcare, sanitation and making safe drinking water available
2. Employment enhancement through training and vocational skill development
3. Income enhancement through farm based and other livelihood opportunities
4. Promoting education & sports
5. Environmental sustainability
6. Infrastructure development

From literature review one can conclude that there exist lots of beneficial factors that can be linked with CSR activities and finance but one must eliminate most obvious choices and should pay proper attention while selecting the characteristic benefits through CSR and areas where they can be implemented. The next portion describes the methodology of AHP approach.

4. Research Methodology
4.1 Multi-Criteria Decision Aid (MCDA)

Multi-criteria decision aid (MCDA) or multiple-criteria decision analysis (MCDA), an advanced field of operations research that explicitly evaluates multiple conflicting criteria in decision making. Research into operations provides many interesting insights into multiple-criteria decision-making (MCDM) and multi-attribute decision tools, which then enhance decision rationality under circumstances where a number of heterogenic objectives must be achieved. The most-popular multi-criteria decision tools include AHP, PROMETHEE, ELECTRE, TOPSIS, and their families [5].

The Analytic Hierarchy Process (AHP), introduced by Thomas Saaty (1980), is an effective tool for dealing with complex decision making, and may aid the decision maker to set priorities and make the best decision. By reducing complex decisions to a series of pairwise comparisons, and then synthesizing the results, the AHP helps to capture both subjective and objective aspects of a decision. In addition, the AHP incorporates a useful technique for checking the consistency of the decision maker's evaluations, thus reducing the bias in the decision making process [6].

Steps involved in AHP are as follows:

1. Identify the problem.
2. Identify the criteria that influences the behavior of the problem-brainstorming and judgement of expert are used to identify the criteria.
3. Structure the hierarchy of criteria and alternatives
4. Remove ambiguity by carefully defining every element in the hierarchy
5. Prioritize the primary criteria with respect to their impact on the overall objective. This is done by first comparing every element to each other, two at a time. This is called pairwise comparison. After the scores are given multiply the scores and raise it to power of 'n' where 'n' is the number of criteria (and alternatives). Add these scores and divide each score with the sum obtained. This way we obtain the normal priority vector.
6. Similarly, prioritize the sub-criteria with respect to respective alternatives.
7. To obtain the overall rankings, multiply the matrix obtained by prioritizing the alternatives and transpose of the priority vector of characteristics.
8. For checking the consistency of the results, first calculate the Principal Eigenvalue (λmax) of the priority vector. For consistency index use the formula:

$$C.I = \frac{\lambda \max - n}{n - 1}$$

For consistency ratio, use the formula:

$$C.R = \frac{C.I}{R.I}$$

A consistency ratio of 0.1 or less is considered acceptable. For our case with n = 6 in both the characteristics matrix and sub-criteria matrix. Therefore R.I for n = 6 is 1.24.

5. Numerical illustration

During this study, through different research papers and by expert opinion we have identified some of the key benefits achieved by company by involving themselves with CSR

activities. Some of those factors are:

N_1- Improves Brand Image, N_2- Brand Awareness by Media Coverage, N_3- Boosts Employee Engagement, N_4- Networking and Maintaining Relationships, N_5- Generating Sales leads, N_6- To increase Web Traffic. Some key areas where CSR activities are focused in automobile industry. These areas are represented as: M_1- Promoting healthcare, sanitation and making safe drinking water available, M_2- Income enhancement through farm based and other livelihood opportunities, M_3- Promoting education & sports, M_4- Employment enhancement through training and vocational skill development, M_5- Environmental sustainability and M_6- Infrastructure development. First step is collection of data from all experts and analyse it and summarize it. Then AHP methodology is followed for generating optimum solution.

Table 1: Pairwise comparison of required characteristic benefits of CSR

	N1	N2	N3	N4	N5	N6
N1	1	0.25	0.5	0.5	0.25	0.25
N2	4	1	2	2	1	1
N3	2	0.5	1	1	0.5	0.5
N4	2	0.5	1	1	0.5	0.5
N5	4	1	2	2	1	1
N6	4	1	2	2	1	1

$\lambda_{max} = 6$, C.I = 0, C.R = 0

Table 2: Pairwise comparison of various areas of CSR based on N1

	M1	M2	M3	M4	M5	M6
M1	1	2	0.25	6	6	6
M2	0.25	1	0.17	4	4	4
M3	4	6	1	8	8	8
M4	0.17	0.25	0.125	1	1	1
M5	0.17	0.25	0.125	1	1	1
M6	0.17	0.25	0.125	1	1	1

$\lambda_{max} = 6.1474$, C.I = 0.02948, C.R = 0.023774

Table 3: Pairwise comparison of various areas of CSR based on N2

	M1	M2	M3	M4	M5	M6
M1	1	3	1	0.25	4	4
M2	0.33	1	0.25	0.17	2	2
M3	1	4	1	0.25	5	5
M4	5	3	4	1	6	6
M5	0.25	0.5	0.5	0.17	1	2
M6	0.25	0.5	0.20	0.17	0.5	1

$\lambda_{max} = 6.2724$, C.I = 0.05448, C.R = 0.043935

Table 4: Pairwise comparison of various areas of CSR based on N3

	M1	M2	M3	M4	M5	M6
M1	1	2	0.25	4	4	3
M2	0.5	1	0.17	1	0.5	0.5
M3	4	6	1	6	4	4
M4	0.25	1	0.17	1	0.5	0.5
M5	0.33	2	0.25	2	1	1
M6	0.33	2	0.25	2	1	1

$\lambda_{max} = 6.2005$, C.I = 0.0401, C.R = 0.032339

Table 5: Pairwise comparison of various areas of CSR based on N4

	M1	M2	M3	M4	M5	M6
M1	1	1	1	2	2	0.20
M2	1	1	1	2	2	0.20
M3	1	1	1	2	2	0.33
M4	0.5	0.5	0.5	1	1	0.25
M5	0.5	0.5	0.5	1	1	0.25
M6	5	5	3	4	4	1

$\lambda_{max} = 6.1193$, C.I = 0.02386, C.R = 0.019242

Table 6: Pairwise comparison of various areas of CSR based on N5

	M1	M2	M3	M4	M5	M6
M1	1	1	1	0.5	2	4
M2	1	1	1	0.5	2	4
M3	1	1	1	0.5	2	4
M4	2	2	2	1	6	8
M5	0.5	0.5	0.5	0.17	1	2
M6	0.25	0.25	0.25	0.125	0.5	1

$\lambda_{max} = 6.0184$, C.I = 0.00368, C.R = 0.002968

Table 7: Pairwise comparison of various areas of CSR based on N6

	M1	M2	M3	M4	M5	M6
M1	1	1	1	1	2	4
M2	1	1	1	1	2	4
M3	1	1	1	1	2	4
M4	1	1	1	1	2	4
M5	0.5	0.5	0.5	0.5	1	2
M6	0.25	0.25	0.25	0.25	0.5	1

$\lambda_{max} = 6$, C.I = 0, C.R = 0

Table 7: Pairwise comparison of various areas of CSR based on N6

	M1	M2	M3	M4	M5	M6
M1	1	1	1	1	2	4
M2	1	1	1	1	2	4
M3	1	1	1	1	2	4
M4	1	1	1	1	2	4
M5	0.5	0.5	0.5	0.5	1	2
M6	0.25	0.25	0.25	0.25	0.5	1

$\lambda_{max} = 6$, C.I = 0, C.R = 0

6. Results and Analysis

From Table 8. It is evident that Employment enhancement through training and vocational skill development should get most priority considering the various characteristic benefits for which company receives while involving with CSR activities. Following Employment enhancement through training and vocational skill development is Promoting education & sports, Promoting healthcare, sanitation and making safe drinking water available and Income enhancement through farm based and other livelihood opportunities. These areas are based on automobile industry expert's point of view.

Table 8: Priority vector of characteristic benefits of CSR and priority matrix of key areas of CSR

	N1 0.0588	N2 0.2352	N3 0.1176	N4 0.1176	N5 0.2352	N6 0.2353	Priority Weight
M1	0.231842908	0.18354273	0.204651845	0.130937654	0.170731382	0.210526316	0.18601
M2	0.125107874	0.07493101	0.06312699	0.130937654	0.170731382	0.210526316	0.13790
M3	0.510352278	0.20742622	0.459427859	0.142573676	0.170731382	0.210526316	0.23935
M4	0.044232313	0.43612998	0.059077899	0.076270728	0.36533516	0.210526316	**0.25664**
M5	0.044232313	0.05461896	0.105264849	0.076270728	0.079787493	0.105263158	0.08035
M6	0.044232313	0.0433511	0.105264849	0.44300956	0.042682845	0.052631579	0.00973

However, this strategic and expert opinion analysis cannot be generalised for all the industries and is limited to content and expert advice of automobile industry only. Also the criteria chosen also was dependent upon the very few experts when compared to whole automobile sector thus the results can be improved with the help of more number of experts and by giving more time to the problem faced by automobile company's during CSR activities.

7. Conclusion
The study of this paper focuses on providing financial decision makers (DMs) and analysts a wide range of methodologies, which are well suited to the complexity of selecting optimum area for CSR in automobile Industry. To find the proper solution of the cited problem Analytical hierarchy approach was used which is tool of MCDA. Out of the chosen areas where a company can invent in CSR activities importance rating order comes as: Employment enhancement through training and vocational skill development (M4) > Promoting education & sports (M3) > Promoting healthcare, sanitation and making safe drinking water available (M1) > Income enhancement through farm based and other livelihood opportunities (M2) > Environmental sustainability (M5) > Infrastructure development (M6). Hence the automobile company will be most benefited if they invest in Employment enhancement through training and vocational skill development.

These studies however have certain limitation. First of all the study was made in context of automobile industry sector and criteria for deciding the optimum area for CSR activity is based on few expert opinion which can vary from business to business. They may also differ within different automobile company.

Availability of data and materials
The study is based on secondary data which have been collected from the annual audited reports of sampled automobile industries. All data generated and analyzed during the study are included in this article. All sample industries that were analyzed for this study maintain company websites and provide information regarding products and services via downloadable PDF format documents of their annual reports.

Authors' contributions
First author conducted the literature review, data collection, tabulation, and empirical studies. Second author assisted to check and interpret the whole paper several times including referencing. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interest.

8. References
1. Corporate Social Responsibility (CSR) by fiinovation(http://www.fiinovation.co.in/corporate-social-responsibility/)
2. Corporate social responsibility (CSR) by iisd (https://www.iisd.org/business/issues/sr.aspx)
3. Zopounidis C, Doumpos M, Multi-Crit J. Decis. Anal. 2002; 11:167-186.
4. Doś A. Managerial Economics. 2017; 18(1):63-86. http://dx.doi.org/10.7494/manage. 2017; 18(1):63.
5. Triantaphyllou E, Shu B, Nieto Sanchez S, Ray T. Encyclopedia of Electrical and Electronics Engineering, (J.G. Webster, Ed.), John Wiley & Sons, New York, NY. 1998; 15:175-186.
6. Saaty TL. The Analytic Hierarchy Process. McGraw-Hill, New York, 1980.

Narrow-Band UVB-Emitting Gd-Doped SrY$_2$O$_4$ Phosphors

VIJAY SINGH,[1,5] K. SWAPNA,[2] SUMANDEEP KAUR,[3] A.S. RAO,[3]
and J.L. RAO[4]

1.—Department of Chemical Engineering, Konkuk University, Seoul 05029, Republic of Korea.
2.—Department of Physics, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur, Andhra Pradesh 522502, India. 3.—Department of Applied Physics,
Delhi Technological University, Bawana Road, New Delhi 110 042, India. 4.—Department of
Physics, S.V. University, Tirupati, Andhra Pradesh 517 502, India. 5.—e-mail:
vijayjiin2006@yahoo.com

Gadolinium (Gd^{3+}) ion-activated SrY$_2$O$_4$ phosphors with varying concentrations of Gd^{3+} ions were prepared using sol–gel synthesis. The photoluminescence spectra were recorded to study the luminescence behavior of the as-synthesized phosphors. Upon excitation at 275 nm, the material displayed two bands, one at 309 nm ($^6P_{5/2} \rightarrow {}^8S_{7/2}$) and the other at 315 nm ($^6P_{7/2} \rightarrow {}^8S_{7/2}$), in the narrowband ultraviolet (UV) spectral regions. The strong emission at 315 nm indicated that the phosphors could be used as a narrowband UV light source. Electron paramagnetic resonance spectra showed a foremost Gd^{3+} center with weak signals. Several resonance signals revealed that Gd^{3+} ions underwent different Zeeman interactions and exhibited varying crystal field strength. The results suggest the potential utility of Gd^{3+} doped-SrY$_2$O$_4$ phosphors for phototherapy lamp applications which require UV emission.

Key words: Sol–gel, XRD, EPR, Gd^{3+} ions, SrY$_2$O$_4$, photoluminescence

INTRODUCTION

Over the past few decades, binary rare-earth (RE) oxides, AR$_2$O$_4$ (A = Ca, Sr, Ba and R = rare earths), have garnered enormous interest from researchers seeking to explore their utility in lighting, sensors and display applications.[1–3] Various hosts, including BaGd$_2$O$_4$, SrEr$_2$O$_4$, SrGd$_2$O$_4$, SrLu$_2$O$_4$, BaY$_2$O$_4$, and SrYb$_2$O$_4$, have been reported for their luminescence characteristics and thermal and chemical stability.[4–7] The thermal properties of SrY$_2$O$_4$ have been studied for its application in thermal barrier coating (TBC).[8] In the process of identifying a new host material with excellent luminescence and thermal stability for application in a variety of fields, SrY$_2$O$_4$ phosphors doped with different RE ions have been prepared and investigated.[9,10]

The f–f and f–d transitions of RE ions are responsible for their radiative properties, which enable their use in diverse fields. Gd^{3+} ions in diamagnetic phosphors offer both scientific and technological opportunities, for example, in laser materials. Gd^{3+} contains seven unpaired electrons in its outermost shell with total spin S = 7/2 and $^8S_{7/2}$ as ground state. The Gd^{3+} ions exhibit luminescence in the ultraviolet (UV) regime of the electromagnetic spectrum. Gd^{3+}-doped materials have attracted research interest owing to their preeminent magnetic, electrical and optical properties. Gd^{3+}-doped phosphor that emits luminescence in the UV region has shown potential for use in phototherapy lamps, water purification, photochemistry, x-ray imaging and low-pressure lamps.[11–13] Electron paramagnetic resonance (EPR) techniques are used to study the nature of paramagnetic ions (having one or more unpaired electrons) as dopants in host materials. EPR study reveals the structural and dynamical information of the doped ions in the host crystals and information regarding the defect centers formed in the lattice responsible for charge compensation and the process of luminescence.[14–16] The EPR spectrum is therefore distinguished by a large number of individual transitions. A number of studies have been carried out on Gd^{3+} ions

incorporated in highly symmetric crystals. In host crystals with low symmetry, a large number of signals can be expected, and spectroscopic analysis will thus be complex and challenging. The Gd^{3+} ion is different from other rare-earth (RE) ions, since its orbital magnetic momentum is zero. Its energy levels may be affected through high-order interactions with the crystalline electric field.[17]

In this report, Gd^{3+}-activated SrY_2O_4 phosphors with different concentrations of Gd^{3+} ions in the host were prepared by a sol–gel method, and their structural and morphological properties were studied. The luminescence characteristics and EPR spectra were analyzed to determine their potential utility as a UV-emitting phosphor.

MATERIAL PREPARATION AND ANALYSIS

A series of SrY_2O_4:Gd^{3+} (SY1-SY6) phosphor samples were prepared by a sol–gel method. High-purity starting materials [$Sr(NO_3)_2$, $Y(NO_3)_3 \cdot 6H_2O$, citric acid and $Gd(NO_3)_3 \cdot 6H_2O$] were used for the synthesis. The compositional details of the starting materials are shown in Table I. Initially, stoichiometric amounts of metal nitrates were placed in a 150-mL glass beaker and then dissolved in 10 mL of distilled water under continuous stirring. After a few minutes, citric acid was added to the solution at a ratio of citric acid/metal ions = 2:1, and stirring was continued for 1 h to obtain a clear solution. The solution was kept at 110°C in an oven overnight to obtain a dry gel. The resulting dry gel was then heated in a furnace at 400°C for 2 h, and a brown residual powder was obtained which was then ground using a mortar and pestle. This residual powder was annealed at 1000°C for 3 h, and the resulting white powder was used for analyzing several characteristics.

To identify the phases of the samples, x-ray diffraction (XRD) patterns were recorded using a Rigaku Miniflex II diffractometer with Cu-Kα radiation ($\lambda = 1.5406$ Å) as an x-ray source. Samples were scanned by a scanning electron microscope (SEM; S-3400, Hitachi, Japan) to obtain the details of the surface morphology. A spectrofluorophotometer (Shimadzu RF-5301PC) equipped with a xenon flash lamp was used to carry out photoluminescence (PL) measurements. A JEOL FE1X electron spin resonance (ESR) spectrometer was used to record the EPR spectra of the sample.

RESULTS AND DISCUSSION

XRD Analysis

Figure 1 shows the powder XRD patterns of SrY_2O_4 and SrY_2O_4:Gd^{3+} (SY1-SY6) compounds prepared through a citrate sol–gel synthesis. The XRD pattern indicates the formation of an orthogonal phase of SrY_2O_4, and the peaks were matched to the standard values of SrY_2O_4 given in JCPDS (no. 74-0264). The diffraction peak also shows one impurity peak present around $2\theta = 29.18°$, which is indicated by an asterisk (*). This impurity peak presents in SrY_2O_4:Gd^{3+} (SY1–SY6) samples. The intensity of this peak increases with increased Gd^{3+} concentration. According to Zhou et al.,[18] this is a characteristic peak of Y_2O_3. They have prepared SrY_2O_4 by the sol–gel method at 1250°C for 12 h, but observed the characteristic peak of Y_2O_3. Pavitra et al.[19] prepared SrY_2O_4 by the sol–gel method at 1400°C, and they also observed the characteristic peak of Y_2O_3. The most intense XRD peak of Gd_2O_3 is at $2\theta = 28.6°$ (JCPDS no. 11-0604), whereas we observed an impurity peak at around $2\theta = 29.18°$, which shows that the impurity peak belongs to only Y_2O_3. This also confirmed that dopant (i.e. Gd^{3+}) ions have no impact on the crystal structure of SrY_2O_4. As shown in the XRD pattern, upon increasing the Gd^{3+} concentration, the intensity of the diffraction peak at $2\theta = 30.52°$ (320) decreases. However, the intensity of the diffraction peak at $2\theta = 31.54°$ (121) increases, probably because of the difference in ionic radius of Gd^{3+} and Y^{3+} ions. The full width at half maximum (FWHM) of the strongest diffraction peak was used to calculate the crystalline size, applying Scherrer's equation $D = 0.9\lambda/\beta\cos\theta$, where λ is the wavelength of incident x-rays, θ is the corresponding Bragg diffraction angle and β is the FWHM of the (320) peak. The crystallite sizes were computed to be in the range of

Table I. Detailed information on sample composition, sample code and starting materials

Sample composition	Sample code	Starting materials			
SrY_2O_4:0.01Gd	SY1	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0090 g
SrY_2O_4:0.03Gd	SY2	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0270 g
SrY_2O_4:0.05Gd	SY3	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0451 g
SrY_2O_4:0.07Gd	SY4	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0631 g
SrY_2O_4:0.09Gd	SY5	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0812 g
SrY_2O_4:0.11Gd	SY6	Sr = 0.4232 g	Y = 1.5320 g	CA = 2.3052 g	Gd = 0.0992 g

Sr = $Sr(NO_3)_2$, Y = $Y(NO_3)_3 \cdot 6H_2O$, CA = citric acid, Gd = $Gd(NO_3)_3 \cdot 6H_2O$.

42.32–47.13 nm and are listed in Table II along with the FWHM for all the samples.

Figure 2 shows the crystal structure of SrY$_2$O$_4$. The SrY$_2$O$_4$ phase belongs to the orthogonal family of the CaFe$_2$O$_4$ structure with a *Pnam* space group. The orthorhombic crystal structure of SrY$_2$O$_4$ comprises a double-octahedral Y$_2$O$_4{}^{2-}$ framework, where the first is the undistorted octahedral site and the second is a substantially distorted site.[20] The Sr^{2+} site resides within the framework. Due to the similar valence of cations and proximity of ionic radii, Gd^{3+} ion prefers to occupy Y^{3+} sites.

Scanning Electron Microscopy (SEM) Analysis

The morphology of the SrY$_2$O$_4$:0.07Gd (SY4) compound was analyzed using SEM micrographs, as shown in Fig. 3. These SEM images show that particles are aggregates of varying shape and size, which consist of several small crystallites (see Fig. 3a and b). The particles are highly interconnected, having particle sizes ranging from 1–2 μm (see Fig. 3c). Since SEM and XRD refer to polycrystals and a single crystal, respectively, a difference occurs in their estimated crystallite sizes. The SEM image shown in Fig. 3d represents the sample that also contains some smaller particles that might have formed during the evolution of a huge amount of gas molecules during the sintering process that generally yields nanosized particles.

Photoluminescence Analysis

Figure 4 shows the PL spectra of Gd^{3+} in SrY$_2$O$_4$ phosphors. Figure 4a shows the excitation spectra of SrY$_2$O$_4$:Gd^{3+} phosphor recorded at λ_{em} = 315 nm. The spectra exhibited bands positioned at 245 nm, 253 nm and 275 nm wavelengths corresponding to $^8S_{7/2} \rightarrow {}^6D_J$ and $^8S_{7/2} \rightarrow {}^6I_J$ transitions of Gd^{3+} ions, respectively.[21–23] Among these transitions, the $^8S_{7/2} \rightarrow {}^6I_J$ transition (275 nm) is very intense. Therefore, the emission spectra of SrY$_2$O$_4$:Gd^{3+} phosphors were recorded in the spectral region from 300 nm to 330 nm, upon excitation at 275 nm (Fig. 4b). The emission spectra exhibit a weak band in the UV region at 309 nm, followed by a sharp intense band at 315 nm, corresponding to $^6P_{5/2} \rightarrow {}^8S_{7/2}$ and $^6P_{7/2} \rightarrow {}^8S_{7/2}$ transitions of the Gd^{3+}, respectively.[21–23] It has been reported that narrow-band UV emission in the range 304–315 nm is useful for phototherapy applications.[24,25] The PL spectra of the prepared phosphors confirm their potential utilization in medical treatment of skin disorders. The various concentrations of Gd^{3+} ions in the prepared phosphors were also studied in order to obtain the optimized Gd^{3+} ion
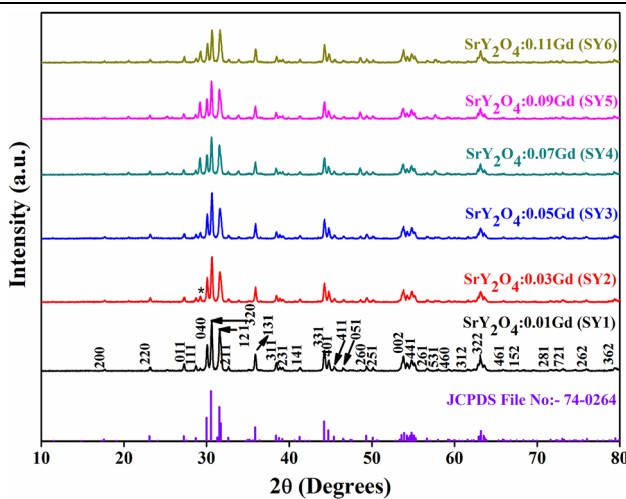
Fig. 1. XRD patterns of SrY$_2$O$_4$:Gd^{3+} (SY1–SY6) phosphors.
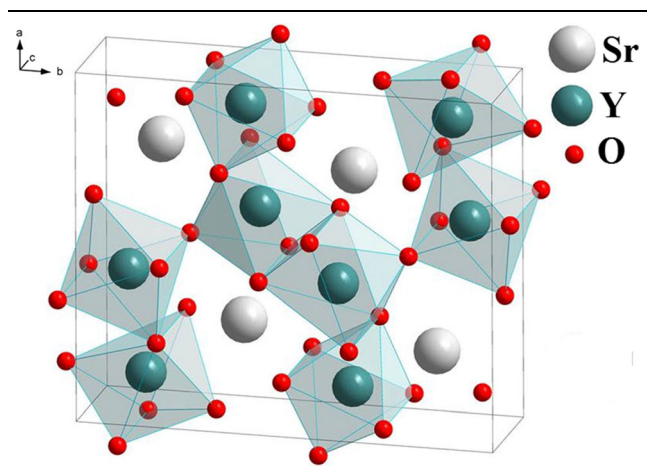
Fig. 2. Schematic illustration of the orthogonal SrY$_2$O$_4$ structure.

Table II. FWHM and crystallite size of SrY$_2$O$_4$:Gd^{3+} (SY1–SY6) phosphor

Sample composition	Sample code	FWHM (°)	Crystalline size (nm)
SrY$_2$O$_4$:0.01Gd	SY1	0.1904	45.20
SrY$_2$O$_4$:0.03Gd	SY2	0.2032	42.36
SrY$_2$O$_4$:0.05Gd	SY3	0.2034	42.32
SrY$_2$O$_4$:0.07Gd	SY4	0.1826	47.13
SrY$_2$O$_4$:0.09Gd	SY5	0.1829	47.06
SrY$_2$O$_4$:0.11Gd	SY6	0.1982	43.43

Fig. 3. SEM images of SrY$_2$O$_4$:0.07Gd (SY4) phosphors at different resolutions: (a) 10 μm scale bar, (b) 5 μm scale bar, (c, d) 1 μm scale bar at different sites.
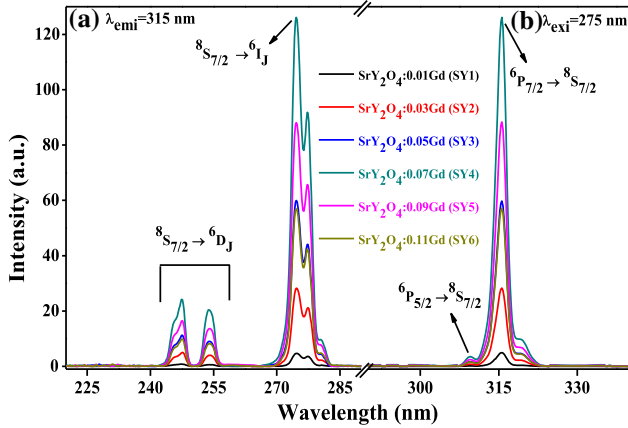
Fig. 4. Photoluminescence spectra of SrY$_2$O$_4$:Gd^{3+} (SY1–SY6) phosphors: (a) excitation spectra (λ_{em} = 315 nm) and (b) emission spectra (λ_{exc} = 275 nm).

Fig. 5. Variation in the emission intensity of strong emission (315 nm) as a function of Gd^{3+} concentration.
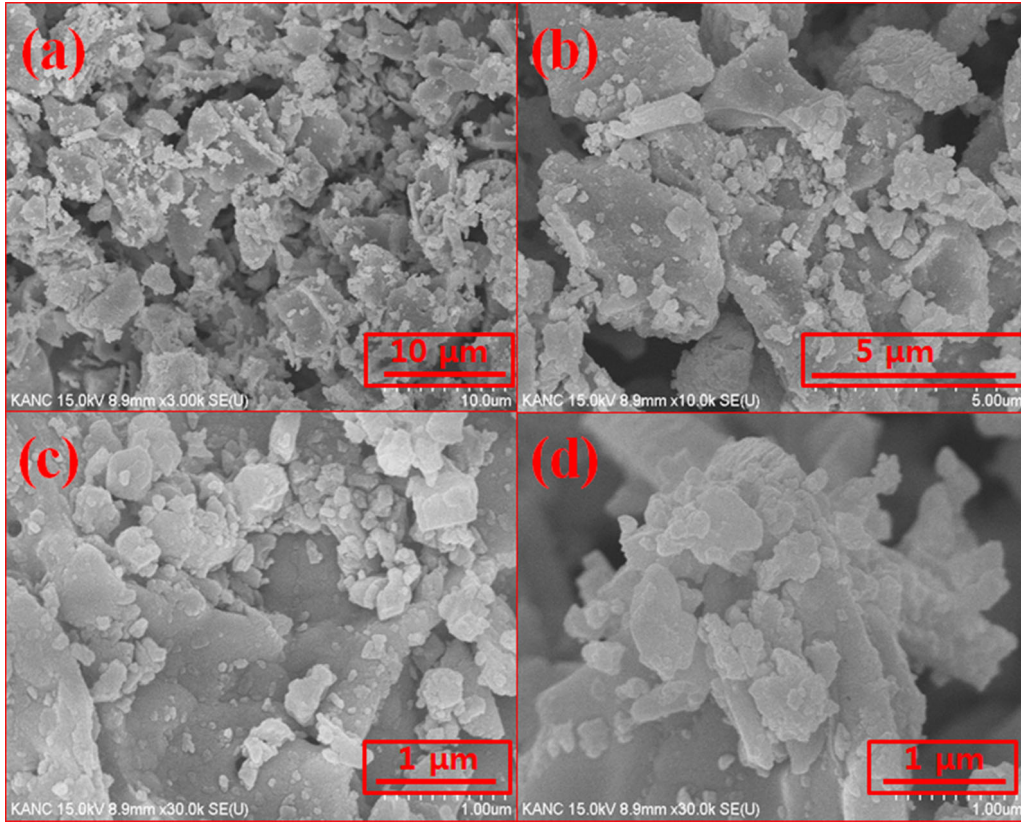
concentration. Figure 5 shows the emission intensity variation for the intense peak with Gd^{3+} ion concentrations from 0.01 mol to 0.11 mol. From Fig. 5 it is evident that the optimal Gd^{3+} concentration in the prepared phosphors is 0.07 mol (SY4 phosphor), and beyond this concentration, quenching occurs due to non-radiative energy transfer between Gd^{3+} ions. This energy transfer process strongly depends on the critical distance between

neighboring Gd^{3+} ions at a critical Gd^{3+} ion concentration. Generally, the energy transfer mechanism may proceed through exchange interaction or multipole–multipole interaction.[26] The critical distance (R_c) can be estimated using the following equation:

$$R_c \approx 2 \left[\frac{3V}{4\pi\chi_c N} \right]^{1/3} \tag{1}$$

where V is the unit cell volume of the SrY$_2$O$_4$ (409.37 Å^3),[27] χ_c is the critical ion concentration of Gd^{3+} ions at which the concentration quenching occurs (0.07), and N is the number of cations in the unit cell.[18] The R_c value obtained was 11.17 Å, which is higher than 5 Å. Therefore, the concentration quenching phenomenon that occurred in SrY$_2$O$_4$:Gd^{3+} (SY1–SY6) compound was mainly due to multipolar interactions. Therefore, 0.07 mol of Gd^{3+} ion in SrY$_2$O$_4$ phosphor is optimal for use in dermatology applications utilizing phototherapy lamps needed to treat several skin disorders. Figure 6 shows the partial energy level diagram of Gd^{3+} ion. After irradiation by UV radiation, the electrons of Gd^{3+} are first excited from the ^{8}S$_{7/2}$ ground state to the ^{6}D$_J$ and ^{6}I$_J$ excited levels. The electrons then return to the lowest excited ^{6}P$_J$ levels through a non-radiative relaxation process. Narrow-band UVB emission can be obtained when electrons in the ^{6}P$_J$ level relax down to the ^{8}S$_{7/2}$ ground state, as shown in Fig. 6.

EPR Analysis

In general, the ground state ^{8}S$_{7/2}$ of Gd^{3+} ions splits into eight $(2S + 1)$ states, thereby leading to seven allowed transitions according to the selection rule, $\Delta M_S = \pm 1$.[28] The non-zero interactions of Gd^{3+} ions are much less susceptible to the crystal field due to Zeeman splitting as compared to other ions. Gd^{3+} exhibits a long spin relaxation time, and is therefore observable at room temperature (RT) and higher temperatures.[29] Figure 7a shows the EPR spectra of SrY$_2$O$_4$:Gd^{3+} phosphors with different Gd^{3+} concentrations at RT. The EPR spectra exhibit signals w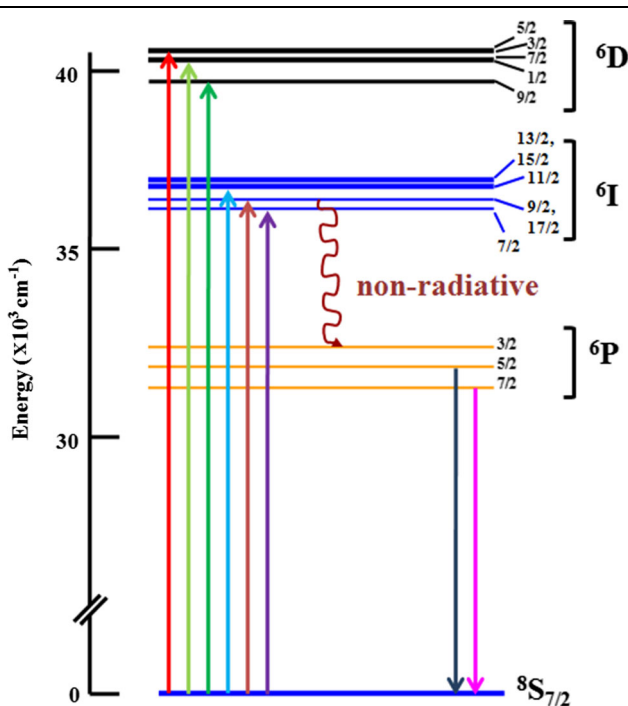ith effective g-factors at $g \approx 9.765$, 6.070, 4.770, 3.403, 2.773, 2.339, 2.067, 1.697, 1.589, 1.445, 1.386, 1.298, 1.252, 1.211, 1.169 and 1.104. Gd^{3+} ions experience a relatively weak crystalline field in these sites, and they are characterized by a coordination number higher than 6.[30] Effective g values for the low and high concentrations of Gd^{3+} ions are shown in Fig. 7b. EPR spectra and the g values obtained are independent of Gd^{3+} ion concentration as well as the phosphor composition. Because of the proximity in ionic radii of Gd^{3+} and Y^{3+} (0.938 and 0.900 Å, respectively), Gd^{3+} ions are most likely to substitute Y^{3+} sites where the valency of cations is similar.[31] The g values obtained are in agreement with the other reported values. The value $g = 6.07$, the characteristic element of the U spectrum, is ascribed to a cubic field.[32] The strong resonance at $g = 2.773$ is attributed the accumulation of Gd^{3+} ions, while the $g = 4.77$ line is a feature of strong crystal fields where the Gd^{3+} ion possesses a lower coordination number in orthorhombic symmetry sites.[32] Since Gd^{3+} ions are sited in a cubic system, we obtained a resonance line at $g = 2.067$. Even though several interpretations of the U spectrum are available, the EPR spectrum of Gd^{3+} ion is considered the most suitable for the depiction of the positions of RE ions in disordered surroundings.
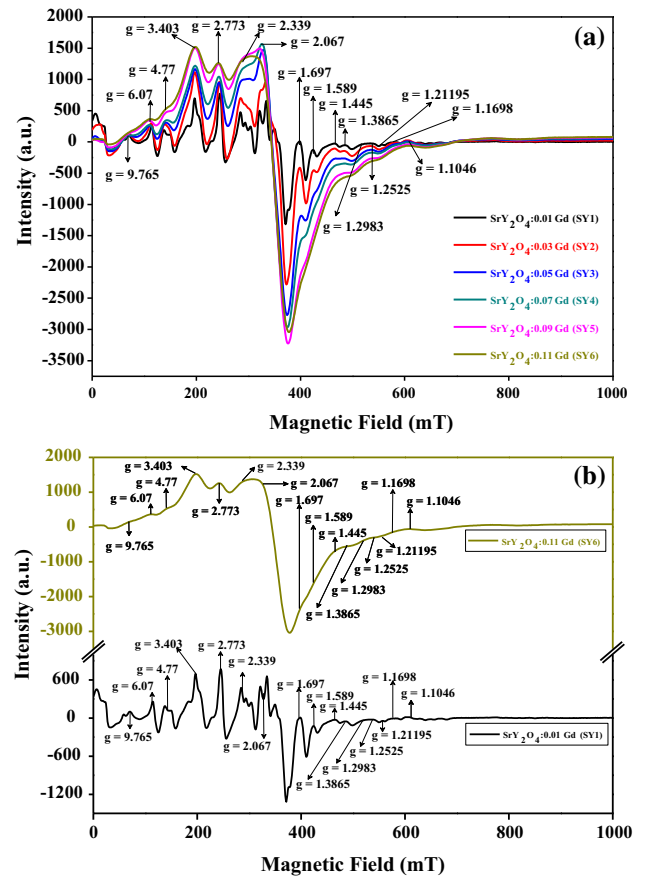

Fig. 7. (a) EPR spectra of SrY$_2$O$_4$:Gd^{3+} (SY1–SY6) phosphors and (b) effective g values marked for the low and high concentrations of dopant Gd^{3+} ions.

Fig. 6. The partial energy level diagram of Gd^{3+} ions.

Such analyses indicate the negligible presence of resonance at $g = 1.697$ which arises due to the Zeeman splitting of S-state ions.[28] These g values manifest the presence of Gd^{3+} ions in octahedral symmetric sites with some distortion. EPR and PL data presented here revealed the presence of Gd^{3+} in the SrY_2O_4 matrix.

CONCLUSIONS

This work reports the successful synthesis of Gd^{3+}-doped SrY_2O_4 phosphors via sol–gel method. The XRD reveal that the orthorhombic phase of SrY_2O_4 could be obtained by this method. The average crystallite size is observed to be in the range of 42.32–47.13 nm. The SEM micrographs demonstrate agglomeration of the particles due to sintering of samples at high temperature. Moreover, the particles possess inhomogeneous morphology with denser distribution of particles. The PL spectra reveal an intense excitation peak positioned at 275 nm indicating high absorption in the UV region. The emission spectra exhibit intense emission at 315 nm with the optimized Gd^{3+} ion concentration of 0.07 mol in the SrY_2O_4 host. The EPR spectra studied for $SrY_2O_4{:}Gd^{3+}$ phosphor exhibit resonance signals as reported earlier for Gd^{3+} ion-doped phosphors. An effective g value of $g \approx 6.07$ is observed, implying a coordination number greater than 6 and a cubic crystal symmetry. The other resonance signal for $g \approx 4.77$ noted indicates the occupancy of Gd^{3+} ions with an orthorhombic symmetry and lower coordination, and the signal at $g \approx 1$ indicates Zeeman splitting. The EPR resonance signals obtained for the prepared phosphors reveal different crystal field strengths and different sites occupied in the host lattice. PL spectra allow us to consider the prepared phosphor (SY4) as a potential candidate for application in phototherapy lamps.

ACKNOWLEDGMENTS

REFERENCES

1. J. Singh and J. Manam, *Ceram. Int.* 42, 18536 (2016).
2. R. Weia, J. Guo, K. Li, L. Yang, X. Tian, X. Li, F. Hu, and H. Guo, *J. Lumin.* 216, 116737 (2019).
3. R. Atkins and A.L. Diaz, *J. Lumin.* 128, 1463 (2008).
4. M. Taibi, E. Antic-Fidancev, J. Aride, M. Lemaitre-Blaise, and P. Porcher, *J. Phys. Condens. Matter* 5, 5201 (1993).
5. S.J. Park, C.H. Park, B.Y. Yu, H.S. Bae, C.H. Kim, and C.H. Pyun, *J. Electrochem. Soc.* 146, 3903 (1999).
6. C. Pyun, C. Park, S. Park, B. Yu, H. Bae, and C. Kim, *J. Rare Earths* 18, 14 (2000).
7. W.L. Xu, W.Y. Jia, I. Revira, K. Monge, and H.M. Liu, *J. Electrochem. Soc.* 148, H176 (2001).
8. K. Kurosaki, T. Tanaka, T. Maekawa, and S. Yamanaka, *J. Alloys Compd.* 395, 318 (2005).
9. R. Wei, Z. Zheng, Y. Shi, X. Peng, H. Wang, X. Tian, F. Hu, and H. Guo, *J. Alloys Compd.* 767, 403 (2018).
10. S. Xianliang, X. Mingming, T. Ying, F. Yao, P. Yong, and L. Xixian, *J. Rare Earth* 34, 458 (2016).
11. H. Jeong, N. Singh, M.S. Pathak, S. Watanabe, T.K.G. Rao, V. Dubey, and V. Singh, *Optik* 157, 1199 (2018).
12. A.B. Gawande, R.P. Sonekar, and S.K. Omanwar, *Mater. Res. Bull.* 60, 285 (2014).
13. S. Tamboli, B. Rajeswari, and S.J. Dhoble, *Luminescence* 31, 551 (2016).
14. W.C. Zheng, S.Y. Wu, D. Hui-Ning, and Z. Jain, *Spectrochim. Acta A* 58, 537 (2002).
15. K. Irmsher and M. Prokjesch, *Mater. Sci. Eng. B* 80, 168 (2001).
16. D.J. Newman and M.M. Ellis, *Phys. Lett.* 23, 46 (1966).
17. B. Bleaney, H.E.D. Scovil, and R.S. Trenam, *Proc. R. Soc. A* 223A, 15 (1954).
18. L. Zhou, J. Shi, and M. Gong, *J. Lumin.* 113, 285 (2005).
19. E. Pavitra, G.S.R. Raju, J.H. Oh, and J.S. Yu, *New J. Chem.* 38, 3413 (2014).
20. E. Pavitra, G.S.R. Raju, and J.S. Yu, *J. Alloys Compd.* 592, 157 (2014).
21. K. Binnemans, C. Gorller-Walrand, and J.L. Adam, *Chem. Phys. Lett.* 280, 333 (1997).
22. S.K. Gupta, R.M. Kadam, V. Natarajan, and S.V. Godbole, *Mater. Sci. Eng. B* 183, 6 (2014).
23. V. Singh, N. Singh, M.S. Pathak, S. Watanabe, T.K. Gundu Rao, P.K. Singh, and V. Dubey, *Optik* 157, 1391 (2017).
24. A.O. Chauhan, N.S. Bajaj, and S.K. Omanwar, *Bull. Mater. Sci.* 40, 1 (2017).
25. K.R. Ashwini, H.B. Premkumar, G.P. Darshan, H. Nagabhushana, S.C. Sharma, S.C. Prashantha, and H.P. Nagaswarupa, *Mater. Today Proc.* 4, 12168 (2017).
26. Z. Fu, S. Zhou, and S. Zhang, *J. Opt. Soc. Am. B* 23, 1852 (2006).
27. Y. Zhang, D. Geng, M. Shang, X. Zhang, X. Li, Z. Cheng, H. Lian, and J. Lin, *Dalton Trans.* 42, 4799 (2013).
28. K. Srinivasulu, I. Omkaram, H. Obeid, A.S. Kumar, and J.L. Rao, *J. Mol. Struct.* 1036, 63 (2013).
29. L.E. Iton and J. Turkevich, *J. Phys. Chem.* 81, 435 (1977).
30. R.C. Nicklin, J.K. Johnstone, R.G. Barnes, and D.R. Wilder, *J. Chem. Phys.* 59, 1652 (1973).
31. S. Simon, *Mod. Phys. Lett. B* 15, 375 (2001).
32. H.J.A. Koopmans, M.M. Peric, B. Niuzenhuijse, and P.J. Gallings, *Phys. Status Solidi (B)* 120, 745 (1983).

Participation of Women in Engineering & Technology Education and Employment

2 authors:

Addissie Melak
Delhi Technological University
5 PUBLICATIONS **5** CITATIONS

SEE PROFILE

Seema Singh
Delhi Technological University
47 PUBLICATIONS **61** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Agricultural Technology Adoption, Poverty and Income Inequality in Ethiopia View project

Women in Engineering View project

Participation of Women in Engineering & Technology Education and Employment

Addissie Melak, Seema Singh

Abstract: Engineering and Technology Education is a means to sustain and accelerate the overall development in a country and it has a direct effect on individuals' productivity and earnings as well. But the problem is there is gender disparity in distribution especially in developing countries among educational branches. In Ethiopia, despite significant improvements in the last couple of years, women are still under-represented in engineering and technology. Since the share of Ethiopian women in science and technology programs in higher education has been low, the governments apply a policy to stream 70% of all university entrants to be in the science and technology track indirectly to increase their entry into these fields. Even if this types of policy is applying, their participation in engineering and technology education and employment is very low proportion. The main purpose of this study is an attempt to assess trends and share of women in the field of engineering and technology regarding to their share of enrollment, employment and professional positions from the period 2000 -12 based on available secondary data collected from different sources. The collected quantitative data were analyzed by using descriptive data analysis techniques. Result from the data shows that there is low share of women in engineering and technology enrollment, employment and professional position status as compared to male in the country. Even if the share of enrollment of women in higher education is on improved status, their share of enrollment in engineering and technology filed is quite low that on average below 30 percent per year. In addition, this grate disparity and low participation of women in engineering and technology invites further investigation regarding to what is the reason behind.

Key words: Women, engineering and technology, enrollment and employment, Ethiopia

I. INTRODUCTION

Ethiopia is found in eastern part of Africa. The total population accounts 102.4 million with life expectancy of 65 years and GNI per capita is $1750 in 2016 [11], [29] and women accounts half of the population.Since education is a means to sustain and accelerate the overall development in a country and it has a direct effect on individuals' productivity and earnings as well [30], the Ethiopian government has adopted a new Education and Training Policy in 1994. The policy focuses on increasing access to educational opportunities with enhanced equity, quality and relevance stared in 1997/98 with the long-term goal of achieving universal primary education [31] and is governed by Ministry

of education. Ethiopia school enrolment in 2015/ 2016 at the age 7-18 is male at primary level 64.3% and at secondary level 5.5% as well as female at primary level is 64.2% and secondary level 5.8%. Tertiary education enrolment rate in 2014 of male is 10.94% and 5.28% female [32]. Total literacy rate in the same year was 64.1 %[12].

Women were totally excluded from education life for long time history of Ethiopia. There were strong social and cultural pressures for girls to marry before 14 years old. Virginity before marriage was highly valued and unmarried girl over 14 year was stigmatized. Girls are socialized to be shy and obedient and not to speak in front of adults, particularly in front of men. The focus is on her future role as respectful wife and good mother, hence they didn't went to school, even they went because of various problem such as harassment , violence [1] and their own natural phenomena of adolescence age , they dropout from school. This situation again leads to low employment participation in countries development. Hence to encourage women in education and employment participation, Ethiopian government policy has supported affirmative action for women since its beginning in 1994. Since the time women started to come out of domestic chores in Ethiopia, they have made significant contribution in all spheres of activities vital to the country development [27]-[28].

However, still, women are underrepresented in engineering and technology education and employment [3], [9], [5], [17], [18]. There are various factors for this under representation. Such as studies by [2], [8], [21], [14] [13], [25] found high school educational background, family education and lack of role models have an impact on decision to study engineering and technology. Studies [16], [19], [17], [24]; [6], [22], [23], [26]found that employment participation of women engineering and technology is low and is male dominated sector. However, there are few study about participation status of women in engineering and technology in Ethiopia. This indicates that there is research gap in the study area which need more research investigation. Thus, the objective of this study is to examine the trends and share of women in engineering and technology education and employment participation with the help of secondary data.
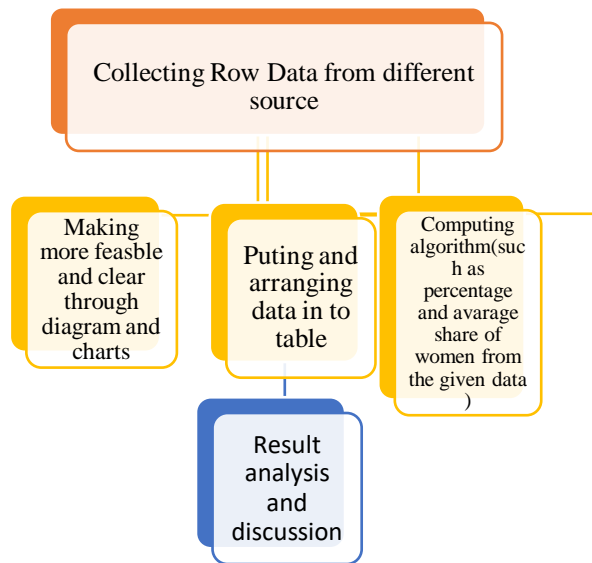
II. METHODOLOGY OF THE STUDY

In this study secondary data from Ethiopia Ministry of Education statistics annual abstract report and National Assessment [10], [28] were taken to analyze the trends and share of women participation in engineering and technology education and employment. Result analysis of this study has done through the following procedure.

The trends of share of female in preparatory school (grade 11 and 12) given in table-I, shows experience of improvement with fluctuation through time. The percentage of female students during 2011/12 has sharply increased and has reached 47% but when we observe the average percentage share of female 33.36% which show the there is some existence of gap.
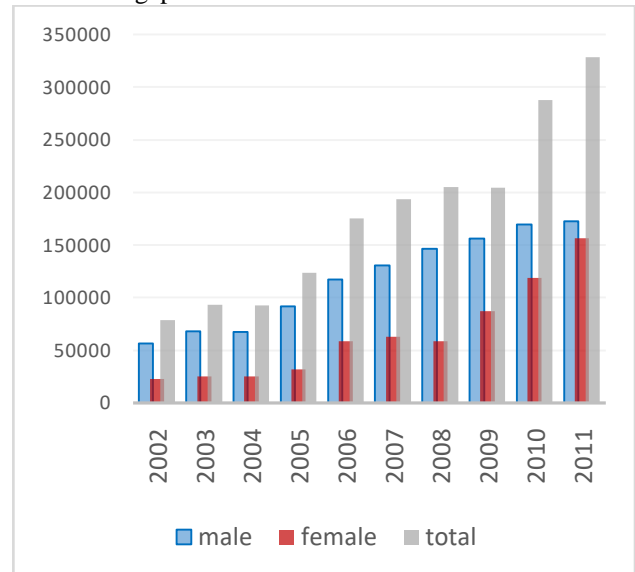
Fig. 1. Women students' enrollment in preparatory school program

Figure 1 clearly show, the share of female students in in preparatory school (grade 11 and 12) improved in the last ten years. This is due to some affirmative action taken for girls able to join in preparatory school.

From table-II, the trends of share of women in higher education is very low in the last 10 years which is an average of 24.98%. This is due to various factors social, cultural, economic and personal factors which is beyond the scope of this study which needs further investigation. However, there was a better improvement of share of women during the year 2001, 2008, and 2011 which is around 28%.

III. RESULTS AND ANALYSIS

Ethiopian Education system is structured by different stages which is kindergarten, alternative basic education, 1st cycle primary education (grade 1-4), 2nd cycle primary level from grade 5-8, 1st cycle secondary level grade 9 and 10, 2nd cycle secondary education level (preparatory school grade 11 and12), Technical and vocational education and training (TVET), Undergraduate degree, Master degree and Ph.D education level. Hence, in this study researchers going to investigate about participation of women in engineering and technology education and employment from preparatory school (grade 11 and 12) share of enrollment since this education level is the base for joining to higher education institutions.

Table-I: Trends of Women Enrollment in Preparatory School Program

Year	Enrollment			
	Boys	**Girls**	**Total**	**% girls**
1995E.C (2002/03)	56367	22274	78641	28.32%
1996E.C (2003/04)	68,330	25,065	93,395	26.83%
1997 E.C (2004/05)	67,413	25,070	92,483	27.1%
1998 E.C (2005/06)	91,889	31,794	123,683	25.7%
1999 E.C (2006/07)	117,000	58,219	175,219	33.23%
2000E.C (2007/08)	130,533	62,911	193,444	32.52%
2001E.C (2008/09)	146,547	58,713	205,260	28.6%
2002E.C (2009/10)	156194	86886	204,308	42.5%
2003E.C (2010/11	169,571	118,645	288,216	41.165 %
2004E.C (2011/12)	172,268	156,724	328,992	47.64%
Average				**33.36%**

Source： [28]

Table-II: Trends of women and Man Enrollment in Ethiopian Universities

Year	Enrolment			
	Man	**Women**	**Total**	**%Women**
1993 E.C (2000/01)	35932	9594	45526	21.07%
1994E.C (2001/02)	32279	13517	46796	28.88%
1995E.C (2002/03)	57321	18676	75997	24.57%
1996E.C (2003/04)	59352	18375	77727	23.64%
1997E.C (2004/05)	70388	21267	91655	23.2%
1998E.C (2005/06)	70388	21267	91655	23.2%
1999E.C (2006/07)	79465	25108	104573	24%
2000E.C (2007/08)	206336	64020	270356	23.67%

2001E.C (2008/09)	229489	91338	320827	28.46%
2002E.C (2009/10)	319327	115332	434659	26.5%
2003E.C (2010/11)	344137	123706	467843	26.44%
2004E.C (2011/12)	353163	138708	491871	28.2%
2005E.C (2012/13)	553848	166141	719989	23%

Figure 2 which is constructed [28] from table-II shows, the enrollment status of trends of women in in Ethiopian higher education which indicates there is high gender disparity in enrollment in this level of education.
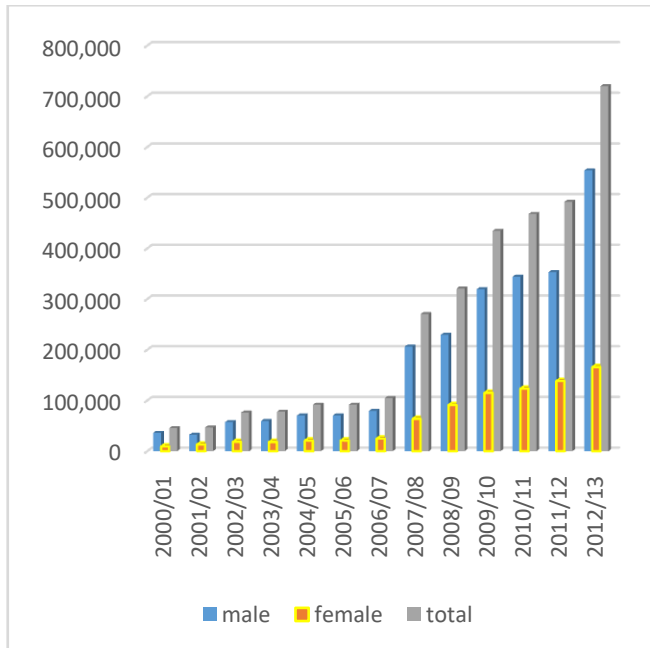
Fig. 2. Share of women' enrollment in Ethiopian Universities
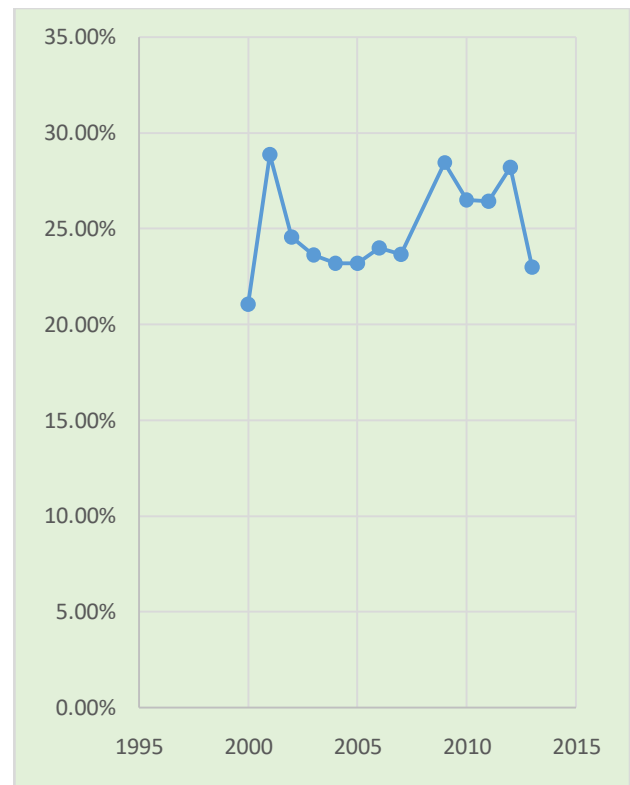
Fig. 3. Percentage of women enrollment in higher education from the total. Source: [28]

Table-III: Trends of Women in Higher Education Enrolment, and Graduation in Undergraduate and Post Graduate Program

Enrollment in Higher Education	Year							
	1997 E.C (2004/05)	1998 E.C (2005/06)	1999 E.C (2006/07)	2000 E.C (2007/08)	2001 E.C (2008/09)	2002 E.C (2009/10)	2003 2010/11	2004 2011/12
1.Under Graduate Degree								
1.1 Enrollment percentage of female students	138159 24.0%	173,901 24.8%	203,399 26.0%	263,001 24.1%	309,092 29.0%	420,387 27.0%	447,693 27.0%	494,110 28.2%
1.2 Graduates Percentage of female graduates	11,535 16.4%	25,335 16.2%	29,845 17.4%	47,979 20.7%	55,770 29.7%	66,999 23.4%	75,348 27.3%	78,144 25.3%
2.Post Graduate								
2.1 Enrolment Percentage of female students	3,604 9.2%	6,385 10.0%	7,057 10.0%	7,355 10.5%	10,125 11.3%	14,272 11.9%	20,150 13.8%	25,660 20.2%
2.2 Graduates Percentage of female graduates	1,126 9.0%	1,388 9.8%	2,671 9.4%	2,664 10.7%	3,257 12.3%	4,873 13.9%	6,250 14.4%	6,162 14.0%

Source: [28]

In figure 3, the percentage share of women enrolment from the total enrollment of higher education had shown increasing trend and falling in 2002, which constantly growing till 2008.

Again in table-III, trends of share of women in higher education enrolment, and graduation in undergraduate and post graduate program. From this the share of enrollment in undergraduate program is not that much improved but the share of graduation in under graduate program shows some improvement trend. Their share of enrollment in post graduate level is increased from 9.2 % to 20.2% in the last 8 consecutive years.

Table-IV: Trends of Share of Women Enrollment and Graduation in Undergraduate Program

Year	Under graduate Enrollment		Undergraduate Graduates	
	Female	*Total*	*Female*	*Total*
2004	33158	138,159	1892	11,535
2005	43127	173,901	4104	25,335
2006	52884	203,399	5193	29,845
2007	63383	263,001	9932	47,979
2008	89637	309,092	16563	55,770
2009	113505	420,387	15678	66,999
2010	120877	447,693	20570	75,348
2011	139339	494,110	19770	78,144
Mean	81989		11712	

Source: [28]

Table-IV shows, the trends of share of women student's enrollment and graduation in undergraduate program in terms of number from the total students which show increment but still low as compared to male students. Figure 4 also shows, share of women in post graduate program that indicates very low participation position.
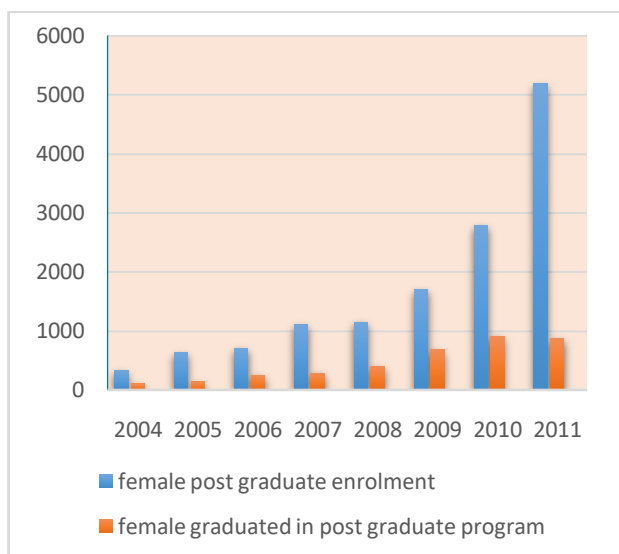
Fig. 4.The share of women students in post graduate program.

Table-V: Women Enrolled in Regular Undergraduate Engineering and Technology Education by Universities

Universities	Gender	Year				
		2007	2008	2009	2010	2011
Adama	Male	1443	Na	4876	5736	5956
	Female	379		865	776	969
	Total	1822		5741	6512	6925
Addis Ababa	Male	1205	3164	2753	2969	7836
	Female	408	1097	955	910	2799
	Total	1613	4261	3708	3879	10635
Adigrat	Male	Na	Na	Na	Na	173
	Female					207
	Total					380
Aksum	Male	38	73	226	857	1794
	Female	8	13	55	362	1027
	Total	46	86	281	1219	2821
Ambo	Male	Na	Na	281	785	785
	Female			54	148	148
	Total			335	933	933
Arba Minch	Male	1789	752	2931	5496	6199
	Female	346	276	711	1092	1362
	Total	2135	1028	3642	6588	7561
Assosa	Male	Na	Na	Na	Na	169
	Female					205
	Total					374
Bahir Dar	Male	2164	Na	2768	3548	4469
	Female	452		749	1087	1358
	Total	2616		3517	4635	5827
Bulle Hora	Male	Na	Na	Na	Na	58
	Female					26
	Total					84
Debre Birhan	Male	Na	122	44	917	1503
	Female		45	6	411	785
	Total		167	50	1328	2288
Debre Markos	Male	Na	Na	368	Na	2077
	Female			54		6112
	Total			420		688

University		Y1	Y2	Y3	Y4	Y5
Debre Tabor	Male	Na	Na	Na	Na	99
	Female					64
	Total					163
Dilla	Male	Na	Na	Na	Na	574
	Female					329
	Total					903
Meda Walabu	Male	Na	Na	38	535	1048
	Female			15	170	3711
	Total			53	7 05	419
Mekele	Male	1276	Na	1753	3497	36757
	Female	268		293	825	904
	Total	1544		2046	4322	465
Mettu	Male	Na	Na	Na	Na	85
	Female					41
	Total					126
Mizan Tepi	Male	Na	Na	482	860	1486
	Female			72	287	6992
	Total			554	1147	185
Sumera	Male	Na	Na	29	Na	109
	Female			8		184
	Total			37		293
Wachemo	Male	Na	Na	Na	Na	114
	Female					64
	Total					178
Wollega	Male	125	Na	332	1093	2257
	Female	20		80	233	7292
	Total	145		472	1326	986
Wollo	Male	Na	139	1127	1572	1477
	Female		25	158	353	2981
	Total		164	1285	1925	775
Defense	Male	Na	Na	531	531	531
	Female			10	10	10
	Total			541	541	541
Dire Dawa	Male	Na	158	806	1413	2313
	Female		37	558	686	10323
	Total		195	1364	2099	345
Ethiopian civil service	Male	328	Na	562	484	474
	Female	39		57	44	62
	Total	367		619	528	536

University		Y1	Y2	Y3	Y4	Y5
Gonder	Male	Na	Na	Na	1054	2014
	Female				459	972
	Total				1513	2986
Haramaya	Male	531	1254	1438	3506	4692
	Female	132	148	205	651	861
	Total	663	1402	1643	4157	5553
Hawassa	Male	838	1899	2214	3833	5070
	Female	175	368	448	673	1161
	Total	1013	2267	2662	4506	6231
Jijiga	Male	Na	Na	Na	851	1406
	Female				186	532
	Total				1,037	1938
Jimma	Male	935	Na	1878	3727	6097
	Female	167		294	477	860
	Total	1102		2171	4204	6957
Wolaita Sodo	Male	Na	Na	Na	694	1665
	Female				192	625
	Total				886	2290
Woldia	Male	Na	Na	Na	Na	79
	Female					84
	Total					163
Wolkte	Male	Na	Na	Na	Na	209
	Female					258
	Total					467

Source: [28], Na = data not available

Table-v show, the share of female student's enrolment in regular undergraduate program in engineering and technology education in 31 public universities from the year 2007-2011 which shows there is grate disparity and under representation of female engineering education participation even if there is some improvement. The pattern of engineering enrolment is various in each universities. Since some of the universities are new established recently there is no data give for them. Especially third generation universities started education 2003/2004 data is not available for them. Hence, Addis Ababa University, Adama University, Aksum university, Mekele University, Haramaya, Hawassa, Jimma, Arba Minch, Bahir Dar, Wollega and Defiance Universes are the old universities and the most Engineering and technology education provider. The summation of Engineering student in this universities are summarized in the following table 6 and figure 4 including enrolment in non- government university and collages.
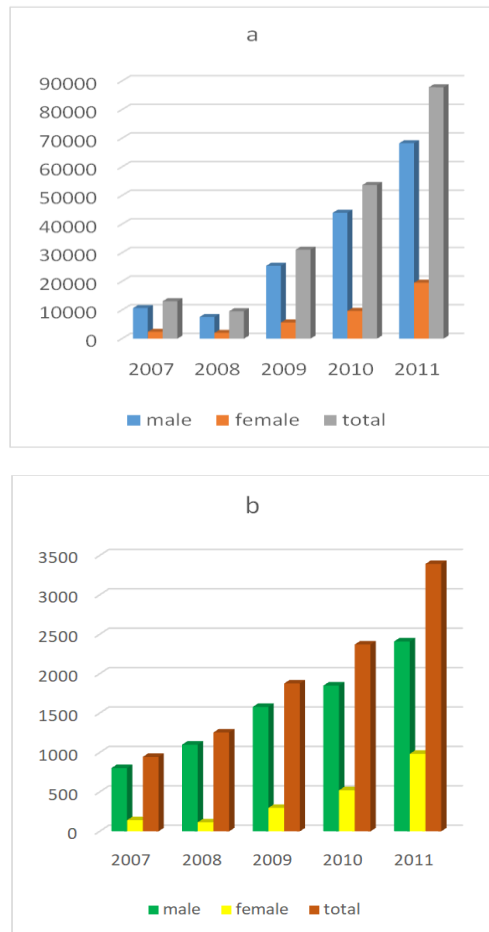
Fig. 5 a and b. trends and share of women regular undergraduate enrolment in engineering and technology education at government and non-government universities respectively.

In figure 5 a and b, on the bases of data [28], the share of women enrolment in engineering and technology education participation is increased from 18% during 2007 to 22.23% during 2011 in government universities and from 15% to 28% in non-government universities and colleges. However, there is gender disparity and women are underrepresented in engineering and technology education.
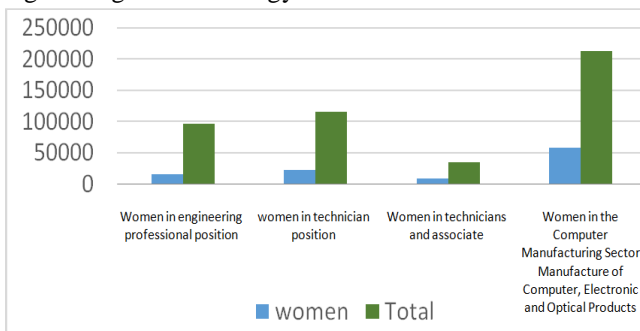

Fig. 6. Share of women from the total employments of engineering and technology fields

Figure 6 is constructed on the bases of data from national assessment [10] shows, the share of women in engineering and technology employment such as, technician position, engineering science, technicians and associate technician, computer manufacturing and others areas according to available data at national assessment 2013. The data shows that the share of women employees in this profession is very low but they are better in manufacturing employment.

IV. CONCLUSION

The finding of this study indicate that, women are underrepresented and there is grate disparities between male and female participation regarding to engineering and technology enrollment, graduation, employment and profession position in the field which is in line with the literature reviewed. This implies that problem needs attention so that give notice for the problem as well as creating awareness about the importance of engineering education starting from lower class through encouraging girls in science subjects and also take experience from other countries. In addition creating suitable environment for engineers at university level is necessary and there should be future research investigation in the field to identify the reason behind underrepresentation of women in engineering education and employment which can suggest sustainable solution for this problem.

REFERENCES

1. Abebayehu Tora (2013) Assessment of Sexual Violence against Female Students in Wolaita Sodo University, Southern Ethiopia, Journal of Intertamporal violence vol.8, issue 11
2. J. Gill et al. (2008) I still want to be an engineer! Women, education and the engineering profession, European Journal of Education, 33:4, 391-402, DOI: 10.1080/03043790802253459
3. Addissie .M and Singh. S (2018) ,Experience of women in engineering education: literature review, Asia Pacific Journal of Research, *Vol: I. Issue LVV, pp174-178*
4. Dimitriadi (2013) Young women in science and technology: the importance of choice, Journal of Innovation and Entrepreneurship vol. 2:No.5pp1-14:http://www.innovation-entrepreneurship.com/content/2/1/5
5. Robsan Margo Egne (2014) Gender Equality in Public Higher Education Institutions of Ethiopia: The Case of Science, Technology, Engineering, and Mathematics, *Discourse and Communication for Sustainable Education, vol. 5, pp. 3ñ21, 2014:* DOI: 10.2478/dcse-2014 0001
6. Tamiru Jote (2017), Exploring Employment Status and Education–Job Match among Engineering Graduates in Ethiopia: Policy Implications
7. Samira I. Islam (2017) Arab Women in Science, Technology, Engineering and Mathematics Fields: The Way Forward, World Journal of Education Vol. 7, No. 6,pp 12-20, doi:10.5430/wje.v7n6p12, ISSN 1925-0746 E-ISSN 1925-0754
8. Choudhury (2015) explaining gender discrimination in the employment and earnings of engineering graduate's in India. Journal of educational planning and administration, vol.XXIX, No.3, pp 225-246
9. Addissie .M and Singh. S (2018), factors affecting academic performance of women in engineering education and their experience of participation in employment: literature review, international Journal of Research, Volume 05 Issue 01 pp 3733-3739
10. Helina Beyene (2015) final report national assessment: Ethiopia gender equality and the knowledge society, women in global science and technology,
11. World bank country study(2005), education in Ethiopia strengthening the foundation for sustainable progress, ISBN-10: 0-8213-6226-7 ISBN-13: 978-0-8213-6226-6, ISSN: 0253-2123 DOI: 10.1596/978-0-8213-6226-6
12. Ethiopia socioeconomic survey(2015/2016) report by A Report by the Central Statistical Agency of Ethiopia in Collaboration with the National Bank of Ethiopia and the World Bank, February 2017
13. A. Powell et al. (2011) A poisoned chalice? Why UK women engineering and technology students may receive more 'help' than their male peers, Gender and Education, 23:5, 585- 599, DOI: 10.1080/09540253.2010.527826
14. Alice E. Smith & Berna Dengiz (2010) Women in engineering in Turkey – a large scale quantitative and qualitative examination, European Journal of Engineering Education, 35:1, 45-57, DOI: 10.1080/03043790903406345

15. Catherine Hill et al. (2010) Women in Science, Technology, and Engineering, and Mathematics, page xiv, Library of Congress Control Number: 2010901076 ISBN: 978-1-879922-40- 2 Published by AAUW Washington, DC 20036 Web: www.aauw.org

16. Carmen García Villa & Elsa M. González y González (2014) Women students in engineering in Mexico: exploring responses to gender differences, International Journal of Qualitative Studies in Education, 27:8, 1044-1061, DOI: 10.1080/09518398.2014.924636

17. C. Riegle-Crumb, C. Moore (2013) Examining Gender Inequality In A High School Engineering Course, American Journal of Engineering Education Special Edition Volume 4, Number 1 pp55- 66, http://creativecommons.org/licenses/by/3.0/

18. Elsa Q. Villa et al.(2016), Engineering Education through the Latina Lens, Engineering Education through the Latina Lens, Journal of Education and Learning; Vol. 5, No. 4; ISSN 1927 5250 E- ISSN 1927-5269, doi:10.5539/jel.v5n4p113

19. Ghiasi G, Larivière V, Sugimoto CR (2015) On the Compliance of Women Engineers with a Gendered Scientific System. PLoS ONE vol.10, No.12, pp1-19: e0145931. doi:10.1371/journal.pone.0145931

20. Kacey D. Beddoes (2012) Feminist Scholarship in Engineering Education: Challenges and Tensions, Engineering Studies, 4:3, 205-232, DOI:10.1080/19378629.2012.693932

21. Maria Udeân (2002) the Impact of Women on Engineering: A Study of Female Engineering Students' Thesis Topics, International Journal of Engineering Education. Vol. 18, No. 4, pp. 458, 2002 0949-149X/91 $3.00+0.00 Printed in Great Britain at TEMPUS Publications

22. Masako Hosaka (2014) Women's experiences in the engineering laboratory in Japan, European Journal of Engineering Education, 39:4, 424-431, DOI: 10.1080/03043797.2014.883363

23. S. Barnard, T. Hassan, B. Bagilhole & A. Dainty (2012) 'They're not girly girls': an exploration of quantitative and qualitative data on engineering and gender in higher education, European Journal of Engineering Education, vol.37:No.2, pp193-204, DOI: 10.1080/03043797.2012.661702

24. Singh. S, S. Fenton (2014) Women Engineers: A Comparative Study between India and Australia International Journal of Advancements in Research & Technology, Volume 3, Issue 7, pp108-121, ISSN 2278-7763

25. Jean Bossart, Neelam Bharti (2017) Women in Engineering: Insight into Why Some Engineering Departments Have More Success in Recruiting and Graduating Women, American Journal of Engineering Education Volume 8, Number 2 pp127-140http://creativecommons.org/licenses/by/3.0/

26. Singh.S (2013) Women in Engineering Education in India http://www.youblisher.com/p/109258-Women-in-Engineering-Education-in-India/

27. https://en.wikipedia.org/wiki/Education_in_Ethiopia)

28. Ethiopian ministry of education educational annual abstract 2000/01-2012, Addis Ababa Ethiopia.

29. World Development Indicators database (2018).

30. Development Indicators of Amhara Region (2011/12), Bureau of Finance and Economic Development, pp80

31. MOFED (2002) ministry of finance and Economic Development, Addis Ababa, Ethiopia pp.

32. UNISCO (2018), Institutes of Statistics,http://uis.unesco.org/country/ET

AUTHORS PROFILE

Addissie Melak[a, b], a Department of Humanities Delhi Technological University, Delhi, India,[b] Department of Economics, Debre Tabor University, Ethiopia.
Ph.D Research Scholar in Economics, Department of Humanities Delhi Technological University, Delhi, India, M.A in Economics at Punjabi University, Patiala, India, Bachelor Art Degree in Economics at Debre Markos University, Ethiopia, , Working staff in Department of Economics at Debre Tabor University, Ethiopia. Four international conferences presentation Participation. Four paper published in the Journals. addmelak24@gmail.com

Professor Seema Singh[a],a Department of Humanities Delhi Technological University, Delhi, India,**INWES Board Member (South Asia) & Vice President (Education & Research) International Network of Women Engineers and Scientists (INWES)**
Hon. Joint Secretary Indian Society of Labour Economics

Hon. Vice President, Women in Science & Engineering &
Hon. Vice President -I, University Women Association of Delhi.
X – HOD of Department of Humanities. More than 15 international conference participation, more than 40 articles published, wrote four book, advisor of more than 10 students…..seemahumanitiesdtu@gmail.com

Path Tracing and Object Avoidance Algorithm for Robotic Manipulators incorporating Constrained Filters

Conference Paper · December 2019

2 authors:

Vipul Garg
Delhi Technological University
2 PUBLICATIONS   **0** CITATIONS

SEE PROFILE

Rastogi Vikas
Delhi Technological University
92 PUBLICATIONS   **303** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Right now, we are working in area of Heavy Vehicle Modeling and control, railway dynamics and ride comfort and turbine blade vibrations View project

Project   Right now, we are working in area of turbine blade vibrations View project

Path Tracing and Object Avoidance Algorithm for Robotic Manipulators incorporating Constrained Filters

Vipul Garg, Vikas Rastogi

Department of Mechanical Engineering
Delhi Technological University
Delhi, India
garg.vipul7@gmail.com, rastogivikas@gmail.com

Abstract— **This paper presents an algorithm for object avoidance while sustaining the trajectory of the end effector of the robotic manipulators. Multiple filters are developed that are based on classified constraints and applied on configuration space to obtain a collision-free path. The recursive and iterative nature of the algorithm makes it possible to acquire results under the permissible zone of error. The identification of desired configurations from millions of choices is achieved to obtain smooth and least deviated movements. MATLAB software is used to design, implement, visualize and test the proposed model.**

Keywords— Robotic manipulator, Configuration space, Object avoidance, Trajectory, Constrained -filters.

I. INTRODUCTION

During the last decade, manipulators of diverse classes have appeared as an imperative part of the automated industries. The flexible and configurable nature of most of the manipulators makes it feasible to stretch their work domain. The very prevalent objective of them is to move the tool located on the manipulator to some goal in 3D (Three Dimensional) space. The task becomes susceptible to failure when it is implemented in a dynamic environment. In such environments, object avoidance and real-time path planning become the fundamental task of the manipulator.

In some applications, it becomes crucial to trace the desired path [1] to reach the goal such as autonomous grout filling of cracks on the dam. The imperialist Competitive Algorithm was used in [2] to solve the path-planning problem. The generation of C-Space (Configuration Space) is very computationally expensive because of the exponential growth in the possible arrangements of serial chain manipulators. Arjang Hourtash and Mahmoud Tarokh who presented their work on path planning decomposition in [3] have solved this issue. The division of the serial chain into sub-chains had been done to lessen the planning time. A* algorithm [4] is used to find the shortest path between the pick and place locations of objects. Most of the approaches were converged to reach the goal through any path between the initial and final state of the manipulator. To solve the trajectory problem, different algorithms were discussed and tested on a robotic tracking control testbed [5]. A nonlinear control design has been proposed in [6] to incorporate the dynamics of the manipulator while performing the tasks which require a trajectory control.

Non-linear behavior of PD (Proportional Derivative) controllers have been proposed in [7]. A robust and novel approach based on multiple constraint filters is proposed in this paper to reduce the complexity and incorporate the general behavior of robotic manipulators. Multiple constraints can be added according to the environment, manipulator's task and the manufacturing constraints.

The organization of the paper is as follows. Section 2 discusses the criteria of the description of rigid bodies over 3-dimensional space. The proposed methodology is presented in detail in Section 3. Illustration of approach is further divided into 4 subsections according to their functionalities. Implementation of the algorithm is discussed in Section 4 and the results of the simulations are explained in Section 5. All simulations have been performed on MATLAB and supported by comparing the desired and actual trajectories. Finally, a conclusion is made in Section 5, which also addresses future work.

II. REPRESENTATION OF RIGID BODIES

A 6-Revolute open-chain anthropomorphic manipulator with six DoF (degrees of freedom) has been scrutinized in this work. Three DoF corresponds to the position of the manipulator in 3D space and others 3 DoF control its orientation. Its mechanism is analogous to the PUMA 560 [8] (refer Fig. 1) in such a way that the first three links (L1, L2, and L3) control the position of the wrist-point, i.e., the intersection of the last three joint axes. The rotary motion of manipulator's links can be achieved using different actuators. These actuators can be purely revolute joint such
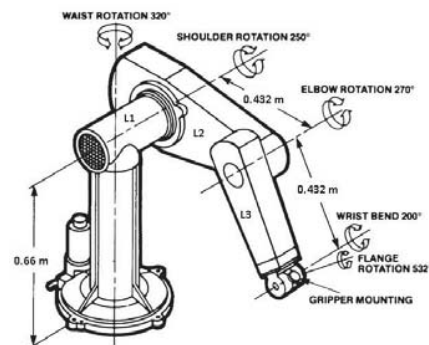

Fig.. 1.   PUMA 560 dimensions [8]

as servo and stepper motors or they can be a part of a complex system. Such a complex mechanism can be found in Yasukawa Motoman L-3, which makes use of two linear actuators coupled to links 2 and 3 with four-bar linkages [9]. For the simplicity of the manipulator, the links are assumed to be actuated through servo motors with a resolution of 1°. The main challenge in the interfacing of servos to industrial manipulators is their high cost, as compared to non-servo motors, for the required load capacity and joint limits.

Considering the complexity of designs of manipulators and objects, components of the environment, modeling of rigid bodies can be achieved either by using superquadrics [10] or considering them as a spherical shell [11]. These spherical shells are the point clouds with a constant distance from their respective geometric primitives. These primitives could be points, line segments or rectangles. A point cloud is generated by forming a sphere of optimum radius around primitive geometry and superimposed. Objects will be described using these point clouds, which fits into it perfectly and with minimum unoccupied space i.e. concave shape. The corresponding geometric shapes of point clouds for points, line segments, and rectangles are the sphere, cylisphere, and box with rounded edges respectively. These simple shapes reduce the computation cost while determining the characteristics of a collision. A simple geometry shape is selected based on the shape of the object to be avoided. Similarly, the individual links of the manipulator are considered as cylinders and superimposed to construct the simplified structure of the manipulator.

MATLAB software has been used to simulate the movement of manipulator because of its versatility in representing objects as a surface or a mesh. Also, the compatibility of the ROS (Robot Operating System) framework in MATLAB can be further utilized to simulate the manipulator in more dedicated simulators like RViz and Gazebo. The primary state, goal state, and the path traced are plotted in 3D space in MATLAB to verify the algorithm. Object-Oriented Programming is employed to store and handle data more efficiently. It reduces the complexity of the model and makes it adaptive to all small constraints.

III. METHODOLOGY

In this paper, a novel approach has been suggested to avoid the objects and disturbances of surroundings while sustaining the desired trajectory. Trajectory represents the path that is required to follow by the end effector. This approach is divided into 4 main steps. Step 1 deals with the generation of the C-Space and Workspace by applying the initial parameters and mechanical constraints. Collision checking and avoidance filters are introduced in step 2. Objects are tested for their interference with the manipulator. Step 3 and step 4 further put more constraints on the collision-free states to obtain a smooth and least deviated path.

A. Workspace and C-Space

A trajectory can be represented through a set of points or an equivalent polynomial equation in its task space. Task space [12] can be defined as a space in which a robot's task is expressed naturally. It should be a subset of the workspace otherwise it will not be possible to trace the trajectory.

Table 1.   Denavit-Hartenberg Parameters for 6 DoF manipulator

i	α_{i-1} (degrees)	a_{i-1}	d_i	θ_i (degrees)
1	0°	0	0	30°
2	-90°	0	5	30°
3	0°	5	-2	30°
4	-90°	5	5	30°
5	90°	0	0	30°
6	-90°	0	0	30°

To find out the possible configurations of a manipulator for a specific point in its workspace, the most popular approach is inverse kinematics. It uses algebraic and numerical methods to solve the configuration problem but on the other side, predetermined C-space solves this problem in a much more efficient way.

Knowledge of only D-H (Denavit-Hartenberg) parameters is not sufficient to solve the non-linear complex equations of Inverse Kinematics. This C-space can also be generated with the help of the forward kinematics by considering the possible states of joints. With a span of 180 degrees for each servo motor, the total number of possible configurations are $span^{DOF}$ i.e 180^6 which requires a lot of computational memory to store but only 180^3 configurations are responsible for the position of the wrist. A vector in (2) representing the position of the wrist is developed by using the D-H notation [13]. This notation uses 4 parameters to describe a robot kinematically by assigning the values for each link. The values of parameters for the 6 R robot used in this work are provided in Table 1. All the parameters except the link twist are given integer values to simplify the matrix calculations. The link length and link offset are left as dimensionless numbers. A 4x4 square link-transformation matrix as in (1) defines the frame {i} in relative to frame {i-1} and acts as a pillar for the kinematics equations.

$$T_i^{i-1} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 & a_{i-1} \\ \sin\theta_i\cos\alpha_{i-1} & \cos\theta_i\cos\alpha_{i-1} & -\sin\alpha_{i-1} & -d_i\sin\alpha_{i-1} \\ \sin\theta_i\sin\alpha_{i-1} & \cos\theta_i\sin\alpha_{i-1} & \cos\alpha_{i-1} & d_i\cos\alpha_{i-1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

In the above equation, T_i^{i-1} is the transformation matrix for two successive links, θ_i is the angle of ith joint, α_{i-1} is the twist of ith link with respect to $(i-1)^{th}$ link, a_{i-1} is the length of ith link, and d_i is the offset between links i and $i-1$.

$$P = \begin{bmatrix} \cos\theta_1\left(5\cos\theta_2 + \cos\theta_{2+3} - 5\sin\theta_{2+3}\right) - 5\sin\theta_1 \\ \cos\theta_1\left(5\cos\theta_2 + \cos\theta_{2+3} - 5\sin\theta_{2+3}\right) - 5\cos\theta_1 \\ -5\cos\theta_2 - \cos\theta_{2+3} - 5\sin\theta_{2+3} \end{bmatrix} \quad (2)$$

The position vector P in (2) is solved for the combination of configurations to obtain the Cartesian space occupied by the wrist-point of the robot. The C-space obtained from above perimeters and constraints is passed through several filters to optimize the path traced.

B. Collision Filter

Object detection [14] and localization [15-17] are the primary tasks prior to the collision check. Localization of objects and mapping them to the simulated environment is not required to discuss at this point. All the generated configurations are iterated to measure the minimum

distance from the object's primitive. The minimum distance is the perpendicular distance between any two geometric primitives. The configurations, which failed to satisfy the condition (3) will be excluded from the original C-space.

$$d(i,o) \geq r(i) + r(o) + th \qquad (3)$$

Where, $d(i,o)$ is the perpendicular distance between the ith link and the object o, $r(i)$ is the radius of the ith link, and $r(o)$ is the radius of object and th is the minimum required separation between the object and the manipulator.

This geometric image is simplistic and less computationally expensive in comparison to superquadratics representation but ends in the exclusion of additional states of the manipulator that are not in a collision. To circumvent this, another image of the object can also be generated by creating a mesh of irregular shape in 3D space and the constraint is modified to (4).

$$\min\left(d\left(i, mesh(o)\right)\right) \geq r(i) + th \qquad (4)$$

In the above equation, $mesh(o)$ is the point cloud of object and $d(i, mesh(o))$ is the distance of ith link from the point cloud. After passing from the collision filter, the remaining configurations are collision-free and can be processed further.

C. Nearest State Filter

The desired trajectory is presented as a set of points. The separation between consecutive points plays an important role in deciding the planning time of the algorithm and the error between the theoretical and calculated trajectories. The former is negatively correlated and the latter is positively correlated with the separation distance. The greater the separation, the lesser will be the total number of points, planning time and more will be the error in tracing the trajectory. An optimum value of separation can be decided on the basis of the computational speed and the toleration zone for error. Fig. 2 shows the relation of the planning time and the number of coordinates i.e. points in 3D space used to represent the trajectory line. It can be inferred from the figure that both of the parameters are directly related to each other. Once the set of points is generated, each point is tested against the collision-free states to obtain the configurations that take the wrist to the path.
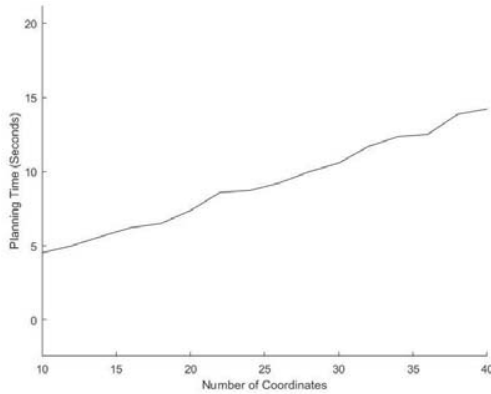

Fig.. 2.   Variation of planning time with the number of points on the trajectory line.

The configurations responsible for crossing the toleration zone of error will be excluded. These are further sorted to determine the least error state. The extracted set of configurations is finally passed through a smoothing filter.

D. Smoothing Filter

The continuous motion of actuators is very important during the action of the manipulator. This continuity enhances the life cycle of an actuator and diminishes the impulses, which generally occurs during jerks in motion. continuous motion can be obtained by collecting the consecutive configurations of a manipulator, which converges towards the task space. Initially, each state is given a common identity. This identity is obtained from eq. (5) by calculating the maximum movement of each actuator to reach the goal state from the initial state.

$$IN = \max([\theta_f] - [\theta_i]) \qquad (5)$$

In the equation, θ_f is the joint vector of goal state and θ_i is the joint vector of the initial state. This number is characterized by an integer value which lies between the joint limits. It can either be positive or negative depending upon the reference state i.e. initial state. IN (Identity Number) of the initial state is unique with a value of zero.

A subset of configuration space is determined with IN ranging from IN of initial state and goal state. Joint states with consecutive IN are extracted, which satisfies the constraint (6) in which θ_{IN} is the joint vector with IN value. The joint vector is a column vector of dimension DoFx1 consisting of the joint states.

$$\min([\theta_{IN+1}] - [\theta_{IN}]) \leq 1 \qquad (6)$$
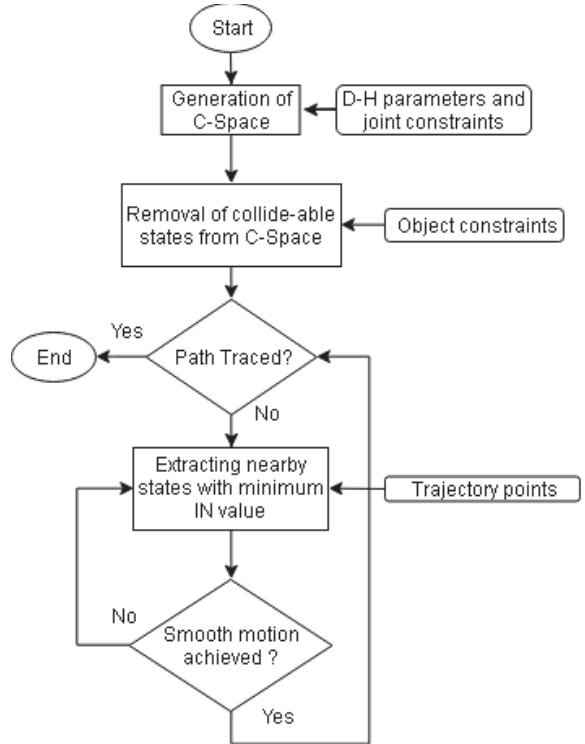

Fig.. 3.   Flow chart of the algorithm

E. *Implementation oF Approach*

The generation of configuration space is achieved during preplanning. Memory required to store the dataset exponentially increases with the joint limits. 4096000 configurations have been obtained for a joint limit [-90°, +90°].

The algorithm is performed iteratively and recursively on this dataset as shown in the flow chart presented in Fig. 3. Fig. 4a shows the desired trajectory as a set of points on a line. A spherical object of unit radius is placed between the desired path (refer Fig. 4b). Initially, Goal state is decided in such a way that it is collision-free, nearest to the first point on the trajectory and takes less time to reach i.e., the difference of IN of goal state with IN of the initial state is minimized. The error between the goal state and the trajectory point is minimized by using the nearest state filter. If the error crosses the tolerance zone, the new goal state is searched with the second least IN value. The process continues until a desirable goal state is found. Initial and primary goal state along with the collision-free C-space is received by the smoothing filter. Joint states which provide continuous motion between initial and goal state are obtained by enforcing the continuity constraint on those states with IN values ranging from $[IN_i, IN_f]$. If any type of discontinuity is detected then it returns back to the nearest filter to pass another goal state.

Once the desired joint movements are achieved, the initial state and goal state are updated. The previous goal state is updated to the initial state of the next iteration and goal state is updated to the next point of the trajectory. Iterations are performed until all points of trajectory are covered.

F. *Results*

The ideal solution is plotted to achieve the desired path (refer Fig. 4b) irrespective of environmental interference. The links of the manipulator are considered to be cylinders with an assumption that end effector doesn't play any role in determining its trajectory. 16% of the path is obstructed because of the object, it required to be avoided without disrupting the trajectory of the end effector.

Multiple filters based iterative algorithm is implemented to find a least deviated, collision-free, and smooth path. Fig. 4c-4f visualize the motion of manipulator for four progressive quarters of the time period. The period is defined by the time required by the end effector to meet the entire trajectory. It is controlled by the characteristics of the actuators such as the joint velocity, acceleration, and resolution. The object is placed near to the reachable workspace of the manipulator and so no collision-free joint configurations are left for the manipulator to overlap the desired position in the obstructed region and hence, the manipulator started to deviate from its mean path and looked for other nearest collision-free configurations at the end of the first quarter (refer Fig. 4c). The other reason for such deviation is the lower resolution of actuators which is negatively correlated with it. High resolution and a wide

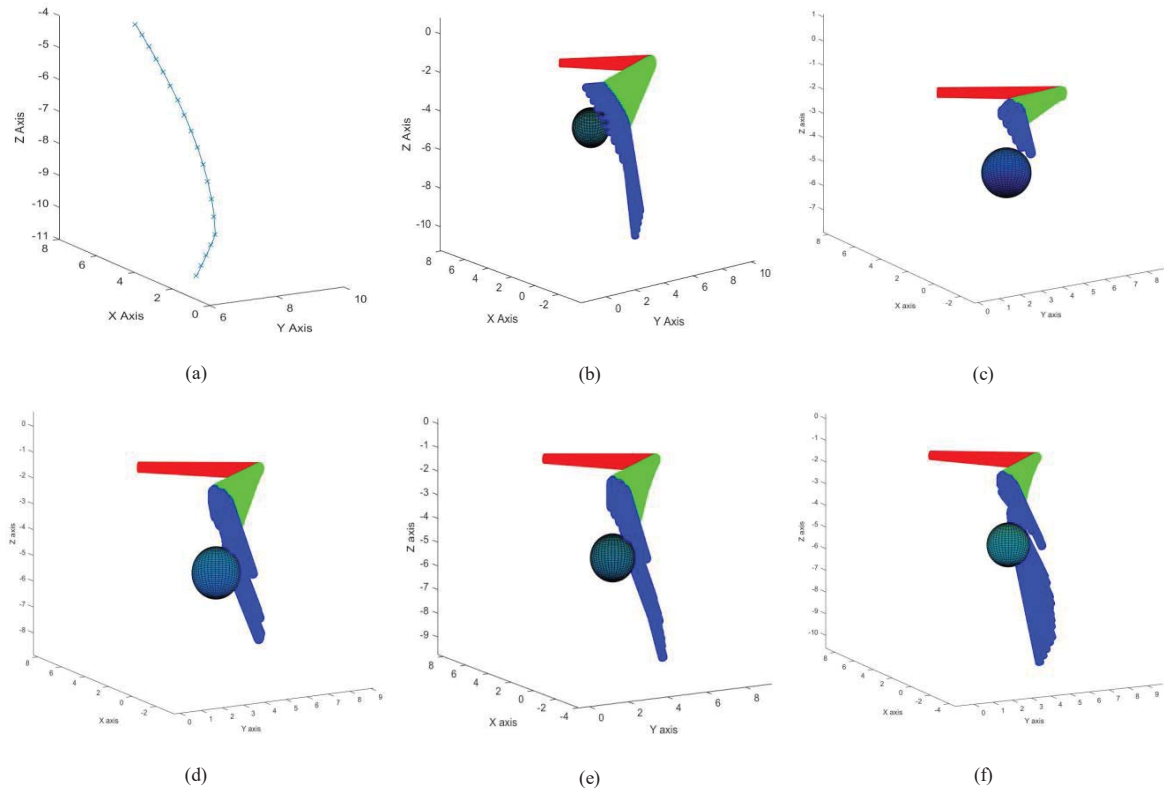| (a) | (b) | (c) |
| (d) | (e) | (f) |

Fig.. 4.   (a) The desired trajectory of the end effector of the manipulator. (b) Manipulator under collision with a spherical object. (c) The path traced during 1/4th of the period. (d) The path traced during 1/2nd of the period. (e) The path traced during 3/4th of the period. (f) The collision-free path traced at the end period.
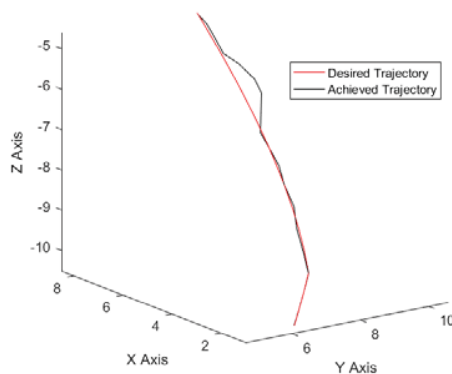
121

Fig.. 5.   Desired and achieved trajectory

the range for joint movements will give the best results at a cost of high computational power. In the second quarter, when the manipulator is crossing the object, a maximum deviation can be observed (refer Fig. 4d) and results into a bump in achieved trajectory as shown in Fig. 5. The manipulator started to converge to the required trajectory at the end of the quarter and the object is completely avoided while maintaining the smoothing action. Fig. 4e-4f shows its continuous motion in the unrestricted environment until the trajectory is completely traced. The required and reached trajectories are compared in Fig. 5 to depict the deviation of the manipulator from the desired path.

G. Conclusion

A novel approach of path tracing and object avoidance algorithm has been presented in this paper. A 6R open-chain manipulator is simulated on MATLAB to trace the trajectory when an object is placed mid-between its path. Forward kinematics is used to generate the configuration space. The manipulator is characterized by parameters of D-H Notation and represented as a superimposition of multiple cylinders to cover the rigid links. Similarly, nearby objects are also treated as a spherical shell. Multiple filters are used to pass only those joint states, which show a smooth, collision-free and minimum deviated path. These filters are robust to adapt to any environment. This algorithm is highly configurable for either the planning time or the deviation. A large number of simulations have been performed and results have shown the good reliability of the proposed approach.

Future objectives of this work are to reduce the planning time such that the algorithm can be implemented in the dynamic environment.
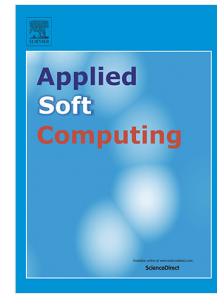
REFERENCES

[1] Tsai, C.S., 2014. Online Trajectory Generation for Robot Manipulators in Dynamic Environment--An Optimization-based Approach (Doctoral dissertation, UC Berkeley).

[2] H. Tabealhojeh and A. Ghanbarzadeh, "Two steps optimization path planning algorithm for robot manipulators using imperialist competitive algorithm," 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), Tehran, 2014, pp. 801-806.

[3] A. Hourtash and M. Tarokh, "Manipulator path planning by decomposition: algorithm and analysis," in IEEE Transactions on Robotics and Automation, vol. 17, no. 6, pp. 842-856, Dec. 2001. doi: 10.1109/70.976006.

[4] e Silva, J.S., Costa, P. and Lima, J., 2013, June. Manipulator path planning for pick-and-place operations with obstacles avoidance: an A* algorithm approach. In International Workshop on Robotics in Smart Manufacturing (pp. 213-224). Springer, Berlin, Heidelberg.

[5] Nagy, P.V., 1989. Trajectory tracking control for industrial robots. Journal of Mechanical Working Technology, 20, pp.273-281.

[6] Oliver, J.P.O., Domínguez-Ramirez, O.A. and Quezada, E.S.E., 2008, September. Trajectory tracking control for robotics manipulators based on passivity. In 2008 Electronics, Robotics and Automotive Mechanics Conference (CERMA'08) (pp. 484-489). IEEE.

[7] Liu, F. and Er, M.J., Trajectory Tracking of Robot Manipulators Using Linear and Nonlinear PD-type Controllers.

[8] H. A. Hendarto, Munadi and J. D. Setiawan, "ANFIS application for calculating inverse kinematics of programmable universal machine for assembly (PUMA) robot," 2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering, Semarang, 2014, pp. 35-40.

[9] John J. Craig. Introduction to Robotics: Mechanics and Control. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1989.

[10] Silva, E.C.E., Costa, M.F., Erlhagen, W. and Bicho, E., 2016, June. Superquadrics objects representation for robot manipulation. In AIP Conference Proceedings (Vol. 1738, No. 1, p. 300004). AIP Publishing.

[11] P. Bosscher and D. Hedman, "Real-time collision avoidance algorithm for robotic manipulators," 2009 IEEE International Conference on Technologies for Practical Robot Applications, Woburn, MA, 2009, pp. 113-122.

[12] Kevin M. Lynch and Frank C. Park. Modern Robotics: Mechanics, Planning, and Control. Cambridge University Press, New York, NY, USA, 1st edition, 2017.

[13] Denavit, J. & Hartenberg, R. S. (1955), 'A kinematic notation for lower-pair mechanisms based on matrices', Trans.ASME E, Journal of Applied Mechanics **22** , 215-221.

[14] J. M. Cho and K. Kim, "Precise object detection using local feature for robot manipulator," 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Jeju, 2017, pp. 497-499.

[15] K. Kim, J. Cho, J. Pyo, S. Kang and J. Kim, "Dynamic Object Recognition Using Precise Location Detection and ANN for Robot Manipulator," 2017 International Conference on Control, Artificial Intelligence, Robotics & Optimization (ICCAIRO), Prague, 2017, pp. 237-241.

[16] J. Kuehnle et al., "6D object localization and obstacle detection for collision-free manipulation with a mobile service robot," 2009 International Conference on Advanced Robotics, Munich, 2009, pp. 1-6.

[17] Yang, Yang & Cao, Qi-Xin. (2012). Monocular vision based 6D object localization for service robot's intelligent grasping. Computers & Mathematics with Applications. 64. 1235–1241. 10.1016/j.camwa.2012.03.067.

Journal Pre-proof

Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN

Deepak Jain, Akshi Kumar, Geetanjali Garg

Please cite this article as: D. Jain, A. Kumar and G. Garg, Sarcasm detection in mash-up language using soft-attention based bi-directional LSTM and feature-rich CNN, *Applied Soft Computing Journal* (2020), doi: https://doi.org/10.1016/j.asoc.2020.106198.

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.

Sarcasm Detection in Mash-up Language using Soft-Attention based Bi-directional LSTM and feature-rich CNN

Deepak Jain[1], Akshi Kumar [2*], Geetanjali Garg[3]

[1]Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China
[2, 3]Department of Computer Science & Engineering, Delhi Technological University, Delhi, India
{*Corresponding author:* akshikumar@dce.ac.in}

Abstract. Analyzing explicit and clear sentiment is challenging owing to the growing use of emblematic and multilingual language constructs. This research proposes sarcasm detection using deep learning in code-switch tweets, specifically the mash-up of English with Indian native language, Hindi. The proposed model is a hybrid of bidirectional long short-term memory with a *softmax* attention layer and convolution neural network for real-time sarcasm detection. To evaluate the performance of the proposed model, real-time mash-up tweets are extracted on the trending political (#government) and entertainment (#cricket, #bollywood) posts on Twitter. The randomly sampled dataset contains 3000 sarcastic and 3000 non-sarcastic bilingual Hinglish (Hindi +English) tweets. Feature engineering is done using pre-trained GloVe word embeddings to extract English semantic context vector, hand-crafted features using subjective lexicon Hindi-SentiWordNet to generate the SentiHindi feature vector and an auxiliary pragmatic feature vector depicting the count of pragmatic markers in tweet. Performance analysis is done to compare and validate the proposed [softAtt]BiLSTM- [feature-rich]CNN model. The model outperforms the baseline deep learning models with a superior classification accuracy of 92.71% and F-measure of 89.05%.

Keywords: Sarcasm; code-switch; mash-up; social media; deep learning

1. Introduction

The number of social network users worldwide continue to grow and is expected to pass 3 billion in 2020[1]. The social, interactive computer-mediated technologies, such as, Twitter, Tumblr, Google+, Facebook, Instagram, Snapchat, YouTube etc. which enable users to create, post and share all types of multimedia text are gaining users globally. Social media serves as a communication channel and a social listening, awareness, activism & feedback tool for stakeholder engagement and cooperation. Keeping tabs on social media sentiment, that is, *'the online mood'* is a key part of social media listening. Sentiment analysis[1, 2] is a key component of the social listening tool for measuring and reporting on the tone or sentiment of your social mention. It is the use of natural language processing (NLP) to analyze social conversations online and determine deeper context as they apply to a mention (topic, brand or theme). Determining the literal exactitude and opinion polarity in text is imperative for real-time sentiment analysis. Sentiment analysis task attains fitting results with literal language where articulation is same as anticipated interpretation. The non-standard vocabulary with informal textual content and noise currently define the content strategy of social media[3,4,5]. The common use of emblematic language markers such as punctuations (great!!!!!!), emojis (☺, 👍, ♥),micro-text [6] which includes wordplay (*suppperrrrr* for *super*), neologisms (*gr8* for *great*), Internet slangs (*FYI* for *for your information*), code-switching [7] and code-mixing [8] makes automated text analytics computational intensive.

As one of the most popular literary genre on social media, comedy defines the use of comic literary devices, which include wit, fun, humour, sarcasm, satire, pun, irony and non-sense. We characterize this as the 'Comedy Cube' (Fig.1).
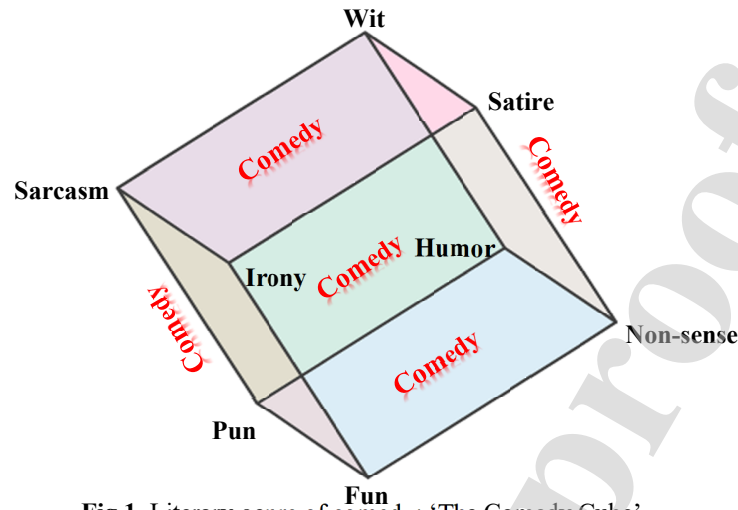
---

[1] https://www.emarketer.com/content/global-social-network-users

Fig.1. Literary genre of comedy: *'The Comedy Cube'*

Satire is a literary genre that relies heavily on irony, wit, and sometimes sarcasm. *Irony* describes the difference between what is expected and what actually happens and Wit is about understanding and intelligence. *Sarcasm* a sharply ironical taunt; sneering or cutting remark. It is the use of wit or ridicule to taunt, mock or contemptuously strike at a target. Irony is a common tool for sarcasm, but it is by no means the only one. Often, a mere laugh can be sarcastic, without any use of words. With reference to sentiment analysis, "Sarcasm is defined as a specific type of sentiment where people express their negative feelings using positive or intensified positive words in the text" [9]. It is a manifestation of considerable conflict between the tangible situation and the statement of expression. In a sample post, *"Yummmmmyy!!! Only if you like uncooked and stale food…go and waste your money!* ☺", this conflict between the actual situation of *"being served uncooked and stale food and is waste of money"* and the expression *"Yummy"* is evident. These inconsistencies within the opinion polarities characterize sarcasm as a special case of sentiment analysis making the performance of sentiment analysis task consequential to detection of sarcasm. Fig. 2 represents an example of a sarcastic tweet misread by the internet bot and consequently unfittingly answered.

Fig.2. Misinterpreted sarcastic tweet

Sarcasm detection is a complex task, as it is highly subjective and contextual [10]. Contextualization of words associates experience with expression to add sense. Basically, context creates meaning by providing precise and useful information. The generic content interpretation depends on three non-trivial elements, namely: the text, context and meaning (Fig.3) .
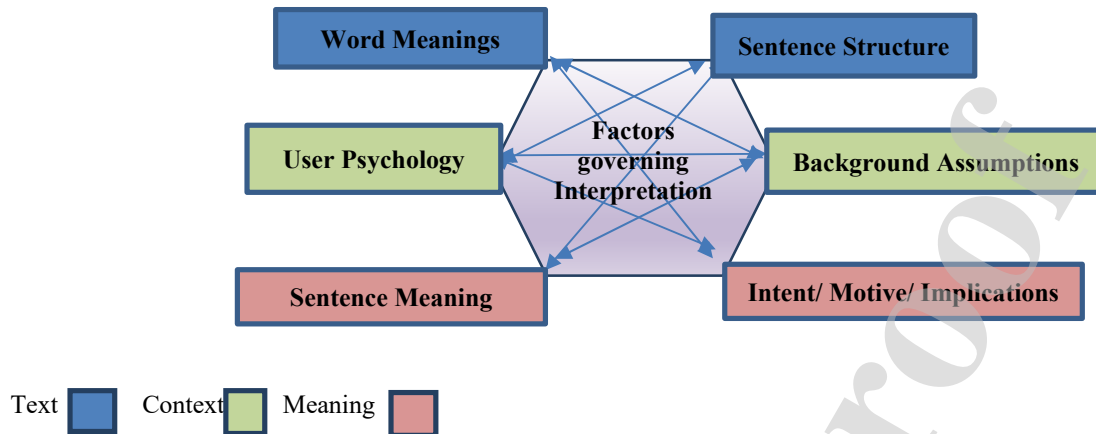
Fig.3. Factors governing Interpretation of content

It is imperative to comprehend supplementary cues from users' linguistic input that are aware of 'context' which aid right interpretation. But, elucidating context is one of the most challenging facets in content moderation. Though, contextual assistance has been studied across pertinent literature, its effectiveness in sentiment analysis to detect sarcasm needs further validation.

Multilinguality is a well-recognized challenge of natural language processing. The cultural diversities, country-specific trending topics and hash-tags on social media and easy availability of native language keyboards for social media applications add to the variety and volume of user-generated content in diverse languages and dialects [11]. The mixture of languages can be observed in text as follows:

- *Transliterated Code-mix*: I looooovvveed the movie!!! Dil aa gaya…ek dum super ❤
- *Literal Code-switch*: I looooovvveed the movie!!! दिल आ गया ...एक दम super ❤

The first case is an example of transliterated bilingual code-mixing where one language/script word (Hindi) is transcribed into a source language (English) such that the source phonetics is preserved. This is also known as phonetic typing. The second example describes a literal bilingual code-switch, where the actual words of one language (Hindi) are mashed up with the other language (English) demonstrating language alternation. Much of the recent research has been conducted on transliterated code-mix using publicly available API[2] or by building subjective lexicons [12]. Sentiment analysis research too has either been conducted on individual languages [13] or transliterated code-mix [14]. Also, specialized sentiments like emotion and sarcasm detection in tweets has been primarily done on monolingual English tweets where the non-ASCII words are removed during the data pre-processing step.

Recent statistics[3] show that Hindi is the fourth most widely spoken language around the world with about 310 million native speakers, coming in only after Chinese, Spanish and English. An upsurge in the use of hybrid of Hindi and English languages has been observed [15]. The availability of keyboards with 'Devanagari' scripts on mobile phones has made it a popular language choice. A research study from 2016 by Parshad et al. [16] shows that people are more fluent in Hinglish than in English alone. The research presented in this paper, focuses on sarcasm detection in the mash-up language (literal code-switch) tweets, specifically the Hinglish, which refers to the juxtaposition of words from Hindi and English language. For example, "*Well played AUSSIES!!!* एक और हार LOL 🖐🖕". The shift in polarities or intensification is analyzed for detecting sarcastic tone in Hinglish mash-up tweets.

As a typical natural language text classification task, automatic detection of sarcastic tone depends on feature engineering and learning model. The hierarchical learning capabilities and generalization offered

---

[2] Google Transliterate
[3]https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers

by deep learning architectures have made them a popular choice within natural language text processing [17, 18]. The proposed ^{softAtt}BiLSTM-$^{feature-rich}$CNN comprises of three primary modules:

- *English language processing module*: BiLSTM with an attention mechanism is used to generate semantic context vectors using pre-trained GloVe embeddings.
- *Hindi language processing module*: n-gram model is built and a subjective lexicon Hindi-SentiWordNet (H-SWN) is used [19] to generate the HindiSenti feature vector which is equivalent to a vector that counts the number of occurrences of each word in the vocabulary in the TF-IDF (term frequency inverse document frequency) vector encoding.
- *Classifier module*: CNN is trained using English feature vector, HindiSenti feature vector and auxiliary pragmatic features. Thus, the CNN comprises of feature-rich convolution layer, down-sampling layer and representation layer. The final dense layer contains an individual node for each category (sarcastic and non-sarcastic), with *softmax* activation function which generates the output predictions between 0-1 for each node.

GloVe, a pre-trained embedding is used to seed the model for English language feature vector generation. This count-based model constructs a co-occurrence matrix (words X context) based on the count that how frequently a word appears in context to represent words by feature vectors. This feature vector is given as input to the bidirectional LSTM (BiLSTM). Using Bi-LSTM, we run the input through two LSTMs, one reading the input left to right and another right to left, and then concatenating the outputs to generate an output vector. This BiLSTM output vector is converted into a weighted sum of the output vectors with attention probabilities to calculate the context vector. The intent of using attention is to amplify the contribution of important words and aggregate useful information. Intuitively, this context vector summarizes the importance of the different words. The English context feature vector along with auxiliary pragmatic features and HindiSenti Feature vector is provided as input to the CNN. Using multiple filter sizes over the data, the first layer performs convolutions to calculate a higher-level representation generating an activation map. Further, nonlinearity is maintained using ReLU activation function which generates a rectified feature map. Next, we down sample the result of the convolution layer into a pooled feature vector using max-pool, add dropout regularization, and classify the result using a fully-connected layer which flattens the high-level features learned by convolution layers and pass it to the output *softmax* layer to predict the input class label. Thus, the key contribution of this research are as follows:

- Building an automated model for detecting sarcasm in bi-lingual Hinglish mash-up tweets, ^{softAtt}BiLSTM-$^{feature-rich}$CNN, a hybrid of bidirectional long short-term memory with soft attention mechanism (^{softAtt}BiLSTM) and feature rich convolution neural network ($^{feature-rich}$CNN)
- Feature engineering using pre-trained word embeddings for English and n-grams with TF-IDF encoding for Hindi hand-crafted features and frequency encoding for punctuations.
- Evaluation using different kinds of discrete baseline deep learning architectures for detecting sarcasm.

The next section describes the related work in automatic sarcasm detection using deep learning architectures. The relevant studies focusing on code-mix and code-switch sarcasm detection are specifically included. Section 3 describes data acquisition and preprocessing of mash-up Hinglish data. Section 4 demonstrates the architecture and working of the proposed ^{softAtt}BiLSTM-$^{feature-rich}$CNN model for sarcasm detection in Hinglish mash-up followed by section 5 which presents the results. Finally, section 6, presents the conclusion and possible future aspects.

2. Related Work

The mounting global interest of users in social media portals has reinforced analytics and sensing-based research areas to discover knowledge from the publicly available content. Tapping the opinion of users within this big pool of user-generated data has found many practical applications within the market and government intelligence domains. Twitter has been the top choice to mine content due to its global presence and accessibility. 'Sentiment Analysis' [20] on all modalities (text, image, video, audio) of social data has been reported in literature. Primary studies with lexicon, machine learning and hybrid approaches are

abundantly available. Literature is well-equipped with reviews and surveys on sentiment analysis [1, 2, 11]. Pertinent studies specify contrast in opinion polarity to characterize sarcasm.

Deep learning models have been popularly used in natural language processing owing to their hierarchical learning and generalization capabilities [21]. Young et al. [22] discuss recent trends in deep learning based natural language processing. For detection of emotions, hybrid of attention based BiLSTM and CNN was explored by Felbo et al. [23]. Amir et al. [24] proposed a CNN trained using user and utterance based embeddings. Joshi et al. [25] used word embedding similarities as training features to detect sarcasm. Ghosh et al. [26] also worked on a hybrid of CNN, RNN and DNN. To explore syntactic and semantic information over tweets Zhang et al. [27] used biLSTM in their work. Poria et al. explored different set of features viz. emotions; sentiment and personality based in their study [28]. Sarcastic sentiment detection has also been studied extensively. Mukherjee et al. [29] applied linguistic styles of authors for sarcasm detection on Naïve Bayes and fuzzy clustering algorithms. Suhaimin et al. [30] proposed a feature extraction process to detect sarcasm in bilingual texts. They used different categories of like feature lexical, pragmatic, prosodic and syntactic. The usage of context with neural network models for sarcasm detection was explored by Ren et al. [31]. They concluded that context-augmented neural model effectively decodes sarcastic clues from contextual information. Hazarika et al. [32] proposed a contextual sarcasm detector that adopts a hybrid approach of both content and context-driven modeling for sarcasm detection. They also explored usage of stylometric and personality features of the users in their work. Majumder et al. [33], presented a multitask learning-based framework using a deep neural network for sarcasm detection. Cai et al. [34], proposed a hierarchical fusion model to deal with the problem of multimodal sarcasm detection. They treated text features, image features and image attributes as three modalities to address this task.

Recently, sentiment analysis in code-mixed languages has attracted much attention owing to the linguistic democratization and deep digital language divide. Sharma et al. [35] included abbreviations, word play, misspelled words and slang words and transliterated them to Romanized English words and analyzed the sentiment of the sentence using lexicon approach. The authors used text normalization before sentiment analysis on FIRE 2013 and FIRE 2014 data which consists of English and Hindi languages to calculate final sentiment score by lexicon lookup in respective dictionaries. Joshi et al. [19] developed Hindi-SentiWordNet, a lexical resource to comprehend sentiments in Devanagari script. A fuzzy logic based approach for polarity detection in Hindi text was given by Rana [36]. Mittal et al. [37] also used Hindi SentiWordNet to classify sentiments using negation and discourse relation. Sharma et al. [38] presented various methods for sentiment analysis of Hinglish language using the dataset of FIRE 2013 and FIRE 2014. Vyas et al. [39] explored POS tagging by using pure Hindi or English tagger and universal tagger in a code-mix script of Hindi and English in their research. Joshi et al. [40] introduced learning sub-word level representations in LSTM architecture for sentiment analysis of Hindi-English code-mixed text. Bhargava et al. [41] proposed sentiment analysis from code mixed sentences for English with combination of four different Indian languages. Vilares et al. [42] presented sentiment analysis for English and Spanish in different environments. Malgaonkar et al. [43] worked on classifying the emotions into various categories like sad, sarcasm, happy, fear, bliss etc using sentences written in a combination of various languages. Konate et al. in [44] used six deep learning models for the sentiment analysis task of code mixed Bambra-French language using the Facebook comments. Patra et al. [15] presented a paper for the task of sentiment identification from Hindi-English and Bengali-English code mixed datasets using the word and character level n- grams as features and SVM for the classification purpose. Abbasi et al. [45] also worked on a code mixed language of English and Arabic language and analyzed sentiment using entropy weighted genetic algorithms. Pakray et al.[46] addressed the difficulties in POS tagging of code-mixed data and introduced a Hidden Markov Model (HMM) based supervised algorithm as the solution of the same. Sane et al.[47] used deep learning techniques for humor detection in Hindi-English code-mixed tweets. Two most recent studies have focused on code-mix English-Hindi datasets for Sarcasm Detection [48], Irony detection [49] and Hate Speech Detection [50]. This research puts forward a sarcasm detection model for bilingual code-switch Hinglish tweets. To the best of our knowledge, this work is the first study on code-mix sarcastic sentiment detection.

3. Data Acquisition and Preprocessing

3.1 Data Acquisition

Real-time mash-up tweets are scrapped from Twitter using the Twitter Python API. We have mined the tweets using #sarcasm, keywords 'sarcasm' and 'sarcastic' and various the trending political (#government) and entertainment (#cricket, #bollywood) hashtags. Sarcasm Detector tool[4] is used to build a randomly sampled dataset of 30000 tweets (12,000 sarcastic and 18000 non-sarcastic tweets). Out of these, the Hinglish mash-up tweets are extracted manually to create a balanced dataset of 3000 sarcastic and 3000 non-sarcastic tweets that are considered as the dataset for this research.

3.2 Data Pre-Processing

Post data acquisition, data cleaning is done. The primary intent of pre-processing is to transform the data for extraction of features [51]. The process includes:

- Removing numeric and empty texts, URLs, mentions, hashtags, stop-words and punctuations.
- Tokenization of tweets is done using the TreebankWordTokenizer of Python Natural Language Toolkit (NLTK)[5] to filter the words, symbols and other elements called tokens [52]. The tokens are converted to lower case. For example, in a mash-up tweet "*Well done Aussies!! एक और हार LOL* ✋👎*"*, the tokens are generated as follows:

| well | done | aussies | ! | ! | एक | और | हार | lol | ✋ | 👎 |

- The slangs and emojis are replaced by their descriptive text using the *SMS Dictionary[6]* and *emojipedia[7]* respectively.

| well | done | aussies | ! | ! | एक | और | हार | laugh | out | loud |

| clapping | hands | thumbs | down |

- After tokenization, word-level language identification is performed using an online language identifier[8]. This is primarily done to annotate each word with a source language tag <HINDI> or <ENGLISH> so that the words can be sent to the respective language processing module. Table 1 depicts this annotation for the example tweet.

Table 1. Word language annotation

Token Word	Language tag
well	<ENGLISH>
done	<ENGLISH>
aussies	<ENGLISH>
एक	<HINDI>
और	<HINDI>
हार	<HINDI>
laugh	<ENGLISH>
out	<ENGLISH>

---

[4]*www.thesarcasmdetector.com*
[5]https://www.nltk.org/
[6]SMS Dictionary. *Vodacom Messaging*. Retrieved 16 March 2012.
[7] https://emojipedia.org/
[8]http://aztekium.pl/languages.py#start

loud	<ENGLISH>
clapping	<ENGLISH>
hands	<ENGLISH>
thumbs	<ENGLISH>
down	<ENGLISH>

- Punctuations are discarded during data pre-processing phase but in casual or informal writing such as text message or online posts, these are used as a technique to add emphasis to written text. Therefore, the count of each punctuation mark (!, ?, ., capitalization, 'x', "x") is extracted as a pragmatic feature set to train the model.

4. Deep Learning Model for Sarcasm Detection in Hinglish

The proposed deep learning model for sarcasm detection in Hinglish includes (i) feature extraction, and (ii) training ^{softAtt}BiLSTM- $^{feature-rich}$CNN model. The following sub-sections present the details.

4.1. Feature Extraction

The extraction of feature for training the model is done separately based on the source language as follows:

4.1.1. English language Feature Extraction

Word embedding is the collective name for a set of language modeling and feature learning techniques in natural language processing (NLP) where words or phrases from the vocabulary are mapped to vectors of real numbers. Embeddings capture the representation of the word in higher dimensional plane. Through embeddings, we create a vector representation of the word which is learned by understanding the context of words. There are many techniques to create word embeddings for the deep neural models, such as, one-hot encoding; TF-IDF encoding, Word2Vec and GloVe. In this research, GloVe, which is a count-based model to capture the semantic relatedness of English words, is used. It learns by constructing a co-occurrence matrix (words X context) that basically count how frequently a word appears in a context.

4.1.2. Hindi language Feature Extraction

The extraction of feature using the Hindi language tokens is a multi-step procedure where the tokens are firstly tagged based on the part-of-speech and then mapped to the Hindi-SentiWordNet [19] to generate polarity and polarity scores and create a HindiSenti feature vector. The details of this feature extraction procedure are as follows:

- All Hindi words are passed to a Hindi POS tagger[9] which assigns tags as adjective, noun, verb, adverb, interjection, common noun, question words etc. to each word. Thus, for the example given, the tags are generated as एक : Noun और: Adverb हार: Noun, Verb
- For each POS-tagged negative score > positive score). If the word is not present in the file then the synonyms of the word are found and assigned the +1 or -1 value. Along with this an attribute that will store the polarity score ($PolScr$) of that word is maintained. The polarity score within the range -1 to <0 shows negative sentiment and within the range >0 to 1 shows positive sentiment. A polarity score of 0 implies a neutral Hindi word, the polarity and polarity score is then determined using the Hindi-SentiWordNet (H-SWN). The H-SWN has various fields like (pos tag, Id, positive score, negative score, and synonyms). It is a subjective lexicon developed by IIT Bombay [19]. It contains words with part of speech and 3 scores positive, negative and objective. Sum of positive, negative and objective score sums to 1. The lexicon assigns single score to a word irrespective of the sense in which it is used.

If the word is present in the file then we will compare the associated positive and negative scores to assign a value of +1 if the positive score is greater than the negative one or the value -1 assigned to that word

---

[9]http://taghindi.herokuapp.com

representative of a negative sentiment(where polarity. Thus, the polarity score (*PolScr*) is calculated as follows:

> *For a word with positive sentiment or (value=+1)*
> PolScr =+1 * positive score
> *whereas, for a word with negative sentiment or (value=-1)*
> PolScr = -1 * negative score

Following this approach the polarity score of every token of a tweet is calculated. A HindiSenti Feature Vector is then constructed by converting words into features using a language-independent n-grams technique with the *Tfidf* Vectorizer. The HindiSenti feature vector also keeps record of the corresponding polarity score for every feature.

4.1.3. Auxiliary Pragmatic Features

Whereas GloVe captures the conceptual relationship within words, pragmatic features necessarily portray the use of language. Punctuations such as exclamation mark, quotation marks, capatalization are used to add emphasis in written informal text and are significant signs which assist to comprehend the context inconsistency or intensity within the text as given below:

- Exclamation mark (!): Act as emotion intensifier without polarity shift. For example: "*She looks beautiful!!!!*" is more intense than "*She looks beautiful.*"
- Wordplay: It involves word lengthening by repetition of alphabets. For example, the long sequence of word, "*sorryyyyy*" intensifies the expression.
- Uppercase word: Explicit use of uppercase alphabets intensifies the semantics. For example, "*I am so SORRY!*"is more persuasive than "*I am so sorry!*"

Moreover, the frequencies of occurrence of punctuations in a tweet strongly suggest sarcasm. Thus, frequency encoding is done for these pragmatic punctuation-based features to create a pragmatic feature vector p with six tuples, <R, E, I, P, U, Q>, where, R is the frequency of recurring alphabetic character, (that is, if recurrence > 2 set R =1, else 0) and E, I, P, U and Q defines the count of exclamation marks, question marks (interrogatory), periods, uppercase letters, single quotes('') or double quotes ("") respectively. The conceptual flow of feature extraction in the proposed model is shown in fig. 4:
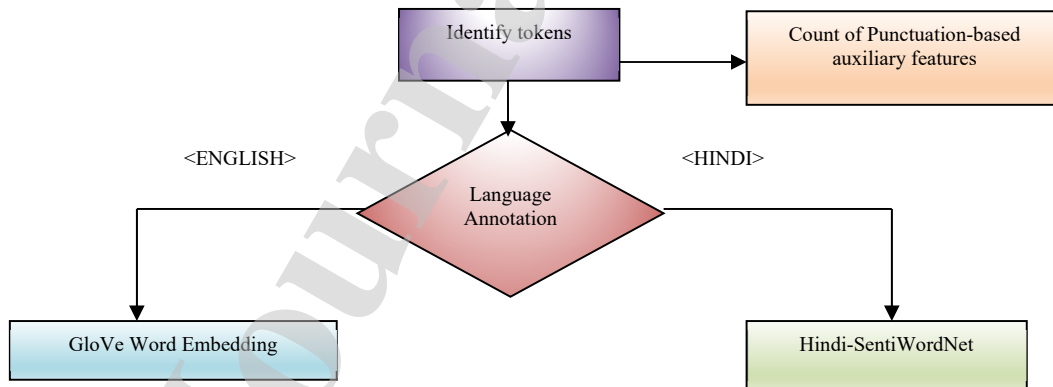

Fig. 4. Feature extraction in the proposed model

4.2. The proposed ^{softAtt}BiLSTM- ^{feature-rich}CNN Model

An automated model for detecting sarcasm in bi-lingual Hinglish mash-up tweets is a hybrid of bidirectional long short-term memory with soft attention mechanism (^{softAtt}BiLSTM) and feature rich convolution neural network (^{feature-rich}CNN)trained using a combination of English, Hindi and auxiliary pragmatic feature vectors. Fig. 5 depicts the architecture of the proposed deep learning model.
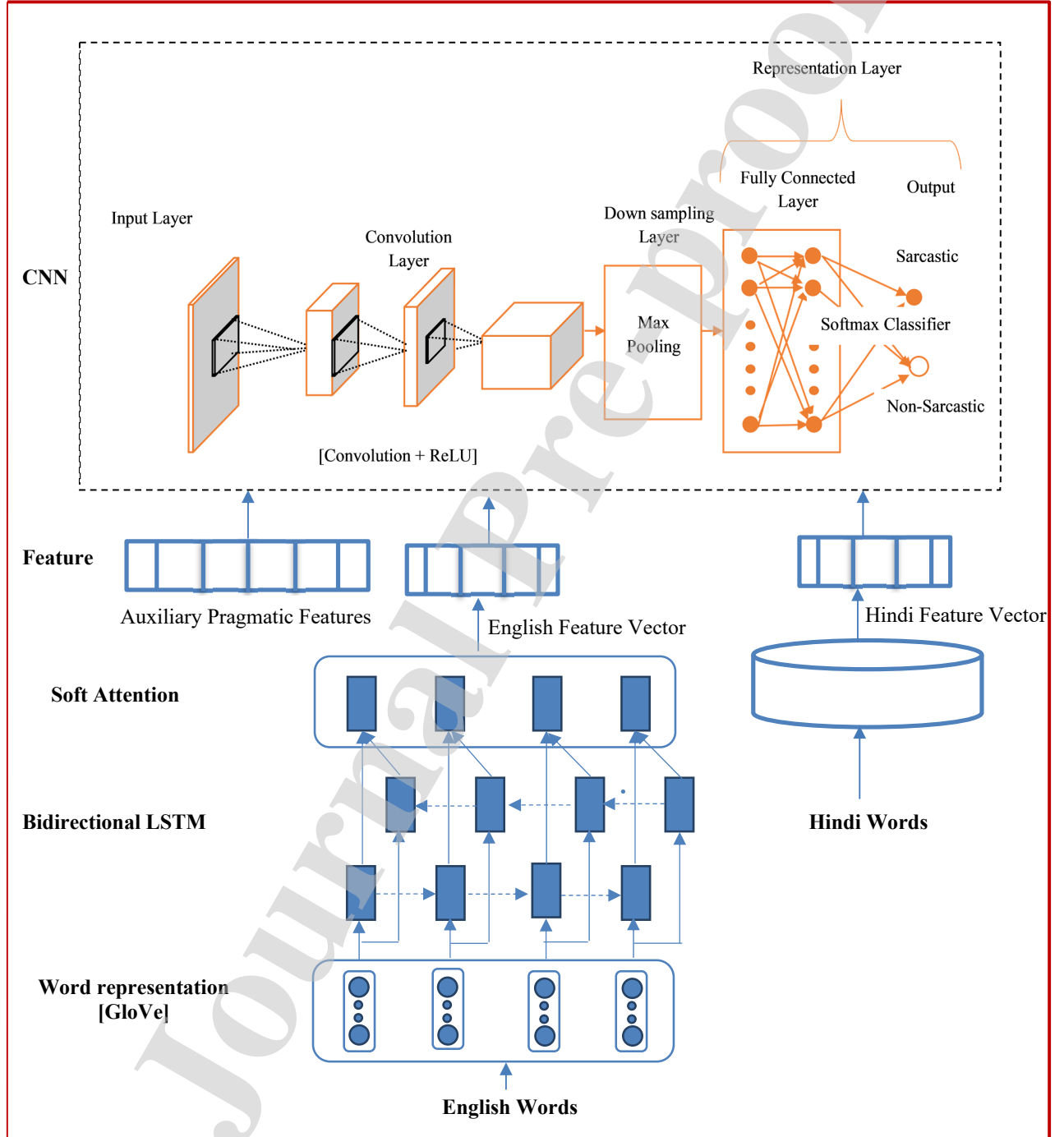

Fig.5. Architecture of the ^{softAtt}BiLSTM- ^{feature-rich}CNN

The model consists of three modules, namely the English language processing module which generates context vectors using attention based BiLSTM; Hindi language processing module which generates TF-IDF encoded HindiSenti feature vector using Hindi-SentiWordNet (H-SWN); and the classifier module where CNN is trained using three feature vectors to generate the output predictions. The following subsections expound the details.

4.2.1. English context feature vector generation using BiLSTM with attention mechanism

This component consists of three layers of the deep learning architecture:

- *Embedding Layer*: Embedding layer maps discrete words into lower dimensional vector space for computational efficiency. In this layer the English language input is mapped to real-valued vectors using pre-trained word embedding. The benefits of extracting features based on learned representations for sarcasm detection are available in literature [52].GloVe has been used to generate 'word vector table' with an embedding dimension of 300 and a batch size of 50.

- *BiLSTM Layer:* The long short-term memory (LSTM) in its core, preserves information from inputs that has already passed through it using the hidden state. Unidirectional LSTM only preserves information of the past because the only inputs it has seen are from the past. During training, errors back-propagate to adjust weights while maintaining the training properties. The calculation for the forget, the input and the output gate along with the input cell state is done using (1) to (6):

$$i_t = \sigma\ (W_i \cdot [h_{t-1}, x_t] + b_i) \tag{1}$$

$$f_t = \sigma\ (W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

$$o_t = \sigma\ (W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$

$$s_t = tanh\ (W_s \cdot [h_{t-1}, x_t] + b_s) \tag{4}$$

$$c_t = f_t * c_{t-1} + i_t * s_t \tag{5}$$

$$h_t = tanh\ (c_t) * o_t \tag{6}$$

where,

- x_t is the *t*-th word vector that is it denotes the word representation of w_t
- $*$ represents the product (element wise)
- $W_i W_f W_o W_s$ are model parameters
- $b_i b_f b_o b_s$ represents the bias vectors
- The σ is the sigmoid function used as the gate activation function
- *tan h* is the hyperbolic tangent function

We use bidirectional LSTM to obtain word features H = (h_1, h_2, ..., h_n) concatenated from both directions. Bi-LSTM has two networks, one which reads information in forward direction and another which reads it in the reverse direction. This way the output is generated from both the past and future context. A forward LSTM processes the sentence (tweet) from x_1 to x_n, while a backward LSTM processes from x_n to x_1. For word x_t, a forward LSTM obtains a word feature as $\overrightarrow{h_t}$ and a backward LSTM obtains the feature as $\overleftarrow{h_t}$. Then, h is calculated using (7) [53]:

$$h_i = (\overrightarrow{h_i} \odot \overleftarrow{h_i}) \tag{7}$$

Where, h_i is the output of the *i*-th word

$\odot$ function is a concatenation function used to combine the two outputs. Generally, different merge modes can be used to combine the outcomes of the Bi-LSTM layers. These are concatenation (default), multiplication, average, and sum.

$\vec{h}$, represents the output sequence of the forward layer which is calculated iteratively using inputs in a positive sequence from time t-n to time t-1,

$\overleftarrow{h}$ represents the output sequence of the backward layer which is calculated using the reversed inputs from time t-n to t-1.

- *Attention Layer*: The concept of attention is based on the hypothesis that words in a sentence (tweet/posts) have different importance quotient while defining the meaning in 'context'. In this work, the differentiable and deterministic *soft*-attention mechanism is used to generate the output as a weighted combination of all the input states rather than just last state (Fig.6).
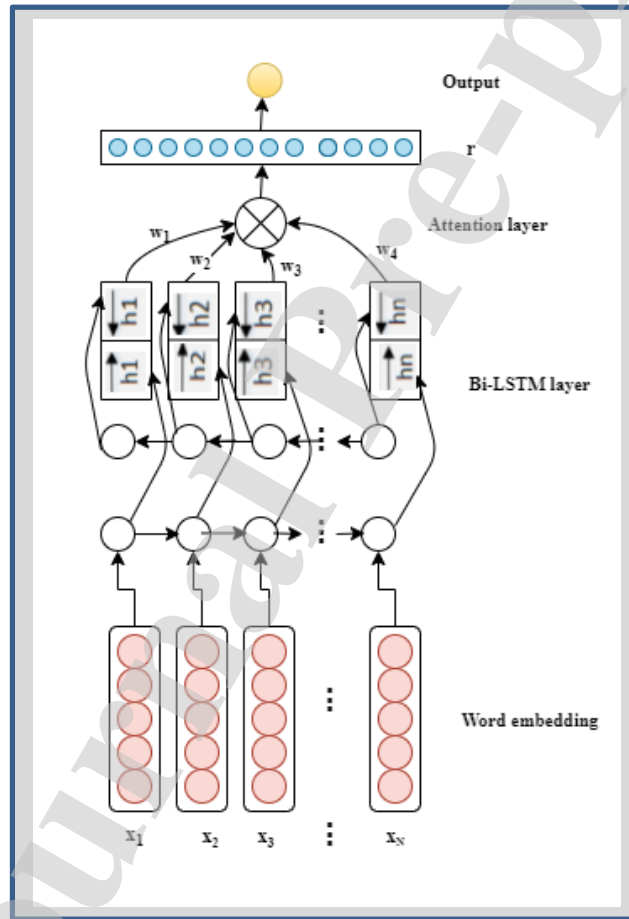

Fig.6. Bi-directional LSTM with Attention

Using this word-level attention mechanism [54], the attention score e_i is given to each word i in the sentence t as given by "(9)".

$$e_i = tanh(W_h h_i + b_h) \qquad (9)$$

Where W_h and b_h are the weight and bias from the attention layer.

And then using attention mechanism, the weight w_i is assigned to each word feature h_i using "(10)".

$$w_i = \frac{\exp(e_i)}{\sum_{t=1}^{N} \exp(e_t)} \quad , \quad \sum_{i=1}^{N} w_i = 1 \qquad (10)$$

The hidden states are finally calculated to produce a hidden sentence feature vector r by a weighted sum function using "(11)".

$$r = \sum_{i=1}^{N} w_i h_i \qquad (11)$$

4.2.2. *Feature-rich* CNN

The CNN is enriched with features from both languages which define the mash-up along with pragmatic marker features. That is, the input to the CNN is a combination of the output features of the [sAtt]BiLSTM, auxiliary punctuation-based features and the HindiSenti feature vector. It comprises the next three layers of the deep learning architecture, that is, the convolution layer, the down-sampling layer and the representation layer (Fig. 7):
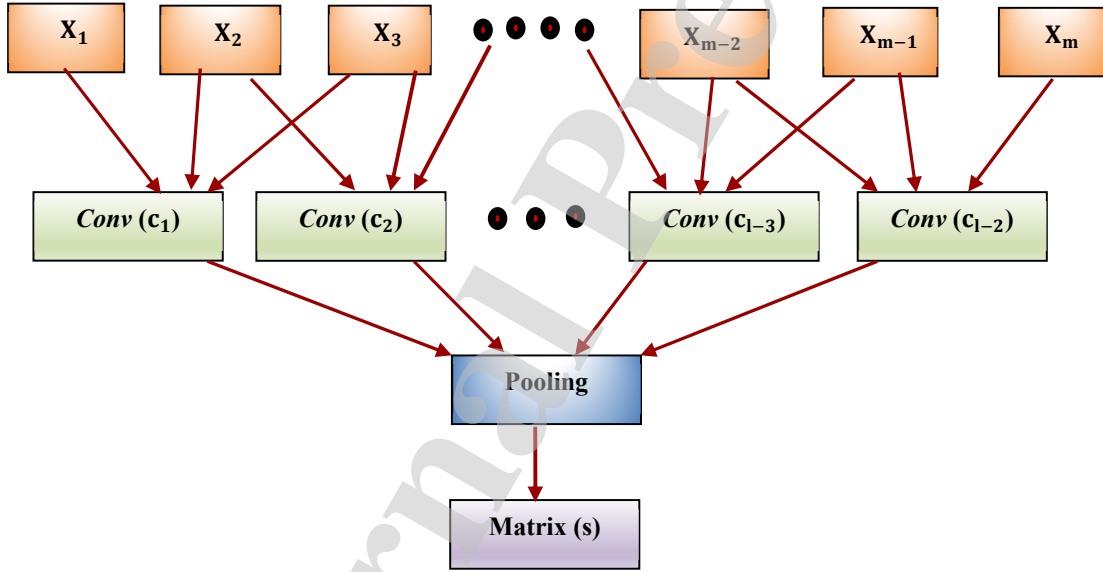

Fig. 7. Convolution and pooling

- *Convolution layer*: Filters perform convolutions and generate (variable-length) feature maps. The model uses three layer convolution architecture with a total of 100 convolution filters each for window size (3, 3). The dropout regularization is set to 0.5 to ensure that that model does not overfit. ReLU activation function is applied to introduce nonlinearities into the model and generate a rectified feature map.
- *Down-sampling layer*: Also known as pooling layer, this layer down samples the feature map and retains the non-trivial features. In this study, max pooling is used, in order capture the maximum value within the rectified feature map.
- *Representation layer*: This is the output layer which consists of a dense layer where every node in the layer is connected to every node in the preceding layer thus forming a fully-connected layer. This layer classifies the output using a *softmax* function to generate a value between 0 (non-sarcastic) and 1 (sarcastic).

5. Results

The discussion of results is done in two parts: (i) performance of the proposed deep learning model (ii) performance comparison with baselines.

5.1. Performance Results

The predictive modeling using 6000 random Hinglish tweets is done. The performance has been evaluated using accuracy, recall, precision, and F-measure. The following Table 2 depicts the results the proposed [softAtt]BiLSTM- [feature-rich]CNN deep learning model.

Table 2. Performance of the [softAtt]BiLSTM- [feature-rich]CNNModel

Performance Quantifier	Value (%)
Accuracy	92.71
Recall	90.67
Precision	89.49
F-measure	89.05

5.2 Comparison with Baselines

The model is evaluated individually for different language modules and hand-crafted features. The results of using CNN for Hindi words and CNN for Hindi words with punctuations are shown in tables 3 and 4 respectively.

Table 3. Performance of CNN-Hindi language

Performance Quantifier	Value (%)
Accuracy	43.44
Recall	42.78
Precision	42.89
F-measure	45.63

Table 4. Performance of CNN-Hindi language with punctuations

Performance Quantifier	Value (%)
Accuracy	54.94
Recall	50.22
Precision	53.9
F-measure	54.64

The performance for English language with and without punctuations is evaluated for both CNN and BiLSTM models individually. The following tables 5, 6, 7 and 8 depict the results.

Table 5. Performance of CNN-English Language

Performance Quantifier	Value (%)
Accuracy	70.35
Recall	69.92
Precision	70.59
F-measure	71.12

Table 6. Performance of CNN-English Language with punctuations

Performance Quantifier	Value (%)
Accuracy	81.28
Recall	80.61
Precision	83.01
F-measure	82.72

Table 7. Performance of BiLSTM-English language

Performance Quantifier	Value (%)
Accuracy	69.75
Recall	68.28
Precision	64.91
F-measure	66.22

Table 8. Performance of BiLSTM-English language with punctuations

Performance Quantifier	Value (%)
Accuracy	71.03
Recall	70.27
Precision	68.77
F-measure	69.45

The graph in fig. 8 depicts the comparison of accuracies of deep models for individual languages with and without punctuation
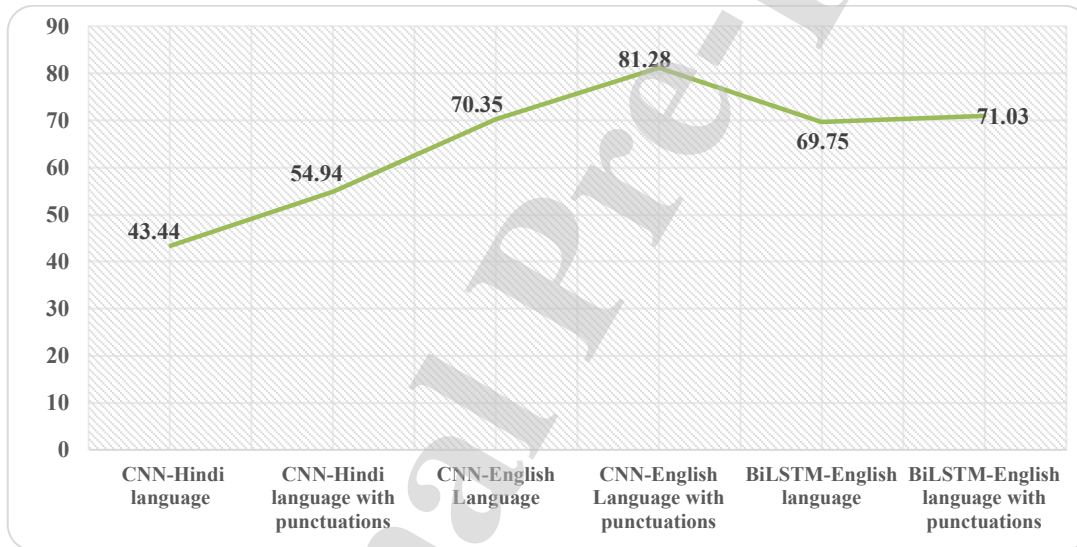
Fig. 8. Comparison of individual deep learning model on language modules

The proposed model is also compared with two deep learning hybrid architectures, namely, LSTM with [feature-rich]CNN and BiLSTM without attention hybrid with [feature-rich]CNN. Table 9 and 10 presents the results. The graph in fig.9 gives the comparison of these two hybrid deep learning models to the proposed model.

- LSTM-[feature-rich]CNN

Table 9. Performance of the LSTM-[feature-rich]CNN Model

Performance Quantifier	Value (%)
Accuracy	81.75
Recall	78.28
Precision	74.91
F-measure	76.45

- BiLSTM-[feature-rich]CNN (BiLSTM without Attention)

Table 10. Performance of the BiLSTM-[feature-rich]CNN Model

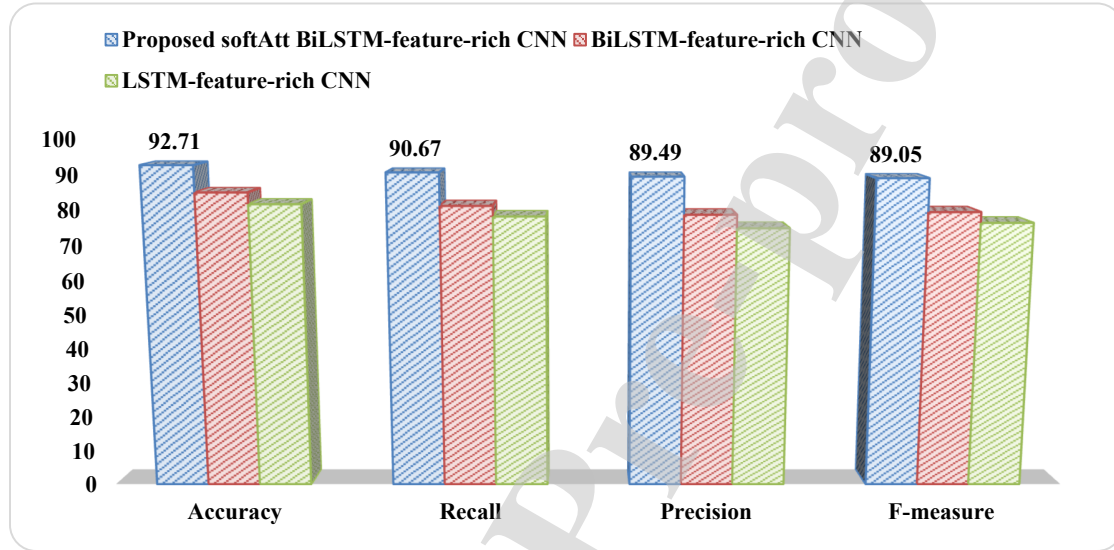| Performance Quantifier	Value (%)
Accuracy	85.03
Recall	81.27
Precision	78.77
F-measure	79.51

Fig.9. Comparison of hybrid deep learning model performance

The results demonstrate that the proposed [softAtt]BiLSTM-[feature-rich]CNNmodel achieves the highest accuracy of 92.71% as compared to the other variants. The order of models from highest to lowest accuracy is, [softAtt]BiLSTM-[feature-rich]CNN>BiLSTM-[feature-rich]CNN>LSTM-[feature-rich]CNN. Fig.10 which depicts the comparison of F-measure obtained for each model.The proposed [softAtt]BiLSTM-[feature-rich]CNN also achieves the highest recall, precision and F-measure.
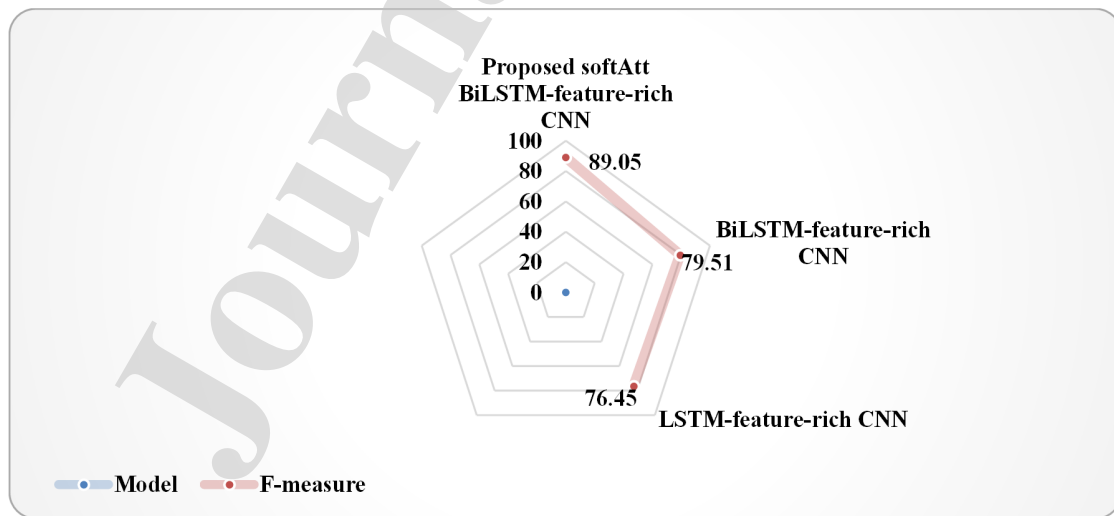

Fig. 10. F-measure of the proposed model with baselines

6. Conclusion

Detecting sarcasm, irony and emotions in real-time, code-switched user-generated text is imperative for fine-grained sentiment analysis. This work presented a foremost learning model for real-time sarcasm detection in Hinglish code-switch dataset. The *softmax* attention based bi-directional LSTM learned semantic context vector for English features from the GloVe word embeddings and forwarded it to the CNN. The HindiSenti feature vector and auxiliary punctuation-based features were also combined with the English feature vector to train the CNN. Performance benchmarking was done to evaluate the effectiveness of the proposed model and high classification accuracy and F-measure were observed. Mining sentiments in Hindi is a challenge as it is a morphologically rich and a free order language as compared to English. This amplifies the intricacies of managing user-generated content. The sarcastic tone classification is thus limited by the capabilities of Hindi-SentiWordNet for extraction of hand-crafted Hindi language features. Moreover, identifying context, performing word sense disambiguation and wordplay are open research problems associated with both English and Hindi languages. Also, the proposed model relies on the efficiency of online language identifier and part-of-speech tagger which have not yet reached the state-of-the-art accuracy.

Conflict of Interest Statement

The authors certify that there is no conflict of interest in the subject matter discussed in this manuscript.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] Pang B. Lee L. (2008). "Opinion mining and sentiment analysis", Foundations and trends in information retrieval, vol. 2, no. 1–2, pp. 1-135.

[2] Wang, K., Bansal, M., & Frahm, J. M. (2018, March). Retweet wars: Tweet popularity prediction via dynamic multimodal regression. In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (pp. 1842-1851). IEEE.

[3] Roy, A., & Todorovic, S. (2018, August). Learning to Learn Second-Order Back-Propagation for CNNs Using LSTMs. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 97-102). IEEE.

[4] Chen J, Yan S, Wong K C. (2018). "Verbal aggression detection on Twitter comments: convolutional neural network for short-text sentiment analysis." Neural Computing and Applications: 1-10.

[5] Zubiaga A., Aker A., Bontcheva K., Liakata M., Procter, R.(2017). "Detection and resolution of rumours in social media: A survey". arXiv preprint arXiv:1704.00656

[6] Satapathy R., Guerreiro, C., Chaturvedi, I., Cambria, E. (2017). "Phonetic-Based Microtext Normalization for Twitter Sentiment Analysis", in: ICDM, 407– 413.

[7] Nilep C. (2006). "Code switching in sociocultural linguistics", Colorado Research in Linguistics, vol. 19, no. 1, pp. 1-22.

[8] Barman U, Das A., Wagner J., Foster J. (2014). "Code Mixing: A Challenge for Language Identification in the Language of Social Media", EMNLP 2014, pp. 13, 2014.

[9] Bharti S.K., Vachha B., Pradhan R K., Babu K S, Jena S K. (2016). "Sarcastic sentiment detection in tweets streamed in real time: a big data approach", Digital Communications and Networks. vol. 2, no. 3, pp. 108-21.

[10] Kumar A., Garg G. (2019). "Systematic Literature Review on Context-Based Sentiment Analysis in Social Multimedia", Multimedia tools and Applications, https://doi.org/ 10.1007/s11042-019-7346-5

[11] Matras Y., Bakker P. (2003). "The mixed language debate: Theoretical and empirical advances", Walter de Gruyter, vol. 145.

[12] P. Arora, (2013). "Sentiment Amalysis for Hindi Language" Master thesis, IIT, Hyderabad.

[13] Zheng W., Qiang Y E. (2009). "Sentiment Classification of Chinese Traveler Reviews by Support Vector

Machine Algorithm", IEEE Third International Symposium on Intelligent Information Technology Application.

[14] Mandal, S., Das, S D., Das D. (2018). "Language identification of Bengali-English code-mixed data using character & phonetic based lstm models". arXiv preprint arXiv:1803.03859.

[15] Patra, B G., Das, D., Das A. (2018). "Sentiment analysis of code-mixed Indian languages: An overview of sail code-mixed shared task@ icon-2017." arXiv preprint arXiv:1803.06745.

[16] Parshad R D., Bhowmick, S., Chand, V. Kumari, N., Sinha, N. (2016). "What is India speaking? Exploring the "Hinglish" invasion." Physica A: Statistical Mechanics and its Applications 449 (2016): 375-389.

[17] Sivanandam, S N., Deepa S N. (2007). "Principles of Soft Computing (With CD)", John Wiley & Sons.

[18] Lei, P., & Todorovic, S. (2016, October). Recurrent temporal deep field for semantic video labeling. In European Conference on Computer Vision (pp. 302-317). Springer, Cham.

[19] Das, A., & Bandyopadhyay, S. (2010, August). SentiWordNet for Indian languages. In Proceedings of the Eighth Workshop on Asian Language Resources, pp. 56-63.

[20] Dave K, Lawrence S, Pennock D M. (2003). "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews." Proceedings of the 12th international conference on World Wide Web. ACM; 519-528.

[21] Wang, R., Frahm, J. M., & Pizer, S. M. (2018). Recurrent neural network for learning dense depth and ego-motion from video. arXiv preprint arXiv:1805.06558.

[22] Young T, Hazarika D, Poria S, Cambria E. (2018). "Recent trends in deep learning based natural language processing". IEEE Computational Intelligence Magazine 13 (3), 55-75.

[23] Felbo B., Mislove A., Søgaard A., Rahwan I., Lehmann S. (2017). "Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm", arXiv preprint arXiv:1708.00524.

[24] Amir, S., Wallace B C, Lyu H., Silva P C M J. (2016). "Modelling context with user embeddings for sarcasm detection in social media", arXiv preprint arXiv:1607.00976.

[25] Joshi A., Bhattacharyya P., Carman M., Saraswati J., Shukla R. (2016). "How do cultural differences impact the quality of sarcasm annotation?: A case study of indian annotators and american text.", In Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities pp. 95-99.

[26] Ghosh A., Veale T. (2016). "Fracking sarcasm using neural network", In Proceedings of the 7th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp. 161-169, 2016.

[27] Zhang M., Zhang Y., Fu G. (2016). "Tweet sarcasm detection using deep neural network", In Proceedings of COLING 2016, The 26th International Conference on Computational Linguistics: Technical Papers, pp. 2449-2460.

[28] Poria S, Cambria E, Hazarika D, Vij P. (2016). "A deeper look into sarcastic tweets using deep convolutional neural networks". Proc. Int. Conf. Computational Linguistics, pp. 1601-1612.

[29] Mukherjee, S., & Bala, P. K. (2017). Sarcasm detection in microblogs using Naïve Bayes and fuzzy clustering. Technology in Society, 48, 19-27.

[30] Suhaimin, M. S. M., Hijazi, M. H. A., Alfred, R., & Coenen, F. (2017, May). Natural language processing based features for sarcasm detection: An investigation using bilingual social media texts. In 2017 8th International Conference on Information Technology (ICIT) (pp. 703-709). IEEE.

[31] Ren, Y., Ji, D., & Ren, H. (2018). Context-augmented convolutional neural networks for twitter sarcasm detection. Neurocomputing, 308, 1-7.

[32] Hazarika, D., Poria, S., Gorantla, S., Cambria, E., Zimmermann, R., & Mihalcea, R. (2018). Cascade: Contextual sarcasm detection in online discussion forums. arXiv preprint arXiv:1805.06413.

[33] Majumder, N., Poria, S., Peng, H., Chhaya, N., Cambria, E., & Gelbukh, A. (2019). Sentiment and sarcasm classification with multitask learning. IEEE Intelligent Systems, 34(3), 38-43.

[34] Cai, Y., Cai, H., & Wan, X. (2019, July). Multi-Modal Sarcasm Detection in Twitter with Hierarchical Fusion Model. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 2506-2515).

[35] Sharma S, Srinivas PY, Balabantaray RC. (2015)."Text normalization of code mix and sentiment analysis". In Advances in Computing, Communications and Informatics (ICACCI), 2015 pp. 1468-1473, IEEE.

[36] Rana S. (2014). "Sentiment Analysis for Hindi Text using Fuzzy Logic". Indian Journal of Applied Research, Vol.4, Issue.8.

[37] Mittal, N., Agarwal B., Chouhan G., Bania N., Pareek P. (2013). "Sentiment Analysis of Hindi Review based on Negation and Discourse Relation." In proceedings of International Joint Conference on Natural Language Processing, pp. 45-50.

[38] Sharma S, Srinivas PY, Balabantaray RC. (2015). "Sentiment analysis of code-mix script." In Computing and

Network Communications (CoCoNet), 2015 pp. 530-534. IEEE.

[39] Vyas Y., Gella S., Sharma J., Bali K., Choudhury m. (2014). "POS Tagging of English-Hindi Code Mixed Social Media Content." In Proceedings of the Con-ference on Empirical Methods in Natural Language Processing (EMNLP), 974-979

[40] Joshi A, Prabhu A, Shrivastava M, Varma V. (2016). "Towards sub-word level compositions for sentiment analysis of Hindi-English code mixed text." In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 2482-2491.

[41] Bhargava R., Sharma Y., Sharma S. (2016). "Sentiment Analysis for Mixed Script Indic Sentences", International Conference on Advances in Computing Communications and Informatics ICACCI 2016, pp. 524-529.

[42] Vilares D., Alonso M A., G´omezRodr´ıguez C. (2015). "Sentiment Analysis on Monolingual, Multilingual and Code-Switching Twitter Corpora." In Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 2–8, Association for Computational Linguistics.

[43] Malgaonkar S., Khan, A., Vichare A. (2017). "Mixed bilingual social media analytics: case study: Live Twitter data." 1407-1412. 10.1109/ICACCI.2017.8126037.

[44] Konate, A. & Du, R. (2018). "Sentiment Analysis of Code-Mixed Bambara-French Social Media Text Using Deep Learning Techniques, Wuhan Univ. J. Nat. Sci. (2018) 23: 237. https://doi.org/10.1007/s11859-018-1316-z

[45] Abbasi, A., Chen, H., & Salem, A. (2008). "Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums." ACM Transactions on Information Systems, 26(3), [12]. https://doi.org/10.1145/1361684.1361685

[46] Pakray, P., Majumder, G., & Pathak, A. (2018, January). An HMM Based POS Tagger for POS Tagging of Code-Mixed Indian Social Media Text. In Annual Convention of the Computer Society of India (pp. 495-504). Springer, Singapore.

[47] Sane, S. R., Tripathi, S., Sane, K. R., & Mamidi, R. (2019, June). Deep Learning Techniques for Humor Detection in Hindi-English Code-Mixed Tweets. In Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (pp. 57-61).

[48] Swami S., Khandelwal A., Singh V., Akhtar S., Shrivastava, M. (2018). "A Corpus of English-Hindi Code-Mixed Tweets for Sarcasm Detection." eprint arXiv:1805.11869, CICLing 2018

[49] Vijay D., Bohra A., Singh V., Akhtar S., Shrivastava M. (2018). "A Dataset for Detecting Irony in Hindi-English Code-Mixed Social Media Text." EMSASW@ESWC, volume 2111 of CEUR Workshop Proceedings, page 38-46, CEUR-WS.org, (2018)

[50] Kamble S., Joshi A. (2018). "Hate Speech Detection from Code-mixed Hindi-English Tweets Using Deep Learning Models. arXiv preprint arXiv:1811.05145, 2018.

[51] Guyon, I. and Elisseeff, A. (2006). "An introduction to feature extraction." In Feature extraction (pp. 1-25). Springer, Berlin, Heidelberg.

[52] Loper E., Bird S. (2002). "NLTK: The natural language toolkit", Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics, Association for Computational Linguistics vol. 1, pp. 63-70, 2002.

[53] Graves A., Jaitly N., Mohamed A R (2013). "Hybrid speech recognition with deep bidirectional LSTM" Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop pp. 273-278.

[54] Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R. and Bengio, Y. (2015). "Show, attend and tell: Neural image caption generation with visual attention." In International conference on machine learning (pp. 2048-2057)

Author Biographies

DEEPAK KUMAR JAIN is an Assistant Professor at Chongqing University of Posts and Telecommunications, Chongqing, China. He received the degrees of Bachelor of Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, India and Master of Technology from Jaypee University of Engineering and Technology, India in 2010 and 2012, respectively. He completed his Ph.D. from University of Chinese Academy of Sciences, Institute of Automation, Beijing, China. He has presented several papers in peer reviewed Conferences, as well as published numerous studies in science cited journals. His areas of research are deep learning, machine learning, pattern recognition, and computer vision.

AKSHI KUMAR is an Assistant Professor in the Department of Computer Science & Engineering at Delhi Technological University (formerly Delhi College of Engineering). She has been with the university for the past 11 years. She has received her Ph.D. in Computer Engineering from Faculty of Technology, University of Delhi in 2011. She completed her Master of Technology with honours in Computer Science & Engineering from Guru Gobind Singh Indraprastha University, Delhi in 2005. She received her Bachelor of Engineering degree with distinction in Computer Science & Engineering from Maharishi Dayanand University, India in 2003. She has presented several papers in international conferences and published work in peer-reviewed and science cited journals. She is a recipient of *"Commendable Research Award for Excellence in Research"* at Delhi Technological University, in 2018 and 2019 and *"VIWA Outstanding Women in Engineering (Computer Science & Engineering)"* Award, in 2020. Dr. Kumar has authored a monograph 'Web Technology: Theory and Practice' published by CRC Press, Taylor and Francis Group and edited a book titled 'A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development', Springer. Dr. Kumar has also authored a monograph 'Web Technology: Theory and Practice' published by CRC Press, Taylor and Francis Group and edited a book titled 'A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development', Springer. She also has an Indian Patent published. She is active reviewer of many top journals of IEEE, IET, Springer and Elsevier.In addition,she is currently serving as a Guest Editor of Energy Systems, Springer, Future Internet, IEEE Internet of Things Magazine and Sustainable Cities & Society, Elsevier. Her research interests are in the area of affective computing, social media analytics, sentiment analysis, natural language understanding, open innovation & smart solutions.

GEETANJALI GARG is an Assistant Professor in the Department of Computer Science & Engineering at Delhi Technological University (formerly Delhi College of Engineering). She is pursuing her Ph.D. from the Department of Computer Science & Engineering at Delhi Technological University. She received her ME degree in Computer Technology & Applications from the University of Delhi in 2006. She received her B.Tech degree in Computer Science & Engineering from Kurukshetra University in 2002. Her research interests include Social media analytics, soft computing and intelligent systems.

Sarcasm Detection in Mash-up Language using Soft-Attention based Bi-directional LSTM and feature-rich CNN

Deepak Jain[1], Akshi Kumar[2*], Geetanjali Garg[3]

[1]Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China

[2, 3]Department of Computer Science & Engineering, Delhi Technological University, Delhi, India

{*Corresponding author:* akshikumar@dce.ac.in}

Conflict of Interest Statement

The authors certify that there is no conflict of interest in the subject matter discussed in this manuscript. 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

Author Statement

Sarcasm Detection in Mash-up Language using Soft-Attention based Bi-directional LSTM and feature-rich CNN

Deepak Kumar Jain: Conceptualization, Reviewing and Editing

Akshi Kumar: Conceptualization, Dataset preparation, Methodology, Implementation, Validation, Writing, Reviewing and Editing

Geetanjali Garg: Dataset preparation, Methodology, Implementation, Validation, Writing

PAPER • OPEN ACCESS

Study of hybrid photovoltaic thermal systems

View the article online for updates and enhancements.

Study of hybrid photovoltaic thermal systems

Ashish Saurabh[1,3], Deepali Atheaya[1], Anil Kumar[2]

[1]Department of Mechanical and Aerospace Engineering,
 Bennett University, Tech Zone – II, Greater Noida - 201310, UP, India
[2]Department of Mechanical Engineering,
 Delhi Technological University, Delhi
[3]Corresponding Author, email: as6723@bennett.edu.in

Abstract. Sun is the primary source of renewable energy. It is abundant, inexhaustible and clean. It plays a very important role in the present energy crisis. Solar energy can be harnessed by hybrid photovoltaic thermal system to generate power and heat. These devices generate thermal and electrical energy simultaneously. Hybrid photovoltaic thermal systems have high efficiency. There is ample scope in this area as much work remains to be done. The hybrid system has huge potential in India where the availability of solar energy is spread throughout the country. The paper focusses on the study of hybrid photovoltaic thermal systems.
Keywords: Photovoltaic thermal; Hybrid solar; PVT performance.

1. Introduction

Climate change, global warming and environmental pollution have forced the world to look for clean energy technology. These problems can be solved by using renewable energy from sun and wind. All renewable power sources have their primary source in Sun. The motivating fact is that solar insolation rate received by earth is greater than the demand of fossil fuels. The earth intercepts about 1.8×10^{11} MW power from sun. This is greater than the present rate of all energy consumption. [1][2] Solar power is abundant, and its conversion is also less costly. Different methods for collecting solar energy in the usable form are being developed. Some methods focus on the use of solar PV, whereas others on thermal conversion. Solar PV works on energy of electrons excited by solar photons. Solar thermal depends on energy of water heated by sunlight. This hot water can be used for heating or electricity generation. [3][4][5] Solar cell converts only a fraction of solar energy (14-20%) to electricity through photovoltaic effect. Remaining energy is dissipated as heat energy. The efficiency of silicon solar cell decreases 0.45 % per ℃ . Panel temperature can reach 70 ℃ in hot summer days. This means panels will produce up to 20 % less power compared to their rated power capacity at 25 ℃. To overcome this, hybrid solar technology was developed. A hybrid solar concentrator system captures this waste energy and thus removes the waste heat from solar module. This also increases the efficiency of the combination remarkably. [6][7] The individual silicon cell has an efficiency of 25% whereas efficiency of solar PVT system can reach up to 70%. [8]

2. History of Solar PVT system

The history of solar PV began in 1839 after the breakthrough discovery of photovoltaic effect by Alexandre-Edmond Becquerel, a French physicist. Solar cells are made up of silicon which captures the photons emitted by sun and generate voltage.[9] The photoconductivity of selenium was then discovered in in 1873 by English electrical engineer Willoughby Smith. The first solar cell was designed in New York by Charles Fritts. It was done by coating selenium with a thin cover of gold. The efficiency of solar cell at the time of invention was mere 1% compared to the efficiency of modern solar cells at 20%. Silicon was found to be more efficient than selenium in 1953 by Bell Laboratories. And in this way production of solar cells took off commercially. The demand for solar power increased as the oil prices rose in 1970.This led to more research in the area and further reduction of cost. [10] The history of solar

thermal dates to 1839. There are records of solar heating in United States. Captain John Ericsson invented solar power engine in 1864. Solar Stirling Engine was driven by a parabolic concentrator. [11] In 1913,the first parabolic trough system was made for generating steam in Maadi, Egypt. The studies in PVT technology started in 1970. PVT technology has become an attractive energy solution for buildings as they require less space and provide more efficiency. [12] Hybrid PVT systems combine both the above idea.
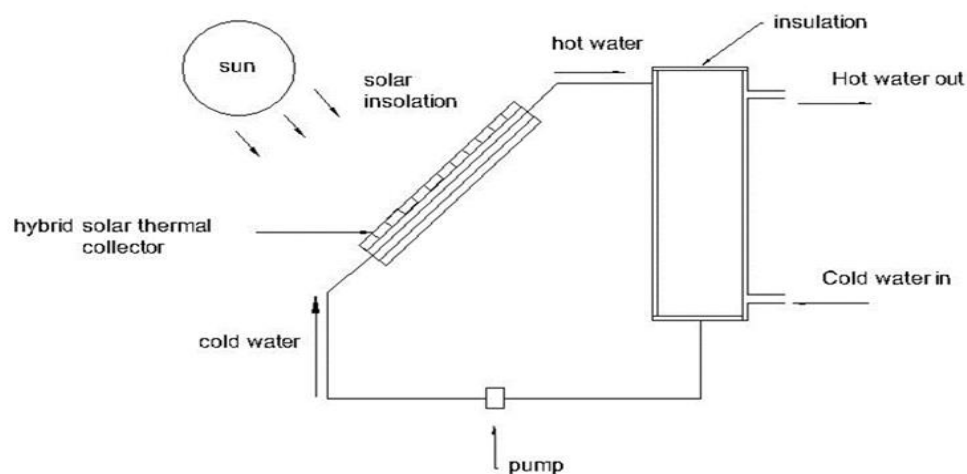

Figure 1. Solar thermal collector system

3. Hybrid solar PVT system

Hybrid photovoltaic solar concentrator is a combination of solar cell and solar thermal collector. The model of hybrid solar thermal collector system is shown in Fig. 1 above. The cold water is supplied from household to the hybrid collector where it gets heated from the waste heat from the hybrid system. Combining thermal and photovoltaic system helps reducing space, on site electricity generation and thermal generation. The detailed view of PV collector is shown in Fig. 2. The cover glass is used for protection and minimizing heat loss. Eva encapsulant is used to protect PV layer from shock and vibration. Thermal insulation is also provided beneath the PV layer to reduce the heat loss from the surface. Fuentes et al. [13] founded that the combined efficiency of hybrid PV/T system can reach 80% during the day. Bhattarai et al. [14] analysed a hybrid solar PVT collector and found that energy savings of hybrid PVT is more compared to conventional PV. Dupeyrat et al. [15] showed 79 % thermal efficiency and 8.7% electrical efficiency for hybrid solar PVT. The electrical efficiency of solar cells was found to be 15% for 90% of solar insolation absorbed. Remaining was rejected as heat. [16] The study on liquid type PVT systems were first done between 1978-1981. [17] Lovvik and Bergene [18] studied the effect of cooling on solar cell. Temperature in the range of 60 – 80 ˚C were found common for solar cell and cooling resulted in increased electrical efficiency of about 10-30 %.
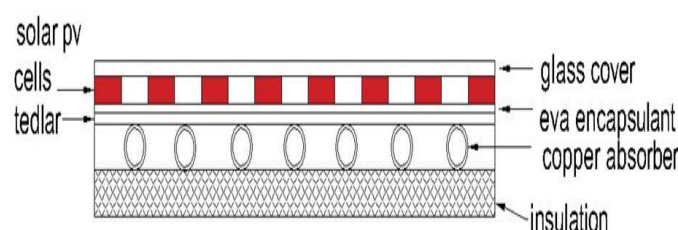

Figure 2. Detailed view of PV collector

4. Application of Hybrid PVT

Solar hybrid PVT liquid collectors have been used for domestic hot water application. Hybrid PVTs have been also integrated with buildings. Chow et al. [19] concluded that building integration for PVT is more popular than

that of water application. A 20 kWp PVT with air as thermal fluid is located at Mantaro Library in Spain. Concordia University, Montreal also developed a PVT system with 24.5 kW and 75 kW electrical and thermal power respectively. Beijing Olympic village also installed a PVT system with 10 kW and 20 kW of electrical and thermal power generation respectively. Fang et al. [20] concluded that the integration of hybrid PVTs with building is going to be most focussed area in the world. Global solar thermal capacity grew from 89×10^6 m^2 (2000) to 675 $\times 10^6$ m^2 (2017) in China. 3.78×10^{10} kg of oil and 1.2×10^{11} kg of CO_2 was saved due to this. Denmark, Sweden, Austria, Greece and Spain have large scale solar thermal plants since 1980. China has also recently installed a large no. of such systems.

4.1 Applications of Hybrid Photovoltaic Thermal Systems:

The major application of Hybrid PVT system is found in building integration. These systems can be used for pool water heating (25 – 35 °C), space heating (up to 60 °C) and domestic hot water. [21] Space heating and DHW constitute for nearly 50 % of heat demand. The hybrid systems also allow for the most efficient utilization of roof space. The total installation cost is also lesser when compared to individual installation of solar thermal unit and solar PV unit separately. The hybrid system can be also used in the following areas:

i. Building integrated photovoltaic thermal – BIPVT capacity has increased due to increase in grid connected photovoltaics. BIPVT is generally mounted on rooftop of building. The orientation of BIPVT should be such that solar insolation can be accessed. Irradiance and PV module temperature are important parameters for such system.

ii. Power generation – Hybrid PVT system can be used for power generation. Also, it was found that PVT plant is capable to produce more power than conventional plant. But the cost would be more. [22]

iii. Agriculture – PV pumps are most suitable application when plenty of sunshine is available. Hybrid PVT collector system with air can be also used for agriculture purpose or crop drying. Fterich et al. [23] experimentally studied a solar drier equipped with solar PV/T air collector. The hybrid PVT drier shortened the drying time and increased the efficiency. Barnawal and Tiwari [24] performed grape drying using hybrid photovoltaic greenhouse drier at solar energy park, New Delhi.

iv. Street lights – Hybrid system can be used for lighting streets, yards, compounds and industrial roads.

v. Concentrating PVT – These systems have been used in hospitals and schools. Swedish manufacturer Absolicon installed concentrating PVTs in a private hospital in Harnosand for power. The Cogenra building installation at the university of Arizona Tech Park supplies 191 kW and 36 kW thermal and electric power respectively. [25]

vi. Solar distillation – The overall thermal efficiency of hybrid solar PVT still designed Kumar and Tiwari [26] was found to be 20 % more than the conventional solar still. The electrical efficiency of the former was also found to be higher. Pounraj et al. [27] found 38% efficiency rise for Peltier based hybrid PVT active solar still as compared to conventional PV still.

5. Performance of Hybrid Solar PVT

Based on working fluid, there are two main types of hybrid collector. These are water-based and air-based hybrid collectors. Air has some advantage over water since it does not freeze and there is no damage if leakage occurs. But the heat capacity and conductivity of air is lower. PVT air collectors are generally used in buildings. The overall performance of the hybrid system is the sum of thermal efficiency and electrical efficiency. [28]

$$\eta_{PVT} = \eta_{th} + \eta_{PV} \tag{1}$$

5.1 Thermal Performance

5.1.1 Solar module temperature

$$t_{mo} = t_{amb} + (NOCT\text{ -}20) \times \frac{G}{800} \tag{2}$$

where NOCT the temperature of module in an environment with solar irradiation of 800 W/m^2, 20°C ambient temperature and wind speed of 1 m/s. The heat can be removed from the cell by different modes of heat transfer. The heat removed can be utilized for different purposes. The thermal efficiency of hybrid system and other parameters can be obtained as:

$$\eta_{th} = \frac{\dot{m} \times c_p \times (t_{out} - t_{in})}{a_{pvt} \times G} \tag{3}$$

The loss of heat from the reservoir to the surrounding is:

$$Q_{loss} = \int U_t (t_{in} - t_{amb}) \, dt \tag{4}$$

Figure 3. Solar thermal systems on houses in Salzburb, Austria [29]

5.2 Electrical efficiency

The electrical efficiency is given as:

$$\eta_{el} = (f') \frac{i_{sc}.U_{oc}}{G.s} \tag{5}$$

$$f' = \frac{P_{max}}{I_{sc}.U_{oc}} \tag{6}$$

The energy conversion efficiency of the PVT system is:

$$\eta_{PVT} = \frac{w+q}{h} = \eta_{el} + \eta_{th} \tag{7}$$

5.3 Energy saving efficiency

The energy-saving efficiency of the pvt system is given by

$$E_{sa} = \frac{\eta_{el}}{\eta_{power}} + \eta_{th} \tag{8}$$

6. Losses in solar thermal system

Losses in solar thermal system are mainly optical and thermal. The solar thermal system efficiency keeps changing due to changing solar and weather condition. [30]

 i. Cosine loss – These losses occur due solar radiation not being perpendicular to reflector.

 ii. Shading and blocking – Losses within the system due to blockage by individual elements

 iii. Reflectance – The reflector does not allow all the incident light to pass through it

 iv. Cleanliness – Dirt and dust from atmosphere accumulate on the surface of reflector and decrease efficiency.

 v. Incident angle modifier (IAM) – Optical parameters generally degrade with increase in incident angle.

 vi. End losses – Some part of receiver does not receive radiation from solar collector.

 vii. Intercept – These losses occur due to geometrical inaccuracies.

 viii. Shielding by bellows – These losses occur due to bellows shielding.

 ix. Transmittance of glass – These losses are depending on the nature of glass material.

 x. Absorbance of receiver – The receiver surface coating absorbs some radiation.

 xi. Thermal losses – This is caused by convection and radiation.

 xii. Thermal losses due to piping – These losses occur due to conduction and convection.

7. Conclusion

There has been huge research in the field of hybrid system since 1970. Hybrid PVT system transforms solar insolation into usable heat and power. The hybrid system is a combination of solar PV and thermal collector. Solar PV converts solar energy into electric power whereas thermal collector removes waster heat from the solar PV. This results in the increase in the efficiency of the hybrid system. The electrical efficiency of solar PV is about 10% without cooling. The conversion efficiency of hybrid PVT systems can reach up to 80 %. The hybrid system can be very useful for developing country like India which has huge potential of solar energy spread throughout the country.

Nomenclature

η_{el}	electrical efficiency	$\dot{m}$	mass flow rate (m^3/hr)
η_{power}	efficiency of electrical power generation	G	Irradiance on collector surface (W/m^2)
t_{out}	outlet temperature of collector (ºC)	PVT	photovoltaic thermal
t_{in}	inlet temperature of collector (ºC)	PV	photovoltaic
c_p	specific heat (J/kg/K)		
η_{th}	thermal efficiency	S	heat absorber plate area (m^2)
IAM	incident angle modifier	t_{amb}	ambient temperature (°C)
NOCT	nominal operating temperature (°C)	t_m	temperature of module (°C)
SRE	standard reference environment	h	energy input (J)
U_t	overall heat transfer coefficient of reservoir (W/m^2K)	w	mechanical work (J)
q	heat (J)	i_{sc}	short circuit current (A)
E_{sa}	energy saving efficiency	f'	fill factor
s	surface area of the unit (m^2)	P_{max}	Maximum power (W)
G	solar irradiance (W/m^2)	a	area (m^2)
U_{oc}	open circuit voltage (V)	η_{power}	electric power generation efficiency
η_{PV}	photovoltaic efficiency	Q_{loss}	heat loss (J)
eva	Ethylene-vinyl acetate		

Subscripts

el	electrical		
out	outlet	amb	ambient
in	inlet	max	maximum
th	thermal	avg	average
mo	module		

References

[1]   Hannah L 2015 Mitigation: Reducing Greenhouse Gas Emissions, Sinks, and Solutions Climate Change *Renewable Energy* **2(1)** 379-402

[2]   Paridaa B, Iniyanb S and Ranko G 2011 A review of solar photovoltaic technologies *Renewable and Sustainable Energy Reviews* **15(3)** 1625-1636

[3]   Thomas N 2003 Field Guide to Appropriate Technology *Energy* 157-275

[4]   Bhatia S 2014 Photovoltaics Energy resources and their utilisation. *Advanced Renewable Energy Systems* **31(6)** 1-31

[5]   Belz K, Kuznik F, Werner K, Schmidt T and Ruck W 2015 Thermal energy storage systems for heating and hot water in residential buildings Woodhead Publishing Series in Energy *Advances in Thermal energy storage system* **12(5)** 441-465

[6]   Abdulyan M and Chiasson A 2017 Thermal Study of Hybrid Photovoltaic-Thermal (PVT) *Energy Procedia* **141(3)** 102–108

[7]   Du D, Darkwa J and Kokogiannakis G 2002 Thermal management systems for Photovoltaics (PV) installations: A critical review *Solar Energy* **97(6)** 238–254.

[8]   Adria E. Brooks 2014 Solar Energy: Photovoltaics. Future Energy Improved, Sustainable and Clean Options for our Planet **4(3)** 383-404

[9]   http://www.pvresources.com/en/introduction/history.php

[10]   https://www1.eere.energy.gov/solar/pdfs/solar_timeline.pdf

[11]   Kongtragool B and Wongwises S 2003 A review of solar-powered Stirling engines and low temperature differential Stirling engines *Renewable Sustainable Energy* **7(4)** 131–54

[12]   Aidroos D, Hasimah B, Rahman A, Zaidi W, Omar W and Obaid S 2015 Historical development of concentrating solar power technologies to generate clean electricity efficiently *Renewable and Sustainable Energy* **41(3)** 996-1027

[13]   Fuentes J, Vivar M, Casa J and Aguilera J 2018 An experimental comparison between commercial hybrid   PV/T and simple PV systems intended for BIPV, *Renewable and sustainable Energy Reviews* **93(1)** 110-120

[14]   Bhattarai S, Ohb J, Euh S, Kaflea G and Kima G 2012 Simulation and model validation of sheet and tube type photovoltaic thermal solar system and conventional solar collecting system in transient states *Sol Energy Matter Sol Cells* **103(5)** 184–93

[15]   Dupeyrat P, Ménézo C, Wirth H and Rommel M 2011 Improvement of PV module optical properties for PV-thermal hybrid collector application *Sol Energy Mater Sol Cells.* **6(1)** 2028–36

[16]   Dupeyrat P, Menezo C and Fortuin S 2015 Study of the thermal and electrical performances of PVT solar hot water system *Energy and buildings* **68(5)** 751-755

[17]   Hendrie S 1982 Photovoltaic/thermal collector development program. Final report

[18]   Bergene T and Lovvik O 1995 Model calculations on a flat-plate solar heat collector with integrated solar cells *Sol. energy* **55 (6)** 453-462

[19]   Chow T, Tiwari G and Menezo C 2012 Hybrid solar: A review on photovoltaic and thermal power integration *International Journal of Photoenergy* 220-227

[20]   Fang X and Li D 2013 Solar photovoltaic and thermal technology and applications *Renewable and Sustainable Energy Reviews* **23** 330-340

[21]   Mellor A et al. 2018 Roadmap for the next-generation of hybrid photovoltaic-thermal solar energy collectors *Solar Energy* **174** 386–398

[22]   Gagliano A et al. 2019 Comparative assessments of the performances of PV/T and conventional solar plants *Journal of Cleaner Production* **219** 304-315

[23]   Fterich et al 2018 Experimental parametric study of a mixed-mode forced convection solar dryer equipped with a PV/T air collector *Solar Energy* **171** 751–760

[24]   Barnwal P and Tiwari G 2008 Grape drying by using hybrid photovoltaic-thermal (PV/T) greenhouse dryer: An experimental study *Solar Energy* **82** 1131–1144

[25]   Good C et al 2015 Hybrid Photovoltaic-thermal Systems in Buildings – A Review. *Energy Procedia* **70** 683-690

[26]   Kumar S and Tiwari A 2010 Design, fabrication and performance of a hybrid photovoltaic/thermal (PV/T) active solar still *Energy Conversion and Management* **51** 1219–1229

[27]   Pounraj et al. 2018 Experimental investigation on Peltier based hybrid PV/T active solar still for enhancing the overall performance *Energy Conversion and Management* **168** 371-381

[28]   Radziemska E 2009 Performance Analysis of a Photovoltaic-Thermal Integrated System *International Journal of Photoenergy* 80-233

[29]   https://www.iea-shc.org/Data/Sites/1/publications/Solar-Heat-Worldwide-2018.pdf
[30]   Hermann et al 2017 The Performance of Concentrated Solar Power (CSP) Systems Modelling, Measurement and Assessment *Solar Energy* 115-165

PAPER • OPEN ACCESS

Two Finger Gesture Imitating Animatronic Hand

View the article online for updates and enhancements.

Two Finger Gesture Imitating Animatronic Hand

Y Verma and V Rastogi

Mechanical, Production and Industrial and Automobile Engineering Department, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi-110042, India

E-mail: yashasviverma22@gmail.com

Abstract. This paper poses the discussion regarding animatronics that has been aided with modern trends in technological innovations. The ease of availability and procurement of present day items incorporate economical factors into research. To accomplish the aforementioned, a simple two fingered hand model has been developed that can be remotely controlled using a glove. The glove can be manipulated via a hand. It has been electronically modulated using Arduino Uno and servo motors. The discussion has been furthered to the amalgamation of future developments that can pave the way for commercialization of such devices. Hardships and challenges faced during building of the model and experimentation of the finalized model are detailed.

1. Introduction

The word "Animatronics" entomologically merges "Animate" and "Electronics". Thus, as suggested it is the technology behind animating objects to imitate humanoid gestures and actions. It can be commonly observed in motion pictures beginning with Disney's Mary Popins. The concept dates back to 3rd Century BC where it finds mention in the script of Liezi (Chinese Scripture). Its first modern day usage was for novelty clocks that were adorned with automated birds in 12th century and with the rapid advances in robotics, anatomy and mechantronics has expanded its scope to a multitude of areas such as prosthetics, industrial robotics and anthropomorphism. Robotics is advancing from remote location usage to being integrated in technology that we interact with on a daily basis. A way to humanize these robot-human interactions is to make robots imitate human gestures. Animatronics has been a step in this direction. There are many prevailing industry situations where human contact can prove to be dangerous or even lethal such as handling of hazardous chemicals, contacting Super-hot or Super-cold fluids or even entering narrow enclosures. A robotic industrial hand can help to improvise this situation to perform human-like actions of inspection or handling of injurious materials. Mishra et. al. [1] discuss the exploitation of Bluetooth feature in smartphones to transmit gesture information for remote control of industrial robots. They cancel the processing time lag enhancing the advantages of distant controllers. For specially-abled persons, animatronics is the basic step towards the development of prosthetic limbs. A hand that is capable of delivering humanoid movement to an artificial limb using electrical sensor can be a boon to many. This is achieved by anthropomorphizing the advanced animatronic limbs and extending the expanse of sensors. Kumar and Kumari [2] present a review of

the control features of animatronically powered prosthetic hands. They elaborate upon the current complications in prosthesis and develop on the various technologies in utilization at present. The utilization of animatronics is expansive, as explained by Oo et. al [3] in their innovative approach to reinforce the aid provided to the deaf and dumb by incorporating a gesture governed robotic arm. A major consideration during such a development is also keeping the financial aspect in check as technology has to be affordable to even small scale industrial houses and economically weaker section of physically challenged people.

There has been present day sophisticated research on such a concept. Pradeep and Paul [4] illustrate the gesture based control of a robotic arm. They achieve this by using accelerometer to sense the gestures that are then processed and transmitted to the robotic arm using an RF transceiver module. They generate the user friendly and accurate model to aid human intervention in industrial processes. Crisman et. al [5] present their innovative technique regarding the design of a robot hand. For this they incorporate the functioning of the tendon structure of a human hand into their mechanical model. They optimize the work freedom available to fingers by developing their model with pulleys and algorithm for control using switches. Ramaswamy and Deborah [6] lay out a comprehensive research on the contemporary robotic hands that are commercially accessible. They include in the descriptive survey and the advancements along with the multitude of problems encountered. Jog et. al [7] worked on a wireless animatronic had that coalesces the modern technology of wireless communication using a XBEE module and the programming ease of Arduino Uno as a microcontroller. The result was a robotic hand with possible distant interaction but with scope of future effort to make movements more precise. Similarly, Gupta et. al [8] talk of the scope of extending the wireless robotic hand technology to a prototype arm. They dwell on the use of XBEE as an interface for wireless communication and involve the sophisticated signal processing of Arduino Uno. Utture et. Al [9] dwell on the concept of animatronics and its timeline. They discuss the amalgamation of robotics and animatronics as well as the prescribed methods for using these rapidly developing innovations. Bilgin et. at [10] has written about the purposeful usage of flexible sensors on finger tips of controller gloves to increase the precision of the animatronic movement. They discuss the use of RF 435 MHz module for wireless functions and embedding of the system with a digital interface for the proper implementation of the same. Chowdhury [11] has worked up the basic model for an animatronically controlled hand. She employs Arduino Uno and servo motors for developing communication between the controller and receiver. She goes on to discuss the economic limitations due to usage of flex sensors, the unnecessary employment of excessive sensors and the loud servo motor sounds.

In this paper, a basic model of an animatronic hand is developed. The basic outline of the logic employed is depicted in figure 1. The controlling glove has been provided with flex sensors to transmit signals corresponding to motion in the fingers of the person who dorns the glove. Syed et. al. [12] have explored in detail the mapping of a hand motion to subsequent robot model. The wireless communication basis has been concluded that enhances the employment of the mentioned robots. The signal communicated via the sensors is converted by an Arduino Uno and sent to the servo motors which control the sample tendon structure mounted on the robotic hand. This paper serves to discuss the socio-economic impact factor with the improvement in present day technology. The challenges faced with the use of flex sensors have been addressed and the ease of making the basic animatronic research available worldwide has been developed.
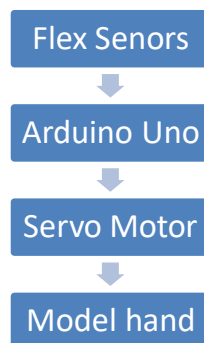
Flex Senors

↓

Arduino Uno

↓

Servo Motor

↓

Model hand

Figure 1. Flow chart of the methodology employed

2. Methodology, Design and Experimental Setup

2.1. Design Specifications

The gesture imitating hand was divided into two sections, one being the controller part or the part consisting of the glove to be worn by the controller and the Arduino Uno microcontroller to convert the flex sensor signals to signals compatible with servo motor movement, the other part being the mimicking part or the model hand that mimics the movements as gestured by the controller and the servo motors that cause the motion in the model fingers. The design of entire system began from the controller part. An everyday usage glove was procured. The glove should be a tight enough fit on the hand of the controller so that the flex sensors can precisely detect any movement of the fingers. For the basic model, two fingers ie. the thumb and the ringer finger were taken into consideration. The flex sensors were attached to the abovementioned fingers. These flex sensors were electrically connected to a bread board and then an Arduino Uno. The microcontroller was fed information from a laptop and was connected to two servo motors through another bread board. The servos were fed two strings which were attached to the model hand. In this case, the two-fingered hand was modeled using cardboard. The fingers were capable of bending in three stages like an anthropoid and the strings were fastened across the finger length.

2.1.1. Flex Sensors. Flex sensors are devices that convert the angle at which they are bent to a change in resistance, with the increase in the bending angle, the resistance of the device increases. It is usually connected across a voltage divider circuit that converts this increase in resistance to a proportionally linear increase in voltage. This voltage signal can be easily manipulated for further usage. For this project, two 2.2 inch flex sensors were procured. The two terminals apart from being present in a voltage divider circuit were connected to the ground terminal and the terminal of a power source present on the Arduino board.

2.1.2. Resistors. Two resistors were employed to successfully convert the resistance signal from the flex sensor to a voltage signal for downstream ease of manipulation. For this, two resistors were acquired with the Red Red Brown Gold colour code that translates to 220 Ohm.

2.1.3. Arduino Uno. This is a microcontroller board that employs 6 input analog pins and 14 digital I/O pins and is easily programmable using the Arduino IDE which is an application that is cross platform and can be worked upon with the basic knowledge of C or C++ languages with just a few changes in the structuring. The board is USB base. It was used to convert the analog voltage signal input to digital signal output compatible with the servo motors.

2.1.4. Bread Boards. They are convenient and cost effective devices to produce circuits using jumper wires without having to solder. This effectively reduces the connection time and also professes safety. The financial aspect is another advantage as they are inexpensive and widely available. They come in

a variety of sizes. For this model, a full (6.5cm X 17cm) and half size (5.5cm X 8.5cm) breadboard were utilized.

2.1.5. Servo Motors. Since the model employs two fingers to configure a basic circuit. Two SG90 servo motors were used. A servo motor has 3 terminal wires which are brown, red and yellow in colour. The brown wire is connected to ground terminal and yellow to the respective output terminals of the digital section of the Arduino Board. The red wire in connected to an external battery source, which in this case was taken as a 9V battery. Two strings that acted as tendons were mounted on the Servos. The servos convert the signal to the desired angle of the blade. This exerts a pull on the strings which subsequently move the fingers of the model hand.

2.2. Methodology
There are two major sectors that required focus, one being the electrical circuit and the other being the program fed into the Arduino.

2.2.1. Electric Circuit. While developing the circuit, the main aim was to manipulate the signals received from the flex sensors into angles of movement of the Servo motors that would invariable pull on the tendonic strings to imitate the movement in the model hand. Starting from the controller section, the basis of the electrical circuit was to join a terminal of both the flex sensors to the ground terminal of the Arduino via the negative terminal of the breadboard. This has been depicted by the black coloured circuit lines emerging from the Flex Sensors in figure 2. The other terminal of the flex sensors were supposed to be connected to a fixed voltage source via the positive strip of the breadboard as shown as the red coloured lines originating from the flex sensors in figure 2. On the same connection, resistance drop had to be added to develop a voltage divider circuit which was then connected to the 2 analog input terminals of the Arduino board. This can be evidently followed along the brown coloured circuit lines concluding in the T1 and T2 input terminals of Arduino Board in figure 2. Moving to the controlled part of the prototype, the digital I/O terminals of the Arduino board were then connected to one terminal of the Servo. These brown connection lines (figure 2) serve the purpose of transmitting the manipulated signals to the Servo motors to move the hand extension. The other two terminals were connected to negative terminal of the battery as well as the ground terminal of the Arduino board and positive terminal of the battery. These can be respectively traced as the black and red coloured circuit lines traversing towards the battery respectively in figure 2. The servos were connected using a tendon functionality thread. A constant power source was required to locomote the modeled two finger hand. This was provided through the Arduino Uno being connected to a power source, abetting the red power lines in figure 2.
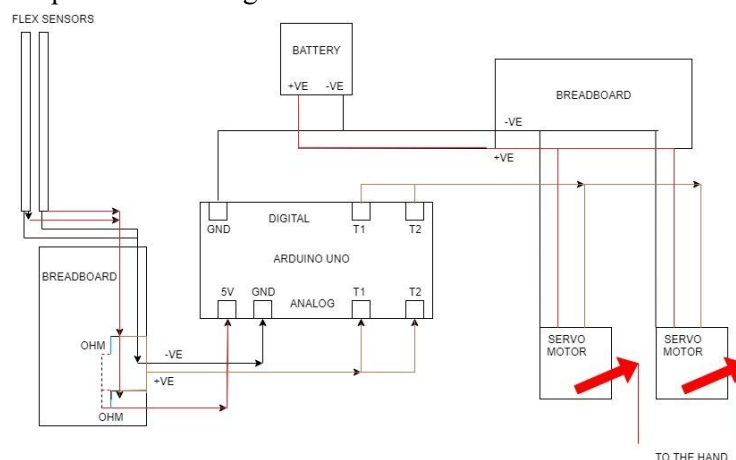

***Figure 2.** Electronic circuit used for development of design*

2.2.2. Arduino Program. Figure 3 shows the code which was fed into the Arduino IDE and uploaded into the microcontroller. The code helped change the voltage analog signal received into a change in angle of the attachment rod mounted onto the servo. As the angle of the rod changed it exerted a pull onto the thread which further caused bending or flexing of the fingers of the model arm. Badamadsi [13] elaborates the functionality principles of Arduino Uno as a microcontroller by expanding on not only the hardware but the software as well. The method of compiling codes for further project development have been thoroughly examined as well.

```
#include <Servo.h>
Servo servo_motor1;
Servo servo_motor2;
int flex1=0;
int flex2=1;
void setup()
{   servo_motor1.attach(9);
  servo_motor2.attach(10);
}
void loop()
{
    int flexpos1, flexpos2, servopos1, servopos2;
    flexpos1 = analogRead(flex1);
    Serial.println(flexpos1);
    servopos1 = map(flexpos1, 1020, 1023, 0, 180);
    servopos1 = constrain(servopos1, 0, 180);
    flexpos1 = analogRead(flex1);
    servopos1 = map(flexpos1, 1020, 1023, 0, 180);
    servopos1 = constrain(servopos1, 0, 180);
    flexpos2 = analogRead(flex2);
Serial.println(flexpos2);

  servopos2=map(flexpos2, 1020, 1023, 0, 180);
    servopos2 = constrain(servopos1, 0, 180);
     servo_motor1.write(servopos1);
    servo_motor2.write(servopos2);
}
```

Figure 3. The Arduino code fed to Arduino IDE

3. Experiments and Results

The controller having access to the glove flexed his thumb and ring finger, first one at a time and then together. The model hand was seen performing the same activities. On checking the angle of the finger flex in case of the controller and the model hand, the difference between the angles was more than 15%. The error in precision of movement could be attributed to the fingers made out of proportion. If the length of the fingers of the animatronic hand are made proportional to real life fingers and the length of each section of the finger too made more anthropomorphic, this could reduce the error down to the scientifically acceptable level. These levels are set by each industry or phase of utility that demand the employment of such technology. A possible solution to such an issue can be the amalgamation of cost effectiveness and anthropomorphism. Bae et. al. [14] embellish such a concept to ensue upon a four fingered hand equipped with actuation. They dwell on the viability of controlling objects and the versatility of the applied science. At present, the resistors had to be replaced with ones having ohmic capacity of 22k Ohm. The reason for replacement being that the low level resistance caused early drainage of the 9V battery. If a long lasting source is available at the same level then 220

ohm resistors can be used. But, for optimized results replacement is advised. Chowdhary [11] expatiated over the probable use of Neoprene bend sensors for cost effective applications. With rapid globalization accompanied with torrid advances in technology, flex sensors are now easily obtainable at much lower costs. They can also be developed at home for more cost saving. Their easily availability denoted ease of procurement and hence time saving. Flex sensors are more sensitive than bend sensors as the later work on sensing pressure. As we are aiming towards a more precise model for industrial applications and enhancing gesture sensitivity, flex sensors should be the way to go. They no longer pose a cost barrier. Using a single flex sensor for more than one finger may not be the proper path because each finger has its own characteristic bend and flex. This bend has to be mimed accurately for furthering the project applications. The Servos were mounted on acoustic foam. This foam is capable of absorbing noise created by the Servos. This effectively solves the issue of lowering noise levels on the industrial floor when the animatronic hand advances to such employment. The project was developed at a low cost and hence suggested that with newer technology, the impact of such a device can be re-evaluated for the better.

4. Discussion

Technology is the science that can be made practically available to its beneficiaries. The gesture imitating basic animatronic hand was developed to check its socio economic standing with the advancements in present day technology. Be it the ease of component availability or the cost of their procurement, the hand was developed cost-effectively. This feature can be exploited in later stages where the hand can be anthropomorphized and made more precise. The time to develop such a project is also optimized with the worldwide obtain-ability of instruments. Any noise in the system is also efficiently reduced by deploying acoustic sponges which are again not a very substantial investment. The breadboards are definitely a superior substitution to soldering in terms of safety and material cost. These implications of the technical project talk bounds about its future evolution into more sophisticated models. The problem of precision in movement persist which remain to be addressed. In conclusion, the desired model and its predicted financial importance were achieved. The subsequent research can deal with the rectification of precision errors bearing in mind the reduction of possible inflation in costs that comes when using futuristic technology. The animatronic hand holds the potential for betterment of human-robot interaction.

5. Conclusion

The paper has developed a basic model for the gesture enabling remote control in an animatronic two-fingered hand. The problems faced during the concentrated effort on socio-economic steer as furnished by modern technology were elaborated upon. Certain possible ameliorations were suggested bearing in mind the basic aim of the research. Furtherance of the model may include the introduction of mechanisms for flat object picking and dynamics of humanoid object transfer.

6. References

[1]   Mishra S, Tamboli S M, Yamgar M M and Prakash R 2017 Smart device based gesture controller for industrial applications *International Conference on Nextgen Electronic Technologies: Silicon to Software* 463-7
[2]   Kumar V V N and Kumari P N 2015 A review on embedded control of animatronic prosthetic hand *International Journal for Technological Research in Engineering* **2** 614-7
[3]   Oo H M, Tun K T and Thant M L 2019 Deaf sign language using automatic hand gesture robot based on microcontroller system *International Journal of Trend in Scientific Research and Development* **3** 2132-6
[4]   Pradeep J and Paul P V 2016 Design and implementation of gesture controlled robotic arm for industrial applications *International Journal of Advanced Scientific Research and Development* **3** 202-9

[5]   Crisman J D, Kanojia C and Zeid I 1996 Graspar: a flexible, easily controllable robotic hand *IEEE Robotics and Automation Magazine* **3** 32-8

[6]   Ramaswamy C V V and Deborah S A 2015 A survey of robotic hand-arm systems *International Journal of Computer Applications* **109** 26-31

[7]   Jog S, Dwivedi A, Ashtankar S and Gautam G 2016 Animatronic hand using wireless module *International Research Journal of Engineering and Technology* **3** 1529-31

[8]   Gupta A, Jain R, Bang D and Kori J 2016 Wireless animatronic arm *International Conference and Workshop on Electronics and Telecommunication Engineering* 105-11

[9]   Utture O, Mane N S, Varute S S and Waingade A B 2017 Multidisciplinary technology of animatronics : a review *Asian Review of Mechanical Engineering* **6** 4-7

[10]   Bilgin S, User Y and Mercan M 2016 Robotic hand controlling based on flexible sensor *International Journal of Engineering and Applied Science* **8** 49-58

[11]   Chowdhury S 2013 Animatronic hand controller 10.13140/RG.2.2.25148.74889

[12]   Syed A, Agasbal Z T H, Melligeri T and Gudur B 2012 Flex senor based robotic arm controller using micro controller *Journal of Software Engineering and Applications* **5** 364-6

[13]   Badamasi Y A 2014 The working principle of an arduino *11th International Conference on Electronics, Computer and Computation* 1-4

[14]   Bae J, Park S, Park J, Baeg M, Kim D and Oh S 2012 Development of a low cost anthropomorphic robot hand with high capability *International Conference on Intelligent Robots and Systems* 4776-82