

PERFORMANCE ANALYSIS OF IMAGE FUSION ALGORITHMS FOR BIO-MEDICAL APPLICATIONS

**A Thesis Submitted in Partial Fulfillment of the Requirements for
the Degree of**

DOCTOR OF PHILOSOPHY
by

TANIMA GHOSH
(2K19/PHDEC/506)

Under the Supervision of

DR. N. JAYANTHI

Department of Electronics and Communication Engineering

Delhi Technological University



DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daultpur, Main Bawana Road, Delhi-110042. India

December, 2024



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I Tanima Ghosh hereby certify that the work which is being presented in the thesis entitled ' Performance Analysis of Image Fusion Algorithms for Bio-Medical Applications' in partial fulfillment of the requirements for the award of the Degree of Doctor of Philosophy, submitted in the Department of Electronics and Communication Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from January 2020 to December 2024 under the supervision of Dr. N. Jayanthi.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Candidate's Signature



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daultpur, Main Bawana Road, Delhi-42

CERTIFICATE

Certified that **TANIMA GHOSH** (2K19/PHDEC/506) has carried out her research work presented in this thesis entitled **“Performance Analysis of Image Fusion Algorithms for Bio-Medical Applications”** for the award of **Doctor of Philosophy** from Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, under my supervision. This thesis embodies results of original work, and studies are carried out by the student herself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University.

Dr. N. Jayanthi
Assistant Professor
(Supervisor)
Department of ECE
Delhi Technological University.
Delhi-110042, India



DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daultpur, Main Bawana Road, Delhi-42

ACKNOWLEDGEMENTS

First and foremost, thanks to the Almighty for giving me strength and inspiration to carry out this research work. I owe a deep sense of gratitude to all the comprehensive soul whose divine light has enlightened my path throughout the journey of my research.

Words cannot express my gratitude towards my supervisor Dr. N. Jayanthi, Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India. I greatly acknowledge her invaluable patience, feedback, guidance, and continuous motivation throughout my journey. I extend my sincere regards to the HoD sir for his constant support. I wish to express my gratitude towards the DRC chairperson, the distinguished DRC members, and the esteemed faculty members for their valuable time and efforts throughout this journey. I am also grateful to all the Senior professors, lab assistants and seniors of the Electronics and Communication Engg. department. I also express my thanks to all the staff members of the department for their continuous support in our academic activities. Lastly, I would like to thank my family for their kindness, support, and patience throughout my journey. Their belief in me has kept my spirits and motivation high during this process.

TANIMA GHOSH

This Thesis is Dedicated to my

Father and Mother.

*For their endless love, support and
encouragement*

ABSTRACT

Image processing plays a vital role in various medical applications, including image-guided surgery, disease progression studies, medical diagnostics, and radiotherapy. Extracting clinically significant and useful information from medical images is crucial across these applications. Medical images provide unique insights into the internal organs of the body. However, the information obtained from a single-modality medical image is often limited and may not fully meet the needs of an accurate clinical diagnosis. To address this limitation, multimodal medical image fusion has emerged as a promising solution, offering information about an organ from different perspectives. Medical image analysis involves four key steps: 1. Filtering the image 2. Segmenting the relevant regions 3. Extracting features and 4. analyzing these features using a pattern recognition system or classifier.

Image preprocessing is a critical step before fusing images from different modalities for reduction of unwanted noise to improve the visual quality of the image. Image denoising after fusion is also crucial for improving the quality, interpretability, and usefulness of the fused images in medical diagnostic and treatment planning.

- The primary objective is to develop multimodal medical image fusion algorithms that enhance visual details for improved clinical diagnosis and to address the problem of requirement of large dataset and computational complexity. Hence, we developed Efficient image fusion models using Dense-ResNet on BraTs 2020 and innovative fusion rules on BraTs 2015, 2018 and Harvard Medical School Brain Dataset. Also, the performance analysis of DWT based image decomposition has been done along with the proposed energy-based coefficient enhancement.
- Denoising of medical images is an essential preprocessing technique for enhancing the performance of the fusion model. Hence, we developed AMT-DWT (Averaging of multiple technique–Discrete Wavelet Transform) based pre-processing. In order to increase the clinical applicability of medical images for diagnosis, custom CNN based denoising has been utilized after fusion. The proposed model is benchmarked against a recent CNN-based image fusion model, TDAN (Two Level Dynamic Adaptive Network), as well as several state-of-the-art image fusion models using the CT-MRI images from the Harvard Medical School Brain dataset. The proposed model outperformed the state-of-the-art image fusion models based on the performance metrics PSNR, RMSE, SSIM, MI, Entropy and $Q^{AB/F}$.

- Additionally, we developed an efficient brain tumor detection and classification model using EADF (Extended Anisotropic Diffusion Filtering), and ResNet-50 with proposed auto thresholding for the classification of medical images based on the specific structures or diseases using the datasets BR35H, BMI-I, BTI, BD_BT, and BraTs. Also, we proposed WMRESNET (Weight Modified Res Net) model for the detection and classification of breast cancer. Promising results with less computational costs indicates the efficacy of the suggested models.

Table of Contents

Candidates' declaration

Certificate

Acknowledgements

Abstract

Table of Contents

List of Figures

List of Tables

List of Abbreviations

1. CHAPTER 1: NTRODUCTION.....	1
1.1 Modalities of Medical Images.....	1
1.2 Feasibility of MMIF.....	3
1.3 Applications of MMIF.....	4
1.4 Image Denoisation.....	4
1.5 Medical Image Classification.....	5
1.6 Research otivation.....	6
1.7 Problem Formulation.....	6
1.8 Methodology.....	7
1.9 Major Contributions.....	9
1.10 Thesis Organization.....	10
 2. CHAPTER 2: LITERATURE SURVEY AND PERF.ORMANCE	
ANALYSIS OF IMAGE FUSIO MODEL.....	11
2.1 Literature Survey on Image Fusion.....	11
2.2 Literature Survey on Image Denoisation.....	16
2.3 Literature Survey on Image Classification.....	17
2.4 Chapter Summary and Gaps in the Study.....	19
2.5 Objectives and Scope of the Work.....	20
2.6 Performance Analysis.....	20
2.6.1 Performance Analysis of different Transform techniques of MMIF.....	21

2.6.2	Calculated Fusion performance metrics	22
2.6.3	Results and discussion.....	24
2.7	Performance analysis of MMIF with the proposed energy-based coefficient enhancement technique.....	26
2.7.1	Methodology.....	26
2.7.2	Results and Discussion.....	29
2.8	Conclusion.....	32
3.	CHAPTER 3: IMAGE DENOISATION.....	33
3.1	Introduction.....	33
3.2	AMT-DWT based Pre-processing with custom CNN based fused image denoisation.....	34
3.3	Methodology.....	35
3.4	Results and Discussion.....	42
3.5	Conclusion.....	56
4.	CHAPTER 4: EFFICIENT IMAGE FUSION.....	57
4.1	Introduction.....	57
4.2	An efficient Dense-ResNet for multimodal image fusion using medical images.....	58
4.2.1	Methodology.....	59
4.2.2	Results and Discussion.....	70
4.3	Multimodal fusion of different medical image modalities using optimized hybrid network	73
4.3.1	Methodology.....	74
4.3.2	Results and Discussion.....	77
4.4	A pair of image fusion rule that can fuse image effectively.....	78
4.4.1	Methodology.....	78
4.4.2	Results and Discussion.....	79
4.5	Conclusion.....	83
5.	CHAPTER 5: DEVELOPMENT OF IMAGE CLASSIFIER.....	84
5.1	Introduction.....	84
5.2	An efficient method for brain tumor detection and classification by Extended Anisotropic Diffusion filtering	85
5.2.1	Methodology.....	85
5.2.2	Results and Discussion.....	91

5.3	An efficient classifier for breast cancer classification by WMRESNET	99
5.3.1	Methodology.....	99
5.3.2	Results and Discussion.....	104
5.4	Conclusion.....	107
6.	CHAPTER 6: CONCLUSION AND FUTURE SCOPE.....	108
6.1	Summary of work done in the thesis.....	108
6.2	Future Scopes	109
	DATA DESCRIPTION.....	110
	PUBLICATION DETAILS.....	112
	REFERENCES.....	113

List of Figures

Fig.1.1. Brain MRI modalities.	3
Fig.1.2. Multimodal Medical Image Fusion methodology design.....	3
Fig.1.3. Fusion of CT and MRI Brain Images	4
Fig.2.1. Block Diagram of Daubechies filter banks.....	21
Fig.2.2. Haar wavelet function.....	22
Fig.2.3. Diagram of Lifting Scheme	22
Fig.2.4. Fused images obtained from Daubechies, Haar and lifting scheme.....	25
Fig.2.5. Comparison of Daubechies, Haar and Lifting scheme based on PSNR value.....	26
Fig.2.6. Calculation of energy of decomposed coefficients from their GLCMs.....	27
Fig.2.7. Detail representation of proposed EBCE model.....	28
Fig.2.8. Flow layout of the suggested EBCE based fusion model	29
Fig.2.9. Fused image obtained from conventional image enhancement and EBCE technique.....	30
Fig.2.10. Fused images obtained from different image fusion techniques .. and the suggested EBCE model.....	31
Fig.2.11. Comparison of EBCE and Conventional Image enhancement techniques for PCA and mean-max fusion rule of image fusion based on PSNR....	32
Fig.3.1. Flow layout of the suggested AMT-DWT based fusion model.....	35
Fig.3.2. Mathematical model used for image sharpening.....	36
Fig.3.3. Sharpened images along with AMT-DWT enhanced images.....	36
Fig.3.4. Pictorial representation of proposed AMT-DWT algorithm.....	38
Fig.3.5. Well-known activation functions.....	41
Fig.3.6. Proposed Activation Functions.....	41
Fig.3.7. Qualitative analysis of proposed AMT-DWT technique with Harvard Medical School Brain Dataset with conventional techniques.....	42
Fig.3.8. Fused images PCA with AMT-DWT pre-processing.....	44

Fig.3.9. Fused images [mean-max [85] fusion rule with AMT-DWT.....	45
Fig.3.10. Output images of the suggested fusion model and existing models.....	50
Fig.3.11. Output images of the suggested model and CELM.....	51
Fig.3.12. Output image of proposed fusion model and existing image.....	
fusion models.....	53
Fig.3.13. Comparison of fused images of proposed fusion model and.....	
Existing models.....	53
Fig.3.14. Comparison of fused images of proposed fusion model with.....	
recently developed image fusion models [140]	55
Fig.3.15. Comparison on computational time required for the suggested fusion	
model with recently developed image fusion models.....	55
Fig.3.16. Comparison of PSNR value of the suggested model with recently	
developed image fusion models.....	56
Fig.4.1. A framework of designed Dense-net Res Net for MMIF fusion.....	59
Fig.4.2. Architecture of ETNet.....	62
Fig.4.3. Block diagram of the designed hybrid Dense Res Net Model.....	63
Fig.4.4. Architecture of Residual model.....	66
Fig.4.5. Structure of DenseNet.....	70
Fig.4.6. Experimental result of Dense-net Res Net model with $T1$, $T1GD$ and	
$T2$	70
Fig.4.7. Experimental image results of Dense-net Res Net model using	
$Flair$, $T1$, and $T1C$ images.....	71
Fig.4.8. Block diagram of the suggested fusion model.....	74
Fig.4.9. Flow diagram of the suggested image fusion model.....	78
Fig.4.10. Fused images of BraTs-2015 dataset by proposed Approximate.....	
fusion rule and conventional max fusion rule.....	80
Fig.4.11. Fused images of BraTs-2018 dataset by proposed Approximate.	
fusion rule and conventional max fusion rule.....	80
Fig.4.12. Fused images of Harvard Medical School Brain dataset by the	
proposed Approximate fusion rule and conventional max fusion rule..	83
Fig.5.1. Flow diagram of the suggested EADF based image fusion model	86
Fig.5.2. The output when no tumor is present.....	96

Fig.5.3. Different Breast Cancer Images of dataset [109] after preprocessing...	100
Fig.5.4. Block diagram of conventional Res Net and WMRESNET.....	101
Fig.5.5. Comparison of weights.....	102
Fig.5.6. Architecture of the suggested WMRESNET based classifier model....	104

List of Tables

Table 1.1. Various Modality images with diagnostic uses [1],[2],[4].....	2
Table 2.1. Research works exist on image denoisation in the literature.....	16
Table 2.2. CNN based image classification techniques used in literature.....	18
Table 2.3. Comparison of different transform methods in MMIF.....	25
Table 2.4. Performance parameters obtained from different enhancement techniques and proposed EBCE technique from Harvard Medical School Brain dataset.....	30
Table 2.5. Performance parameters obtained from proposed model ([mean-max [85]) and different existing image fusion models for image set-I from Harvard Medical School Brain dataset.....	31
Table 3.1. Specifications of custom de-noising CNN.....	40
Table 3.2. Quantitative analysis of the proposed AMT-DWT based enhancement technique on Harvard Medical School Brain dataset.....,	43
Table 3.3. Performance of PCA [13] based image fusion with proposed Preprocessing	44
Table 3.4. Performance of [mean-max [85] based image fusion with proposed Preprocessing.....	45
Table 3.5. Comparative analysis of the proposed model with different activation Functions.....	47
Table 3.6. De-noised output of the proposed model with diff activation functions Harvard Medical School Brain dataset.....	48
Table 3.7. Performance evaluation of suggested model with contemporary models.....	50
Table 3.8. Comparison with MI.....	51
Table 3.9. Comparative analysis of existing algorithms and suggested algorithms.....	52
Table 3.10. Comparative analysis of image fusion model [138],[139] and the proposed model.....	54
Table 3.11. Comparison of proposed model based on $Q^{AB/F}$ with the model Adapter [133].....	54

Table 3.12. Comparison of proposed model with existing CNN based model.....	54
Table 4.1. Performance evaluation of Dense-net-Res Net model using K-value (a) MSE (b) PSNR and (c) RMSE.....	71
Table 4.2. Performance analysis of Dense-net-Res Net model utilizing training data MSE (b) PSNR and (c) RMSE.....	72
Table 4.3. Comparative discussion.....	73
Table 4.4. Comparison of hybrid model 1 and hybrid model 2 for k value	77
Table 4.5. Comparison of Proposed Approx.- Conv. Mean and Conv. [mean-max] Fusion Rule.....	81
Table 4.6. Comparison of proposed Approx.-Proposed Detail and Conv .mean-max Fusion Rule	82
Table 5.1. Performance of EADF with and without haze removal technique.....	87
Table 5.2. Required % of Haze removal for detection accuracy.....	87
Table 5.3. Performance after pre-processing with conventional ADF and EADF.....	99
Table 5.4. Obtained threshold values (th) using the proposed auto thresholding.. from different datasets	92
Table 5.5. Performance of the proposed pre-processing framework with different haze removal techniques on BR35H [106] dataset	92
Table 5.6. Detection without using Detailed image and Principal image..... generation stage.....	93
Table 5.7. Performance comparison of Simple DCP and Approximate DCP	93
Table 5.8. Qualitative performance of the proposed model with different datasets.....	94
Table 5.9. Quantitative performance of the proposed model with different datasets.....	95
Table 5.10. Comparison of the proposed model with the existing CNN model	96
Table 5.11. Time taken in seconds to detect tumor by the proposed model on different Datasets.....	97
Table 5.12. Performance comparison of proposed model with state-of-the-art tumor Classifier models.....	97
Table 5.13. Output of the proposed framework on dataset [101].....	98
Table 5.14. Comparison of the proposed model with the existing ADF based model [101].....	99
Table 5.15. Generation of weight for different values of input	102
Table 5.16. Variation of learning and classification efficiency according to different value of 'P'	102

Table 5.17. Description of the proposed WMRESNET.....	103
Table 5.18. Comparison of proposed model with different CNN models.....	104
Table 5.19. Comparison of proposed model with existing CNN based models.....	105
Table 5.20. Different chest cancer images from dataset [110] after preprocessing.....	106
Table 5.21. Comparison of proposed model with existing CNN based model.....	106
Table 6.1. Detail description of the dataset used in mage fusion and denoisation.....	110
Table 6.2. Detail description of the dataset used in mage classification.....	110

List of Abbreviations

MMIF:	Multimodal Medical Image Fusion
MIF:	Medical Image Fusion
NNST:	Non subsampled Shearlet transform
NSCT:	Non subsampled Contourlet transform
CWT:	Complex Wavelet transform
LWT:	Lifting Wavelet Transform
HWT:	Haar wavelet Transform
DBWT:	Daubechies wavelet Transform
SWT:	Stationary wavelet Transform
EWI:	Empirical Wavelet Transform
DTCWT:	Dual Tree Complex Wavelet Transform
GA:	Genetic Algorithm
EBCE:	Energy based coefficient enhancement
AMT-DWT:	Averaging Multi Techniques-Discrete wavelet Transform
WMRESNET:	Weight Modified Residual Net
CNN:	Convolution Neural Network
PCA:	Principal Component Analysis
IHS:	Intensity Hue Saturation
NSCT:	Non Subsampled Contourlet Transform
DWT:	Discrete wavelet Transform
DCT:	Discrete Cosine Transform
DTCWT:	Dual Tree Complex Wavelet Transform
HWT:	Haar Wavelet Transform
NSST:	Non Subsampled Shearlet Transform
ReLU:	Recti Linear Unit
LReLU:	Leaky Recti Linear Unit
EReLU:	Exponential Recti Linear Unit
PSNR:	Peak Signal to Noise Ratio
RMSE:	Root Mean Square Error
SSIM:	Structural Similarity Measure Index
MI:	Mutual Information
FMI:	Fusion Mutual Information

FF:	Fusion Factor
FS:	Fusion Symmetry
FSIM:	Feature Symmetry Index Measure
$Q^{AB/F}$:	Quality Assessment based on Feature
CT:	Computed Tomography
MRI:	Magnetic Resonance Imaging
SPECT:	Single Photon Emission Computed Tomography
FLAIR:	Fluid Attenuated Inversion Recovery
<i>T1GD</i> :	<i>T1</i> weighted images with Gadolinium contrast
ADF:	Anisotropic Diffusion Filtering
EADF:	Extended Anisotropic Diffusion Filtering
DCP:	Dark Channel Prior
NN:	Neural Network
CNN:	Convolutional Neural Network
SVM:	Support Vector Machine
FL:	Fuzzy Logic
DT:	Decision Tree
KNN:	K-Nearest Neighbour
GA:	Genetic Algorithm
KMC:	K-means Clustering
DRT:	Dimension Reduction Technique
ML:	Machine Learning
DWT:	Discrete Wavelet transform
PSNR:	Peak Signal to Noise ratio
MSE:	Mean Square Error
GLCM:	Gray level Co-occurrence Matrix
LPF:	Low Pass Filter
HPF:	High Pass Filter
MIC:	Medical Image Classifier
PC:	Phase Congruency
POA:	Pelican Optimization Algorithm

CHAPTER 1

INTRODUCTION

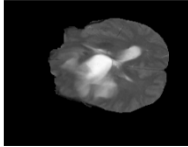

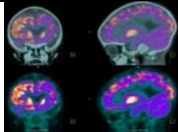
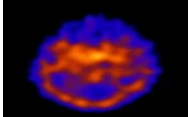


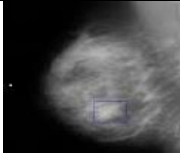
In medical science the multimodal image fusion is in rising demand on account of the availability of information for early diagnosis [1]. Fusion process amalgamate into images with more comprehensive information. Multisource image integration combines functional and structural data from different imaging modalities, providing enhanced visuals that assist physicians, thereby improving diagnostic efficiency [2]. In the present day of technological progress, diagnostic imaging plays a critical role in various medical applications, requiring high-fidelity images with expanded information for more precise and comprehensive diagnosis and treatment [3]. Early detection is critical in managing diseases more effectively, especially for the diseases like cancer and cardiovascular conditions. Disease identification based on the image of an organ obtained from a single source may not be sufficient enough to identify the hidden details. So, images are generated simultaneously by different imaging techniques and then combined together to get the fused image which in turn provides better particulars of an organ and is called Multimodal Medical Image Fusion (MMIF) [4], [5]. Advancements in medical research heavily rely on the progression of MMIF techniques [6]. Images captured by diverse sensors for various purposes such as X-rays, MRI, CT-Scan, etc., are termed as modalities [7].

1.1 Modalities of Medical Images

Medical image fusion (MIF) mainly deals with Ultrasound, Mammography, X-Ray CT (Computed Tomography), MRI (Magnetic Resonance Imaging), PET (Positron Emission Tomography), SPECT (Single Photon Emission Computed Tomography) [8]. Table 1 illustrates different types of medical modalities. These modalities within a specific technology are further categorized based on sensor orientations, tilting, and exposure time [9]. Each modality captures distinct features crucial for disease diagnosis, for instance, *Flair*, *T1*, *T2* and *T1C* are the different modalities of MRI images [10], [11]. MRI aids in diagnosing diseases in soft tissue body parts, especially within the brain, while CT scans are employed for detecting fractures and infections in relatively harder tissues like the skull and lungs [12], [13], [14]. X-rays are utilized for diagnosing ailments in even harder tissues such as bone fractures [4], [6]. The output images from these modality sensors are converted into digital format for further processing and storage by a computer [3]. Digital image processing, a field within

computer engineering, focuses on converting these two-dimensional spatial signals (images) into digital form for subsequent processing [5].

Table 1.1 Various medical modality images with diagnostic uses [1], [4], [5]

Medical Modalities	Image	Diagnostic Use
MRI		Soft tissue diagnosis
CT –Scan		Hard tissue diagnosis
PET		Functional Information
SPECT		Working information of organs
Ultrasound		Stomach and liver anatomy
X-ray		Bone Fractures
Mammography		Extra sensitive soft tissues of breast

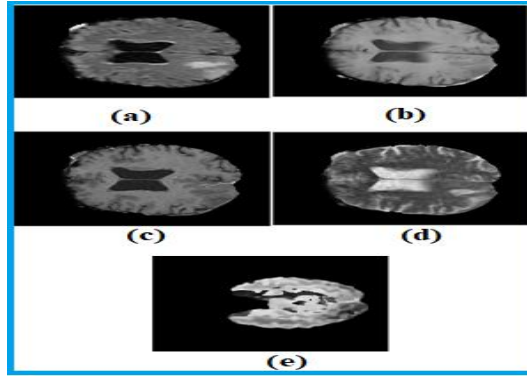


Figure 1.1 Brain MRI Modalities a. *Flair* b. *T1* c. *T1C* d. *T2* e. *T1GD*

MRI images can include sub-modalities such as *T1*, *T2*, *T1C*, *Flair*, and *T1GD*, as depicted in figure 1.1. *Flair* modality offers insights into central anatomy, such as tumors, while *T1* and *T1C* modalities provide details of fluid in the brain in the form of grey information [1].

Modality *T2* reveals fine details around the edges of the brain whereas *T1GD* images show anatomical details where fat appears bright and water appears dark [9]. In figure 1.2 the block diagram of MMIF is illustrated.

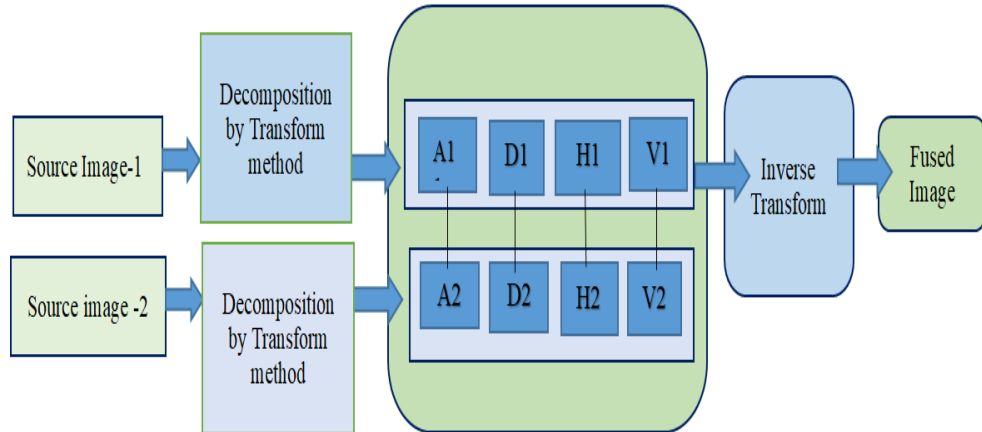


Figure 1.2 Multimodal Medical Image Fusion Methodology Design

1.2 Feasibility of MMIF techniques

There are several advantages for employing the MMIF technique in the realm of medical imaging and diagnostic science. Some of these advantages are outlined below:

- Fused images benefit from the addition of complementary details from both modalities (MRI and CT), is demonstrated in figure 1.3.
- Only a single informative fused image needs to be saved, resulting in saved computer memory.
- Less image data is required to transmit via transmission lines in smart medical systems.

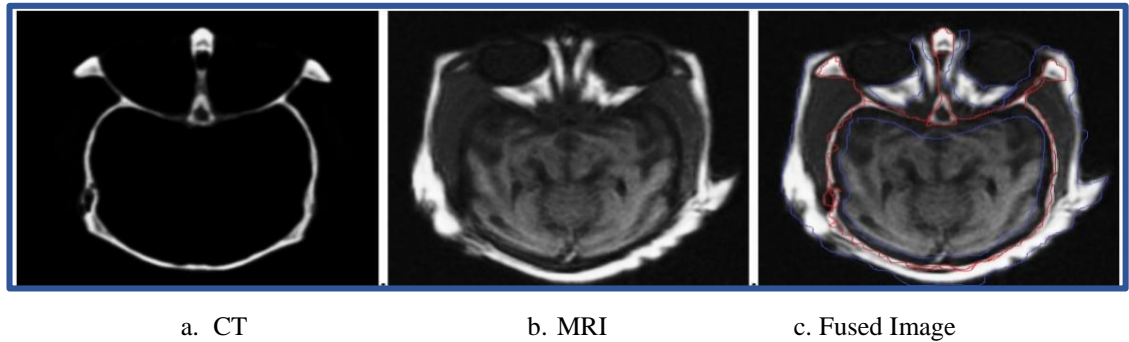


Figure 1.3 Fusion of CT & MRI Brain Images

Although a lot of publications are existing in MMIF field, attaining a fused image of high quality, characterized by minimal noise and minimal loss of source information while maintaining low computational cost remains a formidable challenge.

1.3 Applications of MMIF

By amalgamating different medical imaging techniques like MRI and CT, doctors can get a clearer picture of a patient's condition [1]. This helps in diagnosing diseases such as tumors or fractures more accurately also in diagnosing issues like blockages or structural abnormalities [7]. It assists in locating the exact position and size of tumors, making it easier for doctors to plan treatment.

1.4 Image Denoisation

Image denoising plays a critical role to reduce or eliminate noise from the image to make them more suitable for further analysis [15]. Noise can enter images due to various sources, such as poor lighting conditions, high ISO settings on cameras, imperfections in camera sensors, or environmental disturbances during image capture [15]. Such noise can negatively impact image clarity, obscure fine details, and diminish the effectiveness of tasks like object detection, segmentation, and other applications in computer vision. Denoising techniques seek to remove this noise while securing the image details, such as edges and textures. However, achieving this balance can be challenging, as aggressive denoising can blur essential features, while insufficient denoising leaves visible noise, impacting the image quality.

Conventional denoising methods, including spatial filters like Gaussian and median filters, are simple yet can struggle to retain details in high-noise settings [16]. More advanced methods, such as non-local means, wavelet-based techniques, and deep learning approaches, have shown improved performances. These techniques utilize statistical patterns and self-similarities within images [16].

Advancements in deep learning, with convolutional neural networks (CNNs) and transformer models, have significantly improved the performance of denoising [16]. By training on clean and noisy image pair dataset of large size, these models have

demonstrated impressive abilities to generalize noise patterns and restore high-quality images. Though various de-noising techniques are available in the literature, but de-noising algorithms may be unprotected for adversarial attacks, for which small and delicate information perturbations can mislead the model.

Image denoising plays a pivotal role in medical imaging, enhancing the quality of images for accurate diagnosis and analysis. Image denoising has become increasingly important in fields such as medical imaging, remote sensing, and multimedia, highlighting its value and broad applicability in practical settings.

1.5 Medical Image Classification (MIC)

Image classification performs a vital task to categorise images with different classes, and is widely used in medical science to assist in diagnosis, therapeutic planning and health management. With the advancement of imaging technologies like ultrasound, MRI, CT, and X-rays, vast amounts of image data are generated daily [17]. Physical assessment of these images can be susceptible to errors as well as labour-intensive, emphasizing the need for effective automated classification systems.

The primary responsibility of MIC is to accurately recognise and label particular anatomical structures, tissues, or pathological conditions. This task requires sophisticated techniques capable of recognizing subtle patterns and variations within images, which can be difficult due to factors like low contrast, noise, and anatomical diversity among patients. Automated classification tools have the capability to improve the accuracy of diagnosis, aid decision-making, and lessen the burden on radiologists and other healthcare practitioners [17].

Traditional MIC methods typically rely on manually designed features and statistical models to differentiate between classes [18]. However, these methods could be deficient in capturing the intricate patterns needed for reliable real-world classification. Nowadays, deep learning, notably with convolutional neural networks (CNNs) and more complex architectures, has arisen as a powerful tool for MIC [18]. By realizing complex features from large datasets, these models have achieved notable accuracy in identifying diseases, segmenting organs, and classifying tissues.

The development of accurate and reliable classification models has significant implications across medical fields, including the early detection of status of cardiovascular diseases, neurological disorders and cancer. As healthcare continues to advance digitally, MIC systems are expected to play an increasingly vital role in improving patient outcomes, reducing diagnosis times, and making healthcare more accessible.

MIC plays a major role in identifying and diagnosing various medical conditions, including tumors, infections and fractures from MRI, CT, and X-ray scans. For instance, classification systems can aid radiologists in detecting lung cancer in chest X-rays or recognizing brain tumors in MRIs, enhancing diagnostic accuracy and efficiency. Automated classification tools are essential for early-stage cancer identification, including types such as breast, lung, and skin cancer. By distinguishing

between cancerous and non-cancerous areas, these classification models facilitate early intervention, which is crucial for improving patient outcomes.

1.6 Research Motivation

Efficient Multimodal medical image fusion with less computational cost is the main objective of this thesis. The problem addressed in this thesis are:

- Medical images provide hidden insights of the internal organs of the body. However, the information obtained from a single-modality medical image is often limited and may not fully meet the needs of an accurate clinical diagnosis.
- Denoising of medical images is essential for enhancing the performance of the fusion model. But obtaining a denoising process without losing important image details, such as edges, or small objects, is a significant challenge. Balancing speed and accuracy while handling complex noise pattern is a big challenge. Specially Rician noise is prominent in MRI images which arises due to the complex nature of the MRI signal acquisition especially in low signal to noise ratio (SNR) condition. Similarly, Pepper and Salt noise may appear at image acquisition, transmission or sensor failures in devices like X-rays and CT scans.
- MIC faces many difficulties due to the intricate nature of medical images. Having an image classifier that can be trusted in clinical settings and adapted for use in high-stakes situations, such as diagnosis and treatment planning, is of utmost importance.

1.7 Problem Formulation

Huge efforts are being done to achieve an efficient image fusion model to increase diagnostic treatment and accuracy outcomes. The primary goal of this thesis is to achieve efficient multimodal medical image fusion outcome while minimizing computational costs. This thesis addresses the following problems:

1. The MMIF technique mainly suffers from the problem of redundancy or loss of crucial information of source images. Different imaging techniques can generate images with varying levels of resolution, contrast, and noise. Effectively merging these diverse characteristics remains a significant hurdle. Various imaging modalities may introduce different levels of noise and artefacts. Developing techniques that minimize noise while maintaining key details from each source is crucial for producing high-quality output images. There is a necessity to create more efficient algorithms that lower computational complexity while maintaining image quality. Numerous current algorithms demand considerable processing time and resources, which may restrict their applicability in clinical environments.
2. In general, denoising techniques must effectively reduce noise while preserving important details in the images. Many advanced denoising algorithms require a lot of computational power, making them less practical for use in clinics. There is a need to boost up the performance of the denoising algorithm with minimal execution cost.
3. It's important to know how classifiers reach their conclusions in healthcare. There is a gap in developing models that are easy to interpret, helping clinicians trust and

understand the reasoning behind predictions. Classifiers that perform well on one dataset might not work as effectively on others because of differences in patient demographics or imaging methods. More research is needed to create models that can generalize better across various populations and clinical scenarios. Achieving quick classification while maintaining high-accuracy is again a challenging task.

1.8 Methodology

In this thesis we discussed algorithms developed to address the problem of efficient MMIF, image denoisation, classification along with the performance analysis of existing image fusion models.

Performance Analysis:

The primary objective is to develop multimodal medical image fusion algorithms that enhance visual details for improved clinical diagnosis and to address the problem of requirement of large dataset and computational complexity. Hence, we developed the following:

- We explored various image transformation methods, including Daubechies, Haar, and the Lifting scheme, to decompose different modality MRI images from the BraTS 2015 dataset. We then applied the mean-max [85] fusion rule to combine the corresponding coefficients. Employing performance indicators like PSNR, RMSE, MI, standard deviation, mean, FSIM, SSIM, and entropy, we evaluated and compared the effectiveness of these different transformation methods in MMIF. The results indicated that the Daubechies transform outperformed the others, achieving the lowest RMSE and highest PSNR values.
- The proposed Energy-based Coefficient Enhancement (EBCE) method was applied to enhance medical images from the BraTS-2015, BraTS-2018, and Harvard Medical School Brain datasets and to fuse various imaging modalities. EBCE based enhanced images are fused with a well-suited fusion method (mean-max [85] or PCA [13]). The proposed EBCE method outperformed the traditional image fusion methods with Higher PSNR and lower RMSE.

Image Denoisation

- Image denoising enhances the quality and clarity of images by reducing noise, which consists of irrelevant data that can obscure critical details, allowing for a clearer view of areas of interest. To achieve this, we introduced the AMT-DWT (Averaging of Multi-Technique Discrete Wavelet Transform) image enhancement at the pre-processing stage, followed by an appropriate image fusion method (mean-max [85] or PCA [13]) using MRI and CT images from the Harvard Medical School Brain dataset [78]. The fused images were further refined with a custom CNN utilizing novel activation functions.

Efficient Image fusion model

- The primary goal of this research was to develop robust multimodal medical image fusion algorithms that enhance fused image quality to support doctors in early diagnosis. To achieve efficient fusion, we designed a hybrid model, Dense-ResNet, which combines the strengths of both Dense-net and ResidualNet. The strength of Dense-net is that it can extract the image feature in an efficient manner. The ResNet has a better performance speed due to its skip connection architecture. In the pre-processing phase, noise reduction was accomplished using Median filtering [15], followed by edge enhancement through ET-Net. The proposed fusion model showed superior results compared to existing image fusion models.
- With the continuation of our earlier work, we created another fusion model that combines the Pelican Optimization Algorithm (POA) [20] with Dense-ResNet for the fusion of different modality medical images from BraTs 2020 dataset. Similar to our previous approach, we applied median filtering [15] for pre-processing, used DTCWT for image decomposition, and employed ET-Net for edge preservation. In this model, POA was specifically utilized to optimize the training process of the Dense-ResNet model. The results showed that our proposed model performed better than the earlier one, achieving lower MSE and RMSE values while delivering a higher PSNR.
- Additionally, to mitigate unwanted artefacts and information loss during fusion, we introduced a set of image fusion rules specifically for fusing the corresponding decomposed multimodal medical image coefficients. For performance evaluation, we used the BraTS-2015, BraTS-2018, and Harvard Medical School Brain datasets. The results showed that the proposed fusion rules outperformed state-of-the-art image fusion models.

Efficient Image classifier model

- we proposed an image classifier model based on Enhanced Anisotropic Diffusion Filtering (EADF) combined with ResNet-50. The performance of the proposed classifier model outperformed several image classifier models.
- Additionally, the proposed WMRESNET (Weight-Modified Residual Network) successfully classified breast and chest cancer images with high efficiency. The proposed classifier model has been compared with many image classifier models and it outperformed all of them.

1.9 Major Contributions

1. Performance Analysis of Multimodal Medical Image Fusion Algorithms:

Explored and compared various image transformation techniques (Daubechies, Haar, Lifting) using the BraTS 2015 dataset, leading to the identification of the Daubechies transform as the most effective method for fusion, achieving superior PSNR and RMSE results.

2. Energy-Based Coefficient Enhancement (EBCE):

Introduced the EBCE method, which effectively enhances medical images prior to fusion. The EBCE method demonstrated superior performance over traditional image enhancement techniques, providing higher PSNR and lower RMSE values in the fusion outcomes across multiple datasets, including BraTS-2015 and BraTS-2018.

3. Image Denoising with AMT-DWT:

Developed the AMT-DWT (Averaging of Multi-Technique Discrete Wavelet Transform) for image denoising, which significantly enhances image quality by effectively reducing noise before applying fusion techniques. This model outperformed several existing denoising approaches, leading to clearer fused images for improved diagnostic clarity.

4. Efficient Hybrid Fusion Model (Dense-ResNet):

Designed a hybrid fusion model, Dense-ResNet, integrating the merits of Dense-net for efficient feature extraction and ResidualNet for rapid performance via skip connections. This model demonstrated remarkable improvements in fused image quality, surpassing conventional fusion techniques and several state-of-the-art models.

5. Optimized Hybrid Fusion Model (Dense-ResNet):

We developed an optimized image fusion model by utilizing the Pelican Optimization Algorithm (POA) with Dense-ResNet for the fusion of different modality medical images from BraTs 2020 dataset. The results showed that our proposed model performed better than the earlier one, achieving lower MSE and RMSE values while delivering a higher PSNR.

6. Development of Custom Fusion Rules:

Introduced a set of tailored image fusion rules specifically for the effective combination of decomposed multimodal image coefficients, minimizing the artefacts and information loss during the fusion process. These proposed fusion rules have shown to enhance the performance of the fusion outcomes significantly compared to traditional methods.

7. Development of Image Classifier Models:

Proposed an Enhanced Anisotropic Diffusion Filtering (EADF) approach integrated with ResNet-50 for efficient and rapid brain tumor detection, replacing tedious manual thresholding with an automated method, leading to improved classification accuracy. Developed the WMRESNet (Weight-Modified Residual Network) for classifying breast and chest cancer images, which exhibited high classification efficiency and outperformed numerous existing classifier models in accuracy.

1.10 Thesis Organization

Chapter 1 Introduction:

This chapter introduces the role of multimodal medical image fusion (MMIF), highlighting its advantages and key applications. It delves into the motivation behind the research and outlines its objectives.

Chapter 2 literature review and performance analysis:

In this chapter we reviewed the multimodal image fusion models for enhanced diagnosis. Different state-of-the-art models have been discussed. Also, various insights are derived, and gaps are systematically identified. This chapter also describes the performance analysis of MMIF after decomposing the source images by Doubechies, Haar and Lifting wavelet transform. This chapter further describes the performance analysis of MMIF by the proposed energy-based coefficient enhancement.

Chapter 3 Image denoisation:

This chapter outlines the performance of the MMIF model when combined with AMT-DWT-based pre-processing and custom CNN-based denoising.

Chapter 4 Efficient image fusion:

This chapter explores three efficient image fusion methods employing Dense-ResNet, optimized Dense-ResNet and the proposed fusion rules.

Chapter 5 Texture based image classification:

This chapter presents image classifier models based on anisotropic filtering and WMRESNET.

Chapter 6 Conclusion and future scope:

This section includes the conclusion part including the future scope of this research work.

CHAPTER 2

LITERATURE SURVEY AND PERFORMANCE ANALYSIS OF IMAGE FUSION MODELS

2.1 Literature Survey on Image Fusion

The goal of the present study is to understand the problems of MMIF, its denoisation and image classification in medical field. In the literature, much work has been done on these topics. This chapter review those approaches connected or related to the work presented in this thesis.

Medical imaging is a vital tool in healthcare, helping doctors diagnose, plan treatments, and monitor diseases [1]. Different imaging techniques provide unique details of an organ. However, no single method gives a complete picture because each has its own strengths and weaknesses. MMIF integrates images from different techniques into a single image that contains more relevant information [2]. By merging the strengths of CT and MRI helps doctor to understand complex medical conditions in better way, leading to more accurate treatment plans after proper diagnoses.

Over time, many MMIF techniques have been developed. Early methods, such as simple averaging, guided filtering, IHS, PCA [13] etc. often led to problems like missing important details [3]. Transform-based techniques, like wavelets, improved results by retaining more image details. Recently, artificial intelligence methods, including CNN and GAN models, have made fusion more accurate and automated [4]. Authors of [3] developed a MMIF model using guided filtering approach for securing the edge information of source images. By dividing each medical image into small-scale components, large-scale components, and background components, two fusion strategies have been used to fuse them. The first strategy involves PCNN guided by the structure tensor, and the second strategy is based on a maximum-based approach. Finally, the fused sublayers are combined to generate a single, unified image. While the model efficiently integrates images from different modalities, its performance is entirely reliant on the filter size utilized. Authors of [4] developed a novel method for enhancing structural and spectral features in the NSST domain for MMIF. This research begins with the generation of two pairs of images by using HIS. The NSST is used to obtain low-frequency and high-frequency sub-bands. Structural Information (SI) fusion is used to the Low Frequency Sub-bands (LFSs) to enhance structural elements such as texture and background information. Subsequently, PCA [13] is used for the High Frequency Sub-bands (HFSs) to extract pixel-level information. Finally, the fused image is generated by applying the inverse NSST along with the IHS method.

Though the model can fuse multimodal medical images successfully, it suffers from loss of details information and color distortion.

For sustaining the fine details, the authors of [5], [6], and [7] proposed image transformation-based models. Researchers of [5] introduced a two-stage MMIF approach that combines the Optimal DTCWT (O-DTCWT) and NSST to process distinct imaging modalities. The input images (Image 1 and Image 2) undergo decomposition using the O-DTCWT, separating them into high-frequency (HF) and low-frequency (LF) components. Fuzzy logic based fusion is used in HF components while the Maximum Rule is applied for the LF components and then reconstructed by inverse O-DTCWT. In the second stage, the reconstructed image is further decomposed using the NSST technique. Here, the HF components are fused through an Optimized Deep NN, while the LF components are merged using an averaging rule. At the end, images are merged with high efficiency, while with increased computational complexity. Researchers of [6] used DTCWT to make the model invariant of shifting as well as to acquire information regarding phase. The DTCWT have been utilized at different levels based on multiresolution principle and the complex wavelet coefficients are fused using max selection rule and five quality metrics have calculated to ensure the robustness of the model. This model can fuse three different modality medical images at higher computational cost and with higher noise sensitivity. Authors of [7] used the lifting scheme approach in the HW transform for fusing different modality X-ray images. This approach can't handle the smooth detailed regions of the images due to its blocky approach.

Researchers of [8] developed a new MMIF scheme based on discrete contourlet transformation for capturing edge details along curves. For the fusion process, pixel-level and decision-level rules are used for LF and HF components respectively. Although the proposed method effectively captures curved information, it lacks shift invariance and exhibits high computational complexity. Singh and Gupta [9] introduced an innovative MMIF technique based on structural gradient-based decomposition, for effectively separating structural and textural components and different fusion rules have been used for fusing them. The fusion method is evaluated using an extensive dataset of neurological images. While the model effectively fuses multi-modality medical images, it may encounter a loss of detailed information due to the overlapping of structural and textural data within the frequency band.

Ibrahim et al. [10] used NSCT in the decomposition stage and CNN based fusion for better edge and boundary details. Though this model is effective for grayscale images, its performance on color images has yet to be fully explored. Venkatesan et.al.[11] developed a concise overview of Multimodal Medical Image Fusion (MMIF). Padmavathi, Asha, and Maya [12] used a total variation (TV) decomposition-based image fusion model to enhance edge preservation, reduce noise, and achieve a smooth representation of homogeneous regions, making it effective in removing unwanted artefacts. While the model gives efficient fusion result, high computational complexity and tendency to produce extremely smoothed fused images are significant drawbacks.

The authors of [13] come out with an innovative framework that combines the NSCT with sparse representation (SR). In this approach, each source image undergoes decomposition using NSCT into a LP and multiple HP sub-images. The LP sub-images

are fused using an SR-based method, where a special dictionary is generated by merging compact and informative sub-dictionaries. These sub-dictionaries are derived by PCA [13]. For the HP sub-images, a special morphological operation along with gradient operators are used to extract detailed gradient features effectively. Although it produces fusion image rich in source information, it has the drawback of high computational complexity due to the system's intricate design.

Atwan et al. [14] provided comprehensive reviews on multimodal image fusion. Hussein et al. [15] and Fan et al. [16] represented their review works on image denoisation. Miranda et al. [17] and Tang et al. [18] presented their valuable review work on medical image classification. The authors of [19] introduced a modified Extreme Learning Machine (ELM) model by integrating the conventional CNN framework to mitigate the high computational cost of CNN-based image fusion models. However, the primary limitation of this model is its high sensitivity to noise.

Genetic algorithms (GAs), are widely applied in MMIF to optimize tasks. Several researchers [20]– [27] have explored neural network-based image fusion techniques combined with genetic algorithms. Authors of [20] introduces a new optimization algorithm inspired by nature, called the Pelican Optimization Algorithm (POA). The POA is based on simulating how pelicans behave while hunting. In this method, the search agents represent pelicans looking for food. A mathematical model of the POA is developed to solve various optimization problems. Authors of [21] come out with a new MMIF approach that utilizes spatial domain quality assessment and genetic algorithms, integrating elements of both feature-level and pixel-level fusion. This method involves dividing the source images into blocks and selecting the corresponding blocks with higher quality assessment values to form the final fused image but with a system complexity. The authors of [22] presented a method aimed at generating precise fused images by employing DWT for securing features and GA for optimizing the fusion process. This technique gives better fusion result at the cost of computational complexity.

The authors of [23] developed an innovative MMIF technique that integrates the Curvelet transform with a (GA). The inclusion of GA mitigates artefacts and noise by optimizing the fusion process's overall performance. Though the model effectively preserves data integrity and color fidelity, it suffers from the problem of increased computational time and high parameter sensitivity, which can slow convergence and impact performance quality. Authors of [24]- [29] utilized swarm optimization for obtaining the better fusion result. However, many of these algorithms face challenges with premature convergence and may get trapped in local optima. For better efficiency, boundary constraint handling adjusts candidate solutions at or beyond boundaries during iterations, increasing swarm diversity and reducing premature convergence. While the model effectively fuses multimodal medical images, its performance heavily relies on appropriate parameter selection.

Authors of [30] introduces a cooperative neural fusion (CNF) algorithm for significantly minimizing the contrast information loss in the presence of blind Gaussian noise. But the utilization of multi stage fusion makes the system complex with higher computational cost. A survey on Image Fusion (IF) after enhancing the

edge details is explained by the authors of [31] and [32]. Bhaladhare et al. [33] worked on MMIF based on fractional calculus (FC) approach. Though the model can extract better image features, it suffers from the problem of high computational cost. He. et al. [34] introduced a new MMIF scheme using Fuzzy C-means (FCM) algorithm. Specific fusion strategies are applied to preserve regions and enhance edges, focusing on visual quality and texture details with a higher computational time. In [35], the authors presented a powerful multi-resolution image fusion to imitate the image recognition and interpretation processes of the human visual system. This technique effectively combines SPECT and MRI images, with its superior feature extraction capabilities. Though both the models are efficient for MMIF the increased complexity of the system design leads to a higher computational cost.

Nowadays, hybrid methods are widely used for (MMIF) to incorporate the strengths of different approaches. Alseelawi et al. [36] come out with an approach of hybrid combination of DTCWT and NSCT. This method was tested on images from MRI, PET and CT scans. While this method produces a better fused image, it needs a lot of memory, which can make it difficult to use for real-time medical applications. Authors of [37] designed an MMIF scheme using DWT and Java technology for an efficient fusion and management of large medical image dataset. Ansari and Gupta [38] constructed MMIF model using wavelet Transform (WT), and worked for the reduction of computational cost. However, the introduced artefacts diminish the overall fusion effectiveness. The authors of [39] proposed an extensive review work of MMIF based on DWT.

Authors of [40] designed a novel fusion technique using Anisotropic Diffusion Filtering (ADF) along with thresholding technique for the reduction of computation time and to ease the system implementation and used the visible and infrared images. Though the computational cost is low, this model faces the challenges of accurate parameter selection for achieving a good fusion result.

GAN (Generative Adversarial Network) [41], [42] is widely used nowadays for its ability to produce higher quality outputs, with high efficiency even with a small dataset's but it might only create a small range of outputs instead of capturing the full variety of the target data. The authors of [43] come out with a MMIF scheme for the better reduction of noise. A new fusion rule is introduced for successful fusion of the important regions of the input images but with a high system complexity. The authors of [44] developed a new MMIF technique based on GAN. This method blends the CT's hard tissue information with MRI's soft tissue details. A generator combines data from both CT and MRI, while a discriminator learns to tell apart the fused image from the original one. It helps the generator to keep useful information from both images as possible. However, this method requires high computational power. Bhateja et al. [45] and Ramlal et al. [46] come up with hybrid model of NSCT and SWT. The authors of [45] introduces a framework that combines NSCT and SWT to fuse MRI with CT scans. Although this method efficiently fuses medical images, it struggles with the loss of spatial resolution and finer details from the source images.

In [46], the researchers proposed an improved hybrid fusion technique that also combines NSCT and SWT. Decomposition of the source images is done using NSCT. While this approach improves edge details, it comes with high computational costs.

Researchers of [47] unite NSCT and PCNN in MMIF purpose. The visible image is first divided into IHS transform. Then, the luminance component and the infrared image are decomposed with NSCT to get: LF, pass band (PB) and HP coefficients. The LF sub-band is fused using a weighted-sum method, while the other sub-bands are fused with the PCNN method. But this model is affected by noise. Bhateja et al. [48] established a new method of fusion for multispectral images that combines shearlet transform with NSCT. They utilized the YIQ color space. The LF coefficients are refined using phase congruency (PC), and the HP coefficients are processed with directive contrast (DC) and Shannon entropy. Although this method works well for multimodal medical image fusion, its complexity can be a limitation. Authors of [52] come out with a MMIF method based on local energy in the curvelet transform domain that captures edge details effectively. Loss of global information may produce suboptimal result.

In [53], the researchers developed a fusion framework using NSCT based image disintegration. Two fusion rules based on PC and DC are used for fusing the LF and HF components respectively. Though the model can capture curved information effectively, the computational cost of this circuit is very high. Researchers of [54] used SWT with fuzzy set theory in their model. The input images are broken down using SWT. Gaussian membership functions are used to create fuzzy sets, and local spatial frequency (LSF) is used to extract features from these sets. It gives good fusion results with higher complexity and computational cost. Most image fusion methods use down-sampling (DS) to reduce the load of computation by decreasing the set of parameters to be trained. However, down-sampling can also cause some source information to be lost. For addressing this, Payel et al. [55] suggested an MMIF model that retrieves complementary and edge information without using down-sampling. They used a combination of Laplacian and Gaussian techniques, along with channel attention pooling.

To tackle issues like blurry edges and reduced clarity caused by multiscale transformation-based fusion methods, the authors of [56] suggested a fusion approach using interval gradients with CNN. Although it gives good fusion result, some unwanted artefacts may appear due to the multi-stage fusion process. Kaur and Singh [57] introduced a medical fusion method that uses deep neural networks (DNNs) for automatically extracting image features. Their approach begins by decomposing the image into sub-bands in NSCT domain. The best attributes are selected by the modified version of the multi-objective evolution algorithm. The fusion functions are based on the energy loss and the coefficient of determination. However, this model may face the challenges with spatial distortion. Liang et al. [58] uses a (DCNN) to learn image attributes directly from the source images. This method suffers loss of spatial information even though it is effective for the reduction of noise. Manchanda and Gambhir [59] designed a fusion framework that uses wave atom transform to improve sparsity and time-frequency localization but the system performance is highly dependent on the system parameters. Researchers of [60] proposed an innovative

MMIF model based on two-channel non-subsampled filter banks which are not separable. They addressed filter design issues by introducing a mapping approach framework. The authors of [61], [62], and [63] reviewed MMIF models and the various performance parameters used in the field.

This review work explored different multimodal image fusion methods, highlighting their principles, advantages, challenges, and applications in medicine, while also discussing the latest advances and future potential in this area.

2.2 Literature Survey on Image Denoisation

Image preprocessing is a critical step before fusing images from different modalities for reduction of unwanted noise to ameliorate the image's visual quality [65]. Image denoising after fusion is also crucial for increasing the visual performance, interpretability, and usefulness of the output images in medical diagnostic and treatment planning [66]. Table 2.1 gives a brief information of the research highlighted in the literature. The literature survey on image denoisation reveals that denoising can enhance source images and lead to better fusion results. In contrast excessive denoising may remove essential details, potentially degrading fusion performance.

Table 2.1 Research works exist on image denoisation in the literature

Ref. No.	Author's name/ year	Approach	Advantages	Limitation/remark
[65]	Wangm Tiantian et al. (2024)	Lightweight progressive residual with attention mechanism fusion network	Preserve edge and texture information efficiently even after denoisation	Over fitting problem
[66]	Rajesh Patil et al. (2023)	Wavelet Transform	Computational cost is very low, no training data is required	Degradation of contrast level
[67]	Mukshimova et al. (2023)	Transfer learning method	Reduced computational cost	-
[68]	Yuqin Li et al. (2021)	Conditional GAN	Structural information can be preserved in a better way as training instability reduces	-
[69]	Size Li et al. (2021)	Multiscale Fusion GAN	Denoisation and super pixel reconstruction both can be performed efficiently	Computational cost more

[70]	Moran et.al. (2021)	Neighbor-2-Neighbor CNN (consists of a series of unsupervised and self-supervised models)	Training instability reduces	Vanishing gradient problem exist
[71]	Diego Valsesia et al. (2020)	A denoising method using a Graph Neural Network works on data that is structured as a graph.	The structure of a graph network can effectively model and fit complex noise distributions.	An unstable dynamic topology weakens the network's ability to express features effectively.
[72]	Hepeng Qu et al. (2019)	The Black Widow optimization algorithm uses tent mapping combined with multiple filters to enhance its performance.	Computational cost low	Lack of gradient information
[73]	Moon T. et al. (2019)	Residual network	Vanishing gradient problem reduces	Using too many connections in a neural network can cause overfitting problem.
[74]	Jingwen Chen et al. (2018)	Generative Adversarial Network (GAN) followed by CNN	This method can create realistic noisy images to address the problem of having too few training samples for the model.	The network training is laborious to reach the optimal performance.
[75]	Kai Zhang et al. (2016)	Modified VGG network	Unknown noise called blind Gaussian noise can be removed with promising runtime by VGG implementation	High computational cost

The literature survey on image denoisation reveals that while denoising can enhance source images and lead to better fusion results, in contrast excessive denoising may remove essential details, potentially degrading fusion performance.

2.3 Literature Survey on Medical Image Classification (MIC)

Image classification refers to the process of classifying images based on their features or patterns [80-97]. A concise view of the research works exists in the literature is reported in table 2.2. Literature survey on medical image classification reveals that, many methods already exist in literature but obtaining a model with high classification accuracy with less computational cost is still a big challenge.

Table 2.2 Image classification techniques based on CNN exist in literature

Ref No.	Author and Year	Technology used	Advantages	Dataset Used
[98]	Ganatra et al. (2024)	CNN with auto encoder-based feature extraction followed by SVM based classification	Successfully detect and classify the chest pneumonia and breast cancer images	Ultrasound Breast cancer Images (USBI), and Chest X-ray images-
[99]	Hata et al. (2024)	Probability Prediction model based on ML	Interstitial lung abnormality detected successfully	CT scan from patients in Boston Lung Cancer study
[100]	Bamber et al. (2023)	Deep Learning Network	Diagnosis and tracking of Alzheimer diseases	OASIS-3 dataset of brain MRI images
[101]	Maurya et al. (2022)	Anisotropic diffusion filtering with water shade algorithm	Successfully detect the brain tumor	Brain MRI images
[102]	Joshua et al. (2020)	HE based preprocessing, and deep learning-based classification with opposition-based Crow search optimization	Successfully detect brain tumor, lung cancer and Alzheimer diseases	Brain MRI, Lung CT and Alzheimer MRI images
[103]	Joao et al. (2020)	Low pass digital differentiator, anisotropic diffusion filtering with least square best fit algorithm for classification	Can detect breast cancer successfully	Breast mammography images which contain Benign and malignant cancer images
[104]	Lai et al. 2018	Supervised learning with coding network and multi-layer perceptron	Successfully classify different skin cancer images	HIS 2018 and ISIC 2017
[105]	Jadav et al. (2019)	Deep CNN (Transfer learning, VGG and SVM classifier)	Chest Pneumonia detected successfully	Chest X-ray dataset (normal, bacterial and virus images)

2.4 Chapter Summary and Gaps in the Study

This chapter conducts a detailed examination of the challenges associated with (MMIF), accompanied by a comprehensive review of existing literature. Traditional fusion methods, both in the spatial and transform domains, have created a strong foundation, allowing simple ways to combine images with low computational costs. Nevertheless, these techniques often struggle to effectively preserve both spatial and spectral details and may inadvertently introduce noise or artefacts, which limits their application in more intricate medical contexts.

In contrast, the AI domain of MMIF is notably versatile across various imaging modalities, capable of producing high-quality fused images. However, this advantage

comes with increased system complexity and a substantial demand for large training datasets. Each domain presents distinct strengths and limitations, and the selection of a specific approach is primarily influenced by the particular needs of the medical imaging task, including considerations of speed, accuracy, and the accessibility of computing resources.

In the field of image denoising, enhancing the quality of images from techniques like MRI, CT, Ultrasound [1] etc. is very important. This survey looks at both traditional denoising methods, like spatial filtering and transform-based approaches, as well as newer methods using artificial intelligence, especially deep learning models. Traditional methods often struggle for the reduction of noise on keeping important details intact. In contrast, AI-based methods are highly effective but require huge datasets and a lot of computing power.

MIC has made significant progress, driven by both traditional ML techniques and, more recently, DL models. Traditional methods, such as feature-based approaches like SVMs and decision trees, have provided valuable insights but often struggle with handling high-dimensional and complex data. Deep learning models like CNNs have transformed radically by attaining high precision in classifying various medical images, like X-ray, CT and MRI scans through automatic learning of features at different levels. However, these models result in computational overhead.

An inclusive literature review reveals the challenges listed below:

- The quality attributes of fused images, like the preservation of complementary information and anatomical details, are impacted by unwanted noise and a limited field of view.
- Medical images can contain noise and artefacts due to various reasons such as movement of patients, scanner imperfections, or contrast agent issues.
- Some models were unable to handle such noise and artefacts without compromising the quality of the fused images.
- Numerous de-noising methods are discussed in research, but these techniques can be vulnerable to adversarial attacks. These attacks involve tiny changes to the data that can mislead the model.
- Creating a high-quality fused image with minimal noise, minimal loss of original information, and low computational cost remains a significant challenge.
- The main challenges in MIC are the high computational complexity and the need for large datasets.

2.5 Objectives and Scope of the Work

To achieve high-quality fused images with minimal loss of source information, minimal noise, and low computational costs, the following objectives were framed:

1. Performance analysis of medical image fusion algorithms.

2. Design and development of efficient de-noising method to improve the image quality.
3. Design and development of efficient multimodal image fusion approaches for medical images.
4. Design and development of efficient classifier in the application of medical image processing.

Our first objective is to review and analyse the existing work in the domain of MMIF. This objective provides background information on the topic MMIF, showing the current state of knowledge and where our research fits within the broader academic discourse. It helps to situate our work in the context of existing studies and theories.

Our second objective is to denoise the source images as well as fused images to preserve information acquired from different modality medical images. The objective of denoising is to enhance the quality and clarity of the images for a better comprehensive view of the area of interest by reducing noise, which is any unwanted or irrelevant data that can obscure important details.

Our third objective is to propose an efficient MMIF model that addresses the challenges encountered by current models. The development of an efficient MMIF aims to improve the diagnostic and therapeutic capabilities in medical imaging by amalgamating complementary information from multiple imaging modalities.

The fourth objective is to develop an efficient classifier to analyze and categorize images based on the texture patterns within them.

2.6 Performance Analysis

With the extension of literature survey, the working of different MMIF algorithms has been analyzed and for this purpose the following methodologies have been used:

- Evaluating through execution of different transform techniques for the purpose of MMIF (Daubechies [6], Haar [7] and Lifting scheme [7] with $T1$, $T2$, $T1C$ and *Flair* modality images from BraTs-2015 [76] dataset).
- Performance analysis of MMIF with the proposed energy-based coefficient enhancement.

2.6.1 Performance Analysis of different transform techniques for MMIF

Different image transformation techniques (Daubechies [6], Haar [7] and Lifting scheme [7]) have been utilized to decompose the different modality source images. The detail description of these transform techniques is given below. The corresponding

coefficients of the decomposed coefficients have been fused by the mean-max fusion rule [85]. Different performance metrics [111], [112] have been calculated for weigh up the fused images achieved from different transform methods and the formulas used for this purpose, also have been represented in this section.

Daubechies Wavelet Transform [6]: Daubechies contributed a lot in the analysis of wavelet transform. She has contributed the orthonormal and orthogonal wavelets popularly known as Daubechies family wavelets. According to her if a function is a wavelet function, then it must attractive properties like have a compact support for an interval, it has at least single non vanishing moment and it must translate in orthogonal way. Daubechies wavelets are selected for their maximal number of vanishing moments. The coefficients of Daubechies wavelet transform can be obtained from the digital representation of the image by filtering it with LPF (low pass filter) and HPF (high pass filter) and subsampling the outputs by a factor of 2. The recursive application of the above steps creates many subsamples at the output. The diagram of Daubechies wavelet transform is depicted in figure 2.1.

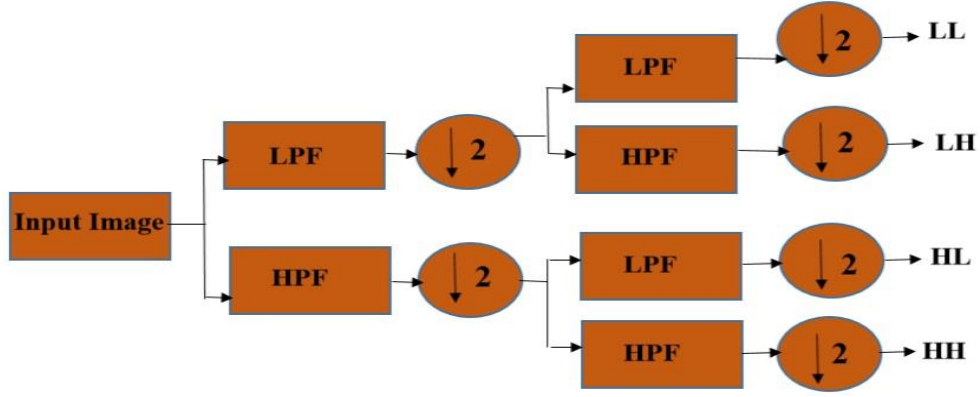


Figure 2.1 Diagram of Daubechies filter banks

Haar Wavelet Transform [7]: Haar Wavelet (HW) transform is considered one of the most effective wavelet functions, particularly in MMIF, because of its simplicity and computational efficiency. It is primarily utilized for image compression purposes. The description of the Haar wavelet function is depicted in Figure 2.2 below.

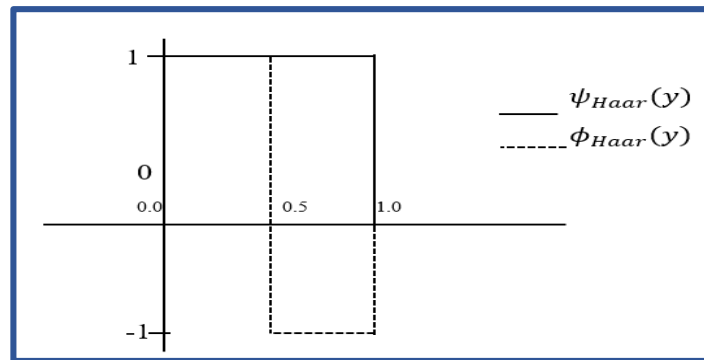


Figure 2.2 Haar Wavelet Function

Lifting Scheme [7]: The lifting scheme simplifies the DWT by breaking it into smaller steps called lifting steps. This process reduces the number of calculations by nearly half and makes it easier to handle the edges of signals. In the traditional discrete wavelet transform, multiple filters are applied to the signal separately. However, in the lifting scheme, the signal is split, and a series of operations are applied to the divided parts. The basic version of a forward wavelet transform in the lifting scheme is shown in Figure 2.3. The "P" represents the prediction step, where a high-pass filter is used to compute the wavelet function. The "U" is the update step, which produces a smoother version of the data. The lifting scheme offers several advantages: (1) Efficient memory usage, (2) The reverse of the forward transform, gives its inverse form (3) Subtraction is replaced with addition. The block diagram of a Lifting scheme in DWT is depicted in the figure 2.3 below.

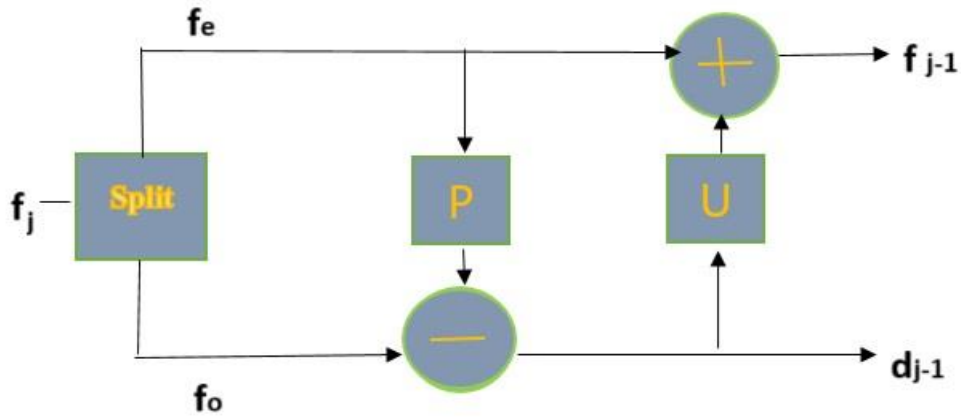


Figure 2.3 Diagram of Lifting Scheme

2.6.2 Calculated Fusion performance metrics

The analysis of different transform methods in the field of MMIF is done with the assistance of following metrics like [15], [111], [112].

Mean Square Error:

$$MSE = \frac{1}{xy} \left(I''(x, y) - I(x, y) \right)^2 \quad (2.1)$$

Where, $I''(x, y)$ is the fused image and $I(x, y)$ is the input image.

Root Mean Square Error:

$$RMSE = \sqrt{MSE} \quad (2.2)$$

Peak Signal to Noise Ratio:

$$(PSNR)_{DB} = (10 \times \log_{10}(Peak\ value)^2)/MSE \quad (2.3)$$

Here Peak value is 255.

Structural Similarity Measure Index:

$$SSIM(x, x_2) = \frac{(2\mu'_{x_1} \times \mu'_{x_2} + v_1)(2\sigma_{x_1, x_2} + v_2)}{(\mu'^2_{x_1} + \mu'^2_{x_2} + V_1)(\sigma_{x_1}^2 + \sigma_{x_2}^2 + V_2)} \quad (2.4)$$

Where, x_1 and x_2 are the input and the pre-processed image respectively, μ' and σ represent the mean and standard deviation respectively, and $V = (R \times S)^2$, where 'S' is a constant whose value is $<<1$ and 'R' is the dynamic range of an image.

Mean Value:

$$Mean = \frac{1}{n} \sum_{j=1}^n x_j \quad (2.5)$$

Variance Value:

$$Variance = \frac{1}{n} \sum_{i=1}^n ((x_i - mean((x_i)))^2 \quad (2.6)$$

Standard Deviation:

$$Standard\ Deviation = \sqrt{Variance} \quad (2.7)$$

Mutual Information:

$$MI(x_1, x_2) = \sum_{x_1, x_2} p(x_1, x_2) \log \frac{p(x_1, x_2)}{(p(x_1) \times p(x_2))} \quad (2.8)$$

Fusion Mutual Information:

$$FMI_F^{x_1, x_2} = \frac{1}{2} (MI(x_1, F) + MI(x_2, F)) \quad (2.9)$$

Quantity Assessment Based on Feature:

$$Q^{AB/F} = \frac{\sum_{m=1}^M \sum_{n=1}^N (Q_{(m,n)}^{AF} W_{(m,n)}^{AF} + Q_{(m,n)}^{BF} W_{(m,n)}^{BF})}{\sum_{m=1}^M \sum_{n=1}^N (W_{(m,n)}^{AF} + W_{(m,n)}^{BF})} \quad (2.10)$$

Where, $Q_{(m,n)}^{AF}$, $Q_{(m,n)}^{BF}$ are the edge preservation value and $W_{(m,n)}^{AF}$, $W_{(m,n)}^{BF}$ are their weights.

$$\text{Fusion Factor: } FF = I_{AF} + I_{BF} \quad (2.11)$$

Where, I_{AF} and I_{BF} are the mutual information with respect to fused image 'F' of the two source images 'A' and 'B'.

$$\text{Fusion Symmetry: } FS = \text{abs}(\frac{I_{AF}}{I_{AF} + I_{BF}} - 0.5) \quad (2.12)$$

$$\text{Entropy: } E = -\text{sum}(p \times \log_2(p)) \quad (2.13)$$

Feature Symmetry Index: FSIM

$S_{PC}(x) = \frac{2PC_1(x) \times PC_2(x) + T_1}{PC_1^2(x) + PC_2^2(x) + T_1}$, where PC stands for phase congruency and represents the feature of an input image 'x' at the point where the Fourier components are maximum in phase.

$S_G(x) = \frac{2G_1(x) \times G_2(x) + T_2}{G_1^2(x) + G_2^2(x) + T_2}$, where $S_G(x)$ represents the gradient magnitude of an image 'x'.

$$FSIM = \sum_{x \in \Omega} \frac{S_L(x) \times PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (2.14)$$

Where, Ω represents the spatial domain image and

$$PC_m(x) = \max(PC_1(x), PC_2(x)) \text{ and } S_L(x) = S_{PC}(x) \times S_G(x)$$

2.6.3 Results and Discussion

The subjective comparison of different transformation techniques is represented in figure 2.4 and the quantitative comparison based on different performance metrics value is presented in table 2.3.

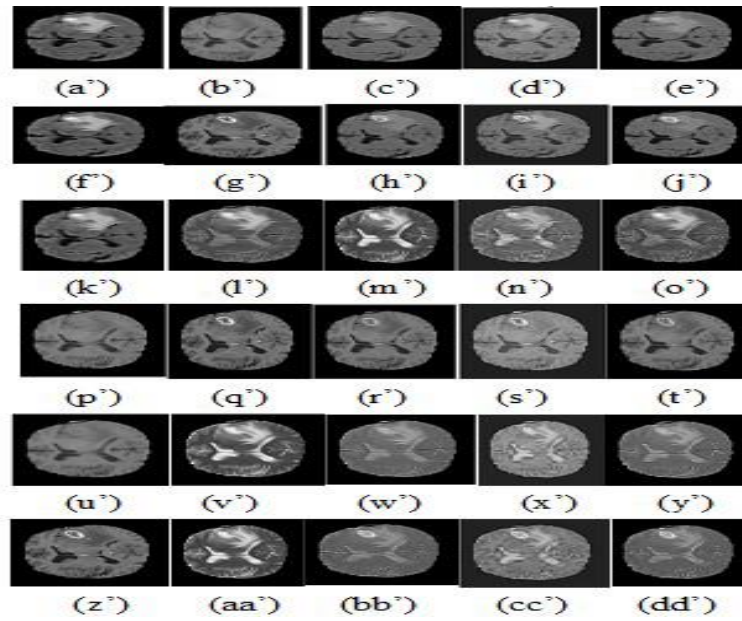


Figure.2.4 Fused images (a'= *Flair*, b'= *T1*, c'=fused by Daubechies, d'=fused by lifting, e'=fused by Haar ,f'= *Flair*, g'= *T1C*, h'=fused by Daubechies, i'=fused by Lifting, j=fused by Haar, k = *Flair*, l'= *T2*,m'= fused by Daubechies , n'=fused by Lifting , o'=fused by Haar , p'= *T1*, q'= *T1C*, r'= fused by Daubechies ,s'= fused by Lifting, t'=fused by Haar, u'= *T1*, v'= *T2*, w'= fused by Daubechies, x'=fused by Lifting, y'=fused by Haar, z'= *T1C*aa'= *T2*, bb'= fused by Daubechies , cc'=fused by Lifting , dd'=fused by Haar)

Table 2.3 Comparison of different transform methods in MMIF

Performance Analysis	Daubechies [6]	Haar [7]	Lifting [7]
PSNR	28.22	25.75	25.59
RMSE	9.9	13.45	13.57
MI	1.27	1.309	1.25
$Q^{AB/F}$	0.89	0.861	0.919
Mean	41.8	47.57	47.3
Variance	1.9e+03	2.44e+03	2.46e+03
medina	25	39	27
St. Deviation	43.59	49.4	49.59
FMI	0.93	0.933	0.928
FF	3.48	3.84	3.387
FS	0.05	0.042	0.016
SSIM	0.91	0.87	0.912
FSIM	0.93	0.92	0.965

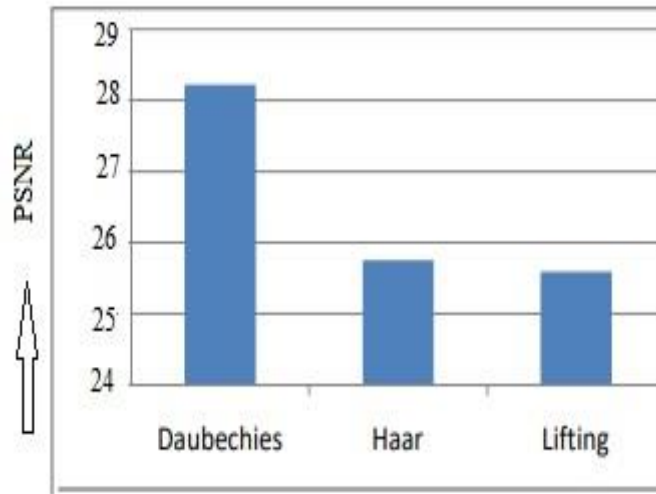


Figure 2.5 Comparison of Daubechies, Haar and Lifting wavelet transform based on the PSNR Value

The quantitative analysis revealed that the Daubechies wavelet transform outperforms the Haar and Lifting wavelet transforms in achieving higher PSNR values, however the result of the Haar wavelet transform is superior to the Lifting scheme.

2.7 Performance analysis of MMIF with the proposed energy-based coefficient enhancement technique.

In this research we have designed a methodology for image enhancement before fusion. This proposed work leverages DWT [22], [28], the novel Energy-based Coefficient Enhancement (EBCE), and the suitable fusion techniques (Conventional mean-max fusion rule [85] or Conventional PCA [13]). The EBCE technique surpasses several conventional image enhancement methods. Additionally, the suggested model's performance has been benchmarked against several cutting-age image fusion models, consistently outperforming them.

2.7.1 Methodology

The proposed MMIF model consists the following steps:

Step-1: Image decomposition by DWT [22]:

At this initial step, the two different modality source images MRI and CT [78] are disintegrated into detailed coefficients and approximate coefficient by DWT [22]. Let's the two source images are : $M_1(p, q)$ and $M_2(p, q)$. After decomposition, we obtain 8 decomposed coefficients: $(a_{11}(p, q), b_{11}(p, q), c_{11}(p, q), d_{11}(p, q))$ and $(a_{22}(p, q), b_{22}(p, q), c_{22}(p, q), d_{22}(p, q))$. Among these coefficients, $a_{11}(p, q)$ and $a_{22}(p, q)$ are the approximate coefficients, and the remaining coefficients are the detail coefficients.

Step 2: Enhancement of coefficients:

The energy of an image characterizes its localized changes [52] which may manifest in magnitude or color brightness over local areas of a pixel. In this research, the energy of each coefficient is calculated using its GLCM [87]. Each element (p ,q) in the GLCM represents the total occurrences of the pixel with value p lying horizontally next to the pixel with value q. The energy of GLCM [87] corresponds to the sum of the squares of each element, also known as its second angular moment. The flow layout of the calculation of energy components is represented in figure 2.6 whereas the detail description of the suggested enhancement process is represented in figure 2.7.

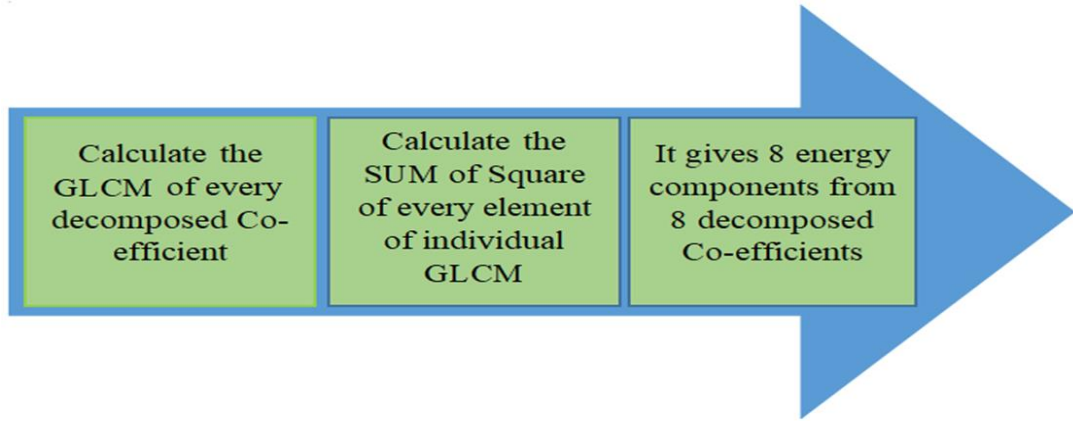


Figure 2.6 Calculation of energy of decomposed coefficients from their GLCMs

The Energy-based Coefficient Enhancement (EBCE) algorithm is elaborated below:

- (i) The energies obtained from the GLCM of each decomposed coefficient primarily signify their mean energies. Hence, (ea_1, eb_1, ec_1, ed_1) and (ea_2, eb_2, ec_2, ed_2) represent the mean energies of the decomposed coefficients.
 - (ii) The mean energy of each source image is also calculated from their respective GLCMs. Thus, $(e_1$ and $e_2)$ represent the mean energies obtained from the two source images $M_1(p, q)$ and $M_2(p, q)$ respectively.
 - (iii) The mean energy of individual coefficient is added to the mean energy of its corresponding source image to compute the effective mean energy. For instance: $(e_{11} = ea_1 + e_1, e_{22} = eb_1 + e_1, e_{33} = ea_3 + e_1, e_{44} = ea_4 + e_1)$ obtained from source image $M_1(p, q)$. And $(e_{55} = ea_5 + e_2, e_{66} = eb_5 + e_2, e_{77} = ea_7 + e_2, e_{88} = ea_8 + e_2)$.
 - (iv) The efficient or boosted coefficients are obtained through dividing the DWT decomposed components by their corresponding effective mean energies. Thus, the energy-efficient or energy-boosted coefficients are as follows: $(a_{111} = a_{11}/e_{11}, b_{111} = b_{11}/e_{22}, c_{111} = c_{11}/e_{33}, d_{111} = d_{11}/e_{44})$ and $(a_{222} = a_{22}/e_{55}, b_{222} = b_{22}/e_{66}, c_{222} = c_{22}/e_{77}, d_{222} = d_{22}/e_{88})$.
- The detail representation of the Energy-based Coefficient Enhancement algorithm is depicted in figure 2.7.

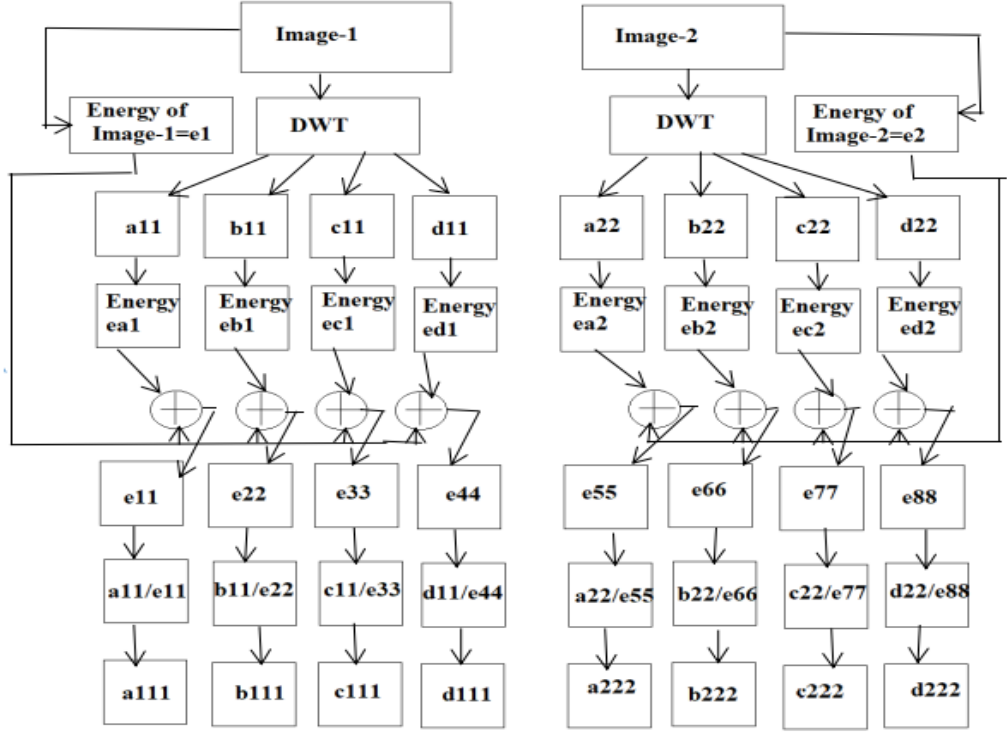


Figure 2.7 Detail representation of EBCE model

Step-3: Fusion of energy-enhanced coefficients: In this phase, for the fusion of the enhanced coefficients, two different conventional techniques have been employed separately: Principal Component Analysis PCA [13] and the mean-max fusion rule [85]. The conventional mean-max fusion rule [85] exhibits superior performance parameters and is thus selected for fusion purposes.

PCA method of image fusion [13]: In the PCA method, the source images are transformed into Eigen space. The principal component of an image contains its essential features with reduced noise corresponding to its major Eigenvalue. The boosted coefficients obtained from the previous stage are fused by the PCA [13] method. After fusion, only four coefficients (a_{1111} , b , c_{1111} , d_{1111}) are retained. The fusion process in the PCA method is governed by equations 2.15, 2.16, 2.17, and 2.18 respectively:

$$a_{1111} = PCA(1) \times a_{111} + PCA(2) \times a_{222} \quad (2.15)$$

$$b_{1111} = PCA(1) \times b_{111} + PCA(2) \times b_{222} \quad (2.16)$$

$$c_{1111} = PCA(1) \times c_{111} + PCA(2) \times c_{222} \quad (2.17)$$

$$d_{1111} = PCA(1) \times d_{111} + PCA(2) \times d_{222} \quad (2.18)$$

The mean-max fusion method [85] combines the approximate coefficients of two input images by calculating their mean. For the detail coefficients, the fusion process selects the maximum value from the corresponding coefficients of the two images. This method is represented by equations 2.19, 2.20, 2.21, and 2.22, where equation 2.19 is dedicated to fusing the approximate coefficients, and equations 2.20, 2.21, and 2.22 are utilized to fuse the detail coefficients. After mean-max fusion only four coefficients ($a_{11111}, b_{11111}, c_{11111}, d_{11111}$) retains.

$$a_{11111} = (a_{111} + a_{222})/2 \quad (2.19)$$

$$b_{11111} = \max(b_{111}, b_{222}) \quad (2.20)$$

$$c_{11111} = \max(c_{111}, c_{222}) \quad (2.21)$$

$$d_{11111} = \max(d_{111}, d_{222}) \quad (2.22)$$

Step-4: The fused coefficients are processed through the IDWT [22], [28] to reconstruct the fused image in the spatial domain separately. The performance comparison of the two different fused images obtained from two different fusion methods is reported in table 2.4.

The flow layout illustrating the suggested model is presented in Figure 2.8.

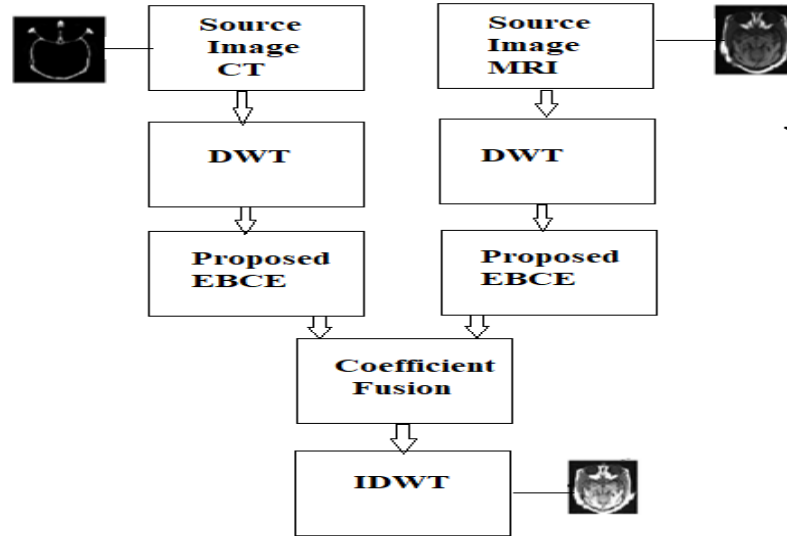


Figure 2.8 Flow layout of the suggested Model

2.7.2 Results and Discussion

The suggested multi-modal medical image fusion model effectively integrates CT and MRI images from the dataset [78]. Various traditional image enhancement methods, including HE [113], AHE [114], and CLAHE [115], are well-documented in existing research. A comparative analysis between the proposed Energy-Based Coefficient

Enhancement (EBCE) method and these conventional techniques is provided in Table 2.4, and the enhanced images are displayed in Figure 2.9. From Table 2.4, it is clear that the EBCE method outperforms traditional enhancement approaches, delivering better output values.

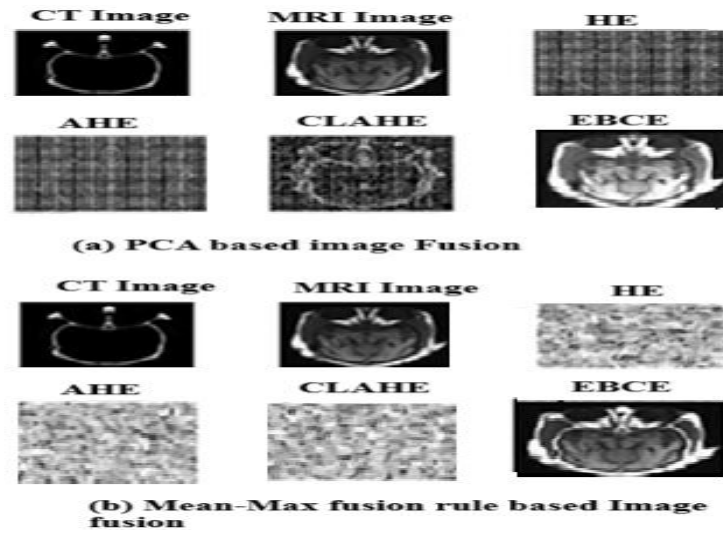


Figure 2.9 Fused images after conventional image enhancement techniques and Proposed EBCE technique

Table 2.4 Performance parameters obtained from different enhancement techniques and proposed EBCE technique from Harvard Medical School Brain data set [78]

Conventional PCA [13]				
Performance parameters	HE [113]	AHE [114]	CLAHE [115]	Proposed EBCE
PSNR	28.787	28.3187	28.35	31.801
RMSE	9.3415	9.3502	9.347	6.776
MI	1.06	1.05	1.01	1.164
FMI	0.732	0.722	0.729	0.786
FSIM	0.510	0.496	0.499	0.547
Entropy	3.51	3.28	3.41	4.52
Conventional [mean-max Fusion Rule [85]]				
PSNR	30.45	30.28	30.34	33.527
RMSE	7.49	7.40	7.45	4.099
MI	1.09	1.06	1.07	1.169
FMI	0.785	0.786	0.784	0.798
FSIM	0.527	0.518	0.521	0.565
Entropy	3.52	3.30	3.40	4.64

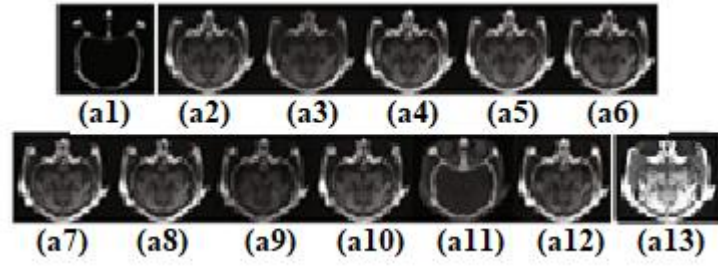


Figure 2.10 Comparison of fused images obtained from different image fusion models and Proposed Model (a1=GFF, a2=NSCT+SR, a3=NSCT+PCNN, a4=NSCT+LE, a5=NSCT+RPCNN, a6=NSST+PAPCNN, a7=DWT, a8=DWT+WA, a9=U-Net, a10=CNN, a11=ESF, a12=ESF+CSF, a13=Propose Model)

Table 2.4 demonstrates that the mean-max fusion rule [85] gives superior performance compared to PCA [13] fusion. So, we mean-max fusion rule [85] based is finally used in this method. The suggested fusion model is evaluated against several existing image fusion models. The quantitative results are reported in Table 2.5, while the qualitative outcomes are illustrated in figure 2.10.

Table 2.5 Performance parameters obtained from Proposed Model-II ([mean-max [85] fusion rule) and different existing image fusion models for image set-1(CT, MRI) from Harvard Medical School Brain Data set [78]

Name of the Models	Performance Parameters	
	PSNR	SSIM
GFF [116]	31.1594	0.4865
NSCT+SR [117]	29.5602	0.4825
NSCT+PCNN [118]	31.2341	0.5043
NSCT+LE [119]	31.6099	0.4861
NSCT+RPCNN [120]	31.6844	0.5002
NSST+PA-PCNN [121]	32.9194	0.4914
DWT [122]	31.972	0.4293
DWT+WA [123]	30.9814	0.4875
U-Net [124]	26.4196	0.3225
CNN [125]	28.9646	0.4751
ESF [126]	30.99	0.485
ESF+CSF [127]	29.6479	0.6483
Proposed Model	33.527	0.705

The foundation of the proposed model lies in its pre-processing stage, where the EBCE method is applied to amplify the energy of decomposed coefficients, thereby enhance the fusion performance. A comparison between EBCE and traditional image enhancement techniques, as shown in table 2.4, reveals that EBCE achieves superior results by delivering higher PSNR and lower RMSE values. Figure 2.10 demonstrates that the fused images produced using EBCE retain more intricate details compared to those generated by conventional methods. This pre-processing step is critical for achieving higher-quality fused images.

In this fusion process, both the traditional PCA [13] technique and the mean-max fusion rule [85] were utilized. According to table 2.4, the mean-max fusion rule outperforms PCA [13] across key performance metrics, including PSNR, RMSE, MI, FMI, FSIM, and Entropy, when applied to the Harvard Medical School Brain dataset [78]. Table 2.5 reports a comparison between the suggested model and state-of-the-art fusion models, highlighting its outstanding performance in terms of PSNR and SSIM. Furthermore, Figure 2.10 illustrates that the fused images produced by the suggested model are sharper and preserve more detailed information compared to those produced by existing methods. The model is also highly efficient, requiring less than 2 seconds to fuse images from different medical modalities, demonstrating its computational effectiveness. The graphical representation of the comparison of the two different fusion rule is depicted in figure 2.11.

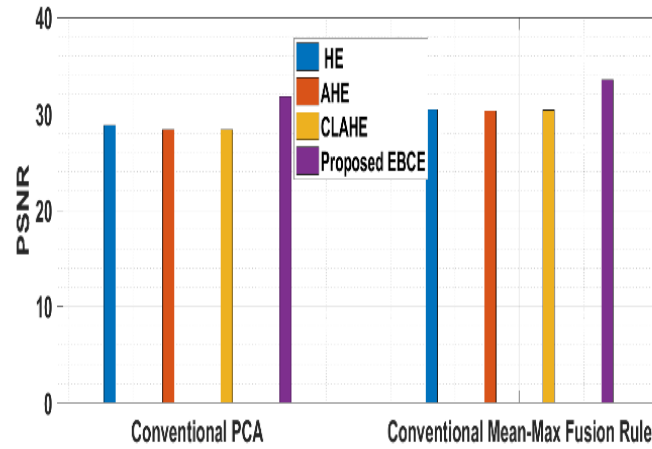


Figure 2.11 Comparison of Proposed EBCE and Conventional Image Enhancement Techniques for PCA [13] and mean-max fusion rule [85] of image fusion based on PSNR value.

2.8 Conclusion

Major progress has been achieved in MMIF, that merges complementary data from different imaging methods like PET, MRI and CT to enhance diagnosis and support clinical decisions. A thorough study of the research on image fusion, denoising, and classification shows that medical images typically contain intricate details and complex structures. Many such models are existing in literature but designing a model that can produce fuse image with all source information with low noise is still a big challenge.

An evaluation of image transform techniques, including Daubechies [6], Haar [7], and the Lifting Scheme [7], indicates that the Daubechies method delivers the most effective results in MMIF. Additionally, implementing the proposed Energy-Based Coefficient Enhancement (EBCE) method, followed by multimodal image fusion, shows that the mean-max fusion rule [85] achieves superior performance on the Harvard Medical School brain dataset.

CHAPTER 3

IMAGE DENOISATION

3.1 Introduction

MMIF amalgamates information from multi-modality images to generate a single fused image, aiding in accurate, efficient and early diagnosis [15]. Preprocessing the images is a key step before fusion, as it helps reduce unwanted noise and enhances the visual quality [16]. Furthermore, applying image denoising after the fusion process is crucial for enhancing the quality, clarity, and clinical relevance of fused images, thereby supporting errorless medical diagnoses and treatment planning [16]. Medical image denoising involves eliminating unnecessary noise from medical images while preserving critical details. Noise can reduce image quality, making it harder for doctors to diagnose conditions or evaluate treatments accurately [69]. Good denoising is important to improve image clarity and ensure essential diagnostic information remains intact [70].

After reviewing research in this area, it is clear that unwanted noise and a limited field of view can impact the quality of fused images, potentially causing important details to be lost. While there are many denoising techniques, algorithms can still be tricked by small, subtle changes. Achieving a high-quality fused image with minimal noise, little loss of information, and low computational cost is still a challenge.

To improve MIF, we suggested a multimodal fusion model. This model follows the important steps like: AMT-DWT-based image enhancement during preprocessing, DWT-based decomposition, PCA [13]-based image fusion, and custom CNN-based denoising. Each step helps reduce noise. Image filtering is a broad term that includes many ways to modify and improve images, while image denoising focuses specifically on removing noise. Both are important in improving image quality, and combining them leads to clearer source images.

The proposed AMT-DWT method processes the same image in different ways to make it clearer. This helps to extract hidden features from the images, which is crucial for medical image analysis, as accurate feature extraction is key for detecting diseases. After enhancement, the images are fused using techniques like PCA [13] and the mean-max fusion rule [85]. Ensuring that input images are clear and free from unnecessary information helps the fusion process, as clean images contribute to a more detailed final image.

Improvement the source images before fusion can produce higher quality fused image. Since medical images often have a lot of data, where each pixel represents a feature, PCA [13] helps reduce the amount of data while keeping the most important information. This makes denoising easier and improves fusion accuracy. The custom CNN, with two special activation functions, is used to reduce White Gaussian Noise (WGN) from the fused image. This stage improves the fused image by removing random noise, resulting in clearer images with better anatomical details, helping doctors interpret the images more easily.

The proposed activation functions work better than six well-known activation functions, and the fusion model shows better performance and efficiency compared to many advanced image fusion models.

Hence for the denoisation we have developed the following:

- AMT-DWT (Averaging of Multi Techniques- Discrete Wavelet Transform) based pre-processing with custom CNN based fused image denoisation has been proposed.

3.2 AMT-DWT based pre-processing with custom CNN based fused image denoisation

This research introduces a MMIF model to improve performance. The model combines key components, including AMT-DWT-based image enhancement in the pre-processing stage, DWT-based decomposition, suitable image fusion methods (PCA [13] or [mean-max [85] fusion rule [85]), and a custom CNN-based denoising process. Each of these components helps reduce noise. While image filtering includes various techniques for improving images, denoising focuses specifically on removing noise. Both approaches are important for enhancing image quality and clarity.

The enhanced images are then fused using the suitable technique of PCA [13] and the [mean-max [85] fusion rule [85]. For successful fusion, the input images must be reliable and free from irrelevant details, as clear images lead to more detailed fused results. Since the input images are enhanced before fusion, the obtained fused image is of higher quality. Medical images typically contain high-dimensional data, where each pixel represents a specific feature. PCA [13] helps reduce the image's dimensionality while preserving important information, which simplifies the denoising process and improves fusion accuracy.

A custom CNN with two proposed activation functions is utilized to remove White Gaussian Noise (WGN) from the fused image. This stage enhances the fused image by minimizing random noise, resulting in clearer and more accurate anatomical representations, which assist healthcare professionals in interpreting the images. The proposed activation functions outperform six well-known ones, and the fusion model can give excellent fusion metrics value and computational efficiency compared to

several state-of-the-art models. Figure 3.1 shows the flow lay-out of the proposed model.

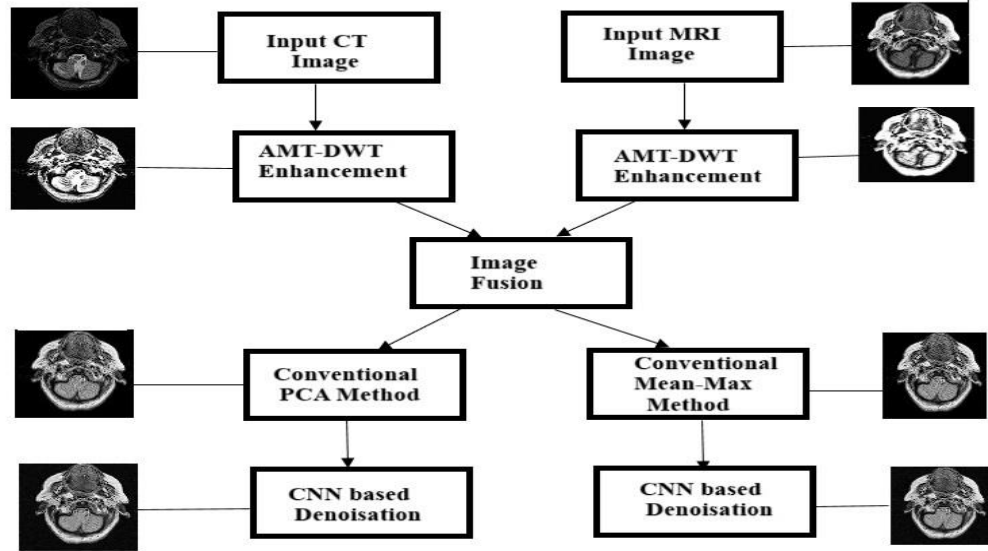


Figure 3.1 Flow layout of the suggested model

3.3 Methodology

This research tackles the challenge of developing an effective image fusion model that balances performance metrics and computational efficiency. The major steps of the suggested MMIF are described below:

Step 1: Intensity Enhancement (Proposed AMT-DWT): In the proposed method, the original image is processed through two parallel procedures. First, it is filtered using a median filter, while also passing through a Gaussian filter [3]. This dual-filtering approach is used because each filter is effective at removing different types of noise, which leads to improved images through two separate pathways. These pathways are then further enhanced using CLAHE [115]. After this, each enhanced pathway undergoes post-processing with unsharp masking. The purpose of unsharp masking is to restore any sharpness lost during the application of the blur filters in the pre-processing phase. Sharpening highlights the intensity transition regions in the image. The mathematical model for sharpening images using unsharp masking [88] is shown in Figure 3.2, and the sharpened images are depicted in Figure 3.3. The steps for sharpening the original image with unsharp masking are outlined in the following algorithm.

1. Construct a blur image $I'(x, y)$ from the original Image $I(x, y)$. (3.1)
2. From the original image, the blur image is subtracted to generate the mask $I_m(x, y)$, represented by equation 3.2 as . $I_m(x, y) = I(x, y) - I'(x, y)$ (3.2)
3. The mask $I_m(x, y)$ is added to the original image to generate the sharpened image $I_s(x, y)$, given in equation 3.3 as : $I_s(x, y) = I(x, y) + I_m(x, y)$ (3.3)

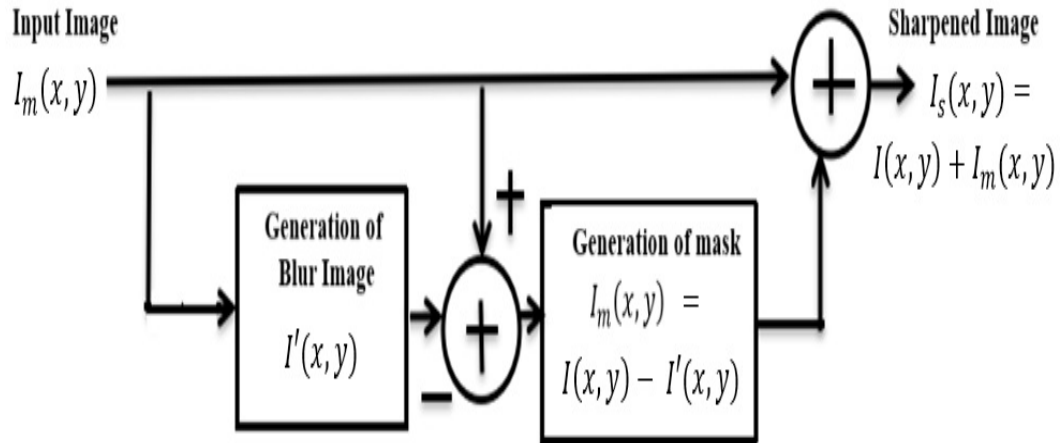


Figure 3.2 Mathematical model used for image sharpening

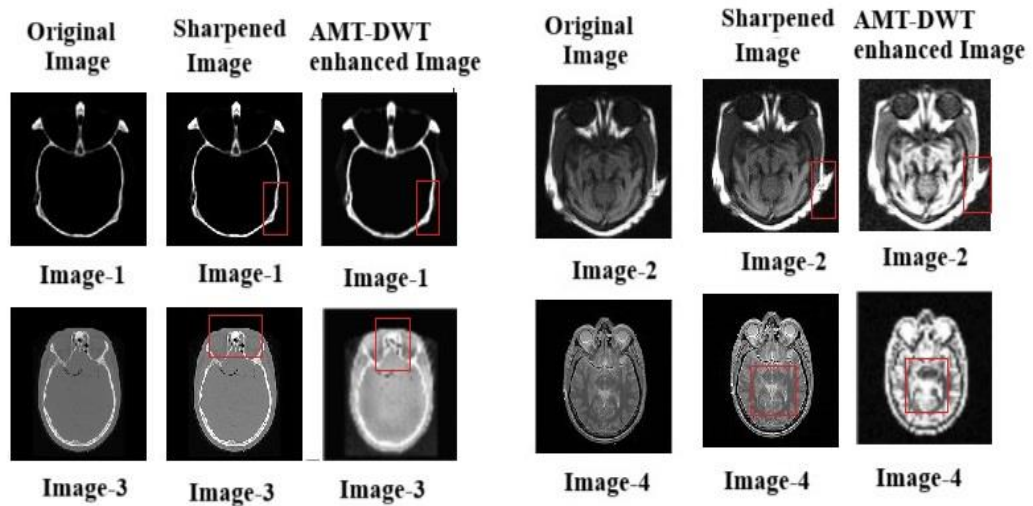


Figure 3.3 Sharpened source images along with AMT-DWT enhanced images

The reason for using different pathways is that each method has its own strengths, and combining their results can lead to better outcomes than using each one separately. CLAHE [115] is applied to prevent excessive contrast enhancement, while DWT [22], [28] helps preserve important image details and prevent noise from being amplified. A median blur filter reduces noise by replacing the centre pixel with the median value of the surrounding pixels. Additionally, a Gaussian blur, which uses a Gaussian kernel to assign different weights to neighbouring pixels, is applied to further reduce noise. Unsharp masking enhances the HF details of the image and make it sharper.

The unsharp masked image is then averaged with the two previous processed images, resulting in an image that is improved compared to the original. The CLAHE [115] enhanced images have higher contrast and are brighter than the original, while the unsharp masked image, although having lower contrast, is sharper and has better contrast than the original. If we group the CLAHE [115]-enhanced images and treat the unsharp masked image separately, they are averaged together in a 2:1 ratio during the fusion process. This balancing approach reduces the extreme contrast and brightness, creating an image that is less over-enhanced than the CLAHE [115] images but still better than the original. The final result is an image with improved contrast, brightness, and sharpness, without the over-enhancement seen in the CLAHE images. At the end of the enhancement process, IDWT is used to transform the image back to the spatial domain.

Let the original image be denoted as $I(p, q)$. The image after applying median filtering [15], CLAHE [115], and unsharp masking is represented as $I_1(p, q)$. The image produced after Gaussian blurring, CLAHE, and unsharp masking is denoted as $I_2(p, q)$. The image resulting from direct unsharp masking of the original image $I(p, q)$ is $I_3(p, q)$. After averaging $I_1(p, q)$, $I_2(p, q)$, and $I_3(p, q)$, the final image is $I_4(p, q)$. The components of $I_4(p, q)$ after DWT are $[CA1, CH1, CV1, CD1]$, while the components of the original image $I(p, q)$ are $[CA, CH, CV, CD]$. In the final step, the corresponding components of $I(p, q)$ and $I_4(p, q)$ are fused using the conventional [mean-max [85] fusion rule [85], [116], meaning CA is fused with CA1, CH with CH1, and so on.

The proposed AMT-DWT enhancement algorithm is applied on the dataset [78], and its performance is compared with existing image enhancement methods. The comparison results are shown in Table 3.2, and the qualitative comparison of images from different techniques is displayed in Figure 3.7.

Algorithm of suggested AMT_DWT:

- Step 1: Input the original image (I) and three other copies of input image have been generated.
- Step 2: Pass one copy through the Median blur filter and another copy through the Gaussian blur filter.
- Step 3: Apply CLAHE on each of the images obtained after completing step 2. Apply unsharp masking for each image.
- Step 4: Out of the remaining two copies of the original image, apply unsharp masking on one of them.
- Step 5: Apply image fusion for images formed from steps 3 and 4. The image fusion method adopted here is averaging.
- Step 6: Apply DWT to decompose the averaged image obtained from step 5 into [CA1 CH1 CV1 CD1].
- Step 7: Apply DWT on the original image and decomposed components are [CA CH CV CD].
- Step 8: Apply conventional [mean-max [85] fusion rule for the fusion purpose of corresponding decomposed coefficients i.e. the CA component fuse with CA1, CH is fused with CH1 and so on.
- Step 9: Apply IDWT to get the fused image in spatial domain.

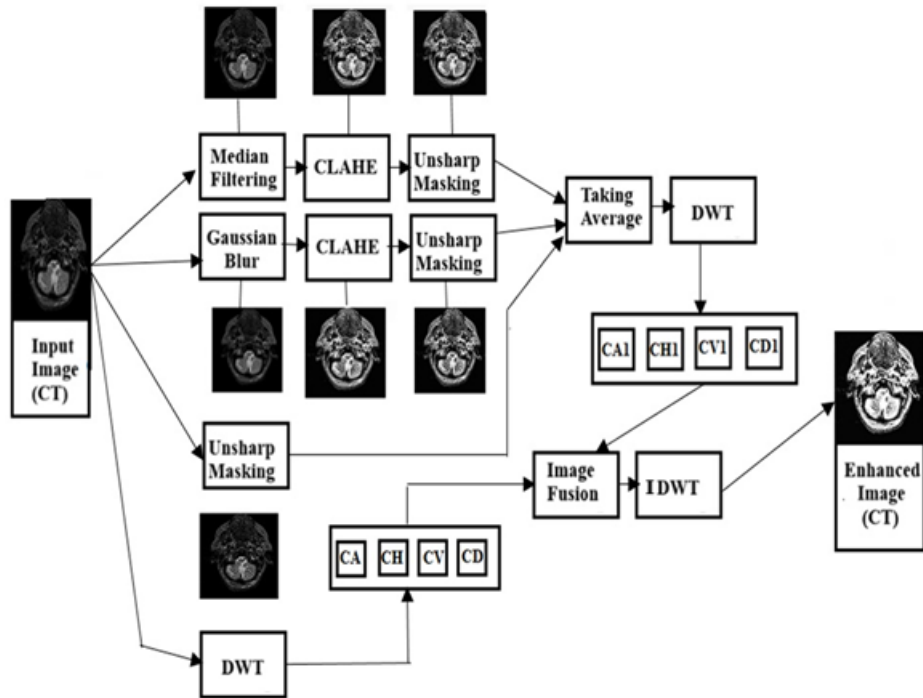


Figure 3.4 Pictorial representation of the suggested AMT-DWT algorithm

Step-2 (Discrete Wavelet Transform): In this stage, a two-dimensional DWT is applied to break down the intensity-enhanced images from different modalities (CT and MRI from the Dataset [78]) into four components, each in the frequency domain. DWT is an effective tool for reducing noise and is commonly used in image processing. It helps localize important features of the image at different scales, producing a sparse representation. The wavelet transform focuses on significant image features by creating large-magnitude coefficients, while noise and unnecessary details are captured in smaller-magnitude coefficients. This makes it easier to remove noise, resulting in a cleaner image. By using DWT in this stage, the image is decomposed into various components, which are then used in the next step. As a result, the decomposed components have less noise, offering a clearer and more informative image.

Step-3 (Fusion of decomposed enhanced images): The corresponding components of the decomposed intensity-enhanced images are fused using appropriate fusion techniques, namely PCA [13] and the [mean-max [85] Fusion rule. The objective is to combine information from the two different modality images into one cohesive fused image. The fusion results of PCA method and the [mean-max fusion rule are reported in Table 3.3 and Table 3.4, respectively, with their quantitative representations shown in Figures 3.8 and 3.9. It is evident that conventional PCA yields better results than the [mean-max [85] fusion rule. This step is crucial for merging complementary information from the two modalities, improving the overall quality of the fused image. PCA is a technique that helps extract the principal components from a raw dataset, reducing its dimensionality. The computational steps followed in PCA are as follows:

- Step 1: The raw image data is organized into columns.
- Step 2: The mean of each column is computed.
- Step 3: Subtraction of mean from each column and generation of covariance matrix.
- Step 4: Calculation of Eigen vectors and Eigen values for the covariance

For two different modality images, ‘A’ and ‘B’, the PCA components are calculated as $P(A)$ and $P(B)$. The output image is then formed by merging $P(A)$ and $P(B)$. PCA acts as a de-correlation technique that transforms the dataset into the principal component domain, keeping only the most important components. This approach effectively eliminates noise and irrelevant details, making PCA a widely used method for image de-noising. By retaining the key components and discarding the less significant ones, PCA improves the overall image quality while reducing noise and unwanted artefacts.

Step-4 (Inverse Discrete Wavelet Transform) IDFT: The four components obtained from the fused image in the previous step are combined using the IDWT, which restores the image to the spatial domain.

Step-5 (De-noising CNN): Image de-noising is essential in medical imaging because it directly affects the diagnostic accuracy, with AWGN being one of the most common types. Research shows that CNN-based de-noising techniques are effective because

CNNs can automatically identify and remove noise from image patches without human intervention. During training, they learn to de-noise images efficiently, and their ability to handle large datasets ensures consistent and reliable results.

While increasing network depth is a common way to improve CNN performance, it also increases system complexity and computational costs. To address this, this paper proposes an alternative approach—improving CNN performance by using stronger nonlinear activation functions, without increasing the network depth, thus reducing both complexity and computational cost. The proposed CNN includes twenty convolution layers and eighteen batch normalization layers, specifically designed for image de-noising. In MATLAB, a ‘denoisingImageDatastore’ object is created using 300 grayscale images at ten different noise levels. Random noise is added to image patches in each epoch, creating a ‘denoisingImageDatastore’ (imds) object with both clean and noisy image patches. The CNN is trained with fifty such patches. The specifications and structure of the proposed CNN are described in Table 3.1.

Activation functions are crucial because they map a node’s output to a range between '0' and '1' or '-1' and '1,' depending on the function. While ReLU is often used for image de-noising, its performance drops when preserving negative input values is important. To solve this, a new activation function is introduced in this research by Equation 3.4, with its derivative in Equation 3.5. Figure 3.6 shows the proposed activation function and its derivative, which is used as the second activation function in the custom CNN. For performance comparison, six well-known activation functions—logistic, rectified linear, identity, tangent hyperbolic, derivative of tangent hyperbolic, and Gaussian—were tested alongside the proposed ones. Figure 3.5 provides plots of these functions, and Tables 3.5 and 3.6 offer a quantitative and qualitative analysis of the de-noising performance, showing the improvements achieved by the proposed activation functions.

Table 3.1 Specifications of custom de-noising CNN

SL. No.	Name of the parameter
1.	Image Input Layer (number in use-1)
2.	Convolution Layer(3x3x1) (number in use-20)
3.	Activation Layer (Proposed Activation Functions) (number in use-19)
4.	Normalization Layer (number in use-18) (Batch Normalization)
5.	Output Layer (number in use-1)
6.	Learning Technique (Residual)
7.	Stride [1 1]
8.	No. of Epoch (number in use-5)
9.	Noise patches(number-50)
10.	Loss Function (MSE)

Algorithm used in De-noising CNN [121]:

1. Considering the input image $I(i, j)$
2. Generation of noisy image $Y(i, j)$ by adding 1% (AWGN) into $I(i, j)$.
3. Prediction of Residual Image $I'(i, j)$, which is difference between $Y(i, j)$ and $V(i, j)$,
Where $V(i, j)$ is cleaned image.
4. Estimation of MSE (Mean Square Error) from the Residual Image $I'(i, j)$.
5. Generation of De-noised Image by using MSE as Loss function.

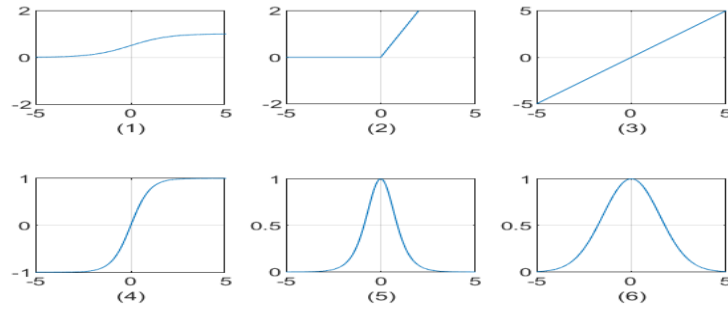


Figure 3.5 Well known activation functions ((1) logistic function (2) rectified linear function (3) identity function (4) tangent hyperbolic function (5) derivative of tangent hyperbolic function (6) Gaussian function)

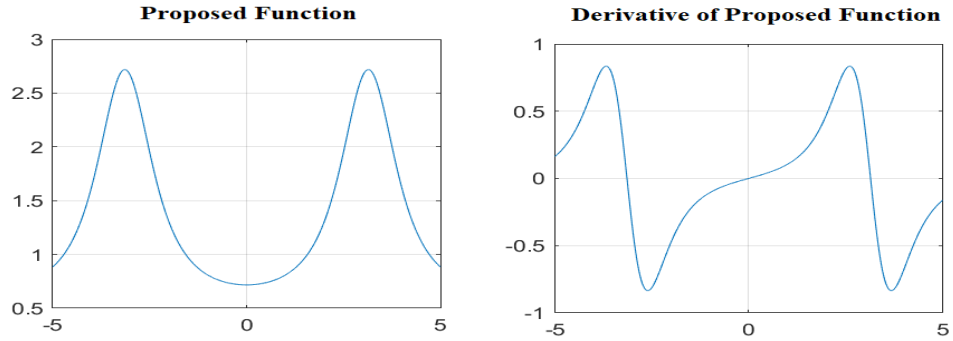


Figure 3.6 Proposed activation functions

$$y = \exp\left(\frac{(1-temp1)}{(1+temp1)}\right) \quad (3.4)$$

Where, $temp1 = 1 + \cos(x)$

The derivative of the proposed activation function-1(equation-1) is also calculated and is represented by equation 3.5.

$$y_1 = y \times \frac{\sin(x)}{(2+\cos(x))} \quad (3.5)$$

3.4 Results and Discussion

The proposed denoising AMT-DWT method is compared with conventional image boosting techniques, including HE, AHE, and CLAHE. This comparison is performed using the Harvard Medical School Brain dataset, evaluating performance using metrics like PSNR, RMSE, and SSIM [111]. The results are shown in Table 3.2 (for image set 1 to 8 for both MRI and CT) and Figure 3.7.

A higher SSIM value indicates better preservation of the image's structure, while a higher PSNR value and lower RMSE value show better noise reduction, leading to improved image quality. The assessment results highlight that the suggested denoising AMT-DWT method outperforms current techniques.

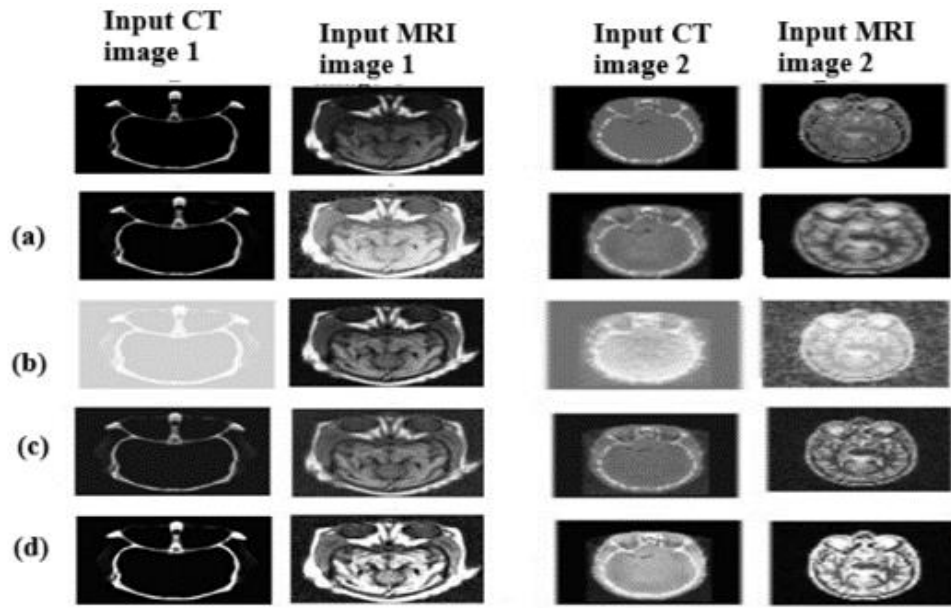


Figure 3.7 Qualitative analysis of proposed AMT-DWT technique (with Harvard Medical School Brain dataset [78] with conventional image enhancement techniques (a) HE (b) AHE (c) CLAHE and (d) Proposed AMT-DWT)

Table 3.2 Quantitative result of the suggested AMT-DWT based enhancement technique on Harvard Medical School Brain dataset [78]

Images (I-Image set)	PSNR value [112]				RMSE value [112]				SSIM value [112]			
	AMT-DWT method	HE [113]	AHE [114]	CLAHE [115]	AMT-DWT method	HE [113]	AHE [114]	CLAHE [115]	AMT-DWT method	HE [113]	AHE [114]	CLAHE [115]
I ₁ (MRI)	30.226	1.853	13.5988	20.49	7.856	206	65.725	24.084	0.976	0.765	0.876	0.904
I ₁ (CT)	21.56	9.5778	13.5944	16.867	21.31	84.6548	65.595	36.574	0.987	0.745	0.843	0.912
I ₂ (MRI)	26.27	7.4694	11.352	20.546	12.384	107.911	120.955	23.94	0.966	0.67	0.868	0.919
I ₂ (CT)	20.25	8.306	11.318	18.463	24.78	98.00	120.9593	30.436	0.978	0.598	0.889	0.909
I ₃ (MRI)	28.07	10.9387	12.91	21.966	10.066	72.377	91.523	20.3356	0.988	0.65	0.897	0.9156
I ₃ (CT)	22.03	6.775	12.80	16.089	20.17	116.8939	91.200	40.003	0.995	0.76	0.945	0.9345
I ₄ (MRI)	22.85	7.2027	9.744	20.84	18.36	111.277	84.31	23.13	0.979	0.745	0.879	0.9563
I ₄ (CT)	21.59	6.0521	9.715	19.60	21.23	127.04	84.57	26.696	0.976	0.786	0.888	0.9236
I ₅ (MRI)	20.93	12.45	9.778	18.109	22.90	60.7865	83.282	31.70	0.998	0.776	0.7999	0.9187
I ₅ (CT)	17.21	9.388	9.7462	16.165	35.158	86.52	83.44	39.65	0.978	0.756	0.8998	0.976
I ₆ (MRI)	16.36	5.5888	10.1633	15.73	38.67	133.998	65.149	41.678	0.969	0.769	0.8798	0.899
I ₆ (CT)	19.65	8.99	10.1933	14.825	26.54	90.48	65.564	46.268	0.972	0.698	0.834	0.959
I ₇ (MRI)	17.02	14.75	6.552	16.90	35.91	46.64	69.296	36.43	0.956	0.699	0.874	0.929
I ₇ (CT)	14.71	18.34	6.542	13.097	46.88	30.86	69.457	56.454	0.949	0.766	0.886	0.976
I ₈ (MRI)	14.80	5.569	12.563	13.508	46.36	134.299	56.66	53.845	0.985	0.709	0.897	0.9129
I ₈ (CT)	19.87	7.95	12.568	18.99	25.88	120.06	56.574	28.62	0.958	0.77	0.890	0.9185
Average value	20.83	8.825	10.8211	17.6366	23.706	101.74	79.64	43.99	0.9756	0.7288	0.877	0.9289

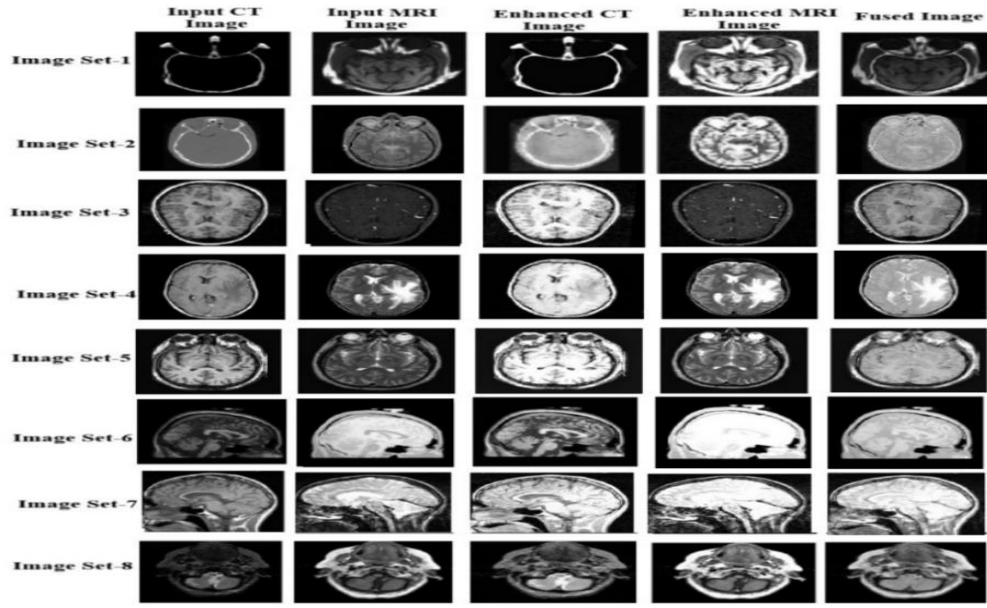


Figure 3.8 Fused images (PCA) [13] with the proposed AMT-DWT pre-processing

Table 3.3 Assessment of PCA [13] based image fusion with proposed pre-processing

Images	PSNR	RMSE	MI	SD	FMI	FSIM	SSIM	Entropy
Image set 1	16.655	37.454	1.0054	0.4563	0.7607	0.7327	0.7723	7.1746
Image set 2	10.774	73.89	1.2	0.4862	0.834	0.7948	0.4149	6.1331
Image set 3	8.5531	95.366	1.2181	0.4992	0.8060	0.4519	0.1748	7.0057
Image set 4	9.184	88.676	1.1973	0.490	0.899	0.6282	0.4692	4.7045
Image set 5	8.07	100.76	1.1337	0.4995	0.7456	0.3858	0.0344	6.4434
Image set 6	14.995	45.358	1.216	0.4986	0.8310	0.5856	0.4539	3.3542
Image set 7	7.6381	105.82	1.068	0.495	0.7012	0.3652	0.0574	6.4596
Image set 8	16.35	38.89	1.095	0.4736	0.675	0.4586	0.4267	5.4694

Table 3.4 Assessment of [mean-max based image fusion [85], [116] with proposed pre-processing

Images	PSNR	RMSE	MI	SD	FMI	FSIM	SSIM	Entropy
Image set 1	13.6	5.662	1.057	0.282	0.8705	0.6553	0.4046	6.546
Image set 2	11.35	6.9376	1.186	0.487	0.8505	0.7095	0.4445	5.56
Image set 3	12.874	6.356	1.135	0.4604	0.6915	0.4186	0.1667	6.7872
Image set 4	9.743	8.336	1.187	0.4884	0.8744	0.6148	0.4817	4.7651
Image set 5	9.781	8.444	1.1234	0.4925	0.7913	0.4556	0.0362	6.2463
Image set 6	10.19	9.1361	1.2072	0.4964	0.869	0.6006	0.434	4.8468
Image set 7	6.55	12.096	1.082	0.4969	0.706	0.3151	0.0362	6.3442
Image set 8	12.58	6.566	1.084	0.415	0.4623	0.4046	0.4106	5.7093

The fusion process is performed using both the conventional PCA [13] method and the mean-max [85] fusion rule. The results of these methods are shown in Figures 3.8 and 3.9, with their corresponding performance metrics listed in Tables 3.3 and 3.4. It was found during experimentation that, when combined with the proposed pre-processing, PCA [13] produces better results than the mean-max fusion rule [85]-based fusions.

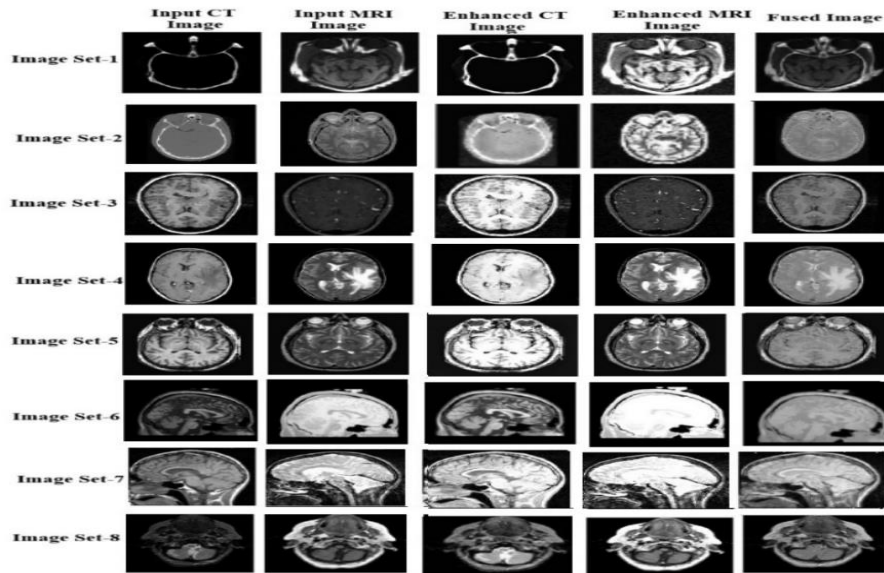


Figure 3.9 Fused images of Mean Max fusion rule [85], [116] with the proposed AMT-DWT preprocessing

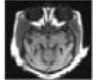



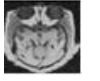


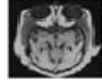

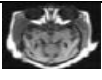






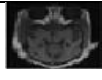























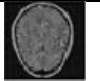








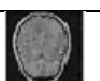


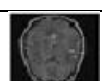








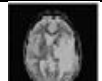

In the post-processing stage, a custom CNN is introduced for de-noising. The assessment of the suggested activation functions is done by comparing with existing ones. The improved results, both quantitative and qualitative, are shown in Tables 3.5 and 3.6. The evaluation result confirms that the proposed framework improves the details of the fused image.

The proposed model is compared with a recent CNN-based MMIF model, TDAN [128] as well as other existing image fusion models using CT-MRI images from the Harvard Medical School Brain dataset [78]. The comparative results are reported in Table 3.7, where the proposed model demonstrates higher SSIM and PSNR values. Additionally, Figure 3.10 presents a qualitative comparison of its image enhancement performance against various state-of-the-art methods. To assess the impact of denoising on SSIM at different noise levels, SSIM values were calculated at low (1% Gaussian noise), medium (10% Gaussian noise), and high (20% Gaussian noise) noise levels. Experimental results show that denoising significantly improved the SSIM of the fused image, particularly at the low noise level (from 0.915 to 0.975). There were also improvements at medium (from 0.85 to 0.90) and high (from 0.70 to 0.86) noise levels. This shows that as the noise level increases, the SSIM before denoising decreases, reflecting lower structural similarity. However, after denoising, SSIM values improve, demonstrating the effectiveness of the denoising process in restoring structural similarity.

Table 3.5 Evaluation of the proposed method in comparison with different activation functions

Activation functions	Proposed model with PCA [13] fusion method								Proposed model with mean-max [85] fusion rule							
	PSNR	RMSE	MI	SD	FMI	SSIM	FSIM	Entropy	PSNR	RMSE	MI	SD	FMI	SSIM	FSIM	Entropy
Logistic	50.5	0.285	1	0	0.814	0.933	0.87	5.278	49.22	0.355	1	0	0.77	0.905	0.912	4.407
ReLu	50.52	0.806	1	0	0.822	0.932	0.86	4.21	49.32	1.057	1	0	0.77	0.91	0.943	3.21
Identity	48.59	0.548	1.43	0.466	0.751	0.896	0.965	6.69	44.17	1.45	1.237	0.495	0.71	0.773	0.946	1.063
Tangent Hyperbolic	44.44	1.61	1.3834	1.57	0.806	0.868	0.961	6.95	54.67	0.295	1.223	0.497	0.704	0.976	0.938	6.044
Derivative of Tangent Hyperbolic	49.4	0.612	1.175	0.373	0.725	0.914	0.928	6.29	48.88	0.89	1.29	0.486	0.714	0.904	0.96	5.08
Gaussian	50.52	0.608	1	0	0.822	0.932	0.91	4.64	49.22	0.76	1	0	0.77	0.91	0.95	3.611
Proposed Activation function-1	57.40	0.1704	1.39	0.459	0.742	0.9895	0.96	6.91	56.85	0.126	1.22	0.284	0.724	0.987	0.91	6.46
Proposed Activation function-2	56.22	0.2066	1.315	0.441	0.722	0.985	0.95	6.56	56.27	0.196	1.17	0.2535	0.7224	0.984	0.899	6.2

Table 3.6 De-noised output of the proposed model with different activation functions on Harvard Medical School Brain dataset [78]

Images	Fusion rule	Fused image	Logistic	ReLu	Identity	Tangent Hyperbolic	Derivative Tangent Hyperbolic	Gaussian	Proposed activation function1	Proposed activation function2
Image set 1	PCA									
	Mean-Max									
Image set 2	PCA									
	Mean-Max									
Image set 3	PCA									
	Mean-Max									
Image set 4	PCA									





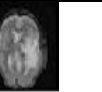


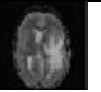
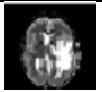
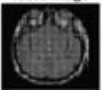


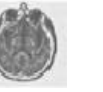



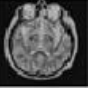








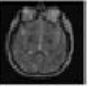

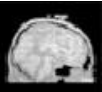








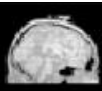








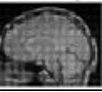
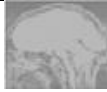







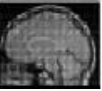


























	Mean- Max									
Image set 5	PCA									
	Mean- Max									
Image set 6	PCA									
	Mean- Max									
Image set 7	PCA									
	Mean- Max									
Image set 8	PCA									
	Mean- Max									

Table 3.7 Performance evaluation of suggested fusion model with contemporary models

Image fusion methods		FMI	SSIM	PSNR (dB)
GFF [116]		0.8664	0.6845	16.9681
RPCNN [120]		0.8767	0.6375	16.0925
LPCNN [107]		0.8771	0.6370	16.6865
NSST+PAPCNN [121]		0.8734	0.6626	16.8626
IFCNN [129]		0.8630	0.6856	17.0069
PMGI [130]		0.8755	0.2060	11.6736
U2fusion [131]		0.8577	0.4128	16.9593
EWC [132]		0.8730	0.6868	17.6090
Adapter [133]		0.8625	0.6619	17.3476
EWC+ Adapter		0.8721	0.6827	17.7391
OWM [134]		0.8773	0.6996	18.1964
TDAN [128]		0.8797	0.7032	18.0048
Proposed fusion models	Proposed model (PCA [13]) with activation function-1	0.742	0.9895	35.17
	Proposed model (PCA [13]) with activation function-2	0.724	0.987	35.09
	Proposed model (mean-max [85]) with activation function-1	0.722	0.985	34.99
	Proposed model (mean-max [85]) with activation function-2	0.7224	0.984	35.00

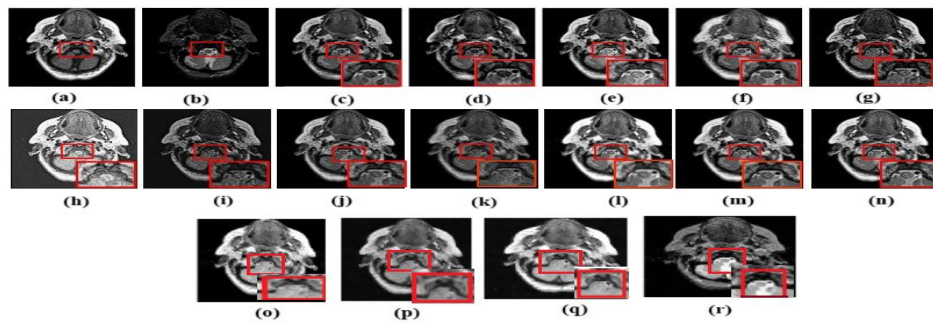


Figure 3.10 Output image of proposed image fusion model and existing models ((a) source image CT (b) source image MRI (c) GFF (d) RPCNN (e) LPCNN (f) NSST+PAPCNN (g) IFCNN (h) PMGI (i) U2 fusion (j) EWC (k) Adapter (l) EWC + Adapter (m) OWMI (n) TDAN (o) proposed model (PCA [13]) with activation function-1 (p) proposed model (PCA [13]) with activation function-2 (q) proposed model (mean-max [85]) with activation function-1 (r) proposed model (mean-max [85]) with activation function-2)

Based on MI, the suggested model is compared with the recently developed CNN-based image fusion model CELM [19]. The results reflect that the suggested model defeats CELM and seven other image fusion models for MRI, CT images. As shown in Table 3.8, the proposed model achieves significantly higher MI values compared to CELM. The qualitative analysis of the fused images is reported in Figure 3.11.

Table 3.8 Comparison of MI

Fusion parameter	CELM [19]	Proposed model (PCA [13]) with activation function-1	Proposed model (PCA [13]) with activation function-2	Proposed model (mean-max [85]) with activation function-1	Proposed model [mean-max [85]) with activation function-2
MI	1.0553	1.39	1.315	1.22	1.17

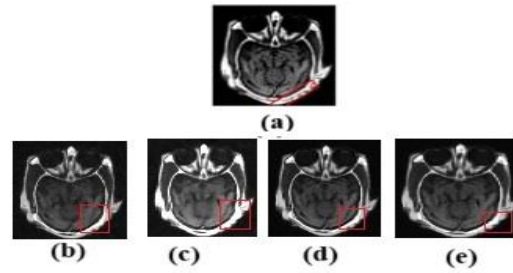


Figure 3.11 Output images of proposed model and CELM [19] ((a) CELM (b) proposed model (PCA [13]) with Ac func-1 (c) proposed model (PCA [13]) with Ac func-2 (d) proposed model (mean-max [85]) with Ac func-1 (e) proposed model ([mean-max [85]) with Ac func-2)

The assessment of the suggested model is evaluated on comparison with the recently developed CNN-based image fusion model CSID [135] and several other existing CNN-based models. The assessment is based on fusion performance parameters, including entropy, FMI, and computational costs, using the Harvard Medical School Brain dataset of CT-MRI images [78]. The results are detailed in Table 3.9, while Figure 3.12 provides a qualitative representation of the output. As shown in Table 3.9, the suggested model attains a higher entropy value while maintaining a similar FMI compared to existing models, but with a significantly lower computational cost.

Table 3.9 Comparative evaluation of existing algorithms and suggested algorithm

Performance Parameters	NSST-PA-PCNN [121]	CSR [136]	CSMCA [137]	CNN [125]	CSID [135]	Proposed model (PCA [13]) with activation function -1	Proposed model (PCA [13]) with activation function -2	Proposed model (mean-max [85]) with activation function-1	Proposed Model (mean-max [85]) with activation func-2
Image set 1									
Entropy	6.95 51	6. 48 71	6.32 74	6.75 41	6.99 71	7.0	6.96	6.846	6.89
FMI	0.45 59	0. 37 12	0.47 51	0.77 12	0.97 81	0.792	0.782	0.774	0.7724
Time in sec	5.08 3	24 .0 37	76.7 00	10.6 96	4.06 5	2.4	2.23	2.45	2.32
Image set 2									
Entropy	7.72 78	6. 08 67	7.61 82	7.84 21	8.01 42	6.78	6.54	7.23	7.22
FMI	0.55 97	0. 56 14	0.57 28	0.74 58	0.88 50	0.9465	0.7905	0.82	0.76
Time in sec	5.14 4	23 .4 41	74.9 94	10.4 47	4.05 1	2.50	2.56	2.36	2.45
Image set 3									
Entropy	5.33 29	5. 03 98	5.00 64	6.94 20	7.59 70	5.9531	6.267	6.5	6.2
FMI	0.55 36	0. 86 57	0.46 79	0.92 24	0.97 44	0.9105	0.7909	0.711	0.799
Time in sec	5.09 0	23 .3 39	76.0 18	9.58 1	4.04 9	2.33	2.45	2.98	3.1

Image set 4									
Entropy	5.05 98	4. 76 95	4.38 96	5.17 48	5.94 59	7.2554	7.3041	5.99	5.676
FMI	0.54 01	0. 84 71	0.49 39	0.84 21	0.98 14	0.7899	0.6174	0.798	0.987
Time in sec	5.23 2	22 .9 98	75.8 02	10.1 13	4.12 2	2.47	2.49	2.78	2.98
Image set 5									
Entropy	4.98 09	5. 02 97	5.93 3	5.99 89	6.28 74	5.2513	6.4186	6.34	6.45
FMI	0.73 60	0. 69 97	0.74 85	0.86 97	0.88 47	0.7337	0.7979	0.767	0.799
Time in sec	5.40 3	23 .4 22	76.1 12	10.6 91	4.04 1	2.33	2.48	2.980	2.95
Image set 6									
Entropy	4.90 38	4. 45 97	4.99 97	5.27 79	5.82 09	6.8996	6.6778	7.2	6.21
FMI	0.71 49	0. 68 39	0.80 97	0.85 27	0.88 17	0.877	0.7027	0.743	0.766
Time in sec	5.11 3	23 .4 83	76.7 72	10.8 34	4.04 7	2.47	2.50	2.34	2.95

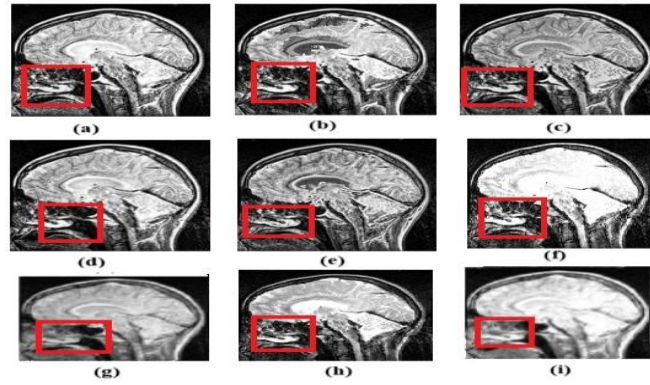


Figure 3.12 Output images of proposed model and existing image fusion models ((a) NSST-PA-PCNN (b) SCR (c) CSMCA (d) CNN (e) CSID (f) proposed model (PCA [13]) with Ac func-1 (g) proposed model (PCA [13]) with Ac func-2 (h) proposed model (mean-max [85]) with Ac func-1 (i) proposed model (mean-max [85]) with Ac func-2)

The suggested method is also compared with recent CNN-based models [138], [139] using the performance metric mean value. The assessment results are depicted in Table 3.10, and the corresponding images are depicted in Figure 3.13.

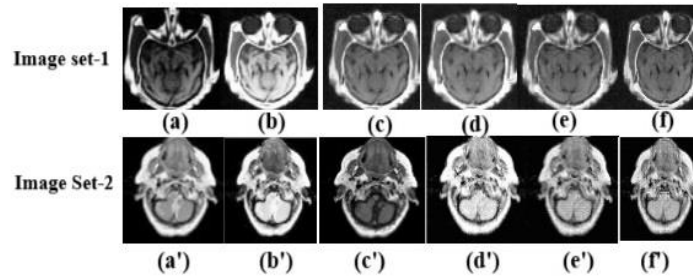


Figure 3.13 Assessment of the suggested model with recent developed MMIF models [138], [139] (((a) (a')) image fusion model [138] ((b) (b')) image fusion model [139] ((c) (c')) proposed model (PCA [13]) with Ac func-1 ((d) (d')) proposed model (PCA [13]) with Ac func-2 ((e) (e')) proposed model(mean-max [85]) with Ac func-1 ((f) (f') proposed model (mean-max [85]) with Ac func-2)

Table 3.10 Comparative analysis of MMIF model [138], [139] and the proposed model

Image fusion model	Mean value (M)	
	Image set-1	Image set-2
[138]	69.87	78.27
[139]	103.3	72.066
Proposed model (PCA [13]) with Activation function-1	99.57	95.29
Proposed model (PCA [13]) with Activation function-2	98.65	93.76

Proposed model (mean-max [85]) with Activation function-1	99.47	94.2
Proposed model (mean-max [85]) with Activation function-2	97.6	92.4

The performance of the suggested model is distinguished from the recently developed CNN-based MMIF model, ADAPTER [133], using the metric $Q^{AB/F}$.

Table 3.11 Comparison of the suggested model based on $Q^{AB/F}$ value with the model ADAPTER [133]

Name of the image fusion model	$Q^{AB/F}$ value
ADAPTER [133]	0.678 ± 0.20
Proposed model (PCA [13]) with Ac function-1	0.789
Proposed model (PCA [13]) with Ac function-2	0.799
Proposed model (mean-max [85]) with Ac function-1	0.765
Proposed model (mean-max [85]) with Ac function-2	0.756

Further the suggested method is distinguished from the MMIF model [140] based on the performance metric SSIM and $Q^{AB/F}$. The comparison values are reported in table 3.12 whereas the images are illustrated in figure 3.14.

Table 3.12 Comparison of proposed model with existing CNN based model [140]

Name of the model	SSIM value			$Q^{AB/F}$ value		
	Image set-1	Image set-2	Image set-3	Image set-1	Image set-2	Image set-3
Fusion model [140]	0.499	0.6814	0.8423	0.3839	0.3575	0.3473
Proposed model (PCA [13]) with Activation function-1	0.9895	0.9899	0.998	0.789	0.76	0.78
Proposed model (PCA [13]) with Activation function-2	0.9896	0.98789	0.9976	0.799	0.768	0.786
Proposed model (mean-max [85]) with Activation function-1	0.985	0.9896	0.9961	0.765	0.765	0.791
Proposed model (mean-max [85]) with Activation function-2	0.984	0.9879	0.9957	0.756	0.759	0.784

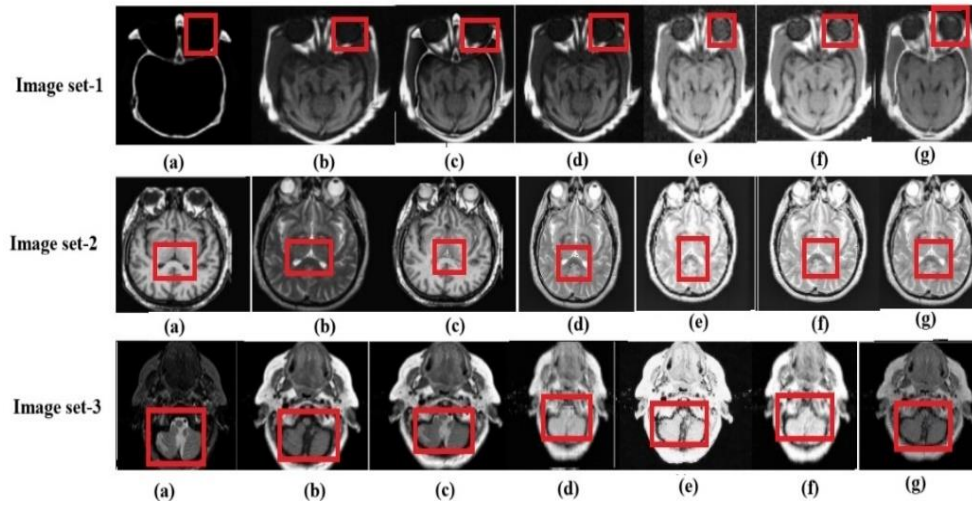


Figure 3.14 Comparison of fused images of suggested model with recently developed MMIF model ((a) source image (CT) (b) source image (MRI) (c) output of fusion model [140] (d) proposed model (PCA [13]) with Ac func-1 (e) proposed model (PCA [13]) with Ac func-2 (f) proposed model (mean-max [85]) with Ac func-1 (g) proposed model (mean-max [85]) with Ac func-2)

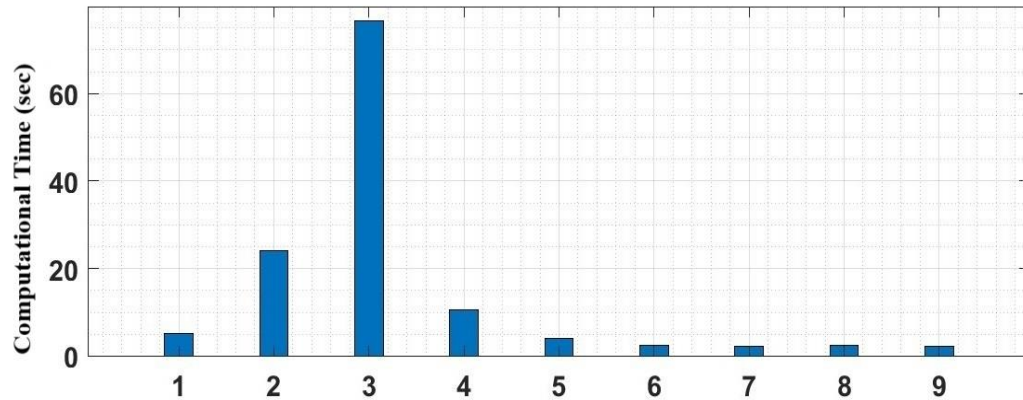


Figure 3.15 Comparison of computational time required for proposed image fusion models and existing image fusion models ((1), (2), (3), (4), (5), (6), (7), (8), (9) are NSST-PAP-CNN, CSR, CSMCA, CNN, CSID, proposed model (PCA [13]) with Ac function-1, proposed model (PCA [13]) with Ac function-2, proposed model (mean-max [85]) with Ac function-1, proposed model (mean-max [85]) with Ac function-2 respectively).

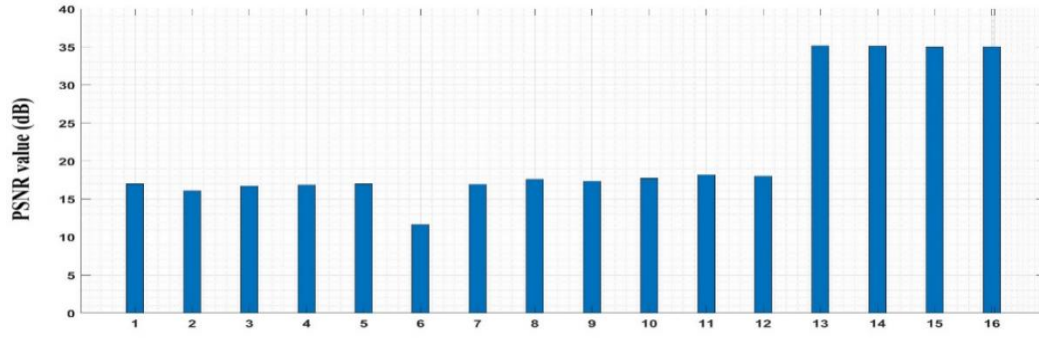


Figure 3.16 Comparison of PSNR Value achieved from proposed image fusion models and existing image fusion models ((1),(2),(3),(4),(5),(6),(7),(8),(9),(10),(11), (12), (13), (14), (15), (16) are GFF, RPCNN, LPCNN, NSNT+PA-PCNN, IFCNN, PMGT, U2Fusion, EWC, Adapter, EWC + Adapter, OWM, TDAN, proposed model (PCA [13]) with Ac function-1, proposed model (PCA [13]) with Ac function-2, proposed model (mean-max [85]) with Ac function-1, proposed model (mean-max [85]) with Ac function-2 respectively)

3.5 Conclusion

The proposed fusion model is systematically evaluated against several existing image fusion models using a range of performance metrics, including FMI, SSIM, PSNR, MI, entropy, mean value, and $Q^{AB/F}$. These metrics assess the fidelity of the de-noised images in comparison to their noisy counterparts, with particular focus on maintaining anatomical and physiological details essential for clinical accuracy. The results consistently show that the suggested model defeats existing methods across all these metrics. This enhanced performance is attributed due to the combination of AMT-DWT-based image enhancement and a custom CNN-based de-noising technique that incorporates a novel activation function. The AMT-DWT stage leverages multiple conventional image enhancement merits for the improvement the quality of source images. Although ReLU is commonly used for image de-noising, it may not be ideal in situations where preserving negative input information is important. To address this, a new activation function is introduced that produces highly de-noised images without increasing system complexity. The de-noised output effectively mitigates artefacts arising either from the imaging process or from image fusion, ensuring clearer, more accurate representations of the original images.

By using advanced techniques to analyze and learn complex patterns within medical images, the proposed model is able to generate more enhanced output compared to traditional methods, thereby improving both the clarity and clinical usefulness of medical image fusion.

CHAPTER 4

EFFICIENT IMAGE FUSION

4.1 Introduction

In medical image processing, the demand for multimodal image fusion is increasing, as it provides valuable details for early diagnosis [17]. In literature, various fusion models are existing to enhance the hidden details present in images from different modalities. But these models often face challenges like incomplete information in the fused images, due to the presence of unwanted artefacts, and issues with texture distortion and gradient preservation, especially in affected regions.

Currently, CNN-based MIF is used to rise the visual quality of images by automatically learning attributes from different sources [18]. Despite its advantages, this approach has some significant downsides, the above factors highlight the need for careful optimization of model parameters and training processes to balance performance and computational efficiency [18].

To address these limitations, this research introduces a composite deep learning model called Dense-ResNet for MMIF. The Dense-ResNet model integrates the benefits of Dense-net and ResNET and performs better in the reduction of vanishing gradient problem with better feature extraction. Initially, images from three distinct modalities are collected and pre-processed individually using a median filter to reduce unwanted noise without losing edge information. The pre-processed spatial domain images are then converted to the spectral domain through the DTCWT to make the system shift invariant. The transformed images are segmented using the Edge-Attention Guidance Network (ET-Net). Finally, the Dense-ResNet model is used to fuse the segmented images, achieving a superior multimodal fusion. The proposed model's performance is validated, showing that Dense-ResNet outperforms existing multimodal image fusion methods. It achieved a MSE of 0.402, RMSE of 0.634, PSNR of 47.136 dB.

Building on our earlier work, we introduced a novel fusion model that integrates the Pelican Optimization Algorithm (POA) with Dense-ResNet to fuse multimodal medical images from the BraTs 2020 dataset. Inspired by the natural hunting behavior of pelicans, POA was deployed to optimize the training procedure of the Dense-ResNet model. As in our earlier research, median filtering was applied for preprocessing, DTCWT was used for image decomposition, and ET-Net was employed for edge preservation. The evaluation results illustrate that the proposed model outperformed the previous one, achieving lower MSE and RMSE values along with a higher PSNR.

Several transformation techniques can decompose multimodal images. After special processing, the decomposed components are combined to achieve better information retention, reduced noise, and improved signal quality in the final image. In transform domain, achieving superior fusion results does not necessarily require a large dataset. Numerous fusion rules have been explored in the literature, with some, such as PCA [13] and mean-max [85], proving effective for fusing images from different modalities. However, these rules often suffer from spatial distortion, artifact introduction, and higher system complexity.

To address these limitations, we suggested a new pair of fusion rules to merge the approximate and detail coefficients of multimodal medical images, resulting in higher fusion efficiency. This chapter discusses the methodology, work carried out, and the results obtained in developing an efficient image fusion model. In medical image processing, MMIF is in rising demand due to the availability of information for early diagnosis. In the recent past, different fusion approaches are used to intensify the hidden information present in different modality images. But insufficient information due to the existence of unwanted artefacts in the fused image remains a big challenge. The methodology, work-done, and results obtained in developing efficient image fusion model is discussed in details in this chapter. Hence for the efficient image fusion, we have developed the followings:

- An efficient Dense-ResNet for multimodal image fusion using medical image
- Multimodal fusion of different medical image modalities using optimized hybrid network
- Proposed a pair of fusion rules which can denoise the fused image effectively.

4.2 An efficient Dense-ResNet for multimodal image fusion using medical image

The proposed (or suggested) approach presents a unique deep learning framework tailored for fusing medical images from multiple modalities. It begins by utilizing images from three different modalities, each subjected to distinct preprocessing techniques, including median filtering [15] to improve image quality. Subsequently, the preprocessed images are transformed into the spectral domain using DTCWT. These spectral domain images are then segmented through the Edge-Attention Guidance Network (ET-Net) [141]. The fusion is executed using a hybrid Dense-ResNet model across three layers, effectively integrating all segmented images. For the assessment of the Dense-ResNet model, it is benchmarked against other multimodal image fusion methods. The evaluation demonstrates that the Dense-ResNet model achieves a minimum MSE of 0.402, a RMSE of 0.634, and a maximum PSNR of 47.136 dB.

4.2.1 Methodology

Development of the Dense-ResNet model for Multimodal Fusion: The proposed Dense-ResNet model facilitates the combination of medical images captured from three distinct modalities. Various medical images, such as $T1$, $T1GD$, and $T2$, are utilized for the fusion process. These images undergo separate preprocessing steps, including denoising via a median filter to mitigate inherent noise. The spatially filtered images are then transformed into the spectral domain using DTCWT [138] and segmented using ET-Net [141] for precise fusion evaluation. Subsequently, the resulting segmented images from the three modalities are fused using the designed Dense-ResNet model, which combines the architectures of Dense-net and Residual Net (ResNet).

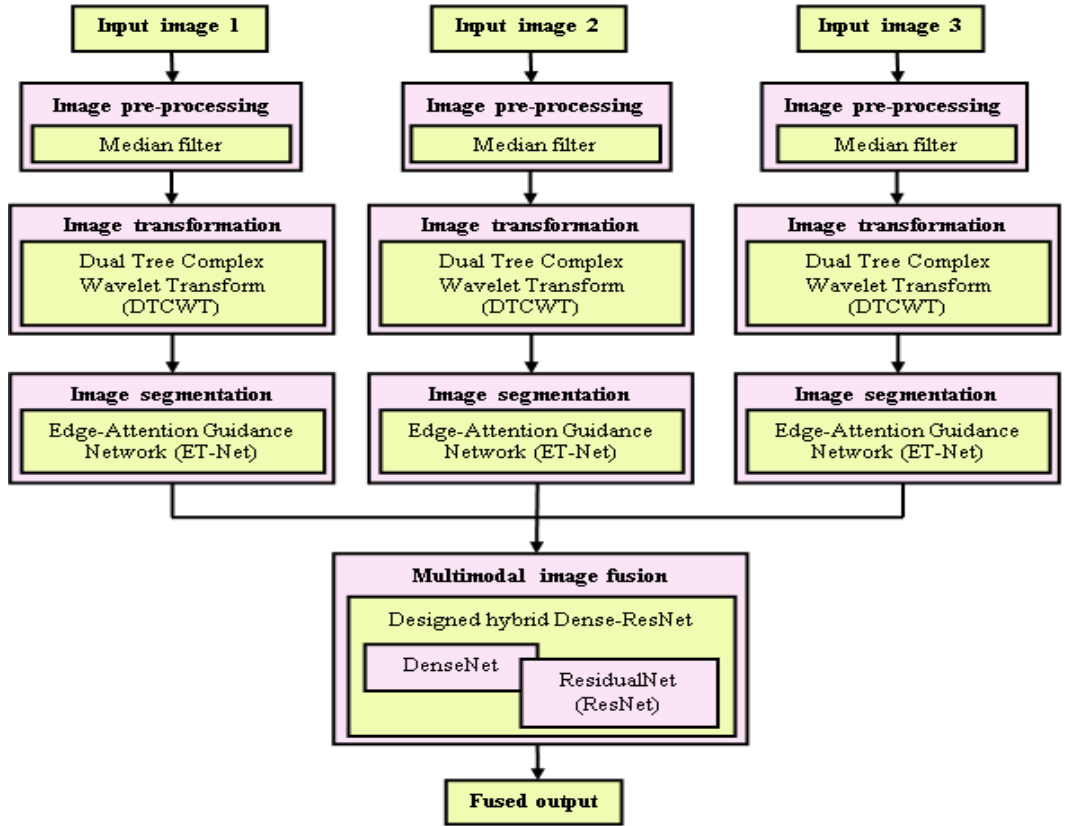


Figure 4.1 A framework of designed Dense-ResNet for MMIF

Image acquisition

The images procured from three modalities in the database for MMIF are represented below:

$$M = \{M_1, M_2, M_3, \dots, M_A, \dots, M_Z\} \quad (4.1)$$

where, the medical database taken for the MMIF is represented as M , the A^{th} Native ($T1$) input image considered for the fusion process is signified as M_A , and Z signifies the number of images available in the dataset as a whole.

Image pre-processing

Image preprocessing is a vital step in the MMIF process, to enhance the quality of the images and make them appropriate for machine learning models [155], [156], [157]. The denoising of the selected $T1$ input image to remove naturally occurring noise occurs during the image pre-processing stage. In this phase, a median filter [158] is employed to preprocess the input image, utilizing statistical methods for effective noise removal. The median filter serves as a nonlinear technique that preserves crucial image information by transforming the original gray pixel values into median gray pixel values. This method effectively eliminates unwanted information (noise) with securing its edges.

Median filtering [15] is utilized primarily for noise reduction. The expression for the median-filtered $T1$ input image is given as follows:

$$M_p = \text{Median} (M_A) \text{ for } A \in R_p \quad (4.2)$$

where the neighborhood sub-image is represented as R_p . The selected $T1$ input image M_A is pre-processed by utilizing a median filter to record the pre-processed output image M_p .

Image transformation

The processed $T1$ input image is then exposed to image transformation to convert it into the spectral domain. In this step, the DTCWT [138] is applied to achieve a spectral domain representation. DTCWT is an advancement of the DWT that utilizes both low and high wavelet filters to derive wavelet coefficients with real and imaginary components. This transformation involves selective and directional shifting of the image across multiple dimensions, rendering the process invariant.

The expressions for the DTCWT transformation are given as follows:

$$\chi(l) = \chi_a(l) + j\chi_b(l) \quad (4.3)$$

$$\delta(l) = \delta_a(l) + j\delta_b(l) \quad (4.4)$$

In these equations, j represents the imaginary term, $\chi(l)$ denotes the complex wavelet, and $\delta(l)$ signifies the complex scaling function. Thus, the pre-processed $T1$ input image M_p is transformed using DTCWT into a spectral domain $T1$ image MT, enabling the subsequent image segmentation process.

Segmentation of medical image

During the image segmentation phase, extracting the target area from the transformed $T1$ image is crucial for achieving accurate and precise multimodal medical image

fusion. In this phase, the Edge-Attention Guidance Network (ET-Net) [141] is employed to extract the target area from the transformed images. ET-Net utilizes a specialized approach to extract primary regions while disregarding edge information to ensure accurate results. It leverages the representation of edge attention in encoding layers and transfers this information to multi-scale decoding layers for effective segmentation.

Structure of ET-Net

ET-Net works as an encoder-decoder network comprising an Edge Guidance Module (EGM) and a Weighted Aggregation Module (WAM). Its main role is to embed edge-attention representation for image segmentation. Within ET-Net, the EGM preserves the characteristics of local edges at the encoding layer, while the WAM aggregates multi-scale side outputs. Specifically, ET-Net uses ResNet-50 network for the purpose of encoding, consisting of four encoding blocks. The input undergoes feature extraction through a stream of convolutional layers, culminating in the final output. ET-Net generates high-level features via residual connections. The decoder part consists of three decoding blocks for enhancing representation capability and managing high-level feature characteristic. The decoding blocks employ convolution based on the depth function to augment the representation of high and low features, followed by convolution to unite the total number of channels. Figure 4.2 illustrates the structure of ET-Net, where the transformed $T1$ image undergoes segmentation via ET-Net to produce the segmented image output.

Let $M_{S(1)}$ denotes the segmented image output obtained when the transformed $T1$ image M_T is segmented using ET-Net. Similarly, the image pre-processing, image transformation, and segmentation using ET-Net are performed for the other two medical images, such as $T1GD$, and $T2$. Here, the segmented image output $M_{S(2)}$ is recorded for input $T1GD$ medical image and $M_{S(3)}$ input $T2$ medical image, during segmentation through ET-Net. The segmented image output obtained from input $T1$ image, $T1GD$ image, and $T2$ image is allowed for multimodal image fusion using the designed Dense-ResNet model.

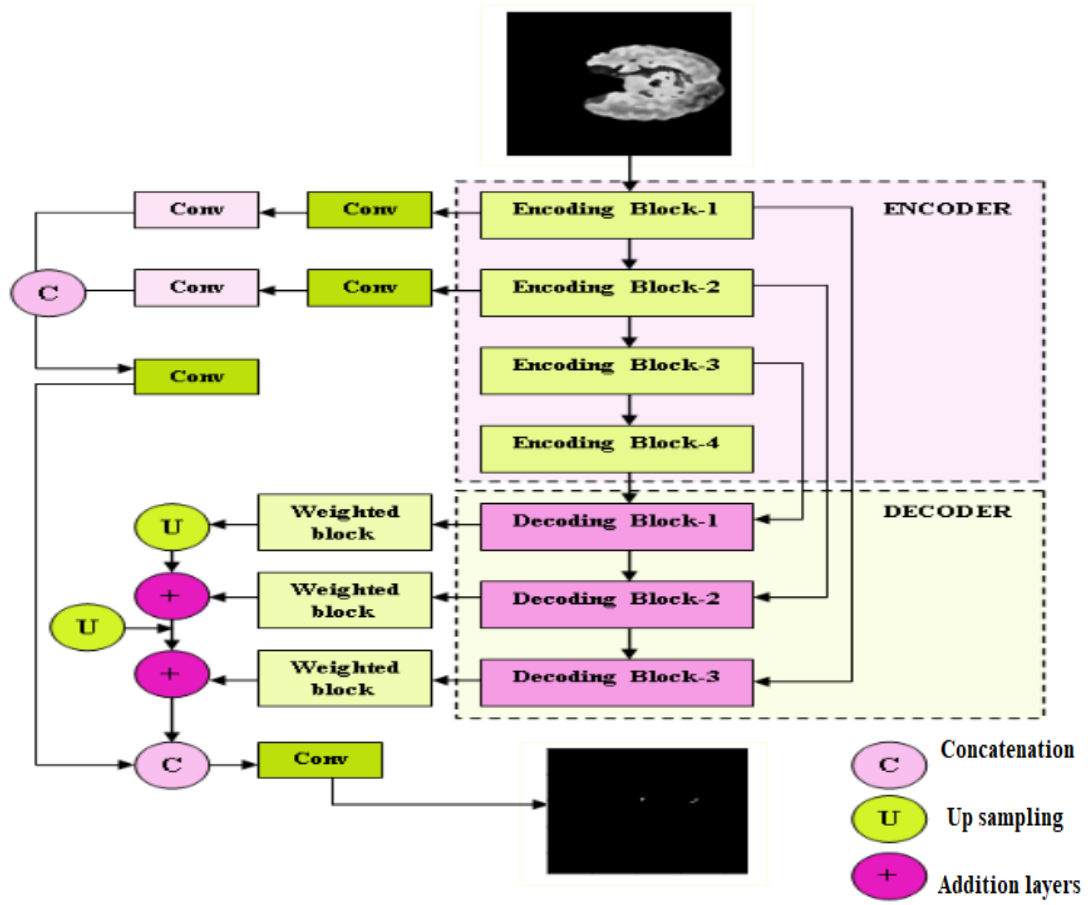


Figure. 4.2 Architecture of ET-Net

Image fusion using Dense-ResNet model

Image fusion in medical imaging is a critical technique aimed at integrating essential details from multiple modalities to give a clearer, more informative output. This process enhances diagnostic accuracy by blending images from different modalities, such as $T1GD$, $T1$, and $T2$, which each provide unique details about the underlying anatomy or pathology. The purpose is to synthesize this information into a single image that retains or enhances the details and clarity of the original images.

In the suggested methodology, the segmented outputs from the input images $M_{S(1)}$ ($T1$), $M_{S(2)}$ ($T1GD$), and $M_{S(3)}$ ($T2$) are fused using a hybrid Dense-ResNet model. This hybrid model integrates the strengths of two powerful neural network architectures: Dense-net [73] and ResNet [41].

Dense-net provides efficient parameter utilization by ensuring that once features are learned, they are not redundantly relearned. This characteristic makes Dense-net highly efficient, requiring fewer parameters to process input data while preserving significant amounts of image information—an advantage particularly relevant to medical images where fine details must be preserved for accurate interpretation. ResNet, in contrast, excels at reducing error rates by allowing deep networks to learn efficiently through the application of residual connections. These connections help

reduce issues like vanishing gradients, resolve it possible to train deeper networks without a loss in performance. This contributes to better feature extraction and ultimately more accurate fusion of multimodal images.

By combining these architectures, the hybrid Dense-ResNet model achieves an ideal alignment between performance and computational complexity, providing a robust and viable remedy for medical image fusion. Dense-net's ability to preserve features complements ResNet's capacity for deep learning, enabling the model to generate a more precise and enhanced fused image. This fusion process is crucial in medical imaging, as the resulting image integrates complementary information from different modalities, leading to improved visualization of anatomical structures or pathological regions. This, in turn, supports more accurate diagnosis and treatment planning.

The hybrid Dense-ResNet model is formulated by employing regression modeling using Fractional Calculus (FC) [33]. It comprises three key modules: the Residual model, Dense-ResNet layer, and Dense-net model. The input medical images $\{M_{A(k)}\}; k = 1, 2, 3$ of dimension $(p \times q)$ are allowed inside the residual model, which comprises ResNet. The dimensionality of the input image can be represented as $(p \times q) \times k$, here k is the total input images. The resultant output achieved from the residual model and the segmented image output $\{M_{S(k)}\}; k = 1, 2, 3$ are fed into the Dense-ResNet layer for fusion and regression. Here, fusion is carried out between the ResNet and Dense-net models to combine the images, and the relationship between the inputs is determined using regression modeling. Subsequently, the output of the Dense-ResNet model is fed into the Dense-net model for fusion. Figure 4.3 depicts the block diagram of the designed hybrid Dense-ResNet model.

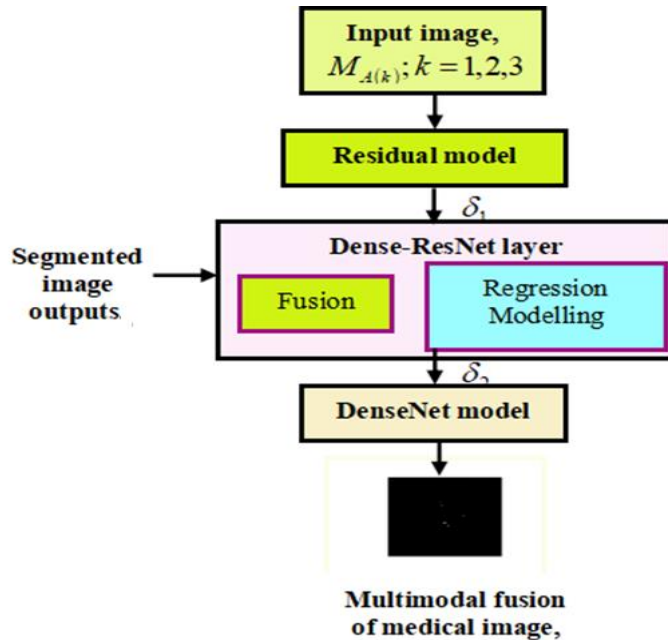


Figure 4.3 Block diagram of the designed hybrid Dense-ResNet model

Residual model

The ResNet (Residual Network) architecture is highly regarded in deep learning, especially for image processing, due to its innovative approach to training very deep neural networks. It effectively addresses challenges such as vanishing gradients, which can occur in traditional deep networks. Here's an overall description the key components and features of the ResNet architecture:

Convolutional layer

This layer is utilized to minimize the number of parameters needed for training while improving performance through weight sharing and local receptive fields. It comprises kernels or filters arranged in sequence within a localized area to process the input data. Each filter moves across the input matrix, calculating dot products at each position concerning the kernel. The computational process in the convolutional layer can be reported as follows:

$$\delta_1 = \sum_{r=0}^{w-1} \sum_{s=0}^{w-1} N_{e,s} * M_{A(k)(g+r),(h+s)} \quad (4.5)$$

Here, the input image is indicated as, $M_{A(k);k=1,2,3}$ the coordinates are represented as g and h , the kernel matrix of size $w \times w$ is denoted as N . The index position of the kernel matrix is signified as r and s also, the operator used for cross-correlation without zero padding is given by $*$.

Average pooling layer

This layer is essential for substantially decreasing the dimensions of the feature map when applied sequentially after the convolutional layers. Furthermore, it helps to alleviate overfitting concerns. The average pooling layer functions independently on each feature map by evaluating the average of the values within the receptive field. The pooling operation carried out in the average pooling layer can be expressed as follows:

$$B_{out} = \frac{B_{in} - t_r}{d} + 1 \quad (4.6)$$

$$H_{out} = \frac{H_{in} - t_s}{d} + 1 \quad (4.7)$$

Here, H_{in} and B_{in} represents the height and width of a two-dimensional input matrix, where the corresponding output is given as B_{out} and H_{out} . The term t_r indicates kernel width and t_s denotes kernel height.

Activation function

This component is crucial for capturing nonlinear features present in the database and determining the nonlinearity of the extracted attributes using a nonlinear activation function. Let $M_{A(k)}$ represent the input feature. Typically, the ReLU is the widely utilized activation function, defined as follows:

$$\text{ReLU}(\delta_1) = \begin{cases} 0 & M_{A(k)} < 0 \\ M_{A(k)} & M_{A(k)} \geq 0 \end{cases} \quad (4.8)$$

Batch normalization

In the batch normalization phase, the training set is partitioned into mini-batches to make a balance between computational complexity and convergence. Batch normalization is then applied to normalize the input layers, reducing the shifting of internal covariates. This process adjusts the activation by improving reliability and training speed.

Residual blocks

Within residual blocks, a shortcut connection is established between inputs and outputs. In certain scenarios where both the dimensions of output and input are alike, the input is directly connected to the output. This connection is expressed as:

$$X = F(M_A) + M_A \quad (4.9)$$

Here, when the dimensions are different, a matching dimension factor is typically devised to ensure alignment between the input and output dimensions. This matching dimension factor is expressed as:

$$X = F(M_A) + \psi M_A \quad (4.10)$$

From the expression, input and output obtained from the residual block is represented as M_A and X . Also, the mapping alliance of input and output is represented by the function F , and the matching factor is indicated as ψ .

Linear classifier: The classification results are determined with the help of the softmax function and fully interconnected layers in the linear classifier. Each neuron is interconnected from one to another using the principles of a multi-layer perceptron in the fully connected layer. The fully connected layer utilizes dot products, which is expressed as:

$$X = \psi_{u \times v} M_{A(p \times q)} + y_{u \times v} \quad (4.11)$$

Here, the weight of the matrix with dimension $u \times v$ is represented as $\psi_{u \times v}$, the input image with $p \times q$ dimension is signified as $M_{A(p \times q)}$, also the bias is denoted by y , The

final output is determined by the fully interconnected layer and cannot obtain classification results directly. The soft-max function is utilized to normalize the input vector of the probability belonging to each class. The utilization of the soft-max function is given by the expression as follows,

$$\chi = \frac{e^{\chi_l}}{\sum_{t=1}^{\rho} e^{\chi_t}} \quad (4.12)$$

One of the output layer elements is indicated as χ_l and the output dimension is given as ρ .

Regression layer: In regression tasks, the loss is obtained using the half-mean-squared error within the regression layer. This layer helps normalize the responses, which accelerates and stabilizes the training process in neural networks. Figure 4.4 demonstrate the architecture of the ResNet layer, and the fused image produced from the convolutional layer is then employed in the fusion process.

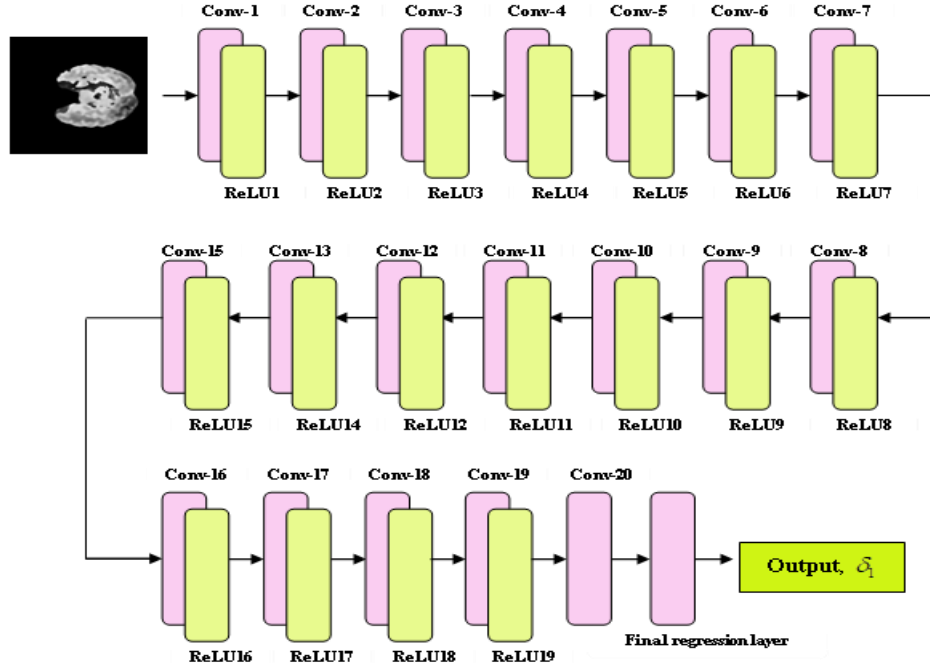


Figure 4.4 Architecture of Residual model

Dense-ResNet layer

Once the output is generated using the Residual model, it is input into the Dense-ResNet model, along with the segmented image outputs, to analyze the relationships between the inputs. Regression modeling is integrated into the system, and regression is performed using Fractional Calculus (FC), [33]. FC is utilized to solve various integral and differential equations using the Laplace transform. The output from the Dense-ResNet layer is obtained by aggregating weighted features over different time intervals. The outputs from the Dense-ResNet layer can be demonstrated as follows:

The output of the Dense-ResNet layer at interval t is expressed as:

$$h_1 = \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^1 * W_{x,y} \quad (4.13)$$

Here, the weight coefficient is denoted as W , and the segmented image input is represented as $M_{S(k)}$, the total rows and columns indicated as x and y . Moreover, Dense-ResNet layer output obtained at $(k-1)^{th}$ interval is given by,

$$h_2 = \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^2 * W_{x,y} \quad (4.14)$$

here, h_2 represents the resultant output. At $(k-2)^{th}$ intervals, the Dense-ResNet layer output is given by,

$$h_3 = \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^3 * W_{x,y} \quad (4.15)$$

Considering the concept of FC [33] for regression modeling, the expression is given by,

$$z(p+1) = \gamma \cdot z(p) + \frac{1}{2} \gamma \cdot z(p-1) + \frac{1}{6} (1-\gamma) z(p-2) + \frac{1}{24} \gamma (1-\gamma) (2-\gamma) \cdot z(p-3) \quad (4.16)$$

Hence, applying the outputs obtained by the Dense-ResNet layer at various time intervals,

$$\delta_2 = \gamma \cdot h_1 + \frac{1}{2} \gamma \cdot h_2 + \frac{1}{6} (1-\gamma) h_3 + \frac{1}{24} \gamma (1-\gamma) (2-\gamma) \cdot \delta_1 \quad (4.17)$$

By substituting the value of z_1 , z_2 , and z_3 , thus the equation becomes,

$$\begin{aligned} \delta_2 = & \gamma \cdot \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^1 * W_{x,y} + \frac{1}{2} \gamma \cdot \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^2 * W_{x,y} + \frac{1}{6} (1-\gamma) \cdot \sum_{x=1}^u \sum_{y=1}^v M_{S(k)_{x,y}}^3 * W_{x,y} \\ & + \frac{1}{24} \gamma (1-\gamma) (2-\gamma) \cdot \delta_1 \end{aligned} \quad (4.18)$$

Here, the order of the derivative is represented by the term γ , and the output obtained by the fusion of the Dense-ResNet layer is given as δ_2 .

Dense-Net model

Dense-net is prevalently used in deep learning classification tasks because it helps solve the problem of vanishing gradients. It creates direct links between each layer and all previous layers, which improves the flow of features and promotes feature reuse.

This approach reduces the overall number of parameters that need to be trained. The architecture of Dense-net consists of several key components:

- a. **Convolutional Layers:** These layers apply filters to the input data to extract relevant features.
- b. **Dense-net Blocks:** Within these blocks, each layer gets inputs from all former layers, advancing feature reuse and mitigating the vanishing gradient problem.
- c. **Transition Layers:** Positioned intermediate of dense-net blocks, these layers perform down-sampling operations, such as pooling, to reduce the extent of feature maps and manage the model's complexity.

Throughout the architecture, ReLU activation functions are used to establish non-linearity, while a Soft-Max function is introduced in the final layer to produce probability distributions for classification tasks. In reference to MMIF, the output from a Dense-ResNet layer can be fed into the Dense-net model to achieve comprehensive integration of details from various imaging modalities.

Convolutional layer: In this layer, activation occurs through the application of filters to the input. The intensity of features at different positions within the input is represented by a feature map. To create these feature maps, multiple filters are employed, which are subsequently processed through an activation function. Typically, the filter size in the convolutional layer is smaller than that of the input data, and a dot product operation is performed to process the two components. The evaluation of the non-linear input can be expressed as:

$$E_{uv}^j = \sum_{C=0}^{f-1} \sum_{D=0}^{f-1} \beta_{CD} F_{(U+C)(V+D)}^j \quad (4.19)$$

where, E_{uv}^j indicates the non-linear input, the coordinates are signified as U and V , the kernel matrix of size $f \times f$ is indicated as β . The term C and D are the index position of the kernel matrix, $*$ is the operator used for cross-correlation without zero padding. The non-linearity assessed by the convolutional layer is given by the expression,

$$F_{uv}^j = \kappa(E_{uv}^j) \quad (4.20)$$

Here, κ is the parameter of the previous layer.

Max pooling layer: The max pooling layer in CNN significantly lessen the dimensionality of the feature map. This layer runs a filter over the feature map, and the pooling filter is used to summarize the features within the region. Let us consider a feature map with dimension $P \times Q \times R$, which indicates the feature map height, width, and channels. The dimension of the feature map, once the max pooling is applied is expressed by,

$$Max.pooling = \frac{(P - \varepsilon + 1)}{k} \times \frac{(Q - \varepsilon + 1)}{k} \times R \quad (4.21)$$

here, ‘ ε ’ represents the filter size and ‘ κ ’ denotes stride.

Dense-net layer: In the dense-net layer, neurons are dense-netly interconnected with neurons from the preceding layer. Inputs are received by the neurons of the dense-net layer and undergo matrix manipulation. This process is expressed as:

$$\mu.K = \begin{matrix} & V_{11} & V_{12} & \cdots & V_{1x} & \varpi_1 \\ & V_{21} & V_{22} & \cdots & V_{2l} & \varpi_2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ & \vdots & \vdots & \cdots & \vdots & \vdots \\ & V_{w1} & V_{w2} & \cdots & V_{wx} & \varpi_w \end{matrix} \quad (4.22)$$

The variable M represents a matrix with dimension $w \times z$, ϖ is a matrix with a dimension $1 \times w$, and K is the parameter of the prior layer, which is modified using back propagation while training. The associated weights are given by the expression,

$$\lambda^{jx} = \lambda^{jx} - \eta \times T\lambda^{jx} \quad (4.23)$$

$$\sigma^{jx} = \sigma^{jx} - \eta \times T\sigma^{jx} \quad (4.24)$$

Where the associated weights of jx layer are represented as λ^{jx} and the associated bias is signified as σ^{jx} , the rate of learning is indicated as η . The loss function partial derivatives are denoted as $T\lambda$ and $T\sigma$, which is computed by using the chain rule given by,

$$T\lambda^{jx} = \frac{\partial \iota}{\partial \lambda^{jx}} = \frac{1}{\alpha} TL^{jx} I^{(jx-1)G} \quad (4.25)$$

$$T\sigma^{jx} = \frac{\partial \iota}{\partial \sigma^{jx}} = \frac{1}{\alpha} \sum_{U=1}^{\alpha} TL^{jx(U)} \quad (4.26)$$

$$TI^{(jx-1)} = \frac{\partial \iota}{\partial I^{(jx-1)}} = Y^{jx} TL^{jx} \quad (4.27)$$

$$TL^{jx} = TI^{jx} \times c \cdot (L^{jx}) \quad (4.28)$$

Here, the linear activation at jx layer is given by L^{jx} , the non-linear function based on L^{jx} is represented as $c \cdot (L^{jx})$, and the non-linear activation function is indicated as I^{jx}

Transition layer: The transition layer in CNN is involved to effectively lessen the complexity of the model. It diminishes the sum of channels in the convolutional layer, thereby reducing the input height and width by half using stride 2 filters. The resultant fused multimodal image output obtained by the Dense-ResNet model is represented as

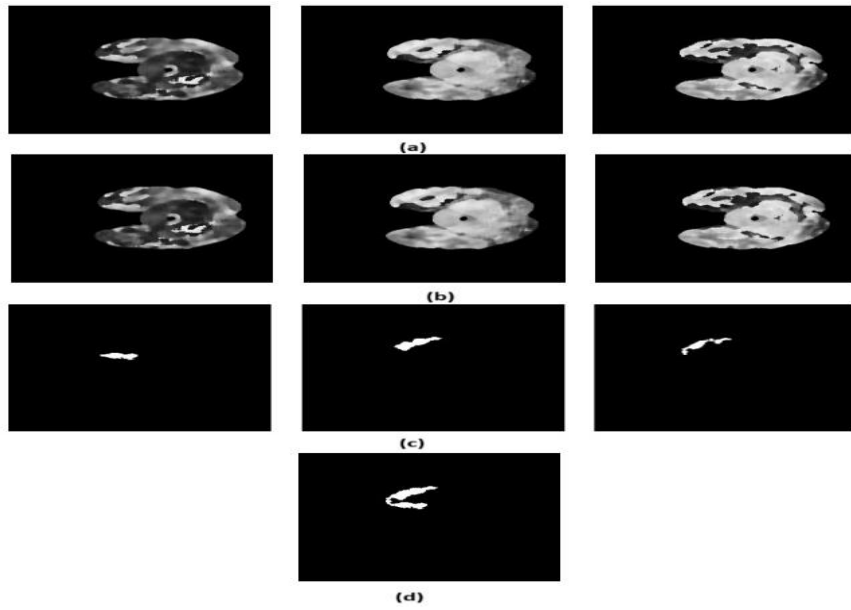


Figure.4.7 Experimental image results of Dense-ResNet model using *Flair*, *T1*, and *T1C* images ((a) Input *Flair*, *T1*, and *T1C* image, (b) Filtered *Flair*, *T1*, and *T1C* image, (c) Segmented *Flair*, *T1*, and *T1C* image, (d) Fused multimodal image)

The performance of the Dense-ResNet model designed for MMIF is analyzed by means of K-value and training data.

By altering K-value

Table 4.1 elucidates the analysis of the performance of Dense-ResNet using K-value for various evaluation indicators. In table 4.1(a), the analysis of MSE measured by Dense-ResNet for the epochs 20, 40, 60, 80, and 100 epochs are 0.885, 0.863, 0.574, 0.147, and 0.058 for the K-value of 8. The evaluation of PSNR is shown in table 4.1(b). The Dense-ResNet obtained PSNR values of 37.105dB, 39.685dB, 43.885dB, 45.897dB, and 47.872dB for 20, 40, 60, 80, and 100 epochs for the 8 K-value. table 4.1(c) shows, the RMSE measured by the Dense-ResNet approach. For K-Value 8, the RMSE achieved by the Dense-ResNet is 0.940 for 20, 0.929 for 40, 0.757 for 60, 0.383 for 80 and 0.240 for 100 epochs.

Table 4.1 Performance evaluation of Dense-ResNet model using K-value (a) MSE, (b) PSNR, and (c) RMSE

Metrics	K-Value	Dense-net ResNet with Epoch20	Dense-net ResNet with Epoch60	Dense-net ResNet with Epoch80	Dense-net ResNet With Epoch100
(a) MSE	5	0.865	0.572	0.145	0.056
	6	0.870	0.574	0.146	0.057
	7	0.881	0.574	0.147	0.058
	8	0.885	0.574	0.147	0.058

	9	0.887	0.574	0.147	0.058
(b) PSNR (dB)	5	30.481	35.553	39.558	41.502
	6	31.225	38.113	42.113	43.407
	7	32.762	41.567	43.571	45.211
	8	37.105	43.885	45.897	47.872
	9	41.199	45.085	47.099	50.395
(c) RMSE	5	0.930	0.757	0.381	0.238
	6	0.933	0.757	0.383	0.240
	7	0.938	0.757	0.383	0.240
	8	0.941	0.758	0.383	0.240
	9	0.942	0.758	0.384	0.240

The performance of the Dense-ResNet technique by altering training data is displayed in table 4.2. Table 4.2(a) shows the evaluation of the Dense-ResNet technique using MSE, for 80% of training data, the MSE measured by the Dense-ResNet for epochs 20, 40, 60, 80, and 100 is 0.815, 0.756, 0.671, 0.597, and 0.473. The PSNR measured by Dense-ResNet is given in table 4.2(b). For 80% training, the Dense-ResNet measured PSNR of 38.251dB for 20 epochs, 39.722dB for 40 epochs, 42.897dB for 60 epochs, 43.001dB for 80 epochs, and 47.242dB for 100 epochs. The RMSE value of the Dense-ResNet is given in table 4.2(c), for 80% of training data, the RMSE value obtained by the Dense-ResNet are 0.903, 0.869, 0.819, 0.772, and 0.688 for here for epochs of 20, 40, 60, 80, and 100, respectively.

Table 4.2 Performance analysis of Dense-ResNet model utilizing training data (a) MSE, (b) PSNR, and (c) RMSE

Metrics	Learning Rate	Dense-ResNet With Epoch20	Dense-ResNet With Epoch40	Dense-ResNet With Epoch60	Dense-ResNet With Epoch80	Dense-ResNet With Epoch100
(a) MSE	0.50	0.756	0.736	0.623	0.563	0.472
	0.60	0.763	0.741	0.659	0.568	0.473
	0.70	0.766	0.746	0.664	0.581	0.473
	0.80	0.815	0.756	0.671	0.597	0.473
	0.90	0.822	0.767	0.682	0.622	0.492
(b) PSNR (dB)	0.50	30.858	32.825	35.202	36.254	39.309
	0.60	31.353	33.943	37.832	39.840	42.294
	0.70	34.233	35.058	39.422	41.422	45.288
	0.80	38.251	39.722	42.897	43.001	47.242
	0.90	42.118	45.956	47.439	49.571	51.176
(c) RMSE	0.50	0.869	0.858	0.789	0.750	0.687
	0.60	0.874	0.861	0.812	0.753	0.688
	0.70	0.875	0.864	0.815	0.762	0.688
	0.80	0.903	0.869	0.819	0.772	0.688
	0.90	0.906	0.876	0.826	0.789	0.701

Comparative discussion

The results obtained from the experiment to analyze the efficiency of the designed Dense-ResNet for MMIF by comparing with prevailing fusion techniques are displayed in table 4.3. The MMIF using the Dense-ResNet approach achieved superior results than prevailing approaches. The results revealed that the designed Dense-ResNet model achieved a minimum of 0.403 MSE, 0.635 RMSE, and a maximum of 47.137dB PSNR than other techniques.

Table 4.3 Comparative discussion

Evaluation parameters	Techniques								
	Deep TL model	GAN	CNN-HOD	G-CNN and fuzzy neural net works	MART	CELM	Dense-net	ResNet	Designed Dense-ResNet
For K-value									
MSE	1.221	1.086	0.894	0.724	0.680	0.651	0.639	0.612	0.443
PSNR (dB)	37.421	39.534	44.145	44.153	40.988	39.568	43.654	44.259	46.155
RMSE	1.105	1.042	0.945	0.851	0.825	0.807	0.800	0.782	0.665
For training data									
MSE	1.110	0.987	0.812	0.658	0.618	0.592	0.581	0.556	0.403
PSNR (dB)	38.217	38.332	45.085	45.092	40.185	41.026	41.259	44.098	47.137
RMSE	1.053	0.994	0.901	0.811	0.786	0.769	0.762	0.746	0.635

4.3 Multimodal fusion of different medical image modalities using optimized hybrid network

This approach introduces a unique deep learning framework designed for fusing medical images from different modalities. This research work is the extension of our previous research work with the application of an optimization process at the training stage of the hybrid network. It starts with images from three modalities $T1$, $T1GD$ and $T2$, each undergoing specific preprocessing steps like median filtering [15] to enhance quality. After preprocessing, the images are recasted into the frequency domain using the Dual-Tree Complex Wavelet Transform (DTCWT). Next, the spectral domain images are segmented using the Edge-Attention Guidance Network (ET-Net). The fusion process is carried out through a hybrid Dense-ResNet model.

4.3.1 Methodology

In this research, POA+Dense-ResNet is developed for multimodal image fusion. Here, the combination of ResidualNet and Dense-net, called Dense-ResNet, which is trained using the POA (Trojovský and Dehghani, 2022). The DTCWT is used to perform the

spatial domain to spectral domain transformation after the input images from various modalities have undergone pre-processing. The ET-Net segments these transformed images. The POA+Dense-ResNet achieve the fusion in the last stage. Figure 1 shows the block diagram of the POA+Dense-ResNet. The block diagram of the suggested model is depicted in figure.4.8.

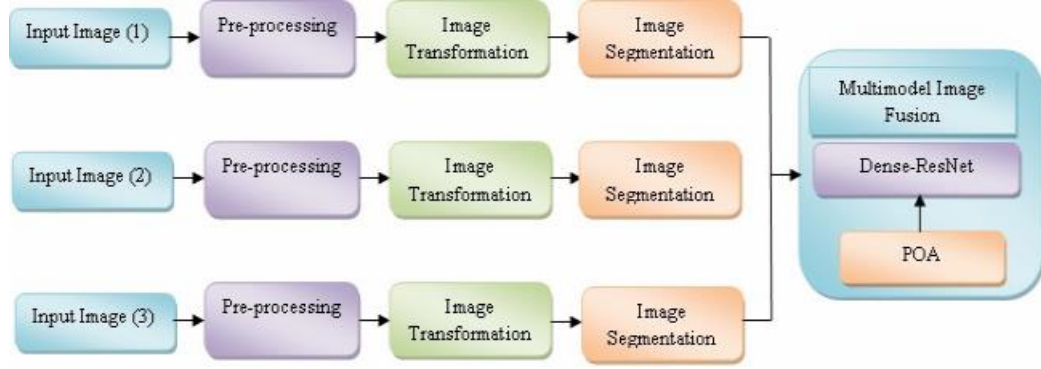


Figure. 4.8 Block diagram of the suggested model

The Dense-ResNet is trained by the POA (Trojovský and Dehghani [20]) to improve the efficiency of Dense-ResNet. The POA is the natural hunting strategy of pelicans to perform optimization in a population-based framework. It uses mechanisms inspired by pelican behaviors for both local and global search, ensuring a balance between exploration and exploitation. POA has potential applications in a wide range of engineering fields, offering a novel and efficient approach to solving complex optimization problems. Pelicans exhibit unique hunting behaviors, such as cooperative hunting and diving to catch fish. The suggested algorithm models these behaviors to effectively explore and make use of the search space. The steps of POA are explained below:

- a. **Initialization:** The POA is a population-based optimization algorithm where the pelicans represent candidate solutions to the optimization problem. The position of each pelican in the search space corresponds to a set of values for the optimization problem's variables. The initial positions of the pelicans are randomly generated within the specified bottom and top bounds of the problem's variables. This ensures that the initial candidate solutions cover a broad area of the search space. The standard formula for initializing positions is given below:

$$y_{j,k} = m_k + r.(up_k - low_k), j = 1, 2, 3, \dots \dots n. \quad 4.29$$

where the j^{th} candidate solution in k^{th} variable value is given as $y_{j,k}$. Here, 'P' is the entire population, n is the entire problem variables, random variable is denoted as r in [0, 1] interval, k^{th} upper bound is denoted as up_k , and k^{th} lower bound is indicated as low_k . The population matrix is given below:

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_j \\ \vdots \\ Z_P \end{bmatrix} = \begin{bmatrix} Z_{1,1} & \dots Z_{1,k} & \dots & Z_{1,n} \\ \vdots & \vdots & \vdots & \vdots \\ Z_{j,1} & \dots Z_{j,k} & \dots & Z_{j,n} \\ \vdots & \vdots & \vdots & \vdots \\ Z_{P,1} & \dots Z_{P,k} & \dots & Z_{P,n} \end{bmatrix}_{P \times n} \quad 4.30$$

Here, row denotes the candidate solution and columns represent the problem variables proposed values. The population matrix is indicated as z and j^{th} pelican is denoted as z_j .

The objective function, is given as:

$$O = \begin{bmatrix} O_1 \\ \vdots \\ O_j \\ \vdots \\ O_P \end{bmatrix} = \begin{bmatrix} O(Z_1) \\ \vdots \\ O(Z_j) \\ \vdots \\ O(Z_P) \end{bmatrix} \quad 4.31$$

where, O is objective function vector, and the j^{th} candidate solution objective function value is given as O_j . The hunting process is done in the exploration and exploitation phases.

- b. **Exploration phase:** in the exploration phase of the POA, pelicans move towards randomly generated prey locations within the search space. This phase is crucial for enhancing the exploration capabilities of the algorithm, ensuring a comprehensive search of the problem-solving space. The movement towards prey, combined with random perturbations, allows the pelicans to scan different areas effectively, avoiding local optima and increasing the likelihood of finding the global optimum. The formula is given as:

$$Z_{j,k}^{Q_1} = \begin{cases} Z_{j,k} + r \cdot (q_k - J \cdot Z_{j,k}), & O_q < Q_j \\ Z_{j,k} + r \cdot (Z_{j,k} - q_k), & \text{else} \end{cases} \quad 4.32$$

where j indicates the random number, k^{th} dimension prey is denoted as q_k , and the updated status of k^{th} dimension j^{th} pelican is given as $Z_{j,k}^{Q_1}$.

A new position for a pelican is accepted if it results in a better objective function value compared to its current position. This ensures that each update leads to an improvement or maintains the current best solution. It is given as:

$$Z_j = \begin{cases} Z_j^{Q_1}, & O_j^{Q_1} < O_j; \\ Z_j, & \text{else} \end{cases} \quad 4.33$$

where j^{th} pelican updated status is denoted as Z^{Q_1} and the exploration phase value of the objective function is given as O^{Q_1} .

- c. **Exploitation phase:** in the local exploitation phase of the POA, pelicans refine their positions by conducting a local search around promising areas identified in the exploration phase. This phase mimics the behavior of pelicans spreading their wings to move fish upwards and collecting them, representing a focused search for good solutions to improve them further. By evaluating the objective function at neighboring points and accepting improvements, the algorithm enhances its exploitation ability, converging to better solutions within the search space. This two-phase approach of exploration followed by exploitation ensures a balance between global search and local refinement, leading to effective optimization and the formula for the same is:

$$Z_{j,k}^{Q_2} = Z_{j,k} + E \left(1 - \frac{iter_{counter}}{iter_{max}} \right) \cdot (Z_r - 1) \cdot Z_{j,k} \quad 4.34$$

Where $Z_{j,k}^{Q_2}$ denotes the exploitation phase k^{th} dimension j^{th} pelican's updated status, the constant is given as E and the value for this is 0.2, iter counter exhibits the iteration counter and iter max is the maximum iteration. The effective updating is based on the rejection and acceptance of the new position of the pelican, which is given in the below equation,

$$Z_j = \begin{cases} Z_j^{Q_2}, & O_j^{Q_2} < O_j; \\ Z_j, & else \end{cases} \quad 4.35$$

where the exploitation phase value of the objective function is given as: $O_j^{Q_2}$.

The POA iteratively updates candidate solutions through exploration and exploitation phases, evaluates their objective function values, and tracks the best solution found so far. This process continues until the stopping conditions are satisfied. The algorithm for the POA is given below.

Algorithm 4.1:

Input of the problem

Evaluation of maximum iterations itermax and population P

Evaluation of objective function

For itercounter = 1: itermax

Random generation of prey

For J = 1: P

Phase 1: Exploration

For k = 1: n

Determination of updated status using equation (4.32)

End

Update population member by equation (4.33)

Phase 2: Exploitation
For $k = 1: n$
Determination of updated status using equation (4.34)
End
Update population member by equation (4.35)
Update optimal candidate result
End
End

4.3.2 Results and Discussion

Our proposed hybrid model in section 4.2.1 defeats several existing image fusion models. Building on this, we introduced optimized hybrid model by incorporating the Pelican Optimization Algorithm (POA). By applying this optimization during training, the model's efficiency improved considerably, resulting in higher PSNR and lower MSE and RMSE values compared to the hybrid model presented in section 4.2. Table 4.4 presents a comparison between hybrid model-1 (Dense-ResNet) and hybrid model-2 (Dense-ResNet+ POA). While hybrid model-2 consistently outperformed hybrid model-1 across all K-values, Table 4.4 highlights the results for K-value 9 across different epochs. Hybrid model-2 outperformed hybrid model-1, which itself surpassed several state-of-the-art image fusion models. This demonstrates the leading performance of hybrid model-2.

Table 4.4 Comparison of hybrid model-1 and hybrid model-2 for the K-value 9

Model	Epoch value-20	Epoch value-40	Epoch value-60	Epoch value-80	Epoch value-100
MSE Value					
Hybrid model-1	0.887	0.865	0.574	0.147	0.058
Hybrid model-2	0.851	0.830	0.855	0.141	0.056
PSNR Value					
Hybrid model-1	41.199	43.058	45.085	47.099	50.395
Hybrid model-2	43.809	43.852	45.634	48.294	51.547
RMSE Value					
Hybrid model-1	0.942	0.930	0.758	0.384	0.240
Hybrid model-2	0.923	0.911	0.743	0.376	0.236

4.4 A pair of image fusion rule that can denoise the fused image effectively

In this research, we suggested two sets of image fusion rules. The first fusion rule is designed for fusing the Approximate coefficients. The second fusion Rule is designed for fusing the detail coefficients obtained from the DWT of two multimodal brain MRI images sourced from the BraTs-2015 [76], BraTs-2018 [77], and Harvard Medical School Brain Datasets [78]. The flow layout illustrating the suggested method is presented in Figure 4.9.

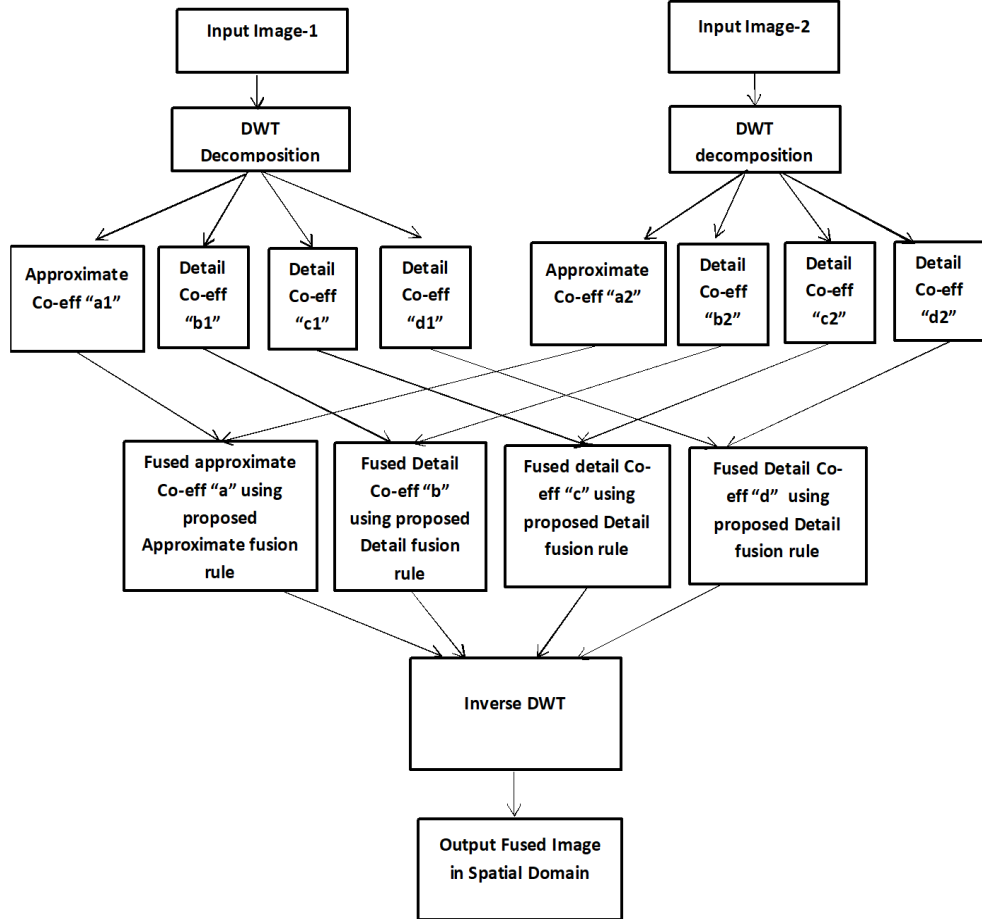


Figure.4.9 Flow layout of Proposed image fusion model

4.4.1 Methodology

The steps utilized in the suggested model are represented here.

Step 1: Two multimodal medical images from different datasets are first decomposed into approximate coefficients and detail coefficients using the DWT. DWT is selected for its effectiveness in enhancing image sparsity by minimizing noise [39]. For two input images, I_1 and I_2 , the DWT decomposition results in four image components for each image: $[a_1, b_1, c_1, d_1]$ and $[a_2, b_2, c_2, d_2]$ where a_1 and a_2 represent the approximate coefficients, while the others are the detail coefficients.

Step 2: Two distinct sets of fusion rules are proposed and applied. The fusion of the approximate coefficients is governed by equation 4.29, while the fusion of the detail coefficients is determined by equations 4.30, 4.31, and 4.32, respectively.

Step 3: After fusing the two images, the IDWT is applied for the recovery of the fused image.

$$a_3(p, q) = \ln \left(\cosh(a_1(p, q)) + \cosh(a_2(p, q)) \right) \quad (4.29)$$

$$b_3(p, q) = \max \left(\tanh(b_1(p, q)), \tanh(b_2(p, q)) \right) \quad (4.30)$$

$$c_3(p, q) = \max \left(\tanh(c_1(p, q)), \tanh(c_2(p, q)) \right) \quad (4.31)$$

$$d_3(p, q) = \max \left(\tanh(d_1(p, q)), \tanh(d_2(p, q)) \right) \quad (4.32)$$

Where, $\cosh(I(p, q)) = \frac{e^I + e^{-I}}{2}$ and $\tanh(I(p, q)) = \frac{e^I - e^{-I}}{e^I + e^{-I}}$ for the image I.

4.4.2 Results and Discussion

In the literature, the commonly utilized image fusion rules in the frequency domain include Mean (Average), Max, and Min [85]. However, the Mean scheme often leads to contrast reduction, while the Max rule exhibits high sensitivity to noise [85]. Conversely, the Min rule tends to yield poor noise reduction. In our study focusing on denoising the fused image, we specifically compared the Mean and Max rules. Table 4.4 shows the performance parameters derived from applying the Proposed Approximate Fusion rule for fusing approximate coefficients, alongside the conventional Max rule for fusing Detail coefficients. The details in table 4.5 exhibits that the suggested approximate fusion rule produces better denoising outcomes. Figures 4.10, 4.11 and 4.12 illustrate the fused images generated from the BraTs-15 [76], BraTs-18 [77], and Harvard Medical School Brain datasets [78], respectively.

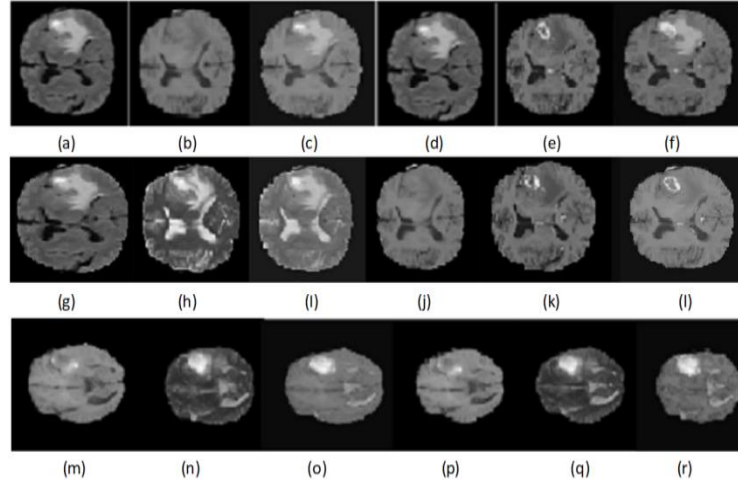


Figure. 4.10 Fused images of BraTs-2015 data set by proposed Approximate fusion rule and conventional max fusion rule. (a= *Flair*, b= *T1*, c=fused image of *Flair* and *T1*, d= *Flair* r, e= *T1C*, f= fused image of *Flair* and *T1C*, g= *Flair*, h= *T2*, I=fused image of *Flair* and *T2*, j= *T1*, k= *T1C*, l=fused image of *T1* and *T1C*, m= *T1*, n= *T2*, o=fused image of *T1* and *T2*, p= *T1C*, q= *T2*, r=fused image of *T1C* and *T1*)

Additionally, we applied the proposed approximate fusion rule and detail fusion rule to fuse two multimodal images, comparing these results with the fusion obtained using the conventional Mean rule and proposed detail fusion rule, as presented in table 4.6. The resulting fused image exhibits higher PSNR coupled with lower RMSE values. The data presented in table 4.5 and table 4.6 demonstrate that the suggested approximate and detail fusion rules yield better results compared to the conventional fusion rules.

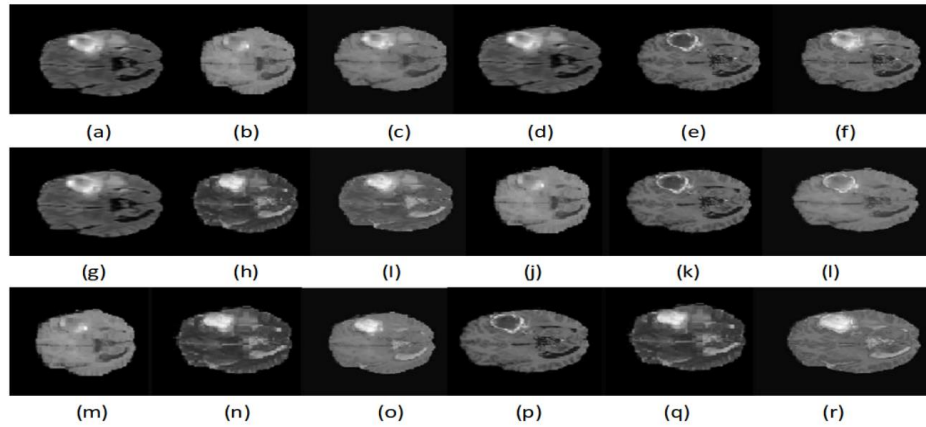


Figure 4.11 Fused images of BraTs-2018 data set by proposed Approximate fusion rule and conventional max fusion rule. (a= *Flair*, b= *T1*, c=fused image of *Flair* and *T1*, d= *Flair* r, e= *T1C*, f= fused image of *Flair* and *T1C*, g= *Flair*, h= *T2*, I=fused image of *Flair* and *T2*, j= *T1*, k= *T1C*, l=fused image of *T1* and *T1C*, m= *T1*, n= *T2*, o=fused image of *T1* and *T2*, p= *T1C*, q= *T2*, r=fused image of *T1C* and *T1*)

Table 4.5 Comparison of Proposed Approx.- Conv. Mean and Conv. mean-max Fusion Rule [85]

BraTs-2015 Dataset [76]										
	Proposed Approximation and Fusion rule					Proposed Detail Conventional Mean and Proposed Detail fusion rule				
Name of source image	Performance Parameters					Performance Parameters				
	PSNR	RMSE	MI	FMI	Entropy	PSNR	RMSE	MI	FMI	Entropy
<i>Flair</i> , <i>T1</i>	19.149	28.3	1.312	0.902	1.19	27.51	10.75	1.258	0.904	1.07
<i>Flair</i> , <i>T1C</i>	19.409	27.3	1.279	0.878	1.187	27.44	10.83	1.254	0.882	1.069
<i>Flair</i> , <i>T2</i>	18.93	28.84	1.306	0.877	1.185	24.61	15.0	1.2799	0.876	1.067
<i>T1, T1C</i>	18.77	29.51	1.321	0.904	1.195	30.2	7.91	1.2897	0.899	1.072
<i>T1, T2</i>	18.91	28.96	1.282	0.89	1.186	23.7	16.66	1.2356	0.909	1.07
<i>T1C, T2</i>	19.112	28.25	1.27	0.876	1.184	23.48	17.08	1.229	0.876	1.073
Average Value	19.046	28.526	1.295	0.888	1.187	26.156	13.04	1.259	0.891	1.047
BraTs-2018 Dataset [77]										
<i>Flair</i> , <i>T1</i>	20.575	24.06	1.383	0.93	1.148	28.90	9.168	1.323	0.93	0.985
<i>Flair</i> , <i>T1C</i>	20.8	23.26	1.353	0.913	1.15	28.66	9.42	1.312	0.915	0.984
<i>Flair</i> , <i>T2</i>	20.6	23.78	1.367	0.912	1.154	27.47	10.79	1.336	0.908	0.985
<i>T1, T1C</i>	20.13	25.26	1.387	0.927	1.146	32.21	6.27	1.344	0.923	0.983
<i>T1, T2</i>	20.42	24.42	1.356	0.921	1.142	25.63	13.33	1.293	0.927	0.982
<i>T1C, T2</i>	20.58	23.85	1.34	0.908	1.15	25.17	14.05	1.285	0.91	0.982
Average value	20.517	23.96	1.364	0.918	1.153	28.00	10.5	1.315	0.919	0.983
Images from Harvard Medical School Brain dataset [78]										
Image-1	16.197	39.51	1.24	0.8728	0.0023	16.425	38.48	1.2257	0.8738	0.001
Image-2	17.35	34.61	1.262	0.855.	1.5196	22.41	19.33	1.25	0.8507	3.861
Image-3	18.135	34.58	1.253	0.879	1.5698	17.66	33.37	1.225	0.8755	1.238
Image-4	15.94	40.73	1.27	0.874	1.62	18.99	28.64	1.2415	0.876	1.2415
Image-5	16.08	41.36	1.23	0.817	0.0493	17.76	32.97	1.196	0.8249	0.025
Image-6	13.89	53.89	1.307	0.88	1.7017	13.98	50.97	1.233	0.88	1.257
Image-7	13.178	57.13	1.148	0.819	0.332	16.32	38.97	1.1324	0.8199	0.1947
Image-8	16.57	38.71	1.228	0.825	2.67	16.99	36.05	1.20	0.8239	1.9126
Average Value	15.92	42.56	1.656	0.837	0.862	17.567	34.84	1.212	0.857	1.217

Table 4.6 Comparison of proposed Approx.-Proposed Detail and conv. mean-max [85] Fusion Rule

	BraTs-2015 Dataset [76]									
	Proposed Approximation and Detail Fusion rule					Conventional mean-max Fusion rule [85]				
Name of source image	Performance Parameters					Performance Parameters				
	PSNR	RMSE	MI	FMI	Entropy	PSNR	RMSE	MI	FMI	Entropy
<i>Flair</i> , <i>T1</i>	25.91	13.25	1.32	0.928	1.249	18.55	30.192	1.282	0.925	1.092
<i>Flair</i> , <i>T1C</i>	25.324	13.838	1.293	0.901	1.253	18.488	30.381	1.273	0.899	1.0918
<i>Flair</i> , <i>T2</i>	23.412	18.668	1.315	0.898	1.234	18.046	32.049	2.292	0.891	1.082
<i>T1</i> , <i>T1C</i>	29.416	8.885	1.314	0.924	1.258	17.862	32.798	1.3025	0.925	1.099
<i>T1</i> , <i>T2</i>	21.13	22.408	1.285	0.915	1.232	17.192	35.674	1.214	0.9122	1.083
<i>T1C</i> , <i>T2</i>	21.037	22.877	1.278	0.893	1.238	17.665	33.542	1.231	0.889	1.087
Average Value	24.37	16.66	1.32	0.91	1.244	17.95	32.39	1.432	0.906	1.082
BraTs-2018 Dataset [77]										
<i>Flair</i> , <i>T1</i>	27.18	11.741	1.379	0.948	1.131	19.99	25.849	1.337	0.946	0.993
<i>Flair</i> , <i>T1C</i>	26.075	12.684	1.355	0.932	1.1384	19.841	26.063	1.328	0.931	0.9934
<i>Flair</i> , <i>T2</i>	25.778	13.961	1.369	0.924	1.134	19.927	26.588	1.347	0.921	0.998
<i>T1</i> , <i>T1C</i>	30.79	7.688	1.375	0.944	1.153	19.085	28.54	1.352	0.942	0.999
<i>T1</i> , <i>T2</i>	22.923	18.314	1.354	0.935	1.133	18.55	30.61	1.289	0.939	0.0068
<i>T1C</i> , <i>T2</i>	22.51	19.279	1.341	0.9199	1.141	19.067	28.631	1.298	0.922	0.0016
Average Value	25.876	13.54	1.352	0.931	1.141	19.134	27.52	1.28	0.932	0.665
Images from Harvard Medical School Brain dataset [78]										
Image-1	15.523	48.68	1.252	0.887	0.0013	16.422	38.497	1.214	0.893	0.001
Image-2	21.08	24.55	1.258	0.8734	3.7844	22.39	19.364	1.246	0.862	3.693
Image-3	19.85	37.63	1.245	0.8992	1.1896	17.688	33.277	1.221	0.897	1.256
Image-4	16.4	39.18	1.284	0.8895	1.569	19.012	28.575	1.241	0.887	1.2168
Image-5	15.56	44.27	1.243	0.8405	0.028	17.75	33.03	1.1956	0.841	0.017
Image-6	22.08	52.64	1.310	0.8955	1.652	13.988	50.95	1.23	0.8897	1.295
Image-7	13.79	53.09	1.159	0.8378	0.2433	16.306	39.015	1.131	0.83	0.314
Image-8	17.47	43.03	1.237	0.8441	2.553	16.98	36.074	1.198	0.834	1.802
Average Value	19.21	42.88	1.242	0.863	1.32	17.462	24.36	1.203	0.87	1.198

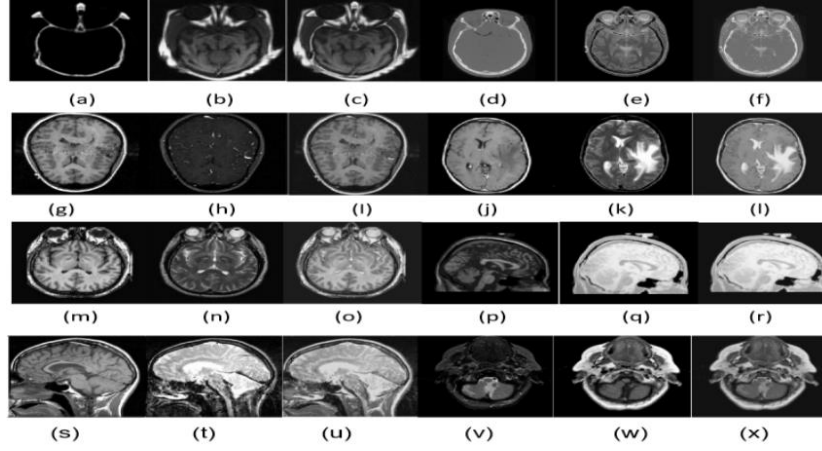


Figure 4.12 Fused Images obtained from Harvard Medical School Brain data set [78] ({a, b source images and c is fused image of a and b}, {d, e source images and f is fused image of d and e}, {g, h source images and i is fused image of g and h}, {j, k source images and l is fused image of j and k}, {m, m source images and o is fused image of m and n}, {p, q source images and r is fused image of p and q}, {s, t source images and u is fused image of s and t}, {v, w source images and x is fused image of v and w})

4.5 Conclusion

We have developed a hybrid Dense-ResNet model (hybrid model-1) to efficiently fuse medical images from three different modalities ($T1$, $T1GD$ and $T2$) of MRI images. In the preprocessing stage, we apply median filtering [15] for noise reduction and utilize ET-Net to preserve edge information. Our model evaluated against several contemporary image fusion models and exhibited superior performance due to its enhanced feature extraction capabilities. With the extension of this research work, we developed the optimized hybrid model. The POA based hybrid model outperformed the hybrid model -1 which itself out performed several contemporary image fusion models. This claims the better performance of the optimized hybrid model at the cost of computational efficiency.

Additionally, we introduced different fusion rules for integrating the approximate and detail coefficients of medical images from two modalities. Performance comparisons with conventional fusion rules reveal that both proposed rules outperform the traditional methods, achieving higher PSNR and lower RMSE values, underscoring the effectiveness of our approach.

CHAPTER 5

DEVELOPMENT OF EFFICIENT IMAGE CLASSIFIER

5.1 Introduction

Image classifiers utilize various algorithms and methodologies, with convolution neural networks (CNNs) being particularly effective in detecting patterns within images. During the image classification process, the model learns to recognize principal attributes of the images, like edges, colors, shapes, textures etc. which assist to distinguish one category from another.

Training an image classifier usually involves providing it with a sufficient number of datasets of labeled images to optimize the hyper parameters. Constructively classifying the medical images is an essential function in support of clinical care and treatment. Deep learning algorithms often need huge amounts of labelled or inferred feedback data. On the other hand, deep learning models provide accurate and consistent results in comparison with the traditional methods. But there are challenges which includes manual annotation of medical images which is very time-consuming process and also there is a chance of oversight. Investigators are also exploring contrastive learning and transfer models for identifying diseases. In medical image classification, Fine-tuning of the pre-trained models is important in achieving better performance. Studies shows that Data augmentation also benefitted in reduction of over fitting.

Detecting, extracting and classifying several types of tumors (or several types of cancer) from MRI images is much needed in the field of medical exploration. Even though many classifiers are existing in the literature with promising results, many of them are suffering from computational complexity. To address this problem, this study dealt with how to improve the performance of the classifier with smaller datasets and with less computational load.

- An efficient method for brain tumor detection and classification by Extended anisotropic diffusion filtering.
- An efficient classifier for breast cancer classification by WMRESNET.

5.2 An efficient method for brain tumor detection and classification by Extended anisotropic diffusion filtering

This research is based on the locating and classification of different types of brain tumors from the MRI images. Image categorization is an important area of research that is receiving increasing attention from the medical imaging community. Even though many classifiers are existing in the literature with promising results, many of them are suffering from the computational complexity. To overcome the problem of computational complexity and the requirement of large number of datasets, a simple three-step process that consists of pre-processing, detection and classification is proposed. The suggested pre-processing stage includes Dark Channel Prior (Haze removal) Extended Anisotropic Diffusion Filtering (EADF) [87] followed by principal image generation, and Edge Enhancement. The accurate area of the tumor can be detected by the suggested auto threshold-based detection technique. By the utilization of preprocessing and all the stages up to the auto thresholding, the proposed model is successful to detect the tumor but it can't classify different types of tumors. So, for the classification purpose, Resnet-50 [91] has been utilized. Thorough investigations are conducted on (BD35H, BMI-1, BTI, BTS, BD_BT dataset) [106] and BraTs-2018 dataset [107]. With very less computational complexity as well as computational cost, the suggested model exhibits satisfactory performance in detection and classification. The flow layout of the proposed model is shown in figure 5.1.

5.2.1 Methodology

Different stages of the proposed methodology are described below:

Stage 1: Pre-processing

Biomedical instruments cannot always give satisfactory images. Often these images are vague and blurry. Researchers and medical practitioners face a lot of problems due to these artefacts. To rectify the quality of the image, different enhancement stages are used and are explained below.

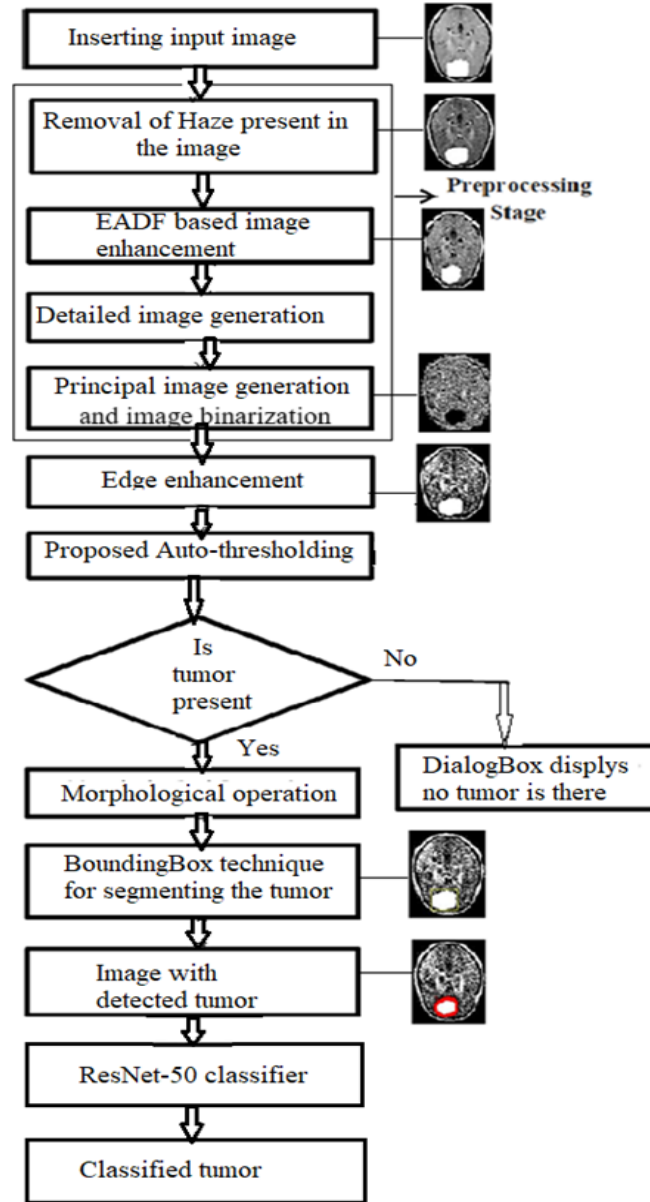


Figure 5.1 Flow layout of Proposed Model

a: Haze Removal

According to [171], any haze image $I(i, j)$ could be represented as

$$I(i, j) = I''(i, j) + \chi(i, j) + \mu(1 - \chi(i, j)) \quad (5.1)$$

Where, $I''(i, j)$ is the haze-free image. $\chi(i, j)$ represents the transmission map, μ represents atmospheric light and (p, q) is the pixel position. Restoration of $I''(i, j)$ is the main objective of haze removal technique. In this paper, the Dark Channel Prior (DCP) is used to remove the haze present in the input image.

$$I^d(i, j) = \min_{(i, j) \in \emptyset(i, j)} (\min_{c \in \{R, G, B\}} (I^c(i, j))) \quad (5.2)$$

Where, I^d represents the DCP of hazy image I and I^c is a color channel of I . $\emptyset(i, j)$ Represents a local patch which is centered at (p, q) . The DCP method encompasses two variations: Simple DCP and Approximate DCP. In the Simple DCP approach, haze estimation is performed on per-pixel basis, while quad tree decomposition is employed to determine the ambient light. In contrast, the Approximate DCP utilizes both per-pixel and spatial blocks for haze estimation but forgoes quad tree decomposition. In this method, the dark channel at coordinates (p, q) is approximated to zero, hence the name. Table 5.11 presents the variation in detection efficiency, PSNR, and RMSE values across different haze removal techniques. The haze removal process enriches the intensity and contrast levels of the input image. Table 5.1 represents the detection efficiency with and without using haze removal technique. Table 5.12 represents the variation of detection efficiency according to the variation of haze removal percentage in different datasets.

Table 5.1 Performance of EADF with and without haze removal technique

Input Image	EADF stage with only Haze removal stage				EADF stage without Haze removal stage	
	Approximate DCP With accuracy 80%		Simple DCP With accuracy 76%			
	PSNR (DB)	RMSE	PSNR (DB)	RMSE	PSNR (DB)	RMSE
Sample Image-1	25.44	1.2	24.63	1.3	Can't detect tumor	
Sample Image-2	26.54	1.1	25.78	1.26	Can't detect tumor	
Sample Image-3	24.62	1.3	23.62	1.5	12.72	3.64
Sample Image-4	23.69	1.4	22.98	1.6	14.01	3.12
Image-5	24.78	1.1	23.84	1.56	14.759	3.11
Average Value	25.014	1.25	24.17	1.28	14.07	3.33

Table 5.2 Required % of Haze removal for detection accuracy

Name of the Dataset						
	BR35H [106]	BMI-I [106]	BTI [106]	BTS [106]	BD_BT [106]	BraTs 2018 [107]
% Haze removal	40%	30%	30%	50%	40%	1%
Achieved Accuracy	93%	93%	96%	90%	100%	97%

b: Extended Anisotropic Diffusion Filtering (EADF)

Perona and Malik are the introducer of ADF [86]. ADF process is applied to minimize noise in the image. This filtering requires no prior information about regions or boundaries; instead, diffusion occurs selectively based on the conduction term generated from local computations [86]. Anisotropic diffusion computes the diffusion in every spatial location and then adds all to get the group diffusion. The mathematical representation of this filtering is described below: We all know that; diffusion is a physical process that brings the equilibrium of concentration difference without disturbing the mass. Fick's law represents the mathematical formula of diffusion as:

$$J = -D \cdot \nabla U \quad (5.3)$$

The above equation describes that the concentration gradient ∇U generates a flux ' J ' to reimburse this gradient. Here ' D ' represents the diffusion function which defines the relation between J and ∇U . The function D is (+ve) where J and ∇U are parallel and is called Isotropic diffusion. D is (-ve) when J and ∇U are not parallel and is called Anisotropic diffusion [96]. The continuity equation tells that the diffusion can only transport mass but cannot destroy or create any mass, and is represented by equation 5.4 as:

$$\frac{\partial}{\partial t} U = -\nabla J \quad (5.4)$$

Where, t signifies the time. Anisotropic diffusion is achieved by merging the continuity equation with Fick's law and can be written as:

$$\frac{\partial}{\partial t} U = -\nabla(-D \cdot \nabla U) \quad (5.5)$$

After applying the Anisotropic diffusion in an image, equation 5.5 becomes equation 5.6.

$$\frac{\partial}{\partial t} I(x, t) = \nabla \cdot (D(x, t) \cdot \nabla I(x, t)) \quad (5.6)$$

Where $I(x, 0)$ is the initial unprocessed image, x is the image coordinate and iteration step is t in second. $D(x, t)$ represents the diffusion function for image ' I ' at time ' t '. For the localization of edge, the gradient of the image is first calculated by (equation 5.7).

$$E(x, t) = \nabla(I(x, t)) \quad (5.7)$$

The gradient magnitude represents the conduction coefficient of diffusion and is given in (equation 5.8).

$$D(x, t) = f(|\nabla I(x, t)|) \quad (5.8)$$

The edge preservation and sharpness depend on the selection of function f . As ' D '

monotonically decreases, according to Perona and Malik (founder of Anisotropic diffusion filtering) the diffusion function can be represented by equations 5.9 and 5.10.

$$D(x, t) = \exp(-(\frac{\|\nabla(x, t)\|}{N})^2) \quad (5.9)$$

$$D(x, t) = \frac{1}{1+(\frac{\|\nabla(x, t)\|}{N})^2} \quad (5.10)$$

Equation 5.9 enhances high-contrast edges but equation 5.10 enhances wide regions over small regions. N denotes the diffusion constant, and it is kept as 45 in our model. The differential relation after discretization can be represented as:

$$I_{i,j}^{t+1} = I_{i,j}^t + \alpha(N_c \cdot \nabla_N I + S_c \cdot \nabla_S I + W_c \cdot \nabla_W I + E_c \cdot \nabla_E I) \quad (5.11)$$

$$I_{i,j}^{t+1} = I_{i,j}^t + \alpha(N_c \cdot \nabla_N I + S_c \cdot \nabla_S I + W_c \cdot \nabla_W I + E_c \cdot \nabla_E I + NE_c \cdot \nabla_{NE} I + NW_c \cdot \nabla_{NW} I + SE_c \cdot \nabla_{SE} I + SW_c \cdot \nabla_{SW} I) \quad (5.12)$$

Equation 5.11 is for the Conventional Anisotropic diffusion filtering where conduction of diffusion is done in four directions Nc, Sc, Wc, Ec, and represents north, south, west and east directions respectively. The equation 5.12 represents EADF [87] with eight directions, where Nc, Sc, Wc, Ec, NEc, SEc, NWc and SWc are the directions of conduction of diffusion and represents north, south, west, east, north-east, south-east, north-west and south-west directions respectively and α represents the step size and is equal to 0.125s to take care of eight direction diffusion.

The performance parameters of Conventional ADF and EADF on multiple datasets is reported in Table 5.3.

Table 5.3 Results after preprocessing with conventional ADF and EADF

Datasets [116]	Conventional ADF Accuracy 91%		EADF Accuracy 93%	
	PSNR (DB)	RMSE	PSNR(D B)	RMSE
BR35H	29.97	0.892	32.69	0.812
BMI-I	26.05	0.965	32.2	0.823
BTI	26.052	0.970	26.81	0.915
BTS	30.34	0.801	34.73	0.684
BD_BT	19.144	1.0002	27.85	0.953
BraTs 2018[107]	33.4	0.753	38.59	0.456

c: Detailed Image Generation

After EADF, the preprocessed image is decomposed using DWT into detail and approximate coefficients. For subsequent processing, the detail coefficients are selected due to their lower noise content.

d: Principal Image Generation

The detailed image is processed through proposed Entropy based intensity enhancement stage and the principal image $I'(p, q)$ is produced by the equation 5.13. The further enhancement of the image detail to differentiate the target region from the detailed image is done efficiently in this stage.

$$I'(i, j) = \frac{I(i, j)}{\text{entropy}(I(i, j))} \quad (5.13)$$

Suggested Entropy-based image enhancement has several advantages that make it a powerful technique for improving image quality. It corrects image contrast by modifying the values of the pixels. This makes the features in the image more distinguishable, particularly for images with poor lighting or low contrast. This method can enhance fine details without introducing significant artefacts.

e: Edge Enhancement

The stages of anisotropic filtering, detailed image generation, and principal image generation primarily serve the purpose of denoising. However, the denoising process tends to smooth image boundaries and reduce image contrast. To solve this issue, an edge enhancement of the principal becomes necessary. Here, an omnidirectional Kernel k , represented in equation 5.14 is convoluted with the principal image to preserve edge information in all directions.

$$K = C * \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (5.14)$$

(where c is a constant that is set to be 0.025 during experimentation)

Stage 2: Detection and classification Stage

This stage consists of:

- (1) Auto thresholding, morphological operation and bounding box representation to identify the location of the tumor. The auto thresholding can detect the tumor successfully but can't classify the tumor.
- (2) ResNet-50 [102] is used to classify the detected tumor.

a: Proposed Intelligent Thresholding and bounding box generations

Threshold based segmentation is one of the fastest methods among other existing segmentation techniques and it is widely used. Perhaps, selection of appropriate threshold value is the main challenge in the process of segmentation. To overcome the process of choosing the appropriate threshold value manually, auto thresholding method is proposed and is given in equation 5.15. The value of threshold is adaptive in nature as it is varying in accordance with the variance of the image along with minimum and maximum intensity of the target image.

$$th = 20 * \ln(\sigma) + \frac{\max(I'(i,j)) + \min(I'(i,j))}{2} \quad (5.15)$$

Here, σ is representing the standard deviation and has a very good impact on overall detection efficiency. Table 5.4 presents the identified threshold value for different data set by the proposed auto thresholding method with different σ values. As the color contrast is high in pseudo colour images, the threshold value is considered as twice the standard deviation. It has been seen that the detection of tumor from pseudo color images is more difficult than the normal color images. But the proposed model can detect tumor even from pseudo colour images in the BD_BT [106] dataset.

Bounding box and morphological operation are utilized to highlight the absolute accurate location of tumor. Image erosion together with solidity function is utilized to map the tumor cell based on the estimation of area and density function. The subtraction of the eroded image from the principal image produced the bounding box around the tumor.

b: ResNet-50 for Tumor Classification:

Pre-trained ResNet-50 [41] classifier is used to classify the tumor category belongs to Glioma, Meningioma or Pituitary tumor. For the assessment, six different data sets namely (BR35H, BMI-1, BTI, BTS, BD_BT) [106] and BraTs-2018 [107] are used. Detail description of these data sets are given in the ‘Data Description’ section.

5.2.2 Results and discussion

In this study, we examined various MRI datasets to identify tumor regions within the images. We began by evaluating our model using the BR35H dataset. The initial phase of our suggested model includes a critical preprocessing step, particularly focusing on haze removal techniques. This pre-processing stage has a very important role in this framework. The variations of different performance parameter values, without haze removal and with haze removal followed by EADF are given in table 5.12 to highlight the importance of haze removal in the pre-processing stage. In the process of haze removal, we have also done experimentation on both Approx. DCP and Simple DCP. The Haze removal has a very good impact on the detection result. We varied the

amount of the percentage haze removal to achieve a betterment in the results which is reported in table 5.2. Even though higher PSNR was obtained after applying Approx. DCP, exact segmentation of tumor region was achieved by Simple DCP method and is highlighted in table 5.7. The output of proposed detection model with the absence of tumor is exhibited in figure 5.2. The suggested model is successful in detecting very small tumor with detection efficiency of 93% on BR35H [106] dataset. The qualitative and quantitative assessment of the suggested model for the dataset BR35H [106] is represented in table 5.8 and table 5.9 respectively.

Table 5.4 Obtained threshold values (th) using the proposed auto thresholding from different datasets

Dataset [106]	BR35H	BMI-I	BTI	BTS	BD_BT	BraTs 2018[107]
Sample Image-1	90	70	75	110	58	150
Sample Image-2	91.2	80.99	75.21	109	149.1	149.1
Sample Image-3	90	75.2	78.3	107	145	145
Sample Image-4	85.4	76.5	79.1	127	168	168
Sample Image-5	89.6	87.4	71.4	113	187	187

Table 5.5 Evaluation of the proposed preprocessing framework with different haze removal techniques on BR35H [106] dataset

Input Image	Approximate DCP Accuracy-93%		Simple DCP Accuracy-93%ple DCP	
	PSNR (DB)	RMSE	PSNR (DB)	RMSE
Sample Image-1	33.7	0.8789	32.16	0.921
Sample Image-2	32.9	0.8991	32.44	0.912
Sample Image-3	33.34	0.8690	32.94	0.902
Sample Image-4	31.64	0.9312	30.67	1.012
Sample Image-5	34.23	0.8560	34.00	1.122
Average Value	33.162	0.8842	32.44	0.974

Table 5.6 Detection without using Detailed image and Principal image generation stage






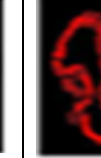





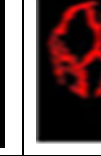
Detection without using Detailed image and Principal image generation stage						
	Input Image	Haze removed Image	Filtered Image	Tumor alone	Tumor outline	Detected Tumor
Sample Image-1						
Sample Image-2						

Table 5.7 Performance comparison of Simple DCP and Approx. DCP


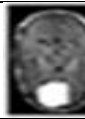
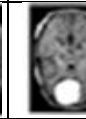
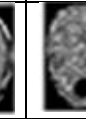
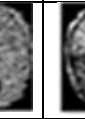
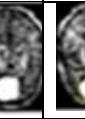
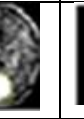




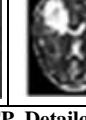
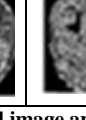
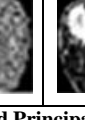
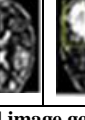



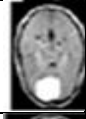
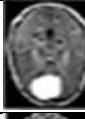
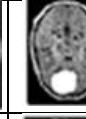
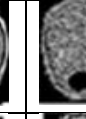
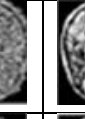
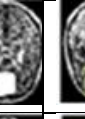




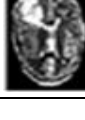
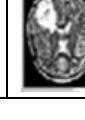
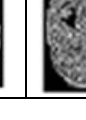
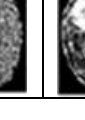












































































Detection with Simple DCP, Detailed image and Principal image generation on BR35H dataset [106]									
	Input Image	Haze removed Image	Filtered Image	Principal Image	Edge enhanced image	Bounding Box generation	Segmented tumor	Tumor boundary	Detected Tumor
Sample Image-1									
Sample Image-2									
Detection with Approx DCP, Detailed image and Principal image generation on BR35H [106] dataset									
Sample Image-1									
Sample Image-2									

Table 5.8 Qualitative assessment of the suggested model with different datasets

Tumor detection by proposed model (BMI-1 dataset [106])									
	Input Image	Haze removed Image	Filtered Image	Principal Image	Edge enhanced image	Bounding Box generation	Segmented tumor	Tumor outline	Detected Tumor
Sample Image-1									
Sample Image-2									
Tumor detection by proposed model (BTI Dataset [106])									
Sample Image-1									
Sample Image-2									
Tumor detection by proposed model (BTS Dataset [106])									
Sample Image-1									
Sample Image-2									
Tumor detection by proposed model (BD_BT Dataset [106])									
Sample Image-1									
Sample Image-2									






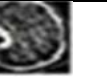


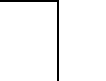




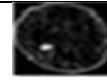
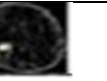

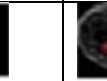

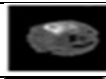
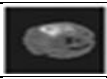















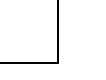
Tumor detection of pseudo color image by proposed model (BD_BT Dataset [106])									
Sample Image-1									
Sample Image-2									
Tumor detection by proposed model (BRATS 2018 Dataset[107])									
Sample Image-1									
Sample Image-2									

Table 5.9 Quantitative performance of the suggested model with different datasets

Name of Image	BMI-I [106]			BTI [106]			BTS [106]			BD_BT [106]			BraTs-2018[107]		
	PSNR	RMSE	Accuracy	PSNR	RMSE	Accuracy	PSNR	RMSE	Accuracy	PSNR	RMSE	Accuracy	PSNR	RMSE	Accuracy
Sample Image-1	32.7	0.946	93%	32.37	1	96%	36.14	0.678	80%	32.97	0.987	99%	38.178	0.523	97%
Sample Image-2	32.26	0.956		32.58	0.965		34.27	0.867		33.198	0.965		38.60	0.497	
Sample Image-3	33.14	0.946		32.12	0.976		37.57	0.567		29.08	1.0006		38.24	0.456	
Sample Image-4	31.64	0.997		32.64	0.955		34.10	0.854		28.59	1.0002		38.94	0.434	
Sample Image-5	34.23	0.897		30.34	1.002		37.97	0.534		30.03	1.0001		38.01	0.521	
Average value	32.79	0.566		32.01	1.0001		36.01	0.687		30.77	0.9998		38.594	0.486	

Table 5.10 Result of Comparison of the suggested model with the existing CNN model [106]

Data Set	Accuracy of the Suggested Model	Accuracy of CNN based Model [106]
BR35H [106]	93%	-
BMI-I [106]	93%	97.08%
BTI [106]	96%	85%
BTS [106]	90%	100%
BD_BT [106]	100%	100%
BraTs 2018 [107]	97%	-

In addition to that, the suggested model is compared with the contemporary CNN based image classification model [106] on different datasets namely BMI-I[106], BTI [106], BTS [106] and BD_BT [106] and is mentioned in table 5.11. Table 5.10 depicts that the suggested model commensurate the contemporary CNN models. The performance of the suggested detection in terms of the computational cost, is compared with the existing classification model [107] and result is presented in table 5.12.

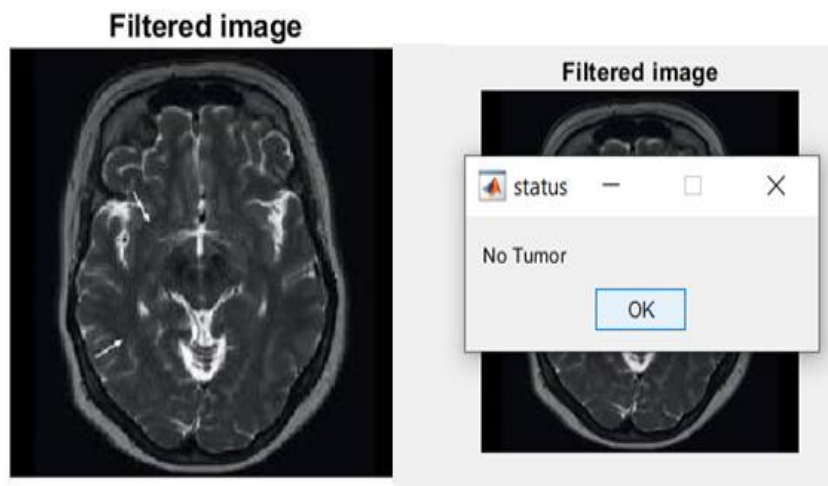


Figure 5.2 The output when no tumor is present (BR35H dataset)

Table 5.11 Time taken in seconds to detect tumor by the suggested model on different datasets

Data Set [106]	Sample Image-1	Sample Image-2	Sample Image-3	Sample Image-4	Sample Image-5	Average time (proposed model)	Average time for existing model [107]
BR35H[106]	6.42	6.36	6.43	6.50	6.49	6.44	-
BMI-I[106]	3.739	3.56	3.66	3.8	3.71	3.69	-
BTI[106]	6.16	6.097	6.199	6.086	6.21	6.1504	-
BTS[106]	8.722	8.9	8.8	8.69	8.75	8.77	-
BD_BT[106]	6.488	6.51	6.46	6.33	6.34	6.426	-
BraTs-2018 [107]	2.776	2.78	2.68	2.81	1.69	2.746	84

To assess the effectiveness of the suggested model in detecting and classifying brain tumors, we conducted a comprehensive comparison with several state-of-the-art models. The results, summarized in Table 5.12 and illustrated in Table 5.13 using dataset [108], indicate that our model consistently defeats the existing methods, achieving superior performance in both tumor detection and classification.

Table 5.12 Performance comparison of suggested model with contemporary tumor classifier models

Models	Technology used	Detection Accuracy in %
Khaliki, et.al 2024 [108]	Efficient Net B4	97
	Inception V3	95
	VGG19	96
	VGG16	98
	CNN	91
Shahajad et.al 2021 [142]	SVM	92
Vankdothu et.al	CNN	89

2022 [143]	LSTM	90.02
	CNN-LSTM	92
Srinivas et.al 2022 [144]	Inception V3	78
	VGG16	96
	ResNet50	95
suggested model	Suggested tumor detection framework	99
	Classification of Detected Tumor by ResNet-50 (without Proposed Preprocessing framework)	88
	Classification of Detected Tumor by ResNet-50 (with Proposed Preprocessing framework)	100

Our suggested model can detect tumors in medical images, but it is unable to classify different types of tumors without the ResNet-50 classifier stage. By incorporating the ResNet-50 [102] classifier at the final stage of the model, it can classify tumors with 100% accuracy. This remarkable efficiency is made possible by the suggested pre-processing stage.

For a more thorough assessment, the proposed detection framework is compared with the ADF based brain tumor detection model [86]. PSNR and SSIM vales have been used to assess the performance of the suggested framework. The authors of [101] also utilized brain MRI images and its performance is compared with the suggested model which is reported in table 5.14.

Table 5.13 Output of the suggested framework on dataset [108]






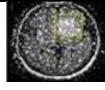












	Input Image	Haze removed Image	Filtered Image	Principal Image	Edge enhanced image	Bounding Box generation	Segmented tumor	Tumor boundary	Detected Tumor
Sample Image-1									
Sample Image-2									

Table 5.14 Comparison of the suggested model with the existing ADF based model [101]

Name of the Model	PSNR (db)	SSIM
Maurya, R., et.al 2022 [101]	17.0195	0.96
Proposed model	33.6	0,9834

Computational Load

The computational load of the proposed model is evaluated against a CNN-based model using various datasets, including BR35H, BMI, BTI, BTS, BD_BT, and BraTs-2018. These datasets, consisting of images with different modalities and resolutions, were analyzed on an Intel i5, 1.8 GHz processor using MATLAB-2024 software and the result obtained is represented in table 5.11. A CNN-based model developed by Ranjbarzadeh, Ramin et al. [107], with the computational complexity mentioned as 84 seconds for the data set BraTs-2018. In comparison, a significantly lower computational cost of 2.69 seconds for the same dataset was achieved by the proposed model. This substantial reduction in computational time highlights the efficiency of the suggested framework.

5.3 An efficient classifier for breast cancer classification by WMRESNET

This study presents an efficient breast cancer classification model, employing weight-modified ResNeT-14 (WMRESNET) to extract features more accurately and rapidly, even with a small training dataset. The suggested model incorporates SVM classifier, that classifies different types of breast cancers based on the feature extracted by WMRESNET achieving a remarkable classification efficiency of 96% on Ultrasound Breast Cancer Images (USBI) dataset [109]. Pre-processing techniques such as Histogram Equalization and CNN-based denoising enhanced the feature extraction capability of WMRESNET during training. The aim of this research is to classify cancer images as malignant, benign or non-cancerous, crucial for appropriate treatment after early detection.

5.3.1 Methodology

The proposed model introduces a Weight Modified Residual Net to classify several types of breast cancer. Feature maps generated by the WMRESNET are utilized to train SVM classifier [142], culminating in the final classification outcome. Specifically, the model comprises 40 residual blocks, each consisting of two convolution layers and one pooling layer. The stepwise working of the proposed methodology is given below:

Pre-processing Stage

To enhance feature information within the training images from dataset [109], a multi-step pre-processing approach is adopted.

Firstly, the training images undergo Histogram Equalization, a technique aimed at standardizing image brightness and contrast, thereby enhancing visual clarity and feature discernibility. Following Histogram Equalization, the images are subjected to denoising using Convolutional Neural Network (CNN).

This denoising process aims to lessen image noise and to raise image quality, thus facilitating more accurate feature extraction during subsequent stages of network training. Figure 5.3 illustrates the transformation of images depicting different breast cancers from dataset [109] following this pre-processing regimen, showcasing the efficacy of these techniques in refining image quality and preparing them for effective network training.

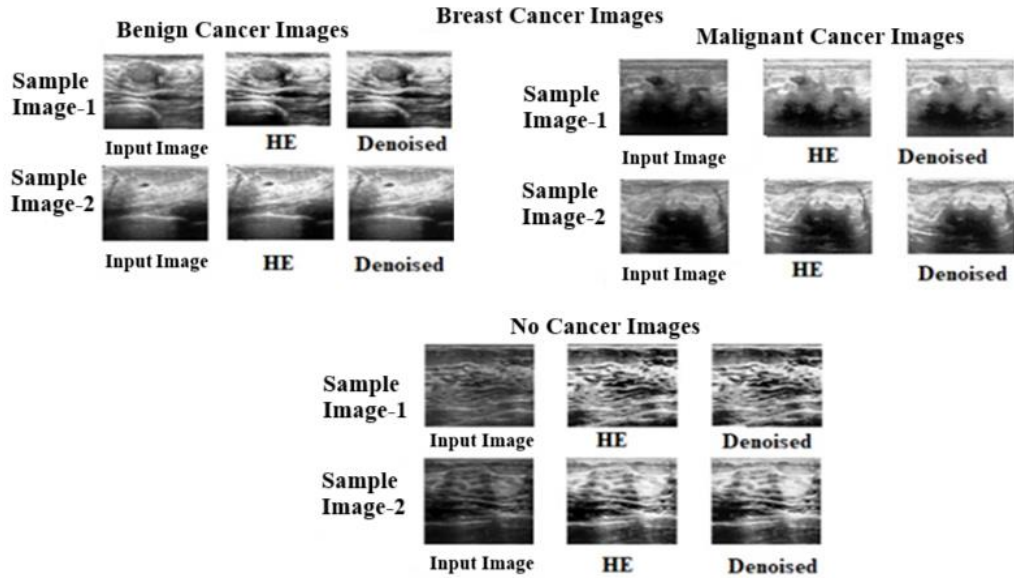


Figure 5.3 Different breast cancer images from dataset [109] after pre-processing

WMRESNET based Image Feature Extraction

At this stage, the attributes of the training images have been extracted by WMRESNET, offering several advantages over conventional ResNet architecture. One crucial aspect of neural networks is the assignment of weights, as weight changes play a significant role in determining the network's overall performance. In Residual Net (ResNet), the output of a residual block is calculated by adding the output of the current node ($F(I)$) and the input of the previous node (I), which enhances network efficiency and performance. In the conventional ResNet architecture, both (I) and $F(I)$ are assigned equal weights and is represented by equation 6.45. However, in Proposed WMRESNET, the weight of residual blocks is modified to significantly increase the learning efficiency compared to conventional ResNet and is represented by equation 5.17. The building block of a conventional ResNet and the proposed WMRESNET is illustrated in figure 5.4. In WMRESNET, 14 such residual blocks are utilized for feature extraction.

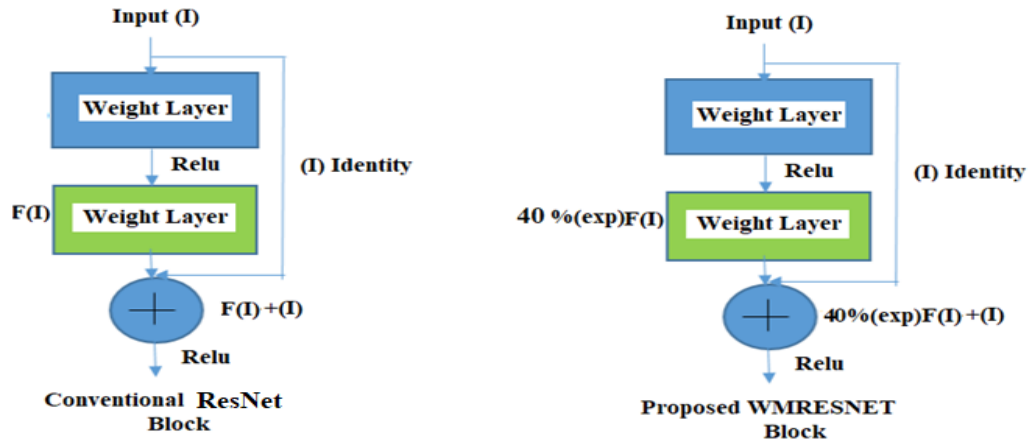


Figure 5.4 Block diagram of conventional ResNet and proposed WMRESNET

In Proposed WMRESNET, the input of the previous residual node (I) is varied exponentially, resulting in higher weights but only a fraction (40%) of the output of the current node ($F(I)$) is considered and is given in equation 6.88. This weight modification leads to a significant increase in learning efficiency, from 61% to 77%, and a corresponding rise in classification efficiency, from 85% to 96% for the (USBI) dataset [109].

In general, higher learning efficiency increases the learning rate of the network at the cost of higher weights. However, excessively high learning rates may lead to increased training losses and erratic network performance. Therefore, in WMRESNET, weight adjustments are made incrementally to optimize learning rates without causing training errors. The examination of weights in a residual node (block) in conventional ResNet versus Proposed WMRESNET is pictorially represented in figure 5.5 with weight values detailed in table 5.15. While the actual weight in any residual node cannot be precisely determined, these visualizations and tables provide insight into the controlled increase in weights within WMRESNET, ensuring optimal learning rates and network performance.

$$(I) + F(I) = \text{Final Output} \quad (5.16)$$

$$(I) + P * F(\exp(I)) = \text{Final Output} \quad (5.17)$$

Where P is a constant. The value of P is considered as 0.4 in this research. The variation of learning efficiency and classification efficiency with the variation of the value of 'P' is represented in table 5.16.

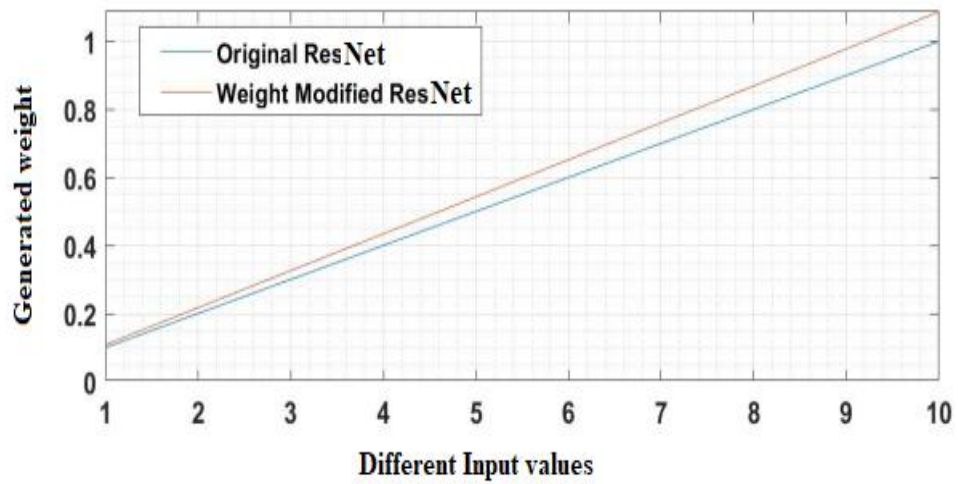


Figure.5.5 Comparison of weights

Table 5.15 Generation of weight for different values of input

SL. No.	Conventional ResNet[102]	WMRESNET
1.	0.1	0.108
2.	0.2	0.217
3.	0.3	0.326
4.	0.4	0.434
5.	0.5	0.543
6.	0.6	0.652
7.	0.7	0.761
8.	0.8	0.869
9.	0.9	0.978
10.	1.0	1.087

Table 5.16 Variation of learning and classification efficiency according to different value of 'P'

SL. No.	Value of 'P'	Learning Efficiency of WMRESNET	Classification Efficiency of SVM
1.	0.3	70.99%	90%
2.	0.4	77%	96%
3.	0.5	72.25%	92%
4.	Conventional ResNet Equation	61.8%	85%

Formulation of the Proposed Weight Modified ResNet (WMRESNET)

Forward Propagation

$$I_{n+1} = I_n + 0.4 * (f(\exp(I_n))) \quad (5.18)$$

Where n represents current residual block where $n+1$ represents its next block. Applying this formula recursively we get:

$$I_{n+2} = I_{n+1} + 0.4 * (f(\exp(I_{n+1}))) \quad (5.19)$$

Substitute the value of I_{n+1} from equation.1. we get

$$I_{n+2} = I_n + 0.4 * (f(\exp(I_n))) + 0.4 * (f(\exp(I_{n+1}))) \quad (5.20)$$

So, equation 5.20 can be represented by its general form as in equation 5.21.

$$I_N = I_n + 0.4 * \sum_{j=n}^{N-1} (f(\exp(I_j))) \quad (5.21)$$

In equation 6.50, (n) represents any previous layer where (N) represents its consecutive layer

Backward Propagation

The partial derivative of generalized Forward Propagation formula with respect to (I_n) is represented by equation 5.22 as below

$$\frac{\partial \alpha}{\partial I_n} = \frac{\partial \alpha}{\partial I_N} \cdot \frac{\partial I_N}{\partial I_n} = \frac{\partial \alpha}{\partial I_N} + \frac{\partial \alpha}{\partial I_N} [0.4 * \sum_{j=n}^{N-1} (f(\exp(I_j)))] \quad (5.22)$$

In the above formula (equation 5.22), the term $(f(\exp(I_j)))$ never becomes zero, so, the total gradient $\frac{\partial \alpha}{\partial I_n}$, never can be vanished. The above equation proved that the proposed weight Modified Residual Net is able to remove vanishing gradient problem efficiently. The architectonic descriptions of WMRESNET are given in table 5.17.

Table 5.17 Architectonic description of the suggested WMRESNET

No. Of Residual Blocks	Max Epoch	Initial Learning Rate	Initial Filter Size	Initial Stride	Polling Technique	No. of Training Images [109]
14	5	0.1	[5 5]	[1 1]	Max Polling	780

SVM Classifier

The feature map generated by WMRESNET has been used to train the SVM [142] for image classification. SVM generates a hyper plane at the decision stage and can classify different category images. In the suggested model the SVM is used to classify different category breast cancers. The architecture of the proposed model is shown in figure 5.6.

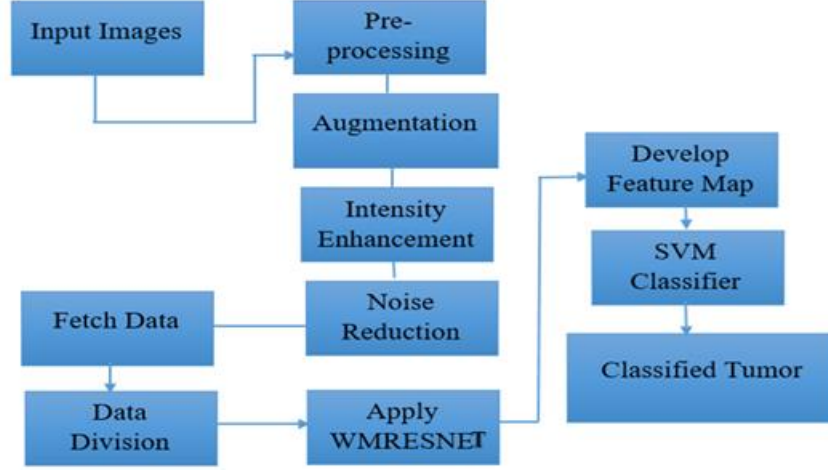


Figure 5.6 Architecture of the suggested model

5.3.2 Results and Discussion

The suggested image classifier is very efficient to classify different breast cancer (Malignant, Benign, and Normal) images with a classification efficiency of 96% with the dataset [109]. The classification efficiency [145] has been calculated by using equation 5.23.

$$\%Classification\ Efficiency = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \quad (5.23)$$

For the assessment of the suggested model, we conducted assessment comparisons with many conventional CNNs with the dataset [109]. During the feature extraction stage, we substituted the proposed WMRESNET with traditional models including ResNet, VGG, and AlexNet individually, and recorded their respective outputs which are detailed in table 5.18 below. Table 5.18 clearly demonstrates that the proposed model outperforms conventional CNNs (like AlexNet, VGG, and ResNet) across learning efficiency, classification accuracy, and computational cost.

Table 5.18 Comparison of proposed model with different CNN models

Name of the CNN							
ResNet		VGG-16		AlexNet		Proposed Model	
% Classification Efficiency	Computational time in minutes	% Classification Efficiency	Computational time in minutes	% Classification Efficiency	Computational time in minutes	% Classification Efficiency	Computational time in minutes
85	35	89	34	92	39	96	30


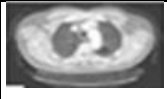
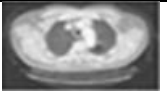
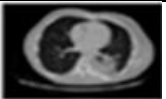





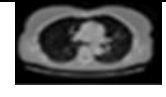
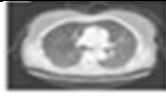
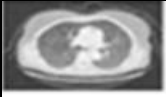

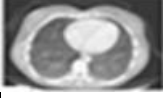
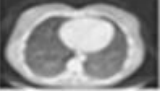







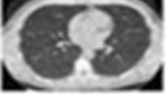

The proposed model has been compared with state-of-the-art CNN-based Breast Cancer Classifier models [109], [146], [147], [148], [149], [150], [151], [152], and [153]. These models similarly utilized (USBI) to discern between benign and malignant cancers from normal images. The comparison is delineated in table 5.19 below.

Table 5.19 Comparison between suggested model and existing CNN based models

CNN based model	Year	Classification Efficiency
CNN model [109]	2023	90%
ViT-patch [146]	2023	89.9%
LeNet CNN [147]	2023	89.91%
BUSIS model [148]	2022	90%
BUViT Net model [149]	2022	95%
Pyramid Trip Deep feature generator [150]	2022	88.67%
Semi supervised GAN based model [151]	2021	90.41%
Variant Enhanced Deep learning model [152]	2021	89.73%
Transfer Learning model [153]	2021	91.5%
Proposed classifier model		96%

To check the proficiency of the suggested model, Lung-cancer image dataset [110] has been used and it has been observed that the training efficiency for this dataset increased from 66% to 76% and the overall classification efficiency increased from 87% to 96%. Different chest cancer images of the dataset [110], after pre-processing is represented in table 5.20.

Table 5.20 Different chest cancer images from dataset [110] after preprocessing

Squamous Cell Carcinoma					
Sample image-1	HE image	Denoised image	Sample image-2	HE image	Denoised image
					
Large Cell carcinoma					
					
Adeno Carcinoma					
					
No cancer					
					

The authors of [154] - [158] employed lung CT images to differentiate between various lung cancer cells. Similarly, the proposed model utilized these CT images to assess its classification accuracy. The performance evaluation of the suggested model against contemporary image classifier models is presented in table 5.21 below.

Table 5.21 Comparison between suggested model and contemporary CNN based model [110]

Classification Model	Year	Classification Accuracy
Deep learning ensemble 2D CNN model [110]	2023	95%
Deep learning-based Support Vector machine [159]	2022	94%
Ensemble of CNN model [160]	2018	87%

Machine Learning model [161]	2018	90%
Deep CNN model [162] (Classified images based on Lung Cancer present or absent)	2018	96%
Multi view CNN model [163]	2016	94%
Proposed Classifier model (Detected and classified three different types of Lung cancers and normal images)		96%

It is clear from table 5.20 and table 5.21 that the proposed model is able to give better classification efficiency related to the contemporary CNN based models.

Computational Complexity

The model was experimented with MATLAB-18 software on an Intel Core i5 system. A relatively small number (780 images) from dataset [109] were utilized to train the WMRESNET, and the depth of the network was not excessively high. Consequently, the training process was completed in a short duration of only 30 minutes. Additionally, the average classification time per image was measured to be 30.36 seconds, which is significantly lower compared to most CNN-based models, indicating the efficiency and speed of the proposed model in both training and classification tasks. During the feature extraction stage, we substituted the proposed WMRESNET with traditional models including ResNet, VGG, and AlexNet individually, and the time taken to train these models were comparatively higher compared to the proposed WMRESNET. The time needed to train these conventional methods were 35mins, 39 mins and 34 mins respectively with the (USBI) [109] dataset.

5.4 Conclusion

Several image classifiers already exist in literature but development of a classifier that can detect and classify very small tumors with low computational cost is still a big challenge. To overcome this problem, we developed an image classifier model that combines Enhanced Anisotropic Diffusion Filtering (EADF) with ResNet-50. We introduced an automatic thresholding method for efficient and rapid detection of brain tumors. The suggested model can detect very small tumors and also can detect tumors from Pseudo colour images. ResNet-50 is used for classification, enabling the model to quickly and accurately differentiate between various types of brain tumors.. Our proposed classifier outperformed several existing image classifiers.

Additionally, the proposed WMRESNET (Weight Modified Residual Network) achieved high classification accuracy on breast and chest cancer images. The Residual Network's skip connection architecture was effectively modified to conquer the vanishing gradient problem significantly by enhancing both model speed and efficiency.

CHAPTER 6

CONCLUSION AND FUTURE WORK

In medical imaging diagnosis, the technique of fusion in the MMIF domain has shown to be highly effective. As a result, instead of having to analyze two separate images, only one image is required, which speeds up diagnosis and reduces computer storage requirements.

In this thesis, we addressed the challenges in MMIF and enhancement algorithms to improve the diagnostic and therapeutic approaches by adding information from multiple imaging modalities and produce more comprehensive images. By integrating data from sources like MRI, CT, $T1$, $T2$, $T1C$, *Flair* and $T1GD$ image fusion provides a richer context, enhancing diagnostic accuracy and better targeted treatments. The proposed algorithms have been compared with contemporary methods and showed promising results when various parameters (included PSNR, RMSE, MI, Entropy etc.) were calculated for all algorithms. We used DWT, DTCWT for the decomposition of different modality source images and PCA [13], mean-max fusion rule [85], Proposed fusion rule etc. for fusing the corresponding coefficients. The addition of enhancement and denoising techniques further refines these fused images, making subtle features more discernible and adding clinicians in identifying potential abnormalities earlier. We also created classification models aimed at early disease detection. Early diagnosis not only saves lives is also crucial for maintaining a healthy life.

6.1 Summary of the work done in the thesis

We conducted a wide range review of current research in MMIF. Expanding this, we analyzed the performance of conventional transformation methods—Haar, Daubechies, and the Lifting scheme—using the BraTs-2015 dataset. This evaluation utilized various performance metrics, including PSNR, RMSE, MI, Mean, Std, FMI, and SSIM. Additionally, we introduced an energy-based coefficient enhancement (EBCE) technique, applied before the fusion with the suitable technique of PCA [13] and the mean-max [85] fusion rule. The EBCE-based enhancement demonstrated its effectiveness in boosting performance compared to several conventional image enhancement techniques.

To reduce unwanted noise in fused images, we employed the proposed AMT-DWT (Average of Multi-Technique-Discrete Wavelet Transform) for preprocessing the source images. Following this enhancement, we applied the suitable technique (PCA [13], mean-max fusion rule [85]) for the fusion of the enhanced images. A custom CNN with Proposed activation functions has been used for denoising, achieving the final refinement of the fused images. The denoising model's performance was then

benchmarked against contemporary MMIF models, where our suggested model exhibited superior results.

For efficient multimodal image fusion, we created a hybrid model, Dense-ResNet, which leverages the strengths of both Dense-net and ResidualNet and three different modality images of MRI ($T1$, $T1GD$ and $T2$) have been fused together for the integration of more source information. During the preprocessing phase, noise reduction was achieved using Median filtering [15], followed by edge enhancement via ET-Net. The suggested fusion model gives enhanced performance when compared to existing image fusion models. Moreover, optimization of the above-mentioned hybrid model is achieved with the help of Pelican Optimization Algorithm (POA). The performance of the optimized model is better than the earlier one, achieving lower MSE and RMSE values while delivering a higher PSNR at the cost of computational time.

Furthermore, to address issues related to unwanted artefacts and information loss during fusion, a set of fusion rules has been introduced for fusing decomposed multimodal medical image coefficients. These proposed rules consistently outperformed traditional image fusion techniques.

We developed an image classifier model using Enhanced Anisotropic Diffusion Filtering (EADF) combined with ResNet-50. We introduced an automatic thresholding method, enabling efficient and rapid detection of brain tumors. In the final stage, ResNet-50 is used for classification, enabling the model to differentiate between various types of brain tumors shortly and accurately.

Our proposed classifier model gives superior performance compared to several existing image classifiers. In addition, the proposed WMRESNET (Weight Modified residual Net) successfully classified breast cancer and chest cancer images with high classification efficiency. The skip connection architecture of the Residual Network has been efficiently modified to reduce the vanishing gradient problem, ultimately enhancing model speed and efficiency.

6.2 Future Scopes

As we explore the vast potential of MMIF techniques in disease detection, we are driven by the belief that the pursuit of knowledge and innovation is a dynamic and ongoing journey, requiring constant investigation, adaptation, and improvement. Many MMIF techniques rely on basic methods, and although issues like spectral distortion, overfitting, and feature extraction have been somewhat tackled, they are still ongoing problems. Creating new algorithms for MMIF continues to be a major challenge in this area. The implementation of innovative algorithms for MMIF continues to pose a major hurdle in this domain. In light of these challenges, the following directions represent key opportunities for future exploration:

1. Difficulties in fusing multimodal images while eliminating background interference.
2. Creating DL methods to address prolonged training times, model overfitting, and challenges in achieving stable convergence.
3. Challenges in extracting and enhancing targeted regions in root images prior to the process of fusion.
4. Challenges in removal of linear and non-linear geometric distortions.

DATA DESCRIPTION

The experimentation for the presented research work was successfully conducted using MATLAB R2018a software on a system equipped with 8 GB RAM and a 3 GHz i5 processor. The descriptions of the datasets are reported in table 6.1 and 6.2.

Table 6.1 Detail description of the dataset used for image fusion and denoisation in this research

SL. No.	Data Set	Type of MRI Images	Number of Images	Total No. of images
1.	BraTs-2015 [76]	<i>T1, T2, T1C, Flair</i>	4 x 274	1096
2.	BraTs-2018 [77]	<i>T1, T2, T1C, Flair</i>	4 x 285	1140
3.	Harvard Medical school Brain dataset [78]	CT, MRI	2 x 8	16
4.	BraTs-2020 [79]	<i>T1, T2, T1GD, Flair</i>	4 x 369	1476

Table 6.2 Detail description of the dataset used for image classification in this research

SL. No.	Data Set	Types of images	Number of Positive Images	Number of Negative Images	Total No. of images
1.	BR35H [106]	<i>T1</i>	86	6	92
2.	BMI-I [106]	<i>T1</i>	86	85	171
3.	BTI [106]	<i>T2</i>	10	10	20
4.	BTS [106]	<i>T2</i>	70	70	140
5.	BD_BT [106]	<i>T2</i>	671	-	671
6.	BraTs-2018 for classifier	<i>T1, T2, T1C, Flair</i>	131	131	1048

	[107]								
7.	Dataset [108]	MRI		2764		440		3204	
8.	(Breast Cancer Ultrasound Images) [109]	No. of Benign cancer Images		No. of Malignant cancer images		No. of normal images		Total no. of images	
		437		210		133		780	
9.	Lung cancer CT Images [110]	No. Of Normal images		Adenocarcinoma cancer images		Large Cell Carcinoma images		Squamous Carcinoma	Total no. of images
		59		322		368		314	1063

Publications Related to the Thesis

Paper Published/Accepted for Publication

Published Journal Papers:

1. T. Ghosh and N. Jayanthi, "An efficient Dense-ResNet for multimodal image fusion using medical image," *Multimedia Tools and Applications*, vol. 2024, pp. 1-28, 2024. <https://doi.org/10.1007/s11042-024-18974-72>.
2. T. Ghosh and N. Jayanthi, "Performance analysis of multimodal medical image fusion using AMT-DWT-based pre-processing and customized CNN for denoising," *Multimedia Tools and Applications*, vol. 2024, pp. 1-33, 2024. <https://doi.org/10.1007/s11042-024-19592-z>.
3. T. Ghosh and N. Jayanthi, "Multimodal fusion of different medical image modalities using optimised hybrid network," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 48, no. 1, pp. 19-33, 2025. <https://www.inderscienceonline.com/doi/abs/10.1504/IJAHUC.2025.14>
4. T. Ghosh and N. Jayanthi, "Impact of preprocessing techniques on MRI-based brain tumor detection," *Multimedia Tools Appl.*, pp. 1-20, 2025. <https://link.springer.com/article/10.1007/s11042-025-20633-4>.

Book chapter:

1. T. Ghosh and N. Jayanthi, "Enhancing sustainable development in medical image fusion using proposed (EBCE) with DWT," in *Proc. Int. Conf. Sustainable Development through Machine Learning, AI and IoT*, Cham, Switzerland, 2023, pp. 236-247.

Conference Papers:

1. T. Ghosh and N. Jayanthi, "Medical image fusion: a critical review," in *Proc. Int. Conf. Advances in Science, Technology and Management (ICSTM-2021)*, 2021.
2. T. Ghosh and N. Jayanthi, "Image fusion and its denoisation by using a new fusion rule," in *Proc. 2022 Third Int. Conf. Intelligent Computing Instrumentation and Control Technologies (ICICT)*, 2022.

Journal Paper in communication:

1. T. Ghosh, and N. Jayanthi "Detection and Classification of Breast cancer by WMRESNET", *Spatial Imaging. (SCIE)*. (under review)

References

- [1] Abhisheka, B., Biswas, S. K., Purkayastha, B., Das, D., & Escargueil, A. (2024). Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. *Multimedia Tools and Applications*, 83(14), 43035-43070.
- [2] Safari, M., Fatemi, A., & Archambault, L. (2023). MedFusionGAN: multimodal medical image fusion using an unsupervised deep generative adversarial network. *BMC Medical Imaging*, 23(1), 203.
- [3] Moghtaderi, S., Einlou, M., Wahid, K. A., & Lukong, K. E. (2024). Advancing multimodal medical image fusion: an adaptive image decomposition approach based on multilevel Guided filtering. *Royal Society Open Science*, 11(4), rsos-231762.
- [4] Khan, S. U., Khan, F., Ullah, S., & Lee, B. (2023). Multimodal medical image fusion in NSST domain with structural and spectral features enhancement. *Heliyon*, 9(6).
- [5] Babu, B. S., & Narayana, M. V. (2023). Two stage multi-modal medical image fusion with marine predator algorithm-based cascaded optimal DTCWT and NSST with deep learning. *Biomedical Signal Processing and Control*, 85, 104921.
- [6] R. Singh and A. Khare, "Fusion of multimodal medical images using Daubechies complex wavelet transform - A multiresolution approach," *Inf. Fusion*, vol. 19, no. 1, pp. 49–60, 2014, doi: 10.1016/j.inffus.2012.09.005.
- [7] Jayant Bhardwaj, Abhijit Nayak, and Kulvinder Singh, "Feature Level Fusion of Gray Dentistry Images using Haar Lifting Wavelet Transform," vol. 8, no. 6, p. MW2C.3, 2017, doi: 10.1364/math.2017.mw2c.3.
- [8] Mankar, R., & Daimiwal, N. (2015, April). Multimodal medical image fusion under nonsubsamped contourlet transform domain. In *2015 International Conference on Communications and Signal Processing (ICCSP)* (pp. 0592-0596). IEEE.
- [9] S. Singh and D. Gupta, "Detail Enhanced Feature-Level Medical Image Fusion in Decorrelating Decomposition Domain," *IEEE Trans. Instrum. Meas.*, vol. 70, no. November, 2021, doi: 10.1109/TIM.2020.3038603.
- [10] Ibrahim, S. I., Makhoul, M. A., & El-Tawel, G. S. (2023). Multimodal medical image fusion algorithm based on pulse coupled neural networks and nonsubsamped contourlet transform. *Medical & Biological Engineering & Computing*, 61(1), 155-177.
- [11] Venkatesan, B., Ragupathy, U. S., & Natarajan, I. (2023). A review on multimodal medical image fusion towards future research. *Multimedia Tools and Applications*, 82(5), 7361-7382.
- [12] K. Padmavathi, C. S. Asha, and V. K. Maya, "A novel medical image fusion by combining TV-L1 decomposed textures based on adaptive weighting scheme," *Eng. Sci. Technol. an Int.J.*, vol. 23, no.1, pp. 225–239, 2020.

- [13] Yang, Y., Ding, M., Huang, S., Que, Y., Wan, W., Yang, M., & Sun, J. (2017). Multi-focus image fusion via clustering PCA based joint dictionary learning. *IEEE Access*, 5, 16985-16997.
- [14] F. E. Z. A. El-Gamal, M. Elmogy, and A. Atwan, "Current trends in medical image registration and fusion," *Egypt. Informatics J.*, vol. 17, no. 1, pp. 99–124, 2016, doi: 10.1016/j.eij.2015.09.002.
- [15] Hussein, T., Omar, H. K., & Jihad, K. (2021). A study on image noise and various image denoising techniques. *Journal of Advanced Research in Engineering and Applied Sciences*, 2(11), 155N.
- [16] Fan, L., Zhang, F., Fan, H., & Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1), 7.
- [17] Miranda, E., Aryuni, M., & Irwansyah, E. (2016, November). A survey of medical image classification techniques. In *2016 international conference on information management and technology (ICIMTech)* (pp. 56-61). IEEE.
- [18] Tang, H., & Hu, Z. (2020). Research on medical image classification based on machine learning. *IEEE access*, 8, 93145-93154.
- [19] Kong, W., Li, C., & Lei, Y. (2022). Multimodal medical image fusion using convolutional neural network and extreme learning machine. *Frontiers in Neurorobotics*, 16, 1050981.
- [20] Trojovský, P., & Dehghani, M. (2022). Pelican optimization algorithm: A novel nature-inspired algorithm for engineering applications. *Sensors*, 22(3), 855.
- [21] J. Zhang, X. Feng, B. Song, M. Li, and Y. Lu, "Multi-focus image fusion using quality assessment of spatial domain and genetic algorithm," *2008 Conf. Hum. Syst. Interact. HSI 2008*, pp. 71–75, 2008, doi: 10.1109/HSI.2008.4581411.
- [22] Arif, M., & Wang, G. (2020). Fast curvelet transform through genetic algorithm for multimodal medical image fusion. *Soft Computing*, 24(3), 1815-1836.
- [23] C. W. Lacewell, M. Gebril, R. Buaba, and A. Homaifar, "Optimization of image fusion using Genetic Algorithms and Discrete Wavelet Transform," *Proc. IEEE 2010 Natl. Aerosp. Electron. Conf. NAECON 2010*, pp. 116–121, 2010, doi: 10.1109/NAECON.2010.5712933.
- [24] Y. Wang, Z. Wan, and Z. Peng, "A Novel Improved Bird Swarm Algorithm for Solving Bound Constrained Optimization Problems," *Wuhan Univ. J. Nat. Sci.*, vol. 24, no. 4, pp. 349–359, 2019, doi: 10.1007/s11859-019-1406-6.
- [25] E. Varol Altay and B. Alatas, *Bird swarm algorithms with chaotic mapping*, vol. 53, no. 2. Springer Netherlands, 2020.
- [26] K. N. V. P. S. Rajesh, R. Dhuli, and T. S. Kumar, "Obstructive sleep apnea detection using discrete wavelet transform-based statistical features," *Comput. Biol. Med.*, vol. 130, no. August 2020, p. 104199, 2021, doi: 10.1016/j.combiomed.2020.104199.

- [27] Velmurugan, S. P., Sivakumar, P., & Rajasekaran, M. P. (2018). Multimodality image fusion using centre-based genetic algorithm and fuzzy logic. *International Journal of Biomedical Engineering and Technology*, 28(4), 322-348.
- [28] Alzahrani, A. A. (2024). Enhanced multimodal medical image fusion via modified DWT with arithmetic optimization algorithm. *Scientific Reports*, 14(1), 19261.
- [29] M. Arif and G. Wang, "Fast curvelet transform through genetic algorithm for multimodal medical image fusion," *Soft Comput.*, vol. 24, no. 3, pp. 1815–1836, 2020, doi: 10.1007/s00500-019-04011-5.
- [30] Y. Xia and M. S. Kamel, "Novel cooperative neural fusion algorithms for image restoration and image fusion," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 367–381, 2007, doi: 10.1109/TIP.2006.888340.
- [31] Saleh, M. A., Ali, A. A., Ahmed, K., & Sarhan, A. M. (2022). A brief analysis of multimodal medical image fusion techniques. *Electronics*, 12(1), 97.
- [32] Nagaraja Kumar, N., Jayachandra Prasad, T., & Satya Prasad, K. (2023). Multimodal medical image fusion with improved multi-objective meta-heuristic algorithm with fuzzy entropy. *Journal of Information & Knowledge Management*, 22(01), 2250063.
- [33] Bhaladhare, P. R., & Jinwala, D. C. (2014). A clustering approach for the l-diversity model in privacy preserving data mining using fractional calculus-bacterial foraging optimization algorithm. *Advances in Computer Engineering*, 2014(1), 396529.
- [34] K. He, J. Gong, L. Xie, X. Zhang, and D. Xu, "Regions preserving edge enhancement for multi-sensor-based medical image fusion," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, doi: 10.1109/TIM.2021.3066467.
- [35] Q. P. Zhang, M. Liang, and W. C. Sun, "Medical diagnostic image fusion based on feature mapping wavelet neural networks," *Proc. - Third Int. Conf. Image Graph.*, pp. 51–54, 2004, doi: 10.1109/ICIG.2004.93.
- [36] Alseelawi, N., Hazim, H. T., & Salim ALRikabi, H. T. (2022). A Novel Method of Multimodal Medical Image Fusion Based on Hybrid Approach of NSCT and DTCWT. *International Journal of Online & Biomedical Engineering*, 18(3).
- [37] L. Chiorean and M. F. Vaida, "Medical image fusion based on discrete wavelet transform using java technology," *Proc. Int. Conf. Inf. Technol. Interfaces, ITI*, no. March 2015, pp. 55–60, 2009, doi: 10.1109/ITI.2009.5196054.
- [38] N. Ansari and A. Gupta, "Image Reconstruction Using Matched Wavelet Estimated from Data Sensed Compressively Using Partial Canonical Identity Matrix," *IEEE Trans. Image Process.*, vol. 26, no. 8, pp. 3680–3695, 2017, doi: 10.1109/TIP.2017.2700719.
- [39] Lei, Y. (2023, May). Research on Image Fusion Algorithm Based on Wavelet Transform. In *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)* (pp. 319-324). IEEE.

- [40] D. P. Bavirisetti and R. Dhuli, "Fusion of Infrared and Visible Sensor Images Based on Anisotropic Diffusion and Karhunen-Loeve Transform," *IEEE Sens. J.*, vol. 16, no. 1, pp. 203–209, 2016, doi: 10.1109/JSEN.2015.2478655.
- [41] D. Xu, Y. Wang, S. Xu, K. Zhu, N. Zhang, and X. Zhang, "Infrared and visible image fusion with a generative adversarial network and a residual network," *Appl. Sci.*, vol. 10, no. 2, 2020, doi: 10.3390/app10020554.
- [42] H. Li, Q. Miao, and H. Shen, "An Improved Medical Image Fusion Algorithm," 2019, pp. 1593–1601.
- [43] G Zhou, T., Li, Q., Lu, H., Zhang, X., & Cheng, Q. (2022). Hybrid multimodal medical image fusion method based on LatLRR and ED-D2GAN. *Applied Sciences*, 12(24), 12758.
- [44] N. Muhammad, N. Bibi, Z. Mahmood, T. Akram, and S. R. Naqvi, "Reversible integer wavelet transform for blind image hiding method," *PLoS One*, vol. 12, no. 5, 2017, doi: 10.1371/journal.pone.0176979.
- [45] V. Bhateja, H. Patel, A. Krishn, A. Sahu, and A. Lay-Ekuakille, "Multimodal Medical Image Sensor Fusion Framework Using Cascade of Wavelet and Contourlet Transform Domains," *IEEE Sens. J.*, vol. 15, no. 12, pp. 6783–6790, 2015, doi: 10.1109/JSEN.2015.2465935.
- [46] S. D. Ramlal, J. Sachdeva, C. K. Ahuja, and N. Khandelwal, "An improved multimodal medical image fusion scheme based on hybrid combination of nonsubsampling contourlet transform and stationary wavelet transform," *Int. J. Imaging Syst. Technol.*, vol. 29, no. 2, pp. 146–160, 2019, doi: 10.1002/ima.22310.
- [47] L. J. Ma and C. H. Zhao, "An effective image fusion method based on nonsubsampling contourlet transform and pulse coupled neural network," *Adv. Mater. Res.*, vol. 756–759, pp. 3542–3548, 2013, doi: 10.4028/www.scientific.net/AMR.756-759.3542.
- [48] V. Bhateja, A. Srivastava, A. Moin, and A. Lay-Ekuakille, "Multispectral medical image fusion scheme based on hybrid contourlet and shearlet transform domains," *Rev. Sci. Instrum.*, vol. 89, no. 8, 2018, doi: 10.1063/1.5016947.
- [49] Goyal, B., Dogra, A., Lepcha, D. C., Koundal, D., Alhudhaif, A., Alenezi, F., & Althubiti, S. A. (2022). Multi-modality image fusion for medical assistive technology management based on hybrid domain filtering. *Expert Systems with Applications*, 209, 118283.
- [50] O. S. Faragallah et al., "A comprehensive survey analysis for present solutions of medical image fusion and future directions," *IEEE Access*, vol. 9, no. Januar. 11358–11371, 2021, doi: 10.1109/ACCESS.2020.3048315.
- [51] Sebastian, J., & King, G. G. (2021). Fusion of multimodality medical images-A review. 2021 *Smart Technologies, Communication and Robotics (STCR)*, 1-6.W.
- [52] R. Srivastava, O. Prakash, and A. Khare, "Local energy-based multimodal medical image fusion in curvelet domain," *IET Comput. Vis.*, vol. 10, no. 6, pp. 513–527, 2016, doi: 10.1049/iet-cvi.2015.0251.

- [53] G.Bhatnagar,Q. M. J. Wu and Z. Liu, "Directive Contrast Based Multimodal Medical Image Fusion in NSCT Domain," in *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1014-1024, Aug. 2013, doi: 10.1109/TMM.2013.2244870.
- [54] Q. Jiang, X. Jin, S. J. Lee, and S. Yao, "A Novel Multi-Focus Image Fusion Method Based on Stationary Wavelet Transform and Local Features of Fuzzy Sets," *IEEE Access*, vol. 5, pp. 20286–20302, 2017, doi: 10.1109/ACCESS.2017.2758644.
- [55] Wankhede, P., Das, M., Gupta, D., Radeva, P., & Bakde, A. M. (2023). A New Multimodal Medical Image Fusion based on Laplacian Autoencoder with Channel Attention. *arXiv preprint arXiv:2310.11896*.
- [56] Gu, X., Xia, Y., & Zhang, J. (2024). Multimodal medical image fusion based on interval gradients and convolutional neural networks. *BMC Medical Imaging*, 24(1), 232.
- [57] M. Kaur and D. Singh, "Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 2, pp. 2483–2493, 2021, doi: 10.1007/s12652-020-02386-0.
- [58] Liang, N. (2024). Medical image fusion with deep neural networks. *Scientific Reports*, 14(1), 7972.
- [59] M. Manchanda and D. Gambhir, "Multifocus image fusion based on waveatom transform," *Sadhana - Acad. Proc. Eng. Sci.*, vol. 44, no. 2, 2019, doi: 10.1007/s12046-018-1010-z.
- [60] L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsamped contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006, doi: 10.1109/TIP.2006.877507.
- [61] Bhavana, V., & Krishnappa, H. K. (2016, March). A survey on multi-modality medical image fusion. In *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (pp. 1326-1329). IEEE.
- [62] S. Katoch, S. S. Chauhan, and V. Kumar, A review on genetic algorithm: past, present, and future, vol. 80, no. 5. *Multimedia Tools and Applications*, 2021.
- [63] Sebastian, J., & King, G. G. (2021). Fusion of multimodality medical images-A review. *2021 Smart Technologies, Communication and Robotics (STCR)*, 1-6.
- [64] K. N. B. Prakash, A. G. Ramakrishnan, S. Member, S. Suresh, and T. W. P. Chow, "Image Feature," *Definitions*, vol. 6, no. 1, pp. 38–45, 2020, doi: 10.32388/jwqi19.
- [65] Tiantian, W., Hu, Z., & Guan, Y. (2024). An efficient lightweight network for image denoising using progressive residual and convolutional attention feature fusion. *Scientific Reports*, 14(1), 9554.

- [66] Patil, R., Bhosale, S., Shirude, S., & Shetty, S. (2023, July). Multi-modal Medical Image Denoising using Wavelet Transform and Principal Component Analysis. In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.
- [67] Muksimova, S., Umirzakova, S., Mardieva, S., & Cho, Y. I. (2023). Enhancing medical image denoising with innovative teacher–student model-based approaches for precision diagnostics. *Sensors*, 23(23), 9502.
- [68] Li, Y., Zhang, K., Shi, W., Miao, Y., & Jiang, Z. (2021). A novel medical image denoising method based on conditional generative adversarial network. *Computational and Mathematical Methods in Medicine*, 2021(1), 9974017.
- [69] Li, S., Qian, P., Zhang, X., & Chen, A. (2021). Research on Image Denoising and Super-Resolution Reconstruction Technology of Multiscale-Fusion Images. *Mobile Information Systems*, 2021(1), 5184688.
- [70] Moran, N., Schmidt, D., Zhong, Y., & Coady, P. (2020). Noisier2noise: Learning to denoise from unpaired noisy data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12064-12072).
- [71] Valsesia, D., Fracastoro, G., & Magli, E. (2020). Deep graph-convolutional image denoising. *IEEE Transactions on Image Processing*, 29, 8226-8237.
- [72] Qu, H., Liu, K., & Zhang, L. (2024). Research on improved black widow algorithm for medical image denoising. *Scientific Reports*, 14(1), 2514.
- [73] Cha, S., & Moon, T. (2019). Fully convolutional pixel adaptive image denoiser. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4160-4169).
- [74] Chen, J., Chen, J., Chao, H., & Yang, M. (2018). Image blind denoising with generative adversarial network based noise modeling. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3155-3164).
- [75] Gondara, L. (2016, December). Medical image denoising using convolutional denoising autoencoders. In *2016 IEEE 16th international conference on data mining workshops (ICDMW)* (pp. 241-246). IEEE.
- [76] Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., ... & Van Leemput, K. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE transactions on medical imaging*, 34(10), 1993-2024.
- [77] Multimodal Brain Tumor Segmentation Challenge database taken from <https://www.med.upenn.edu/sbia/brats2018/data.html>. Accessed on September 2018.
- [78] <http://www.med.harvard.edu/aanlib/home.html>.
- [79] BRATS 2020 dataset taken from <https://www.kaggle.com/datasets/awsaf49/brats2020-trainingdata?select=BraTS20+Training+Metadata.csv>, accessed on July 2023.
- [80] Murugappan V and Sabeenian R. S 2019 Texture based medical image classification

by using multi-scale Gabor rotation-invariant local binary pattern (MGRLBP). *Cluster Computing*, 22-10979:10992.

- [81] Muthevi A and Uppu R B 2017 Leaf classification using completed local binary pattern of textures in: 7th International Advance Computing Conference, pp. 870-874.
- [82] Liang X, Li L, Cheng G and Gao L 2014, Underdeveloped village extraction from high spatial resolution optical image based on GLCM textures and fuzzy classification. In: In 2014 Third International Workshop on Earth Observation and Remote Sensing Applications, pp. 370- 373.
- [83] Kafuo A, Dliaf S, Gonifeda A, Alshatouri Z and Baba A 2017 A Literature survey of Gabor filter and its application. Technical report.
- [84] Yadav S S and Jadhav S M 2019 Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big data*, 6(1):1-18.
- [85] Rocha M M, Landini G and Florindo J B 2023 Medical image classification using a combination of features from convolutional neural networks. *Multimedia Tools and Applications*, 82(13):19299-19322.
- [86] Mohanarathinam A 2020 Enhanced image filtration using threshold based anisotropic filter for brain tumor image segmentation. In: 3rd international conference on intelligent sustainable systems. pp. 308-316.
- [87] Rashid M H O, Mamun M A, Hossain M A and Uddin M P 2018 Brain tumor detection using anisotropic filtering, SVM classifier and morphological operation from MR images. In: International conference on computer, communication, chemical, material and electronic engineering. pp 1-4.
- [88] Palma C A, Cappabianco F A, Ide J S and Miranda P A 2014 Anisotropic diffusion filtering operation and limitations-magnetic resonance imaging evaluation. *IFAC Proceedings Volumes*. 47(3):3887-3892.
- [89] Liang Y, Peng W, Zheng Z J, Silvén O and Zhao G 2021 A hybrid quantum–classical neural network with deep residual learning. *Neural Networks*, 143:133-147.
- [90] Kachouie N N 2008 Anisotropic diffusion for medical image enhancement. *Int. J. Image Process*. 4(4): 436-443.
- [91] Jafar A and Myungho L 2020 Hyperparameter optimization for deep residual learning in image classification. In: International Conference on Autonomic Computing and Self-Organizing Systems Companion, pp. 24-29.
- [92] Maurya R and Wadhwani S 2022 An efficient method for brain image preprocessing with anisotropic diffusion filter and tumor segmentation. *Optik*. 265:169474.
- [93] Dweik M and Ferretti R 2022 Integrating anisotropic filtering, level set methods and convolutional neural networks for fully automatic segmentation of brain tumors in magnetic resonance imaging. *Neuroscience Informatics*. 2 (3):100095.
- [94] Cheng S, Wang L, Du A 2019 An adaptive and asymmetric residual hash for fast

image retrieval. IEEE Access, 7:78942-78953.

- [95] Singh T R, Roy S, Singh O I, Sinam T and Singh K M 2012 A new local adaptive thresholding technique in binarization. arXiv preprint arXiv: 1201:5227.
- [96] Nandihal P, Bhat V S and Pujari J 2018 Adaptive min max thresholds algorithm of microarray image denoising based on no subsampled contourlet transform. IJERT. 6:1-5.
- [97] Shafiq M and Gu Z 2022 Deep residual learning for image recognition: A survey. Applied Sciences, 12(18): 8972.
- [98] Ganatra, V., & Goel, D. (2024). PRECISe: Prototype-Reservation for Explainable Classification under Imbalanced and Scarce-Data Settings. arXiv preprint arXiv:2408.05754.
- [99] Hata, A., Aoyagi, K., Hino, T., Kawagishi, M., Wada, N., Song, J., ... & Hatabu, H. (2024). Automated interstitial lung abnormality probability prediction at CT: a stepwise machine learning approach in the Boston Lung Cancer Study. Radiology, 312(3), e233435.
- [100] Bamber, S. S., & Vishvakarma, T. (2023). Medical image classification for Alzheimer's using a deep learning approach. Journal of Engineering and Applied Science, 70(1), 54.
- [101] Maurya R and Wadhwani S 2022 An efficient method for brain image preprocessing with anisotropic diffusion filter and tumor segmentation. Optik.265:169474.
- [102] Raj, R. J. S., Shobana, S. J., Pustokhina, I. V., Pustokhin, D. A., Gupta, D., & Shankar, K. J. I. A. (2020). Optimal feature selection-based medical image classification using deep learning model in internet of medical things. IEEE Access, 8, 58006-58017.
- [103] João A, Gambaruto A and Sequeira A 2020 Anisotropic gradient-based filtering for object segmentation in medical images. Computer Methods in Biomechanics and Biomedical Engineering Imaging and Visualization. 8(6):621-630.
- [104] Lai, Z., & Deng, H. (2018). Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron. Computational intelligence and neuroscience, 2018(1), 2061516.
- [105] Yadav, S. S., & Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big data, 6(1), 1-18.
- [106] Naseer A, Yasir T, Azhar A, Shakeel T and Zafar K 2021 Computer-aided brain tumor diagnosis: performance evaluation of deep learner CNN using augmented brain MRI. International Journal of Biomedical Imaging. 2021(1):5513500.
- [107] Ranjbarzadeh R, Bagherian Kasgari A, Jafarzadeh Ghouschi S, Anari S, Naseri M and Bendeache M 2021 Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. Scientific Reports. 11(1):1-17.
- [108] Khaliki M Z and Başarslan M S 2024 Brain tumor detection from images and

comparison with transfer learning methods and 3-layer CNN. Scientific Reports. 14(1): 2664.

- [109] Ali MD, Saleem.A, Elahi H, Khan M A, Khan M I and Yaqoob M M 2023 Breast Cancer Classification through Meta-Learning Ensemble Technique Using Convolution Neural Networks. Diagnostics, 13(13): 2242.
- [110] Shah A A, Malik H A M, Muhammad A, Alourani A and Butt Z A 2023 Deep learning ensemble 2D CNN approach towards the detection of lung cancer. Scientific reports, 13(1):2987.
- [111] Kiaei, A. A., Khotanlou, H., Abbasi, M., Kiaei, P., & Bhrouzi, Y. (2019). An objective evaluation metric for image fusion based on Del operator. arXiv preprint arXiv:1905.07709.
- [112] Singh, V. (2020, November). Image Fusion Metrics–A Systematic Study. In Proceedings of the 2nd International Conference on IoT, Social, Mobile, Analytics & Cloud in Computational Vision & Bio-Engineering (ISMAC-CVB 2020).
- [113] Liu, Y., Chen, X., Cheng, J., & Peng, H. (2017, July). A medical image fusion method based on convolutional neural networks. In 2017 20th international conference on information fusion (Fusion) (pp. 1-7). IEEE.
- [114] Zhang, Y., Guo, C., & Zhao, P. (2021). Medical image fusion based on low-level features. Computational and Mathematical Methods in Medicine, 2021.
- [115] Lu, H., Zhang, L., & Serikawa, S. (2012). Maximum local energy: An effective approach for multisensor image fusion in beyond wavelet transform domain. Computers & Mathematics with Applications, 64(5), 996-1003.
- [116] Li, S., Kang, X., Fang, L., Hu, J., & Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. information Fusion, 33, 100-112.
- [117] Liu, Y., Liu, S., & Wang, Z. (2014). Medical image fusion by combining nonsubsampling contourlet transform and sparse representation. In Pattern Recognition: 6th Chinese Conference, CCPR 2014, Changsha, China, November 17-19, 2014. Proceedings, Part II 6 (pp. 372-381). Springer Berlin Heidelberg.
- [118] Das, S., & Kundu, M. K. (2012). NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency. Medical & biological engineering & computing, 50(10), 1105-1114.
- [119] Zhu, Z., Zheng, M., Qi, G., Wang, D., & Xiang, Y. (2019). A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain. Ieee Access, 7, 20811-20824.
- [120] Das, S., & Kundu, M. K. (2013). A neuro-fuzzy approach for medical image fusion. IEEE transactions on biomedical engineering, 60(12), 3347-3353.
- [121] Yin, M., Liu, X., Liu, Y., & Chen, X. (2018). Medical image fusion with parameter-

adaptive pulse coupled neural network in nonsubsampling shearlet transform domain. *IEEE Transactions on Instrumentation and Measurement*, 68(1), 49-64.

- [122] Yang, Y. (2011). A novel DWT based multi-focus image fusion method. *Procedia engineering*, 24, 177-181.
- [123] Shreyamsha Kumar, B. K. (2015). Image fusion based on pixel significance using cross bilateral filter. *Signal, image and video processing*, 9, 1193-1204.
- [124] Fan, F., Huang, Y., Wang, L., Xiong, X., Jiang, Z., Zhang, Z., & Zhan, J. (2019). A semantic-based medical image fusion approach. *arXiv preprint arXiv:1906.00225*.
- [125] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Dense-netly connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [126] Yang, Y., Park, D. S., Huang, S., & Rao, N. (2010). Medical image fusion via an effective wavelet-based approach. *EURASIP journal on advances in signal processing*, 2010, 1-13.
- [127] Du, J., Li, W., & Xiao, B. (2018). Fusion of anatomical and functional images using parallel saliency features. *Information Sciences*, 430, 567-576.
- [128] Zhang, H., Quan, X., & Wang, J. (2022). A two-level dynamic adaptive network for medical image fusion. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-17.
- [129] Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X., & Zhang, L. (2020). IFCNN: A general image fusion framework based on convolutional neural network. *Information Fusion*, 54, 99-118.
- [130] Zhang, H., Xu, H., Xiao, Y., Guo, X., & Ma, J. (2020, April). Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 07, pp. 12797-12804).
- [131] Xu, H., Ma, J., Jiang, J., Guo, X., & Ling, H. (2020). U2Fusion: A unified unsupervised image fusion network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1), 502-518.
- [132] Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences*, 114(13), 3521-3526.
- [133] Houshy, N., Giurgiu, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Gesmundo, A., ... & Gelly, S. (2019, May). Parameter-efficient transfer learning for NLP. In *International conference on machine learning* (pp. 2790-2799). PMLR.
- [134] Zeng, G., Chen, Y., Cui, B., & Yu, S. (2019). Continual learning of context-dependent processing in neural networks. *Nature Machine Intelligence*, 1(8), 364-372.
- [135] Muzammil, S. R., Maqsood, S., Haider, S., & Damaševičius, R. (2020). CSID: A novel

multimodal image fusion algorithm for enhanced clinical diagnosis. *Diagnostics*, 10(11), 904.

- [136] Liu, Y., Chen, X., Ward, R. K., & Wang, Z. J. (2016). Image fusion with convolutional sparse representation. *IEEE signal processing letters*, 23(12), 1882-1886.
- [137] Liu, Y., Chen, X., Ward, R. K., & Wang, Z. J. (2019). Medical image fusion via convolutional sparsity based morphological component analysis. *IEEE Signal Processing Letters*, 26(3), 485-489.
- [138] Haribabu, M., & Guruviah, V. (2023). An Improved Multimodal Medical Image Fusion Approach Using Intuitionistic Fuzzy Set and Intuitionistic Fuzzy Cross-Correlation. *Diagnostics*, 13(14), 2330.
- [139] Haribabu, M., & Guruviah, V. (2023). Enhanced multimodal medical image fusion based on Pythagorean fuzzy set: an innovative approach. *Scientific Reports*, 13(1), 16726.
- [140] Liu, S., Wang, M., Yin, L., Sun, X., Zhang, Y. D., & Zhao, J. (2022). Two-scale multimodal medical image fusion based on structure preservation. *Frontiers in Computational Neuroscience*, 15, 803724.
- [141] Zhang, Z., Fu, H., Dai, H., Shen, J., Pang, Y. and Shao, L., “Et-net: A generic edge-attention guidance network for medical image segmentation”, In *Proceedings of 22nd International Conference, In Medical Image Computing and Computer Assisted Intervention–MICCAI 2019, Shenzhen, China, Part I vol.22*, Springer International Publishing, pp. 442-450, October 13–17 2019.
- [142] Shahajad M, Gambhir D and Gandhi R 2021 Features extraction for classification of brain tumor MRI images using support vector machine. In: *11th International Conference on Cloud Computing, Data Science & Engineering*. pp. 767-772.
- [143] Vankdothu R, Hameed M A and Fatima H 2022 A brain tumor identification and classification using deep learning based on CNN-LSTM method. *Computers and Electrical Engineering*. 101:107960.
- [144] Srinivas C, KS N P, Zakariah M, Alothaibi Y A, Shaukat K, Partibane B and Awal H 2022 Deep transfer learning approaches in performance analysis of brain tumor classification using MRI images. *Journal of Healthcare Engineering*. 3264367.
- [145] Loizidou K, Elia R and Pitris C 2023 Computer-aided breast cancer detection and classification in mammography: A comprehensive review. *Computers in Biology and Medicine*, 106554.
- [146] Feng H, Yang B, Wang J, Liu M, Yin L and Zheng W 2023 Identifying malignant breast ultrasound images using ViT. *Applied Sciences*, 13(6):3489.
- [147] Balasubramaniam, S, Velmurugan Y, Jaganathan, D and Dhanasekaran S 2023 A modified LeNet CNN for breast cancer diagnosis in ultrasound images. *Diagnostics*, 13(17):2746.
- [148] Zhang Y, Xian M, Cheng H D, Shareef B, Ding J and Xu F 2022 BUSIS: a benchmark

for breast ultrasound image segmentation. In: Healthcare, Vol. 10, No. 4, p. 729.

- [149] Ayana G and Choe S W 2022 BUViTNet: breast ultrasound detection via vision transformers. *Diagnostics*, 12(11):2654.
- [150] Kaplan E, Chan W Y, Dogan S, Barua P D, Bulut H T and Tuncer T 2022 Automated BI-RADS classification of lesions using pyramid triple deep feature generator technique on breast ultrasound images. *Medical Engineering & Physics*, 108:103895.
- [151] Pang T, Wong J H D, Ng W L and Chan C S 2021 Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification. *Computer Methods and Programs in Biomedicine*, 203: 106018.
- [152] Ilesanmi A E, Chaumrattanakul U and Makhanov S S 2021 A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning. *Biocybernetics and Biomedical Engineering*, 41(2):802-818.
- [153] Ayana G, Dese K and Choe S W 2021 Transfer learning in breast cancer diagnoses via ultrasound imaging. *Cancers*, 13(4):738.
- [154] Shafi I, Din S, Khan A, Díez I D L T, Casanova R D J P, Pifarre K T and Ashraf I 2022 An effective method for lung cancer diagnosis from ct scan using deep learning-based support vector network. *Cancers*, 14(21):5457.
- [155] Li C, Zhu G, Wu X and Wang Y 2018 False-positive reduction on lung nodules detection in chest radiographs by ensemble of convolutional neural networks. *IEEE Access*, 6:16060-16067.
- [156] Faisal M I, Bashir S, Khan Z S and Khan F H 2018 An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In: 3rd International conference on emerging trends in engineering, sciences and technology, pp. 1-4.
- [157] Sasikala S, Bharathi M and Sowmiya B R 2018 Lung cancer detection and classification using deep CNN. *International journal of innovative technology and exploring engineering*, 8(25):259-262.
- [158] Setio A A A, Ciompi F, Litjens G, Gerke P, Jacobs C and Van Riel 2016 Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE transactions on medical imaging*, 35(5):1160-1169.
- [159] Shafi I, Din S, Khan A, Díez I D L T, Casanova R D J P, Pifarre K T and Ashraf I 2022 An effective method for lung cancer diagnosis from ct scan using deep learning-based support vector network. *Cancers*, 14(21):5457.
- [160] Li C, Zhu G, Wu X and Wang Y 2018 False-positive reduction on lung nodules detection in chest radiographs by ensemble of convolutional neural networks. *IEEE Access*, 6:16060-16067.

- [161] Faisal M I, Bashir S, Khan Z S and Khan F H 2018 An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In: 3rd International conference on emerging trends in engineering, sciences and technology, pp. 1-4.
- [162] Sasikala S, Bharathi M and Sowmiya B R 2018 Lung cancer detection and classification using deep CNN. International journal of innovative technology and exploring engineering, 8(25):259-262.
- [163] Setio A A A, Ciompi F, Litjens G, Gerke P, Jacobs C and Van Riel 2016 Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE transactions on medical imaging, 35(5):1160-1169