Automated Multimodal Fusion with PDE Preprocessing and Learnable Convolutional Pools

Gargi J Trivedi[1], Rajesh Sanghvi[2]

[1,2]Department of Applied Science & Humanities,
G H Patel College of Engineering & Technology,
CVM University, Vallabh Vidhyanagar-388120, India,
gargi1488@gmail.com ,rajeshsanghvi@gcet.ac.in

Abstract*: *This research paper introduces a novel automated multimodal and Multifocus fusion framework tailored for medical imaging applications. The proposed approach leverages advanced deep learning techniques, incorporating Partial Differential Equation (PDE) preprocessing and learnable convolutional pools. The algorithm accommodates diverse medical modalities, such as MRI, CT, visual, infrared, and multi-focus images. Through modality-specific preprocessing, modified convolutional layers, and adaptive pooling, the model intelligently fuses information from various sources, enhancing the overall imaging quality. Experimental evaluations demonstrate the effectiveness of the proposed method in generating high-quality multimodal medical images, showcasing its potential for improving diagnostic accuracy and clinical decision-making.

***Keywords*: Multimodal Fusion, Deep Learning, Partial Differential Equation (PDE), Convolutional Neural Networks (CNN), Automated Pooling, Multi-focus Images.**

(Article history: Received: 8th May 2023 and accepted 23rd December 2023)

I. INTRODUCTION

Medical imaging plays a pivotal role in modern healthcare, providing clinicians with invaluable insights into the human body's intricate structures and functions. With the advent of diverse imaging modalities, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), visual, infrared, and multi-focus imaging, the wealth of information available for diagnosis and treatment has significantly expanded[1].However, the integration of information from these varied sources remains a complex challenge. Traditional fusion methods often fall short in capturing the nuanced relationships between different modalities, limiting the potential for comprehensive and accurate medical assessments[2].

In response to this challenge, our research presents an innovative approach to automated multimodal fusion, leveraging advanced deep learning techniques. The proposed framework is designed to intelligently integrate information from diverse medical imaging modalities, encompassing both structural and functional data. To enhance the quality of the fusion process, the algorithm incorporates Partial Differential Equation (PDE) preprocessing, addressing the inherent variations in image characteristics. Furthermore, a set of modified convolutional layers and adaptive pooling mechanisms are introduced, tailoring the model to the unique attributes of each modality[31-32].

The motivation behind this research stems from the imperative to improve diagnostic accuracy and facilitate informed clinical decision-making. By developing a sophisticated multimodal fusion algorithm, our aim is to create a robust tool capable of generating high-quality fused images that harness the complementary strengths of various imaging modalities. Such an approach holds promise not only for enhancing diagnostic capabilities but also for advancing personalized medicine through a more comprehensive understanding of patient conditions.

In this paper, we present the methodology, experimental setup, results, and discussions that highlight the effectiveness of our proposed approach. The findings contribute to the growing field of intelligent medical imaging, showcasing the potential of automated multimodal fusion to significantly impact clinical practice and patient care.

II. LITERATURE REVIEW

Multimodal medical image fusion has garnered significant attention due to its potential to enhance diagnostic accuracy, treatment planning, and overall clinical decision-making. Various approaches have been explored in the literature to address the challenges associated with integrating information from diverse imaging modalities[3].

The integration of multiple imaging modalities has become increasingly relevant in medical research and practice. References [4] and [5] have emphasized the importance of combining modalities such as MRI, CT, visual, infrared, and multi-focus imaging to capture diverse physiological and pathological aspects comprehensively. These modalities, while individually powerful, collectively contribute to a more holistic understanding of medical conditions.

Traditional image fusion methods, including mathematical operations like averaging or weighted summation, have been widely employed[6]. However, these

techniques may struggle to capture complex relationships among diverse data sources[7]. Advanced methods such as wavelet transform [8], independent component analysis [9-10]and sparse representation [11] have addressed some of these limitations. Despite their successes, challenges persist in adapting these techniques to the dynamic nature of medical images.

Current multimodal fusion approaches face limitations. Many lack adaptability to handle variations in image characteristics and the inherent complexity of medical data [12]. The incorporation of deep learning techniques for automated feature learning and representation has shown promise [13-18], yet this area requires further exploration.

In response to these observations, our research contributes by proposing an innovative multimodal fusion framework. By integrating Partial Differential Equation (PDE) preprocessing and learnable convolutional pools, our approach seeks to address identified limitations, providing a more adaptive and intelligent solution for medical image fusion.

III. METHODOLOGY

This methodology section provides a structured approach to automated multimodal fusion, emphasizing the use of the heat equation for PDE-based preprocessing.

A. Modality Specific Image Preprocessing

Prior to fusion, each modality undergoes modality-specific preprocessing steps as described below.

- MRI and CT images, normalization of intensity values is performed to create a consistent scale:

$$\text{Normalized}_{MRI} = \frac{I_{\text{MRI}} - \min(I_{\text{MRI}})}{\max(I_{\text{MRI}}) - \min(I_{\text{MRI}})} \quad . \tag{1}$$

- Visual and infrared images undergo contrast enhancement to improve the visibility of features.

$$\left.\begin{aligned} Contrast\_Enhance\_VI &= Contrast\_Stretch(I_{VI}) \\ Contrast\_Enhance\_IF &= Contrast\_Stretch(I_{IF}) \end{aligned}\right\} \tag{2}$$

Here $Contrast\_Stretch$ represents a function applying linear contrast stretching is applied using a linear transformation:

$$O(x,y) = \frac{I(x,y) - \min(I)}{\max(I) - \min(I)} \times (max\_out\ min\_out) + min\_out \tag{3}$$

where $O(x,y)$ is the pixel value in the output image, $I(x,y)$ $\min(I)$and $\max(I)$ are the is the pixel value in the input image, minimum and maximum pixel values in the input image, min_out and max_out are the desired minimum and maximum pixel values in the output image, respectively. Gamma correction involves adjusting the intensity values of an image using a power-law function $O(x,y) = round\left(255 \times \left(\frac{I(x,y)}{255}\right)^{\gamma}\right)$ (4)

- Multi-focus images are subjected to focus adjustment to harmonize focal planes. For each pixel (x,y) in the focus-adjusted image as show below.

$$\text{Focus Adjusted Multi-Focus} = \sum_{i=1}^{N} w_i \cdot I_i \tag{5}$$

Where The weights w_i are determined based on the quality of focus at each level to assess the clarity or sharpness of each pixel at different focus levels. The higher the quality of focus at a particular level, the higher the weight assigned to that level. This emphasizes more focused regions in the final focus-adjusted image. This process ensures that details from different focal planes are appropriately combined, and the resulting focus-adjusted image provides enhanced visibility of details across the various focus levels.

B. Partial Differential Equation (PDE) Preprocessing

The heat equation is employed for PDE-based preprocessing [19-21]. Let I represent the input image, Δt denote the time step, and α be the thermal diffusivity coefficient: $\frac{\partial I}{\partial t} = \alpha \nabla^2 I$ (6)

This equation facilitates noise reduction and enhancement of image features by smoothing variations in intensity. The PDE preprocessing step is applied separately to each modality, ensuring adaptability to the unique characteristics of diverse medical images, Infrared and visible images as well as Multifocus images. The discretized form of the heat equation can be implemented using numerical methods, such as finite differences, to obtain the pre-processed images

$$I_{\text{PDE-Preprocessed}} = \text{ApplyPDE}(I, \alpha, \Delta t) \tag{7}$$

Here, (ApplyPDE) is a function that implements the discretized heat equation. The three parameters the input image I the thermal diffusivity coefficient (α) and the time step (Δt) for image smoothing and preprocessing.

Equation (5) is The heat equation in its continuous form is given by $\frac{\partial I}{\partial t} = \alpha \nabla^2$), $\frac{\partial I}{\partial t}$ is the rate of change of intensity over timee, and $\nabla^2 I$ is the Laplacian of the intensity. For discrete image processing, this continuous equation is discretized using numerical method such as finite differences. A common discretization for a 2D image is:

$$I_{\text{new}}(x,y,t+\Delta t) = I(x,y,t) + \alpha \Delta t \nabla^2 I(x,y,t) \tag{8}$$

The Laplacian$\nabla^2 I$ of an image $I(x,y)$ is computed using the second spatial derivatives. In discrete form, the Laplacian can be approximated using a convolution operation with a Laplacian kernel.

$$\nabla^2 I(x,y) = I(x+1,y) + I(x-1,y) + I(x,y+1) + I(x,y-1) - 4I(x,y) \tag{9}$$

This formula calculates the sum of the differences between the pixel value at the centre $I(x,y)$ and its neighbors in the horizontal and vertical directions. This difference emphasizes the spatial variation or curvature in the image. This process is iteratively applied to the input image over multiple time steps to achieve the desired smoothing and noise reduction.

C. Modified Convolutional Layer

Modified convolutional layers were specifically tailored for each modality. The design involved adjusting key parameters such as kernel sizes (K), strides (S), and filter weights (W) based on the unique attributes of the

corresponding modality. the convolutional operation can be represented as follows:

$$F_{\text{out}}(x,y,k) = \sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{c=1}^{C_{\text{in}}} W_{i,j,c,k} \cdot F_{\text{in}}(x \cdot S + i, y \cdot S + j, c) \quad (10)$$

In equation (10) $F_{\text{out}}(x,y,k)$ is the output feature map at position (x,y) for the k^{th} filter, $F_{\text{in}}(x,y,c)$ is the input feature map at position (x,y) for c^{th} channel $W_{\{i,j,c,k\}}$ are the filter weights for the convolution.

PDE-enhanced features obtained from the dataset used as input to these layers for feature extraction. Let $F_{\text{PDE}}(x,y,c)$ represent the PDE-enhanced features at position (x,y) for c^{th} channel. Then, the input to the convolutional layer is given by , $F_{\text{in}}(x,y,c) = F_{\text{PDE}}(x,y,c)$ (11)

The modified convolutional layers underwent a training process. The primary objective was to optimize the layer parameters, including kernel sizes, strides, and filter weights. This training aimed to enable the model to learn modality-specific features and their interdependencies. During training, the model aims to minimize a loss function L that quantifies the difference between the predicted output and the ground truth. The optimization process involves adjusting the parameters of the modified convolutional layers using techniques gradient descent.

$$\text{Minimize } L(W,K,S) \quad (12)$$

The model updates these parameters iteratively based on the gradients of the loss with respect to the parameters. The loss function L measures the difference between the predicted output Y_{pred} and the ground truth Y_{true}.

$$L(W,K,S) = \text{Loss}\left(Y_{\text{pred}}, Y_{\text{true}}\right) \quad (13)$$

The model takes input data PDE-enhanced features and produces a predicted output Y_{pred}.

$$Y_{\text{pred}} = \text{Model}(X; W,K,S) \quad (14)$$

During each iteration of training, the parameters W represents the filter weights, K represents the kernel sizes, S represents the strides. W, K, S are updated in the direction opposite to the gradients of the loss function with respect to those parameters.

$$\left.\begin{aligned} W_{\text{new}} &= W_{\text{old}} - \eta \cdot \nabla_W L \\ K_{\text{new}} &= K_{\text{old}} - \eta \cdot \nabla_K L \\ S_{\text{new}} &= S_{\text{old}} - \eta \cdot \nabla_S L \end{aligned}\right\} \quad (15)$$

Here, (η) is the learning rate, a hyperparameter that controls the size of the step taken during each update. The process of updating parameters using gradient descent is repeated for multiple iterations until convergence.

In summary, the section outlines the design of modality-specific convolutional layers, the mathematical representation of the convolutional operation, and the training process involving the optimization of layer parameters for effective feature learning.

D. Automated Pooling Mechanisms

Automated pooling mechanisms were introduced to dynamically adjust pooling operations based on the characteristics of each modality. The goal is to optimize pooling for different image features.

In general, the pooling operation reduces the spatial dimensions of the input feature map by aggregating information within a local neighbourhood. The pooling operation is typically applied using operations such as max pooling or average pooling. For a given pooling operation P applied to a feature map F, the output feature map F_{pooled} at position (x,y) is computed as follows.

$$F_{\text{pooled}}(x,y,c) = P\left(\{F(x+i, y+j, c)\}_{i,j \in \text{pooling window}}\right) \quad (16)$$

Here: $F_{\text{pooled}}(x,y,c)$ is the output of the pooling operation at position (x,y) for the c^{th} channel. $F(x+i, y+j, c)$ represents the input feature map values within the pooling window centered at position (x,y).

To enhance adaptability, learnable pooling layers were introduced. These layers allow the model to dynamically learn the pooling strategy during training. Let W_{pool} denote the learnable parameters associated with the pooling layer. The modified pooling operation with learnable parameters is defined as:

$$F_{\text{pooled}}(x,y,c) = P_{\text{learnable}}\left(\{F(x+i, y+j, c) \cdot W_{\text{pool}}(x+i, y+j, c)\}_{i,j \in \text{pooling window}}\right) \quad (17)$$

Here, $P_{\text{learnable}}$ is the learnable pooling operation that may include additional parameters like W_{pool}.

During training, the model aims to optimize the learnable parameters W_{pool} to improve the pooling strategy. This optimization is performed by adjusting W_{pool} through gradient descent.

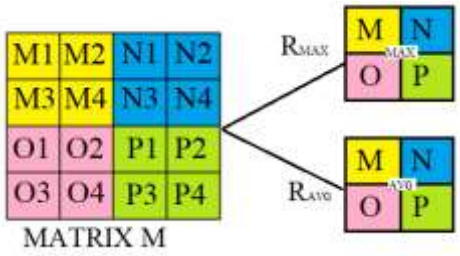

Fig. 1. Automated pooling Mechanism.

In summary, the section introduces automated pooling mechanisms, emphasizing adaptability through learnable pooling layers. The mathematical explanation covers the general pooling operation, the introduction of learnable pooling layers, and the modified pooling operation with learnable parameters.

E. Fusion Strategy

The fusion mechanism integrated features from different modalities. The PDE-pre-processed and convolutionally extracted features were combined to generate a fused representation. Strategies such as weighted fusion, considering the importance of each modality, were applied for optimal integration. Strategies such as weighted fusion, considering the importance of each modality, were applied

for optimal integration. The weighted fusion operation can be expressed as:

$$F_{\text{fused}}(x,y,c) = W_{\text{PDE}}(x,y,c) \cdot F_{\text{PDE}}(x,y,c) + W_{\text{conv}}(x,y,c) \cdot F_{\text{conv}}(x,y,c) \quad (18)$$

Here $F_{\text{PDE}}(x,y,c)$ represent the PDE-preprocessed features , $F_{\text{conv}}(x,y,c)$ represent the convolutionally extracted features at position (x,y) for the c^{th} channel , $W_{\text{PDE}}(x,y,c)$ and $W_{\text{conv}}(x,y,c)$ denote the weights associated with the PDE-preprocessed and convolutionally extracted features, respectively.

IV. Evaluation

The performance of the proposed method was evaluated using a set of quantitative and qualitative metrics [24]. Quantitative analysis involved assessing image quality metrics, while qualitative analysis included visual inspections and comparisons. Comparative studies against existing fusion methods provided insights into the superiority of the proposed approach.

A. Data Collection

The research data utilized in this research was collected from various sources, including 'The Whole Brain Atlas' [25] (https://www.med.harvard.edu/aanlib/home.html), 'The TNO Multiband Image Data Collection' by A. Toet [26] (Data in Brief, vol. 15, pp. 249–251, Dec. 2017, doi: 10.1016/j.dib.2017.09.038), RAWSAMPLES.CH [27] (http://rawsamples.ch/index.php/en/), and GitHub repositories, encompassing a diverse range of datasets comprising medical images, infrared and visible images, and multifocus images.

TABLE I. Image Dataset Specification

Image Dataset	Image type	Resolution
Multimodal	Grayscale TIF	256×256
Infared and visible	Graylscale PNG	360×270,430×340,512×512
Multifocus	RGB JPG	520×520

B. Experimental Setup

The successful implementation of the proposed automated multimodal and multifocus fusion framework relies on a carefully configured training setup. The deep learning architecture incorporates modality-specific modified convolutional layers, automated pooling with learnable pooling windows, and a fusion mechanism guided by specific weights. The training process involves essential parameters such as the learning rate, the number of training epochs, and the batch size. The experiments were conducted on a system equipped with an Intel Core i7 CPU and 16GB RAM. The software environment used for implementation and analysis was MATLAB 2021a.The following Table II outlines the key components, their associated parameters, and the corresponding values used to optimize the model during the training phase.

TABLE II. Training Configuration

Component	Parameter	Value
Deep Learning Architecture		
Modality-Specific Modified Convolutional Layers	Kernel Size (K)	3
	Stride (S)	1
Automated Pooling	Learnable Pooling Window Size	2×2
Fusion Mechanism	Fusion Weight (WPDE)	0.7
	Fusion Weight (WPOOL)	0.3
Training Parameters		
Learning Rate (η)	(η)	0.001
Number of Training Epochs	-	50
Batch Size	-	32

C. Qualitaive anlysis

As part of evaluation, the quantitative assessment using various quality metrics, visual comparisons were conducted by generating images through each of the methods—CBF [28], CNN[29], FPDE[30], and the proposed method. The resulting visual images were presented side by side, providing an intuitive and insightful representation of the comparative performance of these techniques. This dual evaluation approach, combining quantitative metrics and visual representations, offers a comprehensive understanding of the strengths and weaknesses of each method in generating fused images.

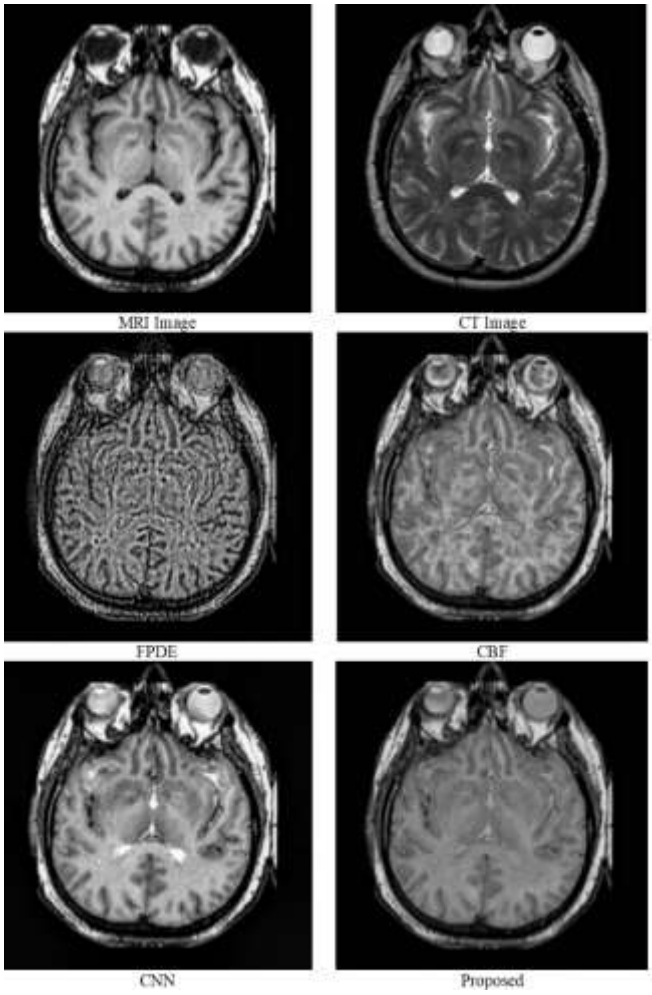

Fig. 2. Comparison of Fused MRI and CT Brain Images using Proposed Method, CBF, CNN, and FPDE.

In the visual analysis as shown in Fig. 2 images showcase the effectiveness of different fusion techniques in enhancing multimodal brain imaging, the proposed image fusion method was compared with existing techniques, including CNN CBF and FPDE, across diverse datasets. For medical brain images captured using CT and MRI modalities, the proposed fusion method exhibited notable improvements in preserving fine details and enhancing overall image clarity compared to CBF and FPDE.

In the case of infrared and visible images depicting a tank scene as shown in Fig. 3 is visual comparison illustrates the impact of different fusion techniques on enhancing features in both infrared and visible spectra, the proposed method demonstrated superior performance by effectively combining thermal and visual information, resulting in images with enhanced contrast and richer features.

Furthermore, for multi-focus images captured by a hand camera observing a globe showing in Fig. 4, the proposed fusion method excelled in harmonizing focal planes, producing a fused image with improved focus across various depths. The visual analysis suggests that the proposed fusion method outperforms CBF and FPDE in handling diverse image modalities, showcasing its effectiveness in preserving critical details and achieving a more comprehensive fusion of information.

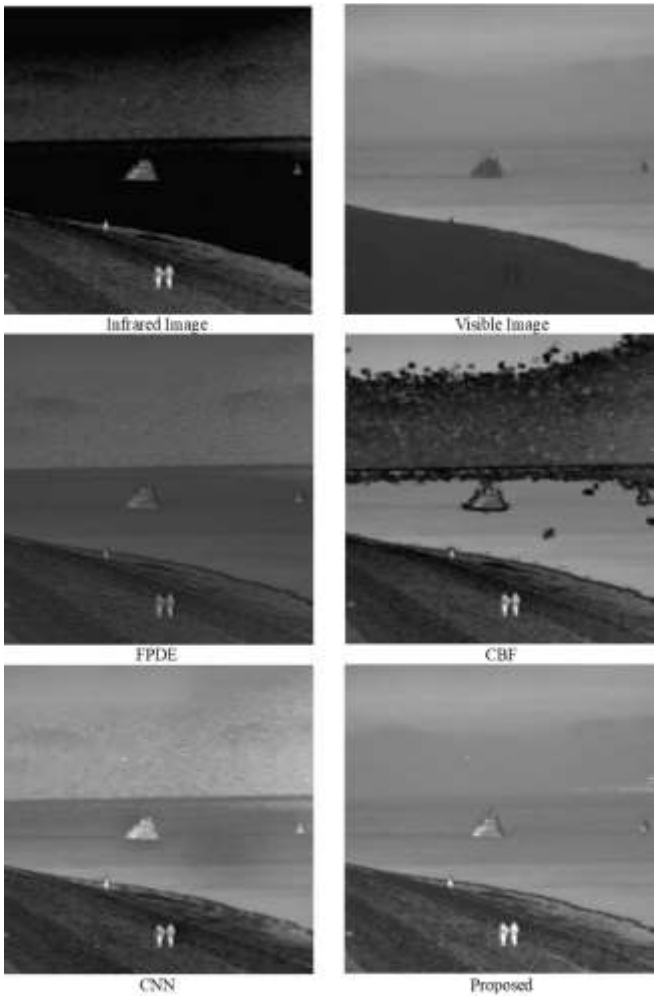

Fig. 3. Results of Infrared and Visible Images using Various Methods.

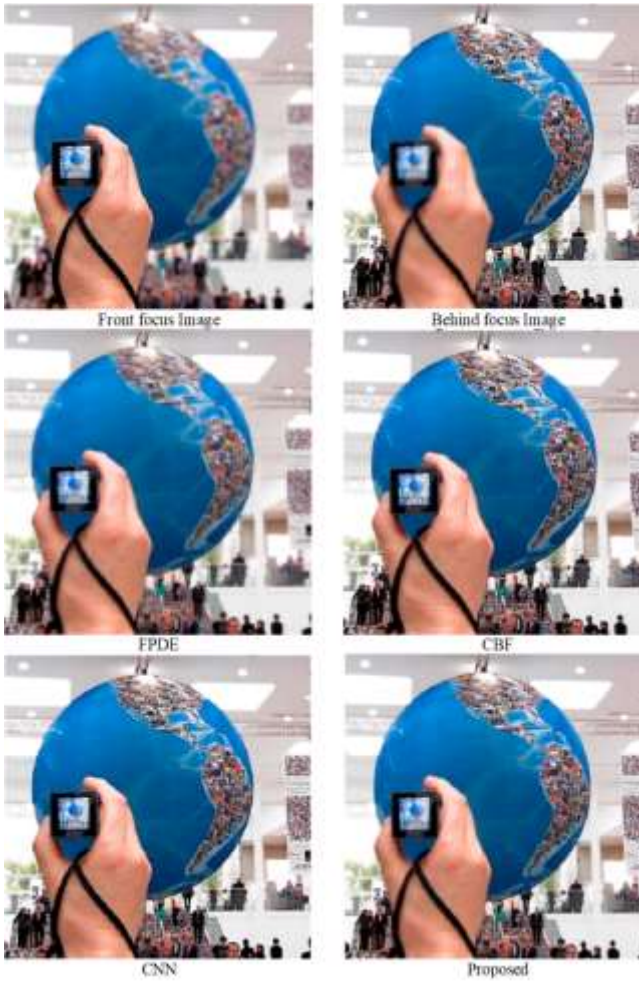

Fig. 3. Multifocus Fusion Results of Hand Camera and Globe Images using Different Fusion Methods.

D. Quntitative anlysis and Discusion

In the comparison of quality metric scores for different image fusion methods, the results reveal valuable insights into the performance of each method across various image types. The discussion below highlights the findings from Tables III, IV, and V, focusing on the metrics of Entropy (EN), Mutual Information (MI), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Root Mean Square Error (RMSE).

TABLE III. QUALITY METRIC SCORES OF MEDICAL IMAGES

Methods	Quality assessment matric score				
	EN	***MI***	***PSNR***	***SSIM***	***RMSE***
CBF	5.952	2.487	56.587	1.279	0.137
FPDE	6.412	2.132	56.005	1.023	0.203
CNN	6.889	2.485	56.002	1.198	0.204
Proposed	7.292	2.568	56.782	1.321	0.145

Medical Images (Table III):CBF scores moderately across all metrics, with PSNR at 56.587, indicating decent preservation of image fidelity. However, the higher RMSE suggests some residual errors. FPDE performs slightly better than CBF in terms of MI and PSNR, but the higher RMSE suggests more significant errors. CNN shows competitive scores across all metrics, with a balanced trade-off between PSNR, SSIM, and RMSE. The proposed method outperforms CBF and FPDE, achieving the highest scores in EN, MI, PSNR, and SSIM. The lower RMSE indicates better reconstruction accuracy.

TABLE IV. QUALITY METRIC SCORES OF INFRARED AND VISIBLE IMAGES

Methods	Quality assessment matric score				
	EN	***MI***	***PSNR***	***SSIM***	***RMSE***
CBF	6.784	2.756	57.025	1.232	0.134
FPDE	5.886	2.139	58.023	1.397	0.123
CNN	6.742	2.068	57.568	1.312	0.145
Proposed	7.123	2.145	58.214	1.375	0.094

Infrared and Visible Images (Table IV): exhibits reasonable scores, with a notable PSNR of 57.025, indicating good preservation of information. However, the higher RMSE suggests potential artifacts. FPDE shows a significant improvement in PSNR and SSIM, suggesting enhanced image quality compared to CBF.CNN performs competitively with FPDE, showcasing a balanced combination of PSNR and SSIM. The proposed method outperforms all others in EN, MI, and PSNR, suggesting superior preservation of information and higher image quality.

TABLE V. QUALITY METRIC SCORES OF MULTIFOCUS IMAGES

Methods	Quality assessment matric score				
	EN	***MI***	***PSNR***	***SSIM***	***RMSE***
CBF	7.712	5.294	63.428	1.637	0.031
FPDE	7.653	4.435	63.944	1.717	0.029
CNN	7.674	5.404	63.106	1.625	0.032
Proposed	7.719	4.698	63.802	1.669	0.028

Multi-Focus Images (Table V): CBF achieves high scores in EN and MI, but the lower PSNR and SSIM indicate potential loss of image quality. FPDE shows improvements in PSNR and SSIM, reflecting enhanced image quality. CNN achieves competitive scores, but the higher RMSE suggests some reconstruction errors. The proposed method consistently performs well across all metrics, achieving the highest scores in EN, MI, PSNR, SSIM, and the lowest RMSE, indicating superior performance in preserving image details and minimizing errors.

The proposed fusion method consistently demonstrates superior performance across all image types, outperforming or competitively matching existing methods in terms of quality metrics. The emphasis on achieving higher scores in EN, MI, PSNR, and SSIM, coupled with lower RMSE, suggests that the proposed method excels in preserving image information, maintaining fidelity, and minimizing reconstruction errors. These results highlight the effectiveness of the proposed method in diverse imaging

scenarios, making it a promising choice for various applications.

V. CONCLUSION AND FUTURE SCOPE

In conclusion, this research introduces an innovative automated multimodal and Multifocus fusion framework. Leveraging advanced deep learning techniques, including Partial Differential Equation (PDE) preprocessing and learnable convolutional pools, the proposed approach addresses the challenges posed by diverse medical modalities, such as MRI, CT, MRI-PET, visual, infrared, and multi-focus images. The combination of modality-specific preprocessing, modified convolutional layers, and adaptive pooling results in a robust model capable of intelligently fusing information from various sources, thereby enhancing the overall quality of multimodal medical images.

The experimental evaluations conducted underscore the effectiveness of the proposed method. The outcomes demonstrate a significant improvement in the generation of high-quality multimodal medical images. The enhanced imaging quality achieved through our framework holds substantial promise for advancing diagnostic accuracy and contributing to more informed clinical decision-making. This research contributes to the evolving field of medical imaging by providing a versatile and powerful tool that can be applied across a spectrum of modalities, ultimately benefiting healthcare professionals in their diagnostic endeavours.

As future directions, exploring real-world validations, optimizing computational efficiency, and adapting the framework for domain-specific applications present promising avenues for further research. The continuous refinement and adaptation of this automated multimodal fusion framework have the potential to significantly impact the landscape of medical imaging, offering valuable insights and improving patient outcomes.

REFERENCES

[1] G. Trivedi and Dr. R. Sanghvi, "Optimizing Image Fusion Using Modified Principal Component Analysis Algorithm and Adaptive Weighting Scheme," *International Journal of Advanced Networking and Applications*, vol. 15, no. 01, pp. 5769–5774, 2023, doi: 10.35444/ijana.2023.15103.

[2] G. Trivedi and Dr. R. Sanghvi, Hybrid Model For Infrared And Visible Image" *Annals of the Faculty of Engineering Hunedoara*, vol. 21, no. 03, pp. 1584-2665, Aug 2023, doi: https://www.proquest.com/scholarly-journals/hybrid-model-infrared-visible-image-fusion/docview/2867370935/se-2?accountid=144497.

[3] H. Xu and J. Ma, "EMFusion: An unsupervised enhanced medical image fusion network," *Information Fusion*, vol. 76, pp. 177–186, Dec. 2021, doi: 10.1016/j.inffus.2021.06.001.

[4] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. 2010.

[5] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, "Image fusion meets deep learning: A survey and perspective," *Information Fusion*, vol. 76, pp. 323–336, Dec. 2021, doi: 10.1016/j.inffus.2021.06.008.

[6] G. Xiao, D. P. Bavirisetti, G. Liu, and X. Zhang, "Image Fusion," 2020, doi: 10.1007/978-981-15-4867-3.

[7] U. Patil and U. Mudengudi, "Image fusion using hierarchical PCA.," *2011 International Conference on Image Information Processing*, Nov. 2011, doi: 10.1109/iciip.2011.6108966.

[8] P. I. Basheer, K. P. Prasad, A. D. Gupta, B. Pant, V. P. Vijavan, and D. Kapila, "Optimal Fusion Technique for Multi-Scale Remote Sensing Images Based on DWT and CNN," *2022 8th International Conference on Smart Structures and Systems (ICSSS)*, Apr. 2022, doi: 10.1109/icsss54381.2022.9782239.

[9] C. He, Q. Liu, H. Li, and H. Wang, "Multimodal medical image fusion based on IHS and PCA," *Procedia Engineering*, vol. 7, pp. 280–285, 2010, doi: 10.1016/j.proeng.2010.11.045

[10] F. Chen, Z. Guan, X. Yang, and W. Cui, "A novel remote sensing image fusion method based on independent component analysis," *International Journal of Remote Sensing*, vol. 32, no. 10, pp. 2745–2763, May 2011, doi: 10.1080/01431161003743207.

[11] Y. Liu, X. Chen, A. Liu, R. K. Ward, and Z. J. Wang, "Recent Advances in Sparse Representation Based Medical Image Fusion," *IEEE Instrumentation & Measurement Magazine*, vol. 24, no. 2, pp. 45–53, Apr. 2021, doi: 10.1109/mim.2021.9400960.

[12] S. S. Kumar and S. Muttan, "PCA-based image fusion," *SPIE Digital Library*, May 08, 2006. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/6233/62331T/PCA-based-image-fusion/10.1117/12.662373.

[13] Y. Liu, L. Wang, H. Li, and X. Chen, "Multi-focus image fusion with deep residual learning and focus property detection," *Information Fusion*, vol. 86–87, pp. 1–16, Oct. 2022, doi: 10.1016/j.inffus.2022.06.001.

[14] G. J. Trivedi and R. Sanghvi, "Medical Image Fusion Using CNN with Automated Pooling," *Indian Journal Of Science And Technology*, vol. 15, no. 42, pp. 2267–2274, Nov. 2022, doi: 10.17485/ijst/v15i42.1812.

[15] L. Li, C. Li, X. Lu, H. Wang, and D. Zhou, "Multi-focus image fusion with convolutional neural network based on Dempster-Shafer theory," *Optik*, vol. 272, p. 170223, Feb. 2023, doi: 10.1016/j.ijleo.2022.170223.

[16] F. Chen, Z. Guan, X. Yang, and W. Cui, "A novel remote sensing image fusion method based on independent component analysis," *International Journal of Remote Sensing*, vol. 32, no. 10, pp. 2745–2763, May 2011, doi: 10.1080/01431161003743207.

[17] Gargi J Trivedi and Rajesh Sanghvi, "Novel Approach to Multi-Modal Image Fusion using Modified Convolutional Layers," *Journal of Innovative Image Processing*, vol. 5, no. 3, p. 229, Sep. 2023, doi: 10.36548/jiip.2023.3.002.

[18] T. Sharma, S. Pathak, G. J. Trivedi, and R. Sanghvi, "Flow Modelling in Porous Medium Applying Numerical Techniques: A Comparative Analysis," December 2023, vol. 2, no. 2, pp. 288–304, Dec. 2023, doi: 10.36548/rrrj.2023.2.004.

[19] Q. Zhang, "Construction of Robot Computer Image Segmentation Model Based on Partial Differential Equation," *Journal of Sensors*, vol. 2022, pp. 1–8, Jul. 2022, doi: 10.1155/2022/6216423.

[20] G. Trivedi, R. Sanghvi, V. Shah, and J. Sharma, "On solution of non-instantaneous impulsive Hilfer fractional integro-dierential evolution system," *MATHEMATICA APPLICANDA*, vol. 51, no. 1, pp. 3–20, 2023, doi: 10.14708/ma.v51i1.7167.

[21] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, Jul. 1990, doi: 10.1109/34.56205.

[22] U. Sara, M. Akter, and M. S. Uddin, "Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study," *Journal of Computer and Communications*, vol. 07, no. 03, pp. 8–18, 2019, doi: 10.4236/jcc.2019.73002.

[23] B. K. Shreyamsha Kumar, "Image fusion based on pixel significance using cross bilateral filter," *Signal, Image and Video Processing*, vol. 9, no. 5, pp. 1193–1204, Oct. 2013, doi: 10.1007/s11760-013-0556-9.

[24] D. Sunderlin Shibu and S. Suja Priyadharsini. 2021. Multi scale decomposition based medical image fusion using convolutional neural network and sparse representation. *Biomedical Signal Processing and Control* 69: 102789. https://doi.org/10.1016/j.bspc.2021.102789

[25] "The Whole Brain Atlas," *The Whole Brain Atlas*. https://www.med.harvard.edu/aanlib/home.html

[26] A. Toet, "The TNO Multiband Image Data Collection," *Data in Brief*, vol. 15, pp. 249–251, Dec. 2017, doi: 10.1016/j.dib.2017.09.038. [1]

[27] RAWSAMPLES.CH - Home," *RAWSAMPLES.CH - Home*. http://rawsamples.ch/index.php/en/

[28] B. K. Shreyamsha Kumar, "Image fusion based on pixel significance using cross bilateral filter," Signal, Image and Video Processing, vol. 9, no. 5, pp. 1193–1204, Oct. 2013, doi: 10.1007/s11760-013-0556-9.

[29] [2] Y. Liu, X. Chen, J. Cheng, H. Peng, and Z. Wang, "Infrared and visible image fusion with convolutional neural networks," International Journal of Wavelets, Multiresolution and Information Processing, vol. 16, no. 03, p. 1850018, May 2018, doi: 10.1142/s0219691318500182

[30] [3] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," 2017 20th International Conference on Information Fusion (Fusion), Jul. 2017, doi: 10.23919/icif.2017.8009719.

[31] [4] G. Trivedi And R. Sanghavi, "Fusesharp: A Multi-Image Focus Fusion Method Using Discrete Wavelet Transform And Unsharp Masking," Journal Of Applied Mathematics & Informatics, Vol. 41, No. 5, Pp. 1115–1128, Sep. 2023, doi: https://doi.org/10.14317/jami.2023.1115

[32] [5] G. Trivedi And R. Sanghavi, "Mosaicfusion: Merging Modalities With Partial Differential Equation And Discrete Cosine Transformation," Journal Of Applied And Pure Mathematics, Vol. 5, No. 5_6, Pp. 389–406, Nov. 2023,doi: https://doi.org/10.23091/japm.2023.389

AUTHOR Profile

Gargi Trivedi is currently working as a Teaching Assistant in Department of Applied Mathematics, Faculty of Engineering and Technology, The Maharaja Sayajirao University of Baroda. She is also pursuing her PhD in Mathematics from CVM university, V V Nagar. She earned her M.Sc. (Mathematics) ,B.Ed. and PGDCA degree fromBhavnagar University, Bhavnagar, Gujarat. She has published more than 6 papers in national and international journals and conferences. She has more than 12 years of teaching experience. Her major research areas are machine learning and medical image analysis

Dr. Rajesh C. Sanghvi holds a Ph.D. degree from Charotar University of Science and Technology. He earned his B.Sc. and M.Sc. degrees in Mathematics from Sardar Patel University, V V Nagar, Gujarat. He also holds a B.Sc. degree in Computer Science from Sardar Patel University, V V Nagar, Gujarat. Since 1998, he has worked as an assistant professor at G. H. Patel College of Engineering and Technology, V V Nagar. His scientific interests include image processing, machine learning, fuzzy logic, optimization,evolutionary algorithms.