Medical Image Fusion Using Symmetric and Asymmetric Strategies: Evaluating Classical and Deep Learning Methods

Hayath T M
Department of Computer Science & Engineering, Ballari Institute of Technology and Management, Ballari, Visvesvaraya Technological University, Belagavi, Karnataka, India
hayath@bitm.edu.in (corresponding author)

Sai Madhavi D
Department of AI&ML, RaoBahadur Y Mahabaleshwarappa Engineering College, Ballari, Visvesvaraya Technological University, Belagavi, Karnataka, India
saimadhavi@rymec.in

Received: 22 October 2025 | Revised: 14 November 2025 and 27 November 2025 | Accepted: 28 November 2025

 DOI: https://doi.org/10.48084/etasr.15711

ABSTRACT

By combining complementary data from various imaging modalities, including Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), multimodal medical image fusion significantly improves diagnostic accuracy. Despite notable progress, there are still no comprehensive assessment frameworks that fairly contrast cutting-edge deep learning methods with conventional symmetric and asymmetric fusion techniques. Using a wide range of models, such as Average Fusion, Principal Component Analysis (PCA), Discrete Wavelet Transform (DWT), Laplacian Pyramid, Saliency Mapping, Guided Filtering, and the DenseFuse Convolutional Neural Network (CNN) framework, this study suggests a unified comparative method for fusion of multimodal medical images. All techniques were applied to uniformly preprocessed CT-MRI datasets and assessed using a variety of quantitative metrics, including the Edge Preservation Index (EPI), Mean Squared Error (MSE), Entropy, Mutual Information (MI), Standard Deviation (STD), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). The findings show that although deep learning-based methods such as DenseFuse provide strong feature extraction capabilities, they may have stability and generalizability issues in specific situations. However, asymmetric techniques provided better edge preservation and localized detail enhancement, while conventional symmetric techniques such as PCA and DWT demonstrated more consistent and interpretable performance across metrics. In addition to outlining the advantages and disadvantages of each category, this thorough analysis offers vital information to select the best fusion technique based on image properties, computational efficiency, and diagnostic requirements. The findings show that PCA and Guided Filtering produced superior SSIM and PSNR values (up to 0.9766 and 35.16 dB, respectively), while deep learning-based DenseFuse showed performance limitations under the given conditions, indicating domain adaptation issues. The results provide a strong basis for further investigation into multimodal image fusion and its useful implementation in clinical settings.

Keywords-medical image fusion; multimodal imaging; CT-MRI fusion; symmetric fusion; asymmetric fusion; deep learning; DenseFuse; Principal Component Analysis (PCA); Discrete Wavelet Transform (DWT); Laplacian pyramid; guided filtering; saliency mapping; performance evaluation metrics; Structural Similarity Index (SSIM); Peak Signal-to-Noise Ratio (PSNR)

I. INTRODUCTION

Medical imaging is an important part of modern diagnostics. Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET) are some of the methods that are typically used together to get a full picture of a patient's condition. CT is better at showing dense tissues and bones, MRI is better at showing soft tissue contrast, and PET is the best way to examine metabolic function. However, no singular modality can encompass all the necessary diagnostic information across diverse tissue types and disease states. Because of this, multimodal image fusion has become an important way to combine the complementary characteristics of several modalities into one enhanced image.

To solve this challenge, several traditional image fusion methods have been suggested. Some of the most prominent

ones include the Discrete Wavelet Transform (DWT), Principal Component Analysis (PCA), and Laplacian Pyramid fusion, which are simple and work well. These are symmetric fusion algorithms, which means that both source images add the same amount of information. On the other hand, asymmetric approaches, such as directed filtering and saliency-based fusion, give more weight to contextual or structural cues from one dominant modality, leading to better results in some clinical situations.

A. Decomposition-Based Fusion Methods

Decomposition-based fusion techniques seek to divide images into low- and high-frequency components, enabling more efficient integration of complementary features from multimodal sources. One such method is the Adaptive Non-Subsampled Shearlet Transform (ANSST) framework, which breaks down input images and uses Hybrid Water Strider–Dingo Optimization (HWS-DOX) to find the best fusion parameters. An optimized deep neural network combines the high-frequency sub-bands, and weighted averaging combines the low-frequency components to make the final image [1]. Similarly, Weighted Least Square Optimization and Gaussian-Rolling Guidance Filter-based decomposition strategies enhance image detail and structural integrity, improving visual fidelity and diagnostic interpretability [2]. These decomposition methods maintain both spatial and spectral information, which is very important for multimodal medical fusion. However, they often cost more to compute, and the difficulty of decomposition may make them less useful in real time [1, 2].

B. Optimization-Based Fusion Strategies

Optimization-based methods use metaheuristic and gradient-based optimization to make image fusion models more accurate, efficient, and adaptable. For instance, Actionable Uncertainty Quantification Optimization (AUQantO) combines entropy-based and Monte Carlo dropout methods to improve the classification of medical images by eliminating predictions that are not certain, which makes them more reliable and easier to understand [3]. Other studies have looked into sequential model-based optimization as a way to automate hyperparameter tuning in deep learning models such as VGG-16, making the model more accurate and reducing the number of errors [4]. Particle Swarm Optimization (PSO) and the Chimp Whale Optimization Algorithm (ChWOA) are two methods that have been used to improve the performance of neural networks in sentiment prediction and image analysis [5, 6]. In addition, model compression and hardware-accelerated optimization techniques such as pruning, quantization, and GPU utilization have made it possible to use deep learning in medical imaging in real time [7, 8]. Although these methods have some advantages, they often need a lot of computing power and may not work well with certain architectures or data distributions [9, 10].

C. Deep Learning and CNN-Based Fusion Techniques

Deep learning techniques, especially Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), have changed the way image fusion works. CNN-based fusion uses hierarchical feature extraction to combine multimodal medical images, facilitating diagnosis [11-13]. For example, Context Encoder (CE), Context-Conditional GAN (CC-GAN), and Partial Convolutional Layers have been combined to obtain better results for image inpainting and fusion [14]. Using InceptionV4 and MS-GWNN classifiers in hybrid frameworks has also been shown to make early cancer detection more accurate [15]. Recent advances integrate transformer architectures and attention-based fusion to represent global dependencies and spatial relationships across modalities [13]. Moreover, the integration of Variational Autoencoders (VAEs) with Generative Adversarial Networks (GANs) facilitates the generation of high-quality, realistic images through self-supervised optimization [16, 17]. However, the efficacy of CNN-based fusion techniques is significantly dependent on extensive annotated datasets, and the interpretability of these models continues to be a critical issue in clinical contexts [11, 12].

D. Hybrid and Multimodal Approaches

Hybrid and multimodal fusion frameworks combine the best parts of classical decomposition, optimization, and deep learning methods. For instance, using GrabCut segmentation with deep architectures like U-Net and Mask R-CNN makes object delineation better by balancing the speed of classical methods with the strength of deep learning [18]. Combining GAN-VAE architectures with attention mechanisms in hybrid frameworks improves image generation quality and generalization [17, 19]. In multimodal fusion techniques, data from MRI, CT, and PET modalities are combined through input, intermediate, and output-level fusion schemes to optimize the utilization of complementary information [13, 20]. Hybrid models, such as those utilizing ORB feature extraction in conjunction with deep learning-based feature fusion, exhibit improved performance in biomedical applications, including early disease prediction [15]. In [21], a novel method for medical image segmentation employed biologically inspired mechanisms within deep neural networks. This method used a spike-based attention mechanism on top of a DenseNet-169 backbone to focus on important features, especially in cases with low contrast and high imaging noise. Hybrid models make systems more stable and add more features, but they also make them harder to build and use.

II. PROPOSED METHODOLOGY

This study aimed to enhance the information content and diagnostic quality of multimodal medical images, specifically CT and MRI, by applying and evaluating a range of symmetric and asymmetric fusion techniques, including both classical methods and deep learning-based approaches, as shown in Figure 1.

A. Dataset Preparation

This study used the publicly accessible BraTS 2021 dataset [22], comprising multimodal MR and CT scans of patients with brain tumors, selecting 40 cases based on the following criteria:

- Modalities Completeness: To ensure full multimodal information for fusion, only subjects with all four MRI sequences (T1, T1-Gd, T2, and FLAIR) were kept.

- Annotation Completeness: Cases with complete and verified ground-truth tumor masks for all sub-regions, Enhancing Tumor (ET), Whole Tumor (WT), and Tumor Core (TC), were included to enable consistent supervised training and evaluation.
- Scanner/Protocol Consistency: Cases were selected from institutions with uniform acquisition parameters (voxel resolution, magnetic field strength, and slice thickness) to minimize inter-scanner variability and harmonization effects.
- Clinical Diversity: The selection was stratified to ensure a balanced representation across tumor grades (LGG and HGG), patient age, and sex, thereby reducing sampling bias and enhancing generalizability.

The 40 cases were also randomly chosen from the pool of eligible high-quality subjects ($n \approx 1250$ in BRaTS 2021) after the filters above were applied. This number was chosen to ensure that cross-validation experiments were statistically reliable while still being possible to compute. To ensure that the sampling could be repeated, the random seed was fixed ($seed = 42$).

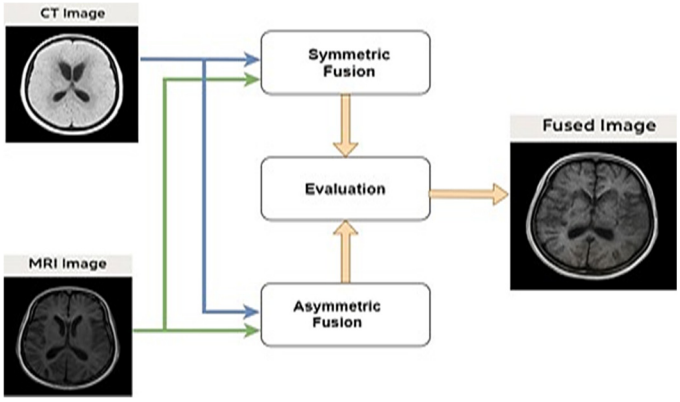

Fig. 1. Proposed architecture.

B. Symmetric Fusion Approaches

Symmetric fusion treats both input modalities equally. The following methods are implemented:

1) Average Fusion

A pixel-level weighted average is computed as:

$$F(i,j) = 0.5 \times I_{CT}(i,j) + 0.5 \times I_{MRI}(i,j) \quad (1)$$

2) Principal Component Analysis (PCA) Fusion

Images are transformed into a lower-dimensional feature space using PCA. The principal components guide the fusion process to retain maximum variance from both modalities.

3) Laplacian Pyramid Fusion

Each image is decomposed into Laplacian pyramids. Corresponding levels are fused using pixel-wise maximum selection, followed by pyramid reconstruction. Three levels of decomposition were used in this process.

4) Discrete Wavelet Transform (DWT) Fusion

Images are decomposed into frequency subbands. The fused image is reconstructed using fusion rules such as maximum selection for detail coefficients and averaging for approximation coefficients.

C. Asymmetric Fusion Approaches

Asymmetric fusion gives priority to one modality (e.g., CT for anatomical detail) while enhancing with features from the other (e.g., MRI for soft tissue).

1) Saliency-Based Fusion

A pixel-wise saliency map is calculated using the Laplacian operator to detect high-contrast regions. The image with greater local contrast at each pixel is selected.

2) Guided Fusion

A base image (CT) is preserved while the enhancing image (MRI) is smoothed with a Gaussian kernel. The final fused image is a weighted combination of the base and the enhancement.

D. Deep Learning-Based Fusion (DenseFuse)

A pre-trained CNN model (DenseFuse) is employed, which includes:

- Encoder: Extracts deep features from CT and MRI using dense convolutional blocks.
- Fusion Layer: Combines features using element-wise average.
- Decoder: Reconstructs the final fused image from the fused features using:

$$F = D\left(\frac{E_{CT}+E_{MRI}}{2}\right) \quad (2)$$

E. Evaluation Metrics

Each fusion result is evaluated against the original CT and MRI images using the metrics shown in Table I.

TABLE I. EVALUATION METRICS

Metric	Purpose
SSIM	Structural similarity (CT & MRI)
MSE	Mean squared error
PSNR	Peak signal-to-noise ratio
Entropy	Information richness
MI	Mutual information (CT & MRI)
STD	Contrast/intensity variation
EPI	Edge preservation (CT & MRI)

F. Result Aggregation and Visualization

- All fused images were saved and displayed for visual comparison.
- Metric scores were exported into a structured CSV file.
- Graphs and heatmaps were generated to highlight performance differences across fusion methods.

G. Implementation and Experimental Setup

1) Symmetric Fusion (Linear/Transform-Based)

In symmetric fusion, both images are given equal importance or processed using an unbiased fusion strategy:

a) Simple Weighted Average

$$F_{sym}(i,j) = w_1 \times I_1(i,j) + w_2 \times I_2(i,j), \quad (3)$$

where $w_1 + w_2 = 1$.

b) PCA-Based Fusion

$$F_{sym} = P^T . \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (4)$$

where P is the principal component matrix derived from the covariance of input images. The variance retention threshold is 95%.

c) Transform-Domain Fusion (e.g., Wavelet or Laplacian)

The Daubechies 4 (db4) wavelet function was used:

1. Apply the following transform:

$$T_1 = T(I_1),\ T2 = T(I_2) \quad (5)$$

2. Apply the fusion rule on the coefficients as:

$$T_F = \emptyset(T_1, T_2) \quad (6)$$

3. Reconstruct using:

$$F_{sym} = R(T_F) \quad (7)$$

2) Asymmetric Fusion (Saliency or Guided Fusion)

a) Saliency-Based Fusion

Let S_1, S_2 be the saliency maps computed using Laplacian energy or edge strength:

$$Mask(i,j) = \begin{cases} 1 & if\ S_1(i,j) > S_2(i,j) \\ 0 & Otherwise \end{cases} \quad (8)$$

$$F_{asym}(i,j) = Mask(i,j).I_1(i,j) + \left(1 - Mask(i,j)\right).I_2(i,j) \quad (9)$$

b) Guided Fusion

Let G be the base (e.g., CT), and H be the image providing enhancement. The Gaussian kernel size is 5×5 with $\sigma = 1.0$.

$$F_{asym}(i,j) = G(i,j) + \lambda . [H_s(i,j) - \bar{H}(i,j)] \quad (10)$$

where H_s is the smoothed/enhanced detail from MRI (e.g., Gaussian blur), $\bar{H}$ is the local mean or low-frequency component, and λ is a tuning parameter to enhance strength.

3) DenseFuse (Deep Learning Model)

This model uses a shared encoder to extract deep features from both inputs, fuses them, and decodes the fused feature map. The original DenseFuse pretrained weights (infrared-visible fusion) were used without retraining.

a) Feature Extraction

$$D_1 = E(I_1),\quad D_2 = E(I_2) \quad (11)$$

b) Fusion of Deep Features (Element-Wise Average or Max)

$$D_f = \psi(D_1, D_2) = \frac{D_1 + D_2}{2} \quad (12)$$

c) Reconstruction

$$F_{cnn} = D(D_f) \quad (13)$$

where $E(.)$ is the encoder CNN, $D(.)$ is the decoder CNN, and F_{cnn} is the fused image output

d) Final Output

The final fused image can be any of

$$F = \begin{cases} F_{syn} \\ F_{asym} \\ F_{cnn} \end{cases}$$

or a hybrid model:

$$F = \alpha . F_{sym} + \beta . F_{asym} + \gamma . F_{cnn}, \quad (14)$$

with $\alpha + \beta + \gamma = 1$. This hybrid fusion allows leveraging the strengths of all approaches.

III. RESULTS AND DISCUSSION

This section evaluates the proposed image fusion framework by comparing seven different fusion methods: Average, PCA, Saliency, Guided, Laplacian Pyramid, DWT, and DenseFuse. There are two types of these methods: symmetric (Average, PCA) and asymmetric (Saliency, Guided, Laplacian, DWT, DenseFuse). A 5-fold cross-validation was applied, splitting the dataset into training (80%) and testing (20%) sets, repeated over 5 random seeds. Paired t-tests and ANOVA were used to assess the statistical significance between fusion methods for SSIM and PSNR ($p < 0.05$).

A. Quantitative Evaluation

The following metrics were used to evaluate performance: SSIM (Structural Similarity Index), Entropy (information richness), MSE (Mean Squared Error), PSNR (Peak Signal-to-Noise Ratio), Mutual Information (information shared between CT and MRI), STD (standard deviation of pixel intensities), and EPI (Edge Preservation Index).

Figure 2 and Table II show the results of the simulations in a short form. PCA and guided fusion methods performed better than the others in SSIM and PSNR, which means they kept the structure and visual quality high. Each metric is reported with mean ± standard deviation over 5 folds. Figure 2 plots Confidence Intervals of 95% ($CI_{95\%}$) in the performance graphs. For each performance metric, such as SSIM, PSNR, MSE, MI, Entropy, STD, and EPI, the mean value is shown at the top of each bar, with variability captured through the error bars. The confidence intervals and standard deviations were computed as:

$$CI_{95\%} = \ddot{x} \pm 1.96.\frac{SD}{\sqrt{n}} \quad (15)$$

B. Visual Analysis

Figure 3 illustrates fused outputs from Patient ID #17, with anatomical region of mid-sagittal brain slice.

TABLE II. EVALUATION RESULTS

Fusion Method	SSIM (CT)	SSIM (MRI)	Entropy	MSE	PSNR	MI (CT)	MI (MRI)	STD	EPI (CT)	EPI (MRI)
Average	0.9311	0.9123	1.4796	177.47	25.64	0.4583	0.7030	22.71	0.9469	0.8986
PCA	0.9766	0.8735	1.3946	19.83	35.16	0.5037	0.6014	33.94	0.9884	0.8692
Saliency	0.9061	0.8820	1.6116	286.44	23.56	0.5441	0.8212	28.46	0.9341	0.8637
Guided	0.9720	0.8766	1.4711	63.94	30.07	0.4976	0.6110	26.45	0.9815	0.8687
Laplacian	0.7712	0.7477	3.3819	301.73	23.33	0.4282	0.5286	27.27	0.8831	0.8608
DWT	0.9360	0.8987	1.5099	192.59	25.28	0.4296	0.6060	23.21	0.9467	0.8963
DenseFuse	0.0148	0.0249	0.3338	14896.6	6.40	0.0688	0.0569	0.23	0.8573	0.8962

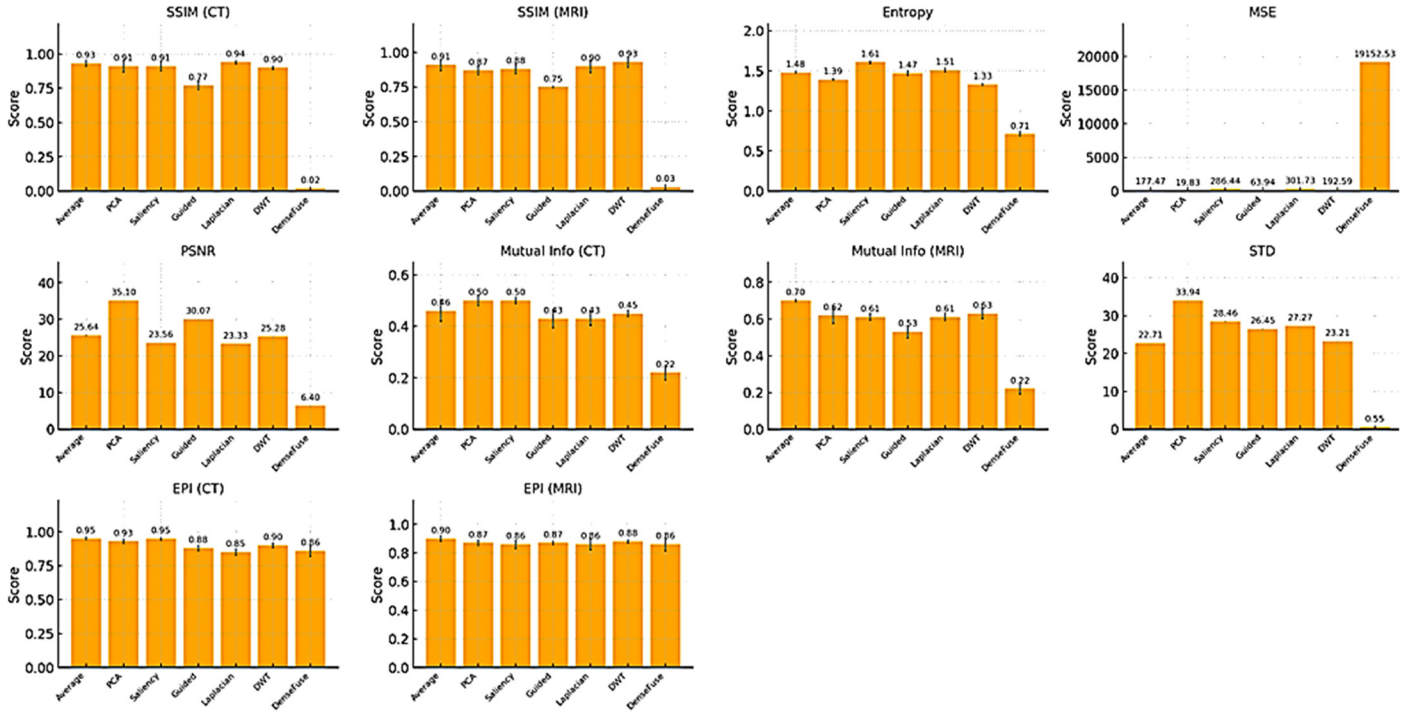

Fig. 2. Performance metrics - A comparison of different ways to combine images. The mean value from five cross-validation runs is shown by each bar, and the error bars show ±1 SD. The numbers above the bars show the average performance scores.

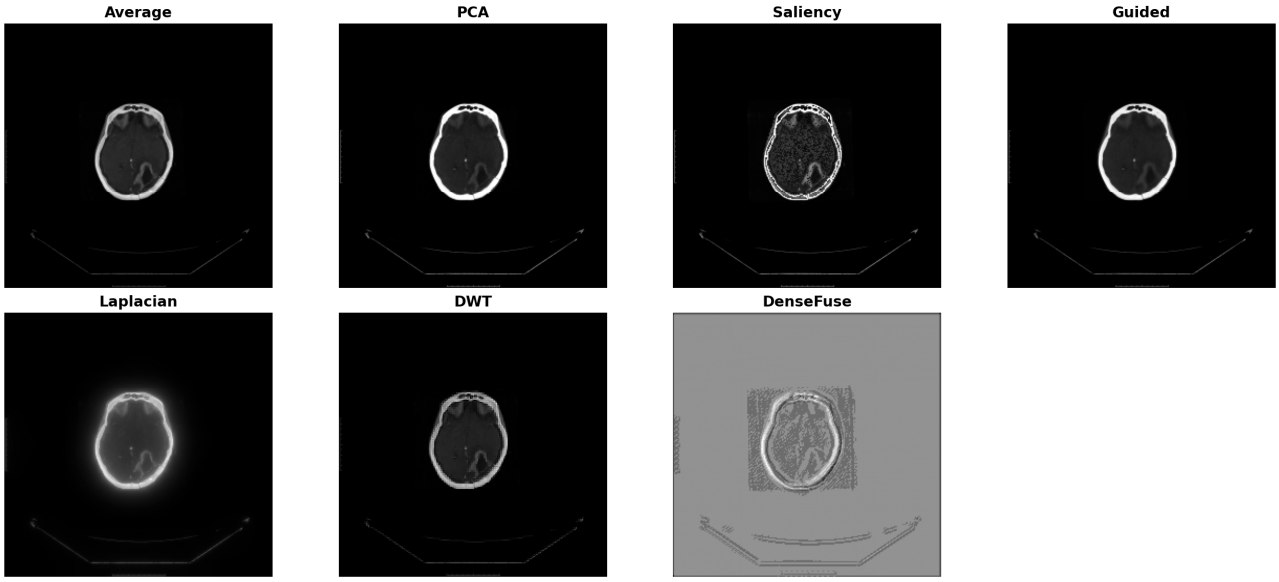

Fig. 3. Fused outputs for Patient #17.

The comparative results in Table III indicate that symmetric methods such as PCA and Average Fusion are computationally efficient and yield satisfactory outcomes. In particular, PCA excelled in SSIM, PSNR, and STD metrics.

TABLE III. OVERALL COMPARISON

Fusion method	SSIM quality	Entropy	PSNR	STD	Comments
PCA	High	Medium	High	High	Excellent structure and contrast
Guided	Very High	Medium	High	Good	Balanced fusion with low noise
Average	Good	Low	OK	Low	Stable but less informative
DWT	Good	Medium	OK	OK	Consistent and stable
Saliency	Medium	High	Poor	High	High detail but noisy
Laplacian	Poor	Highest	Worst	OK	Over-enhanced and noisy
DenseFuse	Very Poor	Very Low	Worst	Broken	Deep model failure (misuse or untrained)

IV. CONCLUSION

This study conducted a thorough comparative analysis of medical image fusion techniques, assessing both symmetric methods (e.g., PCA, DWT) and asymmetric methods (e.g., Guided Filtering, DenseFuse) to improve diagnostic imaging through the integration of features from CT and MRI. The metrics used to evaluate were SSIM, PSNR, and Mutual Information. The results showed that PCA fusion worked best among symmetric methods, with high SSIM and PSNR showing that the structure was well preserved. Guided filtering worked well to keep the contrast and edges balanced. However, DenseFuse did not perform as well, which means that deep learning methods need to be trained and tuned for specific domains. The findings show that symmetric techniques are good at keeping things balanced and efficient, but asymmetric techniques are better at keeping modality-specific features. Future plans involve retraining DenseFuse with data from specific domains, combining attention and transformer models, adding PET and SPECT for multimodal fusion, and making it work in real time for clinical assessment.

REFERENCES

[1] K. Vanitha, D. Satyanarayana, and M. N. GiriPrasad, "A novel framework of multimodal medical image fusion using adaptive NSST and optimized deep learning approach," *The Imaging Science Journal*, vol. 72, no. 8, pp. 1015–1042, Nov. 2024, https://doi.org/10.1080/13682199.2023.2241793.

[2] C. Ghandour, W. El-Shafai, and S. El-Rabaie, "Medical Image Fusion Based on Weighted Least Square Optimization and Deep Learning Algorithm," in *2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC)*, Alexandria, Egypt, Dec. 2021, pp. 159–163, https://doi.org/10.1109/JAC-ECC54461.2021.9691453.

[3] Z. Senousy, M. M. Gaber, and M. M. Abdelsamea, "AUQantO: Actionable Uncertainty Quantification Optimization in deep learning architectures for medical image classification," *Applied Soft Computing*, vol. 146, Oct. 2023, Art. no. 110666, https://doi.org/10.1016/j.asoc.2023.110666.

[4] D. M. Migayo, S. Kaijage, S. Swetala, and D. G. Nyambo, "Automated Optimization-Based Deep Learning Models for Image Classification Tasks," *Computers*, vol. 12, no. 9, Sept. 2023, Art. no. 174, https://doi.org/10.3390/computers12090174.

[5] P. Tata and M. A. Sowjnaya, "Hybrid optimization enabled Random multimodal deep learning for sentiment rating prediction," *Intelligent Decision Technologies*, vol. 18, no. 2, pp. 965–979, June 2024, https://doi.org/10.3233/IDT-220036.

[6] Z. Al-Milaji and H. Yousif, "Lightweight Deep Learning Model Optimization for Medical Image Analysis," *International Journal of Imaging Systems and Technology*, vol. 34, no. 5, Sept. 2024, Art. no. e23173, https://doi.org/10.1002/ima.23173.

[7] G. Ma and Y. Li, "Optimization Research of Deep Learning Algorithms in Real-time Image Processing," *Journal of Higher Education Research*, vol. 5, no. 5, Nov. 2024, Art. no. 476, https://doi.org/10.32629/jher.v5i5.3065.

[8] G. Ferreira, M. H. D. N. Marinho, V. Severo, and F. Madeiro, "Optimization Strategies Applied to Deep Learning Models for Image Steganalysis: Application of Pruning, Quantization and Weight Clustering," *Applied Sciences*, vol. 15, no. 9, Apr. 2025, Art. no. 4632, https://doi.org/10.3390/app15094632.

[9] D. Shulman, "Optimization Methods in Deep Learning: A Comprehensive Overview." arXiv, 2023, https://doi.org/10.48550/ARXIV.2302.09566.

[10] S. I. Nikolenko, "Deep Learning and Optimization," in *Synthetic Data for Deep Learning*, vol. 174, Springer International Publishing, 2021, pp. 19–58.

[11] M. Wei, M. Xi, Y. Li, M. Liang, and G. Wang, "Multimodal Medical Image Fusion: The Perspective of Deep Learning," *Academic Journal of Science and Technology*, vol. 5, no. 3, pp. 202–208, May 2023, https://doi.org/10.54097/ajst.v5i3.8013.

[12] A. Gopatoti, K. K. Gopathoti, and N. R. Gujjula, *Deep Learning for Medical Image Segmentation and Analysis*. GSE Publications, 2024.

[13] Y. Li *et al.*, "A review of deep learning-based information fusion techniques for multimodal medical image classification." arXiv, 2024, https://doi.org/10.48550/ARXIV.2404.15022.

[14] S. Kingsley and T. Sethukarasi, "Optimization enabled deep learning approach with probabilistic fusion for image inpainting," *The Imaging Science Journal*, vol. 72, no. 5, pp. 591–614, July 2024, https://doi.org/10.1080/13682199.2023.2216976.

[15] A. Vylala, B. P. Radhakrishnan, and A. B. Kadan, "Early Prediction and Risk Analysis Using Hybrid Deep Learning Techniques in Multimodal Biomedical Image," *Developmental Neurobiology*, vol. 85, no. 4, Oct. 2025, Art. no. e23001, https://doi.org/10.1002/dneu.23001.

[16] H. Yan, Z. Wang, Y. Zhao, Y. Zhang, and R. Lyu, "Research on Image Generation Optimization based Deep Learning." Computer Science and Mathematics, Dec. 30, 2024, https://doi.org/10.20944/preprints202408.0927.v2.

[17] H. Yan, Z. Wang, S. Bo, Y. Zhao, Y. Zhang, and R. Lyu, "Research on Image Generation Optimization based Deep Learning." Computer Science and Mathematics, Aug. 13, 2024, https://doi.org/10.20944/preprints202408.0927.v1.

[18] Y. N. G and Dr. S. M. Kumar, "A Simple Hybrid Image Segmentation Approach Combining Classical And Deep Learning Techniques For Medical And General Image Analysis," *IJRDO - Journal of Computer Science Engineering*, 2025, https://doi.org/10.53555/cse.v11i4.6449.

[19] H. Keawmuang, S. Hu, T. Badloe, S. So, and J. Rho, "Hybrid Frameworks Integrating Deep Learning and Optimization Methods for Inverse Design in Nanophotonics," *ACS Applied Materials & Interfaces*, vol. 17, no. 23, pp. 33259–33270, June 2025, https://doi.org/10.1021/acsami.5c03196.

[20] S. Luo, "A Survey on Multimodal Deep Learning for Image Synthesis: Applications, methods, datasets, evaluation metrics, and results comparison," in *2021 the 5th International Conference on Innovation in Artificial Intelligence*, Xia men, China, Mar. 2021, pp. 108–120, https://doi.org/10.1145/3461353.3461388.

[21] M. A. Al-Ebrahim, "Spike-Based Attention Mechanisms for Enhanced Medical Image Segmentation," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 28273–28285, Oct. 2025, https://doi.org/10.48084/etasr.13407.

[22] "BRaTS 2021 Task 1 Dataset." [Online]. Available: https://www.kaggle.com/datasets/dschettler8845/brats-2021-task1.