"Bodh", BPIT's International Journal of Technology & Management
BBIJTM, ISSN: 2454-8421, Volume 11,

Medical Image Fusion with Different Image Fusion Rules and denoisation with Custom CNN

[1]Sumit Guin, [2]Sumit Kumar

[1,2]Department of Electrical Engineering, Delhi Technological University,
Bawana Road, Bawana, Delhi-42, India
*Sumitguin_23ee256@dtu.ac.in

Abstract. Multimodal medical image fusion is the process of combining information from different modality medical images into a single image for better diagnostic accuracy and clinical decision-making. Using multimodal medical images, a comparative study of image fusion rules is carried out in this study. Initially images obtained from two different modalities are decomposed into multiple sub bands using the Discrete Wavelet Transform, which enables the separation of approximation and detail components. Various existing fusion rules are then applied to combine the corresponding wavelet coefficients from the source images. In order to reconstruct the final fused image, the inverse transformation is carried out following fusion. The performance of the fusion rules is systematically evaluated using both objective measures and visual assessment to determine their effectiveness in preserving relevant information and structural details. After fusion the fused image is denoised by the custom denoising CNN and the performance of the denoised image is measured with different metrics. This study demonstrates that, despite the computational efficiency of simpler fusion rules, more adaptive approaches typically produce better visual and quantitative results. The findings can assist in selecting suitable methods for multimodal medical image fusion applications and provide a clear understanding of the advantages and disadvantages of various fusion strategies.

Keywords: Multimodal medical image fusion, Fusion rule, Discrete wavelet transform,

1 Introduction

In medical imaging, multimodal image fusion becomes very vital for easy and perfect diagnosis. We all know that CT (Computed tomography) images carry the hard tissue information of human organs. Whereas MRI (Magnetic resonance Imaging) carries soft tissue information of human organs, so if we fuse both the images, the fused image can carry hard as well as soft tissue information which can solve the storage as well as time of inspection two different modality images. With the spreading of image processing, image fusion has been a significant subject in several associated areas such as computer vision, remote sensing, object detection, medical imaging and image classification. Many image fusion methods are already existing in literature. A detailed review of some recent works in this field is reported in table 1.

Table 1. Detail review of multimodal medical image fusion

Ref. No.	Reference with year	Methodology	Application domain	Advantages	Limitations
[1]	Xin et al. 2025	Deep Generative models(GAN based fusion)	CT+MRI fusion	Good structural and functional details preservation	Requires high computational resources
[2]	Hussain, M.et al. 2024	Deep learning based fusion survey	Multimodal medical images	Covers wide range of models	No experimenrtal validation
[3]	Hussain, S.S.et al. 2025	Swin Transformer +CNN fusion	MRI(Parkinson's Detection)	Good performance	-
[4]	Sinha et al. 2024	Improved Dual-channel PCNN	Multimodal medical images	Good edge and detail preservation	Parameter tuning complexity
[5]	Wankhede, P et al. 2023	Laplacian Auto encoder +Attention	CT+MRI fusion	Preserves fine details	Computationally intensive
[6]	Xiang, H et al. 2025	Semantic preserving Multimodal fusion	MRI,CT	Clinically meaningful fusion	Needs annotate semantic data
[7]	Cao et al. 2026	Deep learning based image fusion	CT, MRI	Better spatial alignment	System complexity
[8]	Singh et al. 2025	Aformer (Attention based-Transformer)+ Segmentation aware fusion	Multi modality medical images	Strong Boundary detection	Task specific
[9]	Guo et al. 2025	I scale + Spatial adaptive fusion	Multi modality medical images	D Local-Global balance	Less effective than GAN models
[10]	Liu et al. 2006	DWT + PCA based fusion	CT+MRI	Fast, simple	Image blurring

After a detailed literature review, it has been cleared that, though many image fusion models are there in literature, generating a model that can combine the detail information of source images into the fused image without any information loss is still a big challenge. Fusing of two different modality images are very challenging as pixels are carrying complementary information. So, in fusion loss the information of both the pixels may occur. Many image fusion rules are existing in literature, for example, Mean, Max, Min, etc. [2]. These image fusion rules have applied successfully individually. In this paper we have combined them in pairs and analysis their effects in medical image fusion.

3 Methodology

The flow diagram of the model is depicted in the figure 1. The individual functioning of each block is described below:

Insertion of different modality medical images: The CT and MRI images have been inserted from the Harvard Medical School Brain dataset.

Discrete Wavelet Transform: The input images are decomposed into four coefficients (LL, LH, HL and HH). The (LL) coefficient is called the approximate coefficient, whereas rest of the coefficients are called the detail coefficients.

Image Fusion: The approximate coefficients of two different modality images are fused whereas the corresponding detailed coefficients of CT and MRI images are fused by using different existing image fusion rules. Mainly the Mean-Max fusion rule, where the approximate coefficients are fused with Mean value rule and the detailed coefficients are fused with Max fusion rule. Also the another widely used fusion rule is Max-Min where Max fusion rule and Min fusion rules are used respectively for the fusion of approximate coefficients and detailed coefficients. But in this study we have utilized all possible combinations of Min, Max and Mean rule, and compare their performance to find out the best combination. The Min, Max and Mean fusion rules are represented by the mathematical formulas as:

$$Mean\ Fusion\ Rule = (I(i,j) + K(i,j))/2 \tag{1}$$

$$Max\ Fusion\ rule = Maximum(I(i,j), K(i,j)) \tag{2}$$

$$Min\ Fusion\ Rule = Minimum(I(i,j), K(i,j)) \tag{3}$$

Inverse DWT: The fused images are converted into spatial domain by using inverse Discrete Wavelet Transform. It is just the reverse operation of DWT.

Denoisation of Fused Image: The fused images are denoised by using the custom denoising CNN. The specification of this CNN is given in table 2.

Table 2. Specifications of Custom Denoising CNN

Sl.No.	Parameters	Number/Name
1.	Input Layer	1
2.	Convolution Layer(5X5)	40
3.	Activation Layer	ReLu
4.	Normalization Layer	Batch Normalization
5.	Output Layer	1
6.	Learning Layer	Residual Layer
7.	No. of Epoch	5
8.	Noise Patches	Gaussian Noise(50)
9.	Loss Function	MSE

Dataset Used: In this research, the Harvard Medical School Brain Dataset has been used. (www.med.harvard.edu/aanlib/home.html). It consists of different sets of brain CT and MRI images. This dataset is mainly used for the purpose of multi modal image fusion.

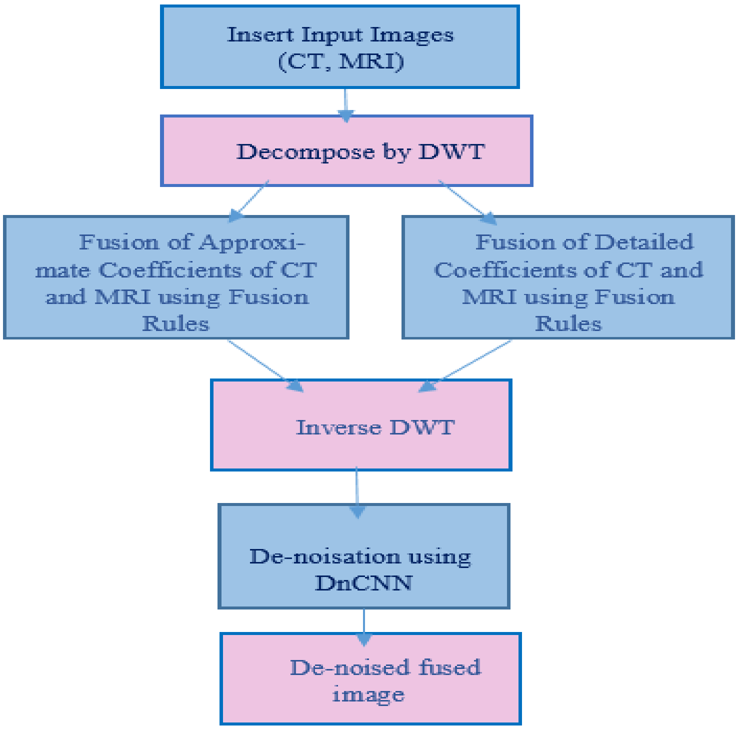

Fig. 1. The flow diagram of the fusion model

Figure 1, described that the CT and MRI images from harvard Medical school Brain Dataset are inserted and are decomposed by DWT. The approximate coefficients are fused with a fusion rule, whereas the detailed coefficients are fused with different fusion rule. The fused images are passing through inverse DWT to get back the images into spatial domain. And at the end, the custom denoising CNN is used to denoise the fused image.

4 Results and Discussion:

To evaluate the performance of the fusion rule, following metrics have been evaluated [2]:

RMSE (Root Mean Square Error): It gives the difference between the processed and original image and its value should be low.

$$RMSE = \sqrt{MSE} \tag{4}$$

Here, MSE is represented by equation 5 given below.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - K(i,j)^2 \tag{5}$$

(Peak Signal to Noise Ratio): This metric gives the image quality in terms of signal vs noise and its value should be high.

$$PSNR = 10\ log_{10}(\frac{MAX^2}{MSE}) \tag{6}$$

Here Max represents the maximum intensity value in an image, and for the gray scale image it's value is 255.

MI (Mutual Information): It represents how much information is shared between two images and its value should be high.

$$MI(I,K) = \sum_i \sum_k p\ IK\ \ (i,k) log\frac{pIK(i,k)}{pI(i).pK(k)} \tag{7}$$

Here, pI(i), pK(k): marginal probability, and pIK (i, k): joint probability distribution.

Entropy: It gives the information content or the randomness in an image. For the case of image denoisation, its value should be reduced.

$$Entropy = -\sum_{i=0}^{L-1} p_i\ log_2\ (p_i) \tag{8}$$

Here: p_i = probability of intensity level I.
L= number of gray levels(e.g., 256 for 8 bit image).

Different combinations of the fusion rules, Mean, Max, Min and the different performance metrics value achieved after and before the application of the Custom de-noising CNN is reported in table 3 below. It is cleared that, the Mean-Mean fusion rule (approximate coefficients are fused with Mean fusion rule and the detailed coefficients are also fused with Mean fusion rule) gives the best performance after fusion. The Min-Mean fusion rule gives the worst performance in fusion. After de-noisation, the PSNR and MI values of the fused images increased, where as RMSE and Entropy values have been decreased.

Table 3. Fusion results without and after using De-noising CNN

	Without using De-noising CNN					After using De-noising CNN			
Sl. No.	**Fusion Rule**	**PSNR (dB)**	**RMSE**	**MI**	**En-tropy**	**PSNR (dB)**	**RMSE**	**MI**	**En-tropy**
1.	Mean-Mean	13.586	56.6652	1.0307	5.94	30.42	23.89	1.346	5.63
2.	Mean-Max	11.313	69.4634	1.0782	5.05	29.46	24.78	1.24	4.98
3.	Mean-Min	12.81	65.4689	1.0408	6.22	30.29	23.99	1.367	6.08
4.	Max-Mean	9.730	83.446	1.1302	4.67	30.04	24.36	1.243	4.57
5.	Max-Max	9.7699	84.4886	1.0999	5.76	28.93	25.87	1.23	5.23
6.	Max-Min	10.1925	91.3578	1.209	4.82	28.85	25.98	1.256	4.32
7.	Min-Mean	6.5229	121.2524	1.0517	7.05	28.58	26.24	1.167	6.98
8.	Min-Max	12.534	65.776	1.0306	4.86	29.32	24.86	1.139	4.45

Table 3, represents the comparison of different fusion rules as well as the effectiveness of the custom denoising CNN. The information reported at table 3 reflects that the Mean-Mean fusion rule gives the best fusion result. The fusion parameter values (PSNR, MI) increased to much higher values, whereas the fusion parameter values (RMSE, Entropy) reduced to much lower values after denoisation of the fused images through the custom denoising CNN. So, it can be concluded that the custom CNN is very much efficient in image denoisation.

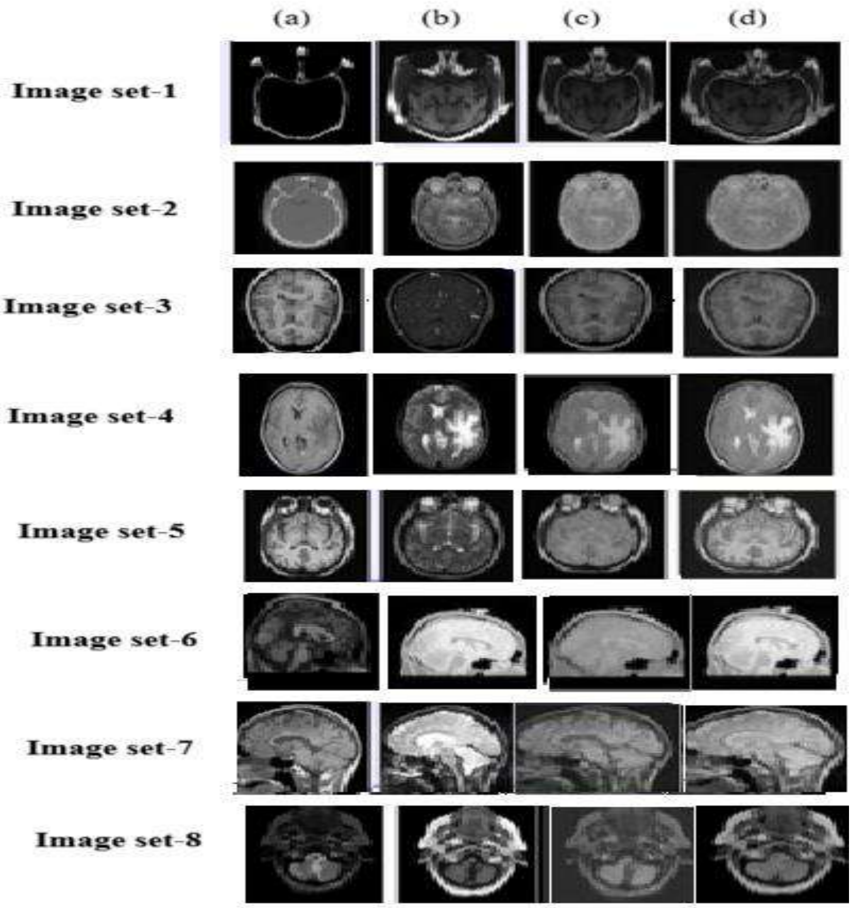

Fig.2. Qualitative analysis of image fusion with Mean-Mean image fusion rule and denoised image((a) input CT image (b) input MRI image (c) Fused image (d) Denoised fused image)

The qualitative results of the proposed fusion model are illustrated in figure 2. The images in (c) represent the fused outputs obtained using the Mean-Mean fusion rule applied to different sets of CT and MRI images. In contrast, the images in (d) show the results after applying a custom CNN-based denoising process to the fused images. It is evident that the denoised images exhibit improved clarity and enhanced details compared to the images prior to denoising. These observations further demonstrate the effectiveness of the proposed custom CNN in improving image through denoisation.

5 Conclusion:

The proposed method is the application of different existing fusion rules in different combinations for the fusion of approximate and detail coefficients achieved after decomposition of source multimodal medical images with DWT. The fused images were denoised with custom denoising CNN implemented for the reduction of Gaussian noise. Different performance metrics have been calculated to identify the best combination of fusion rules. Different performance metrics values also have been calculated after the denoisation of the fused image with custom CNN. The result shows that the fusion performance improved after denoisation. The findings can assist in selecting suitable methods for multimodal medical image fusion applications and provide a clear understanding of the advantages and disadvantages of various fusion strategies.

References

1. Xin, N. I. E. (2025). Multi-Modal Image Fusion for Medical Diagnosis: Combining MRI And CT Using Deep Generative Models. Clinical Medicine And Health Research Journal, 5(03), 1313-1327.
2. Hussain, M., O’Nils, M., Lundgren, J., & Mousavirad, S. J. (2024). A comprehensive review on deep learning-based data fusion. IEEE Access, 12, 180093-180124.
3. Hussain, S. S., Shah, P. M., Dawood, H., Degang, X., Alshamayleh, A., Khan, M. A., & Ghazal, T. M. (2025). A swin transformer and CNN fusion framework for accurate Parkinson disease classification in MRI. Scientific Reports, 15(1), 15117.
4. Sinha, A., Agarwal, R., Kumar, V., Garg, N., Pundir, D. S., Singh, H., ... & Panigrahy, C. (2024). Multi-modal medical image fusion using improved dual-channel PCNN. Medical & biological engineering & computing, 62(9), 2629-2651.
5. 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.
6. Xiang, H., Zhang, H., Cheng, Y., Quan, X., & Huang, W. (2025). SMFusion: Semantic-Preserving Fusion of Multimodal Medical Images for Enhanced Clinical Diagnosis. IEEE Journal of Biomedical and Health Informatics.
7. Cao, D., Wang, J., Yan, J., Chen, Z., Liao, X., & Cheng, H. Neighborhood-Attention-Based Multiscale Alignment and Hierarchical Reconstruction for Multimodal Medical Image Fusion. ACM Transactions on Multimedia Computing, Communications and Applications.
8. Singh, P., Ravi, V., & Bijalwan, A. (2025). Artificial intelligence-assisted medical imaging solutions for integrating pathology and radiology automated systems, volume II. Frontiers in Medicine, 12, 1686859.
9. Guo, W., Wang, L., Zeng, J., Han, Q., Jin, K., & Wang, X. (2025). SMAFusion: Multimodal medical image fusion based on spatial registration and local-global multi-scale feature adaptive fusion. Neurocomputing, 131039.
10. Liu, W., Huang, J., & Zhao, Y. (2006, October). Image fusion based on PCA and undecimated discrete wavelet transform. In International Conference on Neural Information Processing (pp. 481-488). Berlin, Heidelberg: Springer Berlin Heidelberg.