



ISSN: 0067-2904

## A Novel Multi-Level Medical Image Fusion Using Ordered Pixel Mean Filter and Parameter Adaptive Dual Channel Weighted Pulse Coupled Neural Networks

Rajasekhara Reddy Chinta<sup>1\*</sup>, Dr.Raju Kolluri<sup>2</sup>

<sup>1</sup>Jawaharlal Nehru Technological University, Kakinada, Andhra Pradesh, India,

<sup>2</sup>Narasaraopet Engineering College, Andhra Pradesh, India

Received: 12/6/2025

Accepted: 18/1/2026

Published: xx

### Abstract

Medical image fusion is a crucial technique that has been extensively employed in clinical diagnosis and treatment planning, driven by rapid advances in imaging techniques. Though several methods have been introduced in the past for effective fusion of source modalities, there is still room to improve fusion accuracy. In this work, a novel method of image decomposition is introduced using the mean of ordered pixel values filter, and is applied for multi-scale decomposition of input images for fusion. Further, a weight computation mechanism based on singular values of the Singular Value Decomposition (SVD) method is introduced to fuse coarse layers of source images, while the detail layers are fused at multiple levels with the help of Parameter-Adaptive Dual-Channel Weighted Pulse Coupled Neural Networks (PADCWPCNN). The fusion performance of the proposed method is evaluated on 125 slices of diverse brain diseases using both subjective and objective criteria. Experimental results demonstrated the efficacy of the proposed method, as it improved Average Gradient (AG), Standard Deviation (SD), Entropy (H), Spatial Frequency (SF), Edge Strength ( $Q_{ABF}$ ), and Mutual Information (MI) by 57%, 3%, 14.7%, 17.9%, 46%, and 14%, respectively, making it an appealing option for fusing medical images to effectively diagnose tumors and plan treatment accordingly.

**Keywords:** Medical image fusion, mean of ordered pixel values filter, singular values, parameter adaptive dual channel weighted PCNN

### 1. Introduction

Medical image fusion is an important technique that has been extensively employed in clinical diagnosis as well as treatment planning [1]. Medical imaging techniques vary and provide diverse details regarding the organs of the body. Computerized Tomography (CT) offers exceptional resolution, allowing for the analysis of smaller features in tissue, but it is ineffective for depicting overall organ anatomy and disease alterations. Magnetic Resonance Imaging (MRI) provides excellent soft tissue characterization with great spatial resolution; however, it is restricted in detecting fractures. Furthermore, anatomical imaging techniques like CT and MRI cannot capture movement-related information, such as metabolism. Metabolic processes, along with various physiological activities, including blood flow and localized chemical composition absorption, can be seen using Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) [2]. However, the spatial resolution of functional imaging methods like PET and SPECT is poor. In conclusion, a single imaging modality cannot provide all the information about an organ. Multimodal medical image

\*Email: [ch.raja2803@gmail.com](mailto:ch.raja2803@gmail.com)



fusion, which integrates several medical pictures from different modalities to increase quality and reduce duplication of imaging, is a useful technique for increasing the clinical reliability of diagnosis from medical images.

The majority of medical image fusion algorithms rely on the multiscale transform (MST) that employs predefined functions to produce the transformed coefficients after converting the source images into the transform domain. Inverse MST is applied to the processed coefficients to produce the fused image. MST can be separated into three categories depending on the various source image decomposition techniques: wavelet-based techniques [3–4], multiscale geometric analysis (MGA)-based techniques [5–7], and pyramid-based techniques [8–9]. Some crucial aspects of the source images, including edge as well as texture information, might not be adequately represented and retrieved due to the MST-based algorithm's preset function limitations, which drastically lower the fusion performance. Furthermore, misregistration typically affects the MST-based method.

Sparse representation (SR) is a prominent approach for picture fusion that optimizes over complete dictionaries. The technique is more mathematically reasonable and aligns with the human visual system compared to MST. Thus, a multimodal neural image fusion framework that improves adaptability through sparse representation and directed filtering merging is proposed [10]. However, this approach increases computational complexity. [11] introduced LRR theory as a way for investigating information spatial structure in 2010. This technique employs the data matrix, treated as a dictionary, and only analyzes the global structure of an image. As a result, its practical application is limited. To solve this issue, the authors proposed LatLRR, an improved technique based on LRR theory [12]. LatLRR constructs a dictionary using both the observable data matrix, including the unobserved latent data, taking into account the local structure, global structure, and sparse noise of the image information.

An image fusion method that uses Non Subsampled Shearlet Transform (NSST), Convolutional sparse representation (CSR) approach is proposed in [13]. In this approach, the input images were separated into low- and high-frequency sub-bands by the NSST. Next, the high-frequency and low-frequency sub-bands were fused using PAPCNN and CSR techniques. The fused image is reconstructed using inverse NSST. The suggested approach performs well in terms of both objective and visual outcomes. The inability of the quality measurements to combine MR-T1, MR-T2, and SPECT imaging modalities is its drawback. In the NSCT transform, a fusion method based on the Pulse-Coupled Neural Network (PCNN) approach is put forward in [14]. In order to separate the input image into high- and low-frequency bands, this method employed the NSCT approach. The low-frequency bands were then fused using PCNN, and the high-frequency bands were fused using the WSEML technique and guided image filtering (GIF). According to the findings, the method demonstrated performs well in medical image fusion.

To improve the visual quality and diagnostic accuracy of multimodal medical images, a number of deep learning-based and hybrid adaptive fusion techniques have been developed recently. To generate realistic and high-contrast fused outputs, a GAN-based method uses adversarial learning [15]. For CT–MRI image fusion, a multimodal image fusion network (MMIF-Net) and a multi-scale hybrid attention network are suggested [16]. It captures both local and global dependencies. By formulating image fusion as a weighted average process and making suitable assumptions about the noise in the diffusion model, a flexible diffusion model-based unsupervised framework for image fusion, called VDMU Fusion, is developed [17]. This framework integrates the fusion problem into the diffusion sampling process.

Taking motivation from the literature, a new method of multimodal medical image fusion is put forward in this work with the following contributions:

1. A novel method based on Mean of Ordered Pixel Values (MOPV) is introduced for image decomposition.
2. The proposed MOPV is used for multi-level decomposition of multi-modal input images.
3. A parameter adaptive dual-channel weighted PCNN is proposed for the fusion of high-frequency information of source images.
4. A novel method based on the singular values of the SVD method is suggested for the generation of weight maps to fuse low-frequency information of source images.

## 2. Related Concepts

In this work, a novel method of image decomposition is proposed using the concept of Mean Value of Ordered Pixel Values (MOPV). It is a non-linear filtering technique in the spatial domain and can be used to preserve edges in an image while suppressing the noise during a smoothing operation. The working principle of the MOPV filter is as explained as this: Let  $P$  be a gray-scale image of size  $M \times M$ . Select a window  $W$  of size  $w \times w$  for the spatial filtering of each pixel  $(k, l)$  of an image. In traditional averaging filters, the current process pixel  $(k, l)$  of an image is replaced with the average value of its surrounding neighbors covered by the window  $W$ , when the processed pixel  $(k, l)$  occupies the centre of the window. Similarly, in the proposed MOPV filter, the neighbors are selected for a pixel  $(k, l)$  using the window  $W$  as in the conventional method, and then instead of considering all neighbors for averaging, they are sorted, and a subset of sorted pixel values is used for the computation of the mean.

### 2.1 Procedure of the Mean of Ordered Pixel Values

Step 1: For each pixel  $P(k, l)$  of an image, extract the local neighborhood  $W$  of size  $w \times w$ .

Step 2: Sort the local neighborhood.

$$S = \text{sort}(W) \quad (1)$$

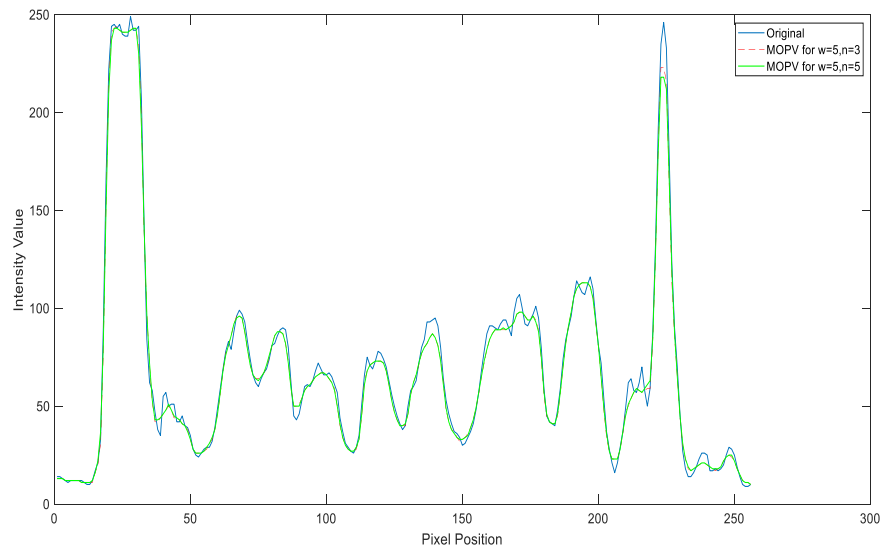
Step 3: A central subset is selected around the median from  $S$

$$S_{\text{subset}} = \{S[N - n], \dots, S[N], \dots, S[N + n]\} \quad \text{where } N = \frac{|S|}{2} \quad (2)$$

Here  $|S|$  implies the number of pixels in the sorted list, and  $N$  indicates the index of the central pixel in the sorted set, and  $n$  specifies the number of pixels to be considered on either side of the center pixel.

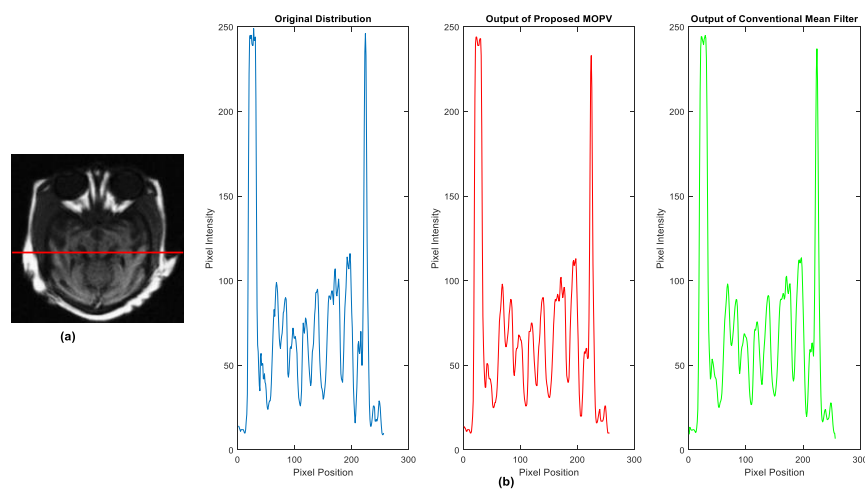
Step 4: The Current processed pixel  $P(k, l)$  is replaced with the mean value of  $S_{\text{subset}}$ . The mean of the selected subset minimizes the influence of outlier pixels (extremely bright or dark). Averaging only the ordered core subset preserves local edges and reduces background noise, balancing smoothing and edge preservation.

$$\text{Mean} = \frac{1}{|S_{\text{subset}}|} \sum_{i \in S_{\text{subset}}} S_{\text{subset}}(i) \quad (3)$$



**Figure 1:** Performance of MOPV filter for selected window size with different  $n$  values

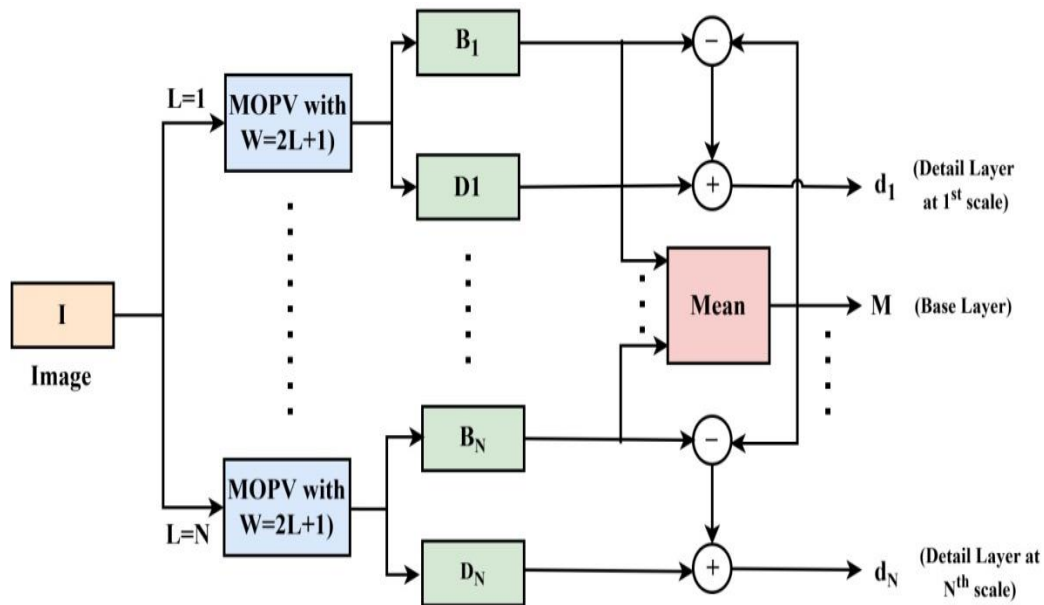
In the proposed work,  $n = 3$  is utilized, and hence a total of 9 pixels in the sorted set  $S$  are used to compute the mean of the required pixel. The effect of  $n$  value in the edge preserved smoothing operation for a given window size  $5 \times 5$  is illustrated in Figure 1 for  $n=3$  and 5. The larger values of  $n$  consider a wider subset of sorted pixels for the smoothing operation. Though this helps to reduce noise, unfortunately, some edge information also gets smoothed. This could eliminate structurally or diagnostically important features from satellite or medical images. The objective of MOPV is not to uniformly perform a smoothing operation and preserve structural details, which is more possible with  $n = 3$ . The superiority of the proposed filtering technique compared to the conventional mean filter is illustrated in Figure 2, where 2(a) is the input image considered for testing. Its intensity distribution for a particular row, highlighted in red, is shown in the first plot of Figure 2(b). The results of the proposed MOPV and conventional mean filter are shown in the second and third plots, respectively, of Figure 2(b). In contrast to the Conventional Mean Filter, which smooths out more and dramatically flattens peaks and valleys, particularly around edges, the proposed MOPV filter preserves most of the original structure, including edges and intensity transitions.



**Figure 2:** Analysis of the edge-preserving ability of the proposed and conventional mean filters

## 2.2 Multilevel Image Decomposition using MOPV filter

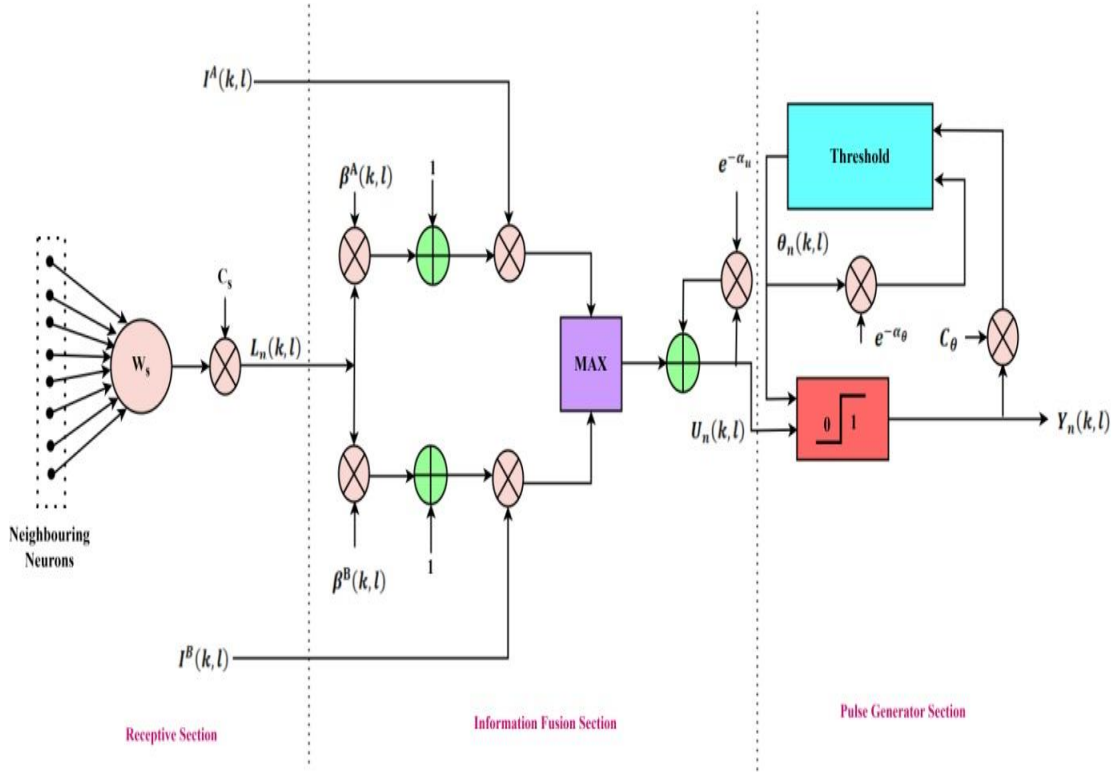
In the case of multi-scale decomposition methods, a window of the same or different dimensions is applied progressively on the output of each stage, whereas in multi-level decomposition, a window of varying dimensions is applied in parallel onto the same input image. The structure of the proposed multi-level decomposition of an image is illustrated in Figure 3. Here, ' $L$ ' indicates the number of levels in the decomposition process, and at each level, the size of the window is progressively increased according to the relation  $w = 2 * L + 1$ . Further, in each level, the output of the filter is treated as low-frequency information of an image, which basically contains homogeneous intensity variation and is often called base or coarse layers ( $B_L$ ) in image fusion terminology. The high-frequency information, which contains abrupt intensity variations like edges of an image, which are also called details ( $D_L$ ), is obtained at each level by computing the difference between the original image and the base layer of that level [18].



**Figure 3:** Multi-level image decomposition mechanism using the proposed MOPV filter

### 2.3 Parameter adaptive dual channel weighted PCNN (PADCWPCNN)

The parameters of conventional PCNN-based image fusion are primarily defined by a large number of experiments or experience. Furthermore, features that perform effectively in one application may not be appropriate in another. The orientation information is employed as the linking strength for automating PCNN settings, while the remaining parameters are determined manually [21]. This variant just implements certain aspects of the automation. This study uses the unit-linking PCNN model and extends it to a dual-channel PCNN model, which drastically reduces the number of parameters. Figure 4 depicts the suggested dual-channel PCNN model with adaptive weights. It has three main sections: the receptive portion, the information fusion section, and the pulse generator. The receptive section accepts stimuli from neighboring neurons in addition to external inputs. The information fusion portion integrates source images based on neuron activity levels. The pulse generator is responsible for producing the output pulse. The suggested model has the same total number of input neurons as there are pixels in an image. Equations (4)–(11) describe the mathematical modeling of the suggested paradigm.



**Figure 4:** Structure of parameter adaptive dual channel weighted PCNN

$$F_n^A(k, l) = I^A(k, l) \quad (4)$$

$$F_n^B(k, l) = I^B(k, l) \quad (5)$$

$$L_n(k, l) = C_s \sum_{i=-1}^1 \sum_{j=-1}^1 W_s(i+2, j+2) Y_{n-1}(k+i, l+j) \quad (6)$$

$$U_n^A(k, l) = F_n^A(k, l) (1 + \beta^A(k, l) L_n(k, l)) \quad (7)$$

$$U_n^B(k, l) = F_n^B(k, l) (1 + \beta^B(k, l) L_n(k, l)) \quad (8)$$

$$U_n(k, l) = e^{-\alpha_u} U_{n-1}(k, l) + \max\{U_n^A(k, l), U_n^B(k, l)\} \quad (9)$$

$$Y_n(k, l) = \begin{cases} 1 & \text{if } U_n(k, l) > \theta_{n-1}(k, l) \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

$$\theta_n(k, l) = e^{-\alpha_\theta} \theta_{n-1}(k, l) + C_\theta Y_n(k, l) \quad (11)$$

Where  $I^A(k, l)$  and  $I^B(k, l)$  are the external inputs of the proposed weight adaptive dual channel PCNN model at  $(k, l)$  coordinates of the input images  $A$  and  $B$ , respectively. Further  $I^A(k, l) \in [0, 1]$  and  $I^B(k, l) \in [0, 1]$ .  $F_n^A(k, l)$  and  $F_n^B(k, l)$  are the symmetrical feeding inputs that hold  $I^A(k, l)$  and  $I^B(k, l)$  respectively, and do not vary in every iteration.  $L_n(k, l)$  is the linking parameter of  $(k, l)^{th}$  neuron in  $n^{th}$  iteration. The neural weights of eight nearby neurons of the  $(k, l)^{th}$  neuron are denoted by the  $3 \times 3$  matrix  $W_s$ .

$$W_s = \begin{bmatrix} 0.5 & 1 & 0.5 \\ 1 & 0 & 1 \\ 0.5 & 1 & 0.5 \end{bmatrix}$$

$U_n^A(k, l)$  and  $U_n^B(k, l)$  characterize the internal state of the  $(k, l)^{th}$  neuron after  $n^{th}$  iteration for input images  $A$  and  $B$ , respectively, both of which taken together make up the non-linear modulation of feeding and linking inputs.  $\theta_n(k, l)$  is the dynamic threshold, and  $Y_n(k, l)$  is the external output of  $(k, l)^{th}$  after  $n^{th}$  iteration, respectively.  $\beta^A(k, l)$  and  $\beta^B(k, l)$  are the linking intensities of  $I^A(k, l)$  and  $I^B(k, l)$ . Linking intensity represents how the human visual system reacts to a pixel and its surroundings, and it cannot be identical for the  $(k, l)^{th}$  pixel of two source images taken using different modalities. Assuming that every pixel in an image has identical linking intensity is not accurate. Unlike the conventional PCNN model, which

employs a single linking strength for every pixel in the image, the proposed PCNN model uses adaptive linking strength and is hence named as parameter adaptive dual channel weighted PCNN (PADCWPCNN). It assigns different values to each pixel. By utilizing equation (12), the linking intensity of the  $(k, l)^{th}$  pixel can be determined.

$$\beta^X(k, l) = (E^X(k, l) + M^X(k, l)) \text{ where } X \in (A, B) \quad (12)$$

Where  $E^X(k, l)$  is the gradient of the image  $X$  that provides the edge map, which is computed using equation (13), while  $M^X(k, l)$  provided in equation (14) computes the morphological map.

$$E^X(k, l) = \sqrt{\left(\frac{\partial X}{\partial x}\right)^2 + \left(\frac{\partial X}{\partial y}\right)^2} \quad (13)$$

$$M^X(k, l) = ((X \oplus S)(k, l) - (X \ominus S)(k, l)) \quad (14)$$

Where  $\oplus$  represents dilation,  $\ominus$  indicates erosion operation, and 'S' is the disk-shaped structuring element. Normalization of the linking term  $\beta^X(k, l)$  is carried out using Equations (15), (17) and (18), and the decay behavior is modeled as in Equation (21).

$$\beta^X(k, l) = \frac{\beta^X(k, l) - \min(\beta^X)}{\max(\beta^X) - \min(\beta^X)} \quad (15)$$

$$\alpha_U = \log\left(\frac{1}{w_1 \sigma_{IA} + w_2 \sigma_{IB}}\right) \quad (16)$$

$$w_1 = \frac{F_A}{F_A + F_B} \quad (17)$$

$$w_2 = \frac{F_B}{F_A + F_B} \quad (18)$$

$$C_\theta = e^{-\alpha_u} + 1 + 6\lambda \quad (19)$$

$$\alpha_\theta = \ln\left(\frac{C_\theta}{(w_1 I^{A'} + w_2 I^{B'})\left(\frac{1 - e^{-3\alpha_u}}{1 - e^{-\alpha_u}} + 6\lambda e^{-\alpha_u}\right)}\right) \quad (20)$$

$$\lambda = \frac{(w_1 I^{B'} I_{max}^A + w_2 I^{A'} I_{max}^B - I^{A'} I^{B'})}{6 I^{A'} I^{B'}} \quad (21)$$

$C_\theta$  indicates the amplitude of the dynamic threshold, while  $e^{-\alpha_u}$  is the decaying coefficient with exponential behavior. A lower  $\alpha_u$  leads to a diverse distribution of  $U_n$ , which is accurately expressed by a logarithmic function [21]. Therefore, equation (16) is employed to evaluate the value of  $\alpha_u$ . Here  $\sigma_{IA}$  and  $\sigma_{IB}$  are the standard deviations of  $I^A$  and  $I^B$ , while  $w_1$  and  $w_2$  are weights derived using fractal dimensions of input images  $A$  and  $B$  evaluated using the DBC method [20]. Following the work of [21], the values of  $\alpha_\theta$  and  $C_\theta$  are simulated using equations (19) and (20).  $I^{A'}$  and  $I^{B'}$  are the Otsu thresholds of  $I^A$  and  $I^B$ , respectively [22]

#### 2.4 Weight Computation based on Singular Values

The weights of the pixels within each base layer are calculated according to their relevance, and they are fused by applying the weighted average rule, as the base layer of any image contains pixels with homogenous grayscale intensities. The local energy of nearby pixels is taken into account while calculating each pixel's weight. A window of size  $w \times w$  is projected so that the pixel of image  $M$  at  $(k, l)$ , represented as  $M(k, l)$ , occupies the center of the window. The area around the pixel that the window covers is taken as matrix  $P$ . Equation (22) is utilized to get the singular values of  $P$  [23].

$$P = USV^T$$

(22)

$S$  is a diagonal matrix specified in equation (23), which contains the singular values of matrix  $P$  obtained by computing the Eigenvalues of  $PP^T$  and  $U, V$  are the singular vectors and  $V^T$  indicates the transpose of matrix  $V$ .

$$S = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_N \end{bmatrix} \quad (23)$$

From singular values, the weight of a pixel located at  $(k, l)$  is taken as the sum of singular values as specified in equation (24), which reflects its local energy. In this way, the weights of each pixel for coarse layers of each of the input images are determined.

$$W(k, l) = \sum_{i=1}^N \sigma_i \quad (24)$$

### 3. Proposed Methodology

If the two source images are grayscale images, such as MRI and CT, the proposed technique applies them immediately to the first stage of the fusion process, which is multilayer source image decomposition. If one of the source images is an RGB color image, such as PET or SPECT, it is first transformed into the YUV color space, and the intensity component 'Y' is used for decomposition during the fusion process. Figure 5 depicts the comprehensive procedure flow of the proposed method for grayscale images, whereas Figure 6 depicts a color image combined with a grayscale image.

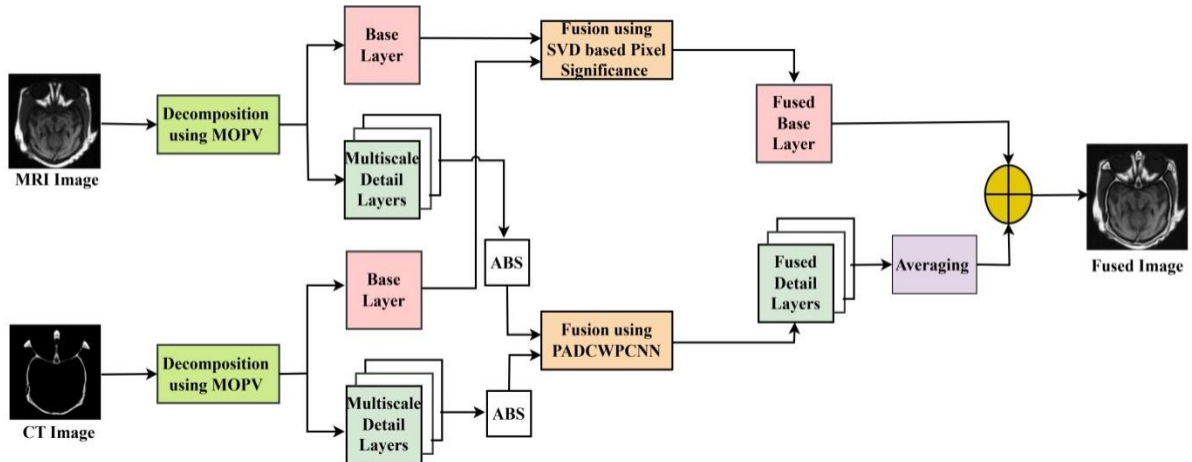


Figure 5: Structure of the CT and MRI image fusion process

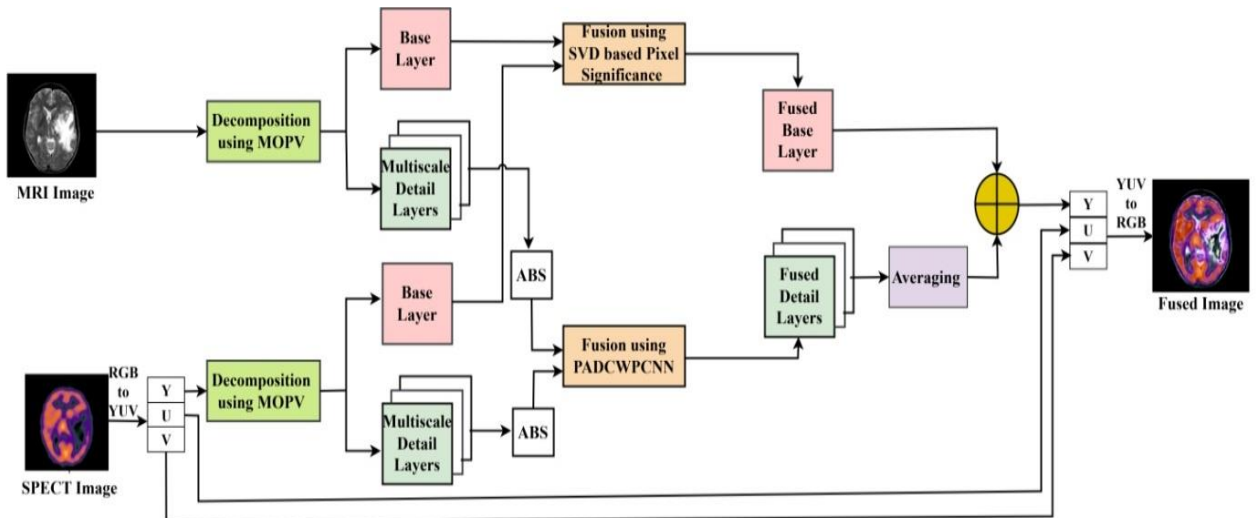


Figure 6: Structure of the PET/SPECT and MRI image fusion process

The step-by-step approach of the fusion process is as follows:

1. The input images are scaled to the same size (256x256) and assigned the labels  $A$  and  $B$ .
2. A four-level decomposition process using the proposed MOPV structure shown in Figure 3 is applied on each input to decompose it into coarse and detail layers.
3. Coarse layers of each level are designated as  $C_{1A}, C_{2A}, C_{3A}$ , and  $C_{4A}$  for image  $A$ , whereas the coarse layers of image  $B$  are designated as  $C_{1B}, C_{2B}, C_{3B}$ , and  $C_{4B}$ .
4. Using the coarse layers at each level, the smoothed coarse layer of each of the inputs is derived using equations (25) and (26), respectively.

$$\text{Coarse layer of input } A, M_A = \frac{1}{4} \sum_{L=1}^4 C_{LA} \quad (25)$$

$$\text{Coarse layer of input } B, M_B = \frac{1}{4} \sum_{L=1}^4 C_{LB} \quad (26)$$

5. Detail layers at each level are derived by subtracting the coarse layer at that level from the respective source image, as indicated in equations (27) and (28).

$$\text{Detail layer at level } L \text{ of input } A, D_{LA} = A - C_{LA} \quad (27)$$

$$\text{Detail layer at level } L \text{ of input } B, D_{LB} = B - C_{LB} \quad (28)$$

6. Using the structure of Figure 3, the modified detail layers at each level are computed using equations (29) and (30).

$$\text{Modified detail layer at level } L \text{ of input } A, d_{LA} = D_{LA} + (C_{LA} - M_A) \quad (29)$$

$$\text{Modified detail layer at level } L \text{ of input } B, d_{LB} = D_{LB} + (C_{LB} - M_B) \quad (30)$$

7. Apply the weight computation procedure discussed in section 2.4 on coarse layers of each input to compute their weights, which are specified in equations (31) and (32).

$$W_A(k, l) = \sum_{i=1}^N \sigma_{MAi} \quad (31)$$

$$W_B(k, l) = \sum_{i=1}^N \sigma_{MBi} \quad (32)$$

8. Coarse layers are fused using the above weights using the weighted combination rule described in equation (33).

$$\text{Fused coarse layer, } M_f(k, l) = (W_A(k, l) * M_A(k, l)) + (W_B(k, l) * M_B(k, l)) \quad (33)$$

9. Detail layers of different levels are fused using the PCNN model discussed in section 2.3. This is accomplished by using detail layers at a specific level of the source image as external stimuli to PCNN, specified in equation (34), and computing the appropriate internal activity  $U_n(k, l)$  as follows:

$$I^A(k, l) = d_{LA}(k, l) \text{ and } I^B(k, l) = d_{LB}(k, l) \quad (34)$$

By equations (6) to (13),  $U_n^A(k, l)$  and  $U_n^B(k, l)$  are evaluated, and detail layers are then fused using the maximum selection rule of fusion described in equations (35) and (36).

$$\text{map}(k, l) = \begin{cases} 1 & \text{if } U_n^A(k, l) \geq U_n^B(k, l) \\ 0 & \text{if } U_n^A(k, l) < U_n^B(k, l) \end{cases} \quad (35)$$

Detail layer at each level of fusion,

$$d_{Lf}(k, l) = \text{map}(k, l) * d_{LA}(k, l) + (1 - \text{map}(k, l)) * d_{LB}(k, l) \quad (36)$$

10. The final fused image of the given inputs is produced by combining the fused base layer and the averaged fused detail layer, as illustrated in equation (37).

$$\text{Fused image}(k, l) = M_f(k, l) + \left( \frac{d_{1f}(k, l) + d_{2f}(k, l) + d_{3f}(k, l) + d_{4f}(k, l)}{4} \right) \quad (37)$$

The proposed MOPV-PADCWPCNN framework was implemented in MATLAB 2022b on an Intel Core i7 CPU with 16 GB RAM. The Mean of Ordered Pixel Values (MOPV) filter employed a 5 x 5 window at the first level of decomposition, increasing by two pixels at each successive level (5x5, 7x7, 9x9, and 11x11). The initial parameters for the Parameter-Adaptive Dual-Channel Weighted Pulse Coupled Neural Network (PADCWPCNN) were empirically optimized using [19] and [21].

#### 4. Results and Discussion

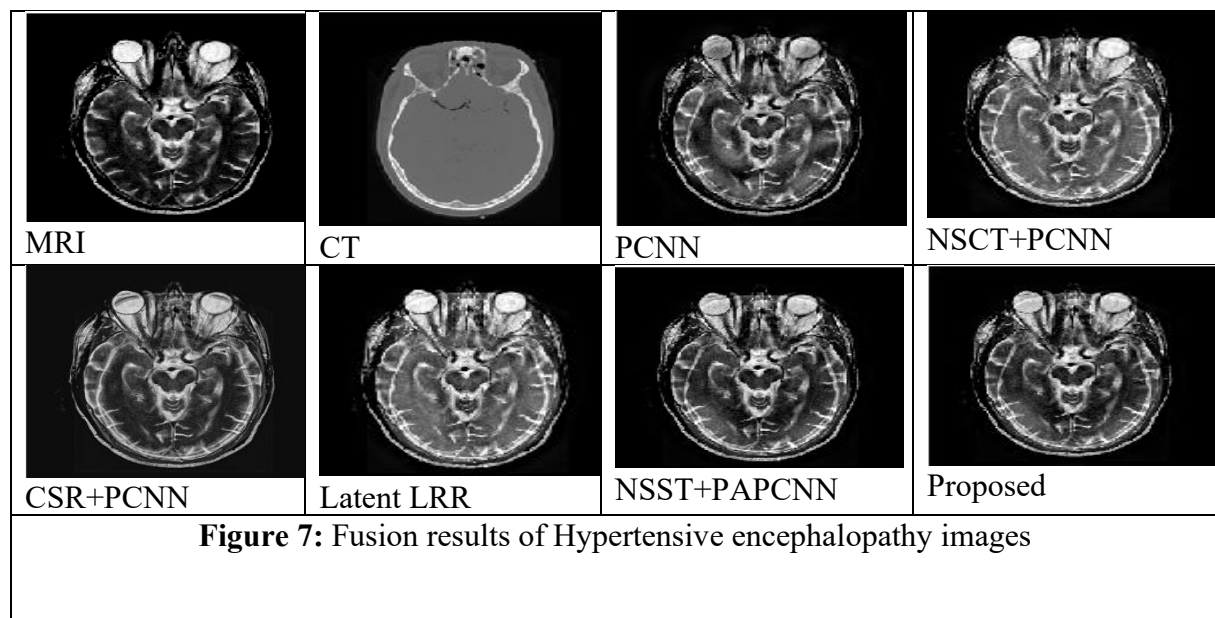
The benchmark images used for experimentation in the proposed work are utilized extensively and have no bearing on the rights of humans or race. Each of these images was taken from the standard repository of images of various brain disorders at <https://www.med.harvard.edu/aanlib/>. Furthermore, each of these images is taken at the same slice and angle, utilizing a variety of imaging methods. To verify the performance of the proposed algorithm, 125 slices of the source imagery from five major brain disorders were selected for testing. From each of the five datasets, 25 slices were chosen. The selected source images were spatially adjusted to  $256 \times 256$  pixels and exactly matched before being used for testing. Hypertensive encephalopathy MRI and CT scans are examined as the first dataset. This illness is primarily caused by excessive blood pressure and falls under the category of Cerebrovascular diseases of the brain.

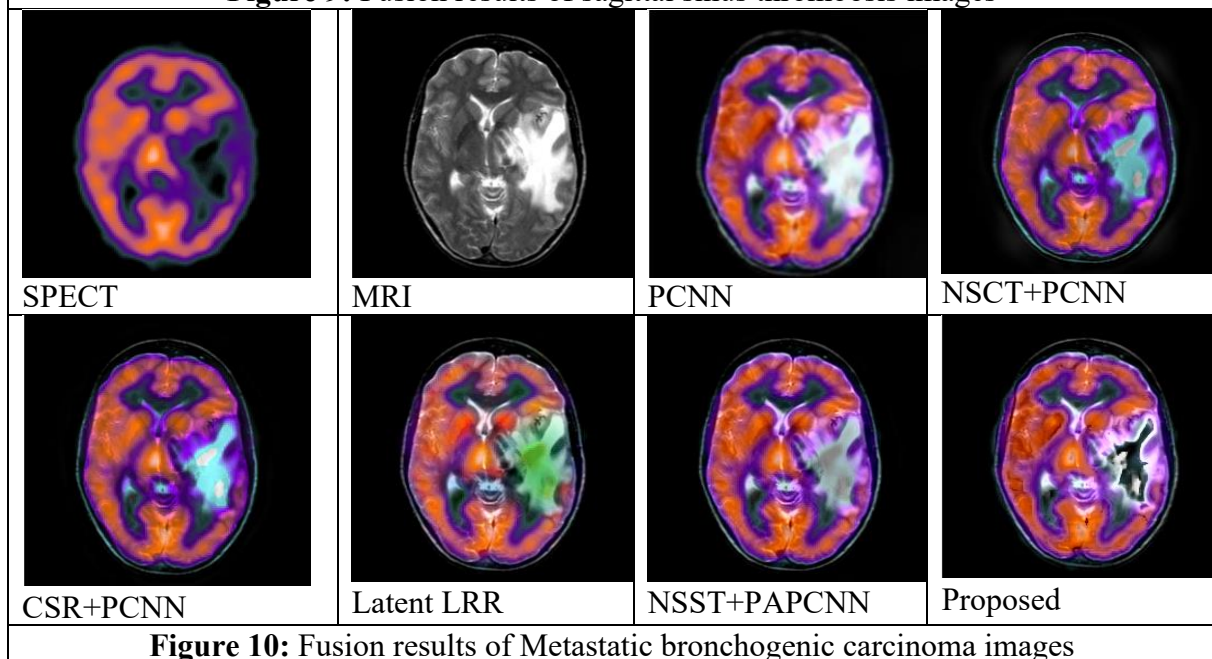
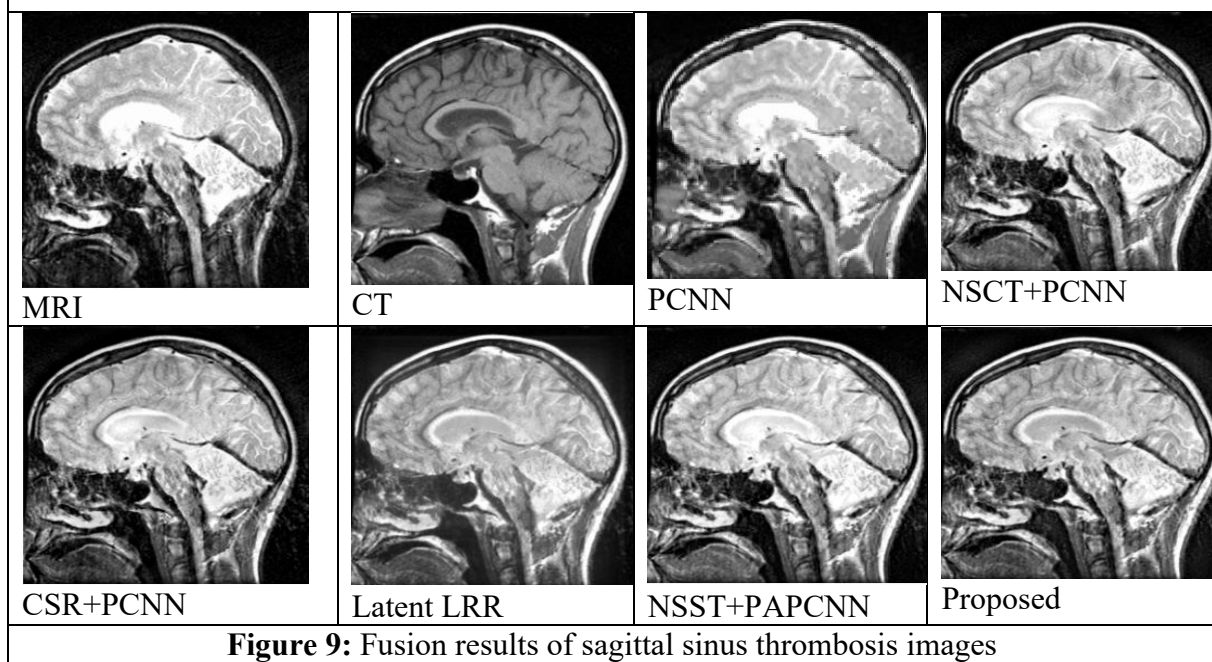
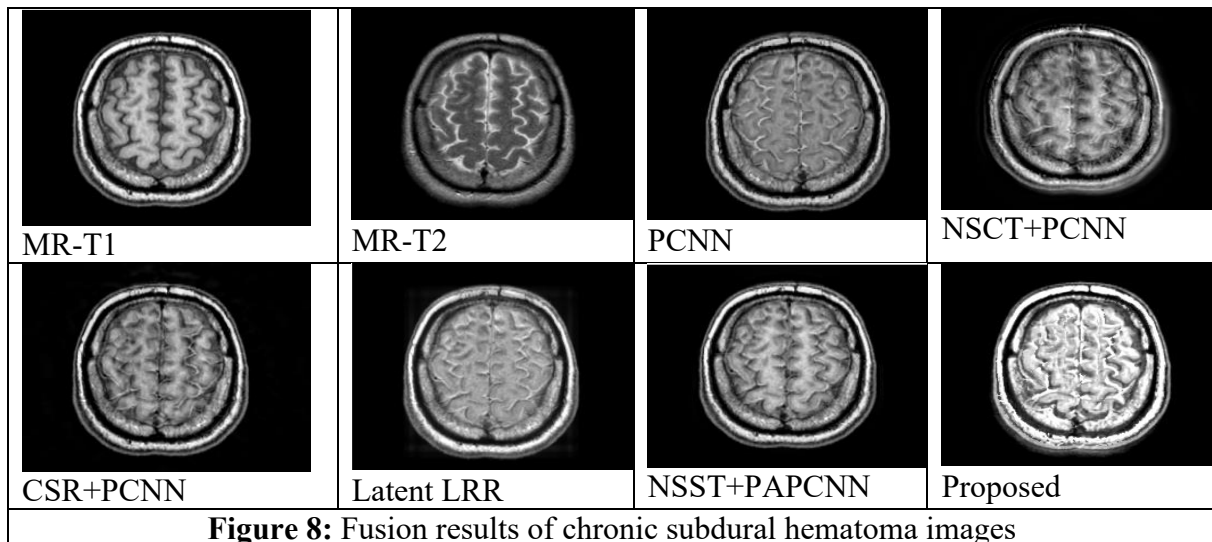
The second dataset consists of CT and MRI scans of sagittal sinus thrombosis. It drains blood from the brain into the venous circulation and is the primary vein in the brain where blood clots. This may lead to an increase in intracranial pressure, which increases the risk of serious neurological side effects like stroke, seizures, and coma.

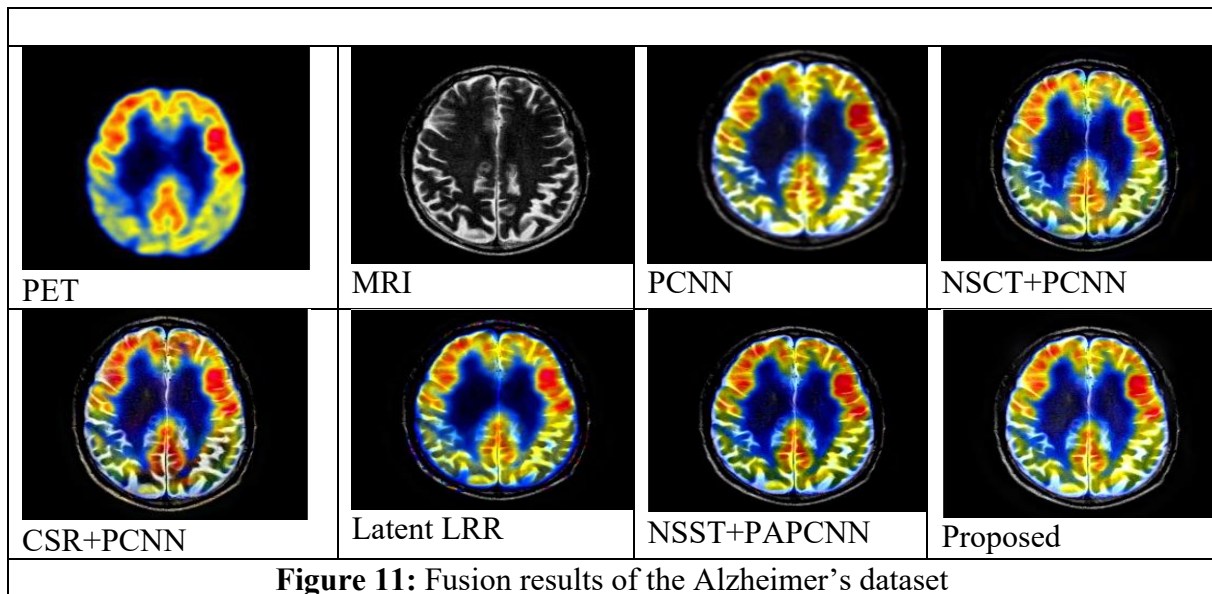
The third dataset is associated with MR-T1 and MR-T2 scans of patients suffering from chronic subdural hematomas. It is a type of brain disease called cerebrovascular disease. It is the progressive deposition of blood and the byproducts of blood breakdown between the neurons and the outer layer.

The fourth dataset relates to SPECT and MR imaging of brain metastases of bronchogenic cancer. It is caused by cancerous cells from lung tumors that travel to the brain. One or more brain tumors may occur as a result of the elevated pressure and disruption of brain tissue function, which can cause a variety of symptoms.

The fifth dataset shows Alzheimer's disease PET and MR images. Memory, reasoning, and other cognitive abilities gradually deteriorate due to this progressive neurodegenerative disease that primarily affects the brain. It is the most common cause of dementia, leading to changes in behavior and psychology and, ultimately, the incapacity to carry out everyday tasks.







#### 4.1 Subjective Analysis

The fusion outcomes of the aforementioned datasets are depicted in Figures 7 through 11. In each figure, the fused images of several established techniques and the suggested method are shown from (c) to (h), while the original images are shown in (a) and (b). From Figure 7, which contains the fusion results of hypertensive encephalopathy images, it can be observed that the PCNN method could provide improved contrast than individual source images. The information from the source images is not fully integrated into the fused image shown in Figure 7(c). As seen in Figure 7(d), NSCT+PCNN could preserve more structural details from source modalities compared to Figure 7(c). It is identified from Figure 7(e) that CSR+PCNN could provide better feature localization than earlier methods. Results of Latent LRR from Figure 7(f) show some slightly blurred details, although it provides balanced detail preservation from source modalities. The NSST+PCNN method provides a fused image with good texture and great structural details, as seen in Figure 7(g). Figure 7(h) contains the fusion result of the proposed method, which has anatomical details of both source modalities integrated comprehensively with improved contrast.

Figure 8 shows the fused images of the MR-T1 and MR-T2 scanned images of a chronic subdural hematoma. Figures 8(a) and 8(b) demonstrate that an MR-T1 scan emphasizes an organ's anatomical features, whereas an MR-T2 scan maintains the organ's pathological details. As seen in Figure 8(c), the PCNN approach still has the blur issue and loses some information of the MR-T2 image in the process of fusion. Even though the NSCT+PCNN approach is superior to the PCNN method at retaining structural information, it is clear from the fusion result shown in Figure 8(d) that the technique produces halos in the deep neural regions during the process of fusion. Although the CSR+PCNN approach produces a fused image with sufficient tissue and structural preservation, it creates intensity inconsistencies that cause the fused image's boundaries to become extremely sharp, as shown in Figure 8(e). The central portion of the fused image appears less realistic due to the artificial brightness, even though the periphery structures are quite well preserved by the comparable Latent LRR approach in Figure 8(f). Even while the NSST-PCNN technique produces a superior fused image, Figure 8(g) shows that the center of the image is slightly smoothed. Compared to other approaches, the fused image of the suggested method, as seen in Figure 8(h), contains fewer artifacts and exceptional preservation of both the pathogenic abnormalities of MR-T2 and the anatomical details of MR-T1.

The fusion results of the sagittal sinus thrombosis images are shown in Figure 9. The PCNN approach introduces blur in the soft tissue region, leading to a low contrast image, although it was able to preserve the soft tissue and bone features of the input images. As seen in Figure 9(d), the NSCT+PCNN approach could provide superior information retention in the central brain region compared to PCNN; nevertheless, it produces some blur as well as halos near edges. It is clear from Figure 9(e) that the CSR+PCNN approach preserves edges far better than previous approaches, but it also dilutes the structural information of the source modalities in the fused image. Figure 9(f) illustrates the way Latent LRR introduces over-enhancement into the fused image, despite the fact that it improves contrast by preserving the soft tissue and bone information. According to Figure 9(g), NSST+PCNN produces an adequate fusion output by smoothing the fused image's periphery while maintaining an appropriate contrast and sufficient fine details. Conversely, the suggested approach balances the preservation of MRI and CT characteristics in the fused image while producing a fused image with acceptable texture retention without excessive sharpening. As seen in Figure 9(h), it also introduces fewer artifacts than other techniques.

Figure 10 presents an analysis of the fusion outcomes of brain metastatic bronchogenic cancer using SPECT and MRI scans. The PCNN approach loses clarity, as seen in Figure 10(c), but it could maintain a respectable percentage of the SPECT picture's functional information in the fused image. Findings from figures 10(d) to 10(g) showed that NSCT+PCNN suffers from artifacts, whereas color distortion is introduced by CSR+PCNN and Latent LRR techniques. Despite outperforming other approaches, the NSST+PCNN method's performance falls short of the suggested method, which guarantees a fused image free of visual artifacts, as seen in Figure 10(h).

Figure 11 illustrates the fusion outcomes from Alzheimer's disease PET and MRI scan images. Figures 11(a) and 11(b) make it clear that a PET scan shows an organ's metabolic activity, but an MRI scan maintains information on soft tissues. The fused image displayed in Figure 11(c) still has blur and low contrast, even though the PCNN approach produces a fused image with a moderate retention of MRI features and a decent conservation of PET details. As seen in Figure 11(d), the NSCT+PCNN approach produces structural distortion but could produce a fused image with improved energy localization. The CSR+PCNN procedure produces output with excessively exposed PET details while lowering the MRI details, as can be observed in the fused image of Figure 11(e). Figure 11(f) makes it clear that the latent LRR approach introduces intensity mixing near edges. As seen in Figure 11(g), the NSST-PCNN technique produces a fused image with constant features but minor softening around the edges. As shown in Figure 11(h), the suggested method produces a merged image with optimal metabolic and anatomical details, sharper edges, and less distortion than other methods.

#### 4.2 Objective Analysis

Before deciding whether a fusion algorithm is effective, its performance must be evaluated both objectively and subjectively. The objective analysis of the fusion algorithms that were compared in the subjective analysis section is presented in this section. Six metrics were taken into consideration for comparison: Average Gradient (AG), Standard Deviation (SD), Image Entropy (H), Spatial Frequency (SF), Edge Preservation Index ( $Q_{ABF}$ ), and Mutual Information (MI). Table 1 provides the mathematical formulas for these measurements as well as their relevance. Evaluated metric values of the proposed method for all considered datasets are provided in Table 2 along with existing methods [24].

The performance of the PCNN method is low across all metrics in all datasets. Lower values of AG and SF indicate their ineffectiveness in retaining the edges. Information content in the fused image produced by this method is also poor, as it provides low values of Entropy and MI compared to other methods. When PCNN is used in combination with NSCT, there is an improvement of 10-20% in AG and SF compared to PCNN alone. However, the values of

entropy and edge strength are well behind the other superior methods. Fusion performance is further improved when PCNN is utilized in the fusion process along with Convolutional sparse representation. This method provided better retention of structural details and intensity information in almost all the datasets. Although the performance of Latent LRR is satisfactory in terms of Entropy and MI, it lacks sharpness and spatial resolution in the fused images, resulting in lower values of other metrics. Among all existing methods, the multi-scale and multi-directional assessment ability of NSST, when combined with PAPCNN, showed a promising result as it provided excellent spatial detail and visual clarity with higher values of AG, SF, and  $Q_{ABF}$ . Activity level retention is also better in this method as it maintains stable MI values across all datasets. However, the proposed method consistently outperforms all existing methods in terms of all objective metrics. It provides higher values of AG and SF, indicating its ability to retain structural details and texture information of source modalities in the fused image. This results in improved anatomical structure visibility, which is essential for clinical diagnosis. It also provides a fused image with rich information content and structural fidelity because of its higher values of Entropy and edge strength compared to existing methods. A higher value of MI in the proposed method indicates its capability to preserve complementary information of source images in the fused image.

**Table 1:** Mathematical formulas of quality evaluation metrics

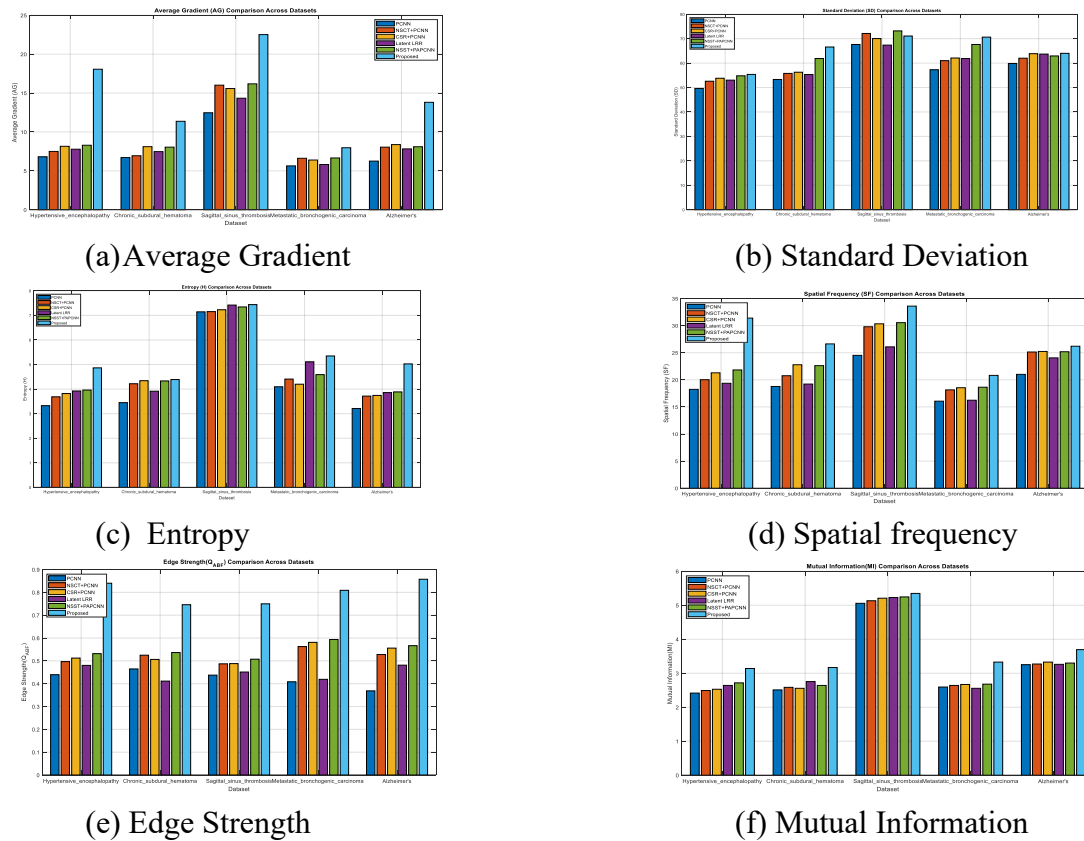
| Metric                                | Formulae                                                                                                                                                                                                                                  | Relevance                                                                                                                                                                                        |
|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Average Gradient (AG)                 | $AG = \frac{\sum_k \sum_l \sqrt{((F(k, l) - F(k + 1, l))^2 + (F(k, l) - F(k, l + 1))^2)}}{MN}$                                                                                                                                            | Evaluates the image's sharpness. A sharper image is indicated by a greater AG value. $F$ is the fused picture of size $MN$ in the equation.                                                      |
| Standard Deviation (SD)               | $SD = \sqrt{\frac{1}{MN} \sum_{k=1}^M \sum_{l=1}^N (f(k, l) - API)^2}$                                                                                                                                                                    | Evaluates the change in contrast within an image. A greater SD produces an image with more contrast.                                                                                             |
| Image Entropy ( $H$ )                 | $H = -\sum_{m=0}^{255} p_m \log_2 p_m$ , where $p_m$ is the probability of occurrence of gray level $m$ in the fused image                                                                                                                | $H$ is a metric that measures the amount of relevant information contained in an image. Images with higher entropy carry more information.                                                       |
| Spatial Frequency (SF)                | $SF(k, l) = \sqrt{ RF(k, l) ^2 +  CF(k, l) ^2}$ $RF(k, l) = \sqrt{\frac{1}{M \times N} \sum_{k=2}^M \sum_{l=2}^N [I(k, l) - I(k, l - 1)]^2}$ $CF(k, l) = \sqrt{\frac{1}{M \times N} \sum_{k=2}^M \sum_{l=2}^N [I(k, l) - I(k - 1, l)]^2}$ | SF assesses intensity changes in an image. An image having a higher SF contains more structural features.                                                                                        |
| Edge preservation index ( $Q_{ABF}$ ) | $Q_{ABF} = \frac{\sum_{k=1}^M \sum_{l=1}^N (Q_{AF}(k, l)W_A(k, l) + Q_{BF}(k, l)W_B(k, l))}{\sum_{k=1}^M \sum_{l=1}^N (W_A(k, l) + W_B(k, l))}$                                                                                           | Calculates the number of edges that are transferred from the input images to the fused image at the time of fusion. Any value around 1 suggests strong edge preservation, whereas a value near 0 |

|                         |                                                                                                                                                                                                                                                                                                                            |                                                                                                                                                                                                                                                                                         |
|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mutual Information (MI) | <p>For two source images <math>A, B</math>, and fused image <math>F</math>, Mutual information is given as <math>MI = MI(A, F) + MI(B, F)</math></p> $MI(A, F) = \sum_{z \in Z} \sum_{y \in Y} p(A, F) \log_2 \frac{p(A, F)}{p(A)p(F)}$ $MI(B, F) = \sum_{z \in Z} \sum_{y \in Y} p(B, F) \log_2 \frac{p(B, F)}{p(B)p(F)}$ | <p>indicates the worst edge preservation. It is assessed in the interval <math>[0,1]</math>. Indicates the amount of information exchanged between the input and fused images. A higher MI value suggests better transmission of active information from inputs to the fused image.</p> |
|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

The average values of each of the metrics taken over 25 slices of each dataset are presented graphically in Figure 12. The effectiveness of the proposed strategy is demonstrated by graphical analysis, which shows that it outperforms existing approaches with an average baseline improvement of 57%, 3%, 14.7%, 17.9%, 46%, and 14% in terms of AG, SD, H, SF,  $Q_{ABF}$ , and MI.

**Table 2:** Metrics for evaluating the quality of various approaches across various datasets

| Dataset                           | Method      | AG    | SD    | H    | SF    | $Q_{ABF}$ | MI   |
|-----------------------------------|-------------|-------|-------|------|-------|-----------|------|
| Hypertensive encephalopathy       | PCNN        | 6.95  | 55.14 | 3.50 | 19.83 | 0.453     | 2.57 |
|                                   | NSCT+PCNN   | 7.65  | 58.42 | 3.88 | 21.76 | 0.512     | 2.65 |
|                                   | CSR+PCNN    | 8.32  | 59.77 | 4.02 | 23.14 | 0.528     | 2.69 |
|                                   | Latent LRR  | 7.94  | 58.91 | 4.13 | 21.05 | 0.495     | 2.81 |
|                                   | NSST+PAPCNN | 8.46  | 60.88 | 4.17 | 23.72 | 0.548     | 2.89 |
|                                   | Proposed    | 18.43 | 61.51 | 5.12 | 34.12 | 0.866     | 3.34 |
| Chronic subdural hematoma         | PCNN        | 6.84  | 59.22 | 3.63 | 20.41 | 0.479     | 2.67 |
|                                   | NSCT+PCNN   | 7.08  | 61.97 | 4.44 | 22.56 | 0.541     | 2.75 |
|                                   | CSR+PCNN    | 8.27  | 62.51 | 4.57 | 24.76 | 0.522     | 2.72 |
|                                   | Latent LRR  | 7.62  | 61.46 | 4.12 | 20.89 | 0.424     | 2.93 |
|                                   | NSST+PAPCNN | 8.21  | 68.75 | 4.56 | 24.58 | 0.553     | 2.81 |
|                                   | Proposed    | 11.60 | 73.98 | 4.62 | 28.94 | 0.769     | 3.37 |
| Sagittal sinus thrombosis         | PCNN        | 12.72 | 75.12 | 7.52 | 26.65 | 0.451     | 5.38 |
|                                   | NSCT+PCNN   | 16.34 | 80.14 | 7.53 | 32.38 | 0.502     | 5.46 |
|                                   | CSR+PCNN    | 15.89 | 77.81 | 7.61 | 32.97 | 0.503     | 5.54 |
|                                   | Latent LRR  | 14.63 | 74.84 | 7.81 | 28.35 | 0.465     | 5.56 |
|                                   | NSST+PAPCNN | 16.51 | 81.30 | 7.73 | 33.20 | 0.523     | 5.58 |
|                                   | Proposed    | 22.97 | 78.98 | 7.83 | 36.52 | 0.773     | 5.69 |
| Metastatic bronchogenic carcinoma | PCNN        | 5.75  | 63.65 | 4.31 | 17.46 | 0.421     | 2.76 |
|                                   | NSCT+PCNN   | 6.75  | 67.77 | 4.64 | 19.72 | 0.580     | 2.81 |
|                                   | CSR+PCNN    | 6.52  | 68.98 | 4.42 | 20.15 | 0.599     | 2.84 |
|                                   | Latent LRR  | 5.93  | 68.74 | 5.38 | 17.65 | 0.432     | 2.72 |
|                                   | NSST+PAPCNN | 6.79  | 75.14 | 4.83 | 20.26 | 0.612     | 2.85 |
|                                   | Proposed    | 8.13  | 78.47 | 5.63 | 22.63 | 0.834     | 3.54 |
| Alzheimer's                       | PCNN        | 6.38  | 66.49 | 3.38 | 22.84 | 0.380     | 3.46 |
|                                   | NSCT+PCNN   | 8.21  | 68.89 | 3.91 | 27.32 | 0.544     | 3.48 |
|                                   | CSR+PCNN    | 8.54  | 70.94 | 3.94 | 27.43 | 0.573     | 3.54 |
|                                   | Latent LRR  | 7.98  | 70.76 | 4.06 | 26.14 | 0.496     | 3.47 |
|                                   | NSST+PAPCNN | 8.26  | 69.91 | 4.09 | 27.37 | 0.584     | 3.51 |
|                                   | Proposed    | 14.09 | 71.10 | 5.29 | 28.48 | 0.884     | 3.93 |



**Figure 12:** Performance analysis of different methods averaged over 25 slices of each dataset. For a pair of  $256 \times 256$  medical images, the average fusion time was roughly 1.42 seconds, as opposed to 2.85 seconds for NSST+PAPCNN and 3.64 seconds for CSR+PCNN.

The adaptive parameter estimation mechanism, which avoids deep model training and repetitive manual tuning, is primarily responsible for the time reduction.

## 5. Conclusion

In this work, a novel method of image decomposition is introduced using the mean of ordered pixel values filter, and is applied for multi-scale decomposition of input images of fusion. Further, a weight computation mechanism based on singular values of SVD is introduced to fuse coarse layers of source images, while the detail layers are fused at multiple levels with the help of parameter adaptive dual channel weighted PCNN (PADCWPCNN). The fusion performance of the proposed method is evaluated on 125 slices of diverse brain diseases using both subjective and objective criteria. Experimental results demonstrated the efficacy of the proposed method, as it improved AG, SD, H, SF,  $Q_{ABF}$ , and MI by 57%, 3%, 14.7%, 17.9%, 46%, and 14%, respectively, making it an appealing option for fusing medical images to effectively diagnose tumors and plan treatment accordingly.

## References

- [1] A. P. James and B. V. Dasarthy, "Medical image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 19, pp. 4–19, 2014, doi: 10.1016/j.inffus.2013.12.002.
- [2] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk, *Positron Emission Tomography*, vol. 2. New York, NY, USA: Springer, 2005.
- [3] S. K. Panguluri and L. Mohan, "Discrete wavelet transform-based image fusion using unsharp masking," *Periodica Polytechnica Electr. Eng. Comput. Sci.*, vol. 64, no. 2, pp. 211–220, 2020.

- [4] J. Bhardwaj, A. Nayak, and D. Gambhir, "Multimodal medical image fusion based on discrete wavelet transform and genetic algorithm," in *Proc. Int. Conf. Innov. Comput. Commun.*, Singapore, 2021, pp. 1047–1057.
- [5] M. Kumar, N. Ranjan, and B. Chourasia, "Hybrid methods of contourlet transform and particle swarm optimization for multimodal medical image fusion," in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Coimbatore, India, 2021, pp. 945–951.
- [6] S. Kollem, K. R. Reddy, and D. S. Rao, "Improved partial differential equation-based total variation approach to non-subsampled contourlet transform for medical image denoising," *Multimedia Tools Appl.*, vol. 80, no. 2, pp. 2663–2689, 2021.
- [7] M. Khare, M. Khare, and R. Srivastava, "Shearlet transform-based technique for image fusion using median fusion rule," *Multimedia Tools Appl.*, vol. 80, no. 8, pp. 11491–11522, 2021.
- [8] Z. Wang, Z. Cui, and Y. Zhu, "Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation," *Comput. Biol. Med.*, vol. 123, art. no. 103823, 2020.
- [9] F. Liu et al., "Medical image fusion method using Laplacian pyramid and convolutional sparse representation," *Concurrency Comput.: Pract. Exp.*, vol. 32, no. 17, art. no. e5632, 2020.
- [10] S. Singh and R. S. Anand, "Multimodal medical image sensor fusion model using sparse K-SVD dictionary learning in non-subsampled shearlet domain," *IEEE Trans. Instrum. Meas.*, vol. 69, pp. 593–607, 2019.
- [11] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Aug. 2010, pp. 663–670.
- [12] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Barcelona, Spain, 2011, pp. 1615–1622.
- [13] J. Xia, Y. Lu, and L. Tan, "Multimodal medical image fusion based on parameter-adaptive pulse-coupled neural network and convolutional sparse representation," *Comput. Math. Methods Med.*, vol. 2020, art. no. 3290136, pp. 1–13, 2020, doi: 10.1155/2020/3290136.
- [14] L. Li and H. Ma, "Pulse-coupled neural network-based multimodal medical image fusion via guided filtering and WSEML in NSCT domain," *Entropy*, vol. 23, no. 5, art. no. 591, 2021, doi: 10.3390/e23050591.
- [15] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "FusionGAN: A generative adversarial network for infrared and visible image fusion," *Inf. Fusion*, vol. 48, pp. 11–26, 2019.
- [16] J. Liu, Y. Li, X. Sun, X. Wang, and H. Luo, "MMIF: Multimodal medical image fusion network based on multi-scale hybrid attention," *Comput. Mater. Continua*, vol. 85, no. 2, pp. 3551–3568, 2025.
- [17] Y. Shi et al., "VDMUFusion: A versatile diffusion model-based unsupervised framework for image fusion," *IEEE Trans. Image Process.*, 2024.
- [18] M. V. Srikanth, A. S. Kumar, B. Nagasirisha, and T. Lakshmi, "Brain MRI and CT image fusion using multiscale local extrema and image statistics," *ECTI Trans. Electr. Eng. Electron. Commun.*, vol. 22, no. 1, 2024, doi: 10.37936/ecti-eec.2024221.249146.
- [19] Y. Chai, H. F. Li, and J. F. Qu, "Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain," *Opt. Commun.*, vol. 283, no. 19, pp. 3591–3602, 2010.
- [20] C. Panigrahy et al., "Differential box counting methods for estimating fractal dimension of gray-scale images: A survey," *Chaos Solitons Fractals*, vol. 126, pp. 178–202, 2019.
- [21] Y. Chen, S.-K. Park, Y. Ma, and R. Ala, "A new automatic parameter setting method of a simplified PCNN for image segmentation," *IEEE Trans. Neural Netw.*, vol. 22, no. 6, pp. 880–892, Jun. 2011.
- [22] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
- [23] V. P. S. Naidu, "Image fusion technique using multi-resolution singular value decomposition," *Defence Sci. J.*, vol. 61, no. 5, pp. 479–484, 2011.
- [24] S. M. V. et al., "Brain tumor detection through image fusion using cross guided filter and convolutional neural network," *ECTI-CIT Trans.*, vol. 18, no. 4, pp. 579–590, Oct. 2024.