Al-Nahrain Journal of Science
ANJS, Vol. 29 (1), March, 2026, pp. 150 – 165

Image Fusion for Brain Tumor Diagnosis Using Fractal Weighted Adaptive Dual Channel Neural Network

Shailaja Mantha[1], Jatothu Brahmaiah Naik[2,*], Kanagala Sateesh Kumar[1], Rajanidevi[3]

[1]Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, India.
[2]Department of Electronics and Communication Engineering, Narasaraopeta Engineering College, India
[3]Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, India

Article's Information

Received: 08.09.2025
Accepted: 24.01.2026
Published: 15.03.2026

Keywords:
Image fusion,
local extrema,
Enhanced Weighted Pulse,
Dual Channel Neural Network,
Fractal Dimension,
Differential Box Counting,
Brain imaging.

Abstract

This research paper proposes a novel image fusion algorithm for assisting effective diagnosis of brain cancers. The method extracts detailed features from source images using a multi-level decomposition strategy that leverages local extrema information. The detailed layers are fused using an Enhanced Weighted Pulse Adaptive Dual Channel Neural Network (EWPADCNN), with fusion weights calculated using the Fractal Dimension with Differential Box Counting (FDDBC) method. In order to preserve low-frequency information, the base layers uses a weighted average method based on pixel significance. Comprehensive tests on 100 slices from four different datasets of brain disorders demonstrates that the proposed approach outperforms current fusion methods in both qualitative and quantitative assessments. These results support the suggested method as an efficient and trustworthy technique for improving brain imaging diagnostic quality. Quantitative evaluation demonstrates average improvements of 41%, 5%, 16%, 23%, 34%, and 4% in terms of AG, H, SF, MI, Q_{ABF}, and SD metrics, respectively, compared to existing fusion methods.

http://doi.org/10.22401/ANJS.29.1.14
*Corresponding author: brahmaiahnaik@gmail.com

1. Introduction

Medical imaging is a vital tool for contemporary medical diagnosis and care. As medical imaging technologies advance, a variety of unique medical images have emerged, produced by various devices and principles. There are distinct features and information contained in various medical imaging modalities. Positron Emission Tomography (PET), Computed Tomography (CT), and Magnetic Resonance Imaging (MRI) are common medical imaging methods. While MRI images give superior soft tissue resolution for evaluating soft tissues, CT images demonstrate transparent anatomical connections and are particularly good at detecting tough structures like bones and calcifications. A single modality image, however, frequently fails to provide enough information. Doctors frequently need to continually review and compare medical images of the same location from multiple methods and sources to make an accurate diagnosis. Their workload is increased, and consultations inefficient as a result of this procedure. This is addressed by multimodal medical image fusion technology, which combines images of the same object from two or more different modalities. By preserving as much information as possible from several modality source pictures, the fused image enhances the visibility of organs or lesion areas [1–3]. The image fusion strategy has

been separated into three levels: feature level, decision level, and pixel level [4]. Although feature-level fusion lowers an image fusion network's computational complexity, some fine-grained information is still lost throughout the fusion process. Similarly, decision-level fusion exhibits greater adaptability and real-time performance, it causes the largest loss of visual information and contributes more errors to experimental outcomes. The most popular image fusion technique currently is pixel-level fusion, which combines information from individual pixels. Deep learning fusion methods and conventional fusion methods are the two categories of image fusion algorithms that are currently available. Conventional fusion techniques often employ mathematical adjustments pertinent to the frequency or spatial domain to achieve image fusion objectives. Multi-scale transformation (MST) is a common conventional image fusion technique. This approach fully utilizes the multiscale information of the image and may adapt to various fusion rules to provide the desired image fusion outcome. Wavelet transform [5], pyramid transform [6], non-subsampled contourlet transform (NSCT) [7], and non-subsampled shearlet transform (NSST) [8] are examples of common MST fusion techniques. The NSCT and NSST techniques split an input image into high and low frequency sub-bands, in contrast to the straight forward wavelet transform decomposition. They successfully capture the unique aspects of images. However, in order to reduce computational complexity while maintaining fundamental visual cortex features, Johnson designed a pulse-coupled neural network (PCNN) model [9–10]. Furthermore, the model's inability to precisely specify parameters severely limited the advancement of PCNN. To efficiently extract images, a framework of NSST (Non-Subsampled Shearlet Transform) is suggested, and PCNN was incorporated into fusion rules [11]. A parametric adaptive PAPCNN model is created by improving the SPCNN model [12]. These studies demonstrated a considerable improvement in the PAPCNN model's convergence speed, and the use in image fusion also produced better experimental results. The global structure of the original image was captured by the suggested low-rank representation (LRR) image fusion technique, but its capacity to retain local structure was limited since LRR does not use dictionary learning [13].In order to concurrently address the fundamental shortcomings of both MST- and SR-based fusion techniques, a general image fusion framework is presented by integrating MST and SR [14]. To determine the low-pass and high-pass coefficients for each of the pre-registered source pictures, MST is applied. After that, an SR-based fusion technique is utilized to merge the low-pass bands, and the absolute values of the coefficients are used to fuse the high-pass bands as a measure of activity level. Finally, by applying the inverse MST to the joined coefficients, the fused image is produced. The non-subsampled contourlet transform (NSCT) in conjunction with parameter-adaptive pulse-coupled neural networks (PA-PCNN) is used to propose a novel multimodal MIF technique. The NSCT separates the images into bands of high and low frequencies. In PA-PCNN, those bands are combined. NSCT's inverse method was employed to produce the fused image [15]. The clarity of fused image detail information was enhanced by a weighted averaging technique and a summation strategy that fused global and local structure information, respectively [16].

Considering the issues with the popular image fusion methods of today, an enhanced fusion technique, the MLE-EWPADCNN-Pixel significance algorithm, is proposed in this study based on the PAPCNN model. The work makes the following contributions:

1. An Enhanced Weighted Pulse Adaptive Dual Channel Neural Network (EWPADCNN) model is developed with weights calculated using the Fractal Dimension using Differential Box Counting (FDDBC) method.
2. A multi-level decomposition strategy based on local extrema information is utilized.
3. The limited precision of free parameter setting in the conventional PCNN model is resolved by introducing the EWPADCNN model into the multi-scale transform high-frequency partial fusion.
4. A weighted average method based on pixel significance is used for fusing base layers.

Unlike conventional PCNN models that employ a single global linking strength for all pixels, and PAPCNN-based methods that rely on heuristic or threshold-based parameter adaptation, the proposed EWPADCNN introduces a pixel-wise adaptive linking mechanism guided by fractal dimension analysis. Furthermore, existing NSST-PCNN approaches focus primarily on transform-domain enhancement, whereas the proposed model integrates adaptive neural dynamics with multilevel

local extrema decomposition. The dual-channel weighted fusion pool in EWPADCNN enables modality-specific information regulation, resulting in superior edge preservation and information fidelity compared to existing PCNN-based fusion frameworks.

2. Related work

2.1. Multiscale Decomposition using Local Extrema

Every image contains pixels with homogeneous intensity, which corresponds to base layers, and pixels with an abrupt change in intensity, which corresponds to detail layers. As seen in Figure 1, the base layer can be produced by taking the mean of the local minimum and maximum envelopes, whilst the detail layers could be imagined as oscillations between such envelopes [17]. For an Image $I(p,q)$ with spatial coordinates(p,q), the process of finding the base and detail layers is as described: Every pixel along four borders of I is regarded as an extrema point with either L_{max} or L_{min} . The extrema envelopes E_{max} and E_{min} are created using Delaunay interpolation on these extrema pixels. Taking the mean of E_{max} and E_{min} yields the base layer M, while subtracting it from the original image I produces the detail layer. Assume that input images A and B were acquired using various imaging modalities. The local extrema envelope method is applied at various scales to split the source images into a base layer and many detail layers [21], while the detail layers depict high-frequency details like textures and edges at various levels of the wx w sample window. The base layer is an estimate of the original image. Figure 2 illustrates the mechanism involved in the multiscale decomposition of an image using local extrema information. The size of the sampling window is taken as $w = 2L + 1$, where $L = 1,2,3\ldots$.N indicates the scale number. At scale L, the window is projected onto the image with the current pixel of the input image being processed occupying the centre of the window, and the response is computed. The process is repeated using a sliding window process to cover all the pixels of an image. Once all the pixels are processed using a window, the resultant image contains only homogeneous pixels, and hence it is called the base layer B_L at scale L. By taking the average of the base layer at different scales, the final base layer of an input image can be generated as specified in equations (1).

Overall base layer,

$$M = \frac{1}{N}\sum_{L=1}^{N} B_L \quad \ldots(1)$$

By subtracting the base layer B_L from the original image, a detail layer D_L at scale L can be obtained. The overall detail layers at each scale can be computed using equation (2).

Detail layer at level L,

$$d_L = D_L + (B_L - M) \quad \ldots(2)$$

Using this mechanism, each source image of the fusion process is decomposed into base and multiscale detail layers in the proposed work.

2.2. Fractal Dimension using Differential Box Counting Method (FDDBC)

A grayscale image $I$of size NxN is divided into non-overlapping strips of size mxm pixels with scale $s = \frac{m}{N}$ in the xy-plane, where $2 \le m \le \frac{N}{2}$ .The process described in [18] is used to calculate the number of boxes n_s required to cover the image surface of each strip at scale s.

$$n_s = L_{max} - L_{min} + 1 \quad \ldots(3)$$

Here, L_{max} and L_{min} are the maximum and minimum gray values in the corresponding strip. Equation (4) is used to get the total number of boxes at scale 's' needed to cover an image I.

$$N_s = \sum n_s \quad \ldots(4)$$

Lastly, linear least squares regression is used to fit the points $\left(log\left(\frac{1}{s}\right), log\left(\frac{1}{N_s}\right)\right) \forall\, s$ into a line, whose slope is referred to as the fractal dimension of an image I.

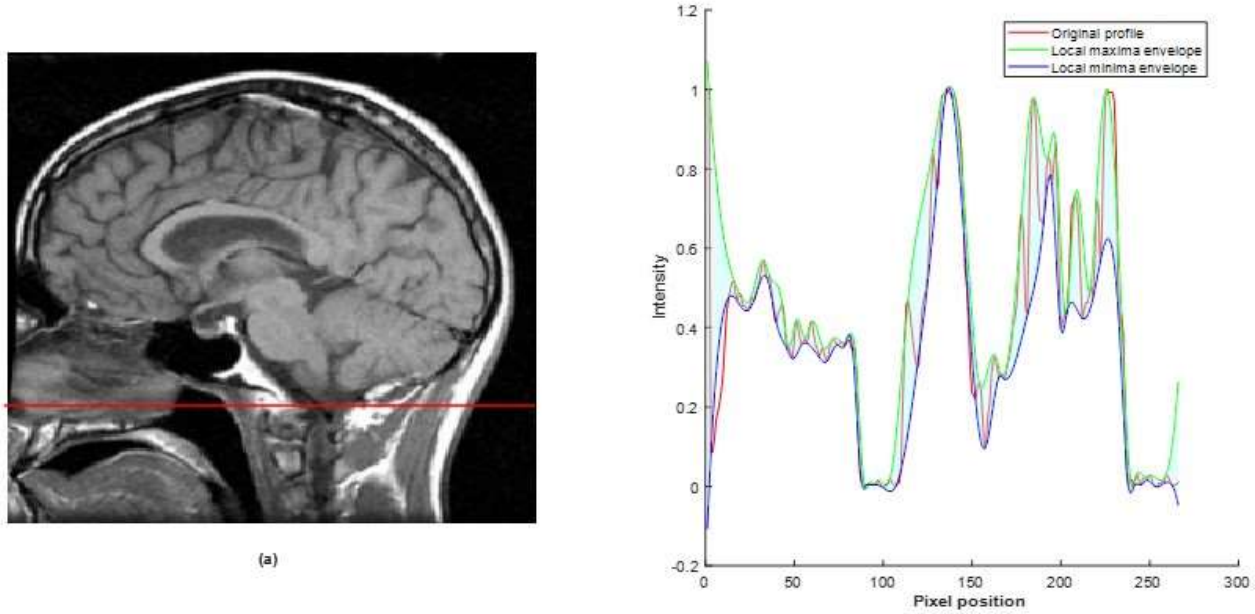

Figure 1. Local extrema distribution of a particular row in an image: (a) Image with particular row highlighted (b) Local extrema distribution of highlighted row

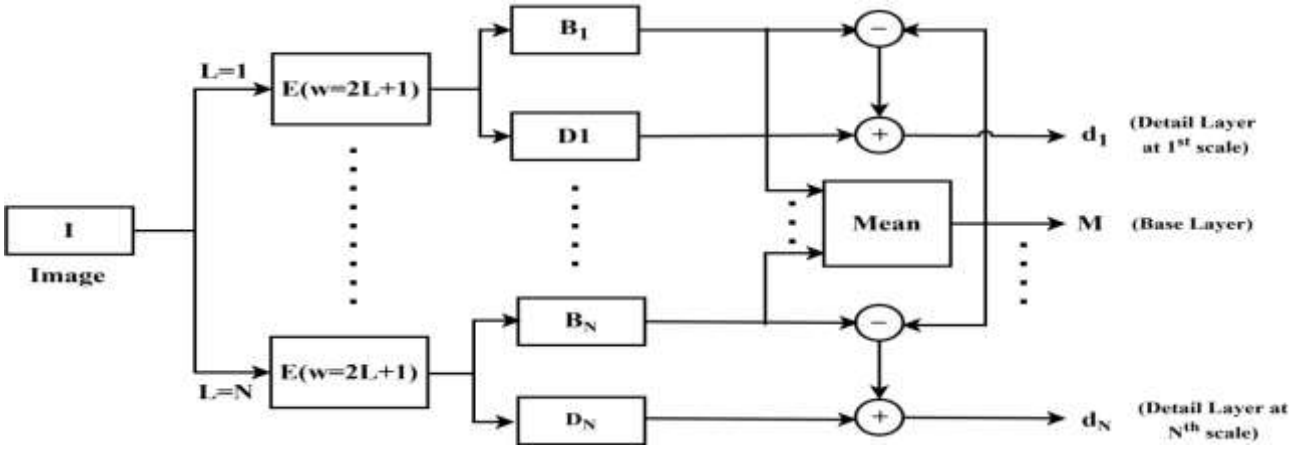

Figure 2. Multi-level decomposition structure using local extrema

2.3. Novel PCNN Model (EWPADCNN)

The proposed dual-channel PCNN model with adaptive weights based on FDDBC is shown in Figure 3. It mainly contains three fields: receptive field, information fusion pool, and pulse generator. The receptive field receives stimuli from nearby neurons and external inputs. The information fusion pool fuses source images based on neuron activity levels. The pulse generator generates the output pulse. The number of input neurons in the proposed model is equal to the number of pixels in an image. Mathematical modeling of the proposed model is described using equations (5) to (12).

$$F_n^A(p,q) = S^A(p,q) \quad \dots (5)$$

$$F_n^B(p,q) = S^B(p,q) \quad \dots (6)$$

$$L_n(p,q) = C_s \sum_{i=-1}^{1} \sum_{j=-1}^{1} W_s(i+2, j+2) Y_{n-1}(p+i, q+j) \quad \dots (7)$$

$$U_n^A(p,q) = F_n^A(p,q)\left(1+\beta^A(p,q)L_n(p,q)\right) \quad \dots(8)$$

$$U_n^B(p,q) = F_n^B(p,q)\left(1+\beta^B(p,q)L_n(p,q)\right) \quad \dots(9)$$

$$U_n(p,q) = e^{-\alpha_u}U_{n-1}(p,q) + max\{U_n^A(p,q), U_n^B(p,q)\} \quad \dots(10)$$

$$Y_n(p,q) = \begin{cases} 1 & if\ U_n(p,q) > \theta_{n-1}(p,q) \\ 0 & otherwise \end{cases} \quad \dots(11)$$

$$\theta_n(p,q) = e^{-\alpha_\theta}\theta_{n-1}(p,q) + C_\theta Y_n(p,q) \quad \dots(12)$$

where $S^A(p,q)$ and $S^B(p,q)$are the external stimuli of the proposed PCNN model at (p,q) location of the source images$A$ and B, respectively. Further$S^A(p,q) \in [0,1]$and$S^B(p,q) \in [0,1]$. $F_n^A(p,q)$ and $F_n^B(p,q)$ are the symmetric feeding inputs that store $S^A(p,q)$ and $S^B(p,q)$ respectively and do not change in every iteration.$L_n(p,q)$is the linking parameter of $(p,q)^{th}$ neuron in n^{th}iteration. The synaptic weights of eight nearby neurons of the $(p,q)^{th}$ neuron are represented by the 3x3 matrix W_s.

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

$U_n^A(p,q)$ and $U_n^B(p,q)$reflect the internal state of the $(p,q)^{th}$ neuron after n^{th}iteration for images$A$ and B, respectively, which together make up the non-linear modulation of linking and feeding inputs.$U_n(p,q)$signifies the internal activity of $(p,q)^{th}$ neuron of the proposed PCNN model after n^{th} iteration. $\theta_n(p,q)$ and $Y_n(p,q)$ are the dynamic threshold and external output of $(p,q)^{th}$ after n^{th}iteration respectively.$\beta^A(p,q)$and$\beta^B(p,q)$ are the linking intensities corresponding to $S^A(p,q)$ and $S^B(p,q)$. The response of the visual system in humans to a pixel and its surrounding areas is represented by linking intensity and cannot be equal for the $(p,q)^{th}$ pixel of two source images acquired with distinct modalities. It is incorrect to assume that all pixels in an image have the same linking intensity. The suggested PCNN model uses adaptive linking strength, assigning distinct values to different pixels in the image, as opposed to the traditional PCNN model that uses a single linking strength for all pixels.

The FDDBC employs $(L_{max} - L_{min} + 1)$ boxes to indicate gray-level variations on a strip. Therefore, this information can be used to calculate the linking intensity. To estimate the linking intensity of a pixel, place the strip on the image such that the pixel occupies the center of the strip while determining $(L_{max} - L_{min} + 1)$. The linking intensity of $(p,q)^{th}$ pixel can be calculated using equation (13).

$$\beta^X(p,q) = 1 - exp\left(-0.02\left(L_{max}^X(p,q) - L_{min}^X(p,q)\right)\right)\ where\ X \in (A,B) \quad \dots(13)$$

C_θ and C_s represent the amplitudes of dynamic threshold and linking input, respectively, while $e^{-\alpha_\theta}$ and $e^{-\alpha_u}$ are the exponentially decaying coefficients. A lower α_u results in a wider distribution range of U_n, which can be accurately represented by a logarithmic function [19]. Therefore, equation (14) can be used to estimate the value of α_u.

$$\alpha_U = log\left(\frac{1}{w_1\sigma_{S^A} + w_2\sigma_{S^B}}\right) \quad \dots(14)$$

Here σ_{S^A} and σ_{S^B} are the standard deviations of S^A and $S^{B,}$ while w_1 and w_2 are weights computed using fractal dimensions (FD) of source images A and B evaluated using the FDDBC method.

$$w_1 = \frac{FD_A}{FD_A + FD_B} \quad \dots(15)$$

$$w_2 = \frac{FD_B}{FD_A + FD_B} \quad \dots(16)$$

Inspired by the work of [19], the values of α_θ and C_θ are modeled using equations (17) and (18). $S^{A'}$and $S^{B'}$ are the Otsu thresholds of S^A and $S^{B,}$ respectively [20]

$$C_\theta = e^{-\alpha_u} + 1 + 6\lambda \quad \dots(17)$$

$$\alpha_\theta = ln\left(\frac{C_\theta}{\left(w_1S^{A'} + w_2S^{B'}\right)\left(\frac{1-e^{-3\alpha_u}}{1-e^{-\alpha_u}} + 6\lambda e^{-\alpha_u}\right)}\right) \quad \dots(18)$$

$$\lambda = \frac{\left(w_1S^{B'}S_{max}^A + w_2S^{A'}S_{max}^B - S^{A'}S^{B'}\right)}{6S^{A'}S^{B'}} \quad \dots(19)$$

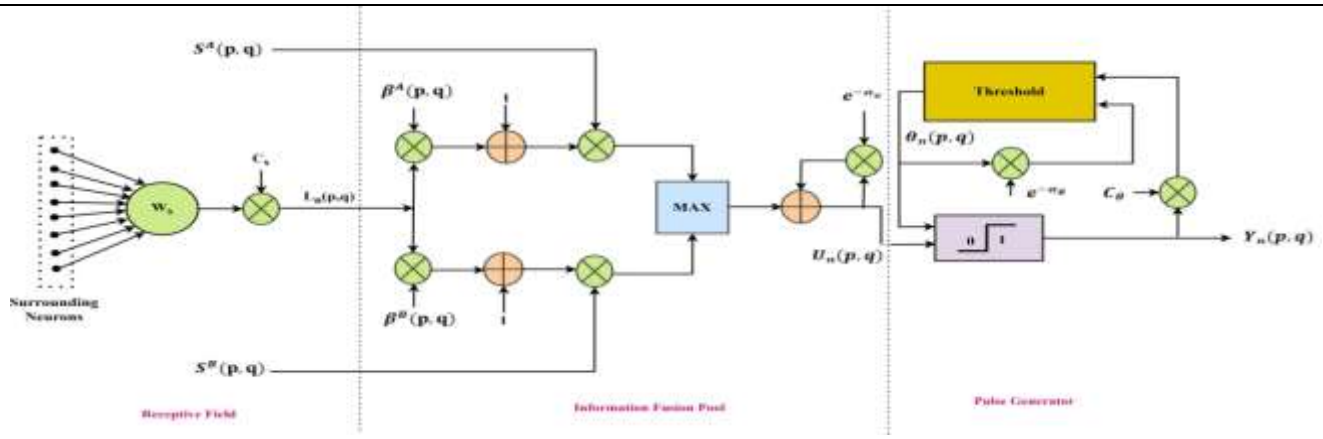

Figure 3. Structure of Enhanced Weighted Pulse Adaptive Dual Channel Neural Network

The suggested EWPADCNN introduces a pixel-wise adaptive linking mechanism that dynamically modifies neuron interactions based on local image complexity, in contrast to the traditional PCNN model, which uses a single globally fixed linking strength for all pixels and thus struggles to adapt to spatially varying image characteristics. By adjusting parameters using threshold or heuristic criteria, PAPCNN-based fusion techniques partially overcome this restriction; nevertheless, these adjustments are still indirectly driven and susceptible to initialization and noise circumstances. On the other hand, the suggested model uses the FDDBC approach to directly extract adaptive linking strengths from fractal dimension information, allowing for a content-aware and mathematically grounded modulation of brain activity. Additionally, while maintaining fixed or semi-adaptive neural parameters, current NSST-PCNN techniques mainly rely on transform-domain decomposition to improve performance. Chombining fractal-guided adaptive dynamics with dual-channel weighted information fusion, the suggested EWPADCNN deviates from this paradigm and enables modality-specific neural response regulation. This combination clearly outperforms PCNN, PAPCNN, and NSST-PCNN based fusion frameworks in terms of both architecture and functionality by greatly enhancing edge preservation, information consistency, and resilience across multimodal medical images.

2.4. Weight Computation based on Pixel Significance

Since the base layer of an image contains pixels with homogeneous gray scale intensities, they are fused using a weighted average rule from the weights of pixels of each base layer computed based on their significance. To compute the weight of each pixel, statistical property like the co-variance of its neighborhood pixels is taken into consideration. For a pixel of image M located at (p, q), which is denoted as $M(p, q)$, a window of size N x N is projected such that the pixel occupies the center of the window. The neighborhood of the pixel covered by the window is represented as a matrix G . The unbiased approximation of the co-variance matrix of G along the horizontal direction, $C_h(p, q)$ is evaluated by considering each row as a statistic while taking the column as a parameter, as illustrated in equations (20) and (21).

$$Cov(G) = E[(G - \bar{G})(G - \bar{G})^T] \quad \ldots(20)$$

$$C_h(p, q) = \frac{\sum_{l=1}^{N}(g_l - \bar{g})(g_l - \bar{g})^T}{(N - 1)} \quad \ldots(21)$$

Here g_l is the l^{th} observation of g and $\bar{g}$ is the mean of N observations. It is noted that a vector of variances for each column of matrix G is provided by the diagonal of matrix $C_h(p, q)$. Now the Eigen values of matrix $C_h(p, q)$ are evaluated, and the sum of these Eigen values directly reflects the horizontal weight of a pixel (p, q). Similarly, unbiased approximation of the co-variance matrix of G along the vertical direction, $C_v(p, q)$, is evaluated by considering each column as a statistic while taking the row as a parameter. The sum of Eigen values of $C_v(p, q)$directly reflects the vertical weight of a pixel (p, q). The overall weight of the pixel (p, q) is the sum of its vertical and horizontal weights [21].

3. Proposed Algorithm of Fusion

In the proposed algorithm, if the two source images are gray-scale images like MRI and CT, they are

directly applied to the first stage of the fusion process, which is the multilevel decomposition of the source images. If one of the source images is an RGB color image like PET or SPECT, then it is first converted from RGB to YUV color space, and the intensity component 'Y' is utilized for the decomposition in the fusion process. The detailed process flow of the proposed method is illustrated in Figure 4 for gray-scale images and in Figure 5 for a color image fused with gray gray-scale image. The step-by-step process of fusion methodology is as follows:

1. Input images are labeled as A and B and are re-sized to the same dimension (256x256)
2. Decompose each input image into base layer and detail layer at multiple levels using a multilevel local extrema procedure (in this, four levels of decomposition are employed, which means L=4 in figure 2).
3. Base layers at each level are labeled as B_{1A},B_{2A},$B_{3A,}$ and B_{4A} for image A, while the base layers of image B are labeled as B_{1B},B_{2B},$B_{3B,}$ and B_{4B}.
4. Detail layers at each level are labeled as D_{1A},D_{2A},$D_{3A,}$ and D_{4A} for image A, while the base layers of image B are labeled as D_{1B},D_{2B},$D_{3B,}$ and D_{4B}.
5. From the base layers at different levels, the overall smoothed base layer of each source image is computed using equation (1).

   Overall base layer of image,

$$M_A = \frac{1}{4}\sum_{L=1}^{4} B_{LA}$$

   Overall base layer of image,

$$M_B = \frac{1}{4}\sum_{L=1}^{4} B_{LB}$$

6. From the detail layers at different levels along with corresponding base layers and overall base layer, the overall saliency detail layer of each source image is computed using equation (2).

   Overall detail layer at level L of image A,

$$d_{LA} = D_{LA} + (B_{LA} - M_A) \;\; for\; L = 1,2,3,4$$

   Overall detail layer at level L of image B,

$$d_{LB} = D_{LB} + (B_{LB} - M_B)$$

7. Fusion of base layers: Compute the weights of each pixel of a base layer using the weight computation mechanism discussed in section 3.4 and label them as W_A and W_B.

$$M_f(p,q) = \frac{(W_A(p,q) * M_A(p,q)) + (W_B(p,q) * M_B(p,q))}{(W_A(p,q) + W_B(p,q))}$$

8. Fusion of detail layers: The proposed novel PCNN model in section 3.3 is used to fuse detail layers at different levels. For this purpose, detail layers at a particular level of the source image are taken as external stimuli to PCNN, and corresponding internal activity $U_n(p,q)$ is computed as follows:

$$S^A(p,q) = d_{LA}(p,q) \quad and \quad S^B(p,q) = d_{LB}(p,q)$$

   Using equations from (5) to (12), $U_n^A(p,q)$ and $U_n^B(p,q)$ are computed, and then detail layers are fused using the maximum rule of fusion.

$$map(p,q) = \begin{cases} 1 \; if \;\; U_n^A(p,q) \geq U_n^B(p,q) \\ 0 \; if \; U_n^A(p,q) < U_n^B(p,q) \end{cases}$$

   Fused detail layer at each level,

$$d_{Lf}(p,q) = map(p,q) * d_{LA}(p,q) + \big(1 - map(p,q)\big) * d_{LB}(p,q)$$

9. The overall fused image is produced by combining the fused base layer and the fused detail layer.

$$Fused\ image(p,q) = M_f(p,q) + d_{Lf}(p,q)$$

In the case of MRI and PET/SPECT fusion, the color channels U and V are directly passed to the output and are added to the fused image to produce a fused image in YUV color space. It is brought back to RGB color space using an appropriate conversion.

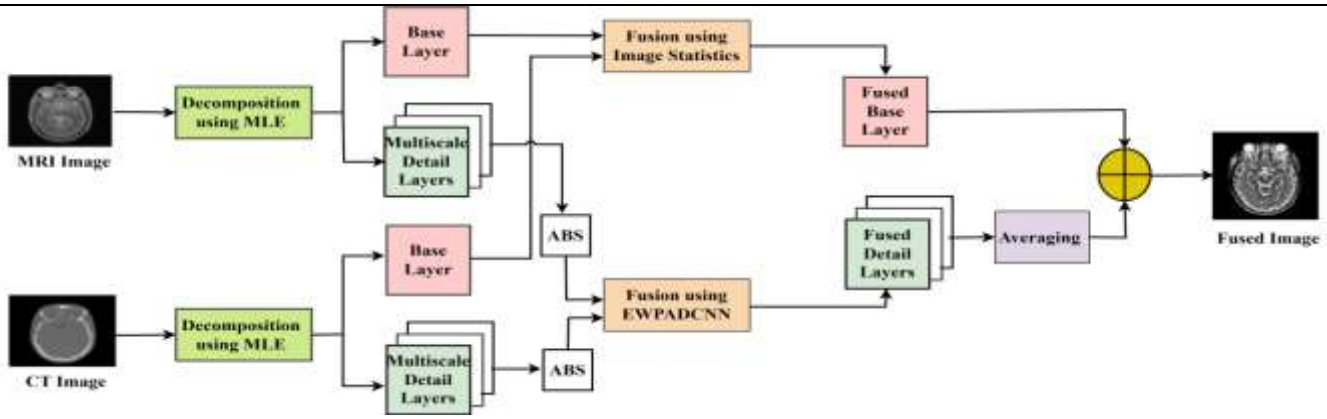

Figure 4. Process flow of fusing CT and MRI images

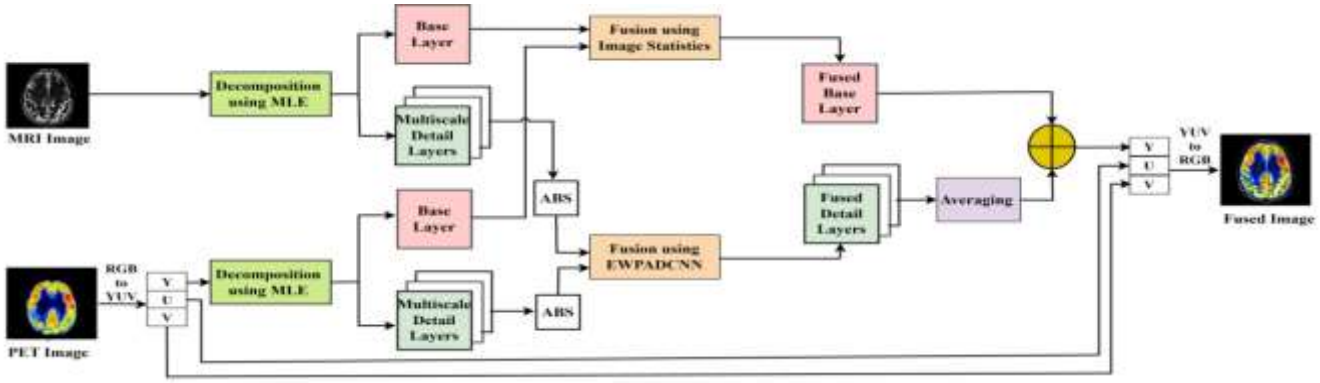

Figure 5. Process flow of fusing MRI and PET/SPECT images

In all experiments, the EWPADCNN model was initialized with identical parameters for fair comparison. The maximum number of iterations was set to 30, and convergence was achieved when the change in neuron firing activity between successive iterations fell below10^{-4}. The decay constantsα_θ and α_L were empirically set to 0.2 and 0.1, respectively. Threshold amplitudes and linking coefficients were adaptively adjusted using FDDBC-based fractal measures, ensuring stability and consistent convergence across datasets.

4. Results and Discussion

The source images utilized for experimentation in the proposed work are benchmark images that are widely used with no impact on human rights or race. All source images were obtained from the standard database of images of various brain illnesses, available at https://www.med.harvard.edu/aanlib/. Additionally, all source images are obtained from the same slice at the same angle using various imaging techniques. To confirm the exceptional performance of the algorithm, 100 slices of source images from four main brain illnesses are chosen for testing. Twenty-five slices are taken from each of the four datasets. Prior to usage, the chosen source images were precisely matched and spatially resolved at 256 x 256 pixels. CT and MRI scanned images of superior sagittal sinus thrombosis (SSST) are included in the first dataset. It is a blood clot that developed in the brain's main vein, the superior sagittal sinus, which empties blood from the brain to the venous circulation. Increased intracranial pressure may result from this, raising the risk of severe neurological side effects such as coma, seizures, and stroke. The second dataset corresponds to MR-T1 and MR-T2 images of chronic subdural hematoma (CSDH) illness. It is a kind of cerebrovascular illness affecting the brain. It is a gradual accumulation of blood and blood breakdown products that accumulates between the brain and its outer covering. The third dataset represents PET and MR images of Alzheimer's disease. It is a progressive neurodegenerative ailment that predominantly affects the brain, resulting in a gradual loss of memory, thinking, and other cognitive skills. It is the

most prevalent cause of dementia, resulting in behavioral and psychological changes which eventually creates the inability to perform basic tasks. The fourth dataset corresponds to SPECT and MR images of Metastatic bronchogenic carcinoma of the brain. It is caused by lung tumor cancer cells spreading to the brain, because of the increased pressure and disturbance of brain tissue function, this may result in the formation of one or more brain tumors, which can produce a range of symptoms.

4.1. Qualitative Analysis

Fusion results of the above datasets are illustrated in Figures 6 to 9. In each figure, the source images are presented in (a) and (b), while the fused images of different existing methods, along with the proposed method are depicted from (c) to (h). Figure 6 contains the fusion outcomes of the superior sagittal sinus thrombosis dataset images. As shown in Figure 6(c). Although the PCNN method could retain soft tissue and bony structures of the input images, it introduces blur in the region of soft tissue, resulting in a low contrast image. The NSCT-PCNN method could provide better retention of information around the central brain area than PCNN, as shown in Figure 6(d). This introduces some blur and halos near edges. From Figure 6(e) it is obvious that the MST-SR method can preserve edges much better than other methods, but at the same time, it dilutes the structural details of source images in the fused image. Even though Latent LRR provides a fused image with improved contrast by retaining the bony and soft tissue information, it introduces over-enhancement into the fused image, as illustrated in Figure 6(f). From the results of Figure 6(g), NSST-PCNN provides balanced fusion output by providing an acceptable contrast with adequate fine details, but it smoothes the peripheral area of the fused image. The proposed method, on the other hand, provides a fused image with good texture retention without introducing excess sharpening and balances the retention of MRI and CT features in the fused image. It also introduces fewer artifacts than other methods, as shown in Figure 6(h). Fusion results of chronic subdural hematoma illness for MR-T1 and MR-T2 scanned images are presented in Figure 7. As seen from figures 7(a) and 7(b), the MR-T1 scan highlights the anatomical details of an organ, while the MR-T2 scan preserves the pathological information of the organ. The problem of blur persists with the PCNN method, along with loss of some information from the MR-T2 image during the fusion process, as shown in Figure 7(c). From the fusion result of NSCT-PCNN as illustrated in Figure 7(d), it is evident that the method introduces haloes in the deep brain regions during the fusion process, even though it is good at preserving structural information better than the PCNN method. MST-SR method provides a fused image with acceptable tissue and structural preservation, but introduces inconsistency in the intensity, resulting in an over-sharpened boundary in the fused image as depicted in Figure 7(e). The fused image in figure 7(f) corresponding Latent LRR method preserves peripheral structures quite well, but the brightness looks artificial, making the center area of the fused image appear less natural. Though the fused image of the NSST-PCNN approach is better, it introduces mild smoothening in the center region of the image, as seen in Figure 7(g). The fused image of the proposed method, as shown in Figure 7(h), contains excellent preservation of anatomical details of MR-T1 as well as pathological features of MR-T2 with fewer artifacts than other methods. Fusion results of PET and MRI scanned images of Alzheimer's disease are shown in Figure 8. As is obvious from Figures 8(a) and 8(b), a PET scan highlights the metabolic activity of an organ, while the MRI scan preserves the soft tissue information. While the PCNN method provides a fused image with moderate preservation of MRI features with decent retention of PET details, the problem of blur and low contrast persists in the fused image shown in Figure 8(c). The NSCT-PCNN method could provide a fused image with better energy localization, but it introduces structural distortion as seen in Figure 8(d). From the fused image of Figure 8(e), it is identified that the MST-SR process provides output with overexposed PET details while reducing the MRI details. The latent LRR method introduces intensity blending near edges, as is obvious from Figure 8(f). The NSST-PCNN approach provides the fused image with consistent details but contains mild softness near edges as observed in Figure 8(g). The proposed method provides a fused image with balanced metabolic and anatomical details with less distortion and sharper edges than other methods, as noticed from figure 8(h). The fusion results of Metastatic bronchogenic carcinoma of the brain are analyzed in Figure 9 with the help of SPECT and MRI scanned images. PCNN method could preserve a decent amount of functional details of SPECT image in the fused image, but lacks sharpness as seen

in figure 9(c). From the results from figures 9(d) to 9(g), it was noticed that NSCT-PCNN suffers from artifacts while MST-SR and Latent LRR methods introduce color distortion. NSST-PCNN method, which performs better than other methods; its performance is under par with than proposed method, which ensures a fused image with no visual artifacts as obvious from figure 9(h).

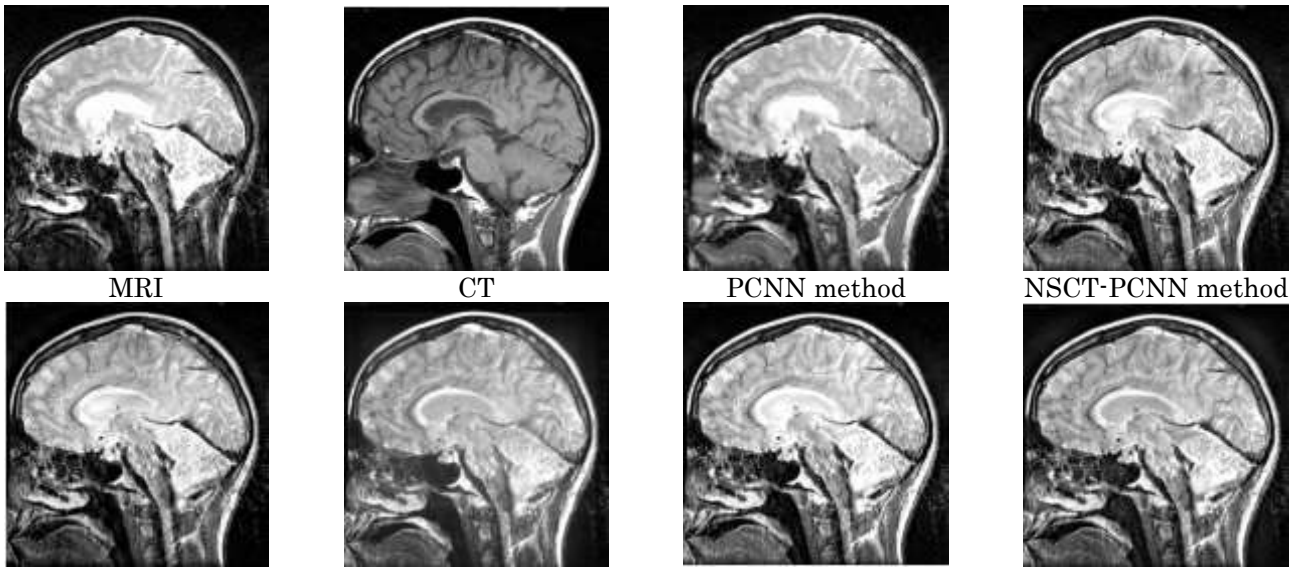

Figure 6. Fusion results of the superior sagittal sinus thrombosis dataset

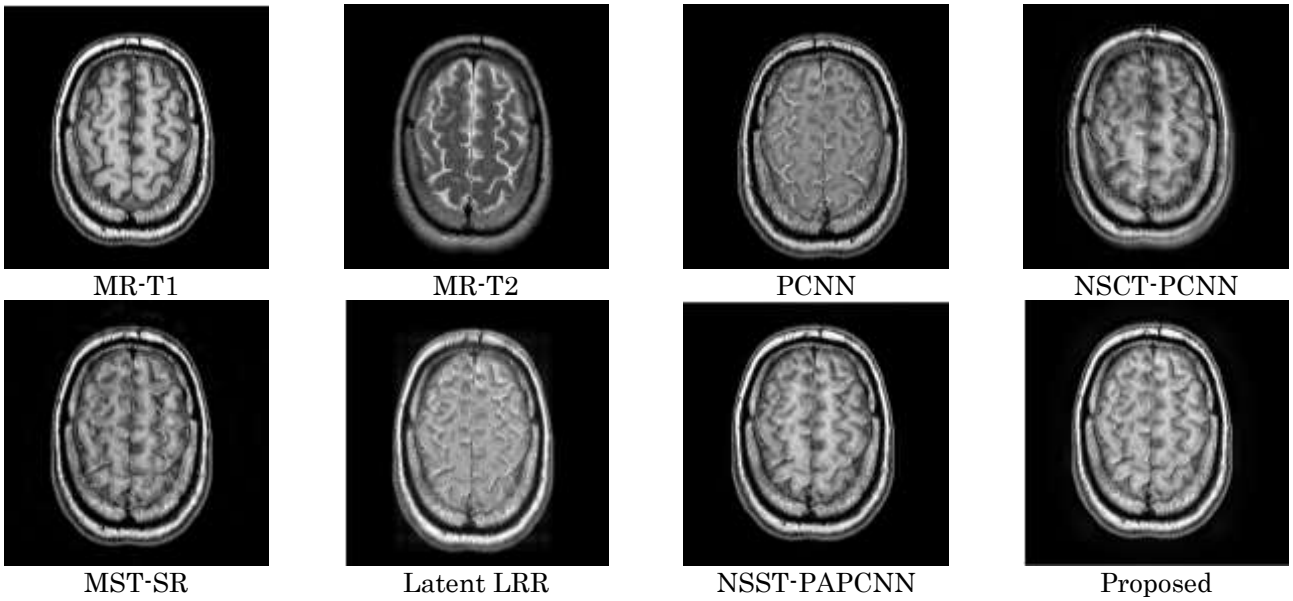

Figure 7. Fusion results of the chronic subdural hematoma dataset

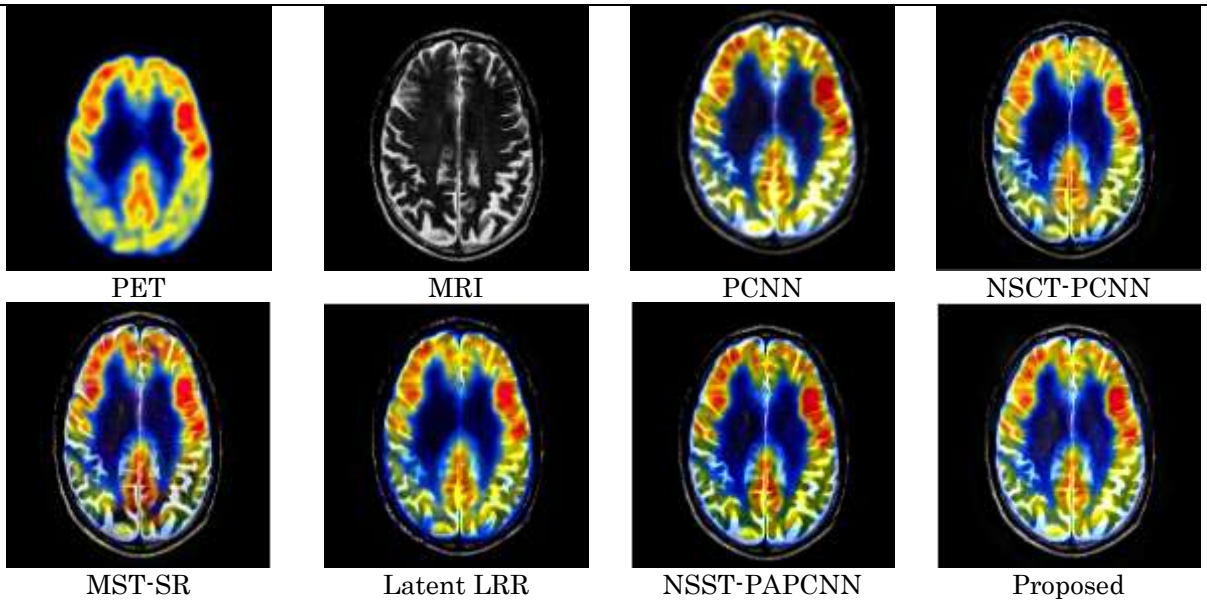

Figure 8. Fusion results of the Alzheimer's dataset

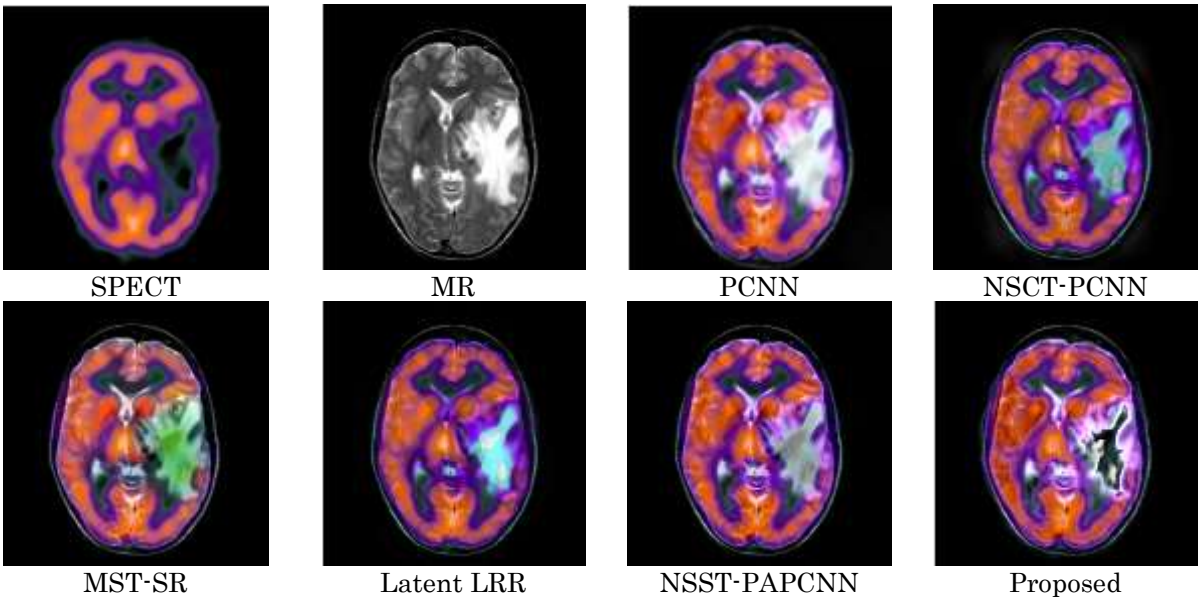

Figure 9. Fusion results of the Metastatic bronchogenic carcinoma dataset

4.2. Quantitative Analysis

Fusion performance of any algorithm must be assessed subjectively as well as objectively. This section presents the quantitative (objective) analysis of fusion algorithms compared in the qualitative analysis. A total of six quality assessment metrics- Average Gradient (AG), Image Entropy (H), Spatial Frequency (SF), Mutual Information (MI), Edge preservation index (Q_{ABF}), and Standard Deviation (SD) are considered for comparison. The mathematical formulae of these metrics, along with their significance, are presented in Table 1. The quantitative results of all methods are presented in Table 2 across four datasets. For each metric, the top three methods are ranked in brackets. For the superior sagittal sinus thrombosis dataset, the

proposed method consistently surpasses the existing methods in terms of information content, activity, and edge level preservation. Methods like NSCT-PCNN and NSST-PAPCNN perform moderately, but do not achieve the sharpness and information fidelity as much as the proposed method does. Results of chronic subdural hematoma illness reveal that the proposed strategy once again rates first in all criteria, indicating its ability to improve anatomical structures while maintaining modest intensity changes. While approaches like as MST-SR and NSST-PAPCNN are reasonably effective, they are outclassed in terms of sharpness, entropy, and mutual information. The results suggest that the proposed method is especially successful for fusing multi-sequence MR images. The problem with the PET-MRI dataset is combining functional information from PET with anatomical characteristics from MRI. The suggested technique provides the best results in all criteria, showing an excellent union of metabolic and structural properties. Other approaches, such as MST-SR and Latent LRR, perform well but suffer from tradeoffs in either color brightness or structural integrity. The proposed method maintains good contrast, rich information, and great edge strength, making it an excellent choice for functional and structural image integration. The SPECT-MRI dataset includes functional-anatomical fusion, which is similar to PET-MRI but with differing resolution and contrast properties. The proposed technique constantly scores first in all quantitative indicators, demonstrating its robustness. Competing approaches, such as NSCT-PCNN and NSST-PAPCNN, perform relatively well but struggle with contrast handling and detail retention. The suggested technique is extremely effective in regulating intensity, edge clarity, and mutual information, making it ideal for clinical applications involving functional imaging. The average metric values of twenty-five slices for each of the four datasets are provided in Figure 10 for comparative analysis. From Figure 10, the efficacy of the proposed method is proved as it surpasses the existing methods with an average minimum improvement of 41%, 5%, 16%, 23%, 34% and 4% in terms of AG, H, SF, MI, Q_{ABF}, and SD.

Table 1. Mathematical formulae of quality assessment metrics

Metric	Formulae	Significance
Average Gradient (AG)	$AG = \frac{\sum_p \sum_q \sqrt{((F(p,q)-F(p+1,q))^2 + (F(p,q)-F(p+1,q))^2)}}{MN}$	Measures the sharpness of an image. A higher value of AG signifies more sharper image. In the formula, F is a fused image of size MxN.
Image Entropy (H)	$H = -\sum_{l=0}^{255} p_l \, log_2 \, p_l$, where p_l is the probability of gray level l in the fused image	It is a measure of useful information content in an image. An image with higher entropy contains rich information content.
Spatial Frequency (SF)	$SF(p,q) = \sqrt{\lvert RF(p,q)\rvert^2 + \lvert CF(p,q)\rvert^2}$ $RF(p,q) = \sqrt{\frac{1}{MxN}\sum_{p=2}^{M}\sum_{q=2}^{N}[I(p,q)-I(p,q-1)]^2}$ $CF(p,q) = \sqrt{\frac{1}{MxN}\sum_{p=2}^{M}\sum_{q=2}^{N}[I(p,q)-I(p-1,q)]^2}$	Evaluates the intensity variations in an image. An image with higher SF has more structural details.
Mutual Information (MI)	For two source images A, B, and fused image F, Mutual information is given as $MI = MI(A,F) + MI(B,F)$ $MI(A,F) = \sum_{z\in Z}\sum_{y\in Y} p(A,F) log_2 \frac{p(A,F)}{p(A)p(F)} x$ $\mathrm{MI}(B,F) = \sum_{z\in Z}\sum_{y\in Y} p(B,F) log_2 \frac{p(B,F)}{p(B)p(F)}$	Signifies the quantity of information shared between the source and fused images. A higher value of MI indicates better transfer of activity

		information from source images to the fused image.
Edge preservation index (Q_{ABF})	$Q_{ABF} = \frac{\sum_{p=1}^{M}\sum_{q=1}^{N}(Q_{AF}(p,q)W_A(p,q) + Q_{BF}(p,q)W_B(p,q))}{\sum_{p=1}^{M}\sum_{q=1}^{N}(W_A(p,q) + W_B(p,q))}$	Measures the number of edges transferred during the fusion process from source images into fused images. It is evaluated in the range [0,1], and any value close to 1 indicates good edge preservation and a value close to 0 indicates the worst edge preservation.
Standard Deviation (SD)	$SD = \sqrt{\frac{1}{MN}\sum_{p=1}^{M}\sum_{q=1}^{N}(f(p,q) - API)^2}$	Measures contrast variations in an image. A higher SD results in a better contrast image.

Table 2. Quality assessment metrics of different methods across different datasets

Dataset	Method	AG	H	SF	MI	Q_{ABF}	SD
Skull	PCNN	12.65	7.51	26.56	5.79	0.449	74.97
	NSCT-PCNN	16.27(3)	7.52	32.97(3)	5.87	0.503(3)	80.04(2)
	MST-SR	15.91	7.59	32.38	5.93	0.502	77.71
	Latent LRR	14.52	7.80(2)	28.34	5.95(3)	0.165	74.48
	NSST-PAPCNN	16.42(2)	7.72(3)	33.00(2)	5.97(2)	0.513(2)	81.20(1)
	Proposed	28.39 (1)	7.83(1)	40.13(1)	5.99(1)	0.796(1)	79.59(3)
MR (T1 and T2)	PCNN	6.85	3.65	20.36	2.68	0.480	59.00
	NSCT-PCNN	7.09	4.42	22.58	2.74	0.538(3)	61.89
	MST-SR	8.28(2)	4.59(2)	24.66(2)	2.70	0.517	62.43(3)
	Latent LRR	7.63	4.10	20.99	2.94(2)	0.218	61.33
	NSST-PAPCNN	8.22(3)	4.58(3)	24.48(3)	2.80(3)	0.546(2)	68.5(2)
	Proposed	14.90(1)	4.62(1)	29.49(1)	3.43(1)	0.810(1)	76.69(1)
PET-MR	PCNN	6.37	3.37	22.79	3.46	0.380	66.39
	NSCT-PCNN	8.20	3.89	27.22	3.47	0.544	68.97
	MST-SR	8.53(2)	3.65	27.33(2)	3.55(2)	0.584(2)	70.81(3)
	Latent LRR	7.97	4.08(2)	26.04	3.45	0.497	70.86(2)
	NSST-PAPCNN	8.25(3)	4.02(3)	27.27(3)	3.49(3)	0.573(3)	69.71
	Proposed	16.17(1)	5.19(1)	29.29(1)	4.26(1)	0.882(1)	71.15(1)
SPECT-MR	PCNN	5.76	4.29	17.46	2.77	0.421	63.55
	NSCT-PCNN	6.76(2)	4.63	20.15(3)	2.80	0.599(3)	67.67
	MST-SR	6.51	4.40	19.72	2.83(3)	0.580	64.98
	Latent LRR	5.92	5.37(2)	17.55	2.71	0.430	68.64(3)
	NSST-PAPCNN	6.78(3)	4.82(3)	20.16(2)	2.84(2)	0.602(2)	75.04(2)
	Proposed	9.02(1)	5.55(1)	21.40(1)	3.57(1)	0.839(1)	75.20(1)

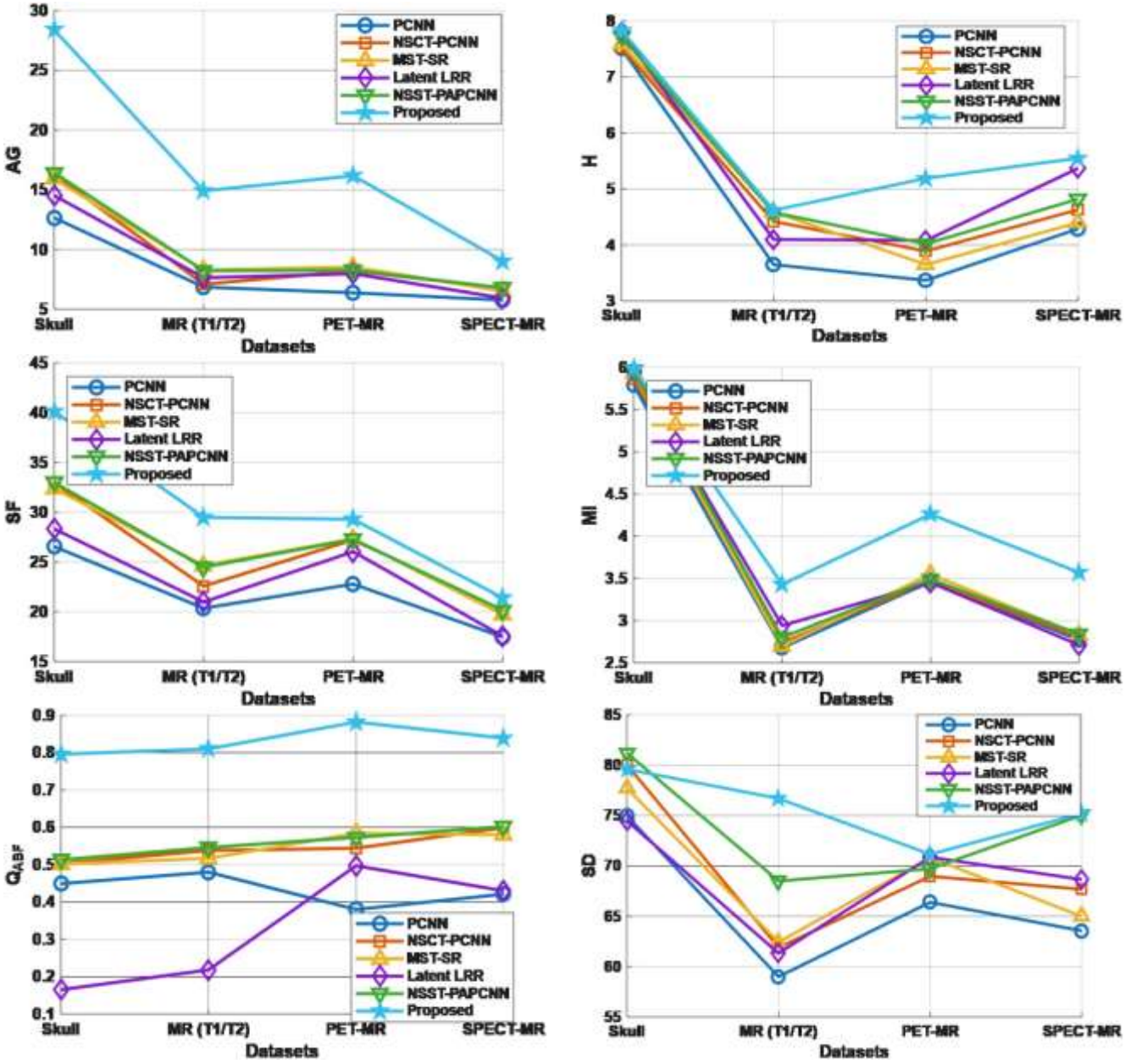

Figure 10. Comparative analysis of average metric values over 100 slices of four datasets

4.3. Computational Complexity and Execution Time

Three steps—multilevel local extrema decomposition, FDDBC-based adaptive weight computation, and iterative EWPADCNN-based detail fusion, have a major impact on the computational cost of the suggested approach. The local extrema decomposition shows a computational complexity of $O(MN)$ per scale for an input image of size M×N. The decomposition complexity in this work is $O(4MN)$ due to the use of four decomposition stages, largely due to the FDDBC-based fractal dimension estimation uses local box-counting operations across the image domain, it thus has linear complexity $O(MN)$. The EWPADCNN fusion step typically takes 25 iterations to converge, requiring neighborhood-based neural updates of complexity $O(MN)$ for each iteration. This results in a total neural processing complexity of $O(25MN)$, *a*s

a result, the suggested method's overall computational complexity is similar to other NSST-PCNN-based fusion techniques and stays linear with respect to image size. All experiments were carried out in MATLAB on a laptop outfitted with an Intel Core i7 CPU (3.2 GHz) and 16 GB RAM. For 256×256 images, the average execution time per fused slice was 0.78 seconds, which included decomposition, adaptive weight computation, neural fusion, and reconstruction. This execution time illustrates that the suggested approach is appropriate for clinical image fusion applications that can be processed offline or in real-time.

5. Conclusion

This work proposed a fusion algorithm for effective diagnosis of brain-related tumors. It utilizes a multi-level decomposition of source images based on local extrema information. Detail layers of source images at different levels are fused using an enhanced weighted pulse adaptive dual channel neural network (EWPADCNN), where the weights are computed using the FDDBC method. The base layers are combined based on individual pixel significance using a weighted average rule. Experiments are conducted on 100 slices of four different data sets of brain disorders and results confirmed that the proposed method conveniently surpassed the existing methods in providing acceptable fusion result both qualitatively and quantitatively, there by achieving an average minimum improvement of 41%, 5%, 16%, 23%, 34% and 4% in terms of AG, H, SF, MI, Q_{ABF} and SD than existing methods, thus making it a viable solution for fusing the medical images to effectively diagnose the tumors, which helps the doctors to plan the treatment accordingly. Despite its superior performance, the proposed method has a restriction that should be considered. The FDDBC-based adaptive weighting process may be sensitive under high noise or low contrast conditions, affecting the accuracy of pixel-wise linking strength estimates. Future work will concentrate on increasing resilience under adverse imaging settings and expanding the suggested fusion architecture to an end-to-end diagnostic systems by including automated tumor segmentation and classification.

Funding Statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment: The authors would like to thank the reviewers and editors for their valuable comments and constructive suggestions, which helped improve the quality of this manuscript.

References

[1] Yin, M.; Liu, X.; Liu, Y.; Chen, X.; "Medical image fusion with parameter-adaptive pulse coupled neural network in non-subsampled shearlet transform domain". IEEE Trans. Instrum. Meas., 68 (1): 49–64, 2019.

[2] Du, J.; Li, W.; Lu, K.; Xiao, B.; "An overview of multi-modal medical image fusion". Neurocomputing, 215: 3–20, 2016.

[3] Guo, K.; Hu, X.; Li, X.; "MMFGAN: A novel multimodal brain medical image fusion based on the improvement of generative adversarial network". Multimedia Tools Appl., 81 (4): 5889–5927, 2022.

[4] Li, S.; Kang, X.; Fang, L.; Hu, J.; Yin, H.; "Pixel-level image fusion: A survey of the state of the art". Inf. Fusion, 33: 100–112, 2017.

[5] Chao, Z.; Duan, X.; Jia, S.; Guo, X.; Liu, H.; Jia, F.; "Medical image fusion via discrete stationary wavelet transform and an enhanced radial basis function neural network". Appl. Soft Comput., 118, 108542,2022.

[6] Yu, N.; Li, J.; Hua, Z.; "Decolorization algorithm based on contrast pyramid transform fusion". Multimedia Tools Appl., 81: 15017–15039, 2022.

[7] Aishwarya, N.; Bennila Thangammal, C.; Praveena, N.G.; "NSCT and focus measure optimization based multi-focus image fusion". J. Intell. Fuzzy Syst., 41: 903–915, 2021.

[8] Kamarthi, V.; Satyanarayana, D.; Ninjappa, G.P.M.; "Multimodal medical image fusion based on intuitionistic fuzzy sets and weighted activity measure in NSST domain". Curr. Signal Transduct. Ther., 17: 22–31, 2022.

[9] Johnson, J.L.; Padgett, M.L.; "PCNN models and applications". IEEE Trans. Neural Netw., 10 (3): 480–498, 1999.

[10] Vanitha, K.; Satyanarayana, D.; Prasad, M.G.; "Multi-modal medical image fusion algorithm based on spatial frequency motivated PA-PCNN

in the NSST domain". Curr. Med. Imaging, 17: 634–643, 2021.

[11] Singh, S.; Gupta, D.; Anand, R.S.; "Non-subsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network". Biomed. Signal Process. Control, 18: 91–101, 2015.

[12] Guo, Z.; Song, Y.; Zhao, Y.; "An adaptive infrared image segmentation method based on fusion SPCNN". Signal Process.: Image Commun., 87: 115905, 2020.

[13] Yang, Y.; Gao, C.; Ming, Z.; Guo, J.; Leopold,E.; Cheng,J.; Zuo,J.; Zhu,M.; "LatLRR-CNN: An infrared and visible image fusion method combining latent low-rank representation and CNN". Multimedia Tools and Applications, 82 (23): 36303-36323, 2023.

[14] Liu, Y.; Liu, S.; Wang, Z.; "A general framework for image fusion based on multi-scale transform and sparse representation". Inf. Fusion, 24: 147–164, 2015.

[15] Ibrahim, S.I.; El-Tawel, G.S.; Makhlouf, M.A.; "Brain image fusion using the parameter-adaptive pulse coupled neural network (PA-PCNN) and non-subsampled contourlet transform (NSCT)". Multimedia Tools Appl., 83 (9): 27379–27409, 2024.

[16] Xiaomin, L.; Haowen, Y.; "An algorithm to generate a weighted network Voronoi diagram based on improved PCNN". Appl. Sci., 12 (12): 6011, 2022.

[17] Venkata Srikanth, M.; Suneel Kumar, A.; Nagasirisha, B.; Lakshmi, T.; "Brain MRI and CT image fusion using multiscale local extrema and image statistics". ECTI Trans. Electr. Eng. Electron. Commun., 22 (1):1-11, 2024.

[18] Panigrahy, C.; Seal, A.; Mahato, N.K.; Bhattacharjee, D.; "Differential box counting methods for estimating fractal dimension of gray-scale images: A survey". Chaos Solitons Fractals, 126: 178–202, 2019.

[19] Chen, Y.; Park, S.K.; Ma, Y.; Ala, R.; "A new automatic parameter setting method of a simplified PCNN for image segmentation". IEEE Trans. Neural Netw., 22 (6): 880–892, 2011.

[20] Otsu, N.; "A threshold selection method from gray-level histograms". Automatica, 11: 23–27, 1975.

[21] Shreyamsha Kumar, B.K.; "Image fusion based on pixel significance using cross bilateral filter".Signal Image Video Process, 9: 1193–1204, 2015.