



ARTICLE

## Multimodality Medical Image Fusion Based on Pixel Significance with Edge-Preserving Processing for Clinical Applications

Bhawna Goyal<sup>1</sup>, Ayush Dogra<sup>2</sup>, Dawa Chyophel Lepcha<sup>1</sup>, Rajesh Singh<sup>3</sup>, Hemant Sharma<sup>4</sup>, Ahmed Alkhayat<sup>5</sup> and Manob Jyoti Saikia<sup>6,\*</sup>

<sup>1</sup>Department of ECE and UCRD, Chandigarh University, Mohali, Punjab, 140413, India

<sup>2</sup>Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, 140401, India

<sup>3</sup>Department of ECE, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India

<sup>4</sup>IES College of Technology, IES University, Bhopal, 462044, India

<sup>5</sup>College of Technical Engineering, The Islamic University, Najaf, 54001, Iraq

<sup>6</sup>Department of Electrical Engineering, University of North Florida, Jacksonville, FL, 32224, USA

\*Corresponding Author: Manob Jyoti Saikia. Email: manob.saikia@unf.edu

Received: 31 October 2023 Accepted: 11 January 2024 Published: 26 March 2024

### ABSTRACT

Multimodal medical image fusion has attained immense popularity in recent years due to its robust technology for clinical diagnosis. It fuses multiple images into a single image to improve the quality of images by retaining significant information and aiding diagnostic practitioners in diagnosing and treating many diseases. However, recent image fusion techniques have encountered several challenges, including fusion artifacts, algorithm complexity, and high computing costs. To solve these problems, this study presents a novel medical image fusion strategy by combining the benefits of pixel significance with edge-preserving processing to achieve the best fusion performance. First, the method employs a cross-bilateral filter (CBF) that utilizes one image to determine the kernel and the other for filtering, and vice versa, by considering both geometric closeness and the gray-level similarities of neighboring pixels of the images without smoothing edges. The outputs of CBF are then subtracted from the original images to obtain detailed images. It further proposes to use edge-preserving processing that combines linear lowpass filtering with a non-linear technique that enables the selection of relevant regions in detailed images while maintaining structural properties. These regions are selected using morphologically processed linear filter residuals to identify the significant regions with high-amplitude edges and adequate size. The outputs of low-pass filtering are fused with meaningfully restored regions to reconstruct the original shape of the edges. In addition, weight computations are performed using these reconstructed images, and these weights are then fused with the original input images to produce a final fusion result by estimating the strength of horizontal and vertical details. Numerous standard quality evaluation metrics with complementary properties are used for comparison with existing, well-known algorithms objectively to validate the fusion results. Experimental results from the proposed research article exhibit superior performance compared to other competing techniques in the case of both qualitative and quantitative evaluation. In addition, the proposed method advocates less computational complexity and execution time while improving diagnostic computing accuracy. Nevertheless, due to the lower complexity of the fusion algorithm, the efficiency of fusion methods is high in practical applications. The results reveal that the proposed method exceeds the latest state-of-the-art methods in terms of providing detailed information, edge contour, and overall contrast.



**KEYWORDS**

Image fusion; fractal data analysis; biomedical; diseases; research; multiresolution analysis; numerical analysis

**1 Introduction**

Numerous developments have been revolutionized in the domain of medical imaging systems for clinical applicability due to technological advancements, including growth in the number of imaging modalities. To accurately diagnose diseases, doctors need details from different imaging sources [1,2]. In particular, computed tomography (CT) images provide information about dense bone structures; however, they cannot provide any details of internal organs. Similarly, the hard tissue contrast is visible in the magnetic resonance imaging (MRI) images, but bone information is missing. As a result of the limitations of a single image, it is incapable of providing all of the necessary information. Thus, medical image fusion plays a pivotal role for doctors or radiologists because it fuses two or multiple medical images of different modalities that provide high-quality images for better disease diagnosis. MRI and CT image fusion allows us to fuse images of connective tissue details obtained from MRI images and details of vertebral anatomy from images, allowing physicians to better interpret and treat patients [3,4]. Image fusion attempts to fuse the requisite details from both images into a single image by taking care not to introduce artifacts in the images [5]. Different fusion algorithms have been developed to solve the issues of medical imaging images over the past several decades [6]. The conventional image fusion strategies in general are grouped into spatial and transform domain techniques. The fusion of images based on blocks, pixels, or regions is known as spatial domain algorithms [7,8]. In addition, transform domain methods are based on the principle of transforming an image into another domain by using the domain's characteristics to accomplish the objective more accurately [9–11]. These techniques, however, are not very good at fusing specific information from images, resulting in the generation of artifacts. Research on pixel-level fusion has been performed over the past 30 years, with over 1000 affiliated scientific papers published during this period. In recent times, deep learning-based techniques [12,13] have made massive advancements in numerous areas of image processing and computer vision, i.e., segmentation, super-resolution, classification, detection, denoising, and others. Numerous algorithms are proposed based on cross-bilateral filters [14,15]. Like a bilateral filter, cross bilateral filter considers both grey level similarities and geometric closeness of the neighboring pixels without smoothing edges, but it uses one image for finding the kernel and another to filter, and vice versa. In these methods, an image is decomposed into an energy layer with pure intensity information and a structure layer with rich details. Then, the visual saliency detection map (VSDM) is used to guide the fusion of the energy layers to take full advantage of the edge contour extraction to improve the edge contour sharpness of the fused image and to retain the image detail information fully. They use average cross-bilateral filtering (ACBF) to decompose the source image into smooth layers (SL) and detail layers (DL). Then, the traditional ABS-MAX method is used to fuse the smooth layer, and the neighborhood energy sum (NES) method is used to fuse the detail layer. Finally, the final fused image is obtained by adding the fused smooth layer and detail layer.

Deep learning algorithms have become an active topic in recent times in the domain of image fusion technology. Wide-ranging deep learning-based algorithms have been developed in recent years. In addition, generative adversarial network (GAN)-based image fusion has recently become

increasingly significant. GANs are a new generative modal proposed by Goodfellow [16], which aims to compute target distributions and generative information that match a point distribution without depending on any prior assumptions. In the past few years, GANs have been often utilized in numerous visual interpretation applications with excellent results [17,18]. While current approaches have achieved optimistic performance in the majority of problems, many features could still be improved. In general, earlier methods needed manual construction of activity-level calculations and fusion regulations that limited fusion performance due to their difficulty in considering all the components in a single manual design process. Secondly, several current algorithms perform multisensory/multifocus image fusion by creating a decision map, which usually provides classification problems based on sharpness detection. Near the boundary lines, these approaches usually fail to distinguish between a focused and unfocused region. Third, when creating a decision map, almost all deep learning-based approaches need post-processing, i.e., consistency checks, which greatly enhances the complexity of techniques. Further, these approaches usually require the manual development of a decision map as ground truth for training the network which further limits their applications.

In addition, transformer-based models are playing a pivotal role in medical imaging analysis [19,20]. Zhang et al. [21] introduced TransFusion, a transformer-based end-to-end architecture for medical picture fusion. The suggested TransFusion fusion approach introduces the transformer as the fusion strategy, which uses its self-attention mechanism to incorporate global contextual information from multimodal aspects and fuse them appropriately. Furthermore, unlike classic parallel multibranch designs or shared networks used for many inputs, they construct branch networks that interact at multiple scales using fusion transformers to better utilize information from different modalities. Yu et al. proposed an end-to-end image fusion network based on the Swin-Transformer mechanism capable of fusing both local and long-range (or global context) image information [22]. The proposed network is trained in two stages. In the first step, it trains a multi-layer transformer-based autoencoder network to extract shallow and multi-scale deep features. In the second phase, numerous features are fused using transformers and convolutional neural networks (CNNs), which have the potential to collect both local and long-range features. For multimodal medical image fusion, a model that combines the CNN module and the transformer module is presented [23]. The CNN module is used to extract image detail texture information, whereas the transformer module is used to extract image pixel intensity distribution information. Extensive experimental results in the Harvard Brain Atlas test dataset show that the proposed method outperforms comparable methods.

In this study, an upgraded version of the image fusion strategy is proposed by integrating all the vital information found in each input image and utilizing a cross-bilateral filter to obtain detailed images after subtracting them from the source images. Then the method employs a combination of a linear lowpass filter and nonlinear processing of residuals to find and reconstruct meaningful edges in lowpass-filtered images. Thus, residuals are then computed and divided into two fractions, one for the positive portion of pixels and the other for the absolute value of the negative portion. To locate the strong edges, i.e., high-contrast ones, thresholding is applied to both fractions. The results of thresholding are then utilized as markers for the morphological reconstruction of residuals. To recover meaningful edges, restored residuals that do not hold meaningful imaging elements are applied to low-pass filtered images. The edge-recovered images are then used for the computation of weights. The estimated weights are integrated with the input images to obtain the final fusion output.

The main contributions were as follows:

1. This study develops a competent image fusion strategy by combining multimodal medical images using pixel significance with edge-preserving processing, which is capable of performing image fusion tasks effectively.
2. This methodology boosts the fusion results by incorporating relevant visual information by estimating weights from the detailed images. Further, this technique is incredibly simple in implementation and computation.
3. In addition, it uses non-linear filtering that considers localizing and recovering relevant edges from the source images, thereby preventing the development of artifacts within the image regions and boosting visual perception.
4. This methodology can be utilized for implementing a manageable edge-aware image blur with contrast enhancement, thus improving the image quality while maintaining a precise amount of information.
5. Lastly, our methodology is compared with other state-of-the-art image fusion algorithms by considering benchmark medical image fusion datasets.

The remainder of the study is structured as follows: In [Section 2](#), we briefly illustrate existing literary works based on image fusion. The proposed method is systematically described in [Section 3](#). [Section 4](#) demonstrates quality evaluation metrics, experimental details, and both qualitative and quantitative findings. [Section 5](#) concludes with the proposed method.

## 2 Related Works

Several image fusion approaches have been introduced in recent decades [24,25]. The majority of prior fusion techniques consist of decomposition and learning-based techniques [9,26]. The multiscale transform is a broadly used decomposition algorithm for the fusion of images that contains three stages in fusion practice: decomposition, fusion, and restoration [27]. The multiscale transform generally employs wavelet-based, pyramid-based, and multi-geometric analysis-based approaches. The non-subsampled contourlet transform (NSCT) [10] and non-subsampled shearlet transform (NSST) [28] domain-based algorithms have shown promising image representations. Moreover, during image transformation, a pivotal problem of the multiscale transform-based fusion approach is the analysis of frequency components. The activity level of high-frequency components has traditionally been determined by their absolute values. It is measured in a window- or pixel-based manner, and the fused coefficient is obtained using a simple function scheme, including a selection of the maximum or weighted average.

Early research found that averaging a coefficient of numerous source images was a prominent low-frequency image fusion strategy. More complex image conversions and fusion approaches have been introduced in recent years [29,30]. Zhang et al. introduced a fusion strategy related to combining convolutional sparse representation and multiscale transformation [31]. Zhu et al. introduced a multimodality decompositions approach related to NSCT that uses local Laplacian energy and phase consistency to combine the low-pass and high-pass subbands [10]. To improve fusion performance, Lepcha et al. [32] introduced a technique based on a cross-bilateral filter (CBF) with a rolling guidance filter that combines with weight normalization to enhance medical images. Dogra et al. [33] developed a method for minimizing distortions or abnormalities in bones and vasculature or arterial image fusion investigations. Goyal et al. [34] recently developed an approach that integrates poor-quality images, maintains less complexity for increasing target detection consistency, and offers a foundation for medical applications. In the NSST domain, Yin et al. incorporated a multimodality fusion of images

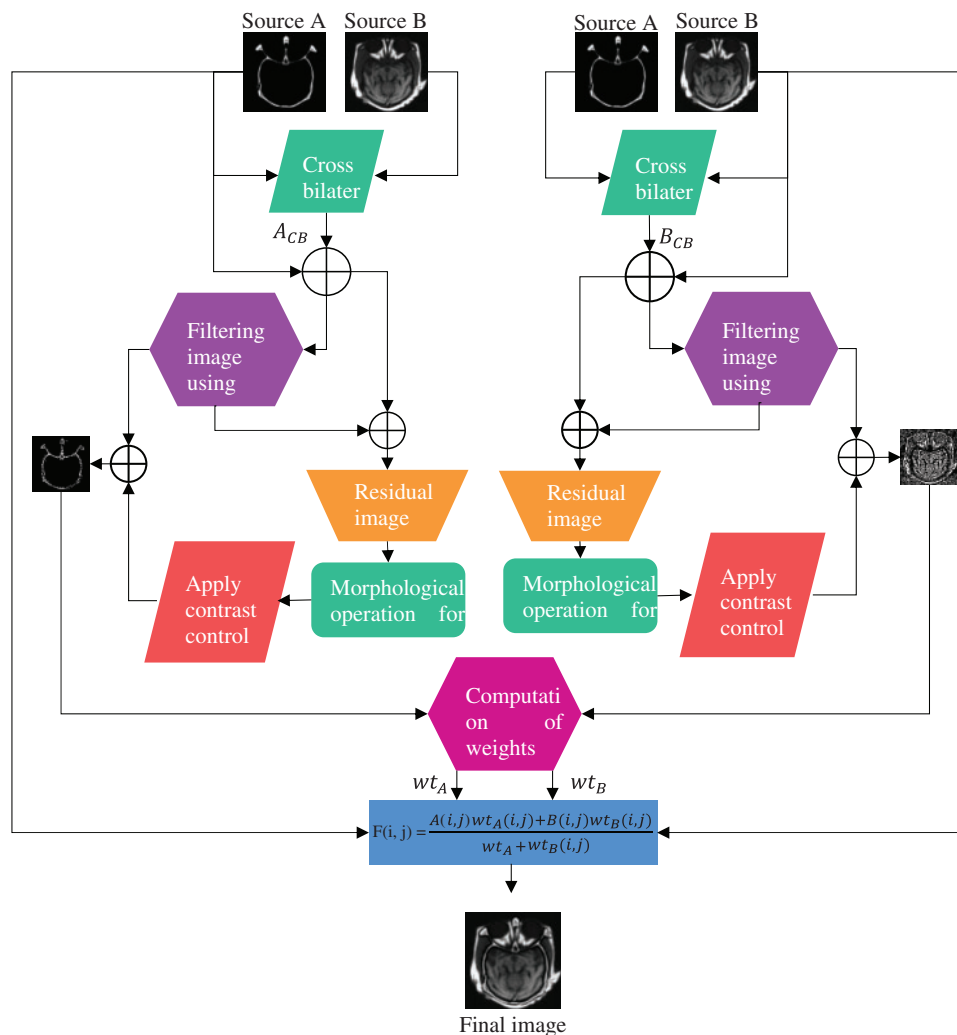
algorithm by including the use of a pulse-coupled neural network (PCCN) [35] for fusing images. Lei et al. developed a new multisensory fusion architecture relating to PCNN and NSST to enhance the image fusion efficiency of multimodal images [36]. A learning-based approach has been widely utilized in fusion strategy in recent times. The most widely used fusion approaches are sparse representation (SR) and deep learning-related fusion methods [37]. The sparse representation-related fusion strategy utilized typical sparse coding modality related to a single image element and local image block in the early stages. The input image is segmented into overlapping image blocks in the initial spatial domain for sparse coding approaches. Most of the current sparse coding-related image fusion approaches try to boost their efficiency in the following ways: by adding detailed constraints, creating a highly effective dictionary learning technique, and representing multiple sub-dictionaries [38].

Kim et al. introduced a dictionary learning approach using the SR-based model based on joint image block clustering. Zhu et al. developed a fusion of medical images recently [39]. However, all of the methods mentioned above involved complicated fusion rules or several sparse representation-based modalities. These particular rules, however, do not generalize to every form of medical image fusion. Deep-learning-related fusion has become a popular research topic with the fast advancement of artificial intelligence [40,41]. In recent years, deep learning has achieved significant success in various fields, including image recognition, natural language processing, speech recognition, and recommendation systems. Some of the popular deep learning architectures include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs). The CNN-based local connection with weight-sharing characteristics will increase the efficiency of the fusion technique by simultaneously reducing network complexity and weights. At the moment, CNNs are becoming more and more strategic in their fusion strategies. Xia et al. combined a multiscale transform and CNN to develop a multimodal medical image fusion modality that utilizes a deeper stacked neural network for dividing input images into low- and high-frequency elements and corresponding image fusion [42].

One of the major problems in current deep learning-based image fusion is computing the weight maps that fuse the pixel activity details from different sources. Most current fusion algorithms contain a two-step process that includes activity-level calculation and weight assignments. The absolute values of decomposition coefficients are employed to calculate its operation initially in conventional transform domain methods. The maximum or weighted average is then selected using a fusing rule, i.e., the choose-max or weighted-average rule. The corresponding weights are eventually assigned to various sources based on the obtained measurements. Different complex decomposition methods and detailed weight assignment methods have been introduced recently to improve fusion efficiency [43,44]. However, constructing an optimal activity-level calculation or weight assignment technique that takes into account all of the important factors is challenging. Jose et al. [28] introduced a novel multimodality algorithm for an NSST-related adolescent identity search (AISA) to achieve image optimization while reducing computational cost and time. Li et al. [45] introduced a fusion technique that helps with medical analysis and increases target recognition precision and fused image quality. NSCT and CNN are combined by Wang et al. [46]. They incorporated the benefits of both NSCT and CNN to boost fusion performance. A hybrid medical image fusion system combines LatLRR and FCNs by Xu et al. [47]. Liu et al. [48] designed an effective technique using a spectral total variation transform (STVT). A STVT algorithm is utilized for decomposing the input images into a set of texture signatures, which refers to the dependent components and deviation components.

### 3 Methodology

This section illustrates the proposed multimodal fusion strategy using pixel significance with edge-preserving processing in detail. Fig. 1 represents a detailed framework for the proposed methodology. It can be observed from Fig. 1 that our fusion algorithm contains three crucial steps: cross-bilateral filtering, low-pass filter with morphological processing of residuals, and pixel-based fusion rule.



**Figure 1:** Methodological flowchart of the proposed algorithm

#### 3.1 Cross-Bilateral Filter (CBF)

CBF [49] represents a nonlinear and non-iterative technique that integrates an edge stopping function and a lowpass filter, which minimizes the filter kernel if the magnitude variation between pixels is greater. Both levels of gray similarity and the geometric closeness of the surrounding pixels are taken into consideration. The weight of filters does not only depend on the Euclidean distance but also depends on the distance of color or gray space. The process benefits as it smooths the edges by conserving the edges while considering neighborhood pixels.

Mathematically, in the case of source image A, the output of the bilateral filter (BF) at position  $p$  could be estimated as [50]

$$A_F(p) = \frac{1}{W} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|) A(q) \quad (1)$$

where  $G_{\sigma_s}(\|p - q\|) = e^{-\frac{\|p - q\|^2}{2\sigma_s^2}}$  is a geometric closeness function.

$G_{\sigma_r}(|A(p) - A(q)|) = e^{-\frac{|A(p) - A(q)|^2}{2\sigma_r^2}}$  is a gray-level similarity.

$W = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|)$  is a normalization constant.

where  $\|p - q\|$  denotes a Euclidean distance between  $p$  and  $q$ ,  $S$  indicates the spatial neighborhood of  $p$ .

Similarly, we can calculate  $B_F$  using Eq. (1) in the case of source image B. Here,  $\sigma_s$  and  $\sigma_r$  manages the performance of BF, the dependence of  $\sigma_s/\sigma_r$  value and derivative of original images on the behavior of BF are demonstrated in [51]. The value of  $\sigma_s$  is selected based on the optimal output of a lowpass filter and blur highly for more  $\sigma_s$ , since it integrates values from highly distant locations [51]. In addition, if the images get scaled down and scaled up,  $\sigma_s$  should be adjusted to obtain optimal outputs. It also defines that, the desired good range for  $\sigma_s$  is (1.5–2.1); similarly, the optimal values of  $\sigma_r$  rest on the quantity of the edges to be preserved. In cases where the image gets attenuated and amplified it,  $\sigma_r$  should be adjusted to obtain a similar output. CBF considers both the level of gray similarity and geometric closeness of neighborhood pixels in image A for shaping the filter kernel and filter image B. The CBF output of image B at a pixel position  $p$  is defined by [52]

$$B_{CBF}(p) = \frac{1}{W} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|) B(q) \quad (2)$$

where  $G_{\sigma_r}(|A(p) - B(p)|) = e^{-\frac{|A(p) - A(q)|^2}{2\sigma_r^2}}$  is a gray-level similarity

$W = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|A(p) - A(q)|)$  is a normalization constant.

Similarly, we can compute  $A_{CBF}$  using Eq. (2) from the corresponding source image. The detailed images are generated through the subtraction of CBF output from the input images. In the case of images A and B, the detailed images can be obtained as  $A_D = A - A_{CBF}$  and  $B_D = B - B_{CBF}$ . In medical images, the unfocused region in image A could be in image B, and using CBF in image B could blur focus regions higher as compared to that of the unfocused region of the image B. Since the unfocused region in the image A somehow seems blurred, having nearly the same gray value in the region therefore forms a kernel filter the same as a Gaussian. However, the concept is to obtain highly focused area information in a detailed image  $B_D$  in such a way that this information would be utilized to determine the weights utilizing weighted averaging. Likewise, in the medical image, the details in image B are not present in image A, and using CBF in the image B could blur the details in image B. Consequently, the details in image A are absent.

### 3.2 Edge-Preserving Processing

This section illustrates edge-preserving processing that is based on lowpass filtering with morphological processing of its residuals (LPMPR) [53]. It is mainly based on residual Gaussian filtering and is expressed as

$$I_{lf} = A_D * \mathcal{L} \quad (3)$$

where  $*$  is a convolutional operator and  $A_D$  is the detailed image obtained after subtracting CBF output from original image  $A$ . While  $\mathcal{L}$  represents a Gaussian filter or any other low-pass filtering mask. Henceforth, a linear filter residual is defined as

$$I_{res} = A_D - I_{lf} \quad (4)$$

The residual is then put through another processing step utilizing functions that are described using positive value images. In that case, the contents, which  $I_{res}$  contain both negative and positive values, are decomposed into two portions  $\sim$  negative and positive:

$$\begin{aligned} I_{res+} &= 0.5 (I_{res} + |I_{res}|); \\ I_{res-} &= 0.5 (|I_{res}| - I_{res}) \end{aligned} \quad (5)$$

The residual fraction follows the related relationship described below:

$$I_{res} = (I_{res+}) + (I_{res-}) \quad (6)$$

Similarly, we can obtain a residual image of detailed image  $B$  using similar procedures.

Both fractions of the residual ( $I_{res+}$ ,  $I_{res-}$ ) are processed to cut off the irrelevant dissimilarity of the residual while maintaining relevant ones. The processing method that selects the most significant residue regions and preserves their original structures is based on morphological function  $\mathcal{M}$  and reconstruction. Lastly, significant residual areas are added in the image for recovering meaningful edges, while simultaneously, images remain blurred across irreverent image regions as

$$I_{out1} = I_{lf} + \mathcal{M}(I_{res+}) - \mathcal{M}(I_{res-}) \quad (7)$$

A function  $\mathcal{M}$  remains similar for both residuals and is described as

$$\mathcal{M}(I) = R_I(\min(I, S_t(I) |_{\{\min(I), \max(I)\}})) \quad (8)$$

when  $R_I(A)$  represents morphological restoration of gray level mask image  $I$  from markers  $A$  and ‘|’ represent mapping operator where the image is recovered by transforming a binary image by substituting primary values of 0’s and 1’s by two stated ones. A ‘*min*’ represents the pointwise minimal function of two images. A binary masker  $S$  that consists of significant regions where contrast can be maintained is the key aspect for the interpretation of  $\mathcal{M}$ . The substitute in these areas is related to the amplitude of the residue as

$$S_t(I) = (I \geq t) \quad (9)$$

where  $S$  stands for the selection operator that extracts regions of  $I$  with an amplitude higher than the given threshold  $t$ . In addition, the number of portions where the original sharpness is reconstructed.



### 3.2.1 Size Criteria for Selection of Meaningful Regains

According to the initial implementation of this strategy, the edges of relevant areas are present in the binary mask that is extracted via thresholding. Using the amplitude residual, meaningfulness is calculated. So far, amplitude is not the only factor that affects how significant an image component is. It is easy to envision an image component with a greater residual amplitude specification that has no impact on how the image should be perceived. For instance, the addition of salt and pepper to the image would result in the generation of several smaller image components with larger amplitudes and the addition of high-amplitude components, which would alter the original image. Subsequently, they could be detected as relevant parts that are not optimal. Thus, another set is incorporated into the original approach to tackle this problem. The binary mask is filtered by the *area opening filter* [54]. The binary image is removed by this filter as are all related components with a size smaller than the given threshold  $s$  (size coefficient). Subsequently, Eq. (9) is expressed as

$$S_{t,s}(I) = (I \geq t) o(s) \quad (10)$$

where  $o(s)$  represents an *area opening* that removes elements smaller than those of the specific size specified by  $s$ . As compared to the initial method, one can notice that in both situations, when the coefficient ( $s$  in terms of the modified method and  $t$  in terms of the initial method) is enhanced, the number of selected regions reduces. On the contrary, in the case of the original situation, when the coefficient  $s$  is employed, the parts are eliminated based to their sizes. It permits thus rejecting smaller pixels and objects from residuals, even if their amplitudes are higher. Finally, it offers to maintain these parts blurry in the output image.

### 3.2.2 Contrast Control

Subtraction (Addition  $\sim$  depending on the components) of high-pass filtering from the original image is one of the most well-known methods for improving contrast in images. It has to do with high pass filtering properties to distinguish local differences in the value of image pixels. An additional method to get high-pass filter results is to compare differences across low-pass filtering and the image itself. High-frequency components of images with amplitude areas exceeding threshold  $t$  are referred to as morphologically processed residuals. To manage the contrast of the final denoised image, a contrast control coefficient is presented in Eq. (7), which provides an updated formulation as

$$I_{out1} = I_{lf} + (\mathcal{M}(I_{res+}) - \mathcal{M}(I_{res-})). c. \quad (11)$$

By depending on  $c$ , can be further enhanced, preserved, and reduced. Also shown are the results of contrast enhancement without morphological processing of its residual. How morphological processing affects this information can be realized in the number of small image features that are visible in the output [53]. The contrast of the relevant regions is improved by this processing method since it enables the unimportant elimination of information. This strategy is controlled by the main three factors  $t$ ,  $s$ , and  $c$ . The amount of information obtained depends on the amplitude of the residual ( $t$ ) and the particle size of the residual ( $s$ ). The parameter  $c$  allows the contrast of an output image to be controlled. Similarly, we can compute detailed image  $B_D$  using a similar process to obtain  $I_{out2}$ .

### 3.3 Pixel Significance-Based Fusion Rule

The fusion rule in [55] is presented to fulfill the completeness of our method. As demonstrated in [55], the weight is estimated employing the statistical characteristics of the neighborhood detailed components instead of wavelet components. The window dimension  $w \times w$  around the detail coefficients

$I_{out1}(i,j)$  or  $I_{out2}(i,j)$  by considering the neighborhood for estimating the weights. The neighborhood is represented by the matrix  $M$ . Every row of the matrix  $M$  is regarded as the observation and the column for computing an unbiased estimate  $C_h^{i,j}$  of its covariance matrix [56] while  $i$  and  $j$  denote the spatial coordinates of the detail components  $I_{out1}(i,j)$  and  $I_{out2}(i,j)$ .

$$\text{covariance}(M) = E[(M - E[M])(M - E[M])^T] \quad (12)$$

$$C_h^{i,j} = \frac{\sum_{k=1}^w (x_k - \bar{x})(x_k - \bar{x})^T}{(w - 1)} \quad (13)$$

where  $x_k$  represents  $k$ th observation of  $w$ -dimensional variable and  $\bar{x}$  denotes an observational mean. It has been discovered that the diagonal of  $C_h^{i,j}$  represents a variance of vector for every column of  $M$ . The eigenvalues of  $C_h^{i,j}$  is estimated and the number of eigenvalues usually depends on the dimension of the matrix  $C_h^{i,j}$ . The summation of these eigenvalues is directly relative to the horizontal detail strength of the neighbors as stated by *HdetailStrength* [55]. In a similar manner, the unbiased covariance estimates  $C_v^{i,j}$  is produced by considering every column of  $M$  as the observation and every row as the variable (i.e., the inverse of how  $C_h^{i,j}$  is calculated), and the summation of eigenvalues of  $C_v^{i,j}$  generates vertical detail strength *VdetailStrength*, which is defined as follows:

$$\begin{aligned} HdetailStrength(i,j) &= \sum_{k=1}^w \text{eigen}_k \text{ of } C_h^{i,j} \\ VdetailStrength(i,j) &= \sum_{k=1}^w \text{eigen}_k \text{ of } C_v^{i,j} \end{aligned} \quad (14)$$

where  $\text{eigen}_k$  denotes the  $k$ th eigenvalue of the unbiased matrix estimate. Moreover, weights for particular detail components have been calculated by combining the corresponding detail strengths. Generally, weights are usually based on the strengths of detail, not on the actual intensity values.

$$wt(i,j) = HdetailStrength(i,j) + VdetailStrength(i,j) \quad (15)$$

After weight computation for all the detailed coefficients based on both assigned source images, the weighted average is employed for obtaining fusion results. Where  $wt_A$  and  $wt_B$  are the weights of the detailed coefficients obtained from the output of edge-preserving processing  $I_{out1}$  and  $I_{out2}$  that belong to corresponding input images  $A$  and  $B$ , respectively. Thus, the weighted average on both detailed images is estimated as a fusion result utilizing Eq. (16) as illustrated below:

$$F(i,j) = \frac{A(i,j) wt_A(i,j) + B(i,j) wt_B(i,j)}{wt_A(i,j) + wt_B(i,j)} \quad (16)$$

The proposed algorithm validates competency using a variety of medical datasets. The proposed method reduces computational complexity and execution time while improving diagnostic computing accuracy. Our method exhibits a major boost in efficiency compared to other competing algorithms. It can be observed that the qualitative interpretation of all the fused images retrieved through the proposed algorithm has superior visual quality.

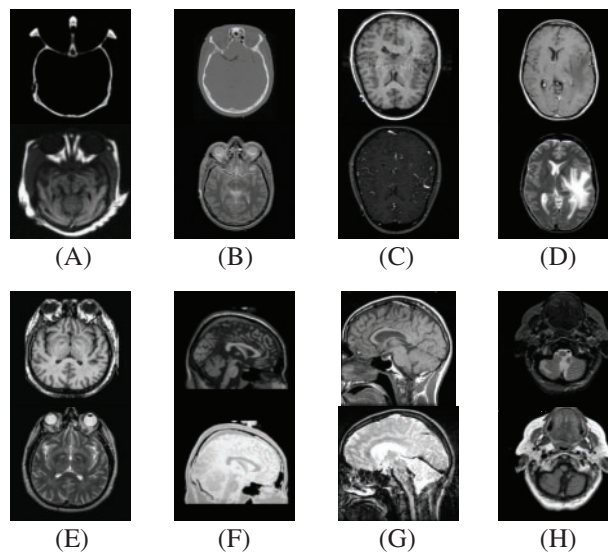
#### 4 Experiments and Results

In this study, eight pairs of medical images are employed during the experiments to validate the practicability and feasibility of our algorithm. A total of eight pairs of CT-MRI images are demonstrated in Fig. 2. All these test datasets contain a spatial resolution of  $256 \times 256$  pixels.

Further, seven existing comparative methods are utilized for comparison to assess the robustness of our approach, including the adolescent identity search algorithm in the NSST domain (AISA-NSST) [28], the medical image fusion method by deep learning (MIFDL) [40], multi-objective differential evolution based deep neural networks (MDE-DNN [41], geometric algebra discrete cosine metrics (GADCT) [43], medical assistive technology management based on hybrid domain filtering (MATM-HDF) [8], FDGNet [57] and sparse representation and Siamese convolutional neural network (SRSCNN [58]. All the experiments are performed on a PC with an Intel (R) Core (TM) i3-7020 CPU @ 2.3 GHz and 8 GB of RAM using MATLAB 2022a. Following [49], the parameter values are fixed at  $\sigma_s \sim 1.8$ ,  $\sigma_r \sim 4$  and kernel size ( $n$ )  $\sim 7$  for some pairs of medical images, and  $\delta_s \sim 1.2$ ,  $\delta_r \sim 0.06$  and kernel size ( $n$ )  $\sim 9$  ccs. The neighboring window of  $11 \times 11$  for the cross-bilateral filter and the neighboring window of  $5 \times 5$  is used to find the detail strengths. Details of parameter consequences are studied subsequently. Further, in [53], the edge-preserving processing parameters are set to amplitude threshold  $t = 0.15$ , threshold size  $s = 5$ , coefficient  $c = 1.2$ , and sigma of the Gaussian filter  $\sigma = 72$ .

#### 4.1 Objective Evaluation Metrics

This study employs twelve quality evaluation metrics for quantitative assessment to determine the validity of our method, namely spatial frequency (SF), average pixel intensity (API), or mean ( $\bar{F}$ ), mutual information (MI), fusion symmetry (FS), entropy (H), average gradient (AG), standard deviation (SD), and correlation coefficient (CC) [49,59]. Along with these, objective fusion metrics [49] based on gradient details are also considered. To provide a detailed study of fusion efficiency through evaluation of overall fusion efficiency ( $Q^{AB/F}$ ), information loss during the fusion process ( $L^{AB/F}$ ) and noise or artifact-produced fusion process ( $N^{AB/F}$  or  $N_m^{AB/F}$ ). The higher values in terms of some metrics and the lower values in terms of other metrics indicate better fusion quality. The process for computing the parameters and their symbolic representation is given below:



**Figure 2:** Source paired images (A–H): CT-MRI image pairs. Top  $\sim$  CT image; Bottom  $\sim$  MRI image [60]

- 1) The standard deviation represents the variance square root and is defined as the spread of information. It is expressed as

$$SD = \frac{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (f(i,j) - \bar{F})^2}}{mn} \quad (17)$$

- 2) The average gradient computes the degree of clarity and sharpness. It is expressed as

$$AG = \frac{\sum_i \sum_j ((f(i,j) - f(i+1,j))^2 + (f(i,j) - f(i,j+1))^2)^{\frac{1}{2}}}{mn} \quad (18)$$

- 3) Mutual information computes the total mutual information between a source and the fused image. It is denoted by

$$MI = MI_{AF} + MI_{BF} \quad (19)$$

where

$$MI_{AF} = \sum_k \sum_l p_{A,F}(k,l) \log_2 \left( \frac{p_{A,F}(k,l)}{p_A(k)p_F(l)} \right)$$

$$MI_{BF} = \sum_k \sum_l p_{B,F}(k,l) \log_2 \left( \frac{p_{B,F}(k,l)}{p_B(k)p_F(l)} \right)$$

- 4) Entropy (H) is the information quantity present in the image.

$$H = - \sum_{k=0}^{255} p_k \log_2(p_k) \quad (20)$$

where  $p_k$  represents the value of the intensity probability  $k$  in the image.

- 5) Spatial frequency computes the total information present in image regions and is denoted as

$$SF = \sqrt{RF^2 + CF^2} \quad (21)$$

where

$$RF = \frac{\sqrt{\sum_i \sum_j (f(i,j) - f(i,j-1))^2}}{mn} \quad (22)$$

and

$$CF = \frac{\sqrt{\sum_i \sum_j (f(i,j) - f(i-1,j))^2}}{mn} \quad (23)$$

- 6) Average pixel intensity (APF) represents the contrast index and is denoted as

$$API = \bar{F} = \frac{\sum_{i=1}^m \sum_{j=1}^n f(i,j)}{mn} \quad (24)$$

where  $m \times n$  represents the image dimension and  $f(i,j)$  represents the intensity of the pixel at  $(i,j)$ .

- 7) The correlation coefficient estimates the relevance of representing fusion output to the original image. It is defined by

$$CC = (r_{AF} + r_{BF})/2 \quad (25)$$

$$\text{where } r_{AF} = \frac{\sum_i \sum_j (a(i,j) - \bar{A})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (a(i,j) - \bar{A})^2) (\sum_i \sum_j (f(i,j) - \bar{F})^2)}} \quad (26)$$

$$\text{and } r_{BF} = \frac{\sum_i \sum_j (a(i,j) - \bar{B})(f(i,j) - \bar{F})}{\sqrt{(\sum_i \sum_j (a(i,j) - \bar{B})^2) (\sum_i \sum_j (f(i,j) - \bar{F})^2)}} \quad (26)$$

8) Fusion symmetry represents the amount of symmetry where the fused image corresponds to the input image. It is denoted as

$$FS = 2 - \left| \frac{MI_{AF}}{MI} - 0.5 \right| \quad (27)$$

9)  $Q^{AB/F}$  = overall information transferred from input images to the fusion result.

10)  $L^{AB/F}$  = overall information loss during the fusion process.

11)  $N^{AB/F}$  = noise or artifacts add up to the fusion result during the fusion process.

In particular,  $Q^{AB/F}$ ,  $L^{AB/F}$  and  $N^{AB/F}$  are complementary, and the summation of all these parameters should result in 1 [49], i.e.,

$$Q^{AB/F} + L^{AB/F} + N^{AB/F} = 1 \quad (28)$$

In particular, for most of the situations, it has been noticed that the above summation may not lead to unity. For this purpose, the technique has been modified, and a revision has been advised for performance parameters for fusion artifacts [49]. The revised computation is denoted by

$$N_m^{AB/F} = \frac{\sum_{\forall i} \sum_{\forall j} AM_{ij} [(1 - Q_{ij}^{AF}) w_{ij}^A + (1 - Q_{ij}^{BF}) w_{ij}^B]}{\sum_{\forall i} \sum_{\forall j} (w_{ij}^A + w_{ij}^B)} \quad (29)$$

where  $AM_{ij} = \begin{cases} 1, & g_{ij}^F > g_{ij}^A \text{ and } g_{ij}^F > g_{ij}^B \\ 0, & \text{otherwise} \end{cases}$  indicates the location of fusion artifacts where fuse-gradient is more powerful than input.

$g_{ij}^A$ ,  $g_{ij}^B$  &  $g_{ij}^F$  represents edge strengths of  $A$ ,  $B$ , and  $F$ , respectively.

$Q_{ij}^{AF}$  and  $Q_{ij}^{BF}$  represent gradient details preservation estimates of original-images A & B, respectively.

$w_{ij}^A$  and  $w_{ij}^B$  represent perceptual weights of original images A & B, respectively.

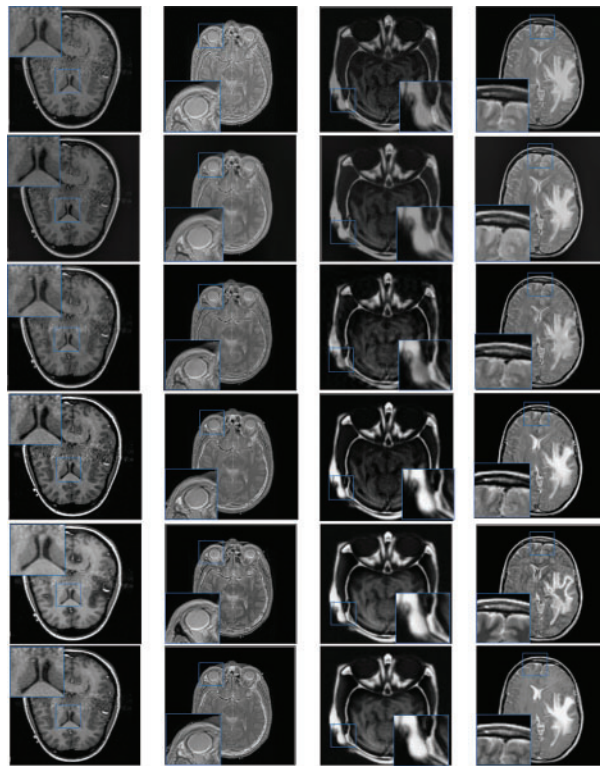
With the revision of fusion artifacts measure  $N_m^{AB/F}$ , we could re-write Eq. (28) as

$$Q^{AB/F} + L^{AB/F} + N_m^{AB/F} = 1 \quad (30)$$

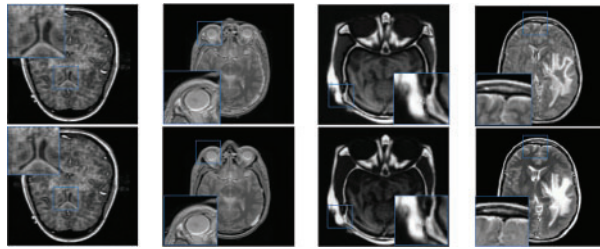
## 4.2 Qualitative and Quantitative Analysis

To investigate the efficiency of our proposed strategy, this study utilizes eight pairs of publicly accessible images from the Whole Brain Atlas [60]. The eight pairs of medical images in Figs. 2A–2H are represented by Set A to Set H, respectively. Statistical numeric indicators to validate the objective fusion performance is considered, including MI, H, API, CC, AG, API, SF, and SD. Further, in addition to the above metrics, an objective fusion characterization based on gradient details is also considered including  $Q^{AB/F}$ ,  $L^{AB/F}$ ,  $N^{AB/F}$  and  $N_m^{AB/F}$  [49]. The fused images of MRI and CT datasets

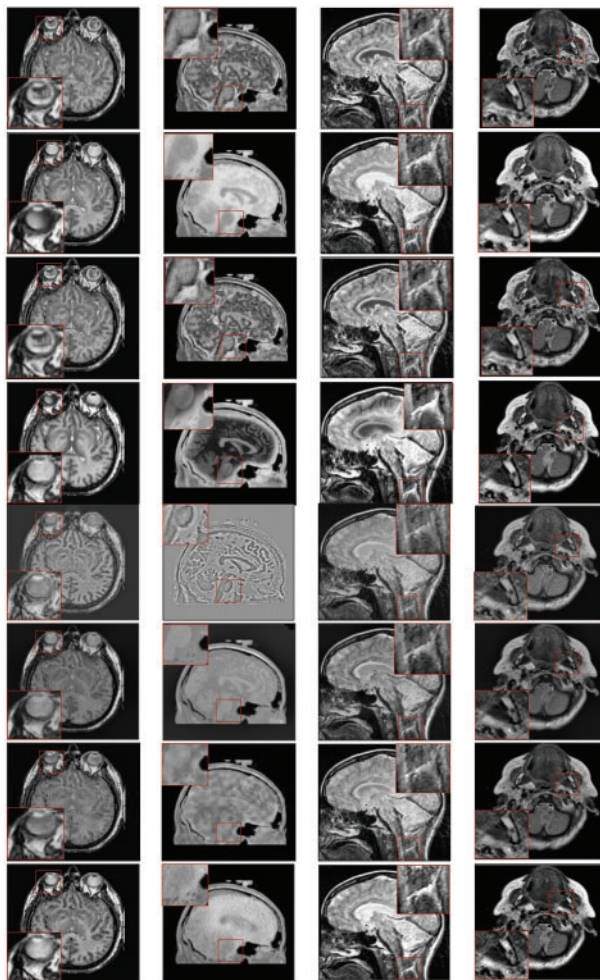
by comparative eight algorithms and our algorithm are demonstrated in Figs. 3 and 4. Figs. 3 and 4 demonstrate the results of Set (A-H) image datasets and observe the visual results of different algorithms. The SRSCNN and the proposed method were satisfactory efficient or productive in extracting the maximum functionality characteristics from input images. In the case of Set A-D; AISA-NSST, MIFDL, and MDE-DNN lose bone structure information; however, MATM-HDF, FDGNet, GADCT, SRSCNN, and our approach presents desired contrast, though it has been observed that the fusion performances of AISA-NSST, MIFDL, and MDE-DNN generate poor contrasts which make it challenging to differentiate tissues inside the brain. The efficiency of GADCT seems to have a superior outcome of contrasting preservation; however, the edge-information throughout the closeup has been distorted, blurred, smeared, and its contrast has been reduced. FDGNet-generated fusion findings have enhanced information about the structure and carry appropriate contrasting intensities. However, certain data and contrast information are lost as displayed in the closeup, and the output generated by MATM-HDF lost particulars of soft tissues (see Figs. 3 and 4). Considering the fusion results of Set C and Set D, we notice that the AISA-NSST, MIFDL, MDE-DNN, and GADCT retain information from two source images somehow, but they maintain poor contrast. However, by highlighting the zoomed region in Set D, the bone-skeletal region is completely noticeable in our proposed method. The objective analysis of several fusion algorithms on eight pairs of CT-MRI datasets has been tabulated in Tables 1 and 2.



**Figure 3:** (Continued)



**Figure 3:** Qualitative results of CT-MRI image pairs on Set (A–D). From top to bottom: results by AISA-NSST, MIFDL, MDE-DNN, GADCT, MATM-HDF, SRSCNN, FDGNet and proposed algorithm



**Figure 4:** Qualitative results of CT-MRI image pairs on Set (E–F). From top to bottom: results by AISA-NSST, MIFDL, MDE-DNN, GADCT, MATM-HDF, SRSCNN, FDGNet and proposed algorithm

**Table 1:** Quantitative indices of different methods based on reference-based metrics for Set (A–D). The best results are marked in bold, and the second-best results are marked underlined

Datasets	Metric	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF [8]	SRSCNN [58]	FDGNet [57]	Proposed
Set-A	API	32.22	28.98	40.77	54.54	49.77	38.65	53.25	56.57
	SD	35.70	45.98	42.80	46.77	52.88	44.06	58.27	59.66
	AG	7.13	10.29	9.76	10.77	10.98	8.43	10.17	11.35
	H	4.58	4.53	4.77	5.88	5.98	6.58	6.70	6.76
	MI	2.87	2.92	3.98	4.88	4.87	4.10	5.42	5.96
	FS	0.67	0.96	1.12	1.34	1.34	1.59	1.58	1.62
	CC	0.49	0.49	0.56	0.57	0.60	0.62	0.63	0.65
	SF	12.63	16.54	16.88	18.87	18.87	19.85	19.30	20.95
Set-B	API	42.39	42.97	42.77	43.17	43.45	43.63	43.87	44.68
	SD	50.10	51.26	51.78	52.77	52.77	53.63	53.23	55.01
	AG	7.40	8.25	8.76	8.7763	8.77	8.46	9.00	9.98
	H	3.50	3.53	4.43	4.65	5.23	5.18	5.20	5.29
	MI	2.76	2.92	3.77	3.77	4.15	4.35	4.60	5.31
	FS	0.67	0.97	1.09	1.36	1.56	1.90	1.89	1.98
	CC	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.98
	SF	16.92	17.13	16.66	16.77	17.78	19.36	18.15	20.97
Set-C	API	45.62	45.86	47.88	49.88	51.98	50.37	52.75	53.36
	SD	45.32	45.90	47.87	49.87	52.28	55.25	57.04	57.23
	AG	9.47	10.69	10.87	13.87	13.98	14.00	13.92	14.13
	H	3.58	4.53	4.87	5.24	5.98	6.47	6.33	6.53
	MI	1.86	2.32	2.98	3.08	3.27	3.76	3.86	3.96
	FS	0.89	0.98	1.07	1.45	1.67	1.89	1.90	1.92
	CC	0.93	0.91	0.93	0.98	1.05	1.02	1.00	1.13
	SF	16.49	19.10	20.87	21.87	23.88	24.38	24.09	25.31
Set-D	API	50.75	51.97	50.87	51.87	52.80	52.89	53.02	53.08
	SD	41.03	45.60	50.87	55.77	64.76	62.85	64.17	65.42
	AG	7.64	8.28	9.76	10.76	10.87	10.90	11.04	11.24
	H	2.64	3.53	3.88	3.89	3.98	4.35	4.24	4.75
	MI	2.58	2.92	2.87	3.17	3.40	3.19	3.35	3.65
	FS	0.87	0.98	1.08	1.11	1.86	1.82	1.84	2.00
	CC	0.70	0.79	0.89	0.90	0.94	0.93	0.95	0.97
	SF	16.88	18.96	20.77	21.88	22.26	22.86	23.09	23.49

**Table 2:** Quantitative indices of different methods based on reference-based metrics for Set (E–H). The best results are marked in bold and the second-best results are marked underlined

Datasets	Metric	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF[8]	SRSCNN [58]	FDGNet [57]	Proposed
	API	59.38	60.76	62.77	64.65	65.66	67.86	67.40	70.10

(Continued)

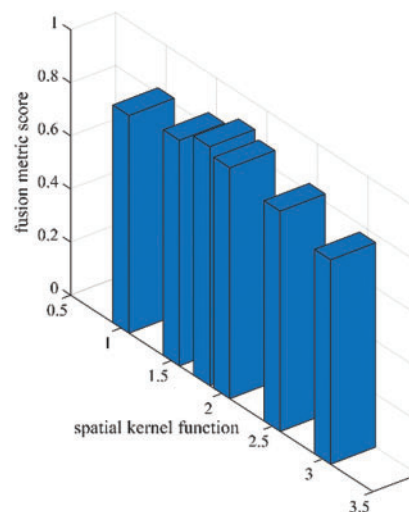


**Table 2 (continued)**

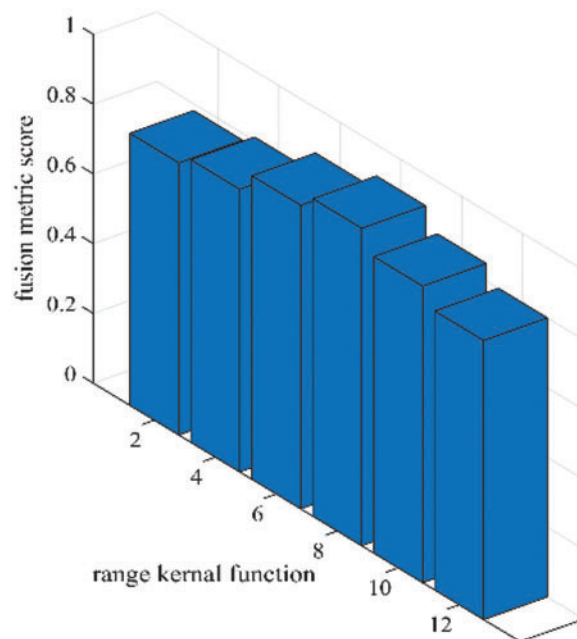
Datasets	Metric	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF[8]	SRSCNN [58]	FDGNet [57]	Proposed
Set-E	SD	57.55	58.77	60.66	61.77	60.87	64.53	63.37	67.34
	AG	9.88	10.76	11.87	11.88	12.76	12.80	13.45	15.67
	H	4.58	4.76	5.12	5.26	5.39	5.45	5.53	6.01
	MI	2.98	3.07	3.24	4.06	4.50	3.99	4.40	4.70
	FS	1.78	1.78	1.72	1.85	1.87	1.84	1.83	1.91
	CC	0.70	0.80	0.82	0.89	0.90	0.89	0.94	0.96
	SF	18.36	20.86	19.76	21.87	21.87	24.35	23.02	26.22
Set-F	API	55.66	55.86	56.77	56.79	57.26	57.42	58.47	58.79
	SD	60.76	61.77	61.56	62.77	62.76	64.34	64.25	65.75
	AG	8.88	9.76	10.76	10.88	11.08	11.21	11.49	11.96
	H	3.87	3.98	3.90	4.01	4.14	4.41	4.26	4.87
	MI	1.98	1.90	2.00	2.12	2.78	2.89	3.13	3.24
	FS	1.78	1.88	1.89	1.90	1.19	1.90	1.89	1.94
	CC	0.70	0.80	0.90	0.97	0.98	1.10	1.00	1.23
SF	16.76	18.76	19.73	20.77	20.98	21.15	21.33	21.85	
Set-G	API	94.96	94.77	95.77	95.66	96.88	97.44	98.44	98.80
	SD	67.95	68.66	68.65	69.76	71.75	71.88	71.22	72.48
	AG	19.23	19.77	20.87	20.88	21.67	21.98	22.29	22.72
	H	6.58	6.77	6.87	7.09	7.21	7.61	7.57	7.77
	MI	2.88	2.99	3.09	3.12	3.22	3.37	3.46	3.78
	FS	0.87	0.98	1.07	1.24	1.67	1.73	1.59	1.89
	CC	0.65	0.71	0.67	0.71	0.85	0.81	0.79	0.90
SF	25.81	26.76	28.75	27.66	29.67	29.84	30.34	30.77	
Set-H	API	39.51	40.28	43.77	44.88	45.87	43.73	47.56	47.81
	SD	47.69	43.57	53.73	55.77	55.77	57.90	57.38	58.84
	AG	9.05	10.11	11.72	11.35	11.56	10.05	12.15	12.32
	H	4.58	4.53	3.98	4.77	5.20	5.06	5.14	5.43
	MI	2.88	2.92	2.97	3.07	3.39	3.24	3.36	3.65
	FS	0.88	1.10	1.09	1.45	1.56	1.85	1.83	1.90
	CC	0.79	0.81	0.83	0.85	0.84	0.87	0.86	0.91
SF	15.10	20.31	17.77	18.77	19.66	18.82	21.13	21.53	

As per the experimental results, the outputs retrieved via the proposed method are very efficient and effective in the case of all quantitative parameters, whereas other algorithms yielded suboptimal values. Tables 1 and 2 show that our technique ranks first in simulation results for all dataset pairs when compared to other competing techniques. Even if a few parameters are higher with smaller margins than other methodologies, the differences in other parameters are higher. Further, the average result of these quantitative measures are shown in Tables 1 and 2 to provide a more intuitive grasp of quantitative evaluations in comparison to existing methodologies. In summary, the fusion results of the proposed algorithm outperform the existing seven comparative approaches in all aspects. Tables 1 and 2 exhibit indices incusing H, AG, SF, API, MI, SD, CC, and FS. At higher values of the above-mentioned metrics, the efficiency of the fused image is appropriate. Similarly, in addition to objective analysis, visual interpretation is more significant. Figs. 3 and 4 show the corresponding fused images to illustrate the performance visually. The proposed algorithm exhibits superior visual interpretation

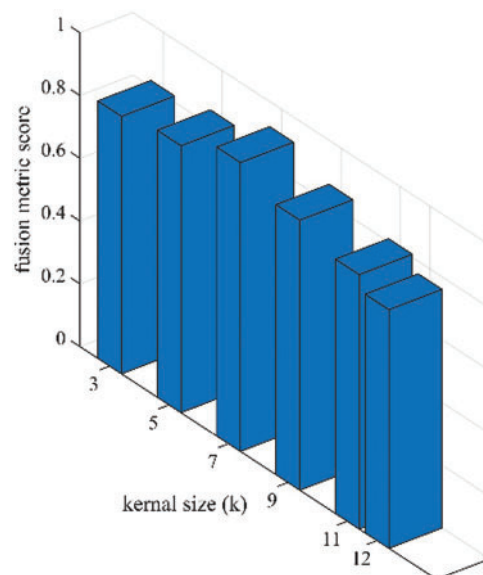
when compared to other methods. It also facilitates the process of testing other algorithms, which have comparably low experimental results. For the assessment of performance and to show the efficiency of our method, it observed that the effects of change in input for  $k$ ,  $\sigma_s$  and  $\sigma_r$  w.r.t the variation in CC, SD, FS, MI, AG, H, SF, and API metrics for different pairs of medical images. We have studied the variation for the value of  $k$ ,  $\sigma_s$  and  $\sigma_r$  and its effects on the performance of our algorithm, as demonstrated in Figs. 5–7. The given parameters in terms of robustness have shown variation for different values of  $k$ ,  $\sigma_s$  and  $\sigma_r$  in a wider range. In addition to the consequences of  $k$ ,  $\sigma_s$  and  $\sigma_r$ , we also observe variation in terms of iteration, threshold size (s), amplitude threshold (t), and contrast coefficient (c) during an experimental operation. For qualitative evaluation for visual analysis is concerned, the fusion results of our algorithm show excellent performance where the details transfer rate is higher. The qualitative results of the proposed algorithm are excellent when compared to other competing algorithms for all image pairs. When compared to existing comparative fusion algorithms, it can be noted that the fusion results for both reference and no-reference-based criteria are excellent. Similarly, the visual interpretation of our method for all the image pairs shows better performance. It can be stated that our algorithm obtains the advantages of normalized weight computation. In addition to that, edge-preserving processing and pixel significance further contribute to computing the optimized weights to generate improved fusion results. Due to this fact, the visibility increases in the fusion results when compared to other approaches, which show minimal information loss and maximum information transfer. In addition, the information loss and noise are substantially higher when using other competing approaches. The proposed method outperforms competing fusion methods across all image pairs in terms of reference-based criteria. In some cases, the quantitative evaluation performance is marginally higher for our algorithm as compared to competing algorithms. Hence, it could be stated that our technique is capable of generating better results than other competing techniques, both in objective evaluation and qualitative interpretation. The proposed method incorporates large-scale details from original images into final fusion results, and small-scale details are recovered in the surrounding of large-scale details. In terms of performance metrics and visual interpretation, it has been validated that our technique has demonstrated exemplary performance compared to other comparative algorithms.



**Figure 5:** Sensitivity analysis of spatial kernel function ( $\sigma_s$ ) in terms of fusion performance on the proposed algorithm



**Figure 6:** Sensitivity analysis of range kernel function ( $\sigma_r$ ) in terms of fusion performance on the proposed algorithm



**Figure 7:** Sensitivity analysis of kernel size ( $k$ ) in terms of fusion performance on the proposed algorithm

#### 4.3 Further Discussion

In addition to the above eight performance metrics, the results of four fusion performance metrics based on gradient details have been discussed in this section [39]. These metrics provide a detailed evaluation of fusion results by evaluating overall details transferred from source images to fusion

result  $Q^{AB/F}$ , overall information loss  $L^{AB/F}$ , artifacts or noise added in fused images during the fusion artefacts measure  $N_m^{AB/F}$ .

Tables 3 and 4 show that our algorithm nearly always gets ideal values on all objective measures as compared to other competing algorithms that intuitively adapt to quality judgments. In summary, the fusion performance of our strategy surpasses all other algorithms in both qualitative interpretation and quantitative assessment. The overall performance measures of evaluation metrics are tabulated in Tables 1 and 2 where the fusion performances of our algorithm are superior with ideal scores in all these eight-objective metrics. The principal aim of image fusion is to produce an output that is highly suitable for human visualization with adequate, comprehensive, and detailed features in the fusion results. Further, visual interpretation is quite important, along with objective performance. The fusion results are demonstrated in Figs. 3 and 4 which interpret performance visually. The proposed approach has demonstrated superior visualization when compared to other comparable methods. The information transfer rate in the proposed technique is the highest when compared to other algorithms. according to the quantitative assessment in the case of gradient-based measures. In addition, the suggested method has a high  $Q^{AB/F}$  and small  $L^{AB/F}$ ,  $N_m^{AB/F}$  across all pairs of images when comparison to other competing algorithms which is ideal for better performance. It can be noticed that the suggested algorithm's non-reference-based measure values are higher than results obtained from other competing approaches. The visual interpretation is perceived to a significant extent with the approach outlined in the manuscript. According to the results obtained from Tables 3 and 4, the fusion results of four methods named AISA-NSST, MIFDL, MDE-DNN, and GADCT show that the fusion performance of the mentioned methods suffered from distortions as well as the soft tissues that were missing (see Figs. 3 and 4). The contrast of the fusion outputs from both MATM-HDF, SRSCNN, and FDGNet is slightly less than our technique (see Figs. 3 and 4). As with AISA-NSST, MIFDL, MDE-DNN, and GADCT, the soft tissues of original images by the SRSCNN and FDGNet methods are slightly missing (see Figs. 3 and 4). The contrast of the MATM-HDF algorithm is slightly low as compared to our method. Table 5 presents the computational time for execution of all methods (i.e., AISA-NSST, MIFDL, MDE-DNN, GADCT, MATM-HDF, SRSCNN, FDGNet and the proposed method kernel size (k)). We can see that the GADCT, MATM-HDF, and SRSCNN methods have the longest running times. This is so because deep learning models take longer processing time during the training and testing phases. Other approaches take less time than these three models. Overall, our method requires much less running time for all of the test images when compared to the other competing methods. Thus, in the case of all evaluation aspects, our algorithm presents desirable fusion results in terms of clinical applicability.

**Table 3:** Quantitative indices of different methods based on gradient-based metrics for Set (A–D). The best results are marked in bold and the second-best results are marked underlined

Datasets	Metrics	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF [8]	SRSCNN [58]	FDGNet [57]	Proposed
Set-A	$Q^{AB/F}$	0.49	0.77	0.73	0.79	0.73	0.82	0.88	0.91
	$L^{AB/F}$	0.23	0.13	0.14	0.13	0.13	0.09	0.11	0.07
	$N_m^{AB/F}$	0.21	0.17	0.13	0.12	0.13	0.14	0.13	0.10
	$N_m^{AB/F}$	0.10	0.14	0.11	0.09	0.06	0.04	0.05	0.01

(Continued)

**Table 3 (continued)**

Datasets	Metrics	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF [8]	SRSCNN [58]	FDGNet [57]	Proposed
Set-B	$Q^{AB/F}$	0.52	0.56	0.55	0.69	0.57	0.71	0.76	0.79
	$L^{AB/F}$	0.28	0.26	0.24	0.20	0.27	0.25	0.24	0.18
	$N^{AB/F}$	0.18	0.23	0.20	0.15	0.14	0.09	0.10	0.07
	$N_m^{AB/F}$	0.13	0.12	0.15	0.07	0.17	0.05	0.06	0.02
Set-C	$Q^{AB/F}$	0.56	0.56	0.57	0.67	0.58	0.83	0.82	0.84
	$L^{AB/F}$	0.23	0.21	0.23	0.16	0.23	0.14	0.17	0.12
	$N^{AB/F}$	0.21	0.23	0.23	0.13	0.16	0.11	0.12	0.09
	$N_m^{AB/F}$	0.16	0.12	0.15	0.08	0.1422	0.05	0.04	0.02
Set-D	$Q^{AB/F}$	0.46	0.55	0.56	0.69	0.56	0.81	0.84	0.85
	$L^{AB/F}$	0.17	0.20	0.24	0.14	0.20	0.18	0.15	0.13
	$N^{AB/F}$	0.22	0.23	0.18	0.14	0.21	0.07	0.08	0.04
	$N_m^{AB/F}$	0.13	0.17	0.12	0.08	0.10	0.04	0.06	0.01

**Table 4:** Quantitative indices of different methods based on gradient-based metrics for Set (E–H). The best results are marked in bold and the second-best results are marked underlined

Datasets	Metrics	AISA-NSST [28]	MIFDL [40]	MDE-DNN [41]	GADCT [43]	MATM- HDF [8]	SRSCNN [58]	FDGNet [57]	Proposed
Set-E	$Q^{AB/F}$	0.39	0.44	0.50	0.65	0.50	0.72	0.76	0.79
	$L^{AB/F}$	0.24	0.24	0.21	0.17	0.22	0.27	0.21	0.16
	$N^{AB/F}$	0.18	0.20	0.22	0.16	0.21	0.11	0.12	0.10
	$N_m^{AB/F}$	0.10	0.15	0.14	0.14	0.17	0.07	0.06	0.03
Set-F	$Q^{AB/F}$	0.53	0.51	0.55	0.67	0.50	0.77	0.80	0.81
	$L^{AB/F}$	0.23	0.23	0.13	0.13	0.25	0.21	0.16	0.14
	$N^{AB/F}$	0.21	0.21	0.25	0.18	0.21	0.17	0.15	0.13
	$N_m^{AB/F}$	0.17	0.14	0.12	0.12	0.16	0.06	0.09	0.04
Set-G	$Q^{AB/F}$	0.49	0.49	0.50	0.61	0.50	0.76	0.75	0.79
	$L^{AB/F}$	0.22	0.23	0.24	0.12	0.22	0.21	0.19	0.18
	$N^{AB/F}$	0.17	0.27	0.26	0.12	0.19	0.09	0.08	0.06
	$N_m^{AB/F}$	0.12	0.12	0.15	0.09	0.12	0.04	0.06	0.01
Set-H	$Q^{AB/F}$	0.49	0.49	0.55	0.67	0.51	0.75	0.76	0.79
	$L^{AB/F}$	0.22	0.23	0.22	0.20	0.23	0.20	0.19	0.17
	$N^{AB/F}$	0.21	0.20	0.18	0.14	0.23	0.12	0.15	0.09
	$N_m^{AB/F}$	0.12	0.12	0.12	0.06	0.14	0.10	0.09	0.03

**Table 5:** Average running times for the different image fusion algorithms

Method	Average running time (s)
AISA-NSST [28]	0.0341
MIFDL [40]	0.0485
MDE-DNN [41]	0.0627
GADCT [43]	0.0727
MATM-HDF [8]	0.0667
SRSCNN [58]	0.0586
FDGNet [57]	0.0378
Proposed	0.0267

## 5 Conclusions

This study presents a competent medical image fusion strategy based on pixel significance with edge-preserving processing. First, the proposed method obtains detailed images obtained from the source images after subtracting the outputs of the cross-bilateral filters. These detailed images are then used for edge-preserving processing that combines a linear low-pass filter with the non-linear technique, which allows the selection of meaningful regions while preserving edges. The relevant restored regions are fused with the result of low-pass filtering for the reconstruction of the original shape of the edges. These images are then used for weight computations. These weights are calculated by estimating the strength of vertical and horizontal details and are directly merged with source images to retrieve the final fusion result. The proposed algorithm validates competency by utilizing eight sets of CT-MRI medical images. The proposed approach decreases computational complexity and execution time while improving diagnostic computing accuracy. Our method exhibits a major boost in efficiency compared to other competing algorithms. It can be observed that the qualitative interpretation of all the fused images retrieved through the proposed algorithm has superior visual quality. The key limitation of the proposed method is that it is not highly appropriate for fused, sophisticated multimodality images. The proposed method could be augmented by using other different types of filters that have been left for future research. Further, image fusion and the importance of this technique have huge prospects for improvement by proposing several image fusion algorithms to reduce noise and artifacts. In addition, it is quite realistic to boost the superiority of medical images by further improving the proposed algorithm in the future.

**Acknowledgement:** We are thankful to the reviewers and editors for their valuable suggestions. We appreciate their insightful comments, which helped us improve the manuscript significantly.

**Funding Statement:** No funding was received to assist with the preparation of this manuscript.

**Author Contributions:** Bhawna Goyal: Methodology, Data curation, Writing–original draft and Formal analysis. Ayush Dogra: Supervision, Writing–original draft and Formal analysis, Validation. Dawa Chyophel Lepcha: Software, Investigation, Writing–review and editing. Rajesh Singh: Investigation and Resources. Hemant Sharma: Formal analysis and Resources. Ahmed Alkhayyat: Formal analysis and editing. Manob Jyoti Saikia: Formal analysis and Validation.

**Availability of Data and Materials:** The dataset used in the manuscript can be accessed from <https://www.med.harvard.edu/aanlib/home.html>.

**Conflicts of Interest:** 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.

## References

- [1] S. Karim, G. Tong, J. Li, A. Qadir, U. Farooq and Y. Yu, "Current advances and future perspectives of image fusion: A comprehensive review," *Inform. Fusion*, vol. 90, pp. 185–217, Feb. 2023. doi: [10.1016/j.inffus.2022.09.019](https://doi.org/10.1016/j.inffus.2022.09.019).
- [2] M. Haribabu, V. Guruviah, and P. Yogarajah, "Recent advancements in multimodal medical image fusion techniques for better diagnosis: An overview," *Curr. Med. Imaging Rev.*, vol. 19, no. 7, pp. 3, Jun. 2022. doi: [10.2174/1573405618666220606161137](https://doi.org/10.2174/1573405618666220606161137).
- [3] C. Kaushal, S. Bhat, D. Koundal, and A. Singla, "Recent trends in computer assisted diagnosis (CAD) system for breast cancer diagnosis using histopathological images," *IRBM*, vol. 40, no. 4, pp. 211–227, Aug. 2019. doi: [10.1016/j.irbm.2019.06.001](https://doi.org/10.1016/j.irbm.2019.06.001).
- [4] P. K. Shukla, J. K. Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary and P. K. Shukla, "Multiobjective genetic algorithm and convolutional neural network based COVID-19 identification in chest X-ray images," *Math. Probl. Eng.*, vol. 2021, pp. 1–9, 2021. doi: [10.1155/2021/7804540](https://doi.org/10.1155/2021/7804540).
- [5] D. C. Lepcha *et al.*, "Multimodal medical image fusion based on pixel significance using anisotropic diffusion and cross bilateral filter," *Hum.-Centric Comput. Inform. Sci.*, vol. 12, pp. 15, 2022. doi: [10.22967/HCCIS.2022.12.015](https://doi.org/10.22967/HCCIS.2022.12.015).
- [6] B. Goyal, A. Dogra, R. Khoond, D. C. Lepcha, V. Goyal and S. L. Fernandes, "Medical image fusion based on anisotropic diffusion and non-subsampled contourlet transform," *Comput. Mater. Contin.*, vol. 76, no. 1, pp. 311–327, Jun. 2023. doi: [10.32604/cmc.2023.038398](https://doi.org/10.32604/cmc.2023.038398).
- [7] J. Chen, L. Chen, and M. Shabaz, "Image fusion algorithm at pixel level based on edge detection," *J. Healthc. Eng.*, vol. 2021, no. 7, pp. 1–10, 2021. doi: [10.1155/2021/5760660](https://doi.org/10.1155/2021/5760660).
- [8] B. Goyal *et al.*, "Multi-modality image fusion for medical assistive technology management based on hybrid domain filtering," *Expert Syst. Appl.*, vol. 209, no. 2, pp. 118283, Dec. 2022. doi: [10.1016/j.eswa.2022.118283](https://doi.org/10.1016/j.eswa.2022.118283).
- [9] J. Chen, X. Li, L. Luo, X. Mei, and J. Ma, "Infrared and visible image fusion based on target-enhanced multiscale transform decomposition," *Inf. Sci.*, vol. 508, no. 4, pp. 64–78, Jan. 2020. doi: [10.1016/j.ins.2019.08.066](https://doi.org/10.1016/j.ins.2019.08.066).
- [10] Z. Zhu, M. Zheng, G. Qi, D. Wang, and Y. Xiang, "A phase congruency and local laplacian energy based multi-modality medical image fusion method in NSCT domain," *IEEE Access*, vol. 7, pp. 20811–20824, 2019. doi: [10.1109/ACCESS.2019.2898111](https://doi.org/10.1109/ACCESS.2019.2898111).
- [11] S. Maqsood and U. Javed, "Multi-modal medical image fusion based on two-scale image decomposition and sparse representation," *Biomed. Signal Process. Control*, vol. 57, no. 2, pp. 101810, Mar. 2020. doi: [10.1016/J.BSPC.2019.101810](https://doi.org/10.1016/J.BSPC.2019.101810).
- [12] L. Jiao and J. Zhao, "A survey on the new generation of deep learning in image processing," *IEEE Access*, vol. 7, pp. 172231–172263, 2019. doi: [10.1109/ACCESS.2019.2956508](https://doi.org/10.1109/ACCESS.2019.2956508).
- [13] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Z. Med. Phys.*, vol. 29, no. 2, pp. 86–101, May 2019. doi: [10.1016/J.ZEMEDI.2018.12.003](https://doi.org/10.1016/J.ZEMEDI.2018.12.003).
- [14] X. Y. Guo, J. Wang, and Z. J. Xu, "Multimodal medical image fusion based on average cross bilateral filtering," in *ACM Int. Conf. Proc. Ser.*, Dec. 2022, pp. 98–105. doi: [10.1145/3582197.3582213](https://doi.org/10.1145/3582197.3582213).
- [15] Y. Zhang *et al.*, "Medical image fusion based on quasi-cross bilateral filtering," *Biomed. Signal Process. Control*, vol. 80, no. 9, pp. 104259, 2023. doi: [10.1016/J.BSPC.2022.104259](https://doi.org/10.1016/J.BSPC.2022.104259).
- [16] Y. Goodfellow *et al.*, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 27, 2014.

- [17] Y. Chen *et al.*, “Generative adversarial networks in medical image augmentation: A review,” *Comput. Biol. Med.*, vol. 144, no. 1, pp. 105382, May 2022. doi: [10.1016/J.COMPBIOMED.2022.105382](https://doi.org/10.1016/J.COMPBIOMED.2022.105382).
- [18] L. Wang, W. Chen, W. Yang, F. Bi, and F. R. Yu, “A state-of-the-art review on image synthesis with generative adversarial networks,” *IEEE Access*, vol. 8, pp. 63514–63537, 2020. doi: [10.1109/ACCESS.2020.2982224](https://doi.org/10.1109/ACCESS.2020.2982224).
- [19] K. He *et al.*, “Transformers in medical image analysis,” *Intell. Med.*, vol. 3, no. 1, pp. 59–78, Feb. 2023. doi: [10.1016/J.IMED.2022.07.002](https://doi.org/10.1016/J.IMED.2022.07.002).
- [20] G. Li, D. Xu, X. Cheng, L. Si, and C. Zheng, “SimViT: Exploring a simple vision transformer with sliding windows,” in *IEEE Int. Conf. Multimed. Expo*, vol. 2022, Jul. 2021, pp. 1–6. doi: [10.1109/ICME52920.2022.9859907](https://doi.org/10.1109/ICME52920.2022.9859907).
- [21] J. Zhang, A. Liu, D. Wang, Y. Liu, Z. J. Wang, and X. Chen, “Transformer-based end-to-end anatomical and functional image fusion,” *IEEE Trans Instrum Meas*, vol. 71, pp. 1–11, 2022. doi: [10.1109/TIM.2022.3200426](https://doi.org/10.1109/TIM.2022.3200426).
- [22] K. Yu, X. Yang, S. Jeon, and Q. Dou, “An end-to-end medical image fusion network based on Swin-transformer,” *Microprocess. Microsyst.*, vol. 98, no. 3, pp. 104781, Apr. 2023. doi: [10.1016/j.micpro.2023.104781](https://doi.org/10.1016/j.micpro.2023.104781).
- [23] W. Li, Y. Zhang, G. Wang, Y. Huang, and R. Li, “DFENet: A dual-branch feature enhanced network integrating transformers and convolutional feature learning for multimodal medical image fusion,” *Biomed. Signal Process. Control*, vol. 80, no. 1, pp. 104402, Feb. 2023. doi: [10.1016/j.bspc.2022.104402](https://doi.org/10.1016/j.bspc.2022.104402).
- [24] H. Kaur, D. Koundal, and V. Kadyan, “Image fusion techniques: A survey,” *Arch. Comput. Methods Eng.*, vol. 28, no. 7, pp. 4425–4447, Dec. 2021. doi: [10.1007/S11831-021-09540-7/FIGURES/7](https://doi.org/10.1007/S11831-021-09540-7/FIGURES/7).
- [25] N. Tawfik, H. A. Elnemr, M. Fakhir, M. I. Dessouky, and F. E. Abd El-Samie, “Survey study of multimodality medical image fusion methods,” *Multimed. Tools Appl.*, vol. 80, no. 4, pp. 6369–6396, 2020. doi: [10.1007/s11042-020-08834-5](https://doi.org/10.1007/s11042-020-08834-5).
- [26] X. Han, J. Yu, J. H. Xue, and W. Sun, “Hyperspectral and multispectral image fusion using optimized twin dictionaries,” *IEEE Trans. Image Process.*, vol. 29, pp. 4709–4720, 2020. doi: [10.1109/TIP.2020.2968773](https://doi.org/10.1109/TIP.2020.2968773).
- [27] G. Li, Y. Lin, and X. Qu, “An infrared and visible image fusion method based on multi-scale transformation and norm optimization,” *Inform. Fusion*, vol. 71, no. 2, pp. 109–129, Jul. 2021. doi: [10.1016/j.inffus.2021.02.008](https://doi.org/10.1016/j.inffus.2021.02.008).
- [28] J. Jose *et al.*, “An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion,” *Biomed. Signal Process. Control*, vol. 66, no. 1, pp. 102480, Apr. 2021. doi: [10.1016/j.bspc.2021.102480](https://doi.org/10.1016/j.bspc.2021.102480).
- [29] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, “Medical image fusion via convolutional sparsity based morphological component analysis,” *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 485–489, Mar. 2019. doi: [10.1109/LSP.2019.2895749](https://doi.org/10.1109/LSP.2019.2895749).
- [30] N. Alseelawi, H. Tuama Hazim, and H. T. S. Alrikabi, “A novel method of multimodal medical image fusion based on hybrid approach of NSCT and DTCWT,” *Int. J. Online Biomed. Eng.*, vol. 18, no. 3, pp. 114–133, Mar. 2022. doi: [10.3991/ijoe.v18i03.28011](https://doi.org/10.3991/ijoe.v18i03.28011).
- [31] C. Zhang, “Multifocus image fusion using multiscale transform and convolutional sparse representation,” *Int. J. Wavelets, Multiresolut. Inf. Process.*, vol. 19, no. 1, pp. 2050061, Oct. 2020. doi: [10.1142/S0219691320500617](https://doi.org/10.1142/S0219691320500617).
- [32] D. C. Lepcha, B. Goyal, and A. Dogra, “Image fusion based on cross bilateral and rolling guidance filter through weight normalization,” *Open Neuroimag J.*, vol. 13, no. 1, pp. 51–61, Jan. 2021. doi: [10.2174/1874440002013010051](https://doi.org/10.2174/1874440002013010051).
- [33] A. Dogra, B. Goyal, and S. Agrawal, “Osseous and digital subtraction angiography image fusion via various enhancement schemes and Laplacian pyramid transformations,” *Future Gen. Comput. Syst.*, vol. 82, no. 1, pp. 149–157, May 2018. doi: [10.1016/J.FUTURE.2017.12.052](https://doi.org/10.1016/J.FUTURE.2017.12.052).
- [34] B. Goyal, D. Chyophel Lepcha, A. Dogra, V. Bhateja, and A. Lay-Ekuakille, “Measurement and analysis of multi-modal image fusion metrics based on structure awareness using domain transform filtering,” *Meas.*, vol. 182, no. 8, pp. 109663, Sep. 2021. doi: [10.1016/J.MEASUREMENT.2021.109663](https://doi.org/10.1016/J.MEASUREMENT.2021.109663).



- [35] M. Yin, X. Liu, Y. Liu, and X. Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampling shearlet transform domain," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 1, pp. 49–64, Jan. 2019. doi: [10.1109/TIM.2018.2838778](https://doi.org/10.1109/TIM.2018.2838778).
- [36] L. Yin, M. Zheng, G. Qi, Z. Zhu, F. Jin and J. Sim, "A novel image fusion framework based on sparse representation and pulse coupled neural network," *IEEE Access*, vol. 7, pp. 98290–98305, 2019. doi: [10.1109/ACCESS.2019.2929303](https://doi.org/10.1109/ACCESS.2019.2929303).
- [37] J. Xia, Y. Chen, A. Chen, and Y. Chen, "Medical image fusion based on sparse representation and PCNN in NSCT domain," *Comput. Math. Methods Med.*, vol. 2018, no. 5, pp. 1–12, 2018. doi: [10.1155/2018/2806047](https://doi.org/10.1155/2018/2806047).
- [38] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, "Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning," *Pattern Recognit.*, vol. 79, no. 3, pp. 130–146, Jul. 2018. doi: [10.1016/j.patcog.2018.02.005](https://doi.org/10.1016/j.patcog.2018.02.005).
- [39] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci.*, vol. 432, pp. 516–529, Mar. 2018. doi: [10.1016/j.ins.2017.09.010](https://doi.org/10.1016/j.ins.2017.09.010).
- [40] Y. Li, J. Zhao, Z. Lv, and J. Li, "Medical image fusion method by deep learning," *Int. J. Cogn. Comput. Eng.*, vol. 2, no. 9, pp. 21–29, Jun. 2021. doi: [10.1016/j.ijcce.2020.12.004](https://doi.org/10.1016/j.ijcce.2020.12.004).
- [41] M. Kaur and D. Singh, "Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 2, pp. 2483–2493, Feb. 2021. doi: [10.1007/s12652-020-02386-0](https://doi.org/10.1007/s12652-020-02386-0).
- [42] K. J. Xia, H. S. Yin, and J. Q. Wang, "A novel improved deep convolutional neural network model for medical image fusion," *Cluster Comput.*, vol. 22, no. 1, pp. 1515–1527, Feb. 2018. doi: [10.1007/S10586-018-2026-1](https://doi.org/10.1007/S10586-018-2026-1).
- [43] R. Wang, N. Fang, Y. He, Y. Li, W. Cao and H. Wang, "Multi-modal medical image fusion based on geometric algebra discrete cosine transform," *Adv. Appl. Clifford Algebr.*, vol. 32, no. 2, pp. 1–23, Apr. 2022. doi: [10.1007/s00006-021-01197-6](https://doi.org/10.1007/s00006-021-01197-6).
- [44] X. Li, F. Zhou, H. Tan, W. Zhang, and C. Zhao, "Multimodal medical image fusion based on joint bilateral filter and local gradient energy," *Inf. Sci.*, vol. 569, no. 7, pp. 302–325, Aug. 2021. doi: [10.1016/j.ins.2021.04.052](https://doi.org/10.1016/j.ins.2021.04.052).
- [45] W. Li, Q. Lin, K. Wang, and K. Cai, "Improving medical image fusion method using fuzzy entropy and nonsubsampling contourlet transform," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 1, pp. 204–214, Mar. 2021. doi: [10.1002/ima.22476](https://doi.org/10.1002/ima.22476).
- [46] Z. Wang, X. Li, H. Duan, Y. Su, X. Zhang and X. Guan, "Medical image fusion based on convolutional neural networks and non-subsampling contourlet transform," *Expert Syst. Appl.*, vol. 171, no. 5, pp. 114574, Jun. 2021. doi: [10.1016/j.eswa.2021.114574](https://doi.org/10.1016/j.eswa.2021.114574).
- [47] Z. Xu *et al.*, "LatLRR-FCNs: Latent low-rank representation with fully convolutional networks for medical image fusion," *Front. Neurosci.*, vol. 14, pp. 1387, Jan. 2021. doi: [10.3389/fnins.2020.615435](https://doi.org/10.3389/fnins.2020.615435).
- [48] Y. Liu *et al.*, "Multimodal medical image fusion based on the spectral total variation and local structural patch measurement," *Int. J. Imaging Syst. Technol.*, vol. 31, no. 1, pp. 391–411, Mar. 2021. doi: [10.1002/ima.22460](https://doi.org/10.1002/ima.22460).
- [49] B. K. Shreyamsha Kumar, "Image fusion based on pixel significance using cross bilateral filter," *Signal, Image Video Process.*, vol. 9, no. 5, pp. 1193–1204, 2013. doi: [10.1007/S11760-013-0556-9](https://doi.org/10.1007/S11760-013-0556-9).
- [50] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth Int. Conf. Comput. Vis. (IEEE Cat. No. 98CH36271)*, Bombay, India, 1998, pp. 839–846.
- [51] M. Zhang and B. K. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2324–2333, 2008. doi: [10.1109/TIP.2008.2006658](https://doi.org/10.1109/TIP.2008.2006658).
- [52] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. Graphics (TOG)*, vol. 23, no. 3, pp. 664–672, Aug. 2004. doi: [10.1145/1015706.1015777](https://doi.org/10.1145/1015706.1015777).

- [53] M. Iwanowski, "Edge-aware color image manipulation by combination of low-pass linear filter and morphological processing of its residuals," in *ICCVG 2020: Computer Vision and Graphics*, 2020, pp. 59–71. doi: [10.1007/978-3-030-59006-2\\_6/COVER](https://doi.org/10.1007/978-3-030-59006-2_6/COVER).
- [54] P. Soille, "Morphological image analysis," *Morphol. Image Anal.*, 2004. doi: [10.1007/978-3-662-05088-0](https://doi.org/10.1007/978-3-662-05088-0).
- [55] P. Shah, S. N. Merchant, and U. B. Desai, "An efficient adaptive fusion scheme for multifocus images in wavelet domain using statistical properties of neighborhood," in *14th Int. Conf. Inf. Fusion*, Chicago, IL, USA, Jul. 2011, pp. 1–7.
- [56] S. J. Devlin, R. Gnanadesikan, and J. R. Kettenring, "Robust estimation and outlier detection with correlation coefficients," *Biometrika*, vol. 62, pp. 531–545, Dec. 1975. doi: [10.1093/BIOMET/62.3.531](https://doi.org/10.1093/BIOMET/62.3.531).
- [57] G. Zhang, R. Nie, J. Cao, L. Chen, and Y. Zhu, "FDGNet: A pair feature difference guided network for multimodal medical image fusion," *Biomed. Signal Process. Control*, vol. 81, no. 3, pp. 104545, Mar. 2023. doi: [10.1016/J.BSPC.2022.104545](https://doi.org/10.1016/J.BSPC.2022.104545).
- [58] A. Sabeeh Yousif, Z. Omar, and U. Ullah Sheikh, "An improved approach for medical image fusion using sparse representation and Siamese convolutional neural network," *Biomed. Signal Process. Control*, vol. 72, no. 3, pp. 103357, Feb. 2022. doi: [10.1016/J.BSPC.2021.103357](https://doi.org/10.1016/J.BSPC.2021.103357).
- [59] C. S. Xydeas and P. Vladimir, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, Feb. 2000. doi: [10.1049/el:20000267](https://doi.org/10.1049/el:20000267).
- [60] E. D. Vidoni, "The whole brain atlas," *J. Neurol. Phys. Ther.*, vol. 36, no. 2, pp. 108, 2012. doi: [10.1097/NPT.0b013e3182563795](https://doi.org/10.1097/NPT.0b013e3182563795).