Mathematical Biosciences and Engineering

MBE, 20(8): 15374–15406.
DOI: 10.3934/mbe.2023687
Received: 14 May 2023
Revised: 14 June 2023
Accepted: 27 June 2023
Published: 21 July 2023

http://www.aimspress.com/journal/MBE

Research article

A novel medical image fusion method based on multi-scale shearing rolling weighted guided image filter

Fang Zhu[1] and Wei Liu[2,*]

[1] Department of Mathematics, Ministry of General Education, Anhui Xinhua University, Hefei 230088, China
[2] College of Mathematics and Computer Science, Tongling University, Tongling 244061, China

* **Correspondence:** Email: lw_feixi@163.com; Tel: 13956051314.

Abstract: Medical image fusion is a crucial technology for biomedical diagnoses. However, current fusion methods struggle to balance algorithm design, visual effects, and computational efficiency. To address these challenges, we introduce a novel medical image fusion method based on the multi-scale shearing rolling weighted guided image filter (MSRWGIF). Inspired by the rolling guided filter, we construct the rolling weighted guided image filter (RWGIF) based on the weighted guided image filter. This filter offers progressive smoothing filtering of the image, generating smooth and detailed images. Then, we construct a novel image decomposition tool, MSRWGIF, by replacing non-subsampled shearlet transform's non-sampling pyramid filter with RWGIF to extract richer detailed information. In the first step of our method, we decompose the original images under MSRWGIF to obtain low-frequency subbands (LFS) and high-frequency subbands (HFS). Since LFS contain a large amount of energy-based information, we propose an improved local energy maximum (ILGM) fusion strategy. Meanwhile, HFS employ a fast and efficient parametric adaptive pulse coupled-neural network (AP-PCNN) model to combine more detailed information. Finally, the inverse MSRWGIF is utilized to generate the final fused image from fused LFS and HFS. To test the proposed method, we select multiple medical image sets for experimental simulation and confirm its advantages by combining seven high-quality representative metrics. The simplicity and efficiency of the method are compared with 11 classical fusion methods, illustrating significant improvements in the subjective and objective performance, especially for color medical image fusion.

Keywords: rolling weighted guided image filter; multi-scale shearing rolling weighted guided image filter; improved local energy maximum; parameter adaptive PCNN model

1. Introduction

Medical imaging mechanisms can capture images of tissues and organs in different parts of the human body using different sensors. For instance, computed tomography (CT) imaging can capture dense bone and implants but fails to clearly depict soft tissue. On the other hand, magnetic resonance imaging (MRI) yields high-resolution imagery featuring detailed anatomical information and clearly depicting soft tissues but has a limited capability towards capturing bone when compared to CT. Additionally, positron emission tomography (PET) imaging can reflect the biological activity of molecules and cells, while single photon emission computed tomography (SPECT) imaging helps capture the metabolic activity of tissues or organs. However, the low resolution of PET and SPECT imaging makes it difficult to provide precise indications in medical diagnoses. Therefore, to accurately describe a lesion, physicians often have to combine and analyze multiple medical images, which can be inconvenient and inefficient in clinical applications. Fortunately, medical images of different modalities share some degree of information complementarity such as CT/MRI, MRI/PET, MRI/SPECT, etc. Using the complementary information from these images to fuse different medical images can provide a more comprehensive, reliable, and accurate description of a lesion. This technique is useful for biomedical research [1–3] and clinical diagnoses [4], helping physicians to locate lesion severity, provide surgical guidance, and plan radiotherapy treatments. Additionally, it can aid in making future healthcare predictions and improving patient care. Moreover, medical image fusion will play a role in other areas, such as ensuring the confidentiality of electronic medical records on the cloud and constructing dummy query sequences to protect location privacy and query privacy in location-based services.

Over the past few decades, there has been a surge in the number of medical image fusion methods, with multi-scale transformation-based fusion methods being particularly prominent. Typically, the core of these methods includes multi-scale decomposition (MSD) and the settings for fusion rules. Researchers have developed several classical MSD tools such as wavelet transform [5], discrete cosine transform [6], curvelet transform (CVT) [7], dual-complex wavelet transform (DTCWT) [8], non-subsampled contourlet transform (NSCT) [9], non-subsampled shearlet transform (NSST) [10,11], etc. Among them, NSCT and NSST-based fusion methods demonstrate significant advantages in image representation. Lately, the application of filtering theory in image fusion has gained interest from scholars. Several researchers use filter decomposition instead of traditional MSD [12–16] to achieve further improvements in the fusion effect. However, classical filters, such as bilateral filter (BF) [12], guided filter [13], and Gaussian filter [12], often result in a loss of structural information, leading to overly smooth images. To address these disadvantages, several improved filters have been proposed, such as the guided bilateral filter (GBF) [14], rolling guided filter (RGF) [15], local Laplacian filtering (LLF) [16], co-occurrence filter [17], weighted guided image filter (WGIF) [18], and side window guided filtering [19]. These improvements have resulted in further enhancements in fusion performance. Recently, scholars have started correlating traditional MSD and filter decomposition to construct novel fusion methods. For instance, Diwakar et al. [20] used the co-occurrence filter and local extrema in the low frequency part of NSST decomposition results. Liu and Wang [21] constructed a novel multi-scale, multi-directional image decomposition tool based on NSST and the non-local guided filter.

Considering the different information characteristics present in the low-frequency subbands (LFS; base layers) and high-frequency subbands (HFS; detail layers), it's crucial to establish cost-effective fusion strategies based on MSD. However, although simple and easy to implement, traditional

strategies, such as the "average" strategy for LFS and the "maximum" strategy for HFS, may be insufficient in terms of extracting details and handling boundaries. To overcome these limitations, several unique strategies have proposed for LFS and HFS based on extensive experimental simulations. Here, we list a few of the latest methods. Yin et al. [10] investigated the information features in both LFS and HFS based on NSST decomposition. This method considers energy conservation and detail presentation in the LFS while combining HFS using a modified PCNN model, which facilitates an adaptive parameter setting and improves computational efficiency. Ullah et al. [11] proposed a multi-modal medical image fusion method in the NSST domain based on fuzzy sets and an improved Laplacian operator. Du et al. [16] used the local Laplacian filter (LLF) to decompose the original image and designed fusion strategies based on the local energy maximum and interest, preserving a large amount of detailed information while minimizing image distortion. In 2019, Meher et al. [22] presented a comprehensive review of region-based fusion methods. Li et al. [23] employed the joint BF to decompose the original image into a base layer and a detail layer. The base layer utilized the absolute maximum strategy, while the detail layer employed the local gradient energy strategy.

Recently, several fusion methods based on sparse representation (SR) [24–27], convolutional neural networks (CNN) [28–29] and deep learning [30–34] have emerged and shown success in visual contrast and clarity. Among them, Maqsood et al. [25] proposed a multi-modal medical image fusion method in the MSD domain based on SR, while Liu et al. [28] introduced CNNs to medical image fusion by obtaining a map containing pixel activity information from the original image. Zhang et al. [29] developed a general image fusion framework based on CNNs that extracts salient features from the original image using convolutional layers and sets the corresponding fusion strategies. Ma et al. [33] constructed a novel adversarial network to implement end-to-end operations. The development of deep learning theory has contributed to the remarkable progress of image fusion, and the powerful feature extraction and reconstruction ability of neural network gives image fusion a broader application prospect. In 2021, Zhang et al. [34] conducted a comprehensive analysis and review of the latest deep learning-based image fusion methods.

In summary, the state-of-the-art medical image fusion methods listed above offer various improvements in fusion performance compared to traditional methods. However, these fusion strategies may have inherent limitations, such as the loss of helpful information and edge blurring, leading to insufficient extraction of useful information in specific regions and unsatisfactory fusion results. For example, filter-based medical image fusion methods [12–23] have potential drawbacks, such as the number of layers of filter decomposition affecting the extraction of details, with too few or too many decomposition layers impacting the efficiency or fusion performance. Additionally, the extraction of more information often lacks directionality. In particular, for color medical images such as PET and SPECT, the choice of fusion rules for the three channels after RGB conversion directly influences the fusion results, and color distortion easily and often occurs. Similarly, the SR-based fusion approach [24–27] relies on over-complete dictionary image training, which suffers from extremely time-consuming and computationally intensive problems. In the same vein, the implementation of deep learning-based fusion methods [28–34] demands robust computer hardware support, extensive time investment, and a complicated sparse reconstruction process that typically involves a precise setting of multiple parameters, putting significant pressure during experimental operations.

Based on the considerations mentioned above, the aim of this paper is to overcome the existing methods' poor MSD capability, insufficient detail and edge extraction, and to control the operation time cost. We introduce the WGIF to create a novel multiscale filter, called the RWGIF, to achieve

progressive image smoothing and the extraction of rich details. Then, we combine the unique characteristics of NSST and replace the non-subsampled pyramid filter in NSST with RWGIF to construct a novel multi-scale multi-directional decomposition tool (MSRWGIF) to achieve the fine decomposition and extraction of the original image. The next step involves using MSRWGIF to decompose the original image into an LHS and multiple HFS. The LFS use the improved local energy strategy ILGM and the HFS use a fast and efficient AP-PCNN strategy. Finally, the inverse MSRWGIF transform is employed to obtain the fused image.

Therefore, the main focus of this paper is to construct MSD tools and design robust LFS fusion rules. The main innovations in the methodology presented in this paper are as follows:

• In this study, a novel filter, called the rolling weighted guided filter (RWGIF), is constructed. This filter can extract abundant details and boundary information layer by layer, which is a great improvement over the existing filters.

• By employing RWGIF in place of the NSPF in the NSST, we have introduced a unique combination of RWGIF and shearing filter bank (SFB), culminating in the development of a multi-scale and multi-directional decomposition tool, or MSRWGIF for short. The MSRWGIF-based method is designed to deliver varying degrees of smoothing on the original image, allowing for the extraction of detail layers across different scales and orientations. Our experimental results demonstrate that MSRWGIF surpasses other decomposition tools in terms of detail extraction, color fidelity, sharpening, and boundary information enhancement, making it an exceptionally effective tool for medical image fusion.

• In terms of LFS, we have made significant improvements to the traditional local energy criterion, obtaining richer and more illuminating energy information. In addition, we devised a novel fusion strategy, called ILGM, which takes the influence of neighboring distances on energy into account, rather than simply summing up the neighboring energy values. This approach enables a more comprehensive consideration of the impact of neighboring distances on the energy output and further optimizes the performance of the fusion strategy, making it a valuable contribution to the field of medical image fusion.

• Regarding the large number of HFS decomposed by MSRWGIF, it's essential to have a fast and efficient fusion strategy for processing. For this reason, we employ the AP-PCNN to realize the fusion of HFS in this study. To test the efficacy of our proposed method, extensive experiments were carried out on over 100 pairs of medical images, encompassing three different types of images (CT/MRT, MRT/PET, MRT/SPECT). We conducted a subjective and objective comparison of the proposed technique against a total of 11 fusion methods consisting of classical and recently proposed methods. Our experimental results confirm the numerous advantages of the proposed method, fully demonstrating its effectiveness over its competitors.

The remainder of this paper is organized as follows. Section 2 offers an overview of the principles of WGIF. Then, we present the construction of MSRWGIF, followed by a detailed discussion of its implementation framework and principles. In Section 3, we apply MSRWGIF to medical image fusion and describe the corresponding LFS and HFS fusion strategies. Furthermore, in Section 4, we present the results of various subjective and objective experimental validation. Finally, in Section 5, we conclude the paper and draw attention to potential directions for future research.

2. RWGIF

The RGF described in literature [15] is a joint BF and Gaussian filter-based approach that aims

to promote edge preservation and enhance the detail extraction of the original image. Thanks to the RGF, artifacts in the fusion results can be effectively avoided. In this section, we aim to construct a more efficient edge-preserving filter decomposition that relies on WGIF, allowing us to achieve gradual smoothing and detail extraction of the original image. Then, we will leverage the directional flexibility of NSST and combine it with our new decomposition tool to construct a multi-scale and multi-directional decomposition tool, referred to as the MSRWGIF method. Below, we provide a brief introduction to MSRWGIF.

2.1. WGIF

Local filtering techniques that focus on edge-preserving smoothing are often subject to certain halo artifacts, such as those caused by BF, which can result in unwanted contours around edges. While the guided image filter (GIF) can help overcome this issue, it poses a new challenge since the regularization parameter in GIF is often fixed, potentially leading to excessive edge smoothing, and thus, halo artifacts. A solution to this problem was proposed by Li et al. [18], who introduced edge-aware weights to GIF and developed the WGIF. This approach combined the strengths of both global and local smoothing techniques and enabled more effective edge-preserving filtering. The primary aim of edge-preserving smoothing is to decompose an image $X(x,y)$ into the following form:

$$X(x,y) = W(x,y) + e(x,y), \tag{1}$$

where $W(x,y)$ is the reconstructed image, also known as the base layer, and $e(x,y)$ is the texture detail image, also called the detail layer. In GIF, assume that $W(x,y)$ is a linear function on $G(x,y)$ within window $\Omega_{\xi}(x',y')$, denoted as follows:

$$W(x,y) = a_{(x',y')}G(x,y) + b_{(x',y')},\ \forall (x,y) \in \Omega_{\xi}(x',y'), \tag{2}$$

where $\Omega_{\xi}(x',y')$ is a square window with pixel (x',y') centered at radius ξ, $a_{(x',y')}$, $b_{(x',y')}$ are two constants obtained by the least squares method, and $G(x,y)$ is a guidance image. The key to WGIF is embodied by designing and minimizing a cost function $E(a_{(x',y')},b_{(x',y')})$, defined as follows:

$$E(a_{(x',y')},b_{(x',y')}) = \sum_{(x,y)\in\Omega_{\xi}(x',y')} [(a_{(x',y')}G(x,y) + b_{(x',y')} - X(x,y))^2 + \frac{\lambda}{\Gamma_G(x',y')} a_{(x',y')}^2], \tag{3}$$

where λ is a regularization parameter and $\Gamma_G(x',y')$ is an edge-aware weighting, defined as follows：

$$\Gamma_G(x',y') = \frac{1}{N}\sum_{(x,y)=1}^{N} \frac{\sigma_{G,1}^2(x',y') + \varepsilon}{\sigma_{G,1}^2(x,y) + \varepsilon}, \tag{4}$$

where $\sigma_{G,1}^2(x',y')$ is the variance in the local region $\Omega_1(x',y')$ (3×3 window), ε is a constant, N is the number of pixels of an image and the complexity of $\Gamma_G(x',y')$ is $O(N)$. The final $W(x,y)$ is calculated in the following form：

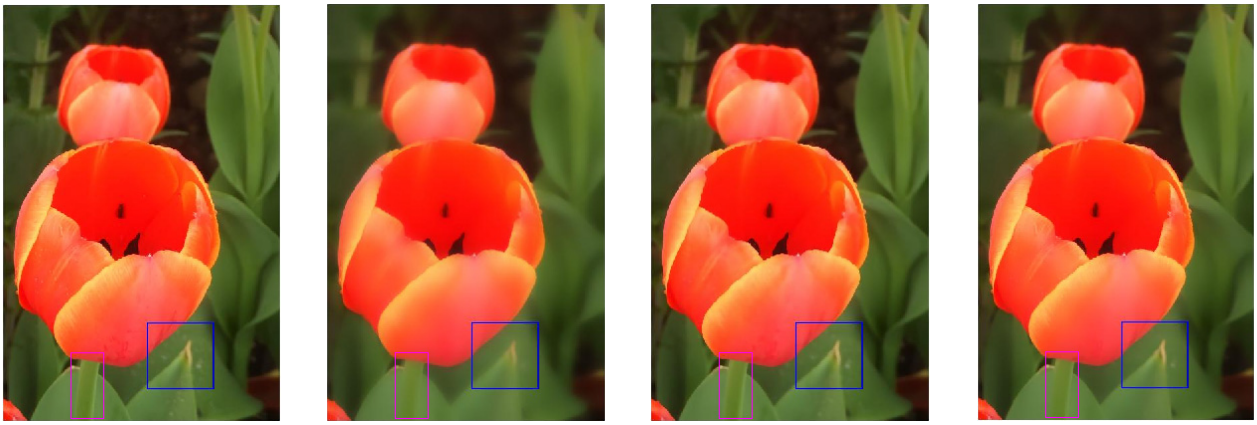
$$W(x,y) = \overline{a}_{(x,y)}G(x,y) + \overline{b}_{(x,y)}, \tag{5}$$

where $\overline{a}_{(x,y)}$, $\overline{b}_{(x,y)}$ are the average values within the local window centered at $a_{(x',y')}$, $b_{(x',y')}$. For a comprehensive theoretical explanation of the WGIF, please refer to the original reference [18].

Building upon the brief description provided above, the results of WGIF smoothing of the image in this paper can be abbreviated as follows:

$$W_B(x,y) = WGIF(X(x,y), G(x,y), \xi, \lambda). \tag{6}$$

To visually illustrate the smoothing effect of WGIF, we have included a comparison between WGIF and two classic filters, namely the guided filter (GF) and the GBF, as shown in Figure 1.



(a) original image (b) GF (c) GBF (d) WGIF

Figure 1. Smoothed results with different filters.

Figure 1 presents the smoothing filtering results for GF, GBF, and WGIF. By closely examining the blue and pink boxes, we observe that the GF result appears excessively smooth, with the edges of the leaves significantly blurred, thus resulting in a loss of important details. The GBF result is very similar to the original image, and consequently, the smoothing effect is poor, leading to relatively weak detail extraction. In contrast, WGIF achieves a balance between the smoothing and preserving edge detail information, resulting in a more desirable smoothing effect compared to the result obtained from GF and GBF.

2.2. RWGIF

To extract useful information for fusion from the medical images of different modalities, efficient image decomposition tools are required. A popular approach in the fusion field is the multi-scale filtering decomposition method, which involves obtaining images of varying degrees of smoothing through filtering, extracting detailed images from the smoothed images, and applying appropriate fusion strategies to process the smoothed and detailed images for the final fusion. While edge-preserving techniques, such as RGF [15], have the ability to preserve edges and contour structures while smoothing images, many existing filters struggle to balance smoothing with detail extraction. To address this issue, we introduce a novel filter decomposition tool, known as the RWGIF, based on the principles of RGF. The RWGIF involves two steps, as illustrated in Figure 2.

• Step 1: Obtain the initial guided image by the Gaussian filter.

The Gaussian filtering of the original image $X(x,y)$ at pixel (x,y) is as follows:

$$G(x,y)=\frac{1}{U}\sum_{(x',y')\in S} exp(-\frac{\left\|(x,y)-(x',y')\right\|^2}{2\sigma_s^2})X(x',y'),$$
$$U=\sum_{(x',y')\in S} exp(-\frac{\left\|(x,y)-(x',y')\right\|^2}{2\sigma_s^2}), \tag{7}$$

where U is used for normalization, S is the set of neighboring pixels of a pixel (x,y), (x',y') is the adjacent pixel of the pixel (x,y), σ_s is Gaussian kernel (minimum standard deviation), and $G(x,y)$ is the result of Gaussian filtering. In this paper, we abbreviate Gaussian filtering as follows:

$$G(x,y)=Gaussian\,(X(x,y),\ \sigma_s). \tag{8}$$

• Step 2: Iterative processing using WGIF and Gaussian filter.

To obtain the filtered and restored image, we subject the original image $X(x,y)$ to WGIF processing under the guidance of the image $G(x,y)$. In each iteration, the guided image is updated as the Gaussian filter result of the previous output image, following the formula:

$$\begin{aligned} G^i(x,y)&=Gaussian\,(W_B^{i\text{-}1}(x,y),\sigma_s^{i\text{-}1}),\ (i=1,2\cdots L),\\ \sigma_s^i&=2\sigma_s^{i-1},\\ W_B^i(x,y)&=WGIF\,(X(x,y),G^i(x,y),\xi,\lambda), \end{aligned} \tag{9}$$

where i is the number of iterations, $G^i(x,y)$ is the result of the $i-th$ Gaussian filter processing, and $W_B^i(x,y)$ is the result of the $i-th$ RWGIF processing. The empirical initialization takes the following value: $W_B^0(x,y)=X(x,y),\ \sigma_s^0=2,\ \xi=8,\ \lambda=1/64$.

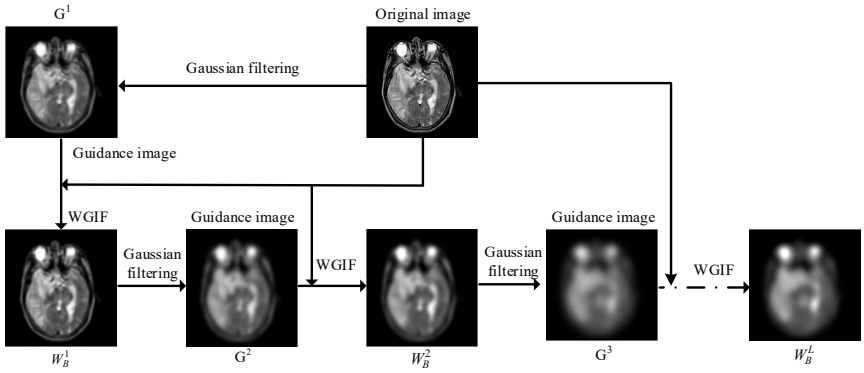


Figure 2. RWGIF.

Figure 2 depicts the detailed iterative filtering process of RWGIF. As shown, the original image is progressively smoothed, leading to an increase blurriness, while the filtering outputs retain a significant amount of original image contour information. Using Eq (9), we can extract the corresponding detail images:

$$W_D^i=W_B^{i-1}-W_B^i,\ (i=1,2,\cdots L,\ W_B^0=X) \tag{10}$$

To showcase the efficacy of RWGIF in terms of smoothing and detail extraction, we conduct a visual comparison with GBF [13] and RGF [15]. The original image underwent processing using both filters, iterated four times to generate four smoothed images and four detailed images. The results are displayed in Figure 3, which presents three smooth images and four detailed images.

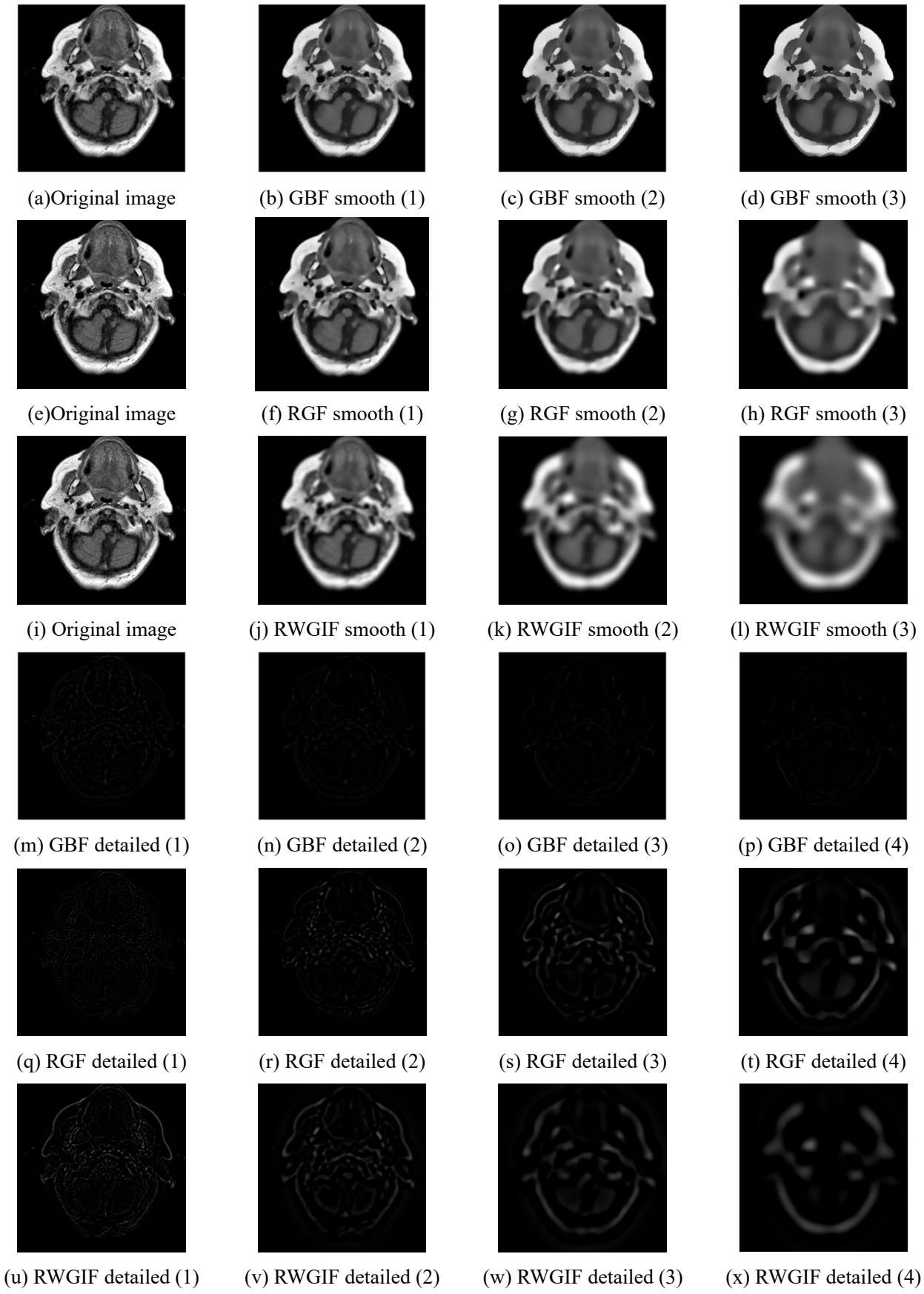


Figure 3. Comparison of smoothed and detail results of GBF, RGF and RWGIF.

Figure 3 showcases the smoothing and detailed extraction performance of GBF, RGF and RWGIF. The first two rows present the results of GBF and RGF smoothing for three iterations, while the third row exhibits the results of RWGIF smoothing after the same number of iterations. Upon closer examination of the comparison, we notice that RWGIF achieves more smoothing and differences compared to GBF and RGF. The last three rows display the detailed images obtained via GBF, RGF and RWGIF, respectively. A careful comparison of these three rows reveals that RWGIF extracts significantly more detailed information compared to GBF and RGF. Thus, a comprehensive evaluation of GBF, RGF and RWGIF indicates that RWGIF is a simple, efficient, and easy-to-implement design that offers unique advantages regarding smoothing and detail extraction. This tool provides a solid foundation for the subsequent MSD of images.

2.3. MSRWGIF

Traditional image decomposition tools have limited effectiveness in extracting the geometric features required for subsequent image fusion due to the richness of unique information present in natural images. Therefore, it is essential to develop a decomposition framework capable of effectively representing the sparse features of an image. The RWGIF filtering proposed in Section 2.2 possesses unique smoothing and detail extraction capabilities and is a promising candidate for application. However, the number of details extracted through filter decomposition is often limited and not directional. In recent years, researchers have constructed decomposition tools for extracting the orientation information of images, including NSCT, NSST, and multi-scale direction BFs [13–21]. Among them, NSST is one of the most widely used conversion tools, providing precise directional detail information with high implementation efficiency. In 2021, Liu and Wang [21] combined the directional flexibility of NSST with the non-local guided filter, resulting in a multi-scale and multi-directional decomposition tool that is effectively used in multi-focus image fusion. Their fusion method has been extensively tested and exhibits high visual effects with a noise-immune and efficient process. Nonetheless, similar multi-scale and multi-directional decomposition tools have not yet been applied in medical image fusion. In this paper, we introduce a novel multi-scale and multi-directional decomposition tool by comprehensively considering the unique advantages of RWGIF and NSST. While NSPF (non-subsampled pyramid filter) and SFB in NSST provide high directional expression capability and flexible operations, isotropic NSPF only achieves MSD of the image; it is not capable of retaining the original image's edge information, leading to edge blurring in the fused image. To address this issue, we propose a novel image representation technique, MSRWGIF, by replacing NSPF with multi-scale RWGIF.

A detailed theoretical explanation of NSST can be found in the literature [10,11]. Figure 4 illustrates the two-level MSRWGIF decomposition framework diagram. The original image undergoes RWGIF decomposition to produce a smooth image and a detailed image. The smooth image undergoes processing with RWGIF while the detail image is processed with SFB, resulting in an LHS and several HFS. In this paper, the steps involved in MSRWGIF decomposition and reconstruction are as follows:

• The original image $X(x,y)$ is decomposed by RWGIF into a smooth image W_B^L and, in turn, L detailed images $W_D^i\ (i=1,2,\cdots,L)$.

• The window function of the shearlet filter is constructed based on the Meyer wavelet on each detail image $W_D^i\ (i=1,2,\cdots,L)$ to obtain the subband coefficients in different directions. Then, the inverse fast Fourier transform (two-dimensional) is performed on the subband coefficients to obtain

the HFS of MSRWGIF in different directions. MSRWGIF is defined as follows:

$$S = \{S^L, \{S^{l,d} \big| 1 \le l \le L, d = 1,2,\cdots,2^{K_l}, K_l \in Z^+\}\}, \tag{11}$$

where $S^L = W_B^L$, $S^{l,d} = \{W_D^l * SFB_d \big| d = 1,2,\cdots,2^{K_l}, K_l \in Z^+\}$, SFB is the SFB, S^L is the LFS of the original image after L MSRWGIF decomposition, and $S^{l,d}$ are HFS of MSRWGIF decomposition in l scale d direction.

• The reconstruction formula of MSRWGIF is as follows:

$$F = S^L + \sum_{l=1}^{L} S_D^l, \tag{12}$$

where $S_D^l = \{S^{l,d} * SFB_d^* \big| d = 1,2,\cdots,2^{K_l}, K_l \in Z^+\}$ and SFB_d^* is the synthesis of the SFB.

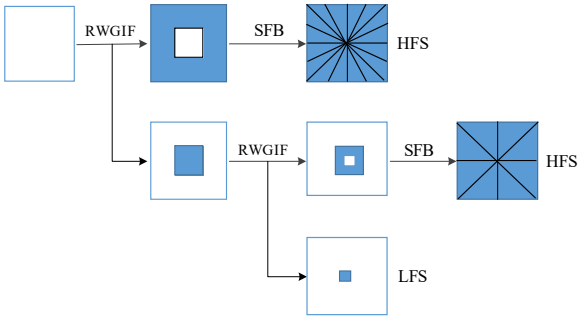


Figure 4. MSRWGIF's decomposition framework.

The flowchart of MSRWGIF decomposition is presented in Figure 5. From the above introduction of MSRWGIF and Figure 5, it is evident that MSRWGIF possesses high directional sensitivity, displacement invariance, and an excellent detail extraction ability. Further subjective and objective experimental comparisons are presented in Section 4.

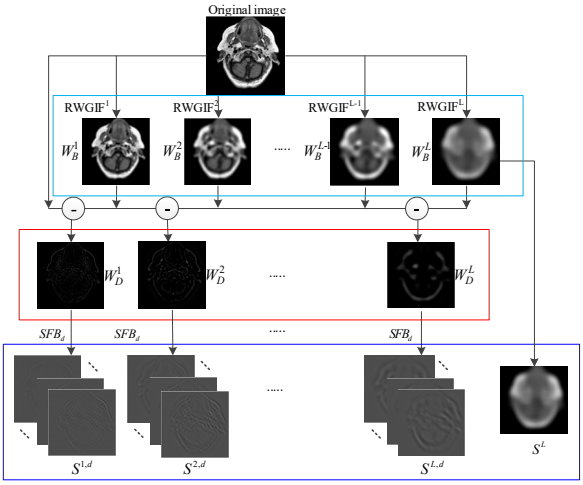


Figure 5. Flow chart of MSRWGIF decomposition.

3. Mdeical image fusion using MSRWGIF

3.1. The basic process of MSRWGIF-based image fusion

The principle and advantages of MSRWGIF have been introduced in detail in Section 2, and we propose a novel medical image fusion method based on MSRWGIF. For ease of study, we denote the two original images as X_A, X_B and the fused image as F.

• The original medical images X_A, X_B are decomposed using MSRWGIF to obtain the LFS and HFS:

$$\begin{aligned}&\{S_A^L, S_A^{l,d} \,|\, 1 \le l \le L, d = 1,2,\cdots,2^{K_l},\ K_l \in Z^+\}\\&\{S_B^L, S_B^{l,d} \,|\, 1 \le l \le L, d = 1,2,\cdots,2^{K_l},\ K_l \in Z^+\},\end{aligned} \tag{13}$$

where the number of MSRWGIF decomposition layers is L, S_A^L, S_B^L are the LFS obtained after decomposition of the two original images, and $S_A^{l,d}$, $S_B^{l,d}$ are the HFS at different scales with different orientations.

• To effectively combine the information features present in LFS and HFS, different fusion strategies are required for each. Specifically, we will employ the ILGM strategy for LFS and the AP-PCNN strategy for HFS. The resulting fused LFS and HFS will be denoted as follows:

$$\{F^L, F^{l,d} \,|\, 1 \le l \le L, d = 1,2,\cdots,2^{K_l},\ K_l \in Z^+\} \tag{14}$$

• The fused LFS and HFS use inverse MSRWGIF to obtain the final merged images F.

By combining the information presented in Figure 5 and the explanation provided in Section 2, it becomes apparent that the LFS contain a significant amount of contour and energy information, while the HFS control detail and edge information of varying levels and directions. Thus, it is imperative to devise distinct fusion strategies for handling the LFS and HFS. The fusion process proposed in this paper is illustrated in Figure 6.

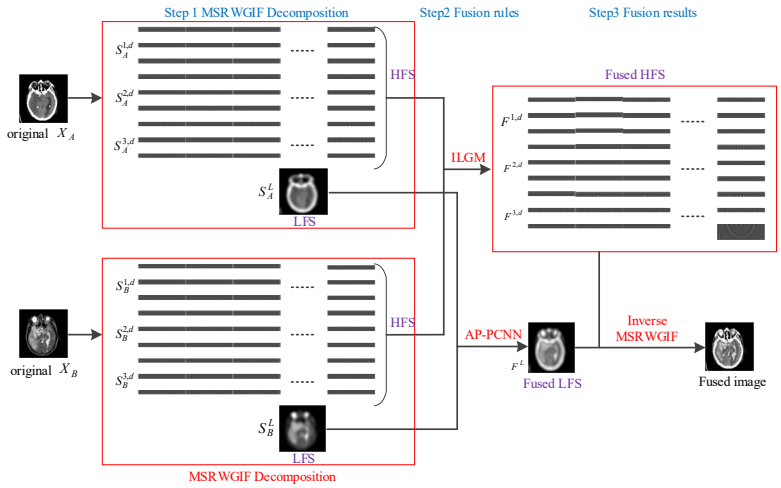


Figure 6. The flow of the proposed method.

3.2. LFS fusion strategy

The majority of traditional LFS or smoothed images employ the "average" strategy, which is a straightforward and easy-to-implement method. However, this method only considers a single pixel and fails to capture the critical features of the image while neglecting the impact of the neighborhood on the central pixel. As a result, this strategy can lead to blurred edges and decreased contrast in the fused image. To overcome this issue, in recent years, several fusion strategies based on energy, gradient, and spatial frequency have been proposed [13,23], which have significantly enhanced the fusion performance. For example, in [13], the local energy at each pixel is calculated, and then the smoothed image is fused based on the local energy maximum strategy. Similarly, the method proposed in [23] employs a more efficient detail fusion strategy based on the local gradient energy. Nonetheless, the energy in these strategies simply adds up the domain pixel values. In contrast, we propose a novel local energy-based strategy (ILGM) to guide LFS fusion, which fully considers the contribution of the spatial distance of the domain pixels to the local energy. The proposed novel local energy is defined as follows:

$$\begin{aligned} GE_t^L(x,y) &= S_t^L(x,y) + S_t^L(x-1,y) + S_t^L(x+1,y) + S_t^L(x,y-1) + S_t^L(x,y+1) \\ &+ \frac{1}{\sqrt{2}} S_t^L(x-1,y-1) + \frac{1}{\sqrt{2}} S_t^L(x-1,y+1) + \frac{1}{\sqrt{2}} S_t^L(x+1,y-1) + \frac{1}{\sqrt{2}} S_t^L(x+1,y+1),\ t \in A,B, \end{aligned} \tag{15}$$

where $GE_t^L(x,y),\ t \in A,B$ are the local energy of $S_A^L,\ S_B^L$ at position (x,y).

To improve the fusion performance of LFS, which typically contains a significant amount of energy and contour information, traditional fusion strategies have been proposed in recent years. However, these strategies only add up the domain pixel values without fully considering their spatial distance. In this paper, we propose a novel local energy-based strategy that considers the contribution of the neighborhood distance to the local energy. The definition of the local energy is provided by Eq (15), where the influence of the neighborhood distance on the local energy is fully considered. The local energy of other windows with varying sizes is similarly defined using the same equation, and in this paper, we select a window size of 3×3. The decision for structural saliency based on Eq (15) is given as follows:

$$M_A^L(x,y) = \begin{cases} 1, & if\ \ GE_A^L(x,y) \geq GE_B^L(x,y) \\ 0, & otherwise \end{cases},\ M_B^L(x,y) = 1 - M_A^L(x,y). \tag{16}$$

Figure 6 demonstrates that LFS acquired through MSRWGIF decomposition contain a considerable amount of contour information. Eq (16) only compares the local energy at a single location, which may result in the loss of crucial information. Therefore, to address this limitation, we propose an improved energy mapping strategy based on local regions as follows:

$$\hat{M}_A^L(x,y) = \begin{cases} 1, & if \sum\limits_{(a,b)\in\Omega} M_A^L(x+a,y+b) > 0.5\times size(\Omega) \\ 0, & otherwise \end{cases},\ \hat{M}_A^L(x,y) = 1 - \hat{M}_B^L(x,y), \tag{17}$$

where Ω is the local region of $T\times T$ and $\sum\limits_{(a,b)\in\Omega} M_A^L(x+a,y+b),\ \sum\limits_{(a,b)\in\Omega} M_B^L(x+a,y+b)$ are the sum of structural saliency mappings in the local region Ω. Thus, the fused LFS is noted as follows:

$$F^L(x,y) = \hat{M}_A^L(x,y) S_A^L(x,y) + \hat{M}_B^L(x,y) S_B^L(x,y). \tag{18}$$

3.3. HFS fusion strategy

Figure 6 showcases that by utilizing MSRWGIF decomposition of the original images, a set of HFS images containing substantial texture boundary information and abundant detail information at various scales and orientations can be generated. Therefore, it is imperative to develop a fusion strategy that can effectively integrate these HFS images while maintaining operational efficiency. However, existing fusion methods such as the local Laplacian energy in [9], sum-modified-Laplacian in [11], and local gradient energy maximum in [16], require the computation of each pixel position in every HFS image, which can result in time-consuming computation and redundancy when handling multiple HFS images. To overcome these restrictions, Yin et al. [10] introduced the AP-PCNN, which provides a fast and efficient approach to extract detailed information from HFS images. Therefore, in this study, we employ the AP-PCNN model to achieve high-quality HFS fusion while maintaining computational efficiency.

PCNN [35] is a biologically inspired neural network model that is based on the cortical model. It establishes a one-to-one correspondence between the input image pixels and the neurons of PCNN. These neurons measure the activity level of neuron coefficients through multiple iterations to achieve the output of neurons, with features such as global coupling and pulse synchronization. However, the traditional PCNN model is often criticized for its slow operation and difficulty in setting parameters. To overcome these issues, Chen et al. [36] proposed a parameter-adaptive simplified PCNN (AP-PCNN) model for image segmentation. The specific model and parameter settings are described as follows:

$$\begin{cases} F_{xy}[n] = S_{xy} \\ L_{xy}[n] = V_L \sum_{kl} W_{xykl} Y_{Kl}[n-1] \\ U_{xy}[n] = e^{-\alpha_f} U_{xy}[n-1] + F_{xy}[n](1+\beta L_{xy}[n]), \\ Y_{xy}[n] = \begin{cases} 1, if\ U_{xy}[n] > E_{xy}[n-1] \\ 0, \quad otherwise \end{cases} \\ E_{xy}[n] = e^{-\alpha_e} E_{xy}[n-1] + V_E Y_{xy}[n] \end{cases} \begin{cases} \alpha_f = \log(1/\sigma(S)) \\ \lambda = \frac{1}{6}((S_{\max}/S')-1) \\ V_E = e^{-\alpha_f} + 1 + 6\lambda \\ \alpha_f = \ln(\frac{V_E / S'}{(1-e^{-3\alpha_f})/(1-e^{-\alpha_f}) + 6\lambda e^{-\alpha_f}}) \end{cases}, \tag{19}$$

where S_{xy} is the input image, $F_{xy}[n]$ and $L_{xy}[n]$ are the feeding input and link input of the input image at position (x, y), n is the number of iterations. $U_{xy}[n]$ is internal activity, $Y_{xy}[n]$ determines the pulse generator status (fired and unfired) of the output module, $E_{xy}[n]$ is the dynamic threshold, V_L is the amplitude of the connection input, $L_{xy}[n]$ is based on V_L and synaptic weights and W_{xykl} associated with the ignition state of the previous eight domain neurons. The parameters $e^{-\alpha_f}$ and β are the attenuation coefficient and connection strength, respectively, which together affect the internal activity $U_{xy}[n]$. The parameters α_e and V_E are the exponential decay coefficient and amplitude of $E_{xy}[n]$, respectively. The initial value of the AP-PCNN model is set as follows: $Y_{xy}[0] = 0$, $U_{xy}[0] = 0$, $E_{xy}[0] = 0$. Since the firing condition $U_{xy}[1] = S_{xy} > 0$, non-zero neurons will be fired in the first iteration. S' and $S_{\max}$ are the Qtsu threshold and maximum intensity after normalizing the image to $[0, 1]$, and $\sigma(S)$ is the standard deviation of the input image. The specific AP-PCNN theoretical explanation and framework can be found in [9].

In this paper, we employ the AP-PCNN model to extract detailed information from each HFS of the original image obtained by MSRWGIF decomposition. The absolute value of the pixel in each HFS serves as the input neuron for the AP-PCNN model ($F_{xy}[n] = \left|S_t^{l,d}\right|,\ t \in \{A, B\}$). Then, the activity level of

the HFS is determined by the total number of firings during the iteration process. To determine the cumulative number of firings at each iteration, we use Eq (19) in conjunction with the following:

$$T_{xy}[n] = T_{xy}[n-1] + Y_{xy}[n], \tag{20}$$

where the total number of firings after n iterations is $T_{xy}[n]$. From Eqs (19) and (20), the number of firings of $S_A^{l,d}$, $S_B^{l,d}$ can be calculated as $T_{A,xy}^{l,d}[n]$, $T_{B,xy}^{l,d}[n]$, and the final fused HFS are as follows:

$$F^{l,d}(x,y) = \begin{cases} S_A^{l,d}(x,y), & if\ T_{A,xy}^{l,d} \geq T_{B,xy}^{l,d} \\ S_B^{l,d}(x,y), & otherwise \end{cases}. \tag{21}$$

The LFS and HFS fusion are completed in Sections 3.2 and 3.3 using different fusion strategies, respectively. The final fused image is obtained by the inverse MSRWGIF:

$$F = F^L + \sum_{l,d} F^{l,d} * SFB_d^*,\ (l = 1,2,\cdots L,\ d = 1,2,\cdots 2^{K_l},\ K_l \in Z^+), \tag{22}$$

where SFB_d^* are the synthesis of the SFB. The following is the main algorithm of the proposed method.

Algorithm 1 Proposed Medical Image Fusion Method

Input: the original images: X_A, X_B.

Parameters: the number of MSRWGIF decomposition levels: L, the number of directions for each level: $K(l), l \in [1, L]$, local region window size T in LFS fusion.

Part 1: MSRWGIF decomposition

01: For original image X_A, X_B

02: Initialize the MSRWGIF parameters: $\sigma_s^0 = 2,\ \xi = 8,\ \lambda = 1/64, L = 3$, the number of directions at each layer is [16,16,16];

03: Performing MSRWGIF decomposition on X_A, X_B yields $\{S_A^L, S_A^{l,d}\}$, $\{S_B^L, S_B^{l,d}\}$, $l \in [1, L],\ k \in [1, K(l)]$;

04: End

Part 2: LFS fusion

05: For each LFS: S_A^L, S_B^L

06: Calculate the local energy of S_A^L, S_B^L using Eq (15) (GE_A^L, GE_B^L);

07: End

08: According to Eq (16)–(18) to obtain the fused low-frequency subband F^L;

Part 3: HFS fusion

09: For each layer $l = 1 : L$

10: For each direction $k = 1 : K(l)$

11: For each high-frequency subband $S_A^{l,d}, S_B^{l,d}$

12: Initialize the PA-PCNN parameters: $Y_{ij}[0] = 0, U_{ij}[0] = 0, E_{ij}[0] = 0, T_{ij}[0] = 0$ and $F_{ij}[n] = \left|S_A^{l,d}\right| / \left|S_B^{l,d}\right|, n \in [1, N]$;

13: Use Eq (19) to determine the parameters of the AP-PCNN;

14: Eq (20) is used to determine the total number of firings after n iterations;

15: Generate the fused high-frequency subbands $F^{l,d}$ according to Eq (21);

16: End

17: End

18: End

Part 4: MSRWGIF reconstruction

19: Perforing inverse MSRWGIF on F^L and $F^{l,d}$ to achieve F;

Output: the final fused image F.

4. Experiments

4.1. Experimental settings

4.1.1. The set of tested images

To demonstrate the superior performance of our proposed method, we conducted extensive experimental simulations and analyses on over 100 pairs of multi-modal medical images. These original images are primarily structural maps of the human brain acquired through multiple imaging mechanisms, including MR-T1/MR-T2, CT/MRI, MRI/PET, and MRI/SPECT, all with a resolution of 256×256. Each pair has been registered. The test image pairs can be downloaded from the Harvard Medical School database website at: http://www.med.harvard.edu/aanlib/home.html. We randomly selected 12 pairs of medical images, covering four types (as depicted in Figure 7) to display. Due to space limitation, we randomly show four types of experimental results to illustrate the generalization effect.

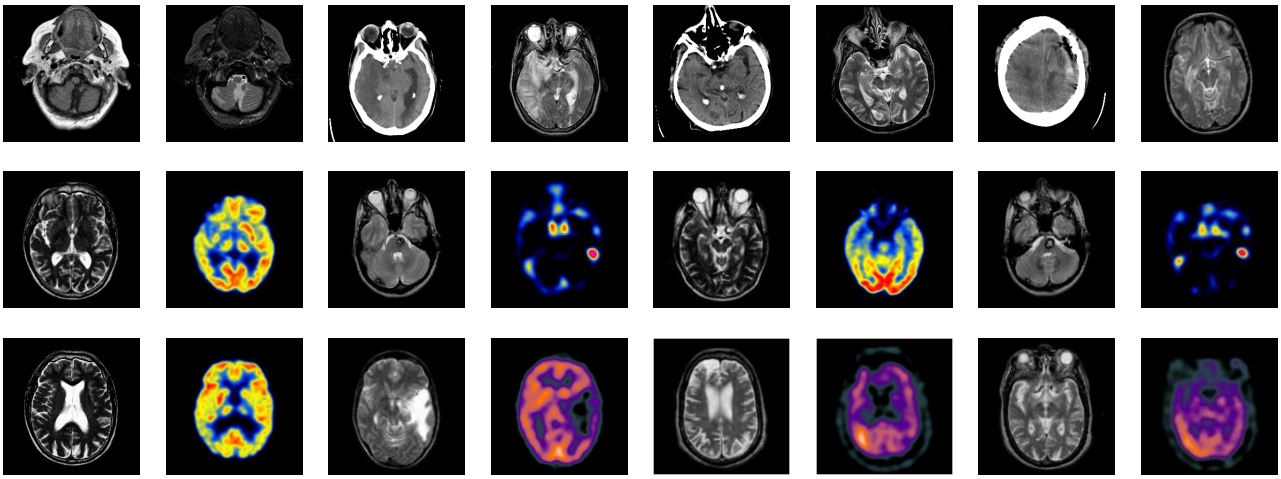


Figure 7. Twelve pairs of multi-modal medical images.

4.1.2. Objective evaluation metrics

Subjective visual analysis alone may not be sufficient to comprehensively analyze and compare fusion results, particularly when dealing with output images that have minor differences. To address this issue, we have chosen six objective metrics to assist with the subjective visual analysis. These metrics include MI [37], $Q_{AB/F}$ based on edge preservation [38], VIF based on visual information fidelity [39], Q_{CB} based on the human visual system [40], NFMI based on feature mutual information [41], and QSR_{SIM} based on spectral residual similarity [42]. Each metric measures different image features from various perspectives. The literature [43] has categorized fusion metrics into four types: information theory-based (MI, $Q_{AB/F}$), human visual perception-based (VIF, Q_{CB}), image feature-based (NFMI), and image structure similarity-based (QSR_{SIM}). Higher values for these six metrics indicate an improved fusion performance. By combining these metrics, the fusion outcomes can be more objectively, reliably, and authoritatively assessed.

4.2. Setting of color medical image fusion

When dealing with color medical image fusion, such as the two types of MRI/PET and MRI/SPECT, directly fusing the three channels in RGB space often leads to unsatisfactory results [44,45]. To address this issue, Kim et al. [46] proposed a color medical image fusion strategy. The strategy involves first converting the input color medical images to the YUV space, fusing the Y channel alone using the proposed method in the article, and then converting the fused Y channel back to the RGB space with the previous U and V channels to obtain the final fusion results. In this paper, we follow the same processing method as [46], and a detailed analysis of the fusion process can be found in their work [35,46].

4.3. Experimental analysis of the proposed method

4.3.1. Parameters setting

After conducting a comprehensive experimental analysis and comparison, we discovered that both the number of layers in the MSRWGIF decomposition and the region size Ω in LFS processing are critical factors in achieving an optimal performance. Further information on the settings of other parameters in the MSRWGIF decomposition, LFS, and HFS can be found in previous studies [10,16,18,35].

1) MSRWGIF decomposition layers L

We conducted a study on the number of decomposition layers for over 100 image pairs, which revealed that using a low number of layers L can result in insufficient extraction of detailed information, while using too many layers may impair the fusion efficiency. Therefore, we set the maximum number of decomposition layers to six. To illustrate the experimental effect, we randomly show the results of the fusion of three pairs of images (CT/MRI, MRI/PET, MRI/SPECT) using the fusion strategy proposed in Section 3, as shown in Figure 8 (T = 15).

Figure 8 displays the fusion results of three types of medical images with different numbers of MSRWGIF decomposition layers. To provide a closer examination, certain regions have been selected and enlarged on the lower right side to highlight the visual effect. It can be observed that the fusion results obtained from a single layer of MSRWGIF decomposition, as depicted in Figure 8 ((c1), (d1), (e1)), lack detailed information, and the color rendering appears dim with poor contrast. In contrast, as the number of decomposition layers increases, the detail and contrast are significantly enhanced, and the blue rendering is gradually reduced, as seen in Figure8(d1)–(d4). However, using an excessive number of decomposition layers, as indicated in Figure8 ((d6), (e4)–(e6)), results in noticeable deterioration and artifacts. From a subjective visual standpoint, we have determined that the optimal number of MSRWGIF decomposition layers is approximately three. To further determine the optimal value of the parameter L, additional analysis of the objective metric for the number of decomposition layers is required. Figure 9 shows the average of the objective metric for the all-test image pairs.

Figure 9 presents a clear trend where MI and QCB metrics gradually decrease as the number of MSRWGIF decomposition layers increases. In contrast, the QAB/F, VIF, NFIM, and QSRSIM metrics show an increasing trend when $L < 3$, but a decreasing or stable trend when $L > 3$. Based on the observations from Figures 8 and 9, we can conclude that the optimal number of decomposition layers is 3.

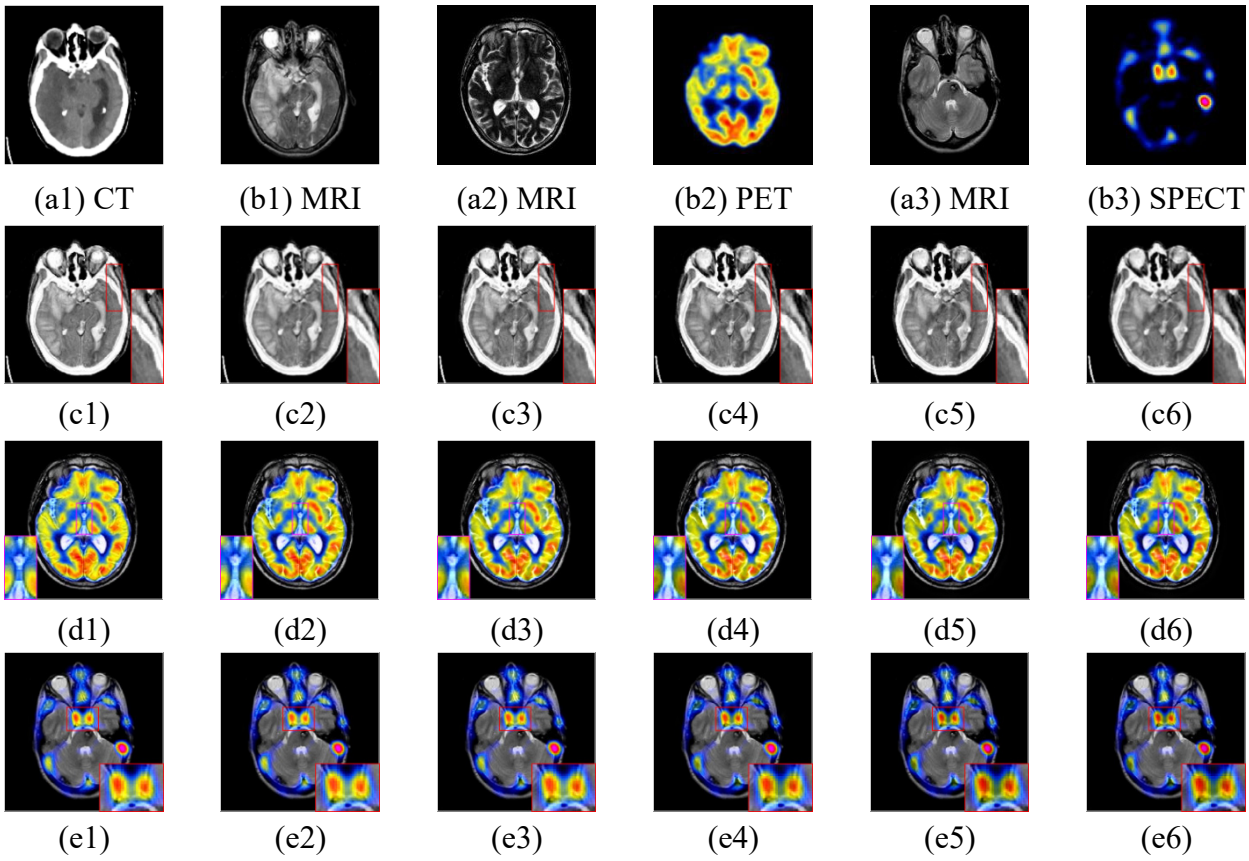


Figure 8. The results of decomposing three pairs of medical images in 6 layers (The first row is the three original image pairs, (c1)–(c6), (d1)–(d6), (e1)–(e6) are the output results of the sequential MSRWGIF decomposition of 6 layers).

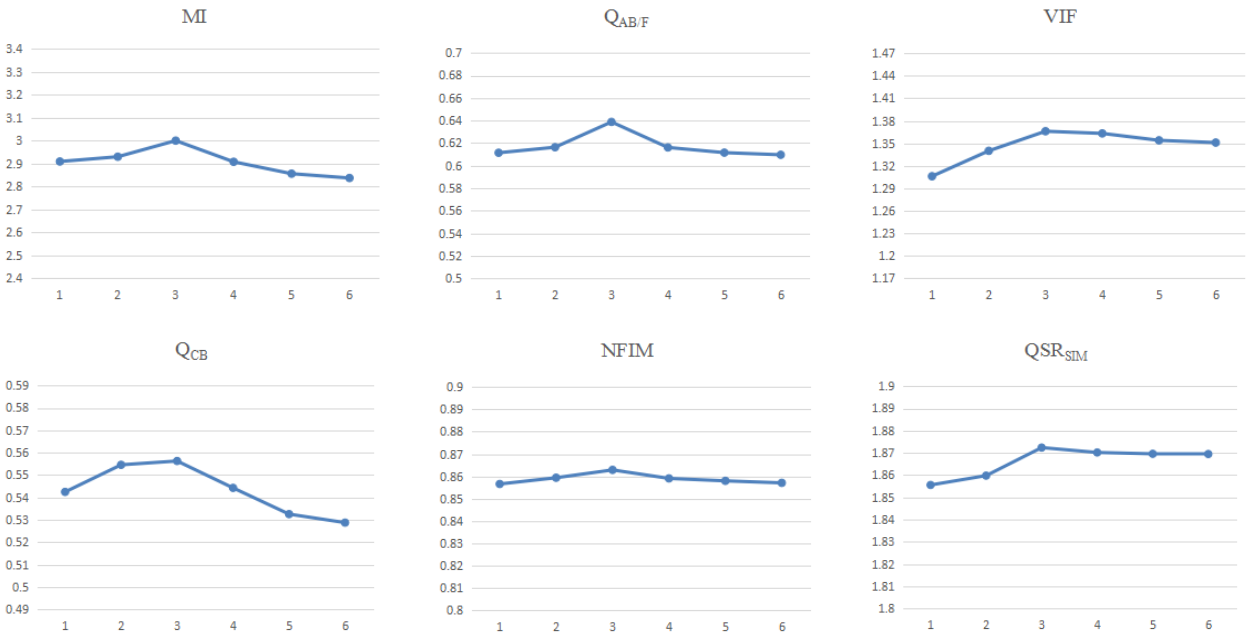


Figure 9. The average metric value of the parameter L.

2) Local region window size T in LFS fusion

After conducting numerous experiments, we discovered that the choice of parameter T in the LFS fusion strategy has a significant impact on the fusion results. Therefore, we performed an experimental analysis on the dataset in Section 4.3.1.1 and set the maximum value of T to 27. However, due to the lack of significant visual contrast, we present the average metric results of the experimental outcomes in Figure 10 (L = 3).

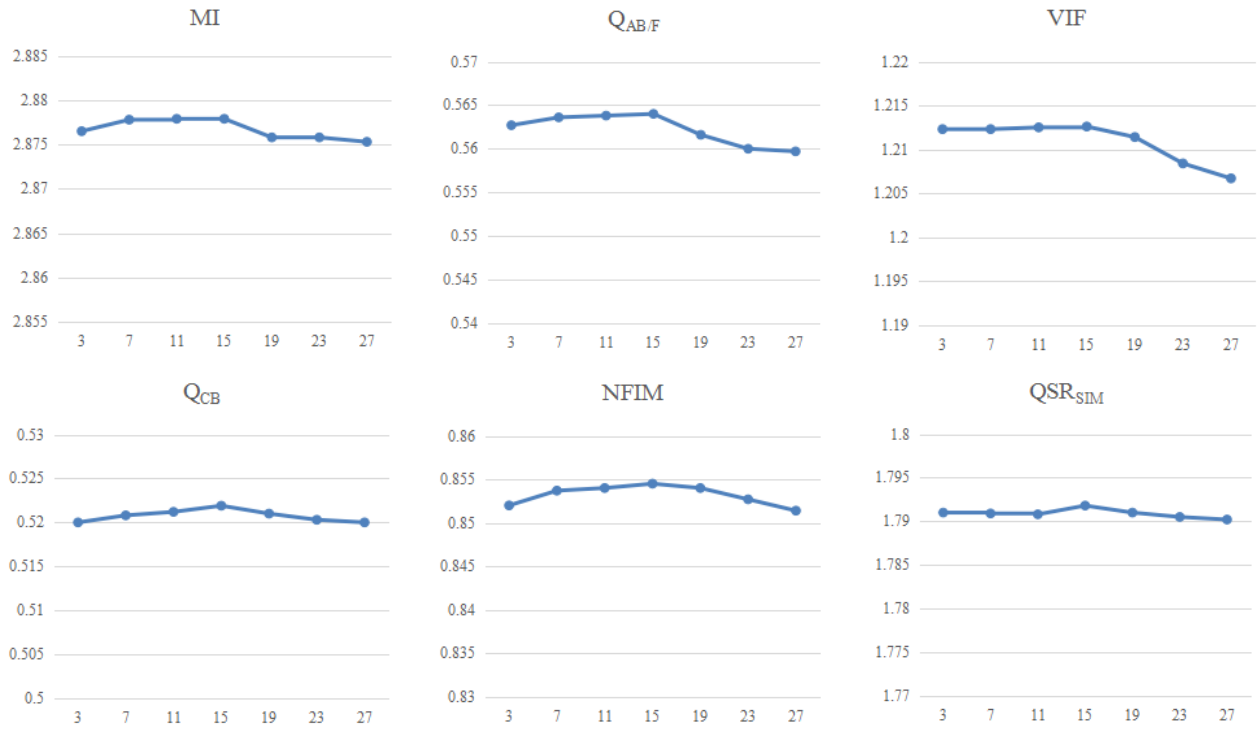


Figure 10. The average metric value of the parameter T.

Figure 10 shows the average metric values for all test images at different T values. The trend plots of the six metrics reveal that when T < 15, the six metrics remain almost unchanged, indicating that the value of T has little impact on the fusion effect. However, when T > 15, the six metrics exhibit a decreasing trend, with $Q_{AB/F}$ and VIF decreasing significantly, indicating a compromise in edge retention and visual effect. Therefore, we have concluded that T=15 is the most appropriate choice for the LFS fusion strategy.

4.3.2. Experimental comparative analysis of MSRWGIF

One of the primary contributions of this paper is the development of a MSD tool, called MSRWGIF, which is discussed in detail in Section 2. In this section, we comprehensively evaluate the practical performance of MSRWGIF by comparing it with six other MSD tools, including DTCWT, CVT, NSCT, NSST, and multi-scale filtering decomposition tools such as LLF and RGF. In addition, the "average" strategy is used for LFS and the "maximum" strategy is used for HFS during the experiments. The subjective effects of the three sets of experiments are randomly displayed, as shown in Figure 11.

Figure 11 displays the fusion results obtained using seven decomposition tools. The fusion results obtained with CVT-based fusion (Figure 11(c1), (d1), (e1)) exhibit the worst performance with blurring and insufficient extraction of details. The fusion results obtained with DTCWT shows improvement compared to CVT but are still slightly inadequate when compared to NSCT and NSST. The fusion results obtained with NSCT and NSST are not directly comparable due to their minor structural differences. The fusion results obtained with the multi-scale filtering decomposition tools LLF and RGF (Figure 11(c5), (d5), (e5), (c6), (d6), (e6)) exhibit obvious shortcomings in detail extraction, resulting in fusion results that are inferior to those obtained with traditional NSCT and NSST. A careful comparison of Figure 11(c5)/(c6), (d5)/(d6), (e5)/(e6) reveals that RGF considers the internal structural features of the image and combines the advantages of Gaussian filter and joint bilateral filtering, which improves detail extraction, contrast, and brightness compared with LLF. Examining each line of fusion results, it is apparent that MSRWGIF-based fusion results outperform other methods regarding sharpness, detail extraction, edge information retention, and color fidelity.

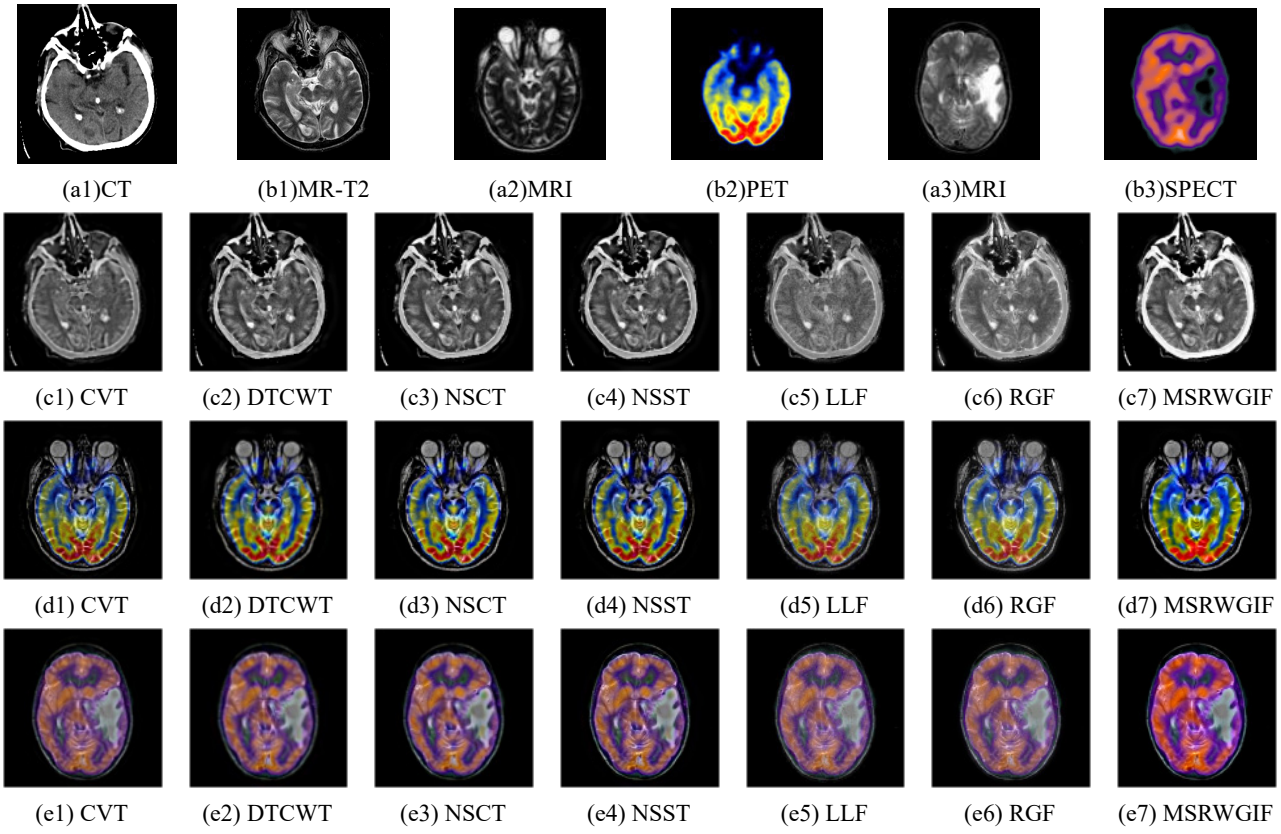


Figure11. Fusion results based on different decomposition tools (The first row shows three different types of medical image pairs, and the next three rows show the fusion results of CVT, DTCWT, NSCT, NSST, LLF, RGF and MSRWGIF, respectively).

Figure 11 displays the fusion results obtained using seven decomposition tools. The fusion results obtained with CVT-based fusion (Figure 11(c1), (d1), (e1)) exhibit the worst performance, with blurring and insufficient extraction of details. The fusion results obtained with DTCWT show an improvement compared to CVT but are still slightly inadequate when compared to NSCT and NSST.

The fusion results obtained with NSCT and NSST are not directly comparable due to their minor structural differences. The fusion results obtained with the multi-scale filtering decomposition tools LLF and RGF (Figure 11(c5), (d5), (e5), (c6), (d6), (e6)) exhibit obvious shortcomings in detail extraction, resulting in fusion results that are inferior to those obtained with traditional NSCT and NSST. A careful comparison of Figure 11(c5)/(c6), (d5)/(d6), (e5)/(e6) reveals that RGF considers the internal structural features of the image and combines the advantages of a Gaussian filter and joint bilateral filtering, which improves detail extraction, contrast, and brightness compared with LLF. Examining each line of fusion results, it is apparent that MSRWGIF-based fusion results outperform other methods regarding sharpness, detail extraction, edge information retention, and color fidelity.

After analyzing the results of the six metrics presented in Table 1, it is apparent that the NSCT and NSST have certain advantages. Although multi-scale filtering decomposition methods LLF and RGF perform well in the MI and QSRSIM metrics, they are notably inferior to traditional decomposition tools in the other four metrics. In contrast, MSRWGIF, proposed in this paper exhibits a superior performance in both subjective and objective metrics.

Table 1. Average metric values based on different decomposition tools.

MRI/PET	CVT	DTCWT	NSCT	NSST	LLF	RGF	MSRWGIF
MI	2.4159	2.3961	2.4372	2.6013	2.6236	2.7235	2.7767
$Q_{AB/F}$	0.6018	0.6104	0.6160	0.6182	0.4246	0.5441	0.6245
VIF	0.9027	1.1243	1.1916	1.1727	0.6420	0.7832	1.1938
Q_{CB}	0.4713	0.4768	0.5109	0.4847	0.5919	0.4710	0.5382
NFIM	0.8435	0.8453	0.8467	0.8476	0.8261	0.8254	0.8516
QSR_{SIM}	1.8215	1.8384	1.8370	1.8313	1.8783	1.8693	1.8817

4.3.3. Confirming the critical role of ILGM

To thoroughly investigate the impact of local region Ω on the fusion results, we fix the other components of the proposed method and solely consider the fusion effect with and without region. We randomly show the experimental results of five pairs of medical images for visual comparison, as shown in Figure 12.

In Figure 12, (a1–e1) and (a2–e2) depict five pairs of medical images, where Figure 12(a3–e3) represents the fusion results of LFS without considering local regions in Eq (17), while Figure 12(a4–e4) show the fusion results by utilizing Eq (17). To intensify the subjective visual contrast, specific regions are magnified and emphasized using red arrows for comparison. A careful examination of the magnified views reveals that Figure 12(a3) appears darker and lacks brightness and detail compared to Figure 12(a4). Similarly, Figure 12(e3) appears bluer and partially obscured in detail compared to Figure 12(e4). To further validate the effectiveness of ILGM in this paper, Figure 13 displays the average metric values of all the test medical images.

The red bars of the six indicators for the three types show varying degrees of improvement over the respective blue bars. This indicates that considering the local region setting in LFS fusion can significantly enhance the fusion performance.

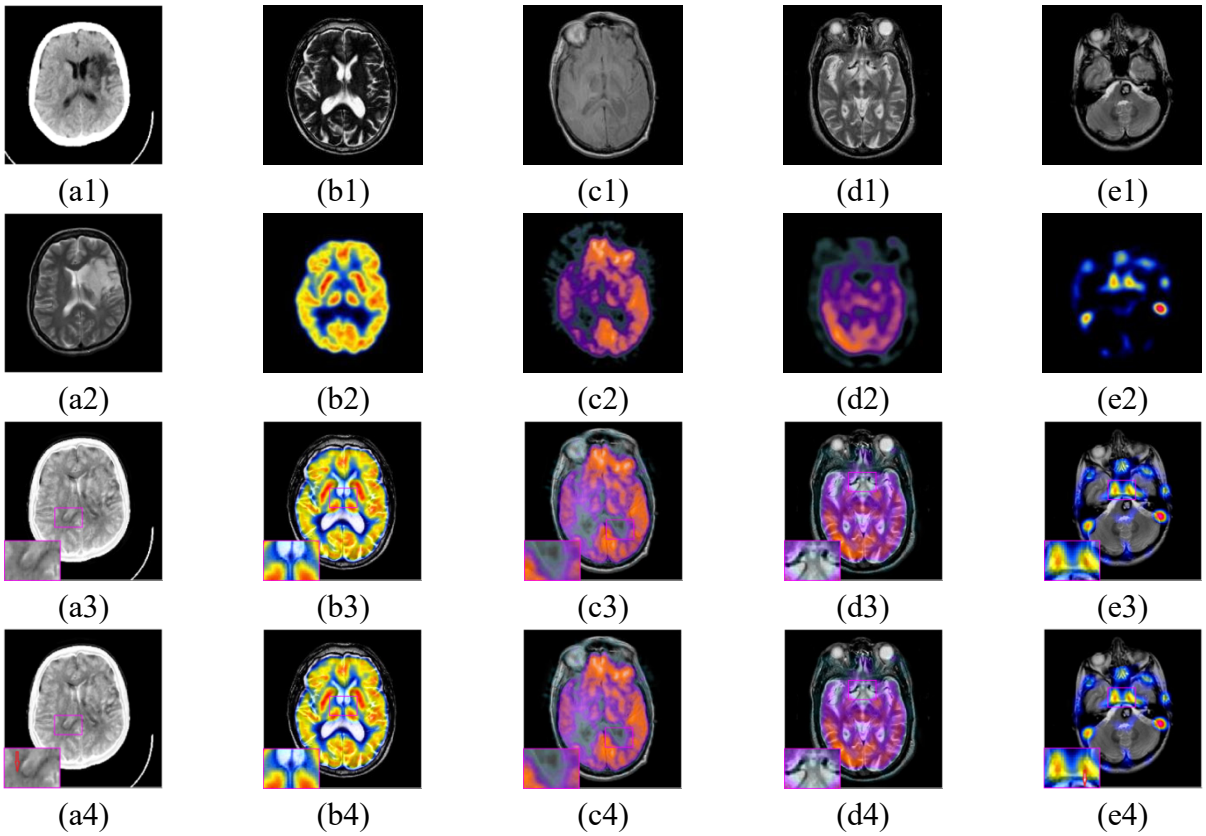


Figure 12. Subjective assessment with and without local region Ω .

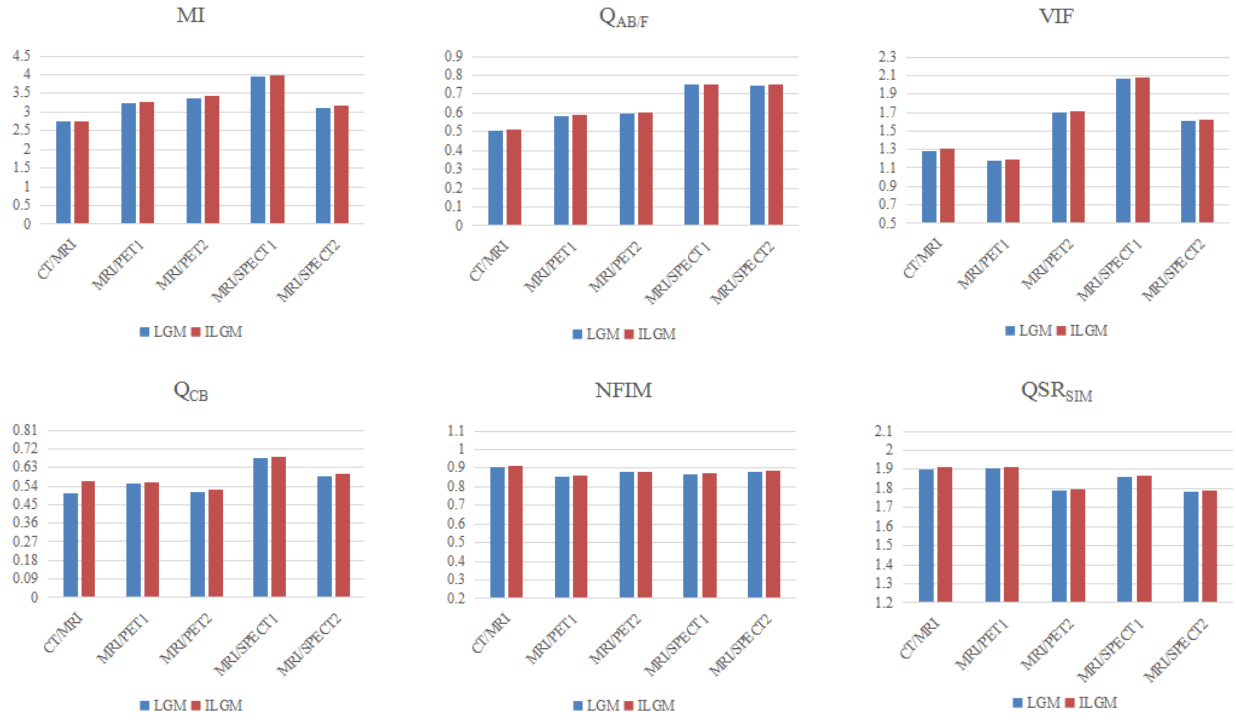


Figure13. Quantitative comparative analysis of LGM and ILGM.

4.4. Experimental comparison analysis of the proposed method with other methods

4.4.1. Selection of comparison methods

Section 4.3 presents a detailed and comprehensive synthesis of the proposed method, including parameters, MSD tools, and LFS fusion strategy. In this section, we aim to demonstrate that the proposed method outperforms existing methods. To achieve this, we selected 11 representative medical image fusion methods for comparison. These methods include the DTCWT-based method [8], the IHS+PCA method based on principal component analysis [47], the NSST and fuzzy set-based method NSST+fuzzy [11], the local extrema-based medical image fusion method LES+DC [48], the sparse theory-based method NSST+SR [25], the local Laplace filter and interest-based method LLF+IOI [49], the deep convolutional neural network-based method CNN [28], the RGF-based method RGF+JBF [15], the NSST domain-based parameter adaptive PCNN method PA-PCNN [10], the co-occurrence filter-based method CoF [20], and the joint BF-based local gradient energy method JBF+LGE [23]. These 11 medical fusion methods cover a wide range of traditional transform domain methods (DTCWT, IHS+PCA, NSST+fuzzy, LES+DC, PA-PCNN), dictionary learning-based fusion method (NSST+SR, CNN), and multi-scale filtering decomposition-based methods (LLP+IOI, RGF+JBF, CoF, JBF+LGE).

4.4.2. Analysis of experimental results

To demonstrate the practical value of the suggested medical image fusion method and its potential influence on biomedical and medical diagnoses, we experimented with all images in the dataset. In addition, a wavelet domain-based contrast metric NCM is added to more fully demonstrate the contrast of the fusion results [50]. Due to space limitations, we randomly selected two groups from the three types of medical images for presentation, as shown in Figures 14–16 and Tables 2–4.

1) The fusion results of CT and MRT

In Figure 14, a visual comparison of 12 fusion methods were employed to integrate two groups of CT/MRI medical images. Through analyzing the fusion outcomes of both sets and the magnified views, a clear evaluation can be made. The worst fusion results are seen in DTCWT (Figure 14(c1), (c2)) and IHS+PCA (Figure 14(d1), (d2)), showing blurred and dark images. NSST+fuzzy, NSST+SR and CNN have unclear textures in magnified views (Figure 14(e1), (g1), (i1)), with poorly rendered black patches resulting in excessive masking of details (Figure 14(e2), (g2), (i2)). LES+DC and AP-PCNN present much improvement compared to previous fusion methods, displaying clear, detailed texture information (Figure 14(f1), (k1)) with the effect of black patches gradually fading (Figure 14(f2), (k2)). However, a closer comparison of the magnified views reveals that the extraction of detailed information and texture processing for black patches are still insufficient. LLF+IOI (Figure 14(h1), (h2)), RGF+JBF (Figure 14(j1), (j2)) and CoF (Figure 14(l1), (l2)) show certain advantages in terms of clarity and detail extraction, but the proposed method still outperforms them in texture processing. JBF+LGE results in a severe loss of detail in Figure 14(m1), (m2) due to extracting only one layer of detail image and producing blurred texture information in flat areas. Overall, subjective analysis reveals that the proposed method outperforms the other nine methods in detail extraction, sharpness, and brightness.

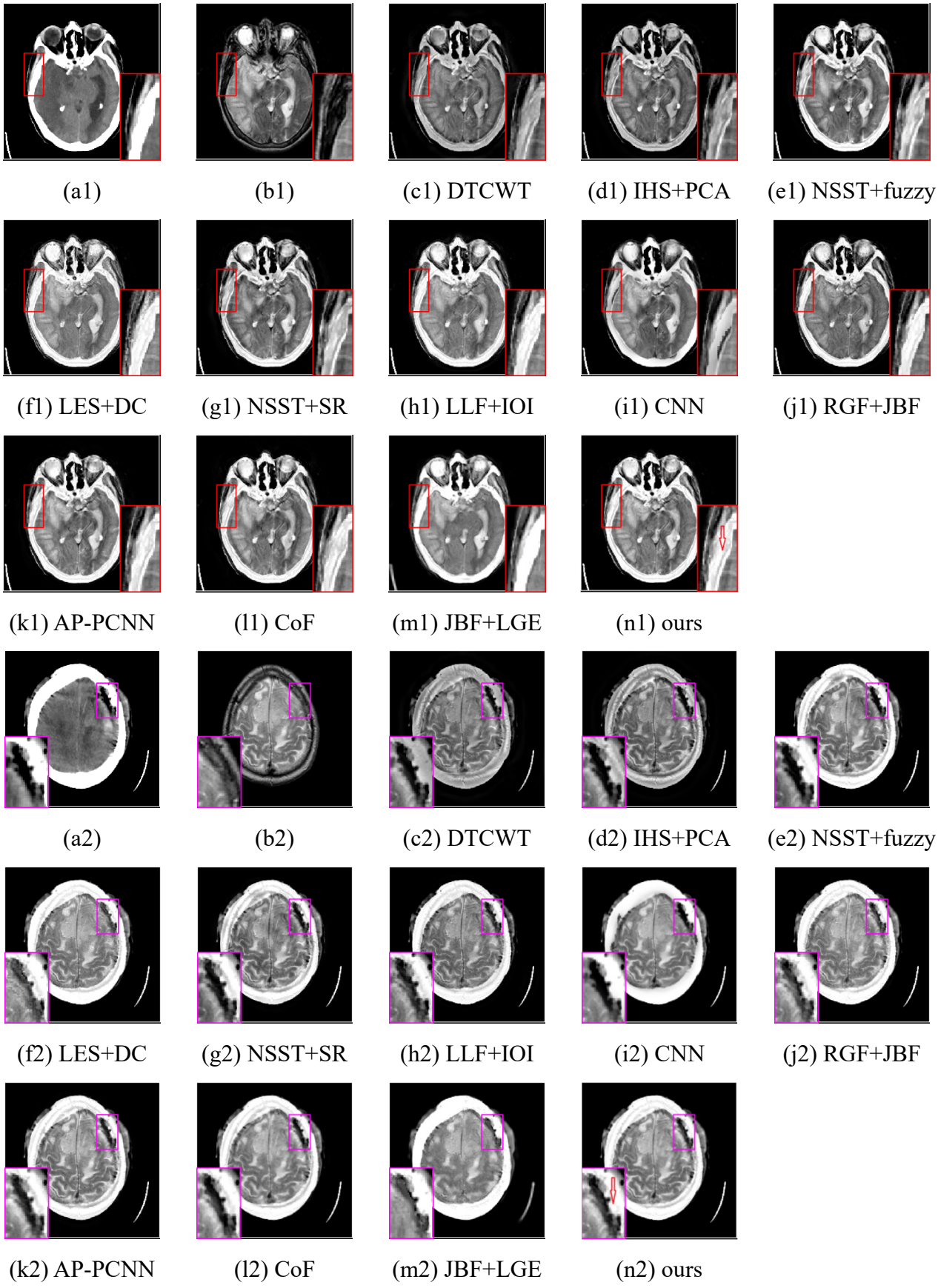


Figure 14. Comparison of subjective results of CT/MRI.

In Figure 14, a visual comparison of 12 fusion methods were employed to integrate two groups of CT/MRI medical images. Through analyzing the fusion outcomes of both sets and the magnified views, a clear evaluation can be made. The worst fusion results are seen in DTCWT (Figure 14 (c1), (c2)) and IHS+PCA (Figure 14(d1), (d2)), showing blurred and dark images. NSST+fuzzy, NSST+SR and CNN have unclear textures in the magnified views (Figure 14(e1), (g1), (i1)), with poorly rendered black patches resulting in excessive masking of details (Figure 14(e2), (g2), (i2)). LES+DC and AP-PCNN present much improvement compared to previous fusion methods, displaying clear, detailed texture information (Figure 14(f1), (k1)) with the effect of black patches gradually fading (Figure 14(f2), (k2)). However, a closer comparison of the magnified views reveals that the extraction of detail information and texture processing for black patches are still insufficient. LLF+IOI (Figure 14(h1), (h2)), RGF+JBF (Figure 14(j1), (j2)) and CoF (Figure 14(l1), (l2)) show certain advantages in terms of clarity and detail extraction, but the proposed method still outperforms them in texture processing. JBF+LGE results in a severe loss of detail in Figure 14(m1), (m2) due to extracting only one layer of detailed images and producing blurred texture information in flat areas. Overall, subjective analysis reveals that the proposed method outperforms the other nine methods in detail extraction, sharpness, and brightness.

Table 2 displays the average metric values for all CT/MRI images. After a detailed examination of each row in the table, it can be observed that the proposed method achieved high scores on all seven objective metrics. Although NSST+fuzzy and AP-PCNN exhibit maximum values on QAB/F and QSRSIM, respectively, the proposed method performs slightly lower on these two metrics, and the fusion results of NSST+fuzzy and AP-PCNN show promising performances as depicted in Figure 14. Nonetheless, in terms of subjective and objective evaluations, the proposed method outperforms all competing methods.

Table 2. Average metric values of CT/MRI.

metrics	DTCWT	IHS+ PCA	NSST+ fuzzy	LES+ DC	NSST+ SR	LLF+ IOI	CNN	RGF+ JBF	AP- PCNN	CoF	JBF+ LGE	ours
MI	2.4035	2.4788	2.8032	3.0530	2.7044	2.8654	2.7456	2.9393	2.9248	2.9964	3.0115	3.4672
$Q_{AB/F}$	0.5644	0.6531	0.6648	0.6462	0.6547	0.5739	0.5467	0.6176	0.6566	0.5467	0.3690	0.6164
VIF	1.4266	1.7128	2.1656	2.1713	2.1717	2.0920	2.0687	2.0573	2.1649	1.8754	1.6587	2.2197
Q_{CB}	0.4893	0.4794	0.5878	0.6007	0.5323	0.5423	0.6875	0.6679	0.5527	0.6678	0.6742	0.7188
NFIM	0.8610	0.8596	0.8615	0.8592	0.8582	0.8654	0.8566	0.8647	0.8592	0.8613	0.8609	0.8644
QSR_{SIM}	1.8852	1.8958	1.8962	1.8883	1.8920	1.8677	1.8745	1.8772	1.8943	1.8845	1.8703	1.8936
NCM	5.2458	4.5879	3.9874	4.2574	4.1264	5.2461	3.1254	3.4562	4.2564	6.2546	8.3254	9.4560

2) The fusion results of MRI and PET

CT/MRI are two sets of grayscale medical images. The subsequent section will present two sets of color medical images, namely MRI/PET. To accentuate the superior visual fusion effect, enlarged views of the fusion outcomes will be directly displayed on the right side, as demonstrated in Figure 15.

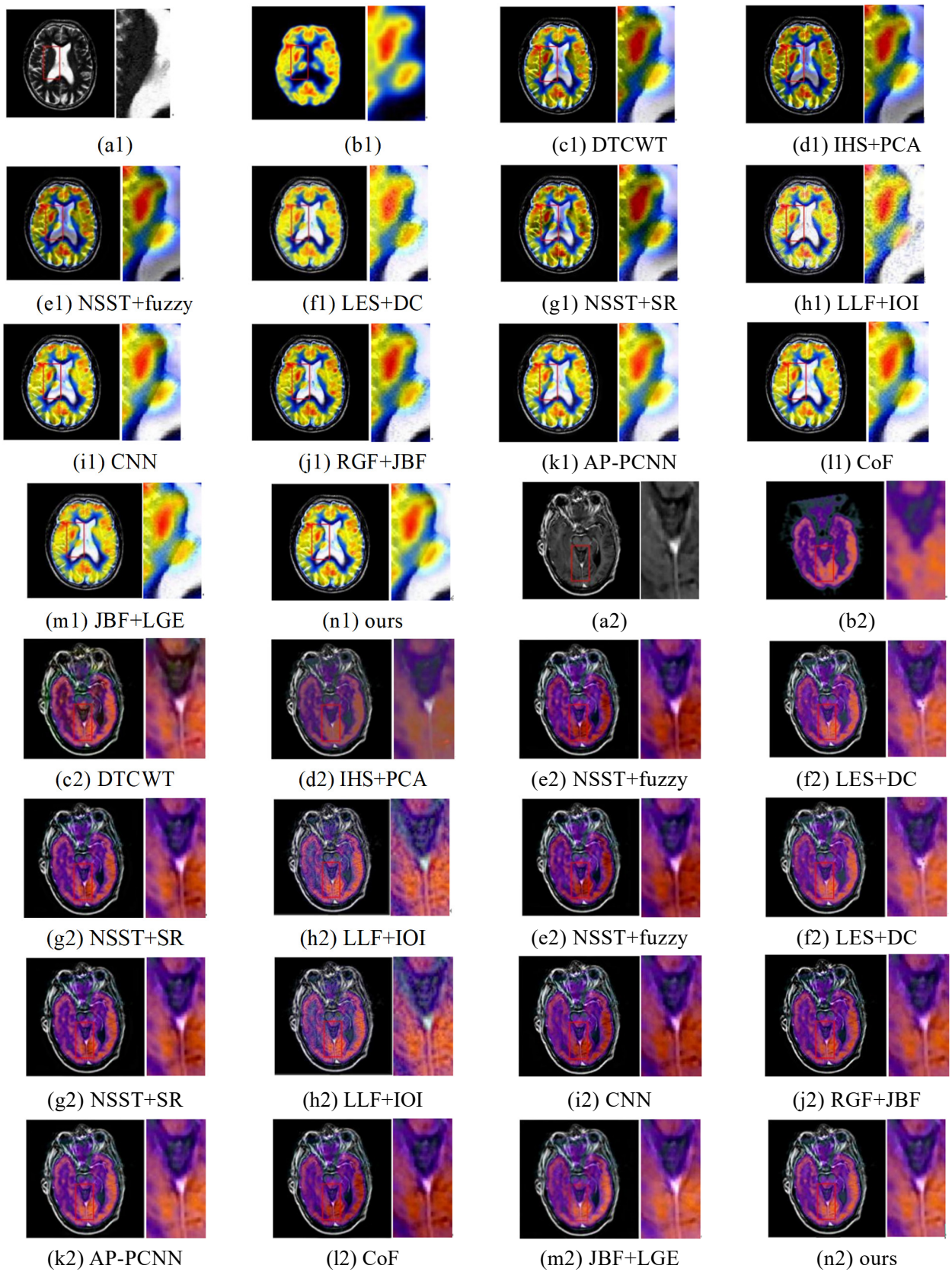


Figure 15. Comparison of subjective results of MRI/PET.

The most significant visual disparities in the fusion results are related to color accuracy, brightness,

and details. Regarding the traditional MSD methods, such as DTCWT, IHS+PCA, NSST+fuzzy, and NSST+SR, the overall visual appearance is inadequate, and the fusion results (Figure 15(c1), (d1), (e1), (g1), (c2), (d2), (e2), (g2)) appear heavily darkened, dull, and blurred. Often, the central portion is heavily displayed in blue. Additionally, the magnified views on the right reveal artifacts. The fusion outcomes based on LES+DC, LLF+IOI and CNN (Figure 15(f1), (h1), (i1), (f2), (h2), (i2)) are visibly superior to the previous results in terms of brightness and color, but upon closer inspection, the details are not smooth enough. Noise-like phenomena are apparent, especially in LLF+IOI, which displays severe color noise. More recent methods proposed, including RGF+JBF, AP-PCNN, CoF and JBF+LGE (Figure 15(j1), (k1), (l1), (m1), (j2), (k2), (l2), (m2)), exhibit satisfactory overall visual effects. However, an in-depth analysis of the magnified views reveals that image sharpness, the extraction of details, and color fidelity are still inferior to the proposed method. To further validate the superiority of the proposed method, Table 3 presents the objective average evaluation results of both sets of images.

Table 3. Average metric values of MRI/PET.

metrics	DTCWT	IHS+ PCA	NSST+ fuzzy	LES+ DC	NSST+ SR	LLF+ IOI	CNN	RGF+ JBF	AP- PCNN	CoF	JBF+ LGE	ours
MI	1.1138	2.2231	3.1505	4.5764	2.9127	3.0844	3.0124	2.8866	3.2568	3.7845	3.4206	4.5952
$Q_{AB/F}$	0.3525	0.6389	0.6897	0.6834	0.6954	0.6204	0.6214	0.6760	0.7035	0.7148	0.7009	0.7239
VIF	0.6290	1.2682	1.5087	1.5954	1.6630	1.6421	1.6257	1.4648	1.7231	1.7245	1.7367	1.6019
Q_{CB}	0.4499	0.6249	0.5726	0.6504	0.5718	0.5931	0.5876	0.5691	0.5812	0.5964	0.5850	0.6623
NFIM	0.7647	0.8894	0.8326	0.8906	0.8934	0.8783	0.8845	0.8814	0.8957	0.8854	0.8958	0.8982
QSR_{SIM}	1.7049	1.7927	1.8326	1.8308	1.8468	1.8460	1.8454	1.8416	1.8376	1.8414	1.8389	1.8490
NCM	3.4598	2.2564	3.2145	4.2364	3.2564	4.2561	3.2654	4.2451	5.2542	4.2451	5.2654	6.2542

Table 3 illustrates the outcomes of the objective evaluation of all MRI/PET images. It is evident that the metric values based on traditional methods, including DTCWT, IHS+PCA, NSST+fuzzy, LES+DC, and NSST+SR, are generally low, while the objective assessment based on LLF+IOI, CNN, RGF+JBF, AP-PCNN, CoF and JBF+LGE have demonstrated improvement. A detailed comparison of each row of data demonstrates that the fusion results based on multi-scale filtering decomposition are notably superior to traditional methods, especially JBF+LGE, which runs with high efficiency and attains the maximum value in the VIF metric. Moreover, other metrics are only slightly lower than the proposed method. From the NCM metrics in the table, it can be seen that the proposed method has the largest value, which indicates that our method has the highest contrast. Hence, from the comprehensive analysis of Figure 15 and Table 3, it can be deduced that the proposed method achieves exceptional fusion performance.

3) The fusion results of MRI and SPECT

MRI delivers a significant amount of textural detail information, while SPECT reveals blood flow in color images. In this study, we experimentally analyze this type of medical image, as illustrated in Figure 16 and Table 4.

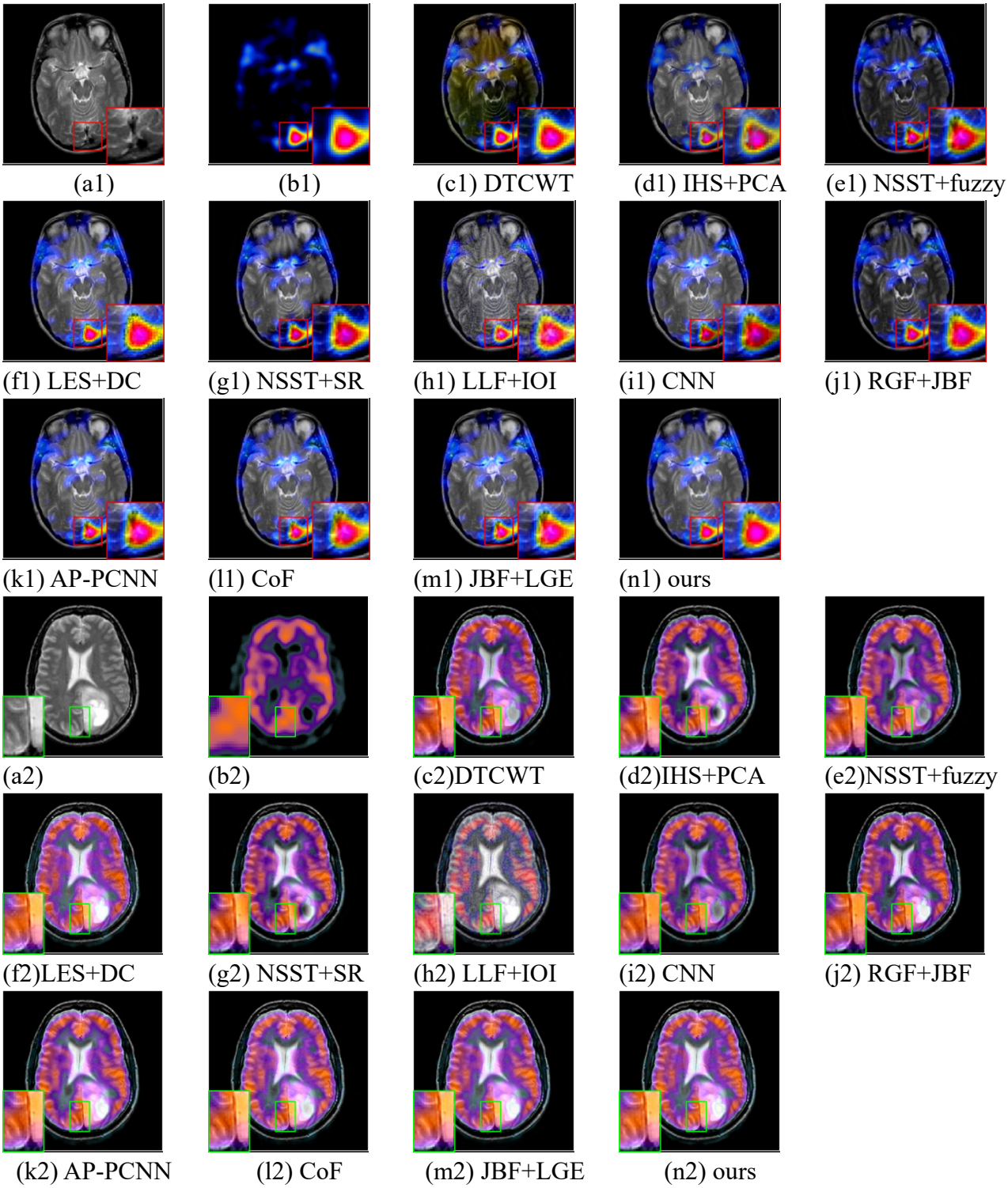


Figure 16. Comparison of subjective results of MRI/SPECT.

Overall, the fusion results based on IHS+PCA and NSST+fuzzy exhibit the worst performance, with severe color distortion visible in the magnified views. The fusion outcomes based on DTCWT and LES+DC appear dim and are inadequate in detail extraction. The fusion results obtained with LLF+IOI and CNN exhibit severe color distortion, and a certain degree of a noise-like phenomenon is also present. Although the results based on NSST+SR and AP-PCNN have improved, some artifacts

in the details are still visible. The fusion outcomes based on RGF+JBF, CoF and JBF+LGE appear visually pleasing. However, upon closer inspection of the comparisons, the detail extraction in flat areas and texture processing at boundaries are not as good as the proposed method. Therefore, the proposed method performs exceptionally well in detail extraction, boundary processing, and color fidelity.

Table 4 displays the objective average evaluation results for all MRI/SPECT images, from which it can be observed that AP-PCNN achieves the highest value in VIF. In conjunction with Figure 16, it indicates that the method yields visually appealing fusion results. Additionally, the proposed method attains the best values for MI, QAB/F, QCB, NFIM, QSRSIM, and NCM metrics, further highlighting the absolute advantages of the proposed method in both visual and quantitative assessments.

Table 4. Average metric values of MRI/SPECT.

metrics	DTCWT	IHS+ PCA	NSST+ fuzzy	LES+ DC	NSST+ SR	LLF+ IOI	CNN	RGF+ JBF	AP- PCNN	CoF	JBF+ LGE	ours
MI	1.0562	1.2823	3.0643	3.5502	3.1591	3.0056	2.9987	2.9894	3.0786	3.3185	3.2505	4.3667
$Q_{AB/F}$	0.5421	0.5625	0.6144	0.5987	0.6270	0.5936	0.6074	0.6387	0.6380	0.6544	0.6388	0.6913
VIF	0.2237	0.2984	1.5319	1.5150	1.7053	1.6982	1.6648	1.5340	1.7148	1.6732	1.6894	1.7016
Q_{CB}	0.2874	0.3037	0.5500	0.6876	0.5602	0.6935	0.6457	0.6245	0.5640	0.6745	0.5945	0.7294
NFIM	0.7956	0.8293	0.8757	0.8739	0.8788	0.8706	0.8744	0.8682	0.8810	0.8854	0.8836	0.8906
QSR_{SIM}	1.5647	1.5465	1.9138	1.9102	1.9241	1.9240	1.9146	1.9225	1.9138	1.9056	1.9101	1.9241
NCM	3.2541	4.2546	3.4587	4.2564	5.1247	3.4589	4.4578	5.2471	7.0012	6.5475	6.4254	7.5987

4) Overall evaluation of the experiments

The experimental analysis of the three different types of medical images mentioned above has enabled us to establish several key findings:

• Traditional MSD-based methods, such as DTCWT, IHS+PCA, and NSST+fuzzy, often produce visually ineffective fusion results for medical images, exhibiting a lack of sharpness and clarity.

• NSST+SR and CNN significantly improve the visual effect; however, there are deficiencies in detail, texture, and boundary processing, and the running time cost is high.

• Multi-scale filtering decomposition methods, such as LLF+IOI, RGF+JBF, CoF, and JBF+LGE, improve operational efficiency and enhance fusion performance. However, LLF+IOI still produces obvious color distortions, and the fusion frameworks of RGF+JBF, CoF and JBF+LGE often suffer from some shortcomings in detail extraction and boundary processing.

• After a comprehensive demonstration and experimental comparison, the proposed method has demonstrated exceptional performance in terms of visual effect, detail extraction, color fidelity, and objective evaluation.

4.4.3. Computational efficiency

Computational efficiency is often an important metric for evaluating fusion methods. Higher running times often lead to excessive costs; therefore, it is desirable to control computational costs as much as possible when designing algorithms. In this study, the computational efficiency of the methods is compared with the 11 methods in Section 4.4.2, and the experiments are conducted on a machine running Matlab 2019a with an Intel(R) Core(TM) i5-6200U CPU @ 2.30 GHz 2.40 GHz. All the

experiments are based on a time complexity of $O(N^2)$ to fuse a pair of $N \times N$ medical images. The average running times of different fusion methods are presented in Table 5.

Table 5. Average running time.

metrics	DTCWT	IHS+ PCA	NSST+ fuzzy	LES+ DC	NSST+ SR	LLF+ IOI	CNN	RGF+ JBF	AP-PCNN	CoF	JBF+ LGE	ours
times	13.376s	11.256s	14.634s	12.134s	23.265s	9.295s	54.248s	15.334s	15.765s	14.254s	4.156s	8.918s

The majority of the methods have running times exceeding 10 seconds. Although JBF+LGE significantly reduces running costs, the extracted detail information is too coarse, resulting in inferior fusion results compared to the proposed method. The proposed method extracts rich and detailed information from various scales and directions, which slightly increases the running costs compared to JBF+LGE. However, we spend some time cost in image fusion to construct a rich detail extraction tool and fuse based on each pixel of subband images in LFS and HFS. Nonetheless, the running times of the proposed method are controlled within 10 seconds. Therefore, a comprehensive analysis of the method design in Section 3 and the experimental part in Section 4 reveals that the design of the proposed method is not complex, visually appealing, and has a low runtime cost.

5. Conclusions and prospect

Medical image fusion is a highly relevant and dynamic research topic with significant practical application value. Nonetheless, achieving a balance between visual impact, quantitative evaluation, and the costs associated with producing fusion results remains a challenging task. This paper presented a novel medical image fusion method that combines MSRWGIF and ILGM. We focused on improving critical aspects of detail preservation and boundary processing given the practical context of medical images. It involves constructing a multi-scale and multi-directional decomposition tool, MSRWGIF, and combining it with ILGM for LFS and AP-PCNN for HFS. More than 100 medical image pairs were selected, and qualitative and quantitative analyses were conducted using seven metrics and eleven representative fusion methods. The results show that the proposed method outperforms current methods regarding color fidelity, detail extraction, and edge processing.

However, the parameters in the proposed method are not adaptive, and the fusion effect will be affected by the parameters to a certain extent, so choosing the optimal and most suitable parameters needs further study. In addition, the experimental dataset in this paper has just over 100 pairs of medical images, which is relatively small compared to the current popular deep learning. Therefore, in the future, the plan is to integrate this method with popular deep machine learning techniques to address unregistered medical images and effective denoising, extending research into biometric recognition and intelligent medical technology.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Data availability statements

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Acknowledgments

The authors of this paper express their sincere gratitude to the editors and anonymous reviewers for their careful review of this article, comprehensive feedback, and invaluable suggestions despite their busy schedules. This work is sponsored by the Key Project of Natural Science Research of Education Department of Anhui Province (Grant No. KJ2021A1154, No. 2022AH051750), in part by the Fundamental Research Funds for the Tongling University (No. 2022tlxyrc11), and the General Program of Domestic Visit and Research for Outstanding Young Talents of Universities and Colleges (Grant No. gxgnfx2022092).

References

1. M. M. Emam, E. H. Houssein, R. M. Ghoniem, A modified reptile search algorithm for global optimization and image segmentation: Case study brain MRI images, *Comput. Biol. Med.*, **152** (2023), 106404. https://doi.org/10.1016/j.compbiomed.2022.106404
2. E. H. Houssein, D. A. Abdelkareem, M. M. Emam, M. A. Hameed, M. Younan, An efficient image segementation method for skin cancer imaging using improved golden jackal optimization algorithm, *Comput. Biol. Med.*, **149** (2022), 106075. https://doi.org/10.1016/j.compbiomed.2022.106075
3. W. Zhu, L. Liu, F. Kuang, L. Li, S. Xu, Y. Liang, An efficient multi-threshold image segmentation for skin cancer using boosting whale optimizer, *Comput. Biol. Med.*, **151** (2022), 106227. https://doi.org/10.1016/j.compbiomed.2022.106227
4. L. Nie, L. Zhang, L. Meng, X. Song, X. Chang, X. Li, Modeling disease progression via multisource multitask learners: A case study with Alzheimer's disease, *IEEE Trans. Neural Networks Learn. Syst.*, **28** (2017), 1508–1519. https://doi.org/10.1109/TNNLS.2016.2520964
5. J. Tang, Q. Sun, Z. Wang, Y. Cao, Perfect-reconstruction 4-tap size-limited filter banks for image fusion application, in *2007 International Conference on Mechatronics and Automation*, (2007), 255–260. https://doi.org/10.1109/ICMA.2007.4303550
6. J. Tang, A contrast based image fusion technique in the DCT domain, *Digital Signal Process.*, **14** (2004), 218–226. https://doi.org/10.1016/j.dsp.2003.06.001
7. E. Candès, L. Demanet, D. Donoho, L. Ying, Fast discrete curvelet transforms, *Multiscale Model. Simul.*, **5** (2006), 861–899. https://doi.org/10.1137/05064182X
8. B. Yu, B. Jia, L. Ding, Z. Cai, Q. Wu, R. Law, et al., Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion, *Neurocomputing*, **182** (2016), 1–9. https://doi.org/10.1016/j.neucom.2015.10.084
9. Z. Zhu, M. Zheng, G. Qi, D. Wang, Y. Xiang, A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain, *IEEE Access*, **7** (2019), 20811–20824. https://doi.org/10.1109/ACCESS.2019.2898111

10. M. Yin, X. Liu, Y. Liu, X. Chen, Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain, *IEEE Trans. Instrum. Meas.*, **68** (2019), 49–64. https://doi.org/10.1109/TIM.2018.2838778
11. H. Ullah, B. Ullah, L. Wu, F. Y. O. Abdalla, G. Ren, Y. Zhao, Multi-modality medical images fusion based on local-features fuzzy sets and novel sum-modified-Laplacian in non-subsampled shearlet transform domain, *Biomed. Signal Process. Control*, **57** (2020), 101724. https://doi.org/10.1016/j.bspc.2019.101724
12. Z. Zhou, B. Wang, S. Li, M. Dong, Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters, *Inf. Fusion*, **30** (2016), 15–26. https://doi.org/10.1016/j.inffus.2015.11.003
13. X. Qiu, M. Li, L. Zhang, X. Yuan, Guided filter-based multi-focus image fusion through focus region detection, *Signal Process. Image Commun.*, **72** (2019), 35–46. https://doi.org/10.1016/j.image.2018.12.004
14. L. Caraffa, J. P. Tarel, P. Charbonnier, The guided bilateral filter: when the joint/cross bilateral filter becomes robust, *IEEE Trans. Image Process.*, **24** (2015), 1119–1208. https://doi.org/10.1109/TIP.2015.2389617
15. L. Jian, X. Yang, Z. Zhou, K. Zhou, K. Liu, Multi-scale image fusion through rolling guidance filter, *Future Gener. Comput. Syst.*, **83** (2018), 310–325. https://doi.org/10.1016/j.future.2018.01.039
16. J. Du, W. Li, B. Xiao, Fusion of anatomical and function images using parallel saliency features, *Inf. Sci.*, **430–431** (2018), 567–576. https://doi.org/10.1016/j.ins.2017.12.008
17. R. J. Jevnisek, S. Avidan, Co-occurrence filter, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 3816–3824. https://doi.org/10.1109/CVPR.2017.406
18. Z. Li, J. Zheng, Z. Zhu, W. Yao, S. Wu, Weighted guided image filtering, *IEEE Trans. Image Process.*, **24** (2015), 120–129. https://doi.org/10.1109/TIP.2014.2371234
19. H. Yin, Y. Gong, G. Qiu, Side window guided filtering, *Signal Process.*, **165** (2019), 315–330. https://doi.org/10.1016/j.sigpro.2019.07.026
20. M. Diwakar, P. Singh, A. Shankar, Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain, *Biomed. Signal Process. Control*, **68** (2021), 102788. https://doi.org/10.1016/j.bspc.2021.102788
21. W. Liu, Z. Wang, A novel multi-focus image fusion method using multiscale shearing non-local guided averaging filter, *Signal Process.*, **166** (2020), 107252. https://doi.org/10.1016/j.sigpro.2019.107252
22. B. Meher, S. Agrawal, R. Panda, A. Abraham, A survey on region based image fusion methods, *Inf. Fusion*, **48** (2019), 119–132. https://doi.org/10.1016/j.inffus.2018.07.010
23. X. Li, F. Zhou, H. Tan, W. Zhang, C. Zhao, Multimodal medical image fusion based on joint bilateral filter and local gradient energy, *Inf. Sci.*, **569** (2021), 302–325. https://doi.org/10.1016/j.ins.2021.04.052
24. C. Xing, Z. Wang, Q. Quyang, C. Dong, C. Duan, Image fusion method based on spatially masked convolutional sparse representation, *Image Vision Comput.*, **90** (2019), 103806. https://doi.org/10.1016/j.imavis.2019.08.010
25. S. Maqsood, U. Javed, Multi-modal medical image fusion based on two-scale image decomposition and sparse representation, *Biomed. Signal Process. Control*, **57** (2020), 101810. https://doi.org/10.1016/j.bspc.2019.101810

26. S. Goyal, V. Singh, A. Rani, N. Yadav, FPRSGF denoised non-subsampled shearlet transform-based image fusion using sparse representation, *Signal Image Video Process.*, **14** (2020), 719–726. https://doi.org/10.1007/s11760-019-01597-z
27. F. Zhou, X. Li, M. Zhou, Y. Chen, H. Tan, A new dictionary construction based multimodal medical image fusion framework, *Entropy*, **21** (2019), 267. https://doi.org/10.3390/e21030267
28. Y. Liu, X. Chen, R. K. Ward, Z. J. Wang, Medical image fusion via convolutional sparsity based morphological component analysis, *IEEE Signal Process. Lett.*, **26** (2019), 485–489. https://doi.org/10.1109/lsp.2019.2895749
29. Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, L. Zhang, IFCNN: A general image fusion framework based on convolutional neural network, *Inf. Fusion*, **54** (2020), 99–118. https://doi.org/10.1016/j.inffus.2019.07.011
30. H. Li, Y. Wang, Z. Yang, R. Wang, X. Li, D. Tao, Discriminative dictionary learning-based multiple component decomposition for detail-preserving noisy image fusion, *IEEE Trans. Instrum. Meas.*, **69** (2020), 1082–1102. https://doi.org/10.1109/TIM.2019.2912239
31. H. Li, M. Yang, Z. Yu, Joint image fusion and super-resolution for enhanced visualization via semi-coupled discriminative dictionary learning and advantage embedding, *Neurocomputing*, **422** (2021), 62–84. https://doi.org/10.1016/j.neucom.2020.09.024
32. Q. Hu, S. Hu, F. Zhang, Multi-modality medical image fusion based on separable dictionary learning and Gabor filtering, *Signal Process. Image Commun.*, **83** (2020), 115758. https://doi.org/10.1016/j.image.2019.115758
33. J. Ma, H. Xu, J. Jiang, X. Mei, X. Zhang, DDcGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion, *IEEE Trans. Image Process.*, **29** (2020), 4980–4995. https://doi.org/10.1109/TIP.2020.2977573
34. H. Zhang, H. Xu, X. Tian, J. Jiang, J. Ma, Image fusion meets deep learning: A survey and perspective, *Inf. Fusion*, **76** (2021), 323–336. https://doi.org/10.1016/j.inffus.2021.06.008
35. K. Zhan, J. Shi, H. Wang, Y. Xie, Q. Li, Computational mechanisms of pulse-coupled neural networks: A comprehensive review, *Arch. Computat. Methods Eng.*, **24** (2017), 573–588. https://doi.org/10.1007/s11831-016-9182-3
36. Y. Chen, S. Park, Y. Ma, R. Ala, A new automatic parameter setting method of a simplified PCNN for image segmentation, *IEEE Trans. Neural Networks*, **22** (2011). https://doi.org/10.1109/TNN.2011.2128880
37. G. Qu, D. Zhang, P. Yan, Information measure for performance of image fusion, *Electron. Lett.*, **38** (2002), 313–315. https://doi.org/10.1049/EL:20020212
38. C. S. Xydeas, V. Petrovic, Objective image fusion performance measure, *Electron. Lett.*, **36** (2000), 308–309. https://doi.org/10.1049/el:20000267
39. Y. Han, Y. Cai, Y. Cao, X. Xu, A new image fusion performance metric based on visual information fidelity, *Inf. Fusion*, **14** (2013), 127–135. https://doi.org/10.1016/j.inffus.2011.08.002
40. Y. Chen, R. S. Blum, A new automated quality assessment algorithm for image fusion, *Image Vision Comput.*, **27** (2009), 1421–1432. https://doi.org/10.1016/j.imavis.2007.12.002
41. M. B. A. Haghighat, A. Aghagolzadeh, H. Seyedarabi, A non-reference image fusion metric based on mutual information of image features, *Comput. Electr. Eng.*, **37** (2011), 744–756. https://doi.org/10.1016/j.compeleceng.2011.07.012

42. L. Zhang, H. Li, SR-SIM: A fast and high performance IQA index based on spectral residual, in *2012 19th IEEE International Conference on Image Processing*, **19** (2012), 6467149. https://doi.org/10.1109/ICIP.2012.6467149
43. Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, W. Wu, Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study, *IEEE Trans. Pattern Anal. Mach. Intell.*, **34** (2012), 94–109. https://doi.org/10.1109/TPAMI.2011.109
44. Z. Zhu, Y. Chai, H. Yin, Y. Li, Z. Liu, A novel dictionary learning approach for multi-modality medical image fusion, *Neurocomputing*, **214** (2016), 471–482. https://doi.org/10.1016/j.neucom.2016.06.036
45. F. Zhou, X. Li, M. Zhou, Y. Chen, H. Tan, A new dictionary construction based multimodal medical image fusion framework, *Entropy*, **21** (2019), 1–20. https://doi.org/10.3390/e21030267
46. M. Kim, D. K. Han, H. Ko, Joint patch clustering-based dictionary learning for multimodal image fusion, *Inf. Fusion*, **27** (2016), 198–214. https://doi.org/10.1016/j.inffus.2015.03.003
47. C. He, Q. Liu, H. Li, H. Wang, Multimodal medical image fusion based on IHS and PCA, *Procedia Eng.*, **7** (2010), 280–285. https://doi.org/10.1016/j.proeng.2010.11.045
48. Z. Xu, Medical image fusion using multi-level local extrema, *Inf. Fusion*, **19** (2014), 38–48. https://doi.org/10.1016/j.inffus.2013.01.001
49. J. Du, W. Li, B. Xiao, Anatomical-Functional image fusion by information of interest in local Laplacian filtering domain, *IEEE Trans. Image Process.*, **26** (2017), 5855–5866. https://doi.org/10.1109/TIP.2017.2745202
50. J. Tang, Q. Sun, K. Agyepong, An image enhancement algorithm based on a new contrast measure in the wavelet domain for screening mammograms, in *2007 IEEE International Conference on Image Processing*, **5** (2007), 16–19. https://doi.org/10.1109/ICIP.2007.4379757

AIMS Press