

# Medical Image Fusion and Denoising Algorithm Based on a Decomposition Model of Hybrid Variation-Sparse Representation

Guofen Wang, Weisheng Li<sup>ID</sup>, Jiao Du<sup>ID</sup>, Bin Xiao<sup>ID</sup>, and Xinbo Gao<sup>ID</sup>, *Member, IEEE*

**Abstract**—Medical image fusion technology integrates the contents of medical images of different modalities, thereby assisting users of medical images to better understand their meaning. However, the fusion of medical images corrupted by noise remains a challenge. To solve the existing problems in medical image fusion and denoising algorithms related to excessive blur, unclean denoising, gradient information loss, and color distortion, a novel medical image fusion and denoising algorithm is proposed. First, a new image layer decomposition model based on hybrid variation-sparse representation and weighted Schatten p-norm is proposed. The alternating direction method of multipliers is used to update the structure, detail layer dictionary, and detail layer coefficient map of the input image while denoising. Subsequently, appropriate fusion rules are employed for the structure layers and detail layer coefficient maps. Finally, the fused image is restored using the fused structure layer, detail layer dictionary, and detail layer coefficient maps. A large number of experiments confirm the superiority of the proposed algorithm over other algorithms. The proposed medical image fusion and denoising algorithm can effectively remove noise while retaining the gradient information without color distortion.

**Index Terms**—Medical image fusion, Image denoising, Hybrid variation-sparse representation, Weighted Schatten p-norm, Alternating direction method of multipliers.

Manuscript received 27 April 2022; revised 15 July 2022 and 28 July 2022; accepted 29 July 2022. Date of publication 5 August 2022; date of current version 7 November 2022. This work was supported in part by the National Key Research and Development Program of China under Grants 2016YFC1000307-3 and 2019YFE0110800, in part by the National Natural Science Foundation of China under Grants U1713213, 62027827, 62176071, and 61972060, in part by the National Science Foundation of Chongqing under Grants cstc2019cxqyljrc-td0270, cstc2020jcyj-zdxdmX0025, and cstc2019jcyj-cxttX0002, in part by the Doctoral Talent Training Project of Chongqing University of Posts and Telecommunications under Grant BYJS202011, and in part by the Guangzhou Science and Technology Project under Grant 202102021243. (*Corresponding author: Weisheng Li*)

Guofen Wang, Weisheng Li, Bin Xiao, and Xinbo Gao are with the Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: d190201020@stu.cqupt.edu.cn; liws@cqupt.edu.cn; xiaboin@cqupt.edu.cn; gaoxb@cqupt.edu.cn).

Jiao Du is with the School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: dujiao19880429@126.com).

Digital Object Identifier 10.1109/JBHI.2022.3196710

## I. INTRODUCTION

THE proliferation of medical imaging modalities has resulted in the widespread use of imaging technologies such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT) to assist clinical diagnosis. However, it is difficult to fully understand the imaging data through single-modality analysis. Fortunately, medical image fusion technology allows simultaneous use of multimodal information to provide a complete understanding of a given pathology [1], [2], [3].

Hitherto, numerous classic medical image fusion methods (traditional and deep learning) have been proposed, such as image fusion based on guided filters [4]; multimodal medical image fusion based on fuzzy discrimination [5]; medical image fusion based on nonsubsampled shearlet transform (NSST) and parameter adaptive pulse coupled neural network [6]; medical image fusion method based on phase congruency and local Laplacian energy in the nonsubsampled contourlet transform domain [7]; image fusion by multiscale transform and sparse representation (SR) [8]; image fusion based on cartoon-texture decomposition and SR [9]; functional and anatomical image fusion based on local Laplacian filtering and information of interest [10]; medical image fusion via coupled neural P systems (CNPS) in NSST domain [11]; image fusion via convolutional neural networks (CNNs) [12]; image fusion by a general CNN framework (IFCNN) [13]; medical image fusion using a CNN and contrast pyramid [14]; medical image fusion through an unsupervised enhanced network (EMFusion) [15]; image fusion using a unified densely connected network (FusionDN) [16]; medical image fusion by generative adversarial network and dual stream attention mechanism (DSGAN) [17]; and a deep multi-cascade framework with classifier-based feature synthesis to fuse multimodal medical images [18]. However, the above-mentioned classic methods are only used for fusing images without noise. In other words, noise removal is not performed at the time of image fusion.

Recently, some algorithms have attempted to fuse images while removing noise. Image fusion and denoising algorithms can be divided into two categories based on the use of the variational model (VM) and SR. On the one hand, image fusion and denoising algorithms based on VM include a new VM based

on  $l_2 - l_p$  norm minimization for simultaneous fusion and denoising of infrared and visible images [19], a multi-focus image fusion and denoising algorithm based on homogeneity similarity [20], fractional order gradient information for image fusion and denoising [21], a fractional order total generalized variation method to remove noise while fusing multimodal images [22], and a medical image fusion and denoising by alternating sequential filter and adaptive fractional order total variation [23]. On the other hand, image fusion and denoising algorithms based on SR are also a research hotspot; these include group SR for medical image denoising and fusion [24], adaptive SR (ASR) for medical image denoising and fusion [25], regularization terms in the SR domain (LR-SR) to realize simultaneous fusion and denoising of medical images [26], and joint image fusion and denoising via three-layer decomposition and SR (TL-SR) [27]. Kim et al. [28] designed a multimodal image fusion method based on patch-clustering dictionary learning. This method introduces a local regression weight of the steering kernel for feature clustering, which can effectively suppress noise during fusion. Qi et al. [29] proposed a Gaussian scale mixture-based SR model for denoising and fusion of texture components, and a spatial domain fusion rule applied to cartoon components. Unfortunately, some of the aforementioned fusion algorithms still require improvement. The ASR method causes loss of brightness and color distortion in the fused image. The LR-SR algorithm is complex and causes fusion image gradient information loss and excessive blur. The TL-SR method has a poor denoising effect when the noise is excessive; however, the speed of the algorithm is an advantage. Moreover, in the method proposed by Kim, the contrast of the fused image is low, and edge information is lost.

Therefore, to address the issues of excessive blur, unclean denoising, gradient information loss, and color distortion in existing image fusion and denoising algorithms, this paper presents a new medical image fusion and denoising algorithm based on a decomposition model of hybrid variation-SR (HVSR). Image layer decomposition is a key step in producing satisfactory results; analysis shows that an image decomposition tool based on VM successfully approximates the major structure without (or with little) noise in a noisy image, whereas an image decomposition tool based on SR is suitable for describing the fine-scale textures of a noisy image. Therefore, the advantages of VM, SR, and weighted Schatten p-norm minimization (WSpNM) are introduced to ensure image layer decomposition with simultaneous removal of noise. The alternating direction method of multipliers (ADMM) is used to solve the model. Ablation experiments demonstrate that the contribution of each item to the HVSR model is essential. The VM term forces noise to exist in the detail layer, the SR term provides a continuous depiction of the details, and the addition of the WSpNM term has a significant influence on noise removal. Subjective and objective experiments were conducted to test the algorithm on images with different noise levels (270 groups) and images without noise (180 groups). The proposed algorithm outperformed classical medical image fusion and denoising algorithms in denoising and fusion tasks.

The main contributions of this study are given:

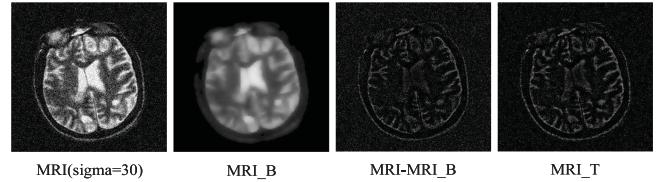


Fig. 1. Decomposition and representation of noisy images.

1) The proposed image layer decomposition model, HVSR, combines the concepts of VM and SR with the advantages of WSpNM to remove noise during image decomposition.

2) During the iterative update process, the dictionary is updated adaptively according to the input.

3) A novel medical image fusion and denoising algorithm is proposed based on the HVSR decomposition model. Extensive experiments demonstrate that, compared with other classical methods, the fusion result image of the proposed algorithm has less color distortion, less noise, and better gradient information.

The remainder of this paper is organized as follows. The proposed image layer decomposition model (HVSR) is described in Section II. In Section III, the proposed fusion and denoising algorithms are depicted in detail. Section IV describes the analysis of the parameters, ablation experiments, and subjective and objective results. Finally, Section V concludes this paper.

## II. PROPOSED IMAGE LAYER DECOMPOSITION MODEL: HVSR

A new decomposition model and its solutions are described for detail in this section.

### A. Decomposition Model: HVSR

A suitable layer decomposition method is an essential component of successful medical image fusion and denoising algorithms. The VM is generally used to obtain the majority of textural information, which indicates that the VM allows all details and noise to exist in the detail layer rather than the structural layer containing the local mean [30]. From Fig. 1, for the noisy image (MRI), the VM can effectively obtain the structure layer without noise (MRI\_B) and the detail layer containing noise and the main detail texture (MRI-MRI\_B); therefore, VM is suitable for capturing noise-free (or slightly noisy) structure layers of noisy images. The SR method may cause some information loss owing to sparse expression when restoring the detail layer; however, when restoring the noisy detail layer, the SR method allows the details of the restored detail layer to be more continuous, as shown in the fourth image (MRI\_T) of Fig. 1. Thus, the SR concept is introduced to describe the detail level. WSpNM generalizes the nuclear norm minimization by allotting weights to different singular values, which can effectively remove noise while maintaining gradient information [31]. Therefore, by fully integrating the advantages of VM, SR, and WSpNM, a new image layer decomposition model is proposed to obtain a noise-free base layer, detail layer

dictionary, and detail layer coefficient map, as follows:

$$\begin{aligned} \underset{D, \alpha, B}{\operatorname{argmin}} & \|D \circ \alpha + B - I\|_F^2 + \lambda_1 \|\nabla B\|_1 \\ & + \lambda_2 \|\alpha\|_1 + \lambda_3 \|D \circ \alpha\|_{w, s_p}^p, \end{aligned} \quad (1)$$

where  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are penalty coefficients.  $\|\cdot\|_F$  and  $\|\cdot\|_1$  separately indicate the Frobenius and  $l_1$  norms.  $\nabla$  is a first-order differential operator, and  $\circ$  represents multiplication operator. The first term  $\|D \circ \alpha + B - I\|_F^2$  considers the original image  $I$  to be divided into structural information without noise  $B$  and detailed information without noise  $D \circ \alpha$ , where  $D$  is the dictionary and  $\alpha$  is the dictionary coefficient. The second term  $\|\nabla B\|_1$  originates from the concept of VM, which allows the main illumination  $B$  to be segmentally smooth and the gradient sparse, thereby allowing most of the noise to exist in the detail layer. The third term,  $\|\alpha\|_1$ , considers the sparsity of the dictionary coefficients. The detail layers are represented by a dictionary and a coefficient map; thus, the coefficients are usually required to be sparse. The aim of the last term  $\|D \circ \alpha\|_{w, s_p}^p$  is to remove noise in the detail layer synthesized from the SR. WSpNM allows noise removal while better approximating the original detail layer.

Thus, the proposed decomposition model can remove noise while separating image layers. If there is no noise interference,  $\lambda_3 = 0$  is set for a noise-free image decomposition model. Indeed, when the other penalty coefficients are zero, the effect of this item is no longer considered; the subsequent ablation experiments demonstrate that the introduction of each item is essential.

### B. Solver of the HVSR Model

To solve the minimization problem of (1), this paper will adopt ADMM, a mature iterative solution framework technique, to iteratively update each variable [26], [32]. When solving a variable in an iteration, fixing other variables as constant invariants is the essential idea of ADMM.

(1) **Updating  $B$ :** keeping only items related to  $B$ , and the following equation describes the new subproblem:

$$\underset{B}{\operatorname{argmin}} \|D \circ \alpha + B - I\|_F^2 + \lambda_1 \|\nabla B\|_1. \quad (2)$$

By substituting  $\nabla B$  with auxiliary variable  $C_1$ , objective (2) can be rewritten as follows:

$$\begin{aligned} \underset{B, C_1}{\operatorname{argmin}} & \|D \circ \alpha + B - I\|_F^2 + \lambda_1 \|C_1\|_1, \\ \text{s.t. } & C_1 = \nabla B. \end{aligned} \quad (3)$$

To solve (3), we need to know its augmented Lagrangian equation, as shown below :

$$\begin{aligned} L(B, C_1, Y) = & \|D \circ \alpha + B - I\|_F^2 + \lambda_1 \|C_1\|_1 \\ & + \langle C_1 - \nabla B, Y_1 \rangle + \frac{\rho_1}{2} \|C_1 - \nabla B\|_F^2, \end{aligned} \quad (4)$$

where  $Y_1$  denotes a Lagrangian dual variable. At iteration  $k+1$ , the values of  $B$ ,  $C_1$ ,  $Y_1$ , and  $\rho_1$  are updated as follows:

$$\begin{aligned} B^{k+1} &= (2 + \rho_1^k \nabla^T \nabla)^{-1} (2(I - D^k \circ \alpha^k) \\ &+ \rho_1^k \nabla^T \left( C_1^k + \frac{Y_1^k}{\rho_1^k} \right)); \\ C_1^{k+1} &= f_{\frac{\lambda_1}{\rho_1^k}} \left( \nabla B^k - \frac{Y_1^k}{\rho_1^k} \right); \\ Y_1^{k+1} &= Y_1^k + \rho_1^k (C_1 - \nabla B^k); \\ \rho_1^{k+1} &= 4\rho_1^k. \end{aligned} \quad (5)$$

where  $f_\alpha(x) = \operatorname{sign}(x) \cdot \max(|x| - \alpha, 0)$  denotes the soft-thresholding calculation formula.

(2) **Updating  $D$  and  $\alpha$ :** deleting the terms unrelated to  $D$  and  $\alpha$ , the subproblem can be expressed as follows:

$$\underset{D, \alpha}{\operatorname{argmin}} \|D^k \circ \alpha^k + B - I\|_F^2 + \lambda_2 \|\alpha\|_1^k + \lambda_3 \|D^k \circ \alpha^k\|_{w, s_p}^p. \quad (6)$$

Notably, the third term denoises the image after dictionary and coefficient reconstruction (noise mainly exists in the detail layer). Thus, let  $M^k = D^k \circ \alpha^k$ ; then, the subproblem becomes

$$\begin{aligned} \underset{M}{\operatorname{argmin}} & \|M^k + B^{k+1} - I\|_F^2 + \|D^k \circ \alpha^k - M^k\|_F^2 \\ & + \lambda_2 \|\alpha\|_1^k + \lambda_3 \|M^k\|_{w, s_p}^p. \end{aligned} \quad (7)$$

At this point,  $M$ ,  $D$ , and  $\alpha$  must be updated. First, we update  $M$ . The above formula can be rewritten as follows by introducing  $M^k = [D^k \circ \alpha^k; I - B^{k+1}]$ :

$$\underset{M}{\operatorname{argmin}} \|\bar{M}^k + M^k\|_F^2 + \lambda_3 \|M^k\|_{w, s_p}^{p, K'}. \quad (8)$$

This is a standard denoising problem and can be solved according to the method provided in [31]. This process involves an internal iteration number  $K'$  such that the larger the value of  $K'$ , the cleaner the denoising.

Second, by updating  $D$  and  $\alpha$ , the subproblem at iteration  $k+1$  becomes an SR model, as follows:

$$\underset{D, \alpha}{\operatorname{argmin}} \|D^k \circ \alpha^k - M^k\|_F^2 + \lambda_2 \|\alpha\|_1^k. \quad (9)$$

The dictionary learning method proposed by Veshki et al. [33] has a shorter single-iteration time than KSVD, and can thus obtain the same sparsity level as the KSVD method in a shorter time. This method updates each column of the dictionary using the following formula:

$$[d]_t = [E]_t [\gamma_t^r]_{\omega_t}^T, \quad (10)$$

where  $[E]_t = [M - \sum_{s \neq t} [d]_s \gamma_s^r]_{\omega_t}$ ,  $\omega_t$  represents the subsets of indices where  $\gamma_t^r \neq 0$ , and  $\gamma_t^r$  is the  $t$ -th row of  $\alpha$ . Then, dictionary  $D^{k+1}$  and coefficient  $\alpha^{k+1}$  are updated using the orthogonal matching pursuit method. Compared to other sparse dictionary learning algorithms, a pre-trained fixed and complete dictionary is required for all tasks. The dictionary used in this study was updated adaptively for each input.

In other words, according to the ADMM concept, (1) is transformed into subproblems ((2) and (6), or (2), (8), and (9)),

which are the variation, denoising, and dictionary-coefficient update problems. The three subproblems can then be solved smoothly using the methods mentioned in relation to the model characteristics. Thus, the proposed HVSR decomposition model has two capabilities: denoising and decomposition. It is worth mentioning that the ADMM concept allows the aforementioned update steps to be reversed. The HVSR model stops updating when the given maximum number of iterations  $K$  is reached in this study, and then outputs the noise-free base layer  $B$ , noise-free detail layer dictionary  $D$ , and noise-free detail layer coefficient map  $\alpha$ .

### III. PROPOSED MEDICAL IMAGE FUSION AND DENOISING ALGORITHM

In view of the decomposition model above, the proposed medical image fusion and denoising algorithm performs the following three steps to obtain a fused noise-free medical image.

#### A. Decomposition

Given  $N$  inputs ( $I_i, i = 1, 2, \dots, N$ ) to be fused, each one is decomposed by the new decomposition model (HVSR) described in the previous section, as follows:

$$B_i, D_i, \alpha_i = HVSR(I_i), i = 1, 2, \dots, N, \quad (11)$$

where  $B_i$ ,  $D_i$ , and  $\alpha_i$  are the structure layer, detail layer dictionary, and detail layer coefficient of  $I_i$ , respectively.

#### B. Fusion Rules

After layer decomposition, different fusion rules are applied to fuse layers containing different feature information.

**1) Structure Layer Fusion:** Considering that the structural layer mainly contains the average energy information, to ensure that the fused structural layers retain their respective energy information, the calculation is as follows:

$$B = \sum_{i=1}^N w_i \cdot * B_i, \quad (12)$$

where  $w_i = \begin{cases} 1 & \text{if } B_i = \max(B_1, B_2, \dots, B_N), \\ 0 & \text{else.} \end{cases}$

**2) Detail Layer Fusion:** Larger coefficients correspond to more salient features. The following formula can acquire the fusion detail layer, which considers the saliency features of different inputs:

$$T = \sum_{i=1}^N W_i \cdot * D_i \circ \alpha_i, \quad (13)$$

where  $W_i = \begin{cases} 1 & \text{if } \alpha_i = \max(\alpha_1, \alpha_2, \dots, \alpha_N), \\ 0 & \text{else.} \end{cases}$

#### C. Reconstruction

The fused image can be received by the following equation:

$$F = B + T. \quad (14)$$

**Fig. 2** depicts the flow of the proposed algorithm (using PET-MRI as an example), where PET and MRI images are

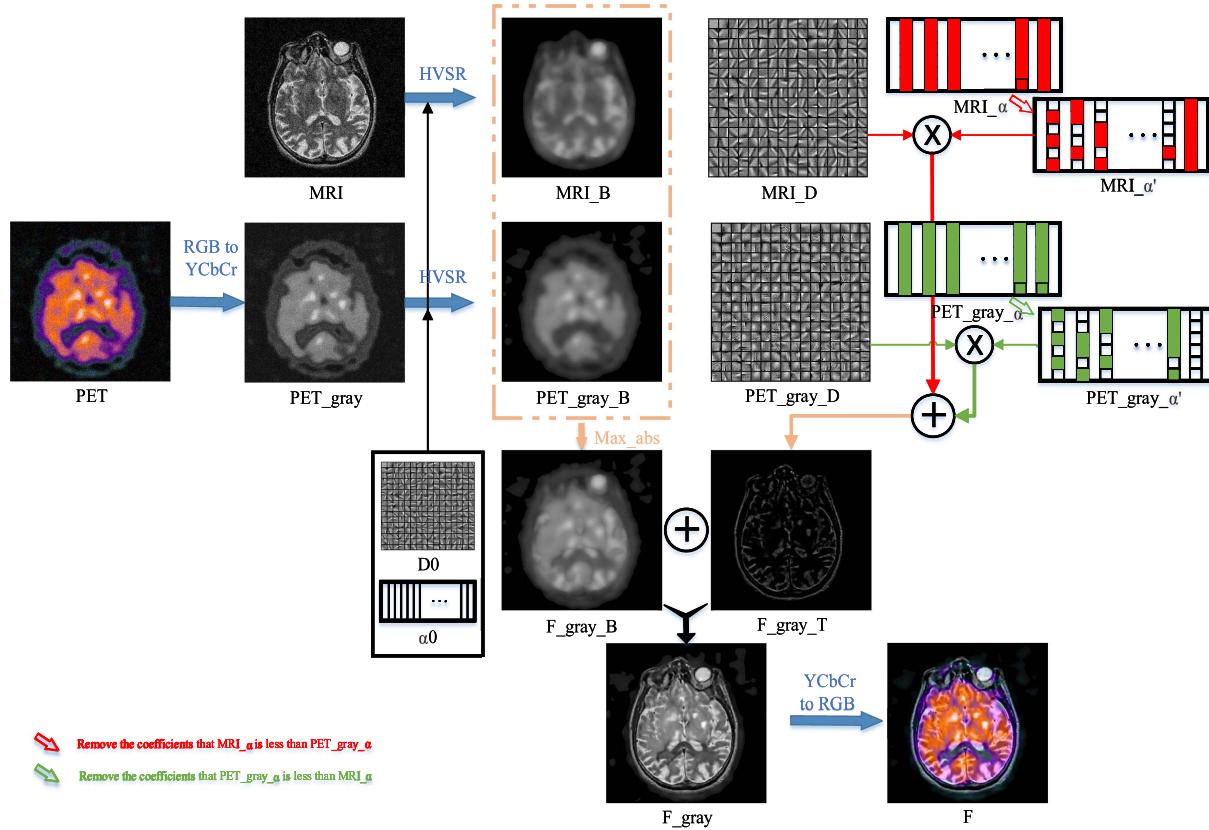
the input; PET\_gray is the gray image of PET; PET\_gray\_B and MRI\_B are the structure layers; PET\_gray\_D and MRI\_D are the detail layer dictionaries; PET\_gray\_α and MRI\_α are the detail layer coefficient maps; F\_gray\_B and F\_gray\_T are the fused gray structure layer and detail layer, respectively; and F is the fused image. In fact,  $\alpha'$  in the figure indicates that  $\alpha' = W \cdot * \alpha$ . To better understand (13), we use red-green arrows to represent this process; the red thick arrow indicates the removal of the coefficients for which MRI\_α is less than PET\_gray\_α, and the green thick arrow indicates the removal of the coefficients for which PET\_gray\_α is less than MRI\_α. The initial dictionary  $D_0$  in the figure randomly selects three images for simple initial training. The initial coefficient is set to  $\alpha_0 = 0$ . It is well known that PET is a three-channel pseudo-color image, and MRI is a single-channel gray image. To avoid losing color information, it is feasible to transform PET into a Y component containing luminance information and two color channels by using a color space conversion method. The YCbCr color space method is more suitable for computer monitors than other color space conversion methods that are suitable for TV sets. This method is widely used in medical image processing [6], [15]. “RGB to YCbCr” means the functional image (PET) is converted from the RGB space to the YCbCr space, where the Y channel is “PET\_gray,” to be fused with MRI. After processing using (11)–(14) described in Section III, the fused gray image (F\_gray) is obtained. Finally, the fused color image (F) is acquired by the inverse color space transformation, that is, “YCbCr to RGB”. The fusion between anatomical images does not require color space conversion.

### IV. EXPERIMENTAL ANALYSIS

This section is organized as follows: *A.* parameter selection of the proposed HVSR model; *B.* ablation experiment; *C.* the medical image fusion and denoising model proposed in this paper is used to fuse medical images disturbed by different levels of noise; and *D.* an extended experiment to fuse medical images without noise. Medical image fusion research lacks ground truth, and therefore, subjective evaluation is usually combined with objective evaluation to evaluate fusion performance. In this study, all medical images from the Harvard Whole Brain Atlas database [34] and eight evaluation indicators were selected to assess the fusion performance from multiple perspectives: human perception-inspired (visual information fidelity for fusion (VIFF) [35], Chen-Blum (CB) [36]), image structural similarity-based (feature similarity (FSIM) [37], structural similarity index (SSIM) [38]), image feature-based (spatial frequency (SF), standard deviation (SD)) [39], and information theory-based (quality-aware clustering (QAC) [11], peak signal-to-noise ratio (PSNR) [38]). The smaller the QAC and the larger the values of the other metrics, the better is the performance.

#### A. Parameter Selection

The proposed HVSR model contains three parameters:  $\lambda_1$  controls the smoothness of the structural layer; the larger the value, the smoother the structural layer and the lesser the noise;  $\lambda_2$  controls the sparsity of sparse coefficients; and  $\lambda_3$  determines the importance of the denoising term.  $K$  implies the number of



**Fig. 2.** Flowchart of the proposed fusion and denoising algorithm (PET-MRI,  $\sigma = 20$ ).

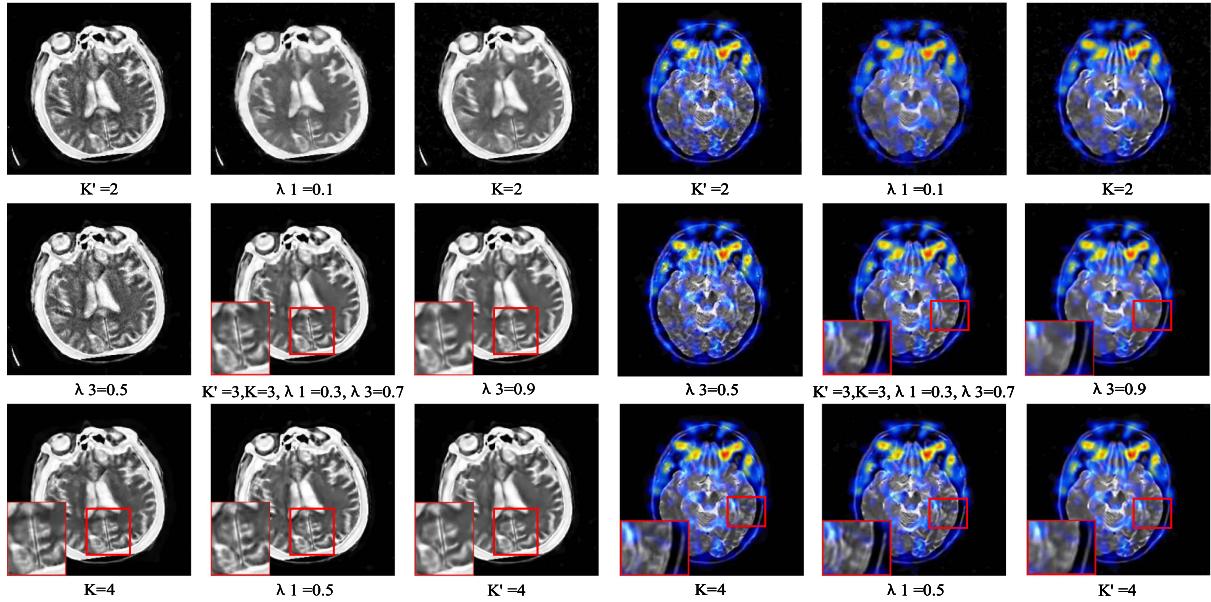
iterations in the HVSR;  $K'$  is the internal iteration number in (8). Evidently, the greater the number of iterations, the greater is the time consumption. According to the literature, it is reasonable to set  $\lambda_2$  to one; therefore, we mainly discuss the values of the other parameters. Sixteen pairs of medical images were used for parameter analysis. According to [38], [39], the indicators for evaluating the fusion algorithm can be divided into four aspects. Here, we chose one indicator for each aspect for the objective evaluation: CB, SSIM, SD, and QAC. Figs. 3 and 4 show the subjective and objective results for all parameter changes. For the convenience of investigation, we use the nine-square grid arrangement for comparison, in which the main diagonal line is used to examine the subjective fusion effect of different  $K'$ , and the sub-diagonal line is used to examine the subjective fusion effect of the number of  $K$ ; the three images in the middle in the horizontal direction are used to examine the influence of the value of  $\lambda_3$  on the fusion result, and the three images in the vertical middle are used to examine how the choice of  $\lambda_1$  affects the fusion result.

It can be observed from Fig. 3 that regardless of which parameter takes a smaller value, a good subjective visual effect cannot be obtained; the larger the  $K'$ , the cleaner the noise removal, and the larger the  $K$ , the greater the gradient edge. However, there are clear differences in time. In this part of the experiment, the average acquisition time of the central image in the nine-square grid is 161.2032 s, whereas the time is 192.1093 s when  $K' = 4$ , and 218.319 s when  $K = 4$ . Furthermore, a larger

$\lambda_1$  allows the fused image to have less noise and greater gradient details, but the artifacts increase. A larger  $\lambda_3$  results in smoother fusion results, but the gradient information is partially lost. Fig. 4 shows a two-group combined line chart of the four evaluation indicators, where each line represents the change in a parameter. When  $K' = 3$ , the best SSIM, CB, and QAC values are obtained, and a larger  $K'$  results in a smaller SD, which is coincided with the previous subjective analysis because of the characteristic information of the SD measurements. In addition, it appears that all the optimal values are obtained when  $K = 4$ ; however, the time consumption cannot be ignored. Similarly, when  $\lambda_1 = 0.3$ , the best SSIM, CB, and QAC values are obtained; however, because a larger  $\lambda_1$  permits the fused image to have greater gradient details, the SD value is not optimal. Finally, a smaller  $\lambda_3$  does not remove the noise of the fused image cleanly; therefore, the SSIM and SD are the best and QAC values are the worst. In addition, a better CB value can be obtained when  $\lambda_3 = 0.7$ . In summary, we determine  $K'$ ,  $K$ ,  $\lambda_1$ , and  $\lambda_3$  as 3, 3, 0.3, and 0.7, respectively.

### B. Ablation Experiments

In the proposed method, a VM was introduced to ensure that the structural layer is sufficiently smooth and to force the noise to remain in the detail layer. The SR better captures detailed information, and WSpNM was introduced for denoising. Ablation experiments were conducted to ascertain the effect of



**Fig. 3.** Subjective results for different parameters.

**TABLE I**  
EVALUATION INDEX MEAN FOR THE ABLATION EXPERIMENTS

Methods \ Metrics	Human perception-inspired			Image structural similarity-based		Image feature-based		Information theory-based		Average ranking
	VIFF	CB	FSIM	SSIM	SF	SD	QAC	PSNR		
without SR	0.4334	0.5852	0.7967	0.5590	27.2481	1.3426	0.4924	14.3432	2.125	
without VM	0.3781	0.3561	0.6178	0.1528	43.8502	1.1478	0.5347	13.5566	3.625	
without WSpNM	0.4179	0.4470	0.6502	0.2217	<b>45.0931</b>	1.3054	0.5469	13.6603	2.875	
Proposed	<b>0.4465</b>	<b>0.5878</b>	<b>0.8017</b>	<b>0.5627</b>	23.9190	<b>1.3471</b>	<b>0.4860</b>	<b>14.4601</b>	<b>1.375</b>	

each term on the results of medical image fusion and denoising. The experimental settings were as follows: 30 sets of data were selected from three datasets (CT-MRI, PET-MRI, and SPECT-MRI) and different levels of noise were added. “Without SR” means  $\lambda_2 = 0$ , and the new detail layer replaced  $D \circ \alpha$ . The fusion rules of the new detail layer were the same as those of the original structure layer. “Without VM” refers to  $\lambda_1 = 0$ , and the fusion rules of the new structural layer were consistent with those of the original structural layer. “Without WSpNM” implies  $\lambda_3 = 0$ , and the fusion rules remained unchanged. Except for the aforementioned changes, the other parameter settings were the same as those of the proposed algorithm.

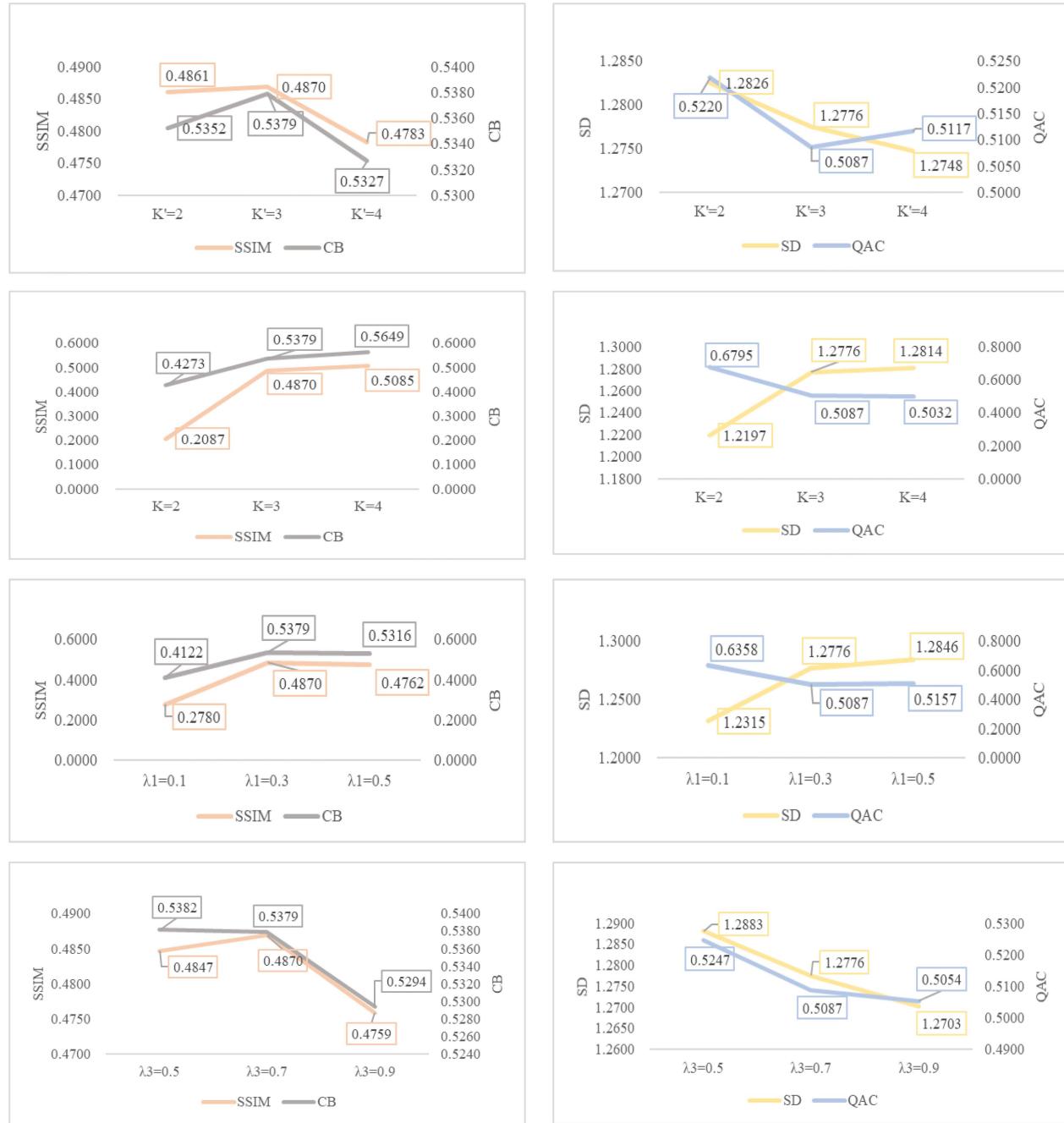
Fig. 5 presents a comparison of subjective effects. The images of the “Without SR” method show some discontinuous information; in the images of the “Without VM” method, the noise removal is unclean; in the images of the “Without WSpNM” method, the noise is not removed. The ablation experiments pinpoint that the proposed model is effective, and the absence of any constituent part leads to the fusion and denoising tasks not being completed together. If there is no variational term to constrain the base layer, the noise will not exist in the detail layer in large quantities, and the noise removal will be unclean. If there is no SR term for the detail layer characterization, the details after denoising may be lost. Furthermore, without the addition of the WSpNM model, the fusion model cannot perform noise removal.

In addition, the eight objective evaluation indices in Table I (the black font in the table displays that the metric of this method has the best value) show that the proposed algorithm can obtain the optimal averaging ranking and seven optimal means. The lowest SF value was obtained owing to the noise removal and continuous gradient details in the fused image of the proposed algorithm. Both subjective and objective experiments proved that the good results of the proposed algorithm can be attributed to a combination of the advantages of the three components.

To investigate the proposed algorithm statistically, the Nemenyi post-hoc test was performed according to the average ranking. Fig. 6 shows the results of the Nemenyi post-hoc test for the ablation experiments. The dots denote the average ranking of the algorithms, and the horizontal lines express the range of critical difference (CD) values. There is a statistical difference between algorithms when the horizontal lines of the algorithms don't overlap, otherwise, there is no significant difference. The proposed algorithm clearly had several advantages.

### C. Medical Image Fusion and Denoising

1) *Experimental Settings:* The proposed algorithm was compared subjectively and objectively with the following classical fusion and denoising algorithms: Kim, ASR, LR-SR, and TL-SR; deep learning methods: DSNGAN and EMFusion; and a



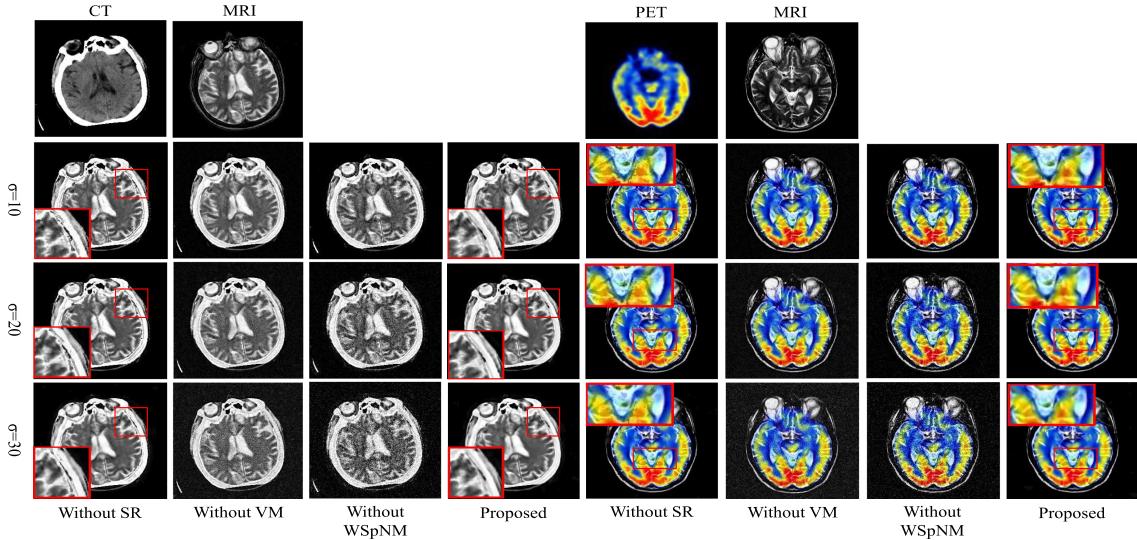
**Fig. 4.** Mean of objective metrics for different parameters.

hybrid traditional-deep learning method: MSPRN [40]. Ninety groups of medical images (CT-MRI, PET-MRI, SPECT-MRI; 30 groups each) and eight objective evaluation indicators were picked for the experimental discuss of fusion and denoising. Different levels of Gaussian noise ( $\sigma = 10$ ,  $\sigma = 20$ , and  $\sigma = 30$ ) were added to each medical image to certify the robustness of our algorithm.

**2) Subjective Evaluation:** Figs. 7–9 give the results of fusion and denoising of input images perturbed by different noise levels using the eight algorithms. The image fused using Kim's method has a loss of brightness and blurred gradient details. In the ASR method, the color distortion of the fused image is considerable. Although the fusion image of the LR-SR method has no color

distortion, the boundary of its core content is diffused; the greater the noise, the more obvious the block effect in the fusion image. The biggest problem with the fusion image of the TL-SR method is that the noise removal is not sufficiently clean; the larger the noise interference, the more obvious it becomes. Evidently, the fusion results of the DSNGAN, EMFusion, and MSPRN methods are not satisfactory under noise interference; a significant amount of information is lost, and the noise is still present. The proposed algorithm successfully avoids these problems; the noise removal is cleaner and the gradient details are clearer.

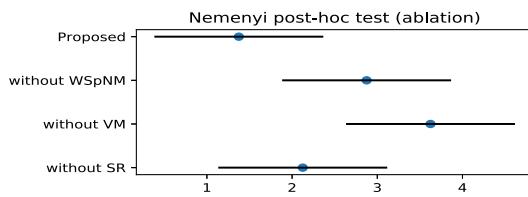
**3) Objective Evaluation:** Table II presents a comparison of the mean of the objective index of the fusion images of input images perturbed by different noise levels  $\sigma$ . The black font



**Fig. 5.** Subjective results of ablation experiments.

**TABLE II**  
INDEX MEAN ON 90 DATASETS WITH DIFFERENT NOISE LEVELS USING EIGHT ALGORITHMS

Noise level	Methods	Metrics		Human perception-inspired		Image structural similarity-based		Image feature-based		Information theory-based		Average ranking
		VIFF	CB	FSIM	SSIM	SF	SD	QAC	PSNR			
$\sigma=10$	Kim	0.3316	0.3181	0.7662	0.1475	12.6435	1.2537	0.6174	16.8356	6.375		
	ASR	0.3852	0.3308	0.7935	0.2158	23.9084	1.3476	0.5439	16.5923	4.125		
	LR-SR	0.5083	0.4142	0.7685	0.2579	18.4005	1.3463	0.6304	14.3814	4.625		
	TL-SR	0.4910	0.3258	0.8032	0.2236	24.1972	1.2968	0.6544	16.0520	4.375		
	DSNGAN	0.3722	0.4287	0.7547	0.3213	20.8609	1.2692	0.6995	16.1037	5		
	EMFusion	0.3672	0.4244	0.7686	0.2223	21.7965	1.3504	0.7825	<b>18.1357</b>	4.375		
	MSPRN	0.4697	0.3624	0.7485	0.2289	24.6090	1.2791	0.8005	15.7653	5.5		
	Proposed	<b>0.5368</b>	<b>0.6524</b>	<b>0.8167</b>	<b>0.6244</b>	<b>25.2087</b>	<b>1.4609</b>	<b>0.4701</b>	15.7668	<b>1.625</b>		
$\sigma=20$	Kim	0.2912	0.2815	0.7324	0.1222	7.6858	1.0877	0.7022	16.4878	6.375		
	ASR	0.3622	0.2988	0.7818	0.1890	20.2843	1.2148	0.6006	16.2334	4.125		
	LR-SR	0.4968	0.3697	0.7345	0.1766	18.3715	1.2671	0.7392	14.1750	5		
	TL-SR	0.4625	0.3053	0.7727	0.2013	22.4004	1.1876	0.7526	15.7662	4.75		
	DSNGAN	0.3585	0.4267	0.7159	0.2723	24.4895	1.2675	0.6898	15.9725	3.5		
	EMFusion	0.3401	0.3503	0.6849	0.1837	24.7455	1.1925	0.7633	<b>17.2540</b>	4.875		
	MSPRN	0.4331	0.3242	0.6553	0.1784	<b>29.6242</b>	1.1623	0.7141	15.2169	5.375		
	Proposed	<b>0.5115</b>	<b>0.6344</b>	<b>0.8156</b>	<b>0.5969</b>	23.1543	<b>1.4495</b>	<b>0.4650</b>	15.6356	2		
$\sigma=30$	Kim	0.2769	0.2686	0.7163	0.1137	7.4061	0.9739	0.7317	16.1700	6.625		
	ASR	0.3302	0.2790	0.7740	0.1661	17.3164	1.1025	0.6327	16.0109	4.375		
	LR-SR	0.4386	0.2929	0.7192	0.1399	14.2093	1.1254	0.7403	13.9323	5.375		
	TL-SR	0.4335	0.2881	0.7470	0.1826	20.8779	1.0881	0.7743	15.4140	5		
	DSNGAN	0.3404	0.3962	0.6760	0.1913	29.6079	1.2595	0.6936	15.7538	3.375		
	EMFusion	0.3124	0.3290	0.6115	0.1566	29.0839	1.0884	0.7065	<b>16.3364</b>	4.5		
	MSPRN	0.3907	0.2988	0.5731	0.1431	<b>38.2594</b>	1.0494	0.5239	14.3954	4.875		
	Proposed	<b>0.4798</b>	<b>0.5543</b>	<b>0.8004</b>	<b>0.5139</b>	21.1700	<b>1.4232</b>	<b>0.4913</b>	15.5272	<b>1.875</b>		

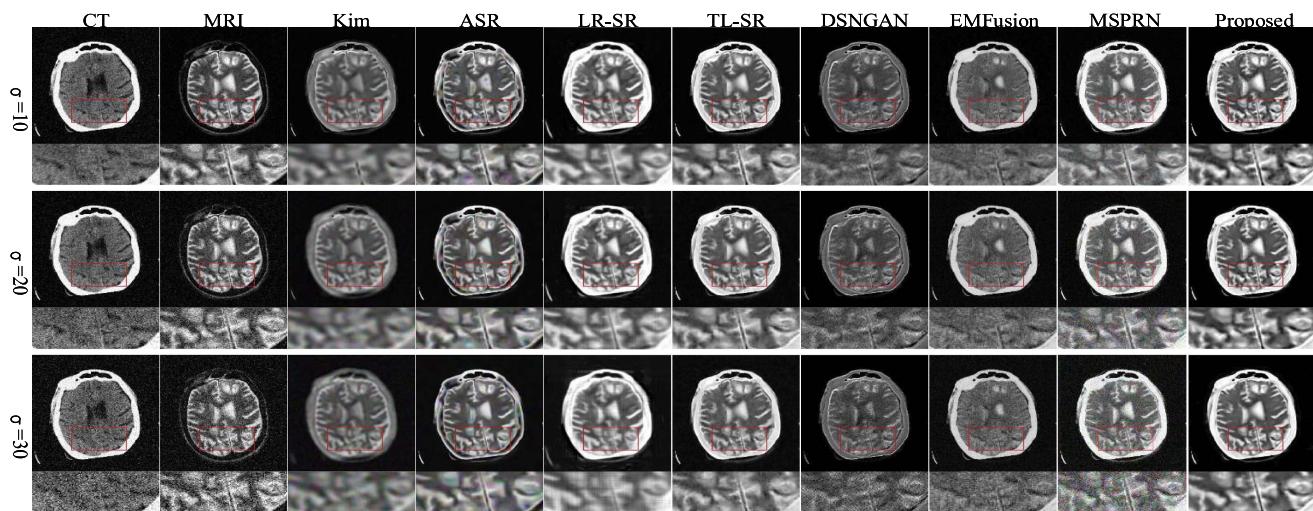


**Fig. 6.** Results of the Nemenyi post-hoc test for images.

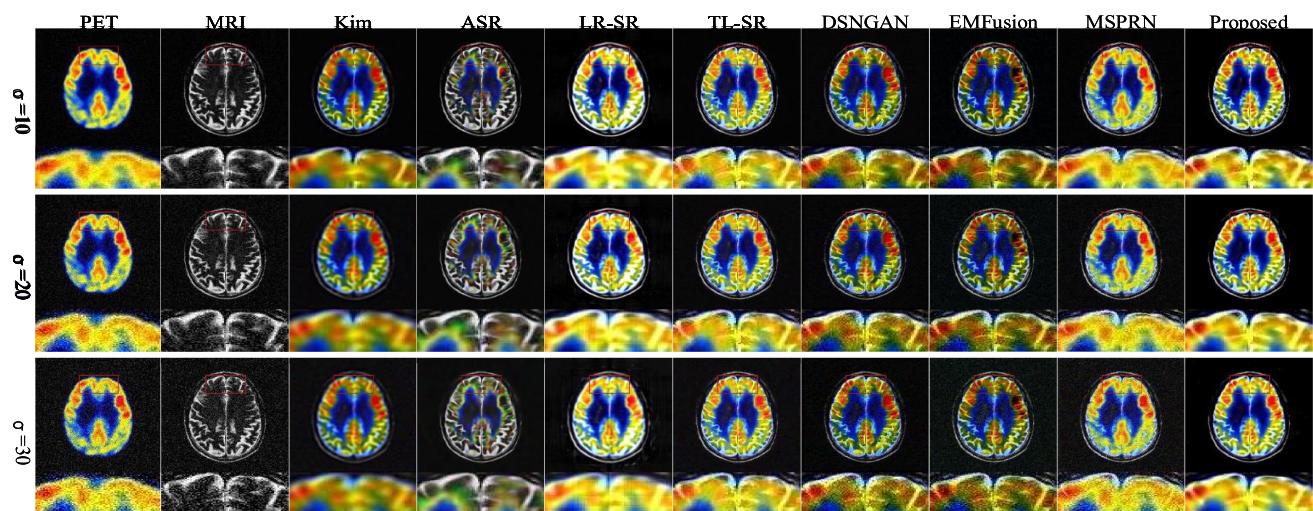
in the table says that the index of this method has the optimal value, and the proposed algorithm obtains the optimal value for multiple indicators. However, the proposed algorithm does not perform as well as other algorithms in terms of the PSNR, mainly because the proposed algorithm removes noise and has a larger

MSE; therefore, this index is not high. In addition, when the noise level increases, the SF value of the proposed algorithm is not as high as those of the DSNGAN, EMFusion, and MSPRN methods. This is mainly because these three methods do not remove noise, and thus the SF value is higher. As the indicators measure the superiority of the algorithm from different perspectives, the average/total ranking best reflects its comprehensive value. The last column in the table shows the mean ranking of all index values of each algorithm; the proposed algorithm achieves the optimal average ranking.

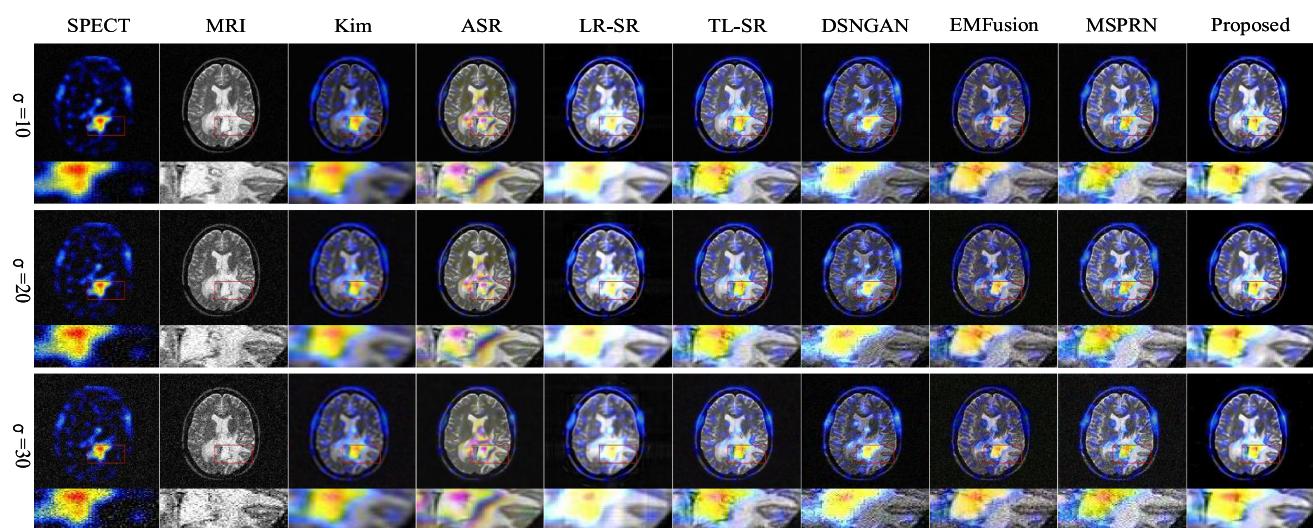
**Fig. 10** shows the results of the Nemenyi post-hoc test for images with noise. The proposed algorithm has the highest average ranking and significantly outperforms the MSPRN, LR-SR, and Kim algorithms at different noise levels. The working



**Fig. 7.** Image fusion and denoising using eight algorithms for CT-MRI with different noise levels.



**Fig. 8.** Image fusion and denoising using eight algorithms for PET-MRI with different noise levels.



**Fig. 9.** Image fusion and denoising using eight algorithms for SPECT-MRI with different noise levels.

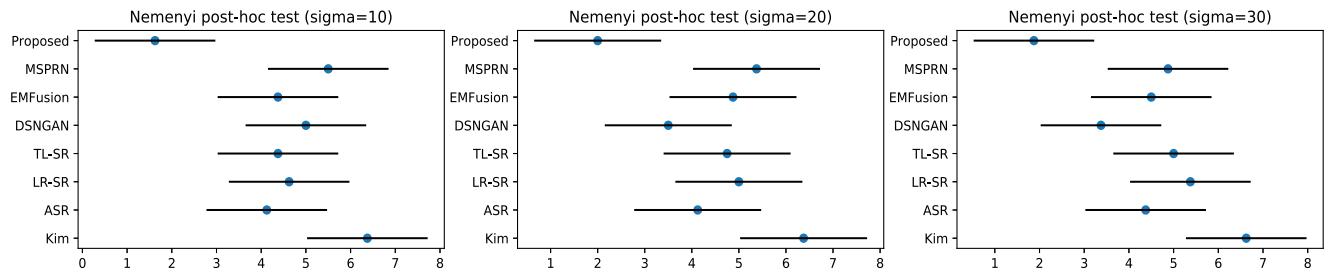


Fig. 10. Results of the Nemenyi post-hoc test for images with noise.

TABLE III  
TIMES OF DIFFERENT FUSION METHODS FOR IMAGES WITH NOISE (UNIT: S)

Methods	Kim	ASR	LR-SR	TL-SR	DSNGAN	EMFusion	MSPRN	Proposed	
$\sigma = 10$	27.2003	61.3501	518.2192	4.2956	0.7184	<b>0.1899</b>	1.8165	140.4122	
Time	$\sigma = 20$	26.9849	54.4092	557.2718	3.5522	0.7288	<b>0.1870</b>	1.1916	140.4429
	$\sigma = 30$	26.9114	55.0461	504.0804	3.3875	0.7258	<b>0.1953</b>	1.1875	174.1100

times of the different fusion methods for noisy images are displayed in Table III. The proposed algorithm only outperforms the LR-SR algorithm in terms of the running time. This issue will be addressed in future studies. Even though the advantage of the running time of the proposed algorithm is not obvious, our subjective and objective results are the best, and the average ranking is significantly higher than those of the other methods.

#### D. Medical Image Fusion Without Noise

1) *Experimental Settings:* Without considering the noise perturbation of the input, the proposed algorithm was compared subjectively and objectively with the following nine classical fusion methods: ASR, LR-SR, TL-SR, and CNPS (traditional algorithms); DSNGAN, EMFusion, FusionDN, and IFCNN (deep learning methods); and MSPRN (a hybrid traditional-deep learning method), most of which have been published in the past two years. A total of 180 sets of medical images (CT-MRI, PET-MRI, SPECT-MRI; 60 groups each) were used for the experiments.

2) *Subjective Evaluation:* Fig. 11 presents three examples of subjective evaluation of the proposed algorithm and other classical fusion algorithms. The color, brightness, edge details, complete information, and noise (shadows) of the fused image are the main bases for the subjective evaluation of quality. In the comparison chart of the fused CT-MRI images, TL-SR, EMFusion, and the proposed algorithm performed better in all aspects. The other methods have different degrees of information loss. In the comparison chart of the fused PET-MRI and SPECT-MRI images, LR-SR, TL-SR, CNPS, and the proposed algorithm performed optimally. In summary, TL-SR and the proposed algorithm performed stably, but the details of the proposed algorithm were sharper, and the texture clearer. Thus, the proposed algorithm preserves the color, brightness, contrast, and gradient information of the fused images.

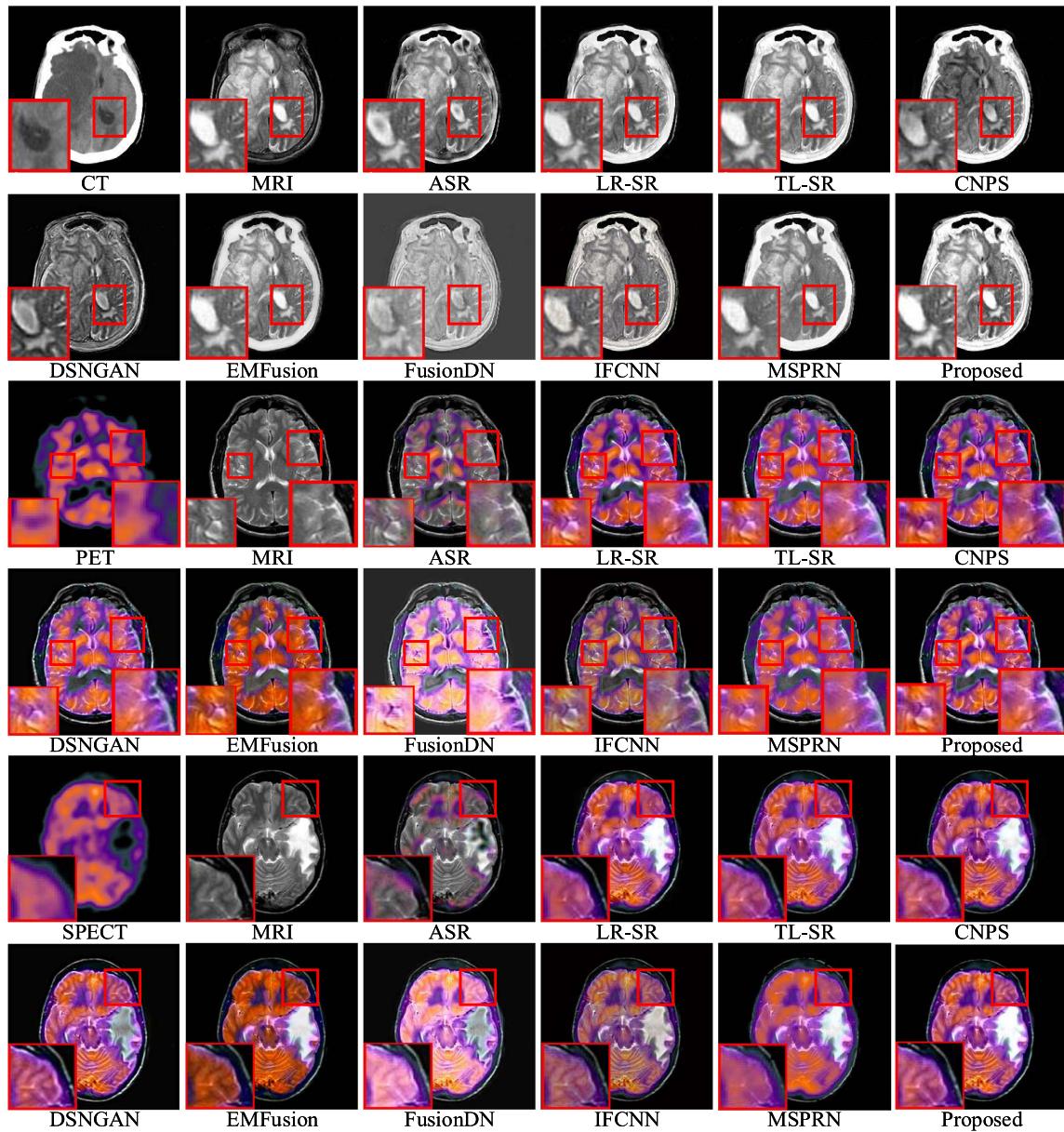
3) *Objective Evaluation:* Similarly, the eight evaluation indicators mentioned above were used for the evaluation, as shown in Table IV. It can be observed that the proposed algorithm obtains the best VIFF, SF, and QAC values, and sub-optimal CB and SD values. Although it did not obtain relatively good rankings for other indicators, the average ranking of the proposed algorithm was the best.

The running times of ten fusion methods for images without noise are listed in Table V. Although the deep learning method can obtain a fast test speed, the computation time of model training is long; the trained model displays poor adaptability to the fused images, and the fused images in these experiments have different degrees of information loss. The proposed algorithm outperformed the ASR and LR-SR algorithms in terms of noise-free image fusion.

#### V. CONCLUSION

A new decomposition model, HVSR, was proposed by combining VM and SR models to exploit their respective advantages. In addition, the WSpNM regularization term was introduced for denoising. With HVSR, any input with noise interference can be decomposed while removing the noise. Subsequently, a novel medical image fusion and denoising algorithm was proposed based on the HVSR decomposition model. After obtaining the structure layer and dictionary coefficients of the images to be fused, the larger coefficient principle was used to fuse each layer. Finally, the performance of the proposed algorithm under different noise levels is verified using numerous medical images. The proposed algorithm obtained optimal objective evaluation index and fusion results, such as clear gradients. The proposed fusion model can also be used on input without noise interference. Subjective qualitative evaluation and objective quantitative assessment demonstrated that the proposed algorithm is competitive with respect to other classical traditional and deep learning methods.

It is worth noting that the existing image fusion denoising algorithms (References 22-29 and the proposed) only consider medical images disturbed by Gaussian noise, but the actual noise distribution of medical images may be more than that. A review of medical image denoising shows that additive Gaussian noise is actually one of the noise types in medical images [41]. Therefore, in future work, the actual noise in medical images of different modalities can be considered and resolved.



**Fig. 11.** Medical image fusion using 10 algorithms.

**TABLE IV**  
AVERAGE EVALUATION INDEX OF DIFFERENT FUSION METHODS WITHOUT NOISE

Metrics \ Methods	VIFF	Human perception-inspired CB	Image structural similarity-based FSIM	Image feature-based SSIM	SF	SD	QAC	Information theory-based PSNR	Average ranking
ASR	0.4366	<b>0.6767</b>	0.8119	<b>0.6503</b>	24.8052	<b>1.4585</b>	0.5151	<b>18.7083</b>	4
LR-SR	0.5343	0.6122	0.8198	0.5732	24.8847	1.4001	0.5266	16.6122	5.5
TL-SR	0.5425	0.6389	0.8264	0.6357	24.4112	1.3759	0.5155	16.9057	5
CNPS	0.5538	0.6252	0.8172	0.6115	24.9122	1.3974	0.5240	15.4450	5.25
DSGAN	0.4384	0.3228	0.7538	0.2074	24.8110	1.1579	0.5657	15.2384	8.625
EMFusion	0.4440	0.6545	<b>0.8315</b>	0.6317	20.7988	1.4174	0.4897	18.5865	4.25
FusionDN	0.4804	0.1780	0.7706	0.1814	24.7993	0.5993	0.5330	11.1999	8.875
IFCNN	0.5043	0.6500	0.8130	0.2893	24.5072	1.4163	0.5072	17.3023	5.375
MSPRN	0.5095	0.6172	0.8287	0.6420	23.1767	1.3646	0.5039	16.3418	5.375
Proposed	<b>0.5873</b>	0.6557	0.8135	0.6392	<b>25.3436</b>	1.4363	<b>0.4823</b>	16.3439	<b>2.75</b>

**TABLE V**  
TIMES OF DIFFERENT FUSION METHODS FOR IMAGES WITHOUT NOISE (UNIT: S)

Methods	ASR	LR-SR	TL-SR	CNPS	DSGAN	EMFusion	FusionDN	IFCNN	MSPRN	Proposed
Time	51.8811	413.8560	17.4051	3.7180	0.0319	0.0898	1.2483	<b>0.0175</b>	1.6556	20.6143

## REFERENCES

- [1] Y.-D. Zhang et al., "Advances in multimodal data fusion in neuroimaging: Overview, challenges, and novel orientation," *Inf. Fusion*, vol. 64, pp. 149–187, 2020.
- [2] S. Wang et al., "Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects," *Inf. Fusion*, vol. 76, pp. 376–421, 2021.
- [3] J. Du, W. Li, and H. Tan, "Three-layer image representation by an enhanced illumination-based image fusion method," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 4, pp. 1169–1179, Apr. 2020.
- [4] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.
- [5] Y. Yang, J. Wu, S. Huang, Y. Fang, P. Lin, and Y. Que, "Multimodal medical image fusion based on fuzzy discrimination with structural patch decomposition," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 4, pp. 1647–1660, Jul. 2019.
- [6] M. Yin, X. Liu, Y. Liu, and X. Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 1, pp. 49–64, Jan. 2019.
- [7] 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.
- [8] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, pp. 147–164, 2015.
- [9] 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, 2018.
- [10] J. Du, W. Li, and B. Xiao, "Anatomical-functional image fusion by information of interest in local Laplacian filtering domain," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5855–5866, Dec. 2017.
- [11] B. Li et al., "Medical image fusion method based on coupled neural P systems in nonsubsampled shearlet transform domain," *Int. J. Neural Syst.*, vol. 31, no. 1, 2021, Art. no. 2050050.
- [12] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36, pp. 191–207, 2017.
- [13] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "IFCNN: A general image fusion framework based on convolutional neural network," *Inf. Fusion*, vol. 54, pp. 99–118, 2020.
- [14] K. Wang, M. Zheng, H. Wei, G. Qi, and Y. Li, "Multi-modality medical image fusion using convolutional neural network and contrast pyramid," *Sensors*, vol. 20, no. 8, 2020, Art. no. 2169.
- [15] H. Xu and J. Ma, "EMFusion: An unsupervised enhanced medical image fusion network," *Inf. Fusion*, vol. 76, pp. 177–186, 2021.
- [16] H. Xu, J. Ma, Z. Le, J. Jiang, and X. Guo, "Fusiondn: A unified densely connected network for image fusion," in *Proc. 34th AAAI Conf. Artif. Intell.*, New York, NY, USA, 2020, pp. 12484–12491.
- [17] J. Fu, W. Li, J. Du, and L. Xu, "DSAGAN: A generative adversarial network based on dual-stream attention mechanism for anatomical and functional image fusion," *Inform. Sci.*, vol. 576, no. 9, pp. 484–506, 2021.
- [18] Q. Zuo, J. Zhang, and Y. Yang, "DMC-fusion: Deep multi-cascade fusion with classifier-based feature synthesis for medical multi-modal images," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 9, pp. 3438–3449, Sep. 2021.
- [19] L. Liu, L. Xu, and H. Fang, "Infrared and visible image fusion and denoising via  $l_2-l_p$  norm minimization," *Signal Process.*, vol. 172, 2020, Art. no. 107546.
- [20] H. Li, Y. Chai, H. Yin, and G. Liu, "Multifocus image fusion and denoising scheme based on homogeneity similarity," *Opt. Commun.*, vol. 285, no. 2, pp. 91–100, 2012.
- [21] J.-J. Mei, Y. Dong, and T.-Z. Huang, "Simultaneous image fusion and denoising by using fractional-order gradient information," *J. Comput. Appl. Math.*, vol. 351, pp. 212–227, 2019.
- [22] S. Goyal, V. Singh, A. Rani, and N. Yadav, "Multimodal image fusion and denoising in NSCT domain using CNN and FOTGV," *Biomed. Signal Process. Control.*, vol. 71, 2022, Art. no. 103214.
- [23] W. Zhao and H. Lu, "Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2283–2294, Sep. 2017.
- [24] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3450–3459, Dec. 2012.
- [25] L. Yu and W. Zengfu, "Simultaneous image fusion and denoising with adaptive sparse representation," *IET Image Process.*, vol. 9, no. 5, pp. 347–357, 2015.
- [26] 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, pp. 130–146, 2018.
- [27] X. Li, F. Zhou, and H. Tan, "Joint image fusion and denoising via three-layer decomposition and sparse representation," *Knowl.-Based Syst.*, vol. 224, 2021, Art. no. 107087.
- [28] M. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Inf. Fusion*, vol. 27, pp. 198–214, 2016.
- [29] G. Qi, G. Hu, N. Mazur, H. Liang, and M. Haner, "A novel multi-modality image simultaneous denoising and fusion method based on sparse representation," *Comput.*, vol. 10, no. 10, 2021, Art. no. 129.
- [30] Z. Liang, J. Xu, D. Zhang, Z. Cao, and L. Zhang, "A hybrid  $l_1-l_0$  layer decomposition model for tone mapping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, 2018, pp. 4758–4766.
- [31] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, "Weighted schatten P-norm minimization for image denoising and background subtraction," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4842–4857, Oct. 2016.
- [32] S. P. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
- [33] F. G. Veshki and S. A. Vorobyov, "A fast dictionary learning method for coupled feature space learning," *IEEE Signal Process. Lett.*, vol. 26, no. 10, pp. 1441–1445, Oct. 2019.
- [34] K. A. Johnson and J. A. Becker, "The whole brain atlas," [Online]. Available: <http://www.med.harvard.edu/aanlib>
- [35] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. Fusion*, vol. 14, no. 2, pp. 127–135, 2013.
- [36] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, 2009.
- [37] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [38] X. Zhang, "Deep learning-based multi-focus image fusion: A survey and a comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 9, pp. 4819–4838, Sep. 2022.
- [39] X. Zhang, "Benchmarking and comparing multi-exposure image fusion algorithms," *Inf. Fusion*, vol. 74, pp. 111–131, 2021.
- [40] J. Fu, W. Li, J. Du, and Y. Huang, "A multiscale residual pyramid attention network for medical image fusion," *Biomed. Signal Process. Control.*, vol. 66, 2021, Art. no. 102488.
- [41] S. V. M. Sagheer and S. N. George, "A review on medical image denoising algorithms," *Biomed. Signal Process. Control.*, vol. 61, 2020, Art. no. 102036.