FCFusion: Fractal Component-wise Modeling with Group Sparsity for Medical Image Fusion

Guoxia Xu, *Member, IEEE,* Xiaoxue Deng, Xiaokang Zhou, *Member, IEEE,* Marius Pedersen, *Member, IEEE,* Lucia Cimmino, Hao Wang, *Senior Member, IEEE,*

***Abstract*—Multimodal image fusion is the process of combing relevant biological information that can be used for automated industrial application. In this paper, we present a novel framework combining fractal constraint with group sparsity to achieve the optimal fusion quality. Firstly, we adopt the idea of patch division and component-wise separation to perceive the fractal characteristics across multi-modality sources. Then, to preserve the spatial information against the redundancy of component-entanglement, the group sparsity is proposed. A dual variable weighting rule is inherently embedded to mitigate the overfitting across the component penalty. Furthermore, the Alternating Direction Method of Multipliers (ADMM) is conducted to the proposed model optimization. The experiments show that our model has a better performance in quantitative visual quality and qualitative evaluation analysis. Finally, a real segmentation application of PET/CT image fusion proves the effectiveness of our algorithm.**

***Index Terms*—medical image fusion, fractal component-wise, structural patch prior, group sparsity, ADMM algorithm**

I. Introduction

IN the fast development medical image devices, secure of data storage of medical image becomes crucial for efficient real industrial application like biometric verification or clinical diagnosis. Sensor fusion today [1], [2] has developed rapidly for verification and treatments. Fusion for fingerprint verification [3], biometric-based efficient medical image watermarking [4], these schemes have applied to real application. Medical image and biometric are inseparable, like magnetic resonance imaging (MRI) for biological fingerprint in patient verification [5]. However, how to preserve the important biological information efficiently has becoming a main challenge. Single modal source provides limited information and cannot meet the requirements of patient verification, disease diagnosis, health monitoring, surgery, and radiation therapy [6]. In Industrial 4.0 era [7], [8], keeping the data with secure and enhancing its usage in operative environment are both important. Heterogeneous data [9] has also attracted a lot attention in many real application. It is necessary to fuse different sources into one image to obtain complementary information to assist aided-diagnosis frameworks [10], [11]. With the growing appeal of image fusion, various fusion algorithms have been developed [12]. The goal of image fusion is to leverage dominated features from multi-modal sources and synthesize the *common-unique* content together.

Currently, a multimodal biometric system is inevitably based on the fusion techniques to solve the entanglement of multiple biometric traits, to make the recognition more robust and distinctive. Like biometric technologies in MRI, it can be treated as the verification and authentication subsystem. However it is hard to ensure the completeness of information about the dynamics of the state of the whole body. It is imperative to complete the accuracy depended on the multimodal source. To deal with the multi-modal information interactions, various strategies have developed with success from many perspectives [13], [14]. From the general description of image fusion development, the feature extraction strategy and feature fusion rule are the main processing alternatives. We further extend the direction of image fusion from our summarized view: The insights of pixel-level image fusion are obtained from ***feature extraction*** and ***redundancy removal***.

In recent years, some works have gradually sought decomposition methods for feature extraction, but they ignored the explicit redundancy removal of the whole framework. The neural network driven medical image fusion methods have become a popular research subject. Liu et al. [15] used convolutional neural network (CNN) to perform medical image fusion by calculating activity level measurement through weight map distribution. However, the overfitting problem would produce the undesirable artifacts in the fusion result with respect to multi-scale representation. Recently, Xia et al. [16] proposed a CNN fusion method that removed the deep stacking of subsampling layers, which could obtain output results of the same size as the input. However, since the input of the training network is the entire image, the inherent local similarity of the image may be ignored. Hermessi et al. [17] leveraged the multi-stream CNN to learn the similarity between MRI/CT patches for image fusion in the proposed shearlet domain, but its fusion ability required a large amount of training data to ensure. Furthermore, different strategies of network design over relevant feature extraction emerged later for medical image fusion like generative adversarial network. Although their algorithm has achieved good performance, it is still not flexible enough. Because, the selection of depth features relies on manual design rules to achieve superior "feature extraction". The deficiency of paired and large scale datasets setbacks the improvement of medical image fusion and is hard

Guoxia Xu, Marius Pedersen and Hao Wang are with Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjøvik, Norway. Xiaoxue Deng is with Jiangsu Province Key Lab on Image Processing and Image Communication, Nanjing University of Posts and Telecommunications, Nanjing 210003, China. Xiaokang Zhou is with the Faculty of Data Science, Shiga University, Hikone, Japan, and also with the RIKEN Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan. Lucia Cimmino is with Department of Computer Science, University of Salerno, Italy. *(Corresponding author: Xiaokang Zhou and Hao Wang) (E-mail: zhou@biwako.shiga-u.ac.jp and hawa@ntnu.no)*

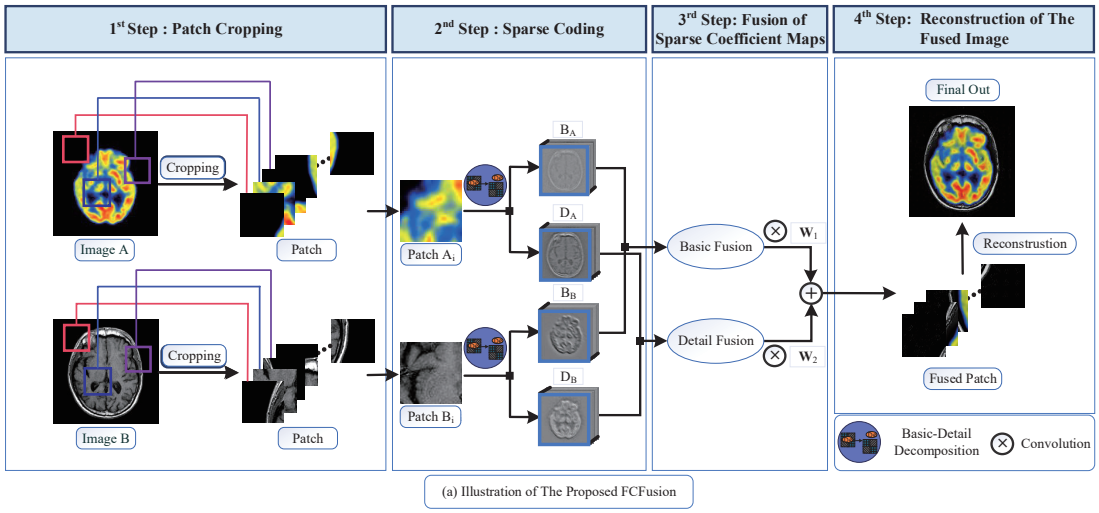


Fig. 1. The image fusion process of our FCFusion model. It includes four steps: patch cropping, sparse coding, fusion of sparse cofficient maps and reconstruction. In step 2, the sourceimages are divided into two components: basic B and detail D, and then process them in the parallel way.

to meet the requirement of Industrial 4.0.

As discussed earlier, feature extraction driven methods have demonstrated that this explicit fusion guideline (extensions of feature extraction capacity only) leads to several issues for medical image fusion. The major problems of these algorithms are that the information is over-completed. Thus, adapting the appropriate balance between feature extraction and redundancy removal is the main solution. In this paper, the fusion quality is determined by the feature extraction and redundancy removal in our proposed fractal component-wise prior and group sparsity model termed as FCFusion. The main flowchart algorithm is shown in Fig. 1. In general, we design the patch division with component-wise separation to perceive the fractal characteristics across the different components in multi-modality. Different from the traditional patch sparse representation (SR) based image fusion [18], [19] in sliding windows manner, they performed feature extraction directly for each patch. Our motivation relies on the proposed fractal constraint for feature extraction. To keep up with the redundancy removal for mitigating the over-smoothing problem, preserving the characteristic information by a group sparsity model is exploited in our proposed model. Unlike the model proposed in [20], to better promote the detail preservation and remove redundancy, we use a fractal variable weighting coefficient strategy to select the features of each patch over the decomposed components. The saliency is reflected from the group weighted sparse coefficient here for medical image fusion to achieve few artifacts. Overall, our designed patch-level component feature extraction and group sparsity mainly focus on how to avoid the over-smoothing from noise interference, color distortion, and artifacts.

The main contributions of our FCFusion model include:

- A new medical image fusion algorithm based on ”fractals” is proposed, which intuitively imposes the patch-level component-wise separation to perceive the fractal characteristic across the different components in multi-modality sources.
- A new strategy of group sparsity for components is proposed to strengthen the detail preservation for medical image fusion, and the dual variable weighting is utilized to mitigate over-smoothing and remove redundancy for characterizing detailed structure and fine components.

The rest of this paper is organized as follows. In Section II, some related work and the motivation of this work are presented. Section III describes the proposed fusion methods in detail. The experimental results and discussions are provided in Section IV. Finally, Section V concludes the paper.

II. Related Work

A. Feature Extraction

No matter in the traditional spatial domain or transform domain methods, which totally design the main two processes: image decomposition and reconstruction. Based on the above summarized points, the multi-scale transformation or multi-scale geometric analysis are conducted from the feature extraction view, as do several transformation-based and pyramid-structured methods [21]. For example, discrete wavelet transformation (DWT) is able to separate high frequency to low frequency information through image decomposition. These decomposition methods suffer from the directional feature distortion and shift invariance property. The non-subsampling paradigm is a popular representative modeling example to solve degeneration of multiple scale features, like non-subsampling contour transformation (NSCT) and non-subsampling shearlet transformation (NSST) [22]. Hybrid schemes are investigated from the perspective of enhancing the “feature extraction” like different extensions of pulse coupled neural network [23].

B. Redundancy Removal

Over the years, the sparse representation (SR) has attracted a lot attention from the natural sparsity of signals in medical image fusion. Thus, the SR with multi-scale transformation [24] and SR with pulse coupled neural network [25] were constructed for mitigating the fixed feature extraction capacity. Jiang et al. [26] proposed a novel multi-component SR-based fusion method via morphological component analysis (MCA) [27], which can obtain the sparse representations of cartoon and texture components of each source image. Sadly, it would bring a significant amount of noise in. The CS-MCA model

[28] integrated the advantages of MCA and convolutional SR (CSR), achieving multi-component and global sparse representation of the source image. However, the over-smoothed issues inevitably would lead to color distortion and weaken the root of SR based medical image fusion. Our analysis indicates that the feature extraction with redundancy removal for the preservation of fine details still remains a critical challenge.

III. The Proposed FCFusion Method

A. Problem Statement of Fractal Component-wise Modeling

A component separation model can simultaneously estimate the extracted components: basic B and detail D of the input image I, which is conducted with a consistent CSR based optimization framework.

$$\min_{D,B} \frac{1}{2} \|I - D * W_1 - B * W_2\|_2^2 + \alpha \|D\|_1 + \beta \|B\|_1 \quad (1)$$

The component separation model uses convolutional sparse coding to achieve the fusion of multiple images, where the W_1 and W_2 denote two sets of given dictionary filters. The D and B are the corresponding sparse coefficient maps. α and β are coefficients that control the degree of significance of the different terms. Unlike the standard SR model which is based on multiple overlapping patches, many component separation based methods [29], [30] have been proposed to grasp multi-component representations of source images. In this paper, we uphold the two principles of *feature extraction* and *redundancy removal*. Motivated by [31], we extend the model in Eq. (1) by learning a decoupled P to preserve the convolutional nature of the problem without other circular boundary conditions.

$$\begin{aligned} &\min_{P,D,B} \frac{1}{2} f_1\{P, D, B\} + f_2\{D, B\} + f_3\{P\}, with \\ &f_1 = \|I - \langle P, D\rangle * W_1 - \langle P, B\rangle * W_2\|_2^2 \\ &f_2 = \alpha_n \|D\|_1 + \beta_n \|B\|_1, f_3 = \lambda_n \|P\|_F \end{aligned} \quad (2)$$

For simplicity, we take $D \in \mathcal{R}^{E\times N\times N}$ as an example. $W_1 \in \mathcal{R}^{M\times N\times N}$ means an over-complete dictionary and M is the number of filters. Matrix $P \in \mathcal{R}^{M\times E}$ is the factorized convolution operator, which acts as a projection matrix from the sparse expression to the original image information. Moreover, E is the number of sparse matrices that can effectively express the image after dictionary learning. $I \in \mathcal{R}^{N\times N}$ denotes the whole image. $\langle P, D\rangle$ can be regarded as a matrix multiplication in different dimensions. And the three-dimensional matrix elements in $\langle P, D\rangle_{(m,i,j)}$ are expressed as follows:

$$\langle P, D\rangle_{(m,i,j)} = \sum_e P_{(m,e)} D_{(e,i,j)} \quad (3)$$

According to the operation rule in Eq. (3), we have

$$\langle P, D\rangle * W = \langle P^T, W\rangle * D \quad (4)$$

To mitigate the over-smoothing issues of the CSR based method, we leverage the fractal analysis to further control the amplified noise. Then, we construct the fractal constraint over the components to realize better image structure information. Eq. (6) is the minimization problem of our fusion model:

$$\begin{aligned} &\min_{P,D,B} \frac{1}{2} f_1\{P, D, B\} + f_2\{P, D, B\} + f_3\{D, B\} + f_4\{P\}, with \\ &f_1 = \|I - \langle P, D\rangle * W_1 - \langle P, B\rangle * W_2\|_2^2 \\ &f_2 = \alpha\{\|\langle P, D\rangle * W_1\|_{fc} + \|\langle P, B\rangle * W_2\|_{fc}\} \\ &f_3 = \beta\{\|D\|_1 + \|B\|_1\}, f_4 = \lambda_n \|P\|_F \end{aligned} \quad (5)$$

Our proposed fractal constraint defines the range of spatial effects and it also essentially confines which pixels would be used to constraint the components. In terms of the spatial effects, the more clear-cut distinction between base and detail would lead to better fusion quality. Instead of using a small local window, the fractal constraint considers a larger window with size $l \times l$. To address the noise variance for accurate noise characterization, we use the group sparsity [32] model to retain feature information of the sparse coefficient map. Further, to leverage the global weighted mechanism, the matrix P is refined into P_1 and P_2. Θ_1 and Θ_2 are fractal structure factors used to better promote the group sparsity. The global group variable weighting strategy is exploited in our model:

$$\begin{aligned} &\min_{P,D,B,\Theta,\nabla} f_1\{P,D,B\} + f_2\{P,D,B\} + f_3\{\Theta, \nabla, D, B\} + f_4\{P\}, with \\ &f_1 = \frac{1}{2} \|I - \langle P_1, D\rangle * W_1 - \langle P_2, B\rangle * W_2\|_2^2 \\ &f_2 = \alpha\{\|\langle P_1, D\rangle * W_1\|_{fc} + \|\langle P_2, B\rangle * W_2\|_{fc}\} \\ &f_3 = \beta\{\|\Theta_1 \odot D_{ij:}\|_2 + \|\Theta_2 \odot B_{ij:}\|_2\} \\ &\quad + \gamma\{\|\nabla_1 \odot D\|_1 + \|\nabla_2 \odot B\|_1\} \\ &f_4 = \lambda\{\|P_1\|_F + \|P_2\|_F\} \end{aligned} \quad (6)$$

where $B_{ij:} = (B_{ij1}, B_{ij2}, \cdots, B_{ijE})^T$, and α, β, γ, and λ are the coefficients that control the degree of significance of the different terms. $\odot$ denotes element-wise multiplication. Then $\|\cdot\|_F$ denotes the Frobenius norm of the matrix, and the symbol $*$ is the convolution operator. ∇_1 and ∇_2 are fractal structure factors used to better promote the l_1 sparsity, respectively. Eq. (6) is so complex that it is hard to find the general optimization, from the perspective of division in components, the optimization can be easily conduced in the parallel way between components and source images.

B. Optimization of Example Detail Component based on ADMM Algorithm

Considering that the component separation are independent of each other, optimal iteration of the minimization problem is achieved by alternately solving two subproblems. Eq. (6) is divided into the following objective function. We will discuss the two parts separately and introduce $J = \langle P_1^T, W_1\rangle$ to get the our final objective function of detail component as an example, another base component is following the same way.

$$\begin{aligned} L(D, Y, P_1, U) = &\frac{1}{2} \left\| \hat{I} - \sum_k \hat{J}_k \cdot \hat{Y}_k \right\|_2^2 + \alpha \sum_k \left\| \hat{J}_k \cdot \hat{Y}_k \right\|_{fc} \\ &+ \beta \sum_{i=1}^{N} \sum_{j=1}^{N} \|\Theta_1 \odot D_{ij:}\|_2 + \gamma \sum_k \|\nabla_1 \odot D\|_1 \\ &+ \frac{\rho}{2} \sum_k \left\| \hat{Y}_k - \sqrt{T} F D_k + \hat{U}_k \right\|_2^2 + \lambda \|P_1\|_F^2 \end{aligned} \quad (7)$$

Here, we introduce auxiliary variables $\hat{Y}_k = \sqrt{T}FD_k$, where F represents the Fourier transform operation of the matrix D. Then we get the augmented Lagrangian function as follows. The $\hat{\cdot}$ denotes Discrete Fourier Transform (DFT), and $\hat{U}_k$ represent the slack variables. Therefore, Eq. (7) can be broken down into the following sub-problems. In here, the parameter $\hat{J} = (\hat{J}_1, \hat{J}_2, \cdots, \hat{J}_E)$, $\hat{Y} = (\hat{Y}_1, \hat{Y}_2, \cdots, \hat{Y}_E)^T$, $\hat{U} = (\hat{U}_1, \hat{U}_2, \cdots, \hat{U}_E)^T$ and $D = (D_1, D_2, \cdots, D_E)^T$.

*1) **solution of** $\hat{Y}^{l+1}$:* The function to be optimized for $\hat{Y}^{l+1}$ from Eq. (7) is:

$$\min \frac{1}{2}\left\|\hat{J}\cdot\hat{Y}-\hat{I}\right\|_2^2+\alpha\left\|\hat{J}^T\cdot\hat{Y}\right\|_{fc}+\frac{\rho}{2}\left\|\hat{Y}-\sqrt{T}FD+\hat{U}\right\|_2^2 \tag{8}$$

where $\|*\|_{fc}$ stands for our proposed fractal regularizer and $\odot$ represents element-wise multiplication. Obviously, the update of $\hat{Y}^{l+1}$ involves a set of small image patches based on fractal analysis.

According to the **Sherman-Morrison formulation** [33], the solution of the linear equation is:

$$\hat{Y}^{l+1}=\frac{\hat{J}^T\cdot\hat{I}+\rho(\sqrt{T}FD-\hat{U})}{(2\alpha+1)\hat{J}\cdot\hat{J}^T+\rho} \tag{9}$$

*2) **solution of** D^{l+1}:* Here, we introduce the auxiliary variable $Z = D$ and the penalty sparse μ. the optimization function of D is presented as follows.

$$\begin{aligned}\min_D \gamma\sum_k\|\nabla_1\odot D\|_1+\beta\sum_{i=1}^N\sum_{j=1}^N\|\Theta_1\odot D_{ij:}\|_2\\ +\frac{\rho}{2}\sum_k\left\|\hat{Y}_k^{l+1}-\sqrt{T}FD_k+\hat{U}_k^l\right\|_2^2\\ \text{s.t.}\quad Z=D\end{aligned} \tag{10}$$

According to the Half Quadratic Splitting algorithm [34], solving the Eq. (10) is equivalent to minimizing the following objective function:

$$\begin{aligned}\min_{D,Z}\gamma\sum_k\|\nabla_1\odot D\|_1+\beta\sum_{i=1}^N\sum_{j=1}^N\|\Theta_1\odot Z_{ij:}\|_2\\ +\frac{\rho}{2}\left\|\hat{Y}-\sqrt{T}FZ+\hat{U}\right\|_2^2+\frac{\mu}{2}\|D-Z\|_F^2\end{aligned} \tag{11}$$

Based on the derivation of Z in Eq. (11), we can get the preliminary solution of Z:

$$Z=\frac{\rho T(Y+U)+\mu D-\beta\Theta_1}{\rho T+\mathbf{1}\cdot\mu} \tag{12}$$

To improve the noise removal ability of the model and increase the fusion quality, the threshold selection operator is used to update Z, as shown below:

$$Z^{l+1}=\text{shrink}_1(\hat{Z},\frac{\beta}{\rho T+\mathbf{1}\cdot\mu}\odot\Theta_1) \tag{13}$$

where $\hat{Z}=\frac{\rho T(Y^{l+1}+U^l)+\mu D^l}{\rho T+\mathbf{1}\cdot\mu}$ and $\Theta_1(i,j,m)=\frac{1}{|\hat{Z}_{(i,j,m)}|+\epsilon}$, where ϵ is a small constant for avoiding the appearance of singularities. The operation of the threshold selection operator is as follows.

$$[\text{shrink}_1(x,\eta)]_{(i,j,m)}=\text{sign}(X_{(i,j,m)})\max(|X_{(i,j,m)}|-\eta,0) \tag{14}$$

η is the set threshold. Similarly, differentiating D in Eq. (11), we can get

$$D=Z-\frac{\gamma}{\mu}\nabla_1 \tag{15}$$

And then, D is further updated to get the final solution:

$$D^{l+1}(i,j,:)=\text{shrink}_{2,1}(\hat{D}(i,j,:),|\nabla_1(i,j)|\cdot\frac{\gamma}{\mu}) \tag{16}$$

where $\hat{D}=Z^{l+1}$, and $\nabla_1(i,j,:)=\frac{1}{\|\hat{D}_{(i,j,:)}\|_2+\epsilon}$.

The operation of the threshold selection on D is as follows.

$$\text{shrink}_{2,1}(D,\eta)=\begin{cases}\frac{\|D\|_2-\eta}{\|D\|_2}\cdot D & \text{if}\quad \eta<\|D\|_2\\ 0 & \text{otherwise}\end{cases} \tag{17}$$

*3) **solution of** P_1:* The problems discussed above do not consider the norm of P_1. To achieve the solution of P_1, we first obtain the minimization loss function in the Fourier domain of P_1 Here the base component is taken as an example.

$$P_1^{l+1}=\min_{P_1}\frac{1}{2}\left\|\hat{I}-\langle P_1,\hat{D}\rangle\cdot\hat{W}_1\right\|_2^2+\lambda\|P_1\|_F^2 \tag{18}$$

This loss function is similar to a matrix factorization problem,thus we can employ Gauss-Newton to linearizing the residuals and use the conjugate gradient descent method to obtain the optimal solution:

$$\min_{\Delta P}\left\|\hat{I}-\hat{D}^{l+1}(P+\Delta P)\hat{W}\right\|_2^2+\lambda\|P+\Delta P\|_F^2 \tag{19}$$

where ΔP is the matrix increment of P, then the solution of the quadratic problem in Eq. (19) can be can be expressed as:

$$\hat{W}^H(\hat{W}P\hat{D}+\hat{W}\Delta P\hat{D}-\hat{I})\hat{D}^H+\lambda\Delta P=-\lambda P \tag{20}$$

In the deduction above, we use the simplified form to obtain the solution of the minimization problem. Then the complete form of Eq. (20) can be defined as follows:

$$\begin{cases}G(\hat{W}\cdot\langle P,\hat{D}\rangle)+\lambda_p\Delta P+G(\hat{W}\cdot\langle\Delta P,\hat{D}\rangle)=G(\hat{I})-\lambda P\\ G(A)=\hat{W}^H\odot(A\cdot\hat{D}^H)\end{cases} \tag{21}$$

where $G(A)$ is an operator produced by derivation.

Let $A(\Delta P)=G(\hat{W}\cdot\langle P,\hat{D}\rangle)+\lambda_p\Delta P+G(\hat{W}\cdot\langle\Delta P,\hat{D}\rangle)$, $b=G(\hat{I})-\lambda P$. Eq. (21) can be written as, $A(\Delta P)=b$. Conjugate gradient descent is used to optimize ΔP. And then, the update of P can be obtained as, $P^{(l+1)}=P^{(l)}+\Delta P$.

IV. EXPERIMENTS

A. Experimental Settings and Metrics

*1) **Experimental setting**:* Experiments are carried out by MATLAB R2016b on a computer with Dual-Core Intel Core i5 processor (1.8GHz) and 8GB 1600 MHz DDR3. For a fair comparison, for both basic B and detail D, we adopt the same fusion strategy as the CSR-based model [29] .We experimentally fix $\alpha = 1.5$, $\beta = 4\times10^{-}7$, $\lambda = 5\times10^{-}8$, $\rho = 50\times\gamma+1$, the patch size is 16×16 and the iteration number is 3. We selecte seven metrics including Average Gradient (AG), Correlation Coefficient (CC), Entropy (EN), Mean Square Error (MSE), Root Mean Squared Error (RMSE),

TABLE I
Objective performance of the fusion methods in 5 modalities. The best each metrics are marked in red.

		$AG\uparrow$	$CC\uparrow$	$EN\uparrow$	$MSE\downarrow$	$RMSE\downarrow$	$MI\downarrow$	$SF\uparrow$	$F_{rank}\downarrow$
MRI-CT	NSST_PAPCNN	0.0233±0.0090	0.8465±0.0379	4.4089±0.9834	0.0337±0.0070	0.1768±0.0164	8.8177±1.9668	-0.2773±0.0718	5.1
	NSCT_PCDC	0.0248±0.0082	0.8068±0.0460	4.4811±0.8200	0.0309±0.0057	0.1698±0.0161	8.9622±1.6401	-0.1635± 0.0310	4.0
	NSCT_RPCNN	0.0239±0.0092	0.8372±0.0384	3.7843 ±0.9456	0.0358±0.0080	0.1772 ±0.0170	7.5686± 1.8911	-0.1525±0.0768	5.5
	LP_SR	0.0234±0.0089	0.8248±0.0504	3.5007±0.8534	0.0341±0.0097	0.1666±0.0174	7.0015±1.7068	-0.1169±0.0677	5.6
	GFF	0.0215±0.0075	0.8309±0.0469	3.8041±0.9386	0.0287±0.0036	0.1611±0.0089	7.6082±1.8773	-0.2447± 0.0622	4.8
	CSMCA	0.0215±0.0080	0.8360±0.0422	3.9825 ±1.0957	0.0269±0.0049	0.1594 ±0.0130	7.9649±2.1914	-0.1770± 0.0251	4.3
	IFCNN	0.0250±0.0091	0.8485±0.0379	3.7312 ±0.8941	0.0272±0.0056	0.1575±0.0142	7.4624±1.7882	-0.1280± 0.0422	3.6
	FCFusion	0.0236±0.0084	0.8388±0.0406	4.5543±1.0752	0.0265±0.0046	0.1592±0.0146	9.1085±2.1504	-0.1904±0.0440	3.1
T1-T2	NSST_PAPCNN	0.0264±0.0096	0.8798± 0.0540	4.7225±0.5682	0.0211± 0.0090	0.1377± 0.0322	9.4451±1.1364	-0.2356±0.0763	5.1
	NSCT_PCDC	0.0287±0.0097	0.8596±0.0650	4.5128±0.6284	0.0189±0.0077	0.1321± 0.0292	9.0255 ±1.2568	-0.1134±0.0276	4.6
	NSCT_RPCNN	0.0284±0.0092	0.8738±0.0505	4.1102±0.6200	0.0226±0.0093	0.1416±0.0321	8.2203± 1.2401	-0.1306±0.0175	5.5
	LP_SR	0.0294±0.0098	0.8599±0.0726	3.9162±0.6122	0.0220±0.0096	0.1356±0.0288	7.8325± 1.2243	-0.0855±0.0188	5.1
	GFF	0.0260±0.0093	0.8667±0.0718	4.0347±0.6632	0.0184±0.0085	0.1234±0.0276	8.0693± 1.3263	-0.1847±0.0372	5.0
	CSMCA	0.0272±0.0097	0.8771±0.0642	3.9033±0.6247	0.0170±0.0072	0.1248 ±0.0269	7.8066±1.2494	-0.1406±0.0355	4.7
	IFCNN	0.0293±0.0098	0.8893±0.0526	3.9086±0.6095	0.0164±0.0069	0.1222±0.0271	7.8172±1.1437	-0.0622±0.0494	3.4
	FCFusion	0.0295±0.0088	0.8754±0.0621	4.8227±0.5718	0.0172±0.0071	0.1260±0.0280	9.6455±1.2189	-0.1061±0.0230	2.6
T2-PD	NSST_PAPCNN	0.0226±0.0046	0.9596±0.0112	5.1315±0.7274	0.0055±0.0026	0.0709± 0.0159	10.2631±1.4547	-0.1779±0.0631	5.2
	NSCT_PCDC	0.0246±0.0058	0.9563±0.0134	4.7843±0.6040	0.0053±0.0025	0.0704±0.0152	9.5685±1.2080	-0.0720±0.0214	4.0
	NSCT_RPCNN	0.0240±0.0054	0.9521±0.0119	4.6876±0.7318	0.0061±0.0029	0.0747±0.0167	9.3753 ±1.4637	-0.1160±0.0272	5.5
	LP_SR	0.0247±0.0057	0.9540±0.0143	4.3775±0.6522	0.0058±0.0028	0.0723±0.0147	8.7549±1.3044	-0.0614± 0.0160	5.6
	GFF	0.0234±0.0055	0.9559±0.0135	4.5152±0.7216	0.0054±0.0026	0.0663±0.0138	9.0303 ±1.4432	-0.1151±0.0296	4.8
	CSMCA	0.0238±0.0055	0.9599±0.0108	4.5190±0.6238	0.0050±0.0022	0.0685±0.0137	9.0379 ±1.2476	-0.0889±0.0219	4.3
	IFCNN	0.0256±0.0063	0.9649±0.0095	4.4285±0.6970	0.0048±0.0021	0.0674±0.0133	8.8569±1.3939	-0.0089±0.0318	3.6
	FCFusion	0.0257±0.0049	0.9581±0.0106	5.1398±0.5203	0.0052±0.0022	0.0698±0.0131	10.2796±1.0406	-0.0480±0.0579	3.1
MRI-PET	NSST_PAPCNN	0.0177±0.0054	0.8700±0.0540	4.4683±0.9984	0.0150±0.0063	0.1171±0.0248	8.9366±1.9968	-0.0747±0.0260	4.1
	NSCT_PCDC	0.0173±0.0055	0.8401±0.0639	4.4677±0.8855	0.0114±0.0053	0.0995±0.0211	8.9355±1.7710	-0.0857±0.0239	4.0
	NSCT_RPCNN	0.0181±0.0057	0.8628±0.0533	4.2220± 1.0236	0.0155± 0.0065	0.1186±0.0254	8.4440±2.0472	-0.0615±0.0243	4.6
	LP_SR	0.0165±0.0053	0.8660±0.0546	4.0162±0.9691	0.0117±0.0052	0.1040±0.0236	8.0325±1.9381	-0.0994±0.0364	5.1
	GFF	0.0165±0.0059	0.8343±0.0987	4.0384±1.0383	0.0123±0.0076	0.0912±0.0187	8.0768±2.0766	-0.1209±0.0760	4.8
	IFCNN	0.0177±0.0056	0.8743±0.0572	4.0480±0.9321	0.0104±0.0047	0.0976±0.0217	8.0961±1.8641	-0.0465±0.0195	3.0
	FCFusion	0.0182±0.0051	0.8634±0.0553	4.6488±0.6528	0.0103±0.0045	0.0983±0.0217	9.2975±1.3057	-0.0548±0.0478	2.4
MRI-SPECT	NSST_PAPCNN	0.0213±0.0070	0.8883±0.0899	4.9682±0.7964	0.0202± 0.0232	0.1123±0.0431	9.9364±1.5928	-0.0528±0.0191	4.2
	NSCT_PCDC	0.0213±0.0071	0.8766±0.0963	4.9421±0.7735	0.0161±0.0214	0.1017±0.0485	9.8843±1.5471	-0.0561±0.0168	4.1
	NSCT_RPCNN	0.0218±0.0073	0.8755±0.0863	4.7035±0.7574	0.0210± 0.0233	0.1141±0.0422	9.4070±1.5148	-0.0436±0.0193	4.6
	LP_SR	0.0210±0.0072	0.8931± 0.0873	4.5611±0.7461	0.0169± 0.0227	0.1017±0.0447	9.1221±1.4922	-0.0546±0.0180	4.6
	GFF	0.0209±0.0076	0.8781± 0.0931	4.6561±0.7458	0.0169±0.0225	0.0933±0.0463	9.3122±1.4915	-0.0703±0.0379	4.5
	IFCNN	0.0213± 0.0075	0.8977± 0.0927	4.6111±0.7720	0.0136±0.0165	0.0956±0.0451	9.2222±1.5439	-0.0464±0.0311	3.6
	FCFusion	0.0218±0.0062	0.8949±0.0746	5.0521±0.5337	0.0131±0.0128	0.1014±0.0430	10.1043±1.0674	-0.0347±0.0589	2.4

Mutual Information (MI), Spatial Frequency (SF) in [35].We also calculate the average rank of these 7 indicators, which is F-rank. Our experiments take use of five different modalities of fusion (magnetic resonance imaging (MRI), positron emission tomography (PET), computed tomography (CT) and single photon emission computed tomography (SPECT) images), including MRI-CT, MRI-PET, MRI-SPECT and T1-T2, T2-PD, where T1, T2 and PD are MRI images based on different weights. The source images used in the experiment are all from the Whole Brain Atlas [36] established by Harvard Medical School.

*2) **Comparison Methods** :* Our FCFusion model compares with 7 existing medical image fusion methods, including NSST-PAPCNN [22], NSCT-PCDC [37], NSCT-RPCNN [23], GFF [38], CS-MCA [29], LP-SR [24] and IFCNN [39], where IFCNN is a deep learning based fusion method. It should be noted that in the fusion of color images, the CS-MCA method is not included in the comparison. All parameters in these methods are set to the default values for unbiased comparison.

B. Fusion Results and Analysis

1) Qualitative Analysis: The comparative experiment is mainly based on the five modalities mentioned above. Fig. 2 shows the fusion results of three randomly selected modalities, MRI-CT, T1-T2 and MRI-PET. And below we conduct a detailed analysis.

Fusion visual effect based on MRI-CT: The experiment based on 8 methods to fuse CT and MRI images is conducted. The comparison results are shown in the first row. It can be seen that the fused image generated by NSST-PAPCNN (c1) is blurred at the boundary. The structure and detail features at the boundary are not preserved. Although the fused images of NSCT-RPCNN (e1) and IFCNN (i1) retain structural similarity, there are deficiencies in the extraction of detail information compared with our FCFusion method. In addition, NSCT-PCDC (d1) has serious energy loss, which reduces the contrast of the fusion image. Finally, our FCFusion has excellent performance in detail and structural information retention and shows a good fusion performance.

Fusion visual effect based on T1-T2: The second row shows the fused comparative experiments. NSST-PAPCNN (c2) and LP-SR (f2) still have the problem of blurred boundaries. The fusion images of NSCT-RPCNN (e2) , GFF (h2) and IFCNN (i2) are insufficient in extracting detailed information. In contrast, the fused image of our method retains the texture and detail features of the source image well and maintains a high image. The enlarged part of the picture shows that our method can well show the gaps and differences between different tissue structures, however other fused images have neutralized this difference and cannot highlight the changes in soft tissue.

Fusion visual effect based on MRI-PET: The results of the fusion are shown in the third row. Since CSMCA cannot perform the fusion of color images, there are only 6 comparison methods in color image fusion. Clearly, the fused image of GFF (g3) is almost the same as the MRI source image. There is no extraction and fusion of PET image information during the fusion process. And the fused images of NSCT-PCDC (d3) have serious color distortion, even affecting the resolution of the image. The symptom of NSST-PAPCNN (c3) is slightly lighter than the other methods, but there is still obvious discoloration. The fusion of IFCNN (h3) has energy loss, and the contrast of the fused image is reduced. The fused images of NSCT-RPCNN (e3) and LP-SR (f3) retain the color

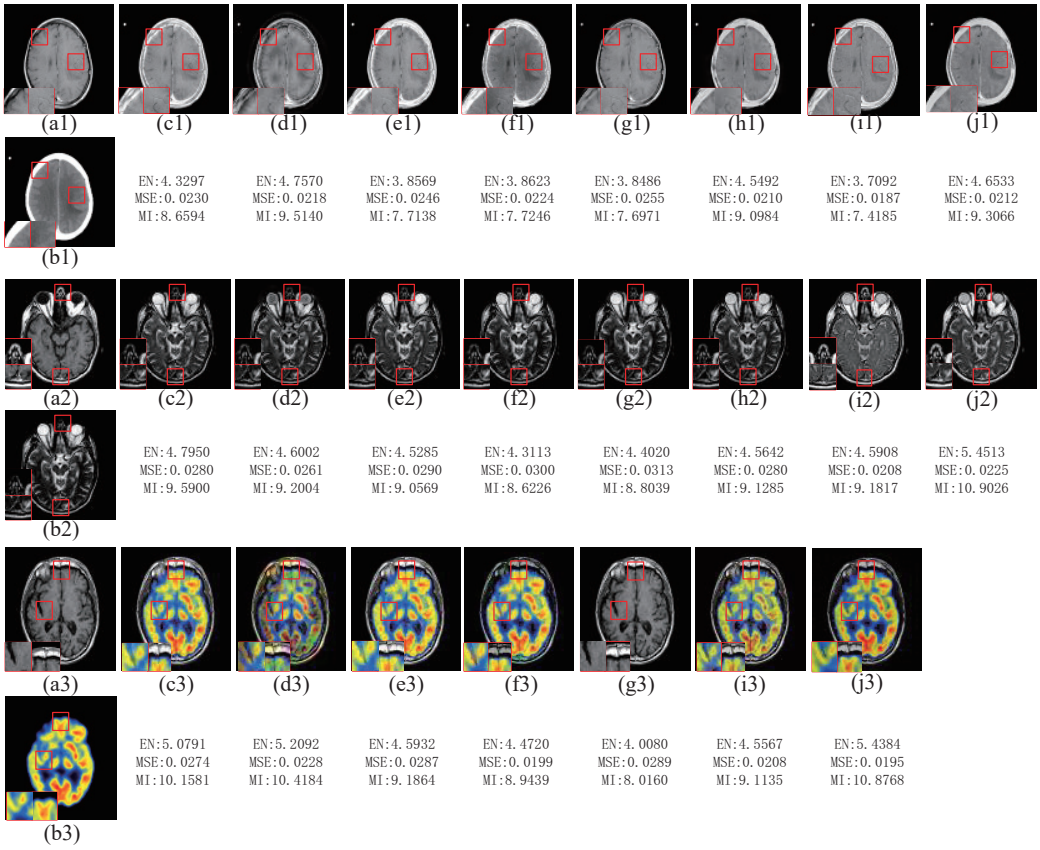


Fig. 2. The fusion performance of various methods, followed by MRI-CT, T1-T2 and MRI-PET. Here, (a1-a3) sourceimage1; (b1-b3) sourceimage2; (c1-c3) NSST-PAPCNN; (d1-d3) NSCT-PCDC; (e1-e3) NSCT-RPCNN; (f1-f3) LP-SR; (g1-g3) GFF; (h1-h3) CSMCA; (i1-i3) IFCNN and (j1-j3) FCFusion. Particularly, CSMCA cannot be used for MRI-PET fusion.

image information, and retain the edge and detail texture of the MRI images.

2) Quantitative Analysis: The comparative analysis of the image fusion effect described above may not be objective enough, then the quantitative analysis of the experiment is added. And the average values to be compared are obtained from the multi-group image fusion index of 5 modalities, including 10 groups of MRI-CT image pairs, 35 groups of T1-T2 image pairs, 46 groups of T2-PD image pairs, 25 groups of MRI-PET and 25 groups of MRI-SPECT. All results are listed in the table I, where the top for each metric is marked with red. The first three groups in the table I are gray-scale fusion modalities, namely MRI-CT, T1-T2 and T2-PD, and the latter two are color fusion modalities, namely MRI-PET and MRI-SPECT. In general, the performance on color fusion modalities is better than that on the grayscale fusion modalities. In MRI-CT and T2-PD, our fusion model ranks first on 4 different metrics. Although there are only 3 best metrics in T1-T2, MSE and RMSE metrics for different methods are very close. In color fusion, almost all indicators of FCFusion model are among the top three. Additionally, we notice that whether it is grayscale or color image fusion, the ranking of our model on EN and MI are the first, which proves that our model can retain the feature information of the source image as much as possible during the fusion process, and maintain a good image structure similarity degree. And our model also leads on AG metric, except for MRI-CT modalities, indicating that our image clarity is better than other methods.

Overall, based on the above analysis results, obviously our model achieves excellent fusion effects in various modalities, both visually and objectively. This means our model can be used for medical image fusion in a variety of situations without being restricted by the modality. Therefore, it can be safely assumed that our proposed method has a good robustness.

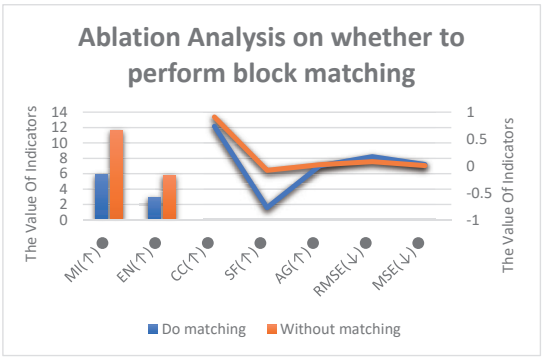


Fig. 3. Ablation study on performing the specific block matching

3) Configuration Analysis: Our model introduces two variable weighted coefficient matrices $\odot$ and ∇ to improve the group sparseness and spatial sparseness of model features. To verify the improvement effect of $\odot$ and ∇ on the fusion effect, we conducted a corrosion comparison experiment. The model does not have the effect of $\odot$ and ∇ is obtained by setting all

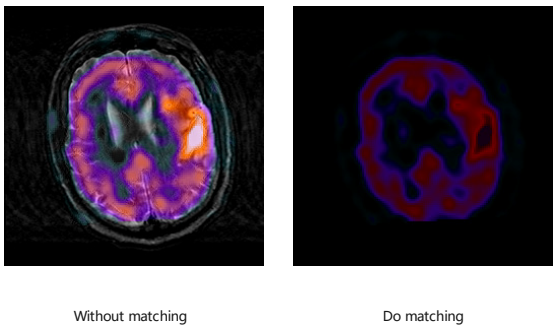


Fig. 4. The comparison result images with or without matching.

matrix elements of $\odot$ and ∇ to 1. The experimental results are shown in Table II. The first row of data of FCFusion-1 corresponds to the model with $\odot$ and ∇, and the second row of data of FCFusion-2 corresponds to the model without $\odot$ and ∇. The better results between the two are marked in red. According to Table II, the model with $\odot$ and ∇ has better results in most metrics, it is believable that the variable weight coefficient matrices $\odot$ and ∇ are helpful to the fusion effect to a certain extent. According to the model construction and solution introduced above, our model uses the idea of dividing patches to maintain the structural similarity of the images. It can be seen from the Fig. 3 the results of our model, in most instances, are better than those doing matches (that is ,the non-local model) except for PSNR and SSIM, which proves that our method is helpful for the fusion effect.The comparison image of the fusion is shown in Fig. 4 below.

C. Ablation Analysis

1) Configuration Analysis: Our model introduces two variable weighted coefficient matrices $\odot$ and ∇ to improve the group sparseness and spatial sparseness of model features. To verify the improvement effect of $\odot$ and ∇ on the fusion effect, we conduct an ablation comparison experiment. The experimental results are shown in Table II. The first row of data of FCFusion-1 corresponds to the model with $\odot$ and ∇, and the second row of data of FCFusion-2 corresponds to the model without $\odot$ and ∇. The better results between the two are marked in red. According to Table II, the model with $\odot$ and ∇ has better results in most metrics. It is believable that the variable weight coefficient matrices $\odot$ and ∇ are helpful to the fusion effect to a certain extent. According to the model construction and solution introduced above, our model uses the idea of dual variable weighting to maintain group sparsity of images. It also can be seen from the Fig. 3 the results of our model, in most instances, are better than those doing block matching process (the non-local model commonly applied [40]), which proves that our method is helpful for the fusion effect. The comparison image of the fusion is shown in Fig. 4.

2) Convergence Analysis: To ensure the fusion effect of the model, we adopt ADMM and the conjugate gradient descent method to achieve the optimal solution of the optimization. As shown in Fig. 6, our fusion model has good convergence in both gray-scale fusion and color fusion. After several iterations, the result is very close to the optimal solution of convergence, and the converged line graph does not show large fluctuations, which means our fusion model has good stability.

TABLE II
The ablation analysis based on $\odot$ and∇

	FCFusion-1	FCFusion-2
AG	0.02741	0.0289
CC	0.8998	0.902
EN	4.375	4.3132
MSE	0.0143	0.0142
RMSE	0.1075	0.1076
MI	8.7501	8.6263
SF	−0.1239	−0.1319
TIME	93.1117	85.2652

D. Real Application Verification of Fusion Algorithm

PET can reflect the genetic, molecular, metabolic, and functional status of the disease. The diseased part usually shows greater 18F-flurodeoxyglucose (FDG) uptake than the normal structure [41]. CT can provide information on the anatomical structure of the part where FDG uptake is abnormal in PET [42]. When performing lesion detection and lesion contour delineation, it is inaccurate to segment PET and CT separately for judgment. PET-CT integrates the metabolic information of PET and the texture features of CT into one image. From the above discussion, we can see that our method has obvious advantages in the fusion of PET and CT. To demonstrate the important value of this advantage, we collect several pairs of PET and CT images obtained from patients suspected of lung cancer for experiments. Our model and five other fusion methods are used to obtain PET-CT images. Then, we segment the PET, CT, and fused PET-CT images to determine the lesion area of lung cancer patients. It can be seen from Fig. 5 that our fusion results can not only detect the disease parts that show no abnormalities in the CT images, but also filter out the non-pathological interference areas of PET based on the knowledge of CT anatomical features, and obtain precise anatomical positioning during segmentation. However, other methods, such as NSCT-PCDC (f1-f3) and NSCT-RPCNN (g1-g3), are disturbed by the high-level metabolism of non-lesion areas in PET images. In their segmentation results, either the location is wrong or the range is enlarged, thereby affecting the patient's follow-up treatment.

V. Conclusion

We propose a medical image fusion method based on fractal prior of components and group sparsity. The proposed mechanism strengthens the connection between different components and improves the fusion performance. The main motivation is the feature extraction and redundancy removal. The global group weighting strategy is adopted in refining the over-smoothing issue. The ADMM algorithm and conjugate gradient descent method are used to solve the model optimization problem. Finally, we conducted a large number of comparative experiments and extracted some results for visualization and analysis. The experimental results in real applications prove that our algorithm has advantages in the performance of fusion and segmentation compared with other algorithms.

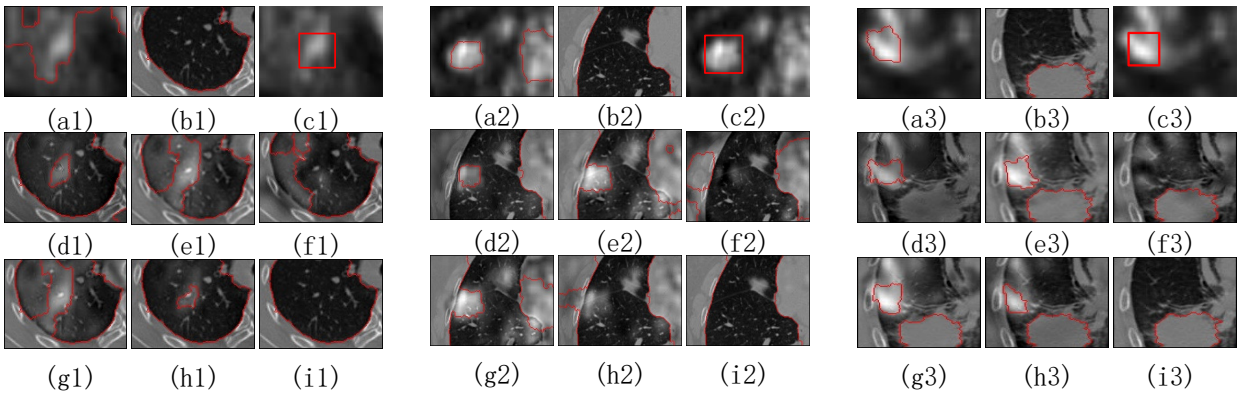
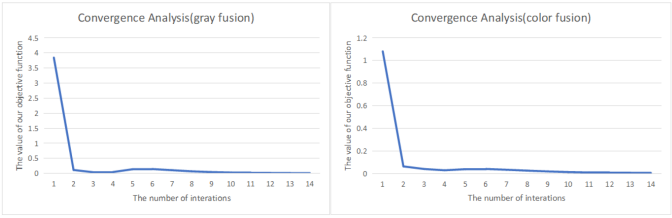


Fig. 5. The segmentation performance of fused image from several algorithms, as is PET, CT, GroundTruth on PET, FCFusionFusion (Our Method), NSST_PAPCNN, NSCT_PCDC, NSCT_RPCNN, LP_SR, GFF.



(a) (b)

Fig. 6. Convergence analysis of FCFusion in gray fusion and color fusion.

In this paper, we conduct experiments using images from the Whole Brain Atlas, and perform a clinical segmentation on PET and CT images. In our future work, we will try more other datasets to fully demonstrate the robustness of the model. And we also hope to explore more real applications of other modalities, thus further illustrating the contribution of our model in the industrial field.

REFERENCES

[1] H. Choi, J. P. Yun, B. J. Kim, H. Jang, and S. W. Kim, "Attention-based multimodal image feature fusion module for transmission line detection," *IEEE Transactions on Industrial Informatics*, 2022.

[2] E. Benli, R. L. Spidalieri, and Y. Motai, "Thermal multisensor fusion for collaborative robotics," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 3784–3795, Jul. 2019.

[3] S. Prabhakar and A. K. Jain, "Decision-level fusion in fingerprint verification," *Pattern Recognition*, vol. 35, no. 4, pp. 861–874, 2002.

[4] P. Aparna and P. V. V. Kishore, "Biometric-based efficient medical image watermarking in e-healthcare application," *IET Image Processing*, vol. 13, no. 3, pp. 421–428, 2019.

[5] Y. Ueda, J. Morishita, S. Kudomi, and K. Ueda, "Usefulness of biological fingerprint in magnetic resonance imaging for patient verification," *Medical & biological engineering & computing*, vol. 54, no. 9, pp. 1341–1351, 2016.

[6] M. Daneshzand, R. A. Zoroofi, and M. Faezipour, "Mr image assisted drug delivery in respiratory tract and trachea tissues based on an enhanced level set method," in *Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education*, 2014, pp. 1–7.

[7] X. Zhou, X. Xu, W. Liang, Z. Zeng, S. Shimizu, L. T. Yang, and Q. Jin, "Intelligent small object detection for digital twin in smart manufacturing with industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 2, pp. 1377–1386, 2021.

[8] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end–edge–cloud surveillance in smart iot," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12 588–12 596, 2021.

[9] X. Zhou, W. Liang, I. Kevin, K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 171–178, 2020.

[10] X. Li, Y. Jiang, M. Li, and S. Yin, "Lightweight attention convolutional neural network for retinal vessel image segmentation," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1958–1967, 2020.

[11] S. Liu, X. Li, Y. Jiang, H. Luo, Y. Gao, and S. Yin, "Integrated learning approach based on fused segmentation information for skeletal fluorosis diagnosis and severity grading," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7554–7563, Jan. 2021.

[12] R. Yang, B. Du, P. Duan, Y. He, H. Wang, Y. He, and K. Zhang, "Electromagnetic induction heating and image fusion of silicon photovoltaic cell electrothermography and electroluminescence," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4413–4422, Jul. 2020.

[13] W. R. Tan, C. S. Chan, P. Yogarajah, and J. Condell, "A fusion approach for efficient human skin detection," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 138–147, Feb. 2012.

[14] X. Xu, H. Peng, M. Z. A. Bhuiyan, Z. Hao, L. Liu, L. Sun, and L. He, "Privacy-Preserving federated depression detection from multisource mobile health data," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4788–4797, Jul. 2022.

[15] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," in *2017 20th International Conference on Information Fusion (Fusion)*, 2017, pp. 1–7.

[16] K.-j. Xia, H.-s. Yin, and J.-q. Wang, "A novel improved deep convolutional neural network model for medical image fusion," *Cluster Computing*, vol. 22, no. 1, pp. 1515–1527, 2019.

[17] H. Hermessi, O. Mourali, and E. Zagrouba, "Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain," *Neural Computing and Applications*, vol. 30, no. 7, pp. 2029–2045, 2018.

[18] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884–892, Apr. 2020.

[19] M. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Information Fusion*, vol. 27, pp. 198–214, 2016.

[20] S. Li, H. Yin, and L. Fang, "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Transactions on biomedical engineering*, vol. 59, no. 12, pp. 3450–3459, 2012.

[21] "Multimodal medical image fusion review: Theoretical background and recent advances," *Signal Processing*, vol. 183, p. 108036, 2021.

[22] 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 Transactions on Instrumentation and Measurement*, vol. 68, no. 1, Jan. 2019.

[23] S. Das and M. Kundu, "A neuro-fuzzy approach for medical image fusion," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, Dec. 2013.

[24] Y. Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion*, vol. 24, pp. 147–164, 2015.

[25] L. Yin, M. Zheng, G. Qi, Z. Zhu, F. Jin, and J. Sim, "A novel image fusion framework based on sparse representation and pulse coupled neural network," *IEEE Access*, vol. 7, pp. 98 290–98 305, 2019.

[26] Y. J. A. B and M. W. A, "Image fusion with morphological component analysis," *Information Fusion*, vol. 18, no. 1, pp. 107–118, 2014.

[27] J. L. Starck, M. Elad, and D. Donoho, "Redundant multiscale transforms and their application for morphological component separation," *Advances in Imaging & Electron Physics*, vol. 132, no. 04, pp. 287–348, 2004.

[28] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Medical image fusion via convolutional sparsity based morphological component analysis," *IEEE Signal Processing Letters*, vol. 26, no. 3, pp. 485–489, 2019.

[29] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Medical image fusion via convolutional sparsity based morphological component analysis," *IEEE Signal Processing Letters*, vol. 26, no. 3, pp. 485–489, March 2019.

[30] X. Li, X. Guo, P. Han, X. Wang, H. Li, and T. Luo, "Laplacian redecomposition for multimodal medical image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6880–6890, 2020.

[31] F. Heide, W. Heidrich, and G. Wetzstein, "Fast and flexible convolutional sparse coding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5135–5143.

[32] Y. Chen, W. He, N. Yokoya, and T. Huang, "Weighted group sparsity regularized low-rank tensor decomposition for hyperspectral image restoration," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 234–237.

[33] W. W. Hager, "Updating the inverse of a matrix," *SIAM review*, vol. 31, no. 2, pp. 221–239, 1989.

[34] R. He, W.-S. Zheng, T. Tan, and Z. Sun, "Half-quadratic-based iterative minimization for robust sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 261–275, 2013.

[35] P. R. Raju, D. S. Rani, and G. Challaram, "Comparison of medical image fusion methods using image quality metrics," in *Proc. 2018 Int. Conf. Commun., Comput. Internet Things (IC3IoT)*, Feb. 2018, pp. 449–454.

[36] J. A. Johnson and J. A. Becker, "The whole brain altas," *Online*, 1997.

[37] G. Bhatnagar, Q. Wu, and Z. Liu, "Directive contrast based multimodal medical image fusion in nsct domain," *IEEE Transactions on Multimedia*, vol. 15, no. 5, Aug. 2013.

[38] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, July 2013.

[39] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "Ifcnn: A general image fusion framework based on convolutional neural network," *Information Fusion*, vol. 54, pp. 99–118, 2020.

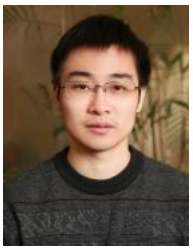
[40] L. Guo, Z. Zha, S. Ravishankar, and B. Wen, "Exploiting non-local priors via self-convolution for highly-efficient image restoration," *IEEE Transactions on Image Processing*, 2022.

[41] S. Kligerman and S. Digumarthy, "Staging of non–small cell lung cancer using integrated pet/ct," *American Journal of Roentgenology*, vol. 193, no. 5, pp. 1203–1211, 2009.

[42] U. Bagci, J. K. Udupa, N. Mendhiratta, B. Foster, Z. Xu, J. Yao, X. Chen, and D. J. Mollura, "Joint segmentation of anatomical and functional images: Applications in quantification of lesions from pet, pet-ct, mri-pet, and mri-pet-ct images," *Medical image analysis*, vol. 17, no. 8, pp.

Guoxia Xu (M'19) received the B.S. degree in information and computer science from Yancheng Teachers University, Jiangsu Yancheng, China in 2015, and the M.S. degree in computer science and technology from Hohai University, Nanjing, China in 2018. He was a research assistant in City University of Hong Kong and Chinese University of Hong Kong. Now, he is pursing his Ph.D. degree in Department of Computer Science, Norwegian University of Science and Technology, Gjøvik Norway. His research interest includes pattern recognition, image processing, and computer vision. E-mail: gxxu.re@gmail.com

Xiaoxue Deng is now pursuing B.E. degree in communication engineering in Nanjing University of Posts and Telecommunications. Her current research interest is medical image fusion based on sparse representation. E-mail: dengxiaoxue1018@163.com



Xiaokang Zhou (M'16) received the Ph.D. degree in human sciences from Waseda University, Japan, in 2014. From 2012 to 2015, he was a Research Associate with the Faculty of Human Sciences, Waseda University. He has been working as a Visiting Researcher with the RIKEN Center for Advanced Intelligence Project (AIP), RIKEN, Japan, since 2017. He is currently an Associate Professor with the Faculty of Data Science, Shiga University, Japan. He has been engaged in interdisciplinary research works in the fields of computer science and engineering, information systems, and social and human informatics. His recent research interests include ubiquitous computing, big data, machine learning, behavior and cognitive informatics, cyber-physical-social systems, and cyber intelligence and security. He is a member of the IEEE CS; ACM, USA; IPSJ; JSAI, Japan; and CCF, China. E-mail: zhou@biwako.shiga-u.ac.jp

Marius Pedersen (M'18) received the B.Sc. degrees in computer engineering and media technology (MiT) from the Gjøvik University College, Norway, in 2006 and 2007, and the Ph.D. degree in color imaging from the University of Oslo, Norway, in 2011, sponsored by Océ. He is currently a Full Professor with the Department of Computer Science, NTNU, Gjøvik, Norway. He is also the Director of The Norwegian Colour and Visual Computing Laboratory (Colourlab). His research interest includes subjective and objective image quality.

Lucia Cimmino received the B.S. and M.S degrees (cum laude) in computer science from the University of Salerno, Fisciano (SA), Italy, in 2018 and 2020, respectively. She is currently pursuing the Ph.D. degree in computer science with the Biometric and Image Processing Laboratory (BIPLAB), University of Salerno, Italy. Her research fields include Machine Learning, Biometrics, Facial Recognition and Image Processing.

Hao Wang (SM'21) received the B.Eng. and Ph.D. degrees, both in computer science and engineering, from South China University of Technology, China in 2000 and 2006, respectively. He is an Associate Professor in the Department of Computer Science, Norwegian University of Science & Technology, Norway. His research interests include big data analytics, industrial internet of things, high performance computing, and safety critical systems. He has published 170+ papers in reputable international journals and conferences including IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Industrial Electronics, Transactions on Industrial Informatics, Internet of Things Journal, Transactions on Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on Geoscience and Remote Sensing and ACM Computing Surveys. He serves as the Editorial Board Member and Guest Editors for several international journals. He served as a TPC Co-Chair for IEEE Transactions on Computational Social Systems 2020, IEEE International Conference on Computer and Information Technology 2017, International Conference on Big Data Intelligence and Computing 2015, and Reviewers for many prestigious journals and conferences. He is a Senior Member of IEEE and a Member of ACM. He is the Chair for Sub TC on Healthcare in IEEE IES Technical Committee on Industrial Informatics. E-mail: hawa@ntnu.no