IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

A novel GA-based optimized approach for regional multimodal medical image fusion with superpixel segmentation

JUNWEI DUAN[1,2], SHUQI MAO[1], JUNWEI JIN[3], ZHIGUO ZHOU[4], LONG CHEN[5], C. L. PHILIP CHEN[6]

[1]College of Information Science and Technology, Jinan University, Guangzhou 511400, China
[2]Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization, Guangzhou 511400, China
[3]School of Artificial Intelligence and Big Data, Henan University of Technology, Henan 475000, China
[4]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[5]Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Macau, China
[6]School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China

Corresponding author: Junwei Duan (e-mail: jwduan@jnu.edu.cn).
Co-corresponding author: Zhiguo Zhou (e-mail: zhiguozhou@bit.edu.cn).

This work was supported in part by the Fundamental Research Funds for the Central Universities, Jinan University under Grant 21619412, in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2021A1515011999, in part by the National Key R&D Program of China under Grant 2018YFC2002500 and in part by Guangdong Provincial Key Laboratory of Traditional Chinese Medicine Informatization under Grant 2021B1212040007.

ABSTRACT For multimodal medical image fusion problems, most of the existing fusion approaches are based on pixel-level. However, the pixel-based fusion method tends to lose local and spatial information as the relationships between pixels are not considered appropriately, which has much influence on the quality of the fusion results. To address this issue, a region-based multimodal medical image fusion framework is proposed based on superpixel segmentation and a post-processing optimization method in this paper. In this framework, the averaging image of the source medical images are firstly obtained by a weighted averaging method. To effectively obtain homogeneous regions and preserve the complete information of image details, the fast linear spectral clustering(LSC) superpixel algorithm is carried out to segment the averaging image and get superpixel labels. For each region of the medical images, log-gabor filter(LGF) and sum modified laplacian(SML) are adopted to extract texture feature and contrast feature for the measurement of region importance. The most important regions are selected and the decision map is generated by comparison. Moreover, to get a more accurate decision map, a new post-processing optimized method based on genetic algorithm(GA) is given. A weighted strategy is applied to the extracted features and the weighting factor can be adaptively adjusted by GA. The effectiveness of the proposed fusion method is validated by conducting experiments on eight pairs of medical images from diverse modalities. In addition, seven other mainstream medical image fusion methods are adopted for comparing the performance of fusion. Experimental results in terms of qualitative and quantitative evaluation demonstrate that the proposed method can achieve state-of-the-art performance for multimodal medical image fusion problems.

INDEX TERMS Multimodal medical image fusion, superpixel segmentation, genetic algorithm, log-gabor filter, sum modified laplacian.

I. INTRODUCTION

As a fundamental and effective supplementary tool, medical images play an increasingly significant role in modern clinical diagnosis and treatments. However, due to the limitation of the imaging mechanism, medical images from a single modality usually cannot provide sufficient information to meet the requirements of complex diagnoses [1]. For instance, computerized tomography (CT) image can provide a clear visualization of dense structures like bones and implants, but it's not good at presenting the soft tissues.

Magnetic resonance imaging (MRI) image can provide high-resolution detailed information of soft tissues, but it is also prone to introduce artifacts when taking photos of bone structures [2]. Functional information of blood flow and metabolic changes can be reflected by positron emission tomography (PET) and single-photon emission computed tomography (SPECT) images, but the spatial resolution is usually very low. Multimodal medical image fusion is an effective technique to solve this problem, which aims to generate fusion images with complementary information contained in medical images from different modalities.

For medical image fusion problems, numerous methods have been proposed which can be roughly divided into three levels: pixel-level, feature-level, and decision-level [3]. Generally, pixel-level image fusion mainly includes two categories: spatial domain methods and transform domain methods. Spatial domain methods such as PCA [4], IHS [5], and averaging fusion select pixels from the source images to construct the final fused image. This kind of fusion methods can completely preserve spatial information and reduce computational complexity [6]. However, they also introduce color distortion and suffer from contrast decrease, which are unacceptable for the fusion of medical images. Different from spatial domain methods, transform domain methods decompose source images into high and low frequency by transform. Multiscale-transform(MST)-based approaches are popular in the field of medical image fusion due to their excellent performance of feature extraction [7]. The transform, namely the decomposition of image, is considered as an important analytical tool that has great effects on the extraction of information and the quality of the fusion results. There are numerous transform methods have been presented including contourlet transform (CT) [8], Laplacian pyramid (LP) transform [9], shearlet transform(ST) [10] and so on. Nevertheless, due to the current level of subband image obtained by the subsampling from these transforms is halved, fusion methods based on them fail to preserve the shift-invariance. To address this problem, the nonsubsampling schemes including nonsubsampling contour transform(NSCT) [11] and nonsubsampling shearlet transform (NSST) [12] are proposed. The nonsubsampling technique can well preserve the shift-invariance property of the decomposition but the fusion strategies adopted are very simple (either average strategy or maximum strategy), which limit the performance to some extent. In recent years, more effective MST-based medical image fusion approaches are proposed by developing more complicated fusion strategies. Sparse representation(SR) [13] and pulse-coupled neural network (PCNN) [14] are two popular fusion strategies used in medical image fusion. SR-based fusion algorithm can accurate describe and reconstruct signal by a linear combination of sparse coefficients. PCNN is a kind of neural network proposed by Eckhorn, which is derived from the cortical model and owns properties of global coupling and pulse synchronization [15]. To completely present the information of the source images, numerous algorithms have been proposed by combining the aforementioned transforms methods and fusion strategies. For example, Xia *et al*. [16] proposed a multimodal medical image fusion method that combined NSCT with SR; Yin *et al*. [17] proposed a medical image fusion method based on NSST and parameter-adaptive PCNN model (NSST-PAPCNN); Zhu *et al*. [18] introduced a medical image fusion algorithm utilizing cartoon-texture decomposition (CTD) and SR to merge the decomposed coefficients(CTD-SR); Li *et al*. [7] introduced a multimodal medical image fusion algorithm based on Laplacian redecomposition (LRD). Although these algorithms can extract more salient features of the source images, several drawbacks of them can also be identified [19]: (1) time-consuming; (2) prone to decrease the contrast of images; (3) sensitive to misregistration and noise.

Researches [20] [21] [22] show that merging regions is further significant than pixels, because regional image processing is more in line with the human visual system and computer vision task as the relationship between pixels are sufficiently considered in it. Compared to pixel-based fusion algorithms, different advantages can be found in region-based fusion methods such as more stable to the noise, better maintain the contrast of the source images, and more efficient due to the reduction of processing units [23]. Basically, the procedure of region-based algorithm includes two steps : (1) segment source images into regions; (2) select the most important regions by considering their properties to construct the fused image. Therefore, accurate image segmentation plays a remarkable role in the performance of medical image fusion. A region-based fusion scheme is initially introduced by Lewis *et al*. in [24], where a dual-tree complex wavelet transform (DT-CWT) [25] is utilized to segment the source images and features of each region are extracted to fuse images region by region. Garg *et al*. [26]presented a region-based medical image fusion algorithm utilizing an evolution algorithm to segment medical images. Luo *et al*. [27] applied the watershed algorithm to segment images into regions for fusion. However, the fusion results of these methods suffer from different degrees of artifacts due to the segmentation algorithms adopted is not precise enough. To address this problem, Normalized cuts(Ncuts) are employed in medical image fusion algorithms [23] [28] to segment images and get homogeneous regions. But it is time-consuming as the conventional eigen-based Ncuts is of high computational complexity. To get a better trade-off between efficiency and accuracy, Meher *et al*. [29] introduced an image fusion method which employed fuzzy c-means (FCM) clustering to segment image. Similarly, Li *et al*. [30] proposed an image fusion algorithm applying entropy rate(ER) superpixel segmentation and get good performance. Nevertheless, the regions segmented by FCM and ER are irregular in shapes and sizes and it is not suitable for the feature extraction of medical images fusion.

Based on the above discussion, it is clear that segmentation is an important factor that determines the fusion results of region-based medical image fusion algorithm. In addition,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

the fusion strategy should be well designed to ensure the most important regions can be correctly selected to construct the final images. As we know, the fusion strategies adopted by pixel-level medical image fuison methods are not suitable for the region-based medical image fusion methods as they work at the pixel level and they are time-consuming as well. Focusing on these two problems, this paper presents a region-based multimodal medical image fusion method that utilizes an effective superpixel segmentation algorithm and a fusion strategy based on feature extraction and optimization algorithm. The main contributions can be summarized as follows.

1) A novel framework for regional image fusion based on fast Linear spectral clustering (LSC) superpixel segmentation is shown in this paper (the general architecture is shown in Fig.1). To the best of our knowledge, this is the first trial to employ superpixel in medical image fusion. Different from conventional region-based medical image fusion methods, the proposed LSC based method makes achievements in preserving the local and spatial information of the source images and get better trade-offs among efficiency, accuracy and fine structure.
2) We present a region-competition-based fusion strategy. Log-Gabor filter(LGF) and Sum Modified Laplacian(SML) are modified to calculate the texture feature value and contrast feature value of each region. The decision map is constructed according to the comparison of these two values for each region. In this way, the important information can be preserved and redundant information can be removed appropriately.
3) In the framework, a new post-processing optimized method based on genetic algorithm(GA) is proposed to optimize the fusion strategy by adaptively adjusted the weights of features. The application of this GA-based post-processing further improves the quality of the fusion results as the importance of the regions is more accruately measured.

The rest of this paper is organized as follows. Section II introduces the related works on superpixel and optimization algorithms. Section III provides a detailed introduction of the proposed region-based medical image fusion method with superpixel segmentation and genetic algorithm. Experimental results and performance evaluation are shown in Section IV. Finally, the conclusion is drawn in Section V and the future work is also discussed.

II. RELATED WORK

A. SEGMENTATION ALGORITHM

As mentioned in the introduction section, accurate image segmentation determines the performance of region-based medical image fusion, because incorrect partition usually leads to unexpected artifacts in fusion results. Convolutional neural networks (CNNs) have been widely used in automatic medical image segmentation in recent years [31] [32]. For instance, Xie *et al*. [33] proposed a medical image segmentation method based on dynamic adaptive residual network

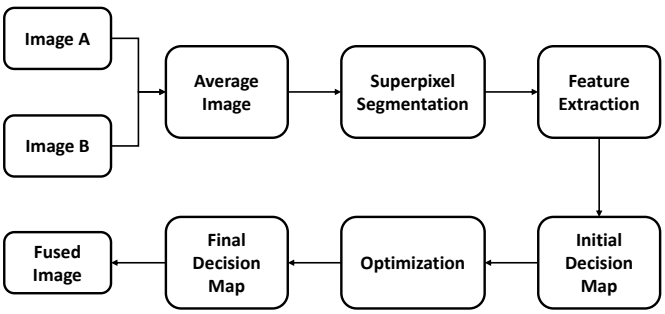

FIGURE 1. The general architecture of the proposed method in this paper.

(DAR-net) to get accuracy segmentation. Xi *et al*. [34] put forward an automated segmentation method for nuclei working with sparse reconstruction and deep convolutional networks. However, these methods are designed to segment the region of interest in medical images (such as the lesion area), which is not suitable for regional image fusion where the homogeneous regions are needed. Superpixel is a technique which not only can be used as an atomic unit for image processing, but also achieves the best trade-off between good performance and high efficiency [35]. Jia *et al*. [36] proposed an effective superpixel-based feature extraction for hyperspectral image fusion, which improves the efficiency for image fusion as the feature extraction is performed on superpixels instead of pixels of the image. Wu *et al*. [37] put forward a superpixel regions extraction for object detection to solve the problems of redundant information and time-consuming in target searching. Zhang *et al*. [38] presented a superpixel-based edge detection in which the clustering-based superpixel methods are centroid updated to improve the accuracy and enhance the robustness. In order to address the wrong selection of similar pixels in remote sensing image fusion, Wang *et al*. [39] applied a superpixel segmentation algorithm to ensure that pixels in the same block have similar properties. These researches prove that superpixels can be flexibly used for many applications of image processing to improve efficiency, accuracy, and robustness.

Although there is no case where superpixel algorithm has been employed in medical image fusion at present, we can find it has been used in other fields of medical images processing. For instance, Achanta *et al*. [40] introduced a simple linear iterative clustering(SLIC) superpixel segmentation which was first applied in medical image segmentation and achieved excellent performance. Wang *et al*. [41] combined SLIC with U-net architecture to segment the lesion area of tuberculosis. However, simple and time-efficient SLIC fails to get a better trade-off between homogeneous regions and fine structures, which could lead to mis-segmentation and generate artifacts in the final medical fusion result. In recent years, many superpixel algorithms have been proposed including TURBO [42], ERS [43], SEEDS [44], and LSC [45]. LSC is an algorithm simply applying K-means clustering in the combined ten dimensional color and coordinate space.

Compared to other superpixel segmentation algorithms, LSC overcomes the shortcoming of SLIC while maintaining simplicity and efficiency. Therefore, LSC is considered to be used in this paper for segmentation.

B. OPTIMIZATION ALGORITHM

Intelligent optimization algorithms such as gray wolf optimization (GWO), modified central force optimization(MCFO), particle swarm optimization (PSO), and genetic algorithms(GA) have been performed effectively in medical image fusion where some kinds of optimizations for parameters are required. Asha *et al*. [46] proposed a medical image fusion method which can adaptively adjust the weights of features by using GWO to minimize the distance between the fused image and the source images. Jin *et al*. [47] proposed an effective image fusion method based on simplified PCNN (S-PCNN) which uses PSO algorithm to set the parameters of PCNN. Hoseny *et al*. [48] introduced an optimal solution for medical image fusion by means of utilizing the MCFO technique to set parameters and improve the quality of medical fused images. Xie *et al*. [49] applied GA to optimize the objective function for a proper image. Compared to traditional optimization algorithms, intelligent optimization algorithms require a more relaxed expression of the objective functions and pay more attention to the speed and efficiency of computation. In this paper, we employ GA to optimize the fusion process as it is the most accurate, and has the best stability among all optimization methods [50].

III. PROPOSED METHOD

The architecture of our proposed fusion method is shown in Fig.2. The proposed method is suitable for the fusion problems of more than two images. Here, the fusion of two medical images is taken as an example. First of all, two source images are weighted averaged to obtain the average image. Then, the average image is segmented by the LSC superpixel segmentation algorithm to obtain the label of the superpixels. And the label is applied to both source images so that the regions of the source images can keep consistent. Subsequently, Log-Gabor filter(LGF) and Sum Modified Laplacian (SML) are used to extract texture feature and contrast feature of each region, respectively. And the decision map is generated by comparing the value of the weighted strategy applied to these two features. The weighting factor is iteratively adjusted by a genetic algorithm. Finally, the fused image can be obtained according to the decision map. In the proposed fusion method, the average image can be obtained by Eq.1, where I_1, I_2 denote the source images, respectively.

$$Avg = (I_1 + I_2)/2 \tag{1}$$

A. LINEAR SPECTRAL CLUSTERING

In medical image fusion, the accuracy and efficiency of segmentation algorithms are of great importance to the performance of region-based image fusion. The linear spectral clustering (LSC) algorithm produces superpixels with the best boundary adhesion in only linear time, which appropriately solves these bottlenecks of image fusion.

For an $M \times N$ medical image I, we define an extensibility mapping to map I to the CIELab color space. In CIELab color space, the value of an pixel $p = (x, y)$ is determined by brightness l and color contrast α and β. These three components are combined with X-Y coordinates to obtain a five dimensional vector (x, y, l, α, β) for each pixel. According to the methodology of LSC, in a well designed ten dimensional feature space, we can simply use weighted K-means clustering to replace the complex operations in Normalized cuts when the Eq.(2) is satisfied.

$$\begin{aligned} \forall p, q \in V, \quad & \Phi(p)\Phi(q) = \frac{D(p,q)}{d(p)d(q)} \\ \forall p \in V, \quad & d(p) = \sum_{q \in V} D(p,q) \end{aligned} \tag{2}$$

Here, each pixel p is assigned with a weight $d(p)$; $D(p, q)$ stands for the similarity between two pixels p and q, and Φ represents the map function which maps the pixel to higher dimensional feature space for improving the linear separability.

In order to measure the similarity between pixels, we first consider widely used Euclidean distance. For any two pixels $p = (x_p, y_p, l_p, \alpha_p, \beta_p)$ and $q = (x_q, y_q, l_q, \alpha_q, \beta_q)$, the formula for measuring their similarity is:

$$\begin{aligned} D(p,q) &= (\frac{d_{xy}}{N_{xy}})^2 + (\frac{d_{lab}}{N_{lab}})^2 \\ d_{xy} &= 1 - (x_p - x_q)^2 + 1 - (y_p - y_q)^2 \\ d_{lab} &= 1 - (l_p - l_q)^2 + 2.55^2[2 - (\alpha_p - \alpha_q)^2 \\ &\quad - (\beta_p - \beta_q)^2] \end{aligned} \tag{3}$$

where d_{xy} and d_{lab} represent the Euclidean distance and the color difference between two pixels, respectively. N_{xy} and N_{lab} are constants that balance the relative importance between color similarity and spatial proximity. A smaller value of N_{xy} or N_{lab} means the more important the corresponding feature is.

Although Eq.(3) has a very specific physical meaning in measuring the similarity of pixels, it can not be used in LSC because it does not satisfy Eq.(2). It can be seen that d_{xy} and d_{lab} both have the form of $1 - u^2, u \in [-1, 1]$, so Eq.(3) can be adapted to Eq.(4).

$$\begin{aligned} D(p,q) =& \frac{1}{N_{xy}^2}[t(x_p - x_q) + t(y_p - y_q)] + \\ & \frac{1}{N_{lab}^2}\{t(l_p - l_q) + 2.55^2[t(\alpha_p - \alpha_q) \\ & + t(\beta_p - \beta_q)]\} \\ t(u) =& 1 - u^2, \qquad u \in [-1, 1] \end{aligned} \tag{4}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

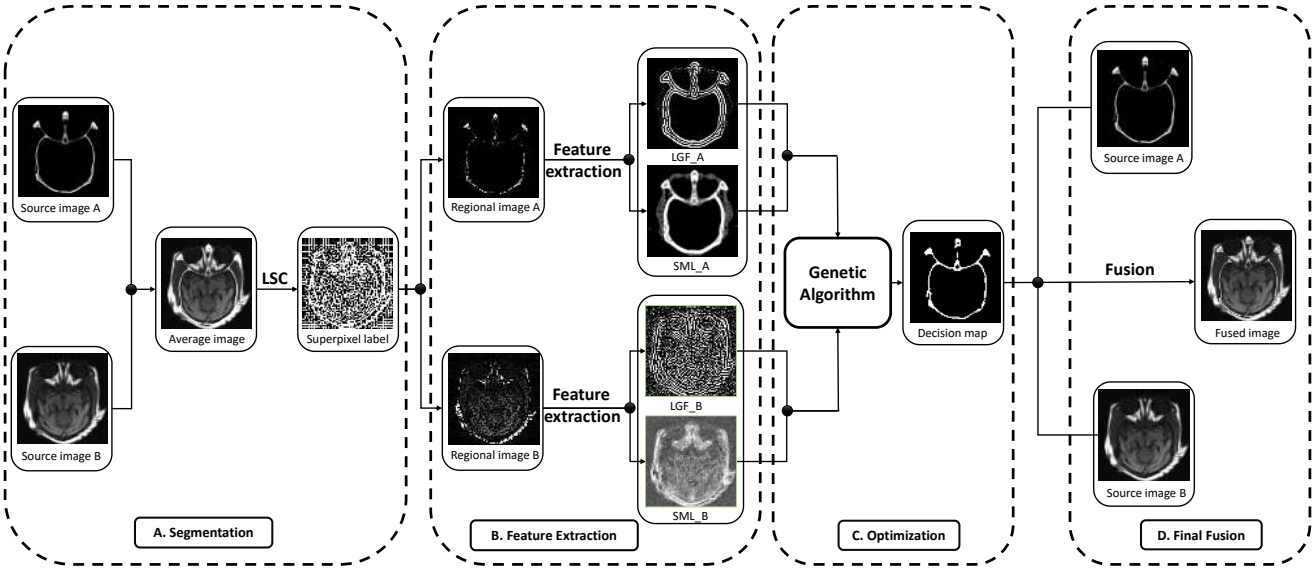

FIGURE 2. The specific architecture of our proposed method in this paper. Part A is the segmentation where the LSC algorithm is used to get superpixel labels. Part B is the feature extraction where LGF and SML are used to get the texture and contrast features of medical images. Part C is an optimization process where GA is used to optimize the fusion process. And Part D is the final fusion.

Furthermore, the fourier series of $t(u)$ is close to $\cos\frac{\pi}{2}u$ in mathematics. In this respect, Eq.(4) is rewritten to Eq.(5).

$$\begin{aligned} D(p,q) =& \frac{1}{N_{xy}^2}[\cos\frac{\pi}{2}(x_p - x_q) + \cos\frac{\pi}{2}(y_p - y_q)] + \\ & \frac{1}{N_{lab}^2}\{\cos\frac{\pi}{2}(l_p - l_q) + 2.55^2[\cos\frac{\pi}{2}(\alpha_p - \alpha_q) \\ & + \cos\frac{\pi}{2}(\beta_p - \beta_q)]\} \end{aligned} \tag{5}$$

Combining the above derivation, the map function Φ is defined as Eq.(6).

$$\begin{aligned} \Phi(p) =& \frac{1}{d(p)}(\frac{1}{N_{lab}}\cos\frac{\pi}{2}l_p, \frac{1}{N_{lab}}\sin\frac{\pi}{2}l_p, \frac{2.55}{N_{lab}}\cos\frac{\pi}{2}\alpha_p, \\ & \frac{2.55}{N_{lab}}\sin\frac{\pi}{2}\alpha_p, \frac{2.55}{N_{lab}}\cos\frac{\pi}{2}\beta_p, \frac{2.55}{N_{lab}}\sin\frac{\pi}{2}\beta_p, \\ & \frac{1}{N_{xy}}\cos\frac{\pi}{2}x_p, \frac{1}{N_{xy}}\sin\frac{\pi}{2}x_p, \frac{1}{N_{xy}}\cos\frac{\pi}{2}y_p, \\ & \frac{1}{N_{xy}}\sin\frac{\pi}{2}y_p) \end{aligned} \tag{6}$$

Above all, the mapping function to a ten dimensional feature space has been designed as Eq.(6). Each pixel of the medical images is mapped using Eq.(6) so that we can simply use K-means clustering to achieve the same or even better segmentation performance of Normalized cuts, which decreases the risk of generating artifacts in pixel-level and the conventional region-based image fusion methods.

B. FEATURES EXTRACTION

1) Superpixel-based Log-Gabor Filters

Although the two-dimensional Gabor filter has good local properties in both the spatial and frequency domains, its even-symmetric filter produces a nonzero DC component when the bandwidth is greater than one times the frequency; the Log-Gabor function is unrestricted in terms of bandwidth and has minimal spatial support. The Log-Gabor function is a Gaussian function on a logarithmic frequency scale, and on a linear frequency scale, the Log-Gabor function is expressed as:

$$LG(f) = \exp(\frac{-[\log(\frac{f}{f_0})]^2}{2[\log(\frac{\beta}{f_0})]}) \tag{7}$$

In Eq.(7), f_0 is the filter center frequency; β is used to determine the radial bandwidth.

According to frequency domain analysis, the two-dimensional Log-Gabor filter is a band-pass filter in a specific direction. For more comprehensive feature extraction, we use multi-channel Log-Gabor filters with different frequency and directions to extract texture features, the specific steps are as follows.

i The image is first filtered by the Log-Gabor filter. u frequency scales and v directions are selected for each channel, and the features of the medical image are extracted using Eq.(8).

$$F_{uv} = LG_{uv} \cdot I(N) \tag{8}$$

where LG_{uv} is the Log-Gabor filter; $I(N)$ is a medical image divided into N regions by superpixels; F_{uv} is the feature extracted.

ii LSC Superpixel segmentation divides a medical image into several regions of uniform size, and calculates each region using Eq.(9).

$$F_{uv}(x,y) = \sum_{n}^{N} I(x+n-\frac{2}{N}, y+n-\frac{2}{N}) \cdot LG_{uv} \tag{9}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

where (x, y) is the center coordinate of each region, and $F_{uv}(x, y)$ is the final texture feature extracted by the Log-Gabor filter.

2) Superpixel-based Sum Modified Laplacian

Sum Modified Laplacian(SML) is a fundamental feature extraction operator widely used in image processing. In this section, we make some corrections for SML at the superpixel level so that we can use it to calculate the contrast feature value of each region in the source image.

In [51], SML calculates the sum of the absolute values of the convolution of an image with modified Laplacian operators(ML), whose expression for the discrete approximation is shown in Eq.(10). Here, $h(i, j)$ is the Laplacian of the pixel and s denotes a variable space. And SML can be calculated as Eq.(11), where S is the parameter determines the window size used to calculate ML, (x, y) represents the center of the window.

$$\begin{aligned} ML_{(i,j)} =& |2h(i,j) - h(i-s,j) - h(i+s,j)| + \\ & |2h(i,j) - h(i,j-s) - h(i,j+s)| \end{aligned} \tag{10}$$

$$SML(x,y) = \sum_{i=x-S}^{i=x+S} \sum_{y-S}^{y+S} ML_{(i,j)} \tag{11}$$

In order to calculate the value of SML for each superpixel, we modify Eq.(11) to Eq.(12).

$$\begin{aligned} SSML_{(x,y)} &= \sum_{i}^{M} \sum_{j}^{N} \mathbf{I}(p_{(i,j)} \in C_{(x,y)}) ML_{(i,j)} \\ \mathbf{I}(p_{(i,j)} \in C_{(x,y)}) &= \begin{cases} 1, & \text{if } p_{(i,j)} \in C_{(x,y)} \\ 0, & \text{if } p_{(i,j)} \notin C_{(x,y)} \end{cases} \end{aligned} \tag{12}$$

Where SSML denotes the SML value of a superpixel whose center is located at (x, y), M, N is the shape of the medical image, $p(i, j)$ denotes the pixel of the image, $C_{(x,y)}$ represents a collection of pixels belonging to a superpixel whose center locates at (x, y), $\mathbf{I}(\cdot)$ is an indicator function when a pixel belongs to a superpixel equal to 1, otherwise equal to 0. Therefore, contrast features of every superpixel region can be obtained from Eq.(12).

C. GENETIC ALGORITHM AND DECISION MAP

In this section, a weighted strategy and an optimization method are used to generate the decision map.

For each superpixel at the same location in two source medical images, the weighted strategy merges the texture and contrast features into a value named hybrid feature(HF). And we select superpixels with larger HF values to generate the decision map. The detail is shown in Eq.(13)-(14).

$$HF(x,y) = \beta \cdot F_{uv}(x,y) + (1-\beta) \cdot SSML(x,y) \tag{13}$$

$$DM(x,y) = \begin{cases} 1, & \text{if } HF(x,y)_{I_1} \geq HF(x,y)_{I_2} \\ 0, & \text{if } HF(x,y)_{I_1} < HF(x,y)_{I_2} \end{cases} \tag{14}$$

Here, (x, y) means the superpixel whose canter locate at (x, y), HF means the hybrid feature obtained by the weighted average of texture featureF_{uv} and contrast feature$SSML$, β is the factor used to control the weight of the features. $HF(x, y)_A$ and $HF(x, y)_B$ represent the HF value of the superpixel (x, y) of source medical image I_1 and I_2, DM is the decision map.

In the proposed fusion framework, other features of the medical images that we consider important can be embedded in the fusion process flexibly. In this case, Eq.(13) can be rewritten as Eq.(15).

$$\begin{aligned} HF(x,y) &= \sum_{i=1}^{n} \beta_i \cdot F_i(x,y) \\ s.t. \quad & \sum_{i=1}^{n} \beta_i = 1 \end{aligned} \tag{15}$$

Here, $F_i(x, y)$ represents the value of feature i of the region centered at (x, y); β_i is the corresponding weight of feature i; $HF(x, y)$ means the weighted sum of all features in the region centered at (x, y). Similarly, all weights can be optimized by genetic algorithm. However, more features involved in fusion also pose problems. On the one hand, genetic algorithms need to spend more time to find the optimal solution for all the feature weights; on the other hand, too many features complicate the process of generating a decision map and increase the probability of region misselection. In this paper, appropriate quantities of image features are involved in the generation of the decision map. In this way, important information of medical images can be detected and maintained in the fusion result while the low computational complexity is achieved in the algorithm.

In the weighted strategy described as Eq.(14), the weighting factor β is related to whether the algorithm can select the superpixel blocks needed for fusion. Since different medical images have different salient features, it is necessary to find the right value of β to maximize the HF value of the superpixel blocks with important information. To address this problem, the genetic algorithm(GA), which is an adaptive probabilistic search technique based on the mechanism of natural selection and natural genetics [52] is adopted. The traditional GA mainly consists of these elements: chromosomes, population, generations, crossover probability, mutation probability, and fitness function. Chromosomes are the individuals in the population, which represent the solutions to the problem at hand. In each generation, chromosomes are first selected by a well-designed fitness function, and then the selected excellent chromosomes crossover and mutate at certain probabilities to form the next generation. Throughout the evolutionary process, the design of the fitness function plays an important role in obtaining the optimal solution. And in this paper, the fitness function designed for the fusion problem is shown as Eq.(16).

$$\begin{aligned} F(i) &= RMSE(DM_i, I_1) \\ RMSE(A,B) &= \sqrt{\frac{1}{N} \sum_{i=1}^{N} (A_i - B_i)^2} \end{aligned} \tag{16}$$

Here, the root mean square error(RMSE) is used to measure the similarity between the decision map and the source image, i is the individual within the population in each generation, N denotes the size of the image. The smaller the fitness function is, the more similar the decision map is to the source image, which means more information about the source image is contained in decision map.

Through the optimization of genetic algorithm, the best weighting factor can be figured out and Eq.(14) can most accurately measure the importance of each region in the medical image. Finally, the fused image can be obtained according to the decision map by Eq.(17).

$$F = DM \cdot I_1 + (1 - DM) \cdot I_2 \tag{17}$$

Where F represents the final fused image; DM is the decision map obtained by Eq.(14); I_1, I_2 denotes the source images, respectively. By applying genetic algorithm to optimize the weighting factor, not only the cumbersome and inaccuracy of manual definition of weight factor are avoided, but also the accuracy of medical image fusion is further improved. More details of the proposed multimodal medical image fusion method is shown in **Algorithm 1**.

Algorithm 1 The proposed Medical image fusion method.

Input: a pair of source medical images I_1 and I_2
Output: The fused image F

1: Weighted averaging two source images to get an average map
2: Using linear spectral clustering to segment the average image and generate superpixels
3: Set generation $G = 10$, individual number $N = 10$, crossover probability $P_c = 0.5$, mutation probability $P_m = 0.1$
4: **for** generation $g = 1 : G$ **do**
5: **for** each superpixel number $k = 1 : K$ **do**
6: Calculate the texture feature and contrast feature with Eqs.(9) and (12)
7: Calculate the HF value of the superpixels and compare to obtain decision map according to Eq.(14)
8: **end for**
9: Calculate the fitness F(i) for each individual in the population according to Eq.(16).
10: **for** individual number $n = 1 : N$ **do**
11: Perform crossover operation on two individuals with crossover probability P_c
12: Perform mutation operation on two individuals with mutation probability P_m
13: **end for**
14: Updating individual populations
15: **end for**
16: Fuse source medical images using decision map.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the setups of experiments are introduced firstly. To verify the performance of our proposed method, it is compared with 7 state-of-the-art medical image fusion methods on eight pairs of multimodal medical images.

A. DATASET AND SETUPS OF EXPERIMENTS

Medical images from different modalities including Computerized Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission tomography (PET), and Single Photon Emission Computed Tomography (SPECT) are used for the diagnosis of various diseases. In the experiments, eight pairs of multimodal medical images are utilized, including three pairs of CT-MRI medical images, two pairs of MRI-SPECT medical images, and three pairs of MRI-PET medical images(see Fig.3). The proposed method is compared to seven other state-of-the-art medical image fusion methods: LP-SR [53], CNN [47], CFL [54], NSST-PAPCNN [17], NSCT-PC-LLE [55], LRD [7] and NSCT-SR [56]. The resolution of all medical images in the experiments was set to 256*256 and all the experiments were conducted in Matlab 2018a.

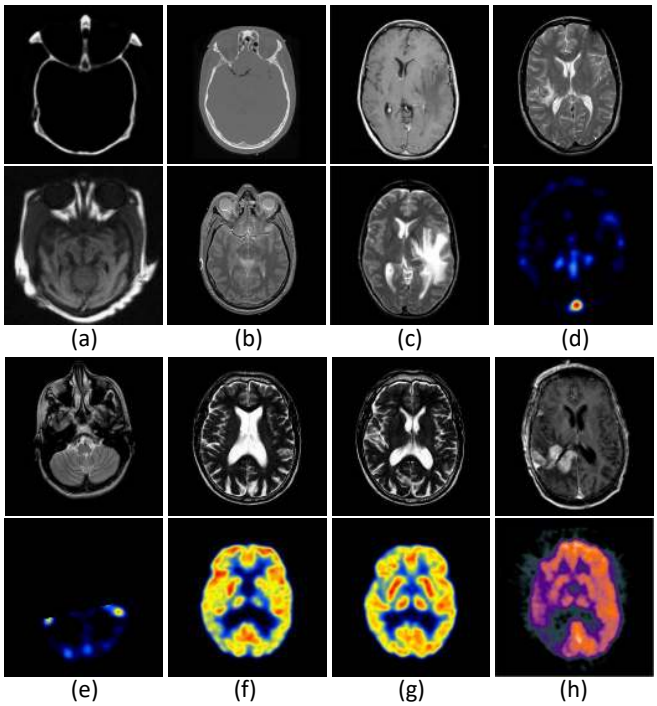

FIGURE 3. Source images for experiments. (a), (b) ,(c)are CT-MRI image pairs; (d), (e)are MRI-SPECT image pairs; (f), (g),(h) are MRI-PET image pairs.

B. FUSION RESULTS

In the experiment, both subjective quality and objective metrics of the fusion results were evaluated.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

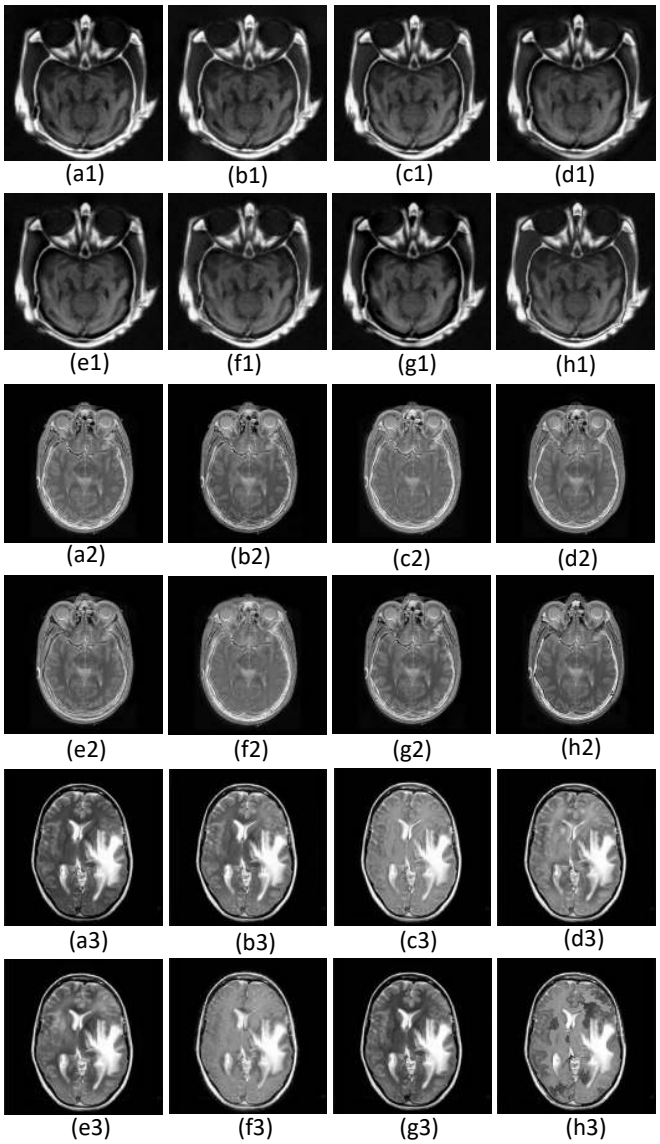

FIGURE 4. Three pairs of CT-MRI medical image fusion results of eight different image fusion methods. Fused images (a1), (a2), (a3) is LP-SR based method; (b1), (b2), (b3) CNN based method; (c1),(c2),(c3) CFL based method; (d1), (d2), (d3) NSST-PAPCNN based method; (e1), (e2), (e3) NSCT-PC-LLE based method; (f1), (f2), (f3) LRD based method; (g1), (g2), (g3) NSCT-SR bAsed method; (h1), (h2), (h3) the proposed method.

1) Qualitative comparison

Fig.4 shows the fusion results of three sets of CT-MRI images. Generally speaking, all methods can retain complementary information of the source images in the fusion results. However, the fusion results of LRD and NSCT-PC-LLE are not clear enough at the boundary, it can be seen that there are ambiguities and adhesions at the edges of the tissues (see (d1),(e1) and (f1) in Fig.4). The fusion results of LP-SR and NSCT-SR lose a large amount of energy information, which leads to a serious reduction of the contrast and intensity of many regions(see (a1),(a3),(g1) and (g3) in Fig.4). The CFL and NSST-PAPCNN methods can preserve the image energy well, but some redundant information is not completely removed from CT image and some artifacts are produced (see (b2) and (c2) in Fig.4). The CNN method and the proposed method perform better in this experiment, edge details and the salient features of the source medical images are preserved well in the fused images of these two methods. But compared to the CNN method, the fusion results of the proposed method are more refined in detail (see (b1), (b2), (b3), (h1), (h2), and (h3) in Fig.4).

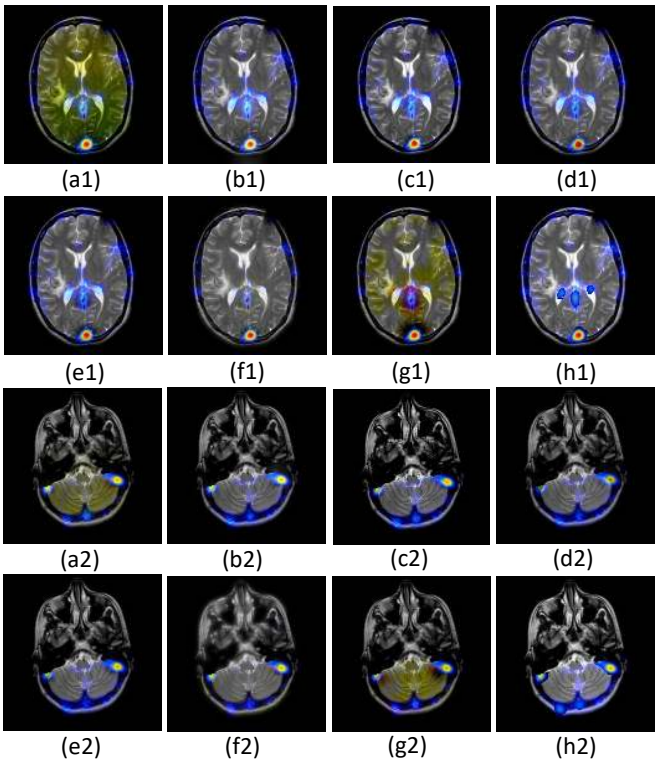

FIGURE 5. Two pairs of MRI-SPECT fusion results of eight different fusion methods.(a1),(a2) LP-SR method;(b1),(b2) CNN method (c1),(c2) CFL method;(d1),(d2) NSST-PAPCNN method;(e1),(e2)NSCT-PC-LLE; (f1),(f2) LRD; (G1),(G2) NSCT-SR; (h1),(h2) the proposed method.

Fig.5 shows the fusion results of two sets of MRI and SPECT images. Compared to the fusion results of CT and MRI images, it is obvious that LP-SR and NSCT-SR fail to keep good performance in color medical image fusion. Color artifacts appear in the fusion results of these two methods, and a large area of green appears in the fused images (see (a1),(a2), (g1) and (g2) in Fig.5). The LRD method performs better on preserving the original color of the images, but the fusion results are blurred in the boundaries and many details. In addition, some information of the SPECT images is weakened or even lost after fusion (see Fig.5 (f1) and (f2)). The fusion results of NSST-PA-CNN and NSCT-PC-LLE preserve most details of the source images, but the energy seems to be lost and the contrast of the fused images decreases (see Fig.5 (d1),(d2), (e1) and (e2)). The CNN and CFL methods generally perform well, but the fusion results of CNN still suffer from slight color distortion(see Fig.5 (b1) and (b2)), and the fusion results of CFL lose some detailed information of SPECT images (see Fig.5 (c1),(c2)). The pro-

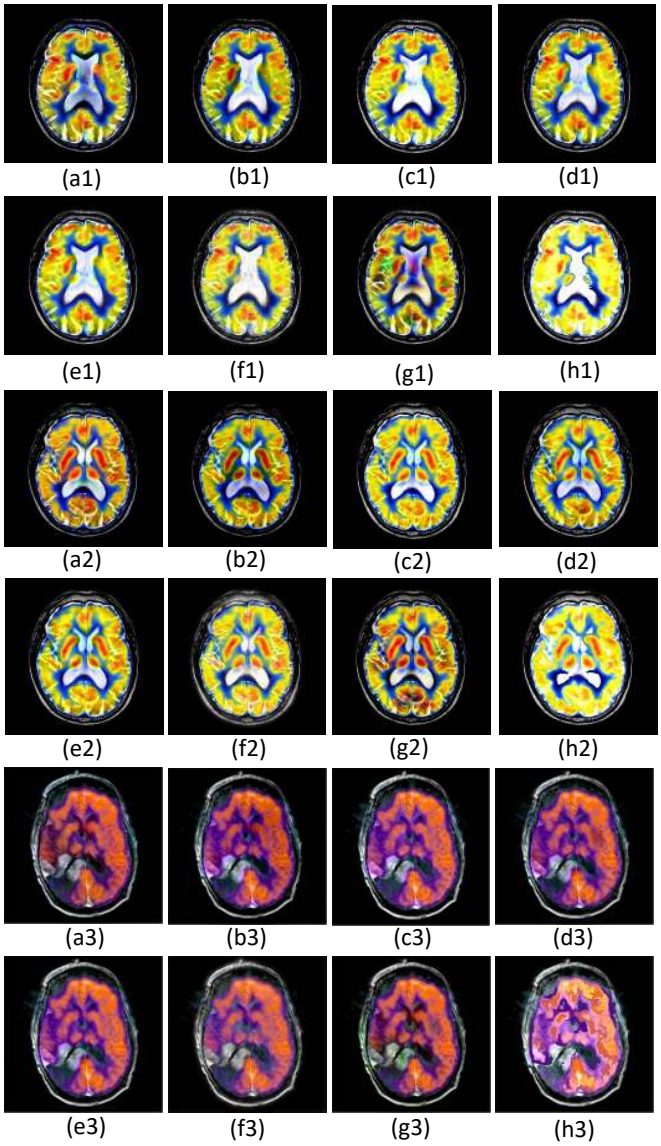

FIGURE 6. Results of three pairs of MRI-PET images fused by eight different image fusion methods. Fused images (a1),(a2),(a3) are LP-SR based method; (b1), (b2), (b3) CNN based method; (c1), (c2), (c3) CFL based method; (d1), (d2), (d3) NSST-PAPCNN based method; (e1), (e2), (e3)NSCT-PC-LLE based method; (f1), (f2), (f3) LRD based method; (g1), (g2), (g3) NSCT-SR bsed method; (h1), (h2), (h3) the proposed method.

posed method performs well in retaining the brightness and contrast of the source medical images. Besides, the clarity of the tissue texture and the edges of the medical images are preserved well in the fusion results (see Fig.5 (h1) and (h2)). But the integration of important information is not perfect enough.

Fig.6 shows the fusion results of three sets of MRI and PET images. Obviously, the fusion results obtained by the LP-SR and NSCT-SR methods have a certain degree of distortion, and the contrast of fused images is much lower because of the loss of energy. The white area of bone in the middle of the MRI image has chaotic colors (see Fig.6 (a1) and (g1)). The other group of images fused by the LP-SR and NSCT-SR methods also have a similar issue (see Fig.6 (a2),(a3),(g2) and (g3)). Such a situation is not conducive to clinical diagnosis. The CNN method suffers from slight contrast distortion. In addition, the outer boundary of the organ is lost mostly in the fusion image (see Fig.6 (b1) and (b2)). The NSCT-PC-LLE and LRD methods perform well in preserving colors and edges, but the contrast of the fusion results of LRD is relatively low compared to that of the NSCT-PC-LLE method, and the results of NSCT-PC-LLE also lose the clarity to some extent (see Fig.6 (e1),(e2),(e3),(f1),(f2) and (f3)). The CFL and NSST-PAPCNN methods perform better in this issue, but color distortion is still exit in the fusion results of NSST-PAPCNN(see Fig.6 (d1),(d2)). The proposed method has a good performance in retaining good information in detail, but the brightness is enhanced slightly (see Fig.6 (h1),(h2), and (h3)).

In order to further verify the performance of the proposed region-based fusion method, two representative examples (CT-MRI and MRI-SPECT) chosen from the eight groups of fusion images are shown in greater detail in Fig.7. As can be seen from the enlarged regions with red borders, in group 1(Fig.7 (a1)-(h1)), the fusion results of LP-SR, CNN, CFL, NSST-PAPCNN, and LRD have the problems of blurring and adhesion in the gap between the white regions. The fusion results of NSCT-PC-LLE and NSCT-SR are clear at the gap, but not dense enough. The proposed method makes the fusion gap fit exactly, which is clear and compact. In group 2 ((a2)-(h2)), the fusion results of the LR-SR, LRD ,and NSCT-SR methods have different degrees of color artifact. The CNN, CFL, NSST-PAPCNN, and NSCT-PC-LLE methods obtained good fusion results, and the proposed method enhances the information of the boundary.

After the above comparative analysis of subjective visual effects, we can see that the proposed method performs satisfactorily in all types of multimodal medical image fusion. It is also evident that the proposed method can precisely retain the important salient features of the source images without producing abnormal details.

2) Quantitative comparison

In this paper, four objective metrics are applied to the evaluation of fusion performance of different methods: average gradient (AG), spatial frequency (SF), mutual information (MI), and petrovic metric $Q^{AB/F}$.

a) Average gradient (AG) [57] refers to the obvious difference in the grayscale near the border point or both sides of the shadow line of the image, which can be used to indicate the clear point of the image. Generally speaking, the larger is the value of AG , the better is the fusion result.

b) Spatial frequency (SF) [58] is used for calculating the general activity level of the space in the image. The value of SF is larger, the riche information is contained in fused image, and the better the fusion method has performed.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

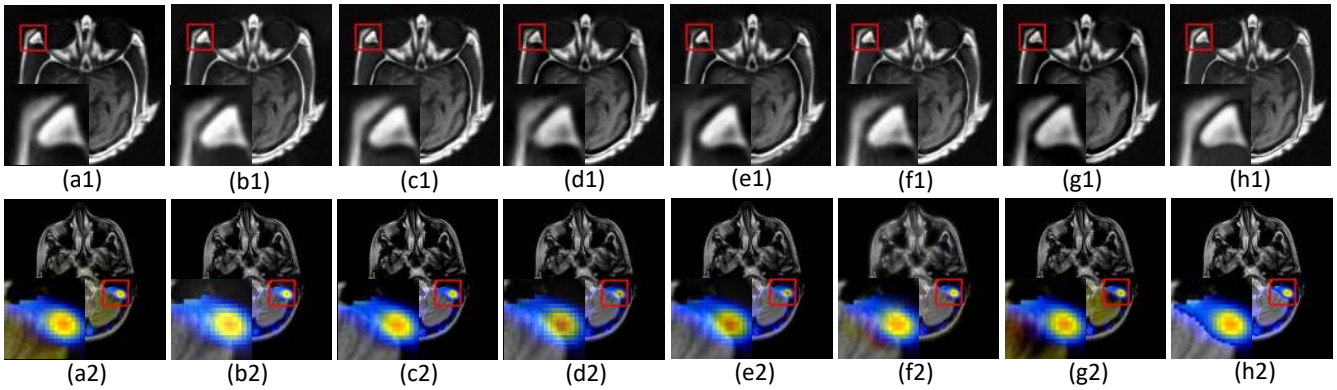

FIGURE 7. The cropping results of three types of fusion images : (a1)-(h1) CT-MRI fused images;(a2)-(h2) MRI-SPECT fused images;(a3)-(h2) MRI-PET fused images. Fused images (a1),(a2)is LP-SR based method;(b1),(b2)CNN based method;(c1),(c2)CFL based method;(d1),(d2) NSST-PAPCNN based method;(e1),(e2) NSCT-PC-LLE based method;(f1),(f2) LRD based method;(g1),(g2) NSCT-SR bsed method;(h1),(h2) the proposed method.

TABLE 1. Average Gradient values for 8 fused images obtained by different fusion methods.

	LP-SR	CNN	NSST-PAPCNN	NSCT-PC-LLE	LRD	NSCT-SR	CFL	Proposed
CT-MRI1	0.0490	0.0489	0.0468	0.0477	0.0457	0.0491	0.0469	**0.0617**
CT-MRI2	0.0580	0.0576	0.0532	0.0526	0.0517	0.0557	0.0561	**0.0682**
CT-MRI3	0.0674	0.0696	0.0675	0.0677	0.0640	0.0682	0.0684	**0.0786**
MRI-SPECT1	0.0535	0.0525	0.0537	0.0537	0.0525	0.0543	0.0587	**0.0688**
MRI-SPECT2	0.0509	0.0499	0.0494	0.0496	0.0402	0.0497	0.0549	**0.0608**
MRI-PET1	0.0817	0.0749	0.0831	0.0739	0.0777	0.0830	0.0896	**0.0946**
MRI-PET2	0.0821	0.0767	0.0865	0.0837	0.0732	0.0847	0.0930	**0.0973**
MRI-PET3	0.0579	0.0576	0.0583	0.0585	0.0484	0.0584	0.0625	**0.0731**
Average Value	0.0626	0.0610	0.0623	0.0609	0.0567	0.0629	0.0663	**0.0754**

TABLE 2. Spatial Frequency values for 8 fused images obtained by different fusion methods.

	LP-SR	CNN	NSST-PAPCNN	NSCT-PC-LLE	LRD	NSCT-SR	CFL	Proposed
CT-MRI1	17.6378	17.6211	16.7959	17.1296	16.5588	17.6342	16.8365	**22.1734**
CT-MRI2	20.8408	20.6734	19.1176	18.8774	18.5867	20.0241	20.1539	**24.5032**
CT-MRI3	24.1974	24.9913	24.2460	24.3321	22.9992	24.5103	24.5601	**28.2179**
MRI-SPECT1	19.2043	18.8538	19.2765	19.2982	18.8731	19.5186	21.0741	**24.7296**
MRI-SPECT2	17.9220	17.9290	17.7624	17.8165	17.9660	17.8578	19.7074	**21.8465**
MRI-PET1	29.3496	26.8962	29.8567	29.4604	27.9157	29.8023	32.1839	**33.9840**
MRI-PET2	29.4981	27.5594	31.0627	30.0771	26.3038	30.4162	33.4079	**34.9658**
MRI-PET3	20.8095	20.6875	20.9564	21.0048	17.4038	20.9844	22.4649	**26.2530**
Average Value	22.4324	21.9015	22.3843	22.2495	20.8259	22.5935	23.7986	**27.0842**

TABLE 3. Mutual information values for 8 fused images obtained by different fusion methods.

	LP-SR	CNN	NSST-PAPCNN	NSCT-PC-LLE	LRD	NSCT-SR	CFL	Proposed
CT-MRI1	3.1260	3.0745	2.4653	3.2902	4.2412	3.0118	4.7364	**6.0565**
CT-MRI2	3.6160	3.7149	3.7144	3.6176	3.9179	3.7702	4.1933	**6.2362**
CT-MRI3	3.8292	3.5254	3.3347	3.4408	3.5892	3.5766	3.5297	**5.6162**
MRI-SPECT1	2.4530	2.7272	3.8347	**3.8930**	2.6105	2.8140	3.7160	3.8372
MRI-SPECT2	3.0017	3.2926	3.2974	3.2801	3.0228	2.8855	2.8650	**3.4792**
MRI-PET1	2.8622	3.0971	2.9967	2.9327	3.1686	2.7918	3.1637	**3.9192**
MRI-PET2	2.6469	2.5615	2.6944	2.6295	2.7070	2.5696	2.9032	**3.7737**
MRI-PET3	3.1343	3.0149	3.3484	3.2920	3.4242	3.2791	3.4811	**4.2459**
Average Value	3.0837	3.1260	3.2108	3.2970	3.3352	3.0873	3.5736	**4.6455**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

TABLE 4. $Q^{AB/F}$ values for 8 fused images obtained by different fusion methods.

	LP-SR	CNN	NSST-PAPCNN	NSCT-PC-LLE	LRD	NSCT-SR	CFL	Proposed
CT-MRI1	0.7682	0.7683	0.6859	0.7325	0.7013	0.7329	0.7688	**0.8056**
CT-MRI2	0.6122	0.6028	0.5914	0.5804	0.5000	0.6210	0.5967	**0.6603**
CT-MRI3	0.6427	**0.6588**	0.5975	0.6424	0.5380	0.6491	0.5960	0.6210
MRI-SPECT1	0.8551	0.7702	0.8665	**0.8677**	0.7172	0.8521	0.8374	0.8431
MRI-SPECT2	0.9040	**0.9064**	0.9006	0.9019	0.8908	0.8943	0.8541	0.8992
MRI-PET1	**0.5377**	0.4751	0.5294	0.5311	0.4632	0.5354	0.5161	0.4543
MRI-PET2	0.5525	0.4937	**0.5565**	0.5484	0.4596	0.5483	0.5463	0.4830
MRI-PET3	0.6721	0.6572	0.6492	0.6567	0.5539	**0.6774**	0.6639	0.5761
Average Value	0.7565	0.7413	0.7284	0.7450	0.6694	0.7499	0.7306	**0.7658**

TABLE 5. Running time of different methods when fusing two source images of size 256x256 pixels (Unit: seconds).

	LP-SR	CNN	NSST-PAPCNN	NSCT-PC-LLE	LRD	NSCT-SR	CFL	Proposed
CT-MRI1	**0.1203**	11.4273	4.9216	2.4360	120.9771	8.8443	7.1660	84.9767
MRI-SPECT1	**0.0282**	12.3585	4.6355	4.2349	118.5621	18.1814	5.1511	82.6332
MRI-PET1	**0.0291**	12.5944	4.8624	4.2799	121.1856	23.6890	8.9292	83.9941
Average Value	**0.0503**	12.0531	4.7972	3.5329	121.4462	16.6433	7.9208	83.6364

c) Mutual information (MI) [59] is used to measure the dependence between the source images and the fused image and perfectly indicate the shared information of the fused and the source images. The value of MI is larger, the fused image contains more information about the source image.

d) Petrovic Metric ($Q^{AB/F}$) [60] refers to the edge information transferred from source images to fused image. Higher is the value of $Q^{AB/F}$, better is the quality of the fused image.

The metric scores of LP-SR, CNN, CFL, NSST-PAPCNN, NSCT-PC-LLE, LRD, and the proposed method for CT-MRI, MRI-SPECT, and MRI-PET image fusion are reported in Tables 1 - 4. For each metric, the best score of different methods is shown in bold indicates, and the averages of the performance for all the testing image sets are shown at the bottom of the table to evaluate the fusion results more objectively.

From Tables 1 - 4, we can see that the proposed method achieves the best values of AG and SF in all test examples. Although not all the fusion results of our method achieve the best values of MI, the only value that is not the best is very close to the best one (only 0.0558 smaller than the best one), and the average value of MI in eight groups is still the largest. Table 4 shows $Q^{AB/F}$ values of different fusion methods. It can be seen that the best score of our proposed method are achieved in CT-MRI image set. Because the $Q^{AB/F}$ is a metric for the evaluation of pixel-level image fusion methods, it is not so suitable for the evaluation of region-based image fusion methods and thus not all of our fusion results get good scores. Nonetheless, the average score of our method in QAB/F is still the best among all methods. Therefore, it can be seen from the results of the four objective quantitative evaluations that the proposed method can achieve the state-of-the-art performance and even more effective than the mainstream fusion methods. This is attributed to the accurate segmentation of the LSC algorithm as well as the effective optimization of the decision map.

3) Time complexity analysis

In this section, the computational efficiencies of the different fusion methods are conducted on the image sets "CT-MRI1", "MRI-SPECT1", and "MRI-PET1", and all the tests were implemented on a computer with a 1.99 GHz CPU and 8 GB of RAM. The running time results are compared in Table 5. The average value is the average running time of all fusion pairs in the experiment. As shown in Table 5, the computational complexity of the LRD method is the highest, and CNN and NSCT-SR are also inefficient. The LP-SR method is the most efficient because the Laplacian pyramid it uses to decompose source images for fusion does not cost too much calculation. In our method, the genetic algorithm is used to obtain better fusion results, but it makes us less efficient as it takes most of the running time. However, the linear computational complexity of the LSC algorithm improves our efficiency at the same time.

C. DISCUSSIONS

In the proposed region-based medical image fusion method, the LSC algorithm is utilized to segment source medical images. To further verify the overall performance of LSC is superior to other segmentation algorithms, a comparative experiment is conducted on three classical superpixel segmentation algorithms (Ncuts, Turbo, and SLIC) and LSC. CT and PET images are used for segmentation.

Fig.8 shows the superpixel segmentation results of the four superpixel segmentation methods. It can be seen that all methods produce uniformly sized superpixels, but Turbo and SLIC tend to produce superpixels which contain pixels of different colors(see Fig.8 (b1),(b2),(c1) and (c2)). And the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3094972, IEEE Access

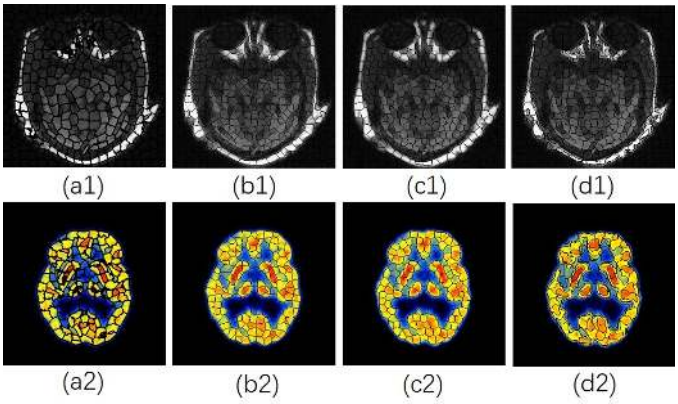

FIGURE 8. Medical images are segmented into 500 superpixels using four superpixel segmentation methods.(a1)-(d1) are segmentation results of CT images; (a2)-(d2) are segmentataion results of PET images; (a1),(a2) is segmented by Ncuts; (b1),(b2) Turbo; (c1),(c2) SLIC; (d1),(d2) LSC .

superpixels produced by Ncuts and LSC is homogeneous(see Fig.8 (a1),(a2),(d1) and (d2)). In terms of objective criteria, edge intensity(EI) is used to measure the edge compactness of the segmentation results and the running time of all segmentation algorithms is also recorded. As is shown in Tabel 6 and Table 7, Ncuts produces the best superpixel edge compactness, and the LSC is in second place. But the computational complexity of Ncuts is much higher than that of LSC (almost 10^2). SLIC has the shortest running time but the quality of segmentation is not as good as the LSC. In summary, although LSC is not the best in some metrics, it achieves the best trade-offs among homogeneous regions, excellent structures, and low computational complexity.

TABLE 6. Edge Intensity of superpixel segmentation results at 500 superpixels

Methods	Ncuts	Turbo	SLIC	LSC
CT	**113.9254**	91.6725	92.1535	100.4443
PET	**105.2669**	77.5464	78.2514	85.9055

TABLE 7. Running time of different superpixel segmentation methods at 500 superpixels

Methods	Ncuts	Turbo	SLIC	LSC
CT	99.6634	1.5381	**0.0889**	0.0939
PET	99.0018	1.7350	**0.0859**	0.0989

In order to make the proposed method achieve the best fusion performance, the appropriate setting of the superpixel number in LSC and population size in GA are of great importance. For example, fewer superpixels mean larger regions, which may lead to an increased probability that the same region contains both required and unrequired information, thereby resulting in an increased probability of region misselection in the fusion process. For the GA, if the population size increases, the likelihood of the algorithm converging to the optimal weights also increases. Therefore, to further investigate the effects of these two parameters on image fusion, we conducted experiments using eight pairs of medical images at different numbers of superpixels and population sizes, respectively. The results are shown in Fig.9 and Fig.10.

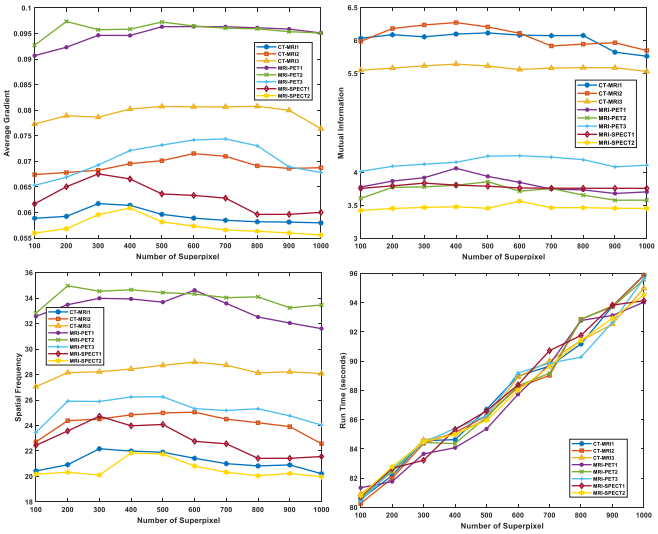

FIGURE 9. Effect of the superpixel numbers on the fusion results of three types of medical images

In Fig.9, the values of AG, SF, and MI increase with the growth of superpixel numbers. The optimal number of superpixels is between 300 and 500. But it does not suggest that the higher the number of superpixels, the better the fusion result. At the stage where the number of superpixels is 800 to 1000, the values of all three evaluation metrics decrease, which is partly due to the fact that too-small region division destroys the structure of the information. Besides, too many regions can lead to low fusion efficiency, which can be seen that the running time of fusion increases linearly with the increase of the number of superpixels.

Fig.10 shows the relationship between the values of AG, SF, MI, and population size, and we can see that most of the fusion results become better as the population size increases. The fusion result remains stable at most population sizes, the larger the population size, the greater the probability of convergence to a more optimal solution. However, the running time also increases when the population size is larger. Therefore, population size between 10 and 20 are optimum.

V. CONCLUSION

In this paper, a novel regional multimodal medical image fusion method based on superpixel segmentation and a post-processing optimized method is proposed. For multimodal medical images, more homogeneous regions can be obtained by the LSC superpixel algorithm. Based on the above regions, Log-Gabor filter and sum modified laplacian are adopted to get texture feature and contrast feature, respectively. Subsequently, a post-processing method based on genetic algorithm is proposed for adaptively adjusting the effects of these

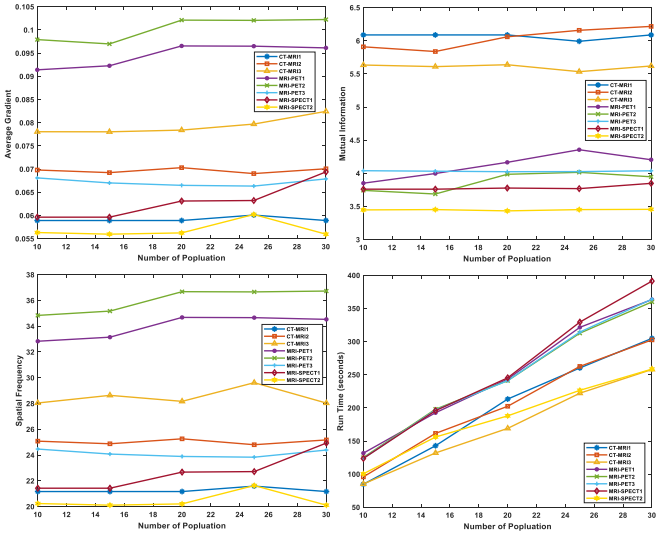

FIGURE 10. Effect of the population size in genetic algorithm on the fusion reuslts of three types of medical images

two features. By comparing the importance information of each superpixel, the final decision map is generated and then the fused image can be obtained. Experiments are conducted on eight groups of multimodal medical images. Compared with seven mainstream fusion methods, the proposed method can achieve better performance in both visual effects and objective evaluation because the segmentation of medical images is more accurate and the detailed information is excellently preserved in the fusion results. In the future, a more accurate segmentation algorithm with deep learning and a faster post-processing optimized method will be considered to further improve the performance of multimodal medical image fusion.

REFERENCES

[1] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4 – 19, 2014, special Issue on Information Fusion in Medical Image Computing and Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1566253513001450

[2] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3 – 20, 2016, sI: Stereo Data. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S092523121630649X

[3] O. S. Faragallah, H. El-Hoseny, W. El-Shafai, W. A. El-Rahman, H. S. El-Sayed, E.-S. M. El-Rabaie, F. E. A. El-Samie, and G. G. N. Geweid, "A comprehensive survey analysis for present solutions of medical image fusion and future directions," *IEEE Access*, vol. 9, pp. 11 358–11 371, 2021.

[4] H. Fu, G. Sun, J. Ren, A. Zhang, and X. Jia, "Fusion of pca and segmented-pca domain multiscale 2-d-ssa for effective spectral-spatial feature extraction and data classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14, 2020.

[5] M. J. Fadhil, "Design and implementation a prototype system for fusion image by using swt-pca algorithm with fpga technique," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 757–766, 2020.

[6] J. A. Bhutto, T. Lianfang, Q. Du, T. A. Soomro, Y. Lubin, and M. F. Tahir, "An enhanced image fusion algorithm by combined histogram equalization and fast gray level grouping using multi-scale decomposition and gray-pca," *IEEE Access*, vol. 8, pp. 157 005–157 021, 2020.

[7] 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.

[8] S. Das and M. K. Kundu, "Fusion of multimodality medical images using combined activity level measurement and contourlet transform," in *2011 International Conference on Image Information Processing*, 2011, pp. 1–6.

[9] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532–540, 1983.

[10] L. Wang, B. Li, and L. Tian, "Multimodal medical volumetric data fusion using 3-d discrete shearlet transform and global-to-local rule," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 1, pp. 197–206, 2014.

[11] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multimodal sensor medical image fusion based on type-2 fuzzy logic in nsct domain," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3735–3745, 2016.

[12] X. Liu, W. Mei, and H. Du, "Multi-modality medical image fusion based on image decomposition framework and nonsubsampled shearlet transform," *Biomedical Signal Processing and Control*, vol. 40, pp. 343–350, 2018.

[13] X. Lan, S. Zhang, P. C. Yuen, and R. Chellappa, "Learning common and feature-specific patterns: A novel multiple-sparse-representation-based tracker," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2022–2037, 2018.

[14] M. M. Subashini and S. K. Sahoo, "Pulse coupled neural networks and its applications," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3965–3974, 2014.

[15] D. Gai, X. Shen, H. Cheng, and H. Chen, "Medical image fusion via pcnn based on edge preservation and improved sparse representation in nsst domain," *IEEE Access*, vol. 7, pp. 85 413–85 429, 2019.

[16] J. Xia, Y. Chen, A. Chen, and Y. Chen, "Medical image fusion based on sparse representation and pcnn in nsct domain," *Computational and Mathematical Methods in Medicine*, vol. 2018, pp. 1–12, 2018.

[17] 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, pp. 49–64, 2019.

[18] 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," *Information ences*, 2018.

[19] W. Li, L. Jia, and J. Du, "Multi-modal sensor medical image fusion based on multiple salient features with guided image filter," *IEEE Access*, vol. 7, pp. 173 019–173 033, 2019.

[20] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119–130, 2007.

[21] J. M. Dolly and N. A.K., "A survey on different multimodal medical image fusion techniques and methods," in *2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, 2019, pp. 1–5.

[22] B. Meher, S. Agrawal, R. Panda, and A. Abraham, "A survey on region based image fusion methods," *Information Fusion*, vol. 48, pp. 119–132, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253517307583

[23] L. Chen, J. B. Li, and C. L. P. Chen, "Regional multifocus image fusion using sparse representation." *Optics express*, 2013.

[24] R. V. Durga, O. Kumari, M. prakash, P. kumar, and Y. Tirupathi, "Region-based image fusion using complex wavelets," *IOSR Journal of Electronics and Communication Engineering*, vol. 9, pp. 23–26, 01 2014.

[25] B. Yu, B. Jia, L. Ding, Z. Cai, Q. Wu, R. Law, J. Huang, L. Song, and S. Fu, "Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion," *Neurocomputing*, vol. 182, pp. 1–9, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523121501543X

[26] S. Garg, K. Kiran, R. Mohan, and U. Tiwary, "Multilevel medical image fusion using segmented image by level set evolution with region competition," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2005, pp. 7680–7683.

[27] X. Luo, J. Zhang, and Q. Dai, "A regional image fusion based on similarity characteristics," *SIGNAL PROCESSING -AMSTERDAM-*, 2012.

[28] Y. L. A, J. J. A, Q. W. A, Y. S. A, and X. D. B, "Region level based multi-focus image fusion using quaternion wavelet and normalized cut - sciencedirect," *Signal Processing*, vol. 97, no. 7, pp. 9–30, 2014.

[29] B. Meher, S. Agrawal, and P. K. Mishro, "A region based optimal multifocus image fusion scheme," in *2018 International Conference on Recent Innovations in Electrical, Electronics Communication Engineering (ICRIEECE)*, 2018, pp. 499–503.

[30] Q. Li, J. Du, F. Song, C. Wang, H. Liu, and C. Lu, "Region-based multi-focus image fusion using the local spatial frequency," in *2013 25th Chinese Control and Decision Conference (CCDC)*, 2013, pp. 3792–3796.

[31] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2017.

[32] W. Zhang, L. Dong, X. Pan, P. Zou, L. Qin, and W. Xu, "A survey of restoration and enhancement for underwater images," *IEEE Access*, vol. 7, pp. 182 259–182 279, 2019.

[33] A. Xx, A. Wz, B. Hw, B. Ll, B. Zf, B. Zw, B. Zw, and C. Xpb, "Dynamic adaptive residual network for liver ct image segmentation," *Computers and Electrical Engineering*, vol. 91.

[34] X. Pan, L. Li, H. Yang, Z. Liu, J. Yang, L. Zhao, and Y. Fan, "Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks," *Neurocomputing*, vol. 229, pp. 88–99, 2017, advances in computing techniques for big medical image data. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231216313765

[35] X. Xiao, Y. Zhou, and Y. Gong, "Content-adaptive superpixel segmentation," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2883–2896, 2018.

[36] S. Jia, M. Zhang, J. Xian, J. Zhuang, and Q. Huang, "Superpixel-based feature extraction and fusion method for hyperspectral and lidar classification," in *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018, pp. 764–769.

[37] Y. Wu, G. Wen, F. Gao, and Y. Fan, "Superpixel regions extraction for target detection," in *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 2015, pp. 1–5.

[38] H. Zhang, C. Wu, L. Zhang, and H. Zheng, "A novel centroid update approach for clustering-based superpixel methods and superpixel-based edge detection," in *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 693–697.

[39] W. Wang, G. Sun, Y. Yao, and A. Zhang, "Superpixel based spatial and temporal adaptive reflectance fusion model," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 2020, pp. 2308–2311.

[40] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[41] H. Wang, H. Liu, Q. Guo, K. Deng, and C. Zhang, "Design of superpiexl u-net network for medical image segmentation," *Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design and Computer Graphics*, vol. 31, no. 6, p. 1007, 2019.

[42] A. Levinshtein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "Turbopixels: Fast superpixels using geometric flows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2290–2297, 2009.

[43] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *CVPR 2011*, 2011, pp. 2097–2104.

[44] M. Bergh, X. Boix, G. Roig, and L. V. Gool, "Seeds: Superpixels extracted via energy-driven sampling," *International Journal of Computer Vision*, vol. 111, no. 3, pp. 298–314, 2015.

[45] Z. Q. Li and t. y. v. n. p. d. J. S. Chen, booktitle=2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[46] C. S. Asha, S. Lal, V. P. Gurupur, and P. U. P. Saxena, "Multi-modal medical image fusion with adaptive weighted combination of nsst bands using chaotic grey wolf optimization," *IEEE Access*, vol. 7, pp. 40 782–40 796, 2019.

[47] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.

[48] H. M. El-Hoseny, Z. E. Kareh, W. A. Mohamed, G. E. Banby, K. R. Mahmoud, O. S. Faragallah, S. El-Rabaie, E. El-Madbouly, and F. A. El-Samie, "An optimal wavelet-based multi-modality medical image fusion approach based on modified central force optimization and histogram matching," *Multimedia Tools and Applications*, vol. 78, no. 18, pp. 26 373–26 397, 2019.

[49] X. Xie and Q. Qin, "Application of genetic algorithm in adaptive medical image fusion," *Computer Engineering and Applications*, vol. 47, no. 3, pp. 181–184, 2011.

[50] T. Bäck and H.-P. Schwefel, "An overview of evolutionary algorithms for parameter optimization," *Evol. Comput.*, vol. 1, no. 1, p. 1–23, Mar. 1993. [Online]. Available: https://doi.org/10.1162/evco.1993.1.1.1

[51] Nagatani and Takashi, "Modified laplacian growth under shear flow," *Journal of the Physical Society of Japan*, vol. 60, no. 8, pp. 2700–2705, 1991.

[52] P. Guo, X. Wang, and Y. Han, "The enhanced genetic algorithms for the optimization design," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, vol. 7, 2010, pp. 2990–2994.

[53] 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, 2015.

[54] F. G. Veshki, N. Ouzir, S. A. Vorobyov, and E. Ollila, "Coupled Feature Learning for Multimodal Medical Image Fusion," *arXiv e-prints*, p. arXiv:2102.08641, Feb. 2021.

[55] 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. 20 811–20 824, 2019.

[56] Y. Yang, S. Tong, S. Huang, and P. Lin, "Log-gabor energy based multimodal medical image fusion in nsct domain," *Computational and Mathematical Methods in Medicine*, vol. 2014, no. 2, p. 835481, 2014.

[57] C. S. Xydeas and P. V. V., "Objective image fusion performance measure," *Military Technical Courier*, vol. 56, no. 4, pp. 181–193, 2000.

[58] C. He, Q. Liu, H. Li, and H. Wang, "Multimodal medical image fusion based on ihs and pca," *Procedia Engineering*, vol. 7, no. none, pp. 280–285, 2010.

[59] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on 'information measure for performance of image fusion'," *Electronics Letters*, vol. 44, no. 18, pp. 1066–1067, 2008.

[60] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Computers and Electrical Engineering*, vol. 37, no. 5, pp. 744–756, 2011, special Issue on Image Processing. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790611001145