

Received April 22, 2020, accepted May 6, 2020, date of publication May 11, 2020, date of current version May 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2993493

MRI and CT Medical Image Fusion Based on Synchronized-Anisotropic Diffusion Model

RUI ZHU^{ID}, XIONGFEI LI^{ID}, (Member, IEEE), XIAOLI ZHANG^{ID}, AND MINGRUI MA^{ID}

Key Laboratory of Symbolic Computation and Knowledge Engineering, Ministry of Education, Jilin University, Changchun 130012, China
College of Computer Science and Technology, Jilin University, Changchun 130012, China

Corresponding author: Xiaoli Zhang (zhangxiaoli@jlu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801190, in part by the Nature Science Foundation of Jilin Province under Grant 20180101055JC, in part by the Outstanding Young Talent Foundation of Jilin Province under Grant 20180520029JH, in part by the China Postdoctoral Science Foundation under Grant 2017M611323, in part by the Industrial Technology Research and Development Funds of Jilin Province under Grant 2019C054-3, in part by the “Thirteenth Five-Year Plan” Scientific Research Planning Project of Education Department of Jilin Province under Grant JJKH20200997KJ and Grant JJKH20200678KJ, and in part by the Fundamental Research Funds for the Central Universities, JLJU.

ABSTRACT Medical image fusion can combine information from multi-modality images and express them through a single image. How to design a fusion method to preserve more information becomes a hot topic. In this paper, we propose a novel multi-modality medical image fusion method based on Synchronized-Anisotropic Diffusion Equation (S-ADE). First, the modified S-ADE model which is more suitable for Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) images is employed to decompose two source images. We get the base layers and texture layers. Next, the “Maximum Absolute Value” rule is used for base layers fusion. On texture layers, the fusion decision map is calculated by New Sum of Modified Anisotropic Laplacian (NSMAL) algorithm which is designed using common decomposition coefficients by anisotropic diffusion. Furthermore, the consistency check is constructed on the decision map to mitigate the staircase effect. After that, the fused image is obtained by a simple linear combination of layers. Finally, the fused MR/CT image is obtained after image correction. Its aim is to eliminate redundant texture information which is from MRI images in the contour part. The extensive experimental results demonstrate that the proposed method preserves much information as well as guarantees image quality and visual effects. It outperforms other state-of-the-art methods in terms of subjective and objective evaluations.

INDEX TERMS Anisotropic diffusion, medical image, multi-modality image fusion, synchronism.

I. INTRODUCTION

With the development of computer science technology, medical imaging plays a vital role in the clinical diagnosis. Among them, Magnetic Resonance Imaging (MRI) images primarily depict soft tissues, such as blood vessels [1]. Computed Tomography (CT) images can clearly reflect the precise localization of dense structures [2]. It is nearly impossible to get them both from any single medical modality as they provide information from different aspects with their own advantages. Therefore, it is necessary to fuse MRI and CT images to meet the requirements of more complex diagnosis, for example, skull base tumor detecting. Medical image fusion has become a widely used tool for creating high-quality images with amounts of information in order to increase the capability

The associate editor coordinating the review of this manuscript and approving it for publication was Shiqi Wang.

of diagnosis of medical problems. It can combine multiple images from single or various imaging modalities into a fused one. Also, the fusion process can reduce the randomness and redundancy. At present, there are several main methods for medical image fusion, such as spatial domain based, transform domain based, neural network based, hybrid methods and others. We review them briefly.

In spatial domain, the pixels of images can be calculated directly. There are many techniques used widely, such as Principal Component Analysis (PCA), Averaging, Simple Maximum/Minimum method, Bovey Transform (BT), Intensity Hue Saturation (IHS), High Pass Filter (HPF). However, it may produce spatial distortion in the fused images when only using spatial domain methods. The hybrid methods [3]–[6] can improve the quality of the fusion results.

The source images can be decomposed to extract features in addition to being processed directly. For example,

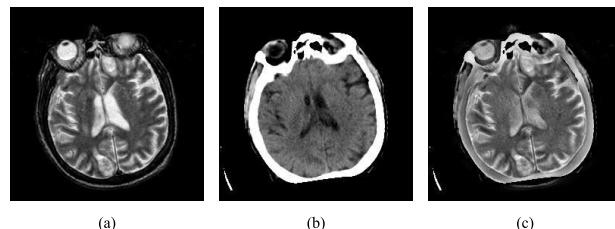


FIGURE 1. An example about medical image fusion. (a) is the source MR-T2 image. (b) is the corresponding CT image. Apparently, (c) is the MR-T2/CT image, the fusion result. It contains both the soft tissue information of (a) and the structural information from (b).

the methods based on Laplacian Pyramid (LP) with multiple features can improve the ability to present outline and contrast of the images [7], [8]. The algorithms using Guided filter (GFF) [9], [10] or Cross Bilateral Filter (CBF) [11] have a good edge preserving property. The multi-level local extrema (LE) schema [12] improves the quality of the fused image. Besides, the methods via multiscale analysis would have the ability to select the frequencies in both space and time because the details in images are usually in the high frequencies. Such as DWT [13]–[15], Curvelet Transform (CVT) [16], [17], Non-Subsampled Contourlet Transform (NSCT) [18]–[20], Non-Subsampled Shearlet Transform (NSST) [21], [22] and Daubechies complex Wavelet Transform [23]. But multiscale transformation methods cannot completely extract the effective information from the base layers of images, so the fusion effect is not ideal in some aspects. Meanwhile, some new image decomposition methods have been proposed for image fusion, such as ones based on Sparse Representation (SR) [24]–[27] and saliency detection [28]–[30].

The transform domain based fusion framework requires the process of image reconstruction, the inverse transformation. Relatively, the end-to-end image processing model is more concise. The methods using Convolutional Neural Networks (CNNs) [2], [31]–[34] to extract features become the research trend gradually. But its interpretability in medicine still needs improvement. There are still limitations of the complicated design of fusion rules for different layers [35]. Whereas the generative adversarial network-based methods can be generalized to fuse images with different modalities and avoid causing functional information blurring or texture detail loss [36], [37]. Besides, some researchers deal with image fusion problem as an optimization model [38]–[41], and use optimization algorithms to combine weighted sub frequency images adaptively. However, these methods have limitations, for example, the need for multiple iterations to find the optimal solution. Approximately, the image fusion process is shown in Fig. 1.

But the white boundary in (b) is not completely preserved. Besides, the tissue information in the center of (a) is blurred after the fusion. To overcome these shortcomings, a novel image fusion method for MRI and CT images is proposed in this paper. The framework is realized using ADE based model

and the features in spatial domain. Compared with the basic scale-space filters above, the anisotropic diffusion is sensitive to texture, and performs well in preserving edge information. It can separate images into base layers and texture layers better [42], [43]. At the same time, the fusion rule for texture layers is designed using the common coefficients obtained by image decomposition. Therefore, the process of image decomposition and fusion has consistency in identifying textures which is beneficial to retaining more textural details. For base layers, we choose the “Maximum Absolute Value” rule. The contributions of this paper are four-fold:

- 1) We propose a preprocessing step for generating a “common image” in order to control the synchronism of the decomposition process of the two source images.

- 2) The fusion rule designed for texture layers is constructed using common coefficients obtained by decomposing the “common image”.

- 3) We establish consistency check on the decision map to mitigate the step effects from anisotropic diffusion.

- 4) We modify the settings and selections of parameters in the proposed algorithm.

The rest of this paper is organized as follows. A survey of related works and motivation are given in Section II. In Section III, the proposed image fusion model is described in detail. The experimental results and analysis are displayed in Section IV. At last, Section V concludes this paper.

II. RELATED WORKS AND MOTIVATION

A. ANISOTROPIC DIFFUSION

The anisotropic diffusion [44] treats images as thermal fields, and each pixel as a heat flow. If the pixel (x, y) is significantly different from others in the neighborhood, it may be retained as boundary instead of being diffused. That is, the filter can smooth a uniform region of a given image. Meanwhile, we use a Partial Differential Equation (PDE) to process non-uniform regions, that is, boundaries [45], which is similar to the bilateral filter. The anisotropic diffusion equation is described as:

$$\begin{cases} I^t = \operatorname{div}(c(x, y, t) \nabla I^{t-1}) \\ = c(x, y, t) \Delta I^{t-1} + \nabla c \cdot \nabla I^{t-1} \\ I^0 = I \end{cases} \quad (1)$$

where the divergence operator is indicated with div , ∇ and Δ represent the Gradient operator and Laplacian operator (divergence of gradient), respectively. $c(x, y, t)$ indicates the rate of diffusion, and t means iteration number. A lower resolution image I^t can be obtained after the anisotropic diffusion function is executed on image I^{t-1} . The solution to Eq. (1) is

$$\begin{aligned} I_{x,y}^t &= I_{x,y}^{t-1} + \lambda \left[c_{Nx,y}^{t-1} \cdot \bar{\nabla}_N I_{x,y}^{t-1} + c_{Sx,y}^{t-1} \cdot \bar{\nabla}_S I_{x,y}^{t-1} \right. \\ &\quad \left. + c_{Ex,y}^{t-1} \cdot \bar{\nabla}_E I_{x,y}^{t-1} + c_{Wx,y}^{t-1} \cdot \bar{\nabla}_W I_{x,y}^{t-1} \right] \end{aligned} \quad (2)$$

where $\lambda \in [0, 1/4]$. $I_{x,y}^t$ calculated from the previous one means the pixel (x, y) in image at t scale. The subscripts N ,

S , E and W are abbreviation for North South East and West. The symbol $\bar{\nabla}$ means nearest-neighbor differences:

$$\begin{aligned}\bar{\nabla}_N I_{x,y} &\equiv I_{x-1,y} - I_{x,y} \\ \bar{\nabla}_S I_{x,y} &\equiv I_{x+1,y} - I_{x,y} \\ \bar{\nabla}_E I_{x,y} &\equiv I_{x,y+1} - I_{x,y} \\ \bar{\nabla}_W I_{x,y} &\equiv I_{x,y-1} - I_{x,y}\end{aligned}\quad (3)$$

In Eq. (2), $c_{Nx,y}^{t-1}$, $c_{Sx,y}^{t-1}$, $c_{Ex,y}^{t-1}$ and $c_{Wx,y}^{t-1}$ indicate the diffusion function $c(x, y, t)$ in the corresponding directions, respectively. They are defined as:

$$\begin{aligned}c_{Nx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x+(1/2),y}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_N I_{x,y}^{t-1}\right\|\right) \\ c_{Sx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x-(1/2),y}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_S I_{x,y}^{t-1}\right\|\right) \\ c_{Ex,y}^{t-1} &= g\left(\left\|(\nabla I)_{x,y+(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_E I_{x,y}^{t-1}\right\|\right) \\ c_{Wx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x,y-(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_W I_{x,y}^{t-1}\right\|\right)\end{aligned}\quad (4)$$

where $c(x, y, t) = g(\|\nabla I(x, y, t)\|)$. The function $g(\cdot)$ will be discussed in Section III in detail.

B. SUM OF MODIFIED LAPLACIAN (SML)

The traditional SML algorithm is able to reflect the edge information of an image as well as the sharpness. But it only calculates the Laplacian value of each pixel in horizontal and vertical directions. SML is defined as:

$$\begin{aligned}ML(x, y) &= |2I(x, y) - I(x - step, y) - I(x + step, y)| \\ &\quad + |2I(x, y) - I(x, y - step) - I(x, y + step)|\end{aligned}\quad (5)$$

$$SML(x, y) = \sum_{i=-N}^N \sum_{j=-N}^N ML(x + i, y + j)\quad (6)$$

where $step$ denotes a variable spacing between pixels. The accumulation window about calculating the value of ML is set to $(2N + 1) \times (2N + 1)$.

C. MOTIVATION OF THIS WORK

In image fusion, how much details are retained in the fused image largely determines the image quality, which will have an impact on clinical diagnostics. The combination of anisotropic diffusion and SML can separate details from the source image more accurately. It is not difficult to find that the anisotropic decomposition model and SML algorithm use similar ideas in the image calculation process. That is, the North South East and West are equivalent to vertical and horizontal in direction. But the selection of only two directions reduces the accuracy and it cannot make use of the strong correlation among adjacent pixels. Therefore, we desire to expand the selections and use more adjacent pixels for calculation. Besides, we need to use parameters to associate the decomposition process with the fusion method in order to improve the consistency in identifying textures of the overall framework. The acquisition of parameters depends on the process of image decomposition. Meanwhile, we hope

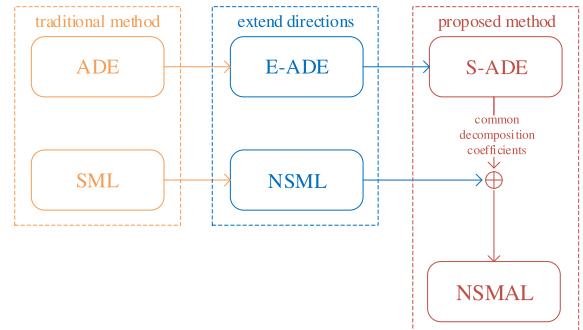


FIGURE 2. The relation between the prior works and proposed model.

that these parameters are obtained through calculating two source images, which makes the whole model more adaptive.

As mentioned above, inspired by some recent advances in anisotropic diffusion method [45], [46] and the importance in synchronized decomposition, the S-ADE image decomposition method (presented in Section III-A) is introduced to decompose medical images. The NSMAL algorithm (proposed in Section III-B.2) is designed to fuse texture layers of medical images. The relation between the prior works and proposed methods is shown in Fig. 2.

III. PROPOSED METHOD

In this section, we describe the proposed model for MRI/CT medical image fusion. Here, we assume that the source MRI and CT images are well registered. There is no need to consider about the problems about the registration. First, we generate a “common image” based on two source images. We can get the common decomposition coefficients (direction parameters) by decomposing the “common image”. After image decomposition via S-ADE, the base layer and texture layer of each image are obtained. Then fuse the base layers using “Maximum Absolute Value” method. The texture layers are fused by NSMAL algorithm. At last, the fusion result can be obtained through the image reconstruction. The framework of the proposed model is presented in Fig. 3.

A. IMAGE DECOMPOSITION USING S-ADE

1) E-ADE

As mentioned in the motivation, the traditional ADE model can be expanded. In detail, the four nearest points around the pixel (x, y) are in the N , S , E and W directions in image. The second nearest points are in the Northwest (NW) Northeast (NE) Southwest (SW) and Southeast (SE) directions. So there are eight points contributing to calculate the energy of the pixel (x, y) . The Eqs. (3) and (4) can be extended to the following:

$$\begin{aligned}\bar{\nabla}_{NW} I_{x,y} &\equiv I_{x-1,y-1} - I_{x,y} \\ \bar{\nabla}_{NE} I_{x,y} &\equiv I_{x-1,y+1} - I_{x,y} \\ \bar{\nabla}_{SW} I_{x,y} &\equiv I_{x+1,y-1} - I_{x,y} \\ \bar{\nabla}_{SE} I_{x,y} &\equiv I_{x+1,y+1} - I_{x,y}\end{aligned}\quad (7)$$

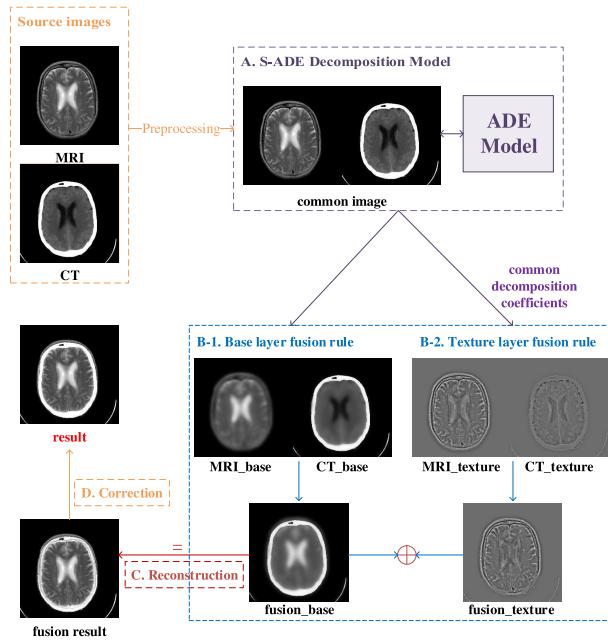


FIGURE 3. The framework of the proposed MRI/CT image fusion model.

$$\begin{aligned} c_{NWx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x+(1/2),y-(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_{NW}I_{x,y}^{t-1}\right\|\right) \\ c_{NEx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x+(1/2),y+(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_{NE}I_{x,y}^{t-1}\right\|\right) \\ c_{SWx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x-(1/2),y-(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_{SW}I_{x,y}^{t-1}\right\|\right) \\ c_{SEx,y}^{t-1} &= g\left(\left\|(\nabla I)_{x-(1/2),y+(1/2)}^{t-1}\right\|\right) = g\left(\left\|\bar{\nabla}_{SE}I_{x,y}^{t-1}\right\|\right) \end{aligned} \quad (8)$$

2) SYNCHRONICITY ESTABLISHMENT

The ADE image decomposition model can be adaptively performed according to the properties of the image itself. So the decomposition bases of different images are different in process. But image fusion is to combine salient features of multi-modality images together, and express it through a single image. We hope that the source images are decomposed under the same condition to meet the fusion consistency. In our opinion, the solution is to generate a “common image”, based on which the decomposition coefficients of different source images are obtained. The preprocessing method should be simple and effective. There are three solutions to the problem: PCA algorithm, average filter and stitching two images. One experiment result about the preprocessing is shown in Fig. 4.

From the figure, it is obvious that the texture information marked by red box in MRI image displayed in Fig. 4(a) is important. Hence, we need to preserve in the fused images. However, the details are partially lost in the fused images obtained using PCA algorithm and average filter. By contrast, the fused image using image stitching preserves more textural information. Extensive experiments are conducted to compare the performance of the three strategies, and the experimental results demonstrate that stitching method performs better than the other two ones. The stitching method is simple but keeping the original information. Furthermore, we need

to generate a “common image” after each iteration diffusion process, which is tedious and time-consuming when using algorithm or filter. But it needs to generate only one initial “common image” using the stitching method. The diffusion iterations are executed based on it.

3) DIFFUSION COEFFICIENT FUNCTION

It needs to determine the diffusion coefficient function $g(\cdot)$ mentioned in Section II-A after obtaining the required “common image”. $g(\cdot)$ is a monotonically decreasing function with $g(0) = 1$. There are many selections meeting the requirements. According to the suggestions in [44], we summarize as follows:

$$g(\nabla I, K, \alpha) = \frac{1}{1 + (\frac{\|\nabla I\|}{K})^{1+\alpha}} \quad (9)$$

where $\alpha > 0$. The function privileges wide regions over the smaller ones. The constant K is called edge threshold, which is important to the function. If K is too large, the base layer information from CT image will be lost. On the contrary, the smooth diffusion on the MRI image stops earlier, resulting in insufficient texture information of the image after decomposition. In order to retain useful information of MRI and CT images effectively, we need to discuss the choices of K and α . Besides, as a non-linear iterative filter, the number of iterations of the S-ADE algorithm determines the degree of separation between texture layer and basic layer. Specifically, the details will be retained from MRI image enough if the number of iterations is large. By contrast, more basic information of CT image will be preserved. The choices of parameters for the diffusion function are shown in Section IV.

4) IMAGE DECOMPOSITION

The image $I_{x,y}^t$ is decomposed by the S-ADE model for obtaining base layer:

$$B_I^t = S - ADE(I_{x,y}^t) \quad (10)$$

where B_I^t is the t -level base layer and $S - ADE(I_{x,y}^t)$ means that the proposed model decompose the t -level image. The texture layer can be obtained by the subtraction:

$$T_I = I - B_I^t \quad (11)$$

In summary, the image decomposition method based on S-ADE model can be summarized as Algorithm 1. The “common image” is generated in Step 1. From Step 2 to 6, we get the “common base layer” B_I^t . After Step 7, the “common texture layer” T_I can be obtained. At last, we get the output layers through Step 8 and 9.

Two groups of results about decomposition on MR-T2 and CT images are shown in Fig. 5. From (e) and (f), it is clear that the textural details separated from MRI images are more than the ones from CT images, and the base layer information in (c) and (d) is well preserved. The results indicate that the proposed image decomposition method is sensitive to the texture information. That the base layer and texture layer are well separated can improve the fusion effect.

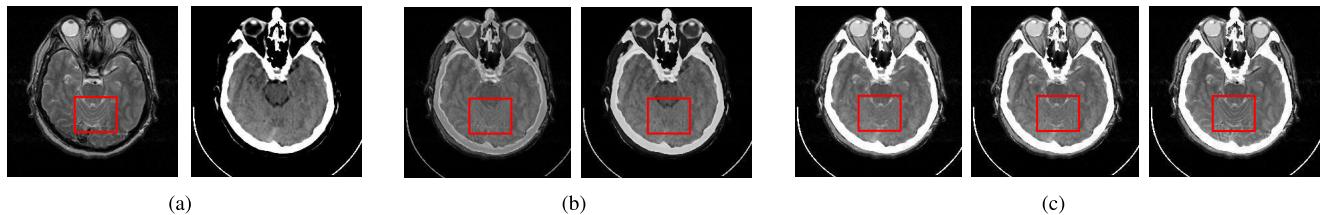


FIGURE 4. The “common image” results and fusion effect from different methods. (a) shows the source images: MR-T2 image (left) and CT image (right). (b) indicates the “common image” by two methods: average filter (left) and PCA algorithm (right). (c) presents the fusion results based on the filter (left), algorithm (middle) and stitching (right) under the same decomposition and fusion condition.

Algorithm 1 The S-ADE Decomposition Procedure

Input: the source images: I_{MRI}, I_{CT} .

Output: the texture layers: T_{MRI}, T_{CT} and the base layers: B_{MRI}, B_{CT} .

- 1: Generate the “common image” I^0 using stitching method.
- 2: **for** $t = 1$:iteration number **do**
- 3: Build the S-ADE model using Eq. (3) Eq. (4) Eq. (7) and Eq. (8);
- 4: Build the diffusion function using Eq. (9);
- 5: Get the base layer B_I^t of the “common image” using Eq. (10).
- 6: **end for**
- 7: Get the texture layer T_I of the “common image” using Eq. (11).
- 8: Get the base layers: B_{MRI}, B_{CT} after segmenting B_I^t .
- 9: Get the texture layers: T_{MRI}, T_{CT} after segmenting T_I .

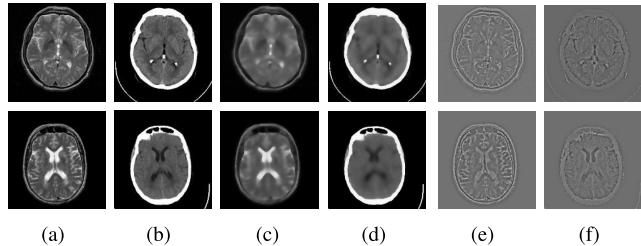


FIGURE 5. The experimental results of image decomposition. (a) MR-T2 image. (b) CT image. (c) Base layer of MR-T2 image. (d) Base layer of CT image. (e) Texture layer of MR-T2 image. (f) Texture layer of CT image.

B. IMAGE FUSION RULES

The base layer and texture layer present different information, and they largely vary in significance for MRI and CT image. For example, texture information of MRI is more useful, and basic information from CT can reflect the bone structure better. Therefore, different fusion methods are designed for different layers.

1) BASE LAYER FUSION RULE

The fusion method for base layer is simpler as the detail information contained in B_I is very limited. In this framework, we use the “Maximum Absolute Value” rule as the fusion

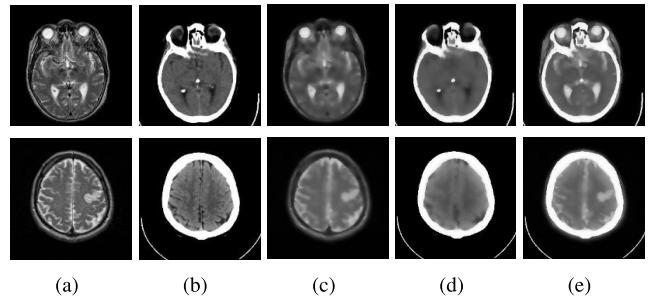


FIGURE 6. Base layer fusion results of two groups of MR-T2 and CT images. (a) MR-T2 image. (b) CT image. (c) MR-T2 image base layer. (d) CT image base layer. (e) Base layer fusion result.

algorithm. We get the fusion map using:

$$map_B = (B_{MRI} \geq B_{CT}) \quad (12)$$

In this way, the base layer B_F of the fused image can be obtained by:

$$B_F = B_{MRI} \times map_B + B_{CT} \times (1 - map_B) \quad (13)$$

The fused base layer results are shown in Fig. 6.

2) TEXTURE LAYER FUSION RULE

On texture layers, the fusion decision map is calculated by NSMAL algorithm which is designed using directional coefficients obtained by decomposing MRI and CT images. Similar to the E-ADE model, the traditional SML algorithm can be extended to NSML, and NML is defined as:

$$\begin{aligned} NML(x, y) &= |(c_1 + c_2)I(x, y) - c_1I(x - step, y) \\ &\quad - c_2I(x + step, y)| + |(c_3 + c_4)I(x, y) \\ &\quad - c_3I(x, y - step) - c_4I(x, y + step)| \\ &\quad + |(c_5 + c_6)I(x, y) - c_5I(x - step, y - step) \\ &\quad - c_6I(x + step, y + step)| + |(c_7 + c_8)I(x, y) \\ &\quad - c_7I(x - step, y + step) - c_8I(x + step, y - step)| \end{aligned} \quad (14)$$

In order to establish the consistency between the decomposition and fusion process, the NMAL is proposed as:

$$\begin{aligned} c_1 &= c_N, \quad c_3 = c_W, \quad c_5 = c_{NW}, \quad c_7 = c_{NE} \\ c_2 &= c_S, \quad c_4 = c_E, \quad c_6 = c_{SE}, \quad c_8 = c_{SW} \end{aligned} \quad (15)$$

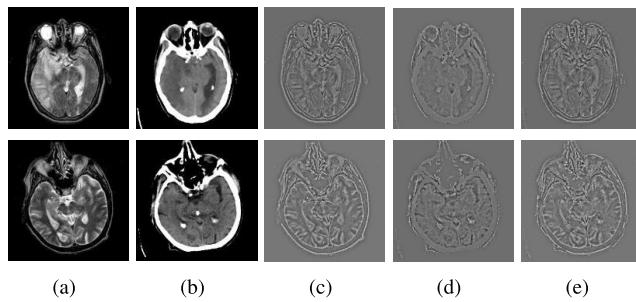


FIGURE 7. Texture layer fusion results of two groups of MR-T2 and CT images. (a) MR-T2 image. (b) CT image. (c) MR-T2 image texture layer. (d) CT image texture layer. (e) Texture layer fusion result.

where the coefficients of Eq. (14) are the diffusion ones obtained from the S-ADE model. After calculating the $NMAL(x, y)$, we get $NSMAL(x, y)$ via $NSMAL(x, y) = \sum_{i=-N}^N \sum_{j=-N}^N NMAL(x+i, y+j)$. The fusion decision map is defined as:

$$map_T = (\omega \times NSMAL_{T_MRI} \geq NSMAL_{T_CT}) \quad (16)$$

where the weight coefficient $\omega > 1$ as we hope to preserve more details from MRI image.

In order to mitigate the staircase effects in fused texture layer, we establish a consistency check on the decision map map_T through a sliding window, which aims to further improve the quality of fused images. The window is set to 3×3 . Similar to the image decomposition process, we calculate the sum of eight values around the pixel in the map. If more than half of the pixels are from the MRI image, the value at pixel (x, y) is set to 1. By contrast, the value is set to 0 (see Eq. (17)). The fused texture layer can be obtained through Eq. (18).

$$map_T(x, y) = \begin{cases} 1 & \text{if } \text{sum} \geq 4 \\ 0 & \text{else} \end{cases} \quad (17)$$

$$T_F = T_{MRI} \times map_T + T_{CT} \times (1 - map_T) \quad (18)$$

The texture layer fusion results are shown in Fig. 7.

3) IMAGE RECONSTRUCTION

The image of low resolution which is treated as base layer is obtained using S-ADE model to decompose the source image. We subtract the low frequency information from the source image to get texture layer according to Eq. (11). The image reconstruction process is the inverse process of decomposition. Thus, the fused image F is given by a linear combination of base layer and texture layer as:

$$F = B_F + T_F \quad (19)$$

Two groups of fusion results images are shown in Fig. 8.

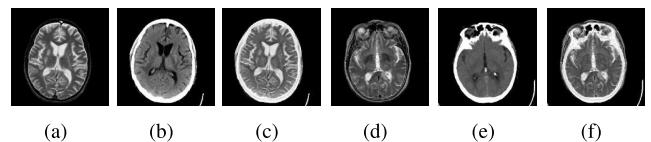


FIGURE 8. The fusion results of two groups of MR-T2/CT images. (a)(d) MR-T2 images. (b)(e) CT images. (c)(f) MR-T2/CT fusion results.

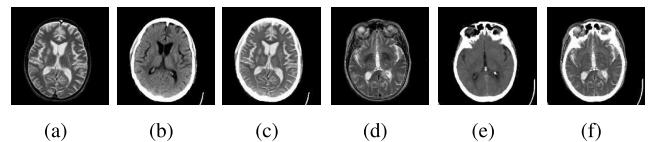


FIGURE 9. The MR-T2/CT final fusion results of two groups of MR-T2 and CT images. (a)(d) MR-T2 images. (b)(e) CT images. (c)(f) MR-T2/CT fusion results after image correction.

4) IMAGE CORRECTION

There are gray lines within the white outline of the fused results by observing the fusion images displayed in Fig. 8. In our opinion, the lines are texture information from MRI image and the white outline is preserved from CT image. The white part of CT image reflects the bone information. But current fused images may mislead a diagnosis as a fracture. Therefore, we add a step to modify it. We get an intensity weight map of CT image:

$$map_I = I_{CT} \geq \theta \quad (20)$$

where θ is an intensity threshold used to extract contour from CT image. Pixels with an intensity greater than θ are retained in white outline of the fused image. The residual texture information of the MRI image in the white outline of fused image is eliminated by:

$$F_{final} = I_{CT} \times map_I + F \times (1 - map_I) \quad (21)$$

The final fused images are shown in Fig. 9.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. DATA SETS

In our experiments, there are 20 pairs of MRI and CT images (see Fig. 10). All of them are from the Whole Brain Atlas database of Harvard Medical School, which is a popular medical image database. (<http://www.med.harvard.edu/AANLIB/home.html>.)

B. EVALUATION

Both subjective visual effect and objective quantitative analysis are necessary when evaluating the fusion results. The subjective evaluation is more important because fused images are read by doctors. The intuitive visual effect may influence the clinical diagnosis.

The objective quantitative indices we select to evaluate the results from two aspects: the amount of salient visual information transferred from source images to their fused image and the visual quality of the fused image.

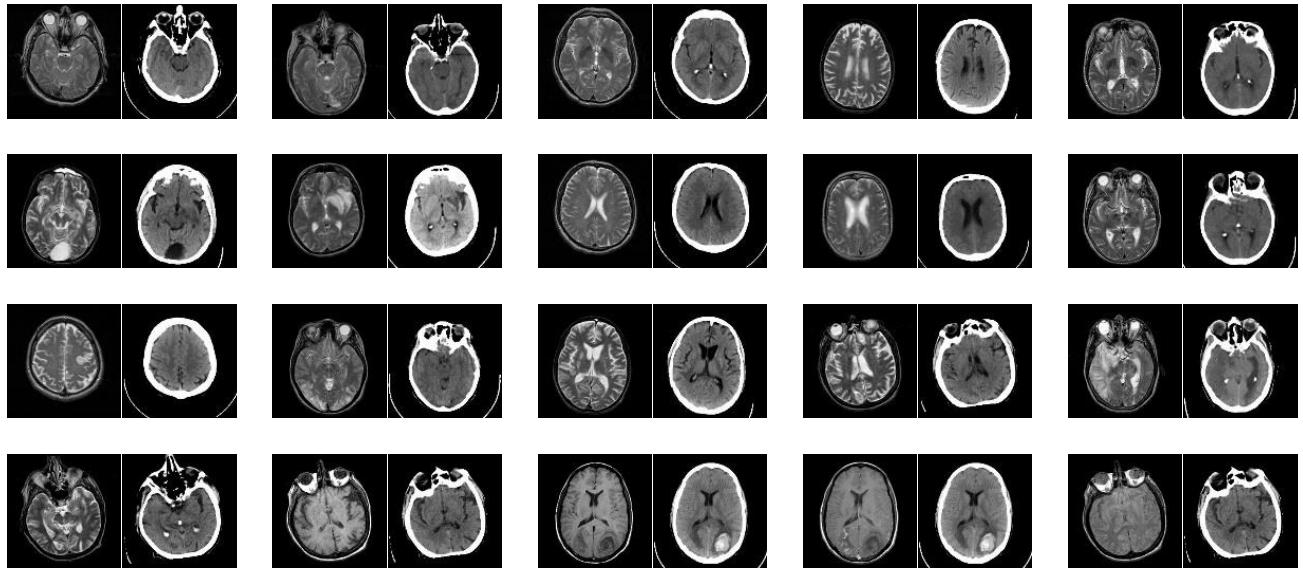


FIGURE 10. The 20 groups of source multi-modality medical images are used in our experiment. Among each pair, MRI is on the left and CT on the right.

Larger values indicate more information and better fusion results. The metrics include:

(1) Mutual Information (MI) [47] measures the mutual dependence of two images which is defined as:

$$MI(U, V) = \sum_{u \in U} \sum_{v \in V} p(u, v) \log_2 \frac{p(u, v)}{p(u)p(v)} \quad (22)$$

where $p(u, v)$ is the joint probability distribution function, $p(u)$ and $p(v)$ are the marginal distribution function. The quality of the fused result with respect to source images can be calculated as:

$$Q_{MI} = 2 \left[\frac{MI(A, F)}{H(A) + H(F)} + \frac{MI(B, F)}{H(B) + H(F)} \right] \quad (23)$$

(2) Universal Quality Index (Q^0) [48] models distortion as a combination of three factors: loss of correlation, luminance and contrast distortion. It is defined as:

$$Q^0 = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (24)$$

where \bar{x} and \bar{y} mean the average, σ_x and σ_y represent the variance, $\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$.

(3) Edge-Based Similarity Measure ($Q^{AB/F}$) [49] measures the similarity between the fused image and source images by evaluating the amount of edge information transferred in the fusion process. The $Q^{AB/F}$ is defined as:

$$Q^{AB/F} = \frac{\sum_{i=1}^M \sum_{j=1}^N (Q_{i,j}^{AF} w_{i,j}^A + Q_{i,j}^{BF} w_{i,j}^B)}{\sum_{i=1}^M \sum_{j=1}^N (w_{i,j}^A + w_{i,j}^B)} \quad (25)$$

where $w_{i,j}^X$ is the weights for $Q_{i,j}^{XF}$, ($X = A, B$). The definition of $Q_{i,j}^{XF}$ is given as $Q_{i,j}^{XF} = Q_{g,i,j}^{XF} \cdot Q_{\alpha,i,j}^{XF}$ where $Q_{k,i,j}^{XF}$, ($k = g, \alpha$) presents the similarity of the width and direction between X and F at location (i, j) .

(4) Nonlinear Correlation Information Entropy (NCIE) [50] has excellent mathematical properties as a measure for the nonlinear type of correlation of concerned variables:

$$NCIE = 1 + \sum_{i=1}^K \frac{\lambda_i^R}{K} \log_b \frac{\lambda_i^R}{K} \quad (26)$$

where λ_i^R ($i = 1, \dots, K$) are the eigenvalues of the nonlinear correlation matrix R .

(5) Visual Information Fidelity Fusion (VIFF) [51]:

$$VIFF(I_1, \dots, I_F) = \sum_k p_k \cdot \frac{\sum_b FVID_{k,b}(I_1, \dots, I_F)}{\sum_b FVIND_{k,b}(I_1, \dots, I_F)} \quad (27)$$

where $FVID_{k,b}$ and $FVIND_{k,b}$ denote fusion visual information with distortion and fusion visual information without distortion in the b th block at k th sub-band, respectively.

(6) Feature Similarity Index (FSIM) [52] focuses on low-level features. It calculates the similarity between images using phase congruency (PC) and gradient magnitude (GM) information:

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (28)$$

where $S_L(x) = S_{PC}(x) \cdot S_G(x)$. The similarity measure for $PC_1(x)$ and $PC_2(x)$ is defined as $S_{PC}(x) = \frac{2PC_1(x) \cdot PC_2(x) + T_1}{PC_1^2(x) + PC_2^2(x) + T_1}$ where $T_1 > 0$. Similarly, $S_G(x) = \frac{2G_1(x) \cdot G_2(x) + T_2}{G_1^2(x) + G_2^2(x) + T_2}$.

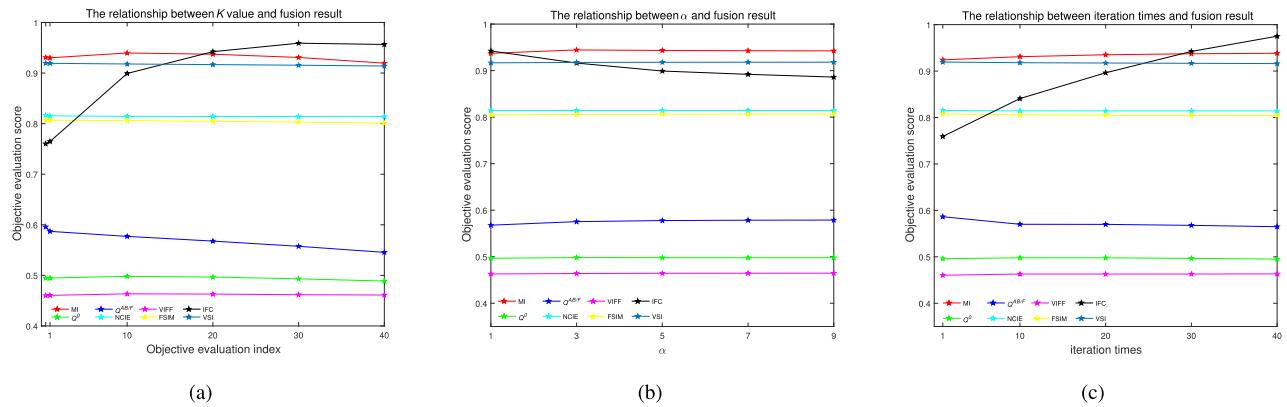


FIGURE 11. The evaluation of results about experiments on the relationship between parameters and fusion results. (a) The evaluation results about the relationship between K value and the fusion result where iteration = 30, $\alpha = 1$. (b) The evaluation results about the relationship between α and fusion result where $K = 20$, iteration = 30. (c) The evaluation results about the relationship between iteration times of S-ADE algorithm and fusion result ($K = 20$, $\alpha = 1$).

TABLE 1. The subjective evaluation criteria.

score	subjective evaluation
1-2	image blur, texture confusion
3-4	poor visual effect, distortion and information loss
5-6	ordinary, slight distortion and information loss
7-8	good visual effect, slight texture information loss
9-10	excellent visual effect, texture clear

(7) Information Fidelity Criterion (IFC) [53]:

$$IFC = \sum_{k \in \text{subbands}} I(C^{N_k, k}; D^{N_k, k} | s^{N_k, k}) \quad (29)$$

where $C^{N_k, k}$ denotes N_k coefficients from the RF C^k of the k th sub-band, and similarly for $D^{N_k, k}$ and $s^{N_k, k}$.

(8) Visual Saliency induced Index (VSI) [54] employed as a weighting function is closely related to image quality assessment which reflects the importance of a local region:

$$VSI = \frac{\sum_{x \in \Omega} S(x) \cdot VS_m(x)}{\sum_{x \in \Omega} VS_m(x)} \quad (30)$$

where $S(x) = S_{VS}(x) \cdot [S_G(x)]^\alpha \cdot [S_C(x)]^\beta$, α and β are two parameters used to adjust the relative importance of visual saliency (VS), GM and chrominance features, $S_{VS}(x) = \frac{2VS_1(x) \cdot VS_2(x) + C_1}{VS_1^2(x) + VS_2^2(x) + C_1}$ and $S_C(x) = \frac{2M_1(x) \cdot M_2(x) + C_2}{M_1^2(x) + M_2^2(x) + C_2}$, $\frac{2N_1(x) \cdot N_2(x) + C_3}{N_1^2(x) + N_2^2(x) + C_3}$ where M and N denote channels of images.

The subjective evaluation criteria are described in TABLE 1. We invite five experts to evaluate the proposed method with other methods subjectively. All images are displayed under the same conditions.

C. PARAMETERS SETTINGS

In order to make the S-ADE model more suitable for the decomposition of MRI and CT images, we discuss the settings of diffusion function parameters. To improve the computational efficiency of the model, we choose constants

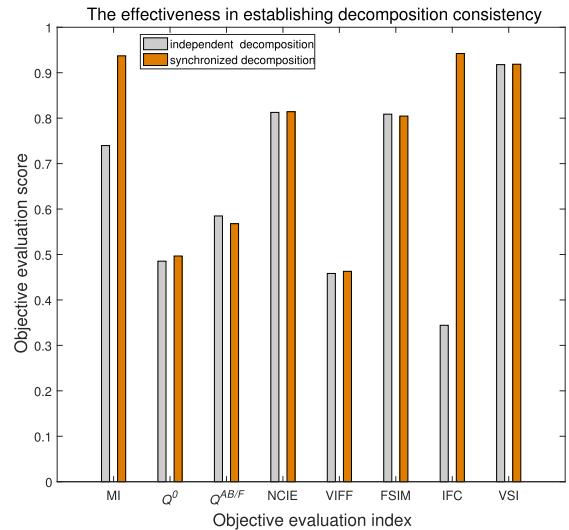


FIGURE 12. The results under synchronized decomposition and independent decomposition.

as hyperparameters. The experiments are performed on ten pairs of images. The objective evaluation about experimental results is shown in Fig. 11.

It is obvious that most scores rise and stabilize gradually from $K = 0.5$ to $K = 20$, then they slide down slightly except the IFC and $Q^{AB/F}$ in Fig. 11(a). IFC keeps rising and $Q^{AB/F}$ decreases. Both of them flatten gradually, that is, the slope decreases. Because the value of $Q^{AB/F}$ decreases more in CT image than it rises in MRI image, the average one goes down. So we decide that $K = 20$ in the S-ADE model after analyzing the objective quantitative evaluation.

In order to maintain the properties of the diffusion function, α is an odd number. It is set to 1 after analyzing the sensitivity of the indices to changes of α in Fig. 11(b). IFC decreases more obviously than MI, VSI and $Q^{AB/F}$ which rise a little in $\alpha \in [1, 3]$.

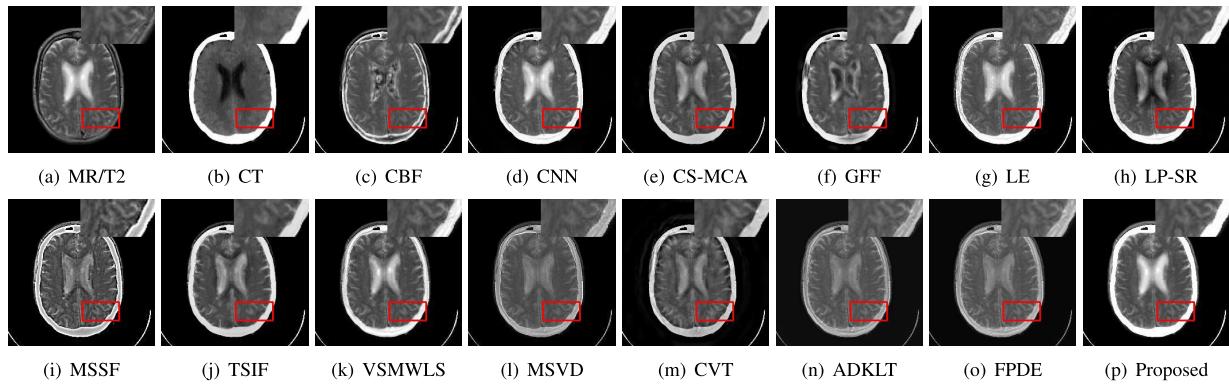


FIGURE 13. The fusion results on first group of MR-T2 image and CT image under 14 different methods.

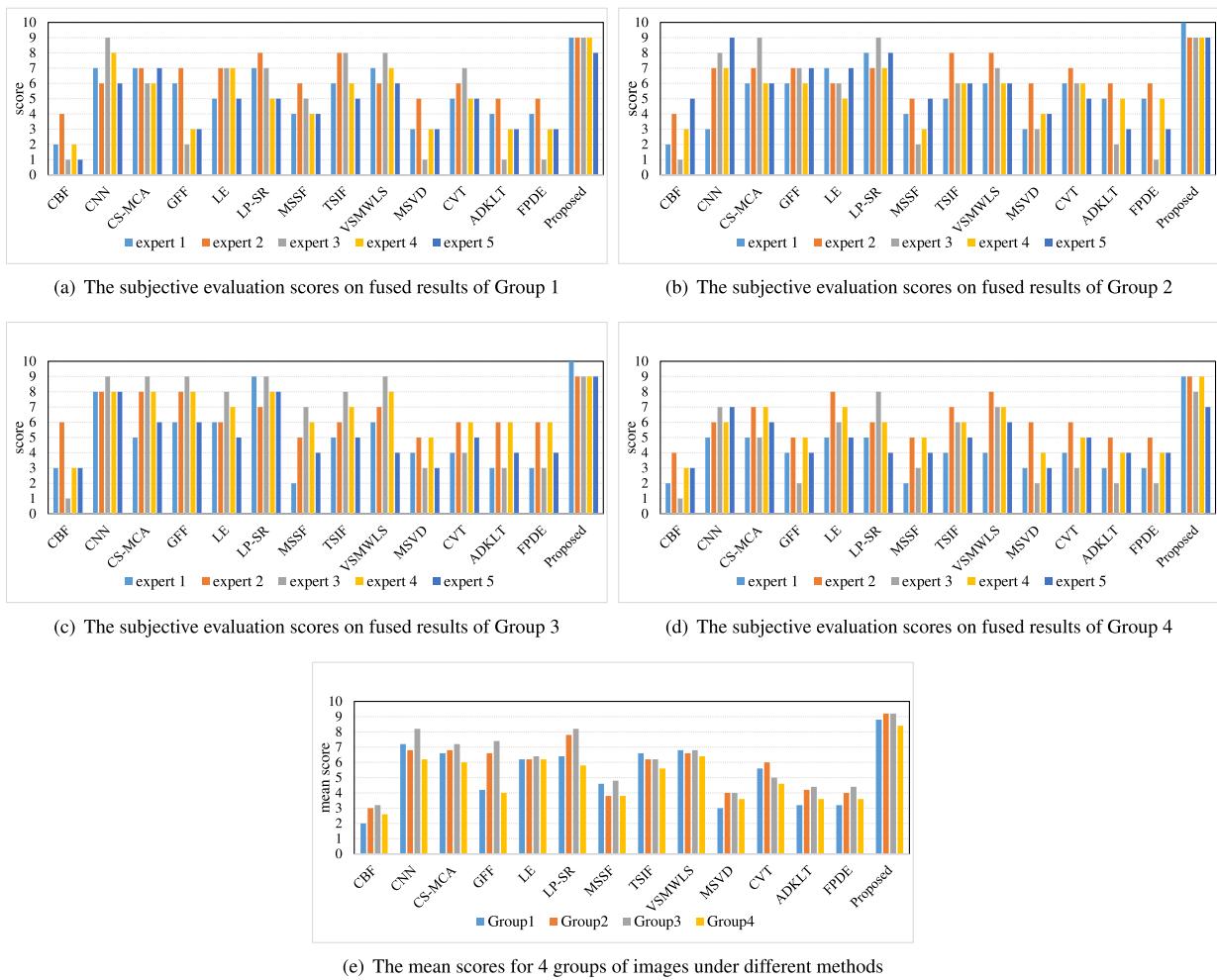


FIGURE 14. The subjective evaluation on different groups of images.

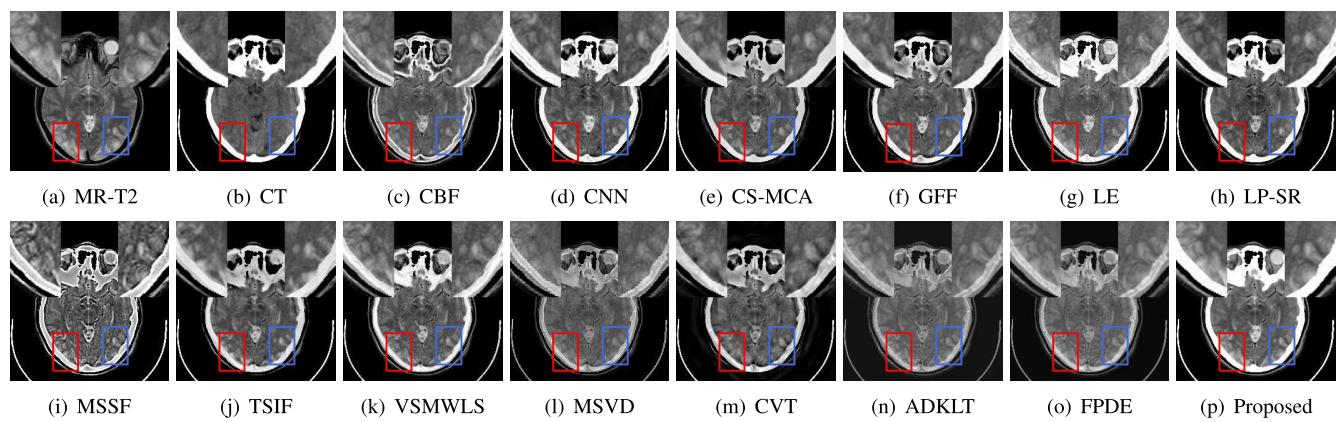
The evaluation results in Fig.11(c) indicate that most of the indices rise up slowly before 20 iterations, decline until 30 iterations. Because the details of MRI image increase less than the lost basic information of CT image during the iteration process, the scores of FSIM and $Q^{AB/F}$ reduce a

little, but not enough to affect the conclusion, the number of iterations = 30.

Fig. 12 shows the effectiveness in establishing decomposition synchronism using stitching on ten pairs of images in terms of objective quantitative evaluation. The scores in

TABLE 2. The objective evaluation of the fusion results of the first group of images.

Groups	Metrics Methods	MI	Q^0	$Q^{AB/F}$	NCIE	VIFF	FSIM	IFC	VSI
Group 1 MR-T2/CT	CBF	1.7883	0.3192	0.5455	0.8084	0.3733	0.7977	1.5660	0.8602
	CNN	1.6267	0.3912	0.5657	0.8082	0.5525	0.8393	1.4753	0.8842
	CS-MCA	1.4475	0.3565	0.5403	0.8075	0.5107	0.8528	1.7072	0.8808
	GFF	1.4327	0.3261	0.5481	0.8096	0.4718	0.8335	2.0467	0.8761
	LE	2.1146	0.4135	0.5665	0.8105	0.5860	0.8292	1.9041	0.8915
	LP-SR	1.3655	0.3541	0.5526	0.8079	0.5675	0.8394	1.5541	0.8841
	MSSF	1.5288	0.3343	0.4806	0.8074	0.5127	0.7793	0.9572	0.8627
	TSIF	1.4973	0.3524	0.5028	0.8077	0.5567	0.8440	1.1878	0.8800
	VSMWLS	1.7080	0.3854	0.5379	0.8082	0.5778	0.8291	1.1607	0.8842
	MSVD	1.6982	0.2770	0.4498	0.8082	0.4268	0.8387	1.0248	0.8917
	CVT	1.3434	0.2956	0.4673	0.8070	0.4467	0.8117	1.1521	0.8824
	ADKLT	1.7845	0.2322	0.4153	0.8082	0.3903	0.8269	1.3836	0.8912
	FPDE	1.8408	0.2528	0.4199	0.8085	0.4270	0.8311	1.2762	0.8901
	Proposed	2.8439	0.4230	0.5992	0.8137	0.6132	0.8542	5.4533	0.9245

**FIGURE 15.** The fusion results on second group of MR-T2 image and CT image under 14 different methods.**TABLE 3.** The objective evaluation of the fusion results of the second group of images.

Groups	Metrics Methods	MI	Q^0	$Q^{AB/F}$	NCIE	VIFF	FSIM	IFC	VSI
Group 2 MR-T2/CT	CBF	1.8034	0.4379	0.6321	0.8082	0.3186	0.7217	2.2416	0.8524
	CNN	1.3500	0.4378	0.6174	0.8072	0.3873	0.7685	1.4979	0.8541
	CS-MCA	1.4383	0.4327	0.5823	0.8069	0.3557	0.7824	1.6249	0.8667
	GFF	1.4479	0.4353	0.6265	0.8069	0.3502	0.7629	1.6366	0.8596
	LE	2.1287	0.4878	0.5986	0.8107	0.4433	0.7605	2.2463	0.8714
	LP-SR	1.2633	0.4404	0.6228	0.8081	0.3819	0.7752	1.8082	0.8721
	MSSF	1.4299	0.4073	0.5062	0.8065	0.3860	0.6845	1.0682	0.8400
	TSIF	1.3710	0.4203	0.5510	0.8068	0.4016	0.7723	1.2271	0.8649
	VSMWLS	1.5251	0.4568	0.5981	0.8071	0.4414	0.7568	1.1990	0.8611
	MSVD	1.5894	0.3439	0.4813	0.8072	0.3596	0.7642	1.0806	0.8866
	CVT	1.2195	0.3593	0.5141	0.8062	0.3028	0.7429	1.2184	0.8627
	ADKLT	1.7287	0.2469	0.3788	0.8074	0.3050	0.7594	1.4382	0.8903
	FPDE	1.7801	0.2853	0.4373	0.8077	0.3492	0.7630	1.3649	0.8898
	Proposed	2.7974	0.4895	0.6322	0.8133	0.4508	0.7877	4.1540	0.9062

the figure have been normalized in the same way for the convenience of display. The MI, Q^0 and IFC get improved obviously, NCIE, VIFF and VSI rise a little, but $Q^{AB/F}$ and FSIM reduce slightly. Overall, generating “common image” can ameliorate the fusion effect.

D. COMPARATIVE ANALYSIS AND RESULTS DISCUSSION

In this subsection, we will show seven groups of experimental results about medical multi-modality images fusion.

It consists of four groups of MR-T2 and CT images (MR-T2/CT), a group of MR-T1 and CT images (MR-T1/CT), a group of MR-Gad and CT images (MR-Gad/CT) and a group of MR-PD and CT images (MR-PD/CT). The proposed method is compared with 13 fusion methods: CBF [11], CNN [31], CS-MCA [27], GFF [9], LE [12], LP-SR [24], MSSF [28], TSIF [29], VSMWLS [30], MSVD [55], CVT [17], ADKLT [45] and FPDE [46]. The parameters of the methods above are the default ones from the provided codes.

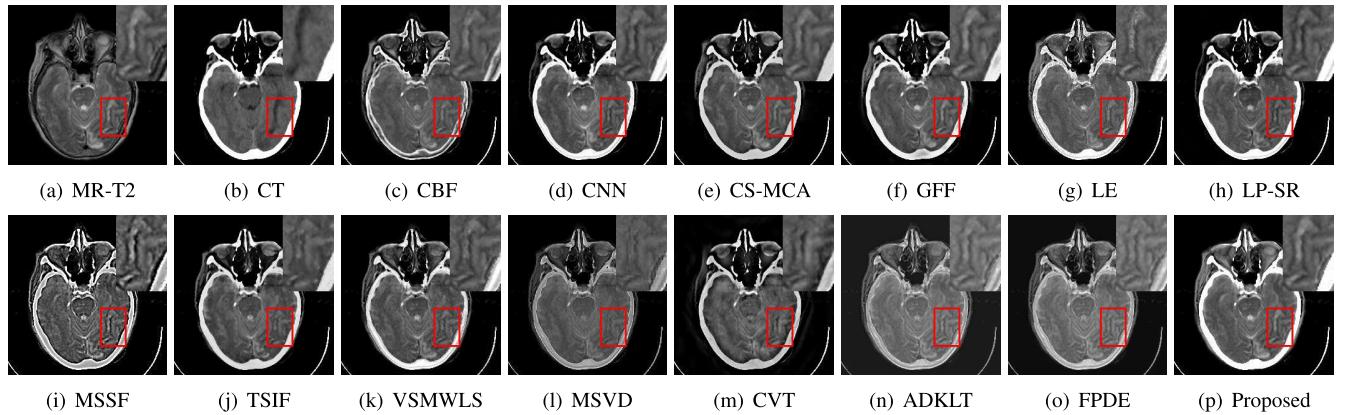


FIGURE 16. The fusion results on third group of MR-T2 image and CT image under 14 different methods.

TABLE 4. The objective evaluation of the fusion results of the third group of images.

Groups	Metrics	MI	Q^0	$Q^{AB/F}$	NCIE	VIFF	FSIM	IFC	VSI
Methods									
Group 3 MR-T2/CT	CBF	1.8365	0.4492	0.6111	0.8081	0.3661	0.7360	2.7569	0.8715
	CNN	1.2690	0.4515	0.6133	0.8070	0.4111	0.7726	1.8062	0.8847
	CS-MCA	1.3653	0.4604	0.6046	0.8066	0.3924	0.7822	2.0139	0.8870
	GFF	1.4403	0.4805	0.6309	0.8070	0.4112	0.7778	2.5531	0.8862
	LE	2.2514	0.4866	0.5767	0.8107	0.4566	0.7529	2.8262	0.8921
	LP-SR	1.1988	0.4545	0.6128	0.8080	0.4023	0.7796	1.9917	0.8986
	MSSF	1.3996	0.4110	0.4945	0.8061	0.4213	0.6868	1.0803	0.8550
	TSIF	1.4097	0.4546	0.5668	0.8068	0.4429	0.7771	1.4678	0.8859
	VSMWLS	1.5359	0.4652	0.5658	0.8069	0.4448	0.7584	1.3974	0.8858
	MSVD	1.6410	0.3384	0.4661	0.8073	0.3503	0.7732	1.1161	0.9015
	CVT	1.1290	0.3632	0.5302	0.8058	0.3263	0.7594	1.4534	0.8816
	ADKLT	1.7360	0.2564	0.4461	0.8073	0.3333	0.7645	1.5106	0.9009
	FPDE	1.8189	0.3020	0.4847	0.8078	0.3899	0.7630	1.3534	0.8995
	Proposed	3.0158	0.4884	0.6134	0.8138	0.4859	0.7904	5.0230	0.9183

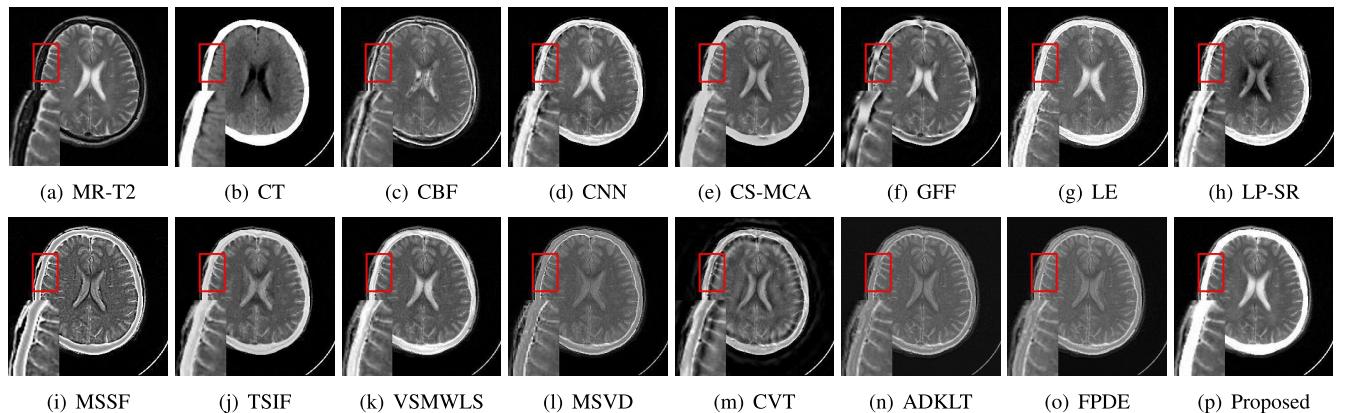


FIGURE 17. The fusion results on fourth group of MR-T2 image and CT image under 14 different methods.

The first group of MR-T2 image and CT image fusion results from 14 methods are shown in Fig. 13. First, we observe the information obtained from CT image, which highlights bone information. The information in (c) and (f) is lost a lot because CBF and GFF methods perform well in details preservation but bad in base information retention. The results MSVD, ADKLT and FPDE have low contrast as the fuse rule they use for base layers perform not well. Then

we pay attention to the central of (a). It is obvious that CS-MCA, LP-SR, MSSF, TSIF and CVT have poor performance. Combining the above two aspects, there are relatively excellent fusion visual effects in CNN, LE and proposed method whereas VSMWLS method has slight color distortion in center. The texture details in the red box are enlarged in the upper right corner of each image. There is faint texture information in the white outline in (d) and (g) which may

TABLE 5. The objective evaluation of the fusion results of the fourth group of images.

Groups	Metrics Methods	MI	Q^0	$Q^{AB/F}$	NCIE	VIFF	FSIM	IFC	VSI
Group 4 MR-T2/CT	CBF	2.6935	0.7036	0.5309	0.8121	0.2520	0.7264	5.5730	0.8777
	CNN	1.6231	0.5111	0.5073	0.8081	0.3768	0.7646	1.6165	0.8685
	CS-MCA	1.5324	0.4836	0.4947	0.8079	0.3460	0.7965	1.9216	0.8725
	GFF	2.0143	0.6563	0.5594	0.8089	0.2815	0.7584	3.7424	0.8903
	LE	3.3002	0.7904	0.5289	0.8166	0.4251	0.7659	6.2037	0.8792
	LP-SR	1.4764	0.4681	0.5057	0.8084	0.4092	0.7690	1.6658	0.8697
	MSSF	1.8797	0.4857	0.4412	0.8084	0.3819	0.6868	1.3537	0.8509
	TSIF	1.7107	0.5159	0.4647	0.8083	0.3921	0.7761	1.4553	0.8699
	VSMWLS	2.0310	0.6347	0.4812	0.8092	0.4190	0.7574	1.6192	0.8689
	MSVD	2.0980	0.4937	0.4434	0.8095	0.3236	0.7732	1.4756	0.8983
	CVT	1.3569	0.3375	0.4174	0.8070	0.2998	0.7414	1.1006	0.8657
	ADKLT	2.2019	0.3097	0.3878	0.8094	0.2742	0.7587	2.2830	0.8984
	FPDE	2.2570	0.3650	0.4033	0.8099	0.3059	0.7585	2.1646	0.8959
	Proposed	4.0211	0.7779	0.5664	0.8217	0.4307	0.8008	11.0180	0.9193

**FIGURE 18.** The fusion results about MR-T1 image (a1) MR-Gad image (a2) MR-PD (a3) and CT image (b) from 14 different methods. (c) CBF (d) CNN (e) CS-MCA (f) GFF (g) LE (h) LP-SR (i) MSSF (j) TSIF (k) VSMWLS (l) MSVD (m) CVT (n) ADKLT (o) FPDE (p) Proposed.

lead to misdiagnosis as we explained in the previous section. Besides, (d) and (p) retain more detail structure than (g). The subjective evaluation results are shown in Fig. 14(a). The bold

in TABLE 2 indicates the highest score in terms of objective evaluation. It is clear that the fused image obtained by the proposed method is better than others.

TABLE 6. The statistical results about objective evaluation indices on different fusion methods.

Metrics Methods	MI	Q^0	$Q^{AB/F}$	NCIE	VIFF	FSIM	IFC	VSI
CBF	2.0204±0.3976	0.4380±0.1526	0.5472±0.0449	0.8094±0.0014	0.2722±0.0412	0.7381±0.0310	2.7238±1.7663	0.8746±0.0154
CNN	1.5771±0.1633	0.4398±0.0567	0.5532±0.0325	0.8080±0.0006	0.4011±0.0627	0.7746±0.0241	1.6255±0.2655	0.8726±0.0151
CS-MCA	1.5594±0.1106	0.4270±0.0613	0.5442±0.0326	0.8078±0.0006	0.3688±0.0554	0.7953±0.0246	1.8536±0.1886	0.8821±0.0147
GFF	1.7390±0.2893	0.4410±0.1336	0.5841±0.0294	0.8083±0.0008	0.2983±0.0663	0.7724±0.0257	2.6104±1.0957	0.8866±0.0145
LE	2.3661±0.4697	0.5055±0.1543	0.5181±0.0485	0.8125±0.0020	0.4253±0.0563	0.7736±0.0257	2.9526±1.8839	0.8852±0.0120
LP-SR	1.5404±0.2895	0.4310±0.0469	0.5524±0.0320	0.8085±0.0007	0.4139±0.0619	0.7720±0.0377	1.6862±0.2185	0.8758±0.0141
MSSF	1.5377±0.1543	0.3854±0.0695	0.4324±0.0394	0.8074±0.0006	0.3790±0.0466	0.6895±0.0366	1.1277±0.1437	0.8516±0.0109
TSIF	1.5943±0.1092	0.4265±0.0609	0.4911±0.0359	0.8079±0.0006	0.4035±0.0550	0.7822±0.0233	1.3884±0.1672	0.8776±0.0136
VSMWLS	1.7362±0.1565	0.4752±0.0965	0.5050±0.0436	0.8084±0.0007	0.4198±0.0542	0.7636±0.0278	1.4730±0.2133	0.8744±0.0115
MSVD	1.7234±0.1888	0.3395±0.0811	0.4242±0.0498	0.8081±0.0006	0.3400±0.0397	0.7841±0.0280	1.1975±0.1891	0.9010±0.0124
CVT	1.3561±0.1145	0.3407±0.0292	0.4611±0.0302	0.8069±0.0006	0.3135±0.0539	0.7466±0.0244	1.1919±0.1149	0.8746±0.0100
ADKLT	1.7993±0.2417	0.2538±0.0440	0.3985±0.0350	0.8083±0.0007	0.2950±0.0420	0.7728±0.0255	1.6360±0.4187	0.9000±0.0120
FPDE	1.8447±0.2510	0.2838±0.0635	0.4058±0.0404	0.8085±0.0007	0.3281±0.0416	0.7720±0.0246	1.5331±0.3745	0.8990±0.0124
Proposed	3.0144±0.4917	0.5102±0.1496	0.5581±0.0422	0.8150±0.0030	0.4374±0.0608	0.8050±0.0258	4.9106±2.6251	0.9175±0.0088

The second group of MR-T2/CT fusion results are shown in Fig. 15 which reveals that most fused images preserve the outline information except for CBF, MSSF, MSVD, CVT, ADKLT and FPDE. It is because that multiscale transform based or filter based methods cannot completely extract the effective information from the base layers. In addition, the fuse rules used for base layers in these algorithms do not perform well. Observing the eyes, CNN, CS-MCA, LE, TSIF, VSMWLS and proposed method perform better. Further, the texture information marked by red and blue boxes is enlarged to the corresponding directions. The details lost in (d), (e), (j) and (g) successively increase, that is, CNN<CS-MCA<TSIF<LE. However, the proposed method can preserve more details as the anisotropic diffusion function is sensitive to the texture and the fuse algorithm using common coefficients has excellent performance. The “Maximum Absolute Value” rule extracts information from base layers effectively. The objective evaluation (in TABLE 3) and subjective evaluation (in Fig. 14(b)) indicate that the fused image obtained by proposed method performs better.

The third group of fused images are shown in Fig. 16. The result (c) has relatively poor performance in preserving the information from CT image. The white part, bone information is lost a lot. There is color distortion in (l), (m), (n) and (o) where the skeleton part appears gray in visual. The texture details in red box are enlarged. It is obvious that there is noise in LE and MSSF feels blurry. The lines in contours of CNN, LP-SR and VSMWLS may lead to misdiagnosis. By observing closely at the enlarged part, our method retains more details than CS-MCA and TSIF. The evaluation indices in TABLE 4 illustrate that GFF performs best in $Q^{AB/F}$ but the information at bottom in fused image is partly lost. The proposed algorithm retains effective information though it ranks second. In other aspects, our method performs better. The subjective evaluation in Fig. 14(c) indicates that the proposed method has stable fusion effect.

The fourth group of MR-T2/CT image fusion results are shown in Fig. 17. The result (m) has artifacts which reduce image quality in visual. The (l), (n) and (o) have lower contrast. Considering about the central part of the

TABLE 7. Average run time of different fusion methods on data set.

Methods	CBF	CNN	CS-MCA	GFF	LE	LP-SR	MSSF
Time/s	6.6671	15.6649	126.8838	0.1924	14.6510	0.0199	0.3877
Methods	TSIF	VSMWLS	MSVD	CVT	ADKLT	FPDE	Proposed
Time/s	0.2575	0.3023	0.1169	0.5083	0.1763	0.2073	9.7103

image, (c), (e), (f), (h), (i) and (j) lose information to varying degrees. The details in red box can be observed carefully. The skeleton information from CT image are retained better in proposed method, than CNN, LE and VSMWLS. The objective evaluation in TABLE 5 shows that LE gets the highest score in Q^0 . The proposed method is in second position. The subjective evaluation in Fig. 14(d) shows that the proposed algorithm has excellent performance in visual.

The average subjective evaluation scores on four groups of experiments are shown in Fig. 14(e). It is clear that the proposed method has relatively stable performance with good visual effects. Besides the MR-T2 images, which are used widely, there are other MRI images, such as MR-T1, MR-Gad and MR-PD. The fused images between them and CT images are shown in Fig. 18. It shows that the proposed method can perform well.

The statistical results about objective evaluation indices on 20 groups experimental data are concluded in TABLE 6. Except for the $Q^{AB/F}$, the proposed method obtains the highest scores. For $Q^{AB/F}$, GFF performs best, and our method takes the second place. In our opinion, the GFF method is good at preserving the edge information due to its principle of decomposing images. So it has better performance in $Q^{AB/F}$ but worse in Q^0 than our method. In general, the proposed method based on S-ADE model has the ability to fuse MRI images and CT images effectively, maintaining more useful information as well as ensuring the quality of the fused images.

E. COMPUTATIONAL EFFICIENCY ANALYSIS

All the algorithms are implemented using the MATLAB 2016b, the simulations are run on Dell Inspiron 3668 with Intel Core CPU 3.0 GHz and 8GB of RAM.

The computational efficiencies of methods compared above are concluded in TABLE 7. Among CNN, CS-MCA, GFF, LE and LP-SR methods, which have excellent performance in fused images, the CS-MCA method is a time-consuming one as shown in the table. The CNN model needs amounts of time to train medical images to improve the quality of fusion results. The LP-SR method uses a trained dictionary to realize fusion work. But the time of their preliminary work is not counted in the running time. The GFF has higher computational efficiency than our method, which is higher than LE method. The other fusion algorithms have shorter running time. However, they perform not well in fusion results. On the whole, the proposed fusion method has excellent fusion results with acceptable computational efficiency.

V. CONCLUSION

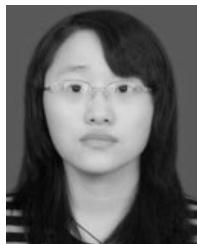
In this paper, we propose a medical image fusion method based on S-ADE model. First, the MRI images and CT images are decomposed into two layers: base layer and texture layer. Then, the “Maximum Absolute Value” rule is used for the base layer fusion. On the texture layer, we design the NSMAL fusion rule using common decomposition coefficients in order to retain more details. In this way, the image decomposition and fusion process have consistency in identifying textural information, which is beneficial to improve the quality of the fused image. Simultaneously, the consistency check using a sliding window is designed for decision map on texture layer to mitigate the staircase effect. Finally, correct the fusion results after image reconstruction. In the experiments, we select subjective evaluation and eight different objective evaluation metrics including MI, Q^0 , $Q^{AB/F}$, NCIE, VIFF, FSIM, IFC and VSI. Compared with 13 excellent algorithms, the proposed method gets the highest scores except for $Q^{AB/F}$. The $Q^{AB/F}$ uses the local metric to estimate the performance of the significant information from the source images in the fused one. The higher the $Q^{AB/F}$ value, the more the edge information of the fused image. The GFF method performs better because it belongs to a filtering algorithm that can maintain edges. Our method is in the second place with a small score gap. At the same time, the computational efficiency is acceptable. Overall, the experimental results demonstrate that the proposed method can maintain more information from the two source images and preserve the visual effect of the image compared with 13 fusion algorithms used widely.

In the future, we will work on the research about the diffusion function design and the combination of neural networks method to realize image fusion. At the same time, we are going to try to improve the computational efficiency. In addition, the situation of anatomical and function image fusion would be helpful for clinical diagnosis. The proposed method can be extended to fuse anatomical and function imaging modalities via some color space transform methods. The approach that avoids introducing structural and color distortion during the fusion process needs further study.

REFERENCES

- [1] L. Chang, X. Feng, X. Zhu, R. Zhang, R. He, and C. Xu, “CT and MRI image fusion based on multiscale decomposition method and hybrid approach,” *IET Image Process.*, vol. 13, no. 1, pp. 83–88, Jan. 2019.
- [2] R. Hou, D. Zhou, R. Nie, D. Liu, and X. Ruan, “Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model,” *Med. Biol. Eng. Comput.*, vol. 57, no. 4, pp. 887–900, Apr. 2019.
- [3] A. Krishn, V. Bhateja, Himanshi, and A. Sahu, “Medical image fusion using combination of PCA and wavelet analysis,” in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2014, pp. 986–991.
- [4] V. Bhateja, A. Krishn, and A. Sahu, “Medical image fusion in curvelet domain employing pca and maximum selection rule,” in *Proc. ICT*, Hyderabad, India, 2016, pp. 1–9.
- [5] S. Madanala and K. J. Rani, “PCA-DWT based medical image fusion using non sub-sampled contourlet transform,” in *Proc. Int. Conf. Signal Process., Commun., Power Embedded Syst. (SCOPES)*, Oct. 2016, pp. 1089–1094.
- [6] J. Reena Benjamin and T. Jayasree, “Improved medical image fusion based on cascaded PCA and shift invariant wavelet transforms,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 2, pp. 229–240, Feb. 2018.
- [7] J. Du, W. Li, B. Xiao, and Q. Nawaz, “Union Laplacian pyramid with multiple features for medical image fusion,” *Neurocomputing*, vol. 194, pp. 326–339, Jun. 2016.
- [8] A. Sahu, V. Bhateja, A. Krishn, and Himanshi, “Medical image fusion with Laplacian pyramids,” in *Proc. Int. Conf. Med. Imag., m-Health Emerg. Commun. Syst. (MedCom)*, Nov. 2014, pp. 448–453.
- [9] S. Li, X. Kang, and J. Hu, “Image fusion with guided filtering,” *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.
- [10] 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. 173019–173033, 2019.
- [11] B. K. Shreyamsha Kumar, “Image fusion based on pixel significance using cross bilateral filter,” *Signal, Image Video Process.*, vol. 9, no. 5, pp. 1193–1204, Jul. 2015.
- [12] Z. Xu, “Medical image fusion using multi-level local extrema,” *Inf. Fusion*, vol. 19, pp. 38–48, Sep. 2014.
- [13] R. Vijayarajan and S. Muttan, “Discrete wavelet transform based principal component averaging fusion for medical images,” *AEU Int. J. Electron. Commun.*, vol. 69, no. 6, pp. 896–902, Jun. 2015.
- [14] V. Bhavana and H. K. Krishnappa, “Multi-modality medical image fusion using discrete wavelet transform,” *Procedia Comput. Sci.*, vol. 70, pp. 625–631, Jan. 2015.
- [15] X. Xu, Y. Wang, and S. Chen, “Medical image fusion using discrete fractional wavelet transform,” *Biomed. Signal Process. Control*, vol. 27, pp. 103–111, May 2016.
- [16] F. E. Ali, I. M. El-Dokany, A. A. Saad, and F. E. Abd El-Samie, “A curvelet transform approach for the fusion of MR and CT images,” *J. Modern Opt.*, vol. 57, no. 4, pp. 273–286, Feb. 2010.
- [17] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, “Remote sensing image fusion using the curvelet transform,” *Inf. Fusion*, vol. 8, no. 2, pp. 143–156, Apr. 2007.
- [18] G. Bhatnagar, Q. M. J. Wu, and Z. Liu, “Directive contrast based multi-modal medical image fusion in NSCT domain,” *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1014–1024, Aug. 2013.
- [19] Z. Zhu, M. Zheng, G. Qi, D. Wang, and Y. Xiang, “A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain,” *IEEE Access*, vol. 7, pp. 20811–20824, 2019.
- [20] M. M. I. Ch, M. M. Riaz, N. Iltaf, A. Ghafoor, and A. Ahmad, “Weighted image fusion using cross bilateral filter and non-subsampled contourlet transform,” *Multidimensional Syst. Signal Process.*, vol. 30, no. 4, pp. 2199–2210, Oct. 2019.
- [21] X. Liu, W. Mei, and H. Du, “Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion,” *Neurocomputing*, vol. 235, pp. 131–139, Apr. 2017.
- [22] P. Ganasala and A. D. Prasad, “Medical image fusion based on frei-chen masks in NSST domain,” in *Proc. 5th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2018, pp. 619–623.
- [23] R. Singh and A. Khare, “Fusion of multimodal medical images using daubechies complex wavelet transform—A multiresolution approach,” *Inf. Fusion*, vol. 19, pp. 49–60, Sep. 2014.
- [24] Y. Liu, S. Liu, and Z. Wang, “A general framework for image fusion based on multi-scale transform and sparse representation,” *Inf. Fusion*, vol. 24, pp. 147–164, Jul. 2015.

- [25] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, "Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning," *Pattern Recognit.*, vol. 79, pp. 130–146, Jul. 2018.
- [26] J.-J. Zong and T.-S. Qiu, "Medical image fusion based on sparse representation of classified image patches," *Biomed. Signal Process. Control*, vol. 34, pp. 195–205, Apr. 2017.
- [27] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Medical image fusion via convolutional sparsity based morphological component analysis," *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 485–489, Mar. 2019.
- [28] D. P. Bavirisetti and R. Dhuli, "Multi-focus image fusion using multi-scale image decomposition and saliency detection," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 1103–1117, Dec. 2018.
- [29] D. P. Bavirisetti and R. Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," *Infra. Phys. Technol.*, vol. 76, pp. 52–64, May 2016.
- [30] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infra. Phys. Technol.*, vol. 82, pp. 8–17, May 2017.
- [31] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Xi'an, China, Jul. 2017, pp. 1–7.
- [32] K.-J. Xia, H.-S. Yin, and J.-Q. Wang, "A novel improved deep convolutional neural network model for medical image fusion," *Cluster Comput.*, vol. 22, no. S1, pp. 1515–1527, Jan. 2019.
- [33] J. Wang, X. Li, Z. Wang, H. Duan, and X. Zhang, "Exposure correction using deep learning," *J. Electron. Imag.*, vol. 28, no. 3, May 2019, Art. no. 033003.
- [34] Z. Wang, X. Li, H. Duan, X. Zhang, and H. Wang, "Multifocus image fusion using convolutional neural networks in the discrete wavelet transform domain," *Multimedia Tools Appl.*, vol. 78, no. 24, pp. 34483–34512, Dec. 2019.
- [35] J. Ma, P. Liang, W. Yu, C. Chen, X. Guo, J. Wu, and J. Jiang, "Infrared and visible image fusion via detail preserving adversarial learning," *Inf. Fusion*, vol. 54, pp. 85–98, 2020.
- [36] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "FusionGAN: A generative adversarial network for infrared and visible image fusion," *Inf. Fusion*, vol. 48, pp. 11–26, Aug. 2019.
- [37] J. Ma, H. Xu, J. Jiang, X. Mei, and X.-P. Zhang, "DDeGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 4980–4995, 2020.
- [38] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Inf. Fusion*, vol. 31, pp. 100–109, Sep. 2016.
- [39] E. Daniel, J. Anitha, and J. Gnanaraj, "Optimum Laplacian wavelet mask based medical image using hybrid cuckoo search—Grey wolf optimization algorithm," *Knowl.-Based Syst.*, vol. 131, pp. 58–69, Sep. 2017.
- [40] 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. 40782–40796, 2019.
- [41] V. S. Parvathy and S. Pothiraj, "Multi-modality medical image fusion using hybridization of binary crow search optimization," *Health Care Manage. Sci.*, pp. 1–9, Jul. 2019.
- [42] R. R. Nair, E. David, and S. Rajagopal, "A robust anisotropic diffusion filter with low arithmetic complexity for images," *EURASIP J. Image Video Process.*, vol. 2019, no. 1, p. 48, Dec. 2019.
- [43] M. Al-nasrawi, G. Deng, and W. Waheed, "Structure extraction of images using anisotropic diffusion with directional second neighbour derivative operator," *Multimedia Tools Appl.*, vol. 78, no. 5, pp. 6385–6407, Mar. 2019.
- [44] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.
- [45] D. P. Bavirisetti and R. Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loeve transform," *IEEE Sensors J.*, vol. 16, no. 1, pp. 203–209, Jan. 2016.
- [46] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," in *Proc. 20th Int. Conf. Inf. Fusion (Fusion)*, Xi'an, China, Jul. 2017, pp. 1–9.
- [47] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electron. Lett.*, vol. 38, no. 7, p. 313, 2002.
- [48] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [49] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, Feb. 2000.
- [50] Q. Wang, Y. Shen, and J. Jin, "Performance evaluation of image fusion techniques," *Image Fusion, Algorithms Appl.*, vol. 19, pp. 469–492, Jun. 2008.
- [51] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. Fusion*, vol. 14, no. 2, pp. 127–135, Apr. 2013.
- [52] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [53] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.
- [54] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.
- [55] V. P. S. Naidu, "Image fusion technique using multi-resolution singular value decomposition," *Defence Sci. J.*, vol. 61, no. 5, p. 479, 2011.



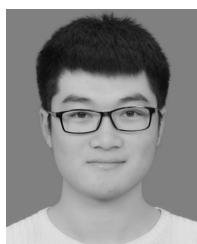
RUI ZHU received the B.S. degree in software engineering and the master's degree from Jilin University, in 2016, where she is currently pursuing the Ph.D. degree. Her research specializes in image processing and deep learning.



XIONGFEI LI (Member, IEEE) received the B.S. degree in computer software from Nanjing University, in 1985, the M.S. degree in computer software from the Chinese Academy of Sciences, in 1988, and the Ph.D. degree in communication and information system from Jilin University, in 2002. Since 1988, he has been a member of the faculty of the Computer Science and Technology, Jilin University, Changchun, China. He is currently a Professor of computer software and theory with Jilin University. He has authored more than 60 research articles. His research interests include data mining, intelligent network, image processing, and analysis.



XIAOLI ZHANG received the B.S. degree in computer science and engineering from Dalian Nationalities University, in 2010, and the Ph.D. degree in computer science and technology from Jilin University, in 2016. He is currently an Associate Professor with the College of Computer Science and Technology, Jilin University. He has authored more than 30 research articles. His research interests include image fusion and machine learning.



MINGRUI MA received the B.S. degree from Jilin University, in 2018, where he is currently pursuing the master's degree. His research interests include image fusion and image enhancement.