

Research Article

An Image Decomposition Fusion Method for Medical Images

Lihong Chang , Wan Ma, Yu Jin, and Li Xu

School of Mathematics and Computer Science, Ningxia Normal University, Guyuan 756000, Ningxia Province, China

Correspondence should be addressed to Lihong Chang; changlihong-1999@126.com

Received 19 November 2019; Revised 25 March 2020; Accepted 17 April 2020; Published 29 July 2020

Academic Editor: Alessandro Mauro

Copyright © 2020 Lihong Chang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A fusion method based on the cartoon+texture decomposition method and convolution sparse representation theory is proposed for medical images. It can be divided into three steps: firstly, the cartoon and texture parts are obtained using the improved cartoon-texture decomposition method. Secondly, the fusion rules of energy protection and feature extraction are used in the cartoon part, while the fusion method of convolution sparse representation is used in the texture part. Finally, the fused image is obtained using superimposing the fused cartoon and texture parts. Experiments show that the proposed algorithm is effective.

1. Introduction

With the development of imaging equipment and technology, there are different modality medical images which may reflect different organ or tissue information. For example, the computed tomography (CT) image can precisely exhibit dense structures such as bones and implants, while the magnetic resonance (MR) images detect enough soft tissue information with high-resolution anatomical details but are less sensitive to the diagnosis of fractures than CT. To obtain sufficient image information for accurate diagnosis, a doctor often needs to sequentially analyze different modality images, but this one by one manner way still brings inconvenience in many cases. The aim of the medical image fusion method is to generate a single comprehensive image contained in multiple medical images with different modalities, which are more suitable for doctor diagnosis. In fact, the fusion method can not only provide diagnosis information for doctors but also provide auxiliary treatment information [1, 2].

Over the last few years, a variety of medical image fusion methods have been proposed in various clinical applications. According to a recent survey [3], there are three categories based on multiscale decomposition (MSD) methods, based on learning representation methods, and based on combination of different methods. Classical MSD-based fusion methods [4–7] assume that the salient information of the

source images is contained in the decomposition coefficients. Obviously, the selection of transform methods and decomposition levels are very important. Li et al. [8] proposed a comparative study of different MSD-based methods, where they found that the fusion method based on NSCT can generally obtain the best fusion effects. Based on learning representation methods include sparse representation (SR) [9, 10], parameter-adaptive pulse-coupled neural network (PAPCNN) [11], convolutional sparse representation (CSR) [12], convolutional neural network (CNN) [13], convolutional sparsity-based morphological component analysis (CSMCA) [14], and deep learning (DL) [15, 16]. These methods represent the information of the source image by learning dictionary or learning model. Compared with the MSD-based fusion methods, the DL-based methods can achieve better results. The methods based on combination of different methods overcome the shortcomings of the single method. For an example, the fusion method based on MSD (the high-pass bands are merged with the “max-absolute” rule, while the low-pass bands are fused using the “averaging” rule) has two main drawbacks: the loss of contrast and the difficulty of selecting level.

Yang et al. [9] first introduced the SR theory into image fusion field. The early fusion methods based on SR use the standard sparse coding which is applied local patches. After this practice, there are many image fusion methods based on SR. These fusion methods generally try to improve fusion

performance by adding constraints [17, 18] and designing the effective strategies of learning dictionary [19]. The standard SR methods often have three defects: the details tend to be smoothed, spatial inconsistency, and the low computational efficiency. To solve these problems, there are many effective improved algorithms which aim at learning a compact and efficient dictionary. Qiu et al. [20] learnt a discriminative dictionary using the mutual information rule. In order to improve the localization and recognition of multiple objects, Siyahjani and Doretto [21] proposed a context aware dictionary. Qi et al. [22] learnt an integrated dictionary which used an entropy-based algorithm for informative block selection. They used the online dictionary learning algorithm to extract discriminative features from high-frequency components, which enhances the accuracy and efficiency of the fusion result.

Convolutional sparsity representation (CSR) model, unlike the standard SR model [7], is based on single image and overlapping patches in the original spatial domain for sparse coding; it is a global SR model of the source image. In addition, the sparse coding of the CSR model is performed over the entire image rather than on overlapping patches. Therefore, the fusion method based on CSR achieves better representations.

The representation of different components of the image has become a hot topic in recent research. Any image f can be decomposed into a cartoon part and a textural part. The cartoon part is composed of the image contrasted shapes such as strong edges, while the textural part consists of the oscillating patterns. In other words, the cartoon part contains low-frequency components, and the texture part contains middle and high frequencies. For different components, the fusion strategies can be more effective designing. In conventional transform methods, the fusion strategies were selected by weighted-average or choose-max in the cartoon part and the textural part, respectively. These fusion rules often reduce the contrast of the image. To improve the shortcomings of the traditional fusion rules, there are many improved fusion methods. Zhu et al. [23] proposed a medical image fusion method based on cartoon-texture decomposition and selected sparse representation as fusion rules. Yin et al. [8] proposed a parameter-adaptive pulse-coupled neural network (PAPCNN) model in the high-frequency bands based on the non-subsampled shearlet transform domain. Specially, in the PAPCNN model, the fusion rule of the low-frequency bands selected weighted local energy (WLE) and weighted sum of eight-neighborhood-based modified Laplacian (WSEML). WLE and WSEML have the effect of energy preservation and detail extraction. This point just makes up for the deficiency that the texture details are often remained in the cartoon part. Liu et al. [13] proposed a fusion method based on CSMCA for the medical image. In the CSMCA model, the cartoon and texture components of each source image are not only obtained by prelearned dictionaries but also the cartoon and texture components are fused by the preset dictionary in the process of fusion. Therefore, all representations of image information are affected by the quality of the dictionary in the process of information representation and fusion. In

order to reduce the influence of the dictionary on the fusion results, we try to use an improved fast cartoon-texture decomposition (IFCTD) [10] instead of dictionary decomposition. It is confirmed in [10] that IFCTD is more effective using image decomposition.

A fast cartoon-texture decomposition (FCTD) [24] applied a pair of low-high pass filters; therefore, it is fast and simple. However, it blurs strong edges and retains certain textures in the cartoon part. One of the reasons for these results is that the edge maps computed are used as a local gradient computation which utilizes a few pixels around the central pixel. The local gradient operator is inaccurate for the noise image. In order to improve the stability of gradient, we use the global sparse gradient (GSG) [25] instead of local operators to improve FCTD. GSG uses more information around the central pixel. It is more stable for noise. Figure 1 shows an example of various gradient operators and the GSG on a noisy image ($\sigma = 20$).

Figure 2 shows an example of cartoon + texture decomposition for medical images using the FCTD method and the IFCTD method, respectively. Figures 2(c)–2(f) show the decomposition results of FCTD, and 2(g)–2(j) give the decomposition results of IFCTD, respectively. 2(c) and 2(g) are cartoon parts of CT images; 2(e) and 2(i) are texture parts of CT images; 2(d) and 2(h) are cartoon parts of MR images; and 2(f) and 2(j) are texture parts of MR images. From the experimental results, the result of IFCTD can extract details better than the result of FCTD. In the amplification texture parts of the MRI, the result of the IFCTD (Figure 2(j)) contains more texture details than the result of the FCTD (Figure 2(f)).

In addition to using the IFCTD tool instead of the prelearned dictionaries, in order to better protect the energy of the cartoon part in the fusion process, we use the fusion rule of energy protection (WEL and WSEML) in the cartoon part and use the fusion rule of the texture part of CSR.

The main contribution of this paper is to introduce the IFCTD into the cartoon + texture decomposition for medical images and combine the energy protection method and CSR method to improve the fusion effects.

The rest of this paper is organized as follows. In Section 2, the CSR model is briefly introduced. Section 3 describes the proposed method in detail. Section 4 presents experiments and discussion. Finally, the conclusions are reported in Section 5.

2. CSR Model

SR-based image fusion method was first introduced by Yang and Li [9]. In this model, source images were divided into a large number of overlapping patches using the sliding window technique, and then the “max L1-norm” of the sparse coefficient vector was selected as the activity level measurement. SR has been widely used in image fusion. These methods have achieved great success. However, it is worthwhile to notice that these methods have some defects, such as (1) the SR-based methods are only shift-invariant when the stride of patches is one pixel in both vertical and horizontal directions, (2) the fine details in source images

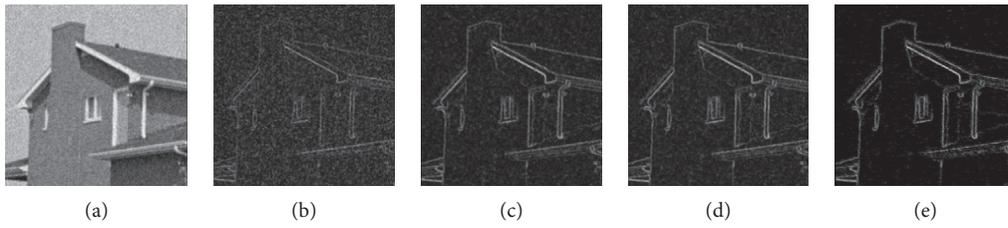


FIGURE 1: Comparison of various gradient operators and the GSG on the noisy image ($\sigma = 20$). (a) Noisy image. (b) Roberts. (c) Prewitt. (d) Sobel. (e) GSC.

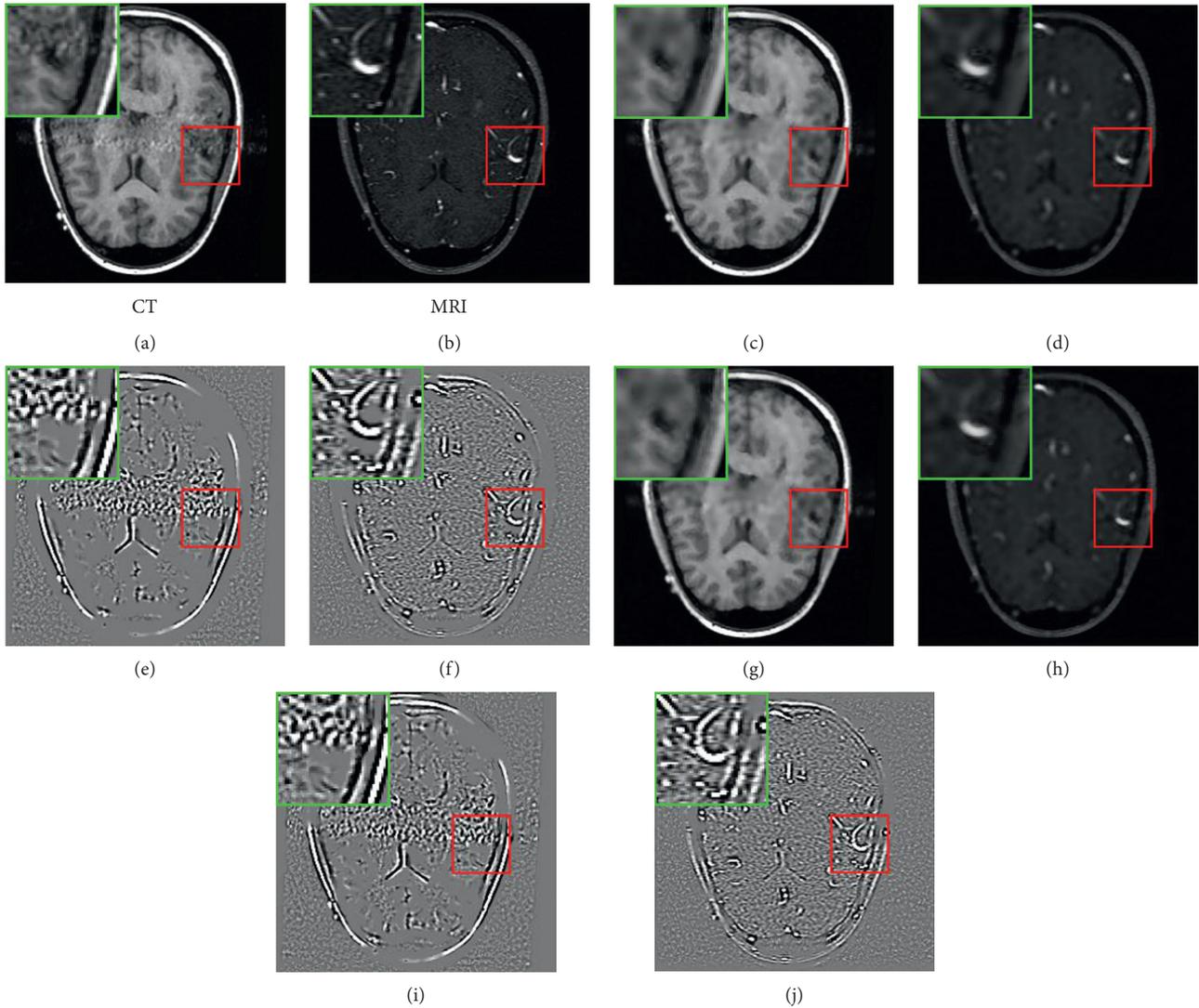


FIGURE 2: An example of cartoon + texture decomposition of FCTD and IFCTD: (c), (g) and (e), (i) the cartoon part of CT and texture part of CT; (d), (h) and (f), (j) the cartoon part of MR and texture part of MR, respectively.

like textures and edges tend to be smoothed, (3) the “max-L1” rule may cause spatial inconsistency in the fused results for different modality images, and (4) the computational efficiency is low because the sliding window’s step length should be small enough in the sparse coding technique. In a word, the reasons for these defects are caused by patch-based coding which is performed on overlapping patches to

achieve better representations. In order to solve these problems, CSR mode is introduced by Liu et al. [12]. The sparse coding is a global sparse representation, which is performed over the entire image.

CSR model can be seen as an alternative representation to SR using the convolutional form. This model can be formulated as the sum over a set of M convolutions between

global sparse coefficients $\{X_m\}$ and local dictionary filters $\{d_m\}$ [26]:

$$\min_{X_m} \frac{1}{2} \left\| I - \sum_{m=1}^M d_m * X_m \right\|_2^2 + \lambda \sum_{m=1}^M \|X_m\|_1, \quad (1)$$

where I is an entire image, λ is a regularization parameter, and $*$ denotes a convolution operator. CSR is translation/shift-invariant sparse representation [27]. In the process of optimizing representation, CSR is single valued over the entire image; therefore, the details should be preserved from source images.

$$|\nabla t| = \arg \min_{t=(t_1, t_2)} \left\{ \frac{1}{|\Omega|} \iint_{\Omega \times \Omega} \delta_{xy}^s (h(x) - h(y)) + t(x) \cdot (y - x)^2 dy dx + \beta \int_{\Omega} |t(\mathbf{x})| d\mathbf{x} \right\}, \quad (2)$$

where $t(\mathbf{x}) = (t_1(\mathbf{x}), t_2(\mathbf{x}))$, $\delta_{xy}^s = \exp(-|y - x|^2/\rho^2)$, and ρ and $\beta > 0$ are parameters.

The local total variation (LTV) is obtained by convolution with the gradient norm of f and $L_\sigma * f$ and the Gauss kernel G_σ :

$$LTV_\sigma(x)(f) := G_\sigma * |\nabla f|(x). \quad (3)$$

Set $\lambda(x) := (LTV_\sigma(x)(f) - LTV_\sigma(L_\sigma * f))/LTV_\sigma(x)(f)$.

(2) We obtain the cartoon image $u(x)$ and the texture image $v(x)$:

$$u(x) = \omega(\lambda(x))L_\sigma * f + (1 - \omega(\lambda(x)))f, v(x) = f - u(x), \quad (4)$$

where $\omega(x): [0, 1] \rightarrow [0, 1]$ is the soft threshold function (Figure 3).

In these two parts, one of the key elements is how to choose the appropriate fusion rules to improve the fusion effect.

3.1. Fusion Ruler of Cartoon Parts. In IFCTD, I_{1c} and I_{2c} are obtained by the low-pass filters; therefore, the main energy of the source image is concentrated in the cartoon parts of medical images. Because the imaging mechanism of CT and MR images is different, the intensity of their same location is different. If the averaging strategy is applied to the cartoon parts of CT and MR images, the brightness of some areas will decrease dramatically, which will reduce the visual perception ability. In addition, due to some factors, there always exists a limitation on cartoon and texture decomposition. In other words, the cartoon part still contains some texture information. In order to better extract texture details and protect energy, the WLE and WSEML are selected by Yin et al. [10]. The WLE and WSEML are defined as

3. Proposed Fusion Method

Suppose that there are two pre-registered source images denoted as I_1, I_2 . The cartoon parts (I_{1c}, I_{2c}) and the texture parts (I_{1t}, I_{2t}) of I_1, I_2 are obtained using IFCTD. The IFCTD method uses a pair of low-high pass filters to decompose the image. Its main process is as follows:

- (1) Use the low-pass filter L_σ on the original image f , and calculate the Euclidean norm of the gradients of f and $L_\sigma * f$. The gradient is calculated as follows:

$$WLE_s(i, j) = \sum_{m=-r}^r \sum_{n=-r}^r W(m+r+1, n+r+1) \cdot C_s(i+m, j+n)^2, \quad (5)$$

$$WSEML_s(i, j) = \sum_{m=-r}^r \sum_{n=-r}^r W(m+r+1, n+r+1) EML_s \cdot (i+m, j+n), \quad (6)$$

$$\begin{aligned} EML_s(i, j) = & |2S(i, j) - S(i-1, j) - S(i+1, j)| \\ & + |2S(i, j) - S(i, j-1) - S(i, j+1)| \\ & + \frac{1}{\sqrt{2}} |2S(i, j) - S(i-1, j-1) - S(i+1, j+1)| \\ & + \frac{1}{\sqrt{2}} |2S(i, j) - S(i-1, j+1) - S(i+1, j-1)|, \end{aligned} \quad (7)$$

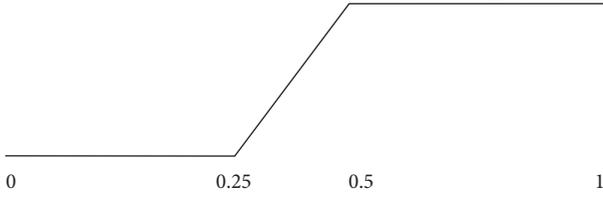
where $s \in \{I_{1c}, I_{2c}\}$, w is a $(2r+1) \times (2r+1)$ weighting matrix with radius r . The value of the weighting matrix is set to 2^{2r-d} (d is its four-neighborhood distance to the center). In this paper, set $r = 1$ and $d = 0, 1, 2$, and we select the 3×3 normalized eight neighboring version of w . It is

$$\text{expressed as } W = 1/16 \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Finally, the strategy of cartoon part fusion is

$$I_{fc}(i, j) = \begin{cases} I_{1c}, & \text{if } WLE_{1c(i,j)} * WSEML_{1c(i,j)} \\ & \geq WLE_{2c(i,j)} * WSEML_{2c(i,j)}, \\ I_{2c}, & \text{otherwise.} \end{cases} \quad (8)$$

3.2. Fusion Ruler of Textural Parts. Suppose a set of dictionary filters $d_{t,m}$ ($m \in \{1, \dots, M\}$) have been learned by the dictionary learning method in [22]. For the textural parts I_{1t} and I_{2t} , their sparse coefficient maps $X_{I_{1t},m}$ and $X_{I_{2t},m}$ are obtained by solving the CSR model with the method in [22]:

FIGURE 3: The soft threshold function $\omega(x)$.

$$\min_{X_{I_{1t},m}} \frac{1}{2} \left\| I_{1t} - \sum_{m=1}^M d_{t,m} * X_{I_{1t},m} \right\|_2^2 + \lambda_{1t} \sum_{m=1}^M \|X_{I_{1t},m}\|_1, \quad (9)$$

$$\min_{X_{I_{2t},m}} \frac{1}{2} \left\| I_{2t} - \sum_{m=1}^M d_{t,m} * X_{I_{2t},m} \right\|_2^2 + \lambda_{2t} \sum_{m=1}^M \|X_{I_{2t},m}\|_1. \quad (10)$$

Let $X_{k,1:M}(x, y)$ denote the contents of $X_{k,m}$ ($k = I_{1t}, I_{2t}$) at the position (x, y) in the textural part. Thus, the activity level map is selected as window-based averaging strategy:

$$\bar{A}_n^k(x, y) = \frac{\sum_{p=-r_n}^{r_n} \sum_{q=-r_n}^{r_n} \|X_{n,1:M}^k(x, y)\|_1}{(2r_n + 1)^2}, \quad n \in \{I_{1t}, I_{2t}\}, \quad (11)$$

where r is the size of the window.

Then, the “choose-max” rule is applied to achieve the fused coefficient maps:

$$X_{n,1:M_n}^f(x, y) = X_{n,1:M_n}^{k^*}(x, y), \quad k^* = \arg \max_k \left(\bar{A}_n^k(x, y) \right). \quad (12)$$

The fused texture part can be expressed as

$$I_{ft} = \sum_{m=1}^{M_t} d_{t,m} * X_{t,m}^f. \quad (13)$$

3.3. Final Fusion Results. In the process of cartoon and texture decomposition, the texture part is obtained by subtracting the cartoon part from the source image, so the final fusion result is still obtained by simply stacking the fused cartoon part and texture part.

$$I_f = I_{fc} + I_{ft}. \quad (14)$$

Figure 4 shows the fusion flowchart of the proposed method.

4. Experimental Results and Analysis

4.1. Testing Images. In our experiments, the eight pairs of medical images are used as test images which are collected from Yu Liu’s personal homepage (<http://home.ustc.edu.cn/~liuyul/>) and the website web page: <http://www.imagefusion.org/>, Figure 5. The first line is the CT images, and the second line is the corresponding MR images. We assumed that each pair of source images is pre-registered.

4.2. Objective Evaluation Metrics of Image Fusion Effect.

To measure the performance of the algorithm, five popular objective metrics are applied to evaluate the fusion results from different aspects. They are the entropy (EN), standard deviation (SD), normalized mutual information (MI) [28], gradient-based fusion metric Q_G [29], and phase congruency metric Q_p [30]. EN is used to measure the amount of information in the fused image. SD is used to measure the overall contrast of the fused image. MI represents the amount of information the fused image obtains from the source image. Q_G computes the amount of gradient information injected into the fused image from the source images. Q_p measures the extent that the salient features in the source images are preserved. When the value of the five metrics is greater, the fusion effect is improved.

4.3. Experimental Discussion.

Because the proposed method, namely, CTCSSR, is mainly aimed at the improvement of the SR-based methods, the comparison algorithm chooses the fusion method based on SR, such as standard SR [9], a hybrid cartoon texture sparse representation method (BSR) [10], PAPCNN [11], CSR [12], and CSMCA [14]. All the parameters are set to the recommended values as reported in [9–12, 14]. Apart from PAPCNN, the dictionaries of other SR-based methods have 256 atoms and are learned by the K-SVD method from natural image patches. For the CTCSSR method, the spatial size of each dictionary filter is set to 8×8 , which is the same size of the dictionary atom in other SR-based methods. Specially, dictionary filters $\{d_{t,m}\}_{m=1}^M$ are learned from the textural parts of 20 high-quality natural images (the size of the image is 256×256) using the method in [22]. In our experiments, the number of dictionary filters M is set 32.

All of the fusion methods are implemented on the platform of the HP-Z600 workshop (Four Core 2.4 GHz CPU and 8G RAM), Matlab R2017b programming environment in the Windows 7 operating systems.

Figures 6–9 show four examples with different fusion methods for CT and MR images. We magnify the marked red rectangle part of the experimental results and show them the marked green rectangle part.

In Figure 6, it can be seen that the performances of the SR and BSR methods suffer from obvious undesirable visual artifacts. The effects of the PAPCNN, CSR, and CTCSSR methods are basically the same; these methods enhance the anatomical details (in the enlarged part). The CSMCA method is relatively low contrast.

Figure 7 shows a set of C1 and C2 image fusion results. Because structural details are mainly contained in the MR image, almost all of these methods can extract details well. But, from the partially enlarged image, the details are seriously blurred in the SR method. The BSR, PAPCNN, and CSR methods lose a lot of information of I_2 , while CSMCA and CTCSSR methods keep more details of I_2 . From the perspective of visual perception, the fusion result of the proposed method is better.

Figure 8 shows a set of E1 and E2 image fusion results. The BSR, PAPCNN, and CTCSSR methods have almost the

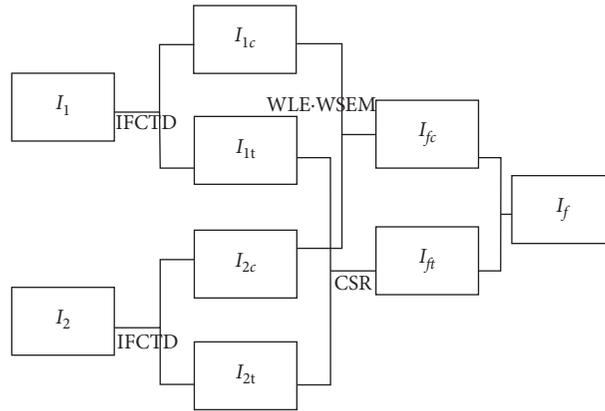


FIGURE 4: The fusion flow chart of the proposed method.

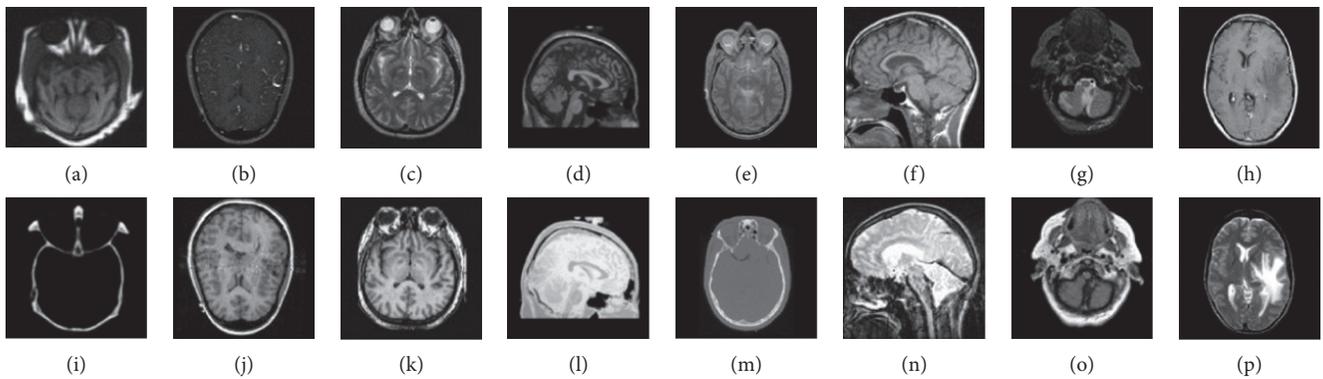


FIGURE 5: Testing images. (a) A1. (b) B1. (c) C1. (d) D1. (e) E1. (f) F1. (g) H1. (h) G1. (i) A2. (j) B2. (k) C2. (l) D2. (m) E2. (n) F2. (o) H2. (p) G2.

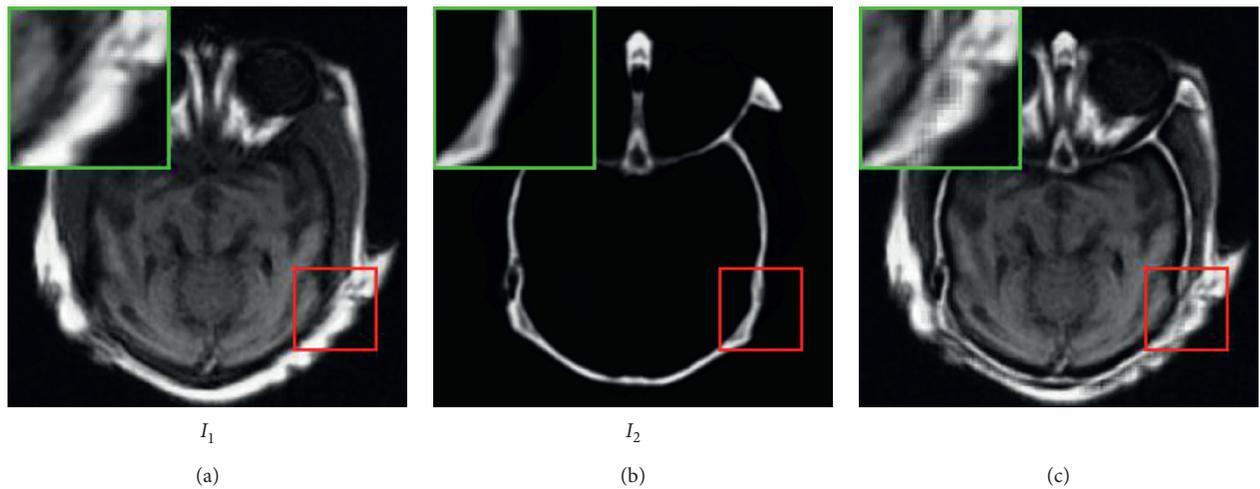


FIGURE 6: Continued.

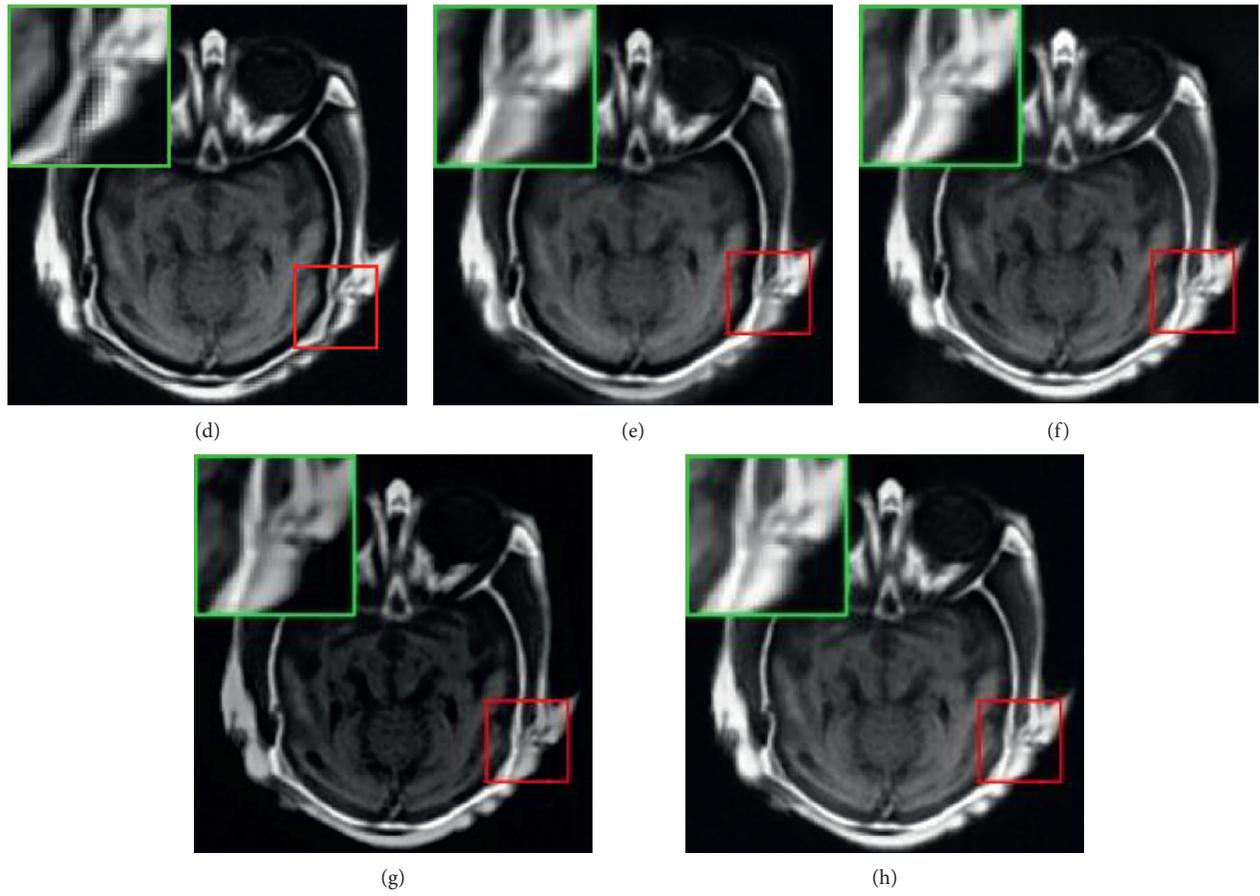


FIGURE 6: An example for the fusion results of the different methods. (c) SR. (d) BSR. (e) PAPCNN. (f) CSR. (g) CSMCA. (h) CTCRS.

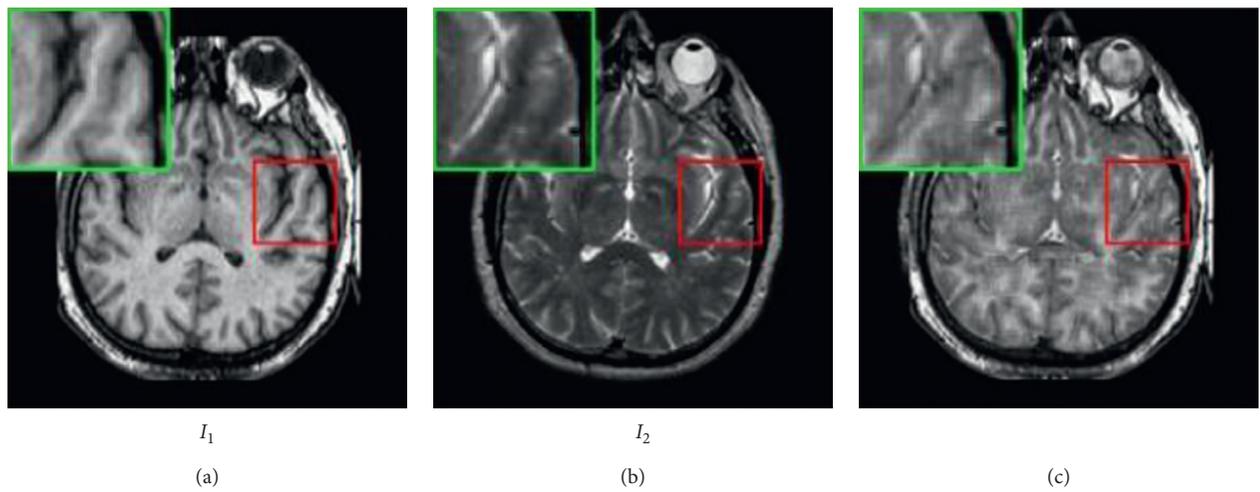


FIGURE 7: Continued.

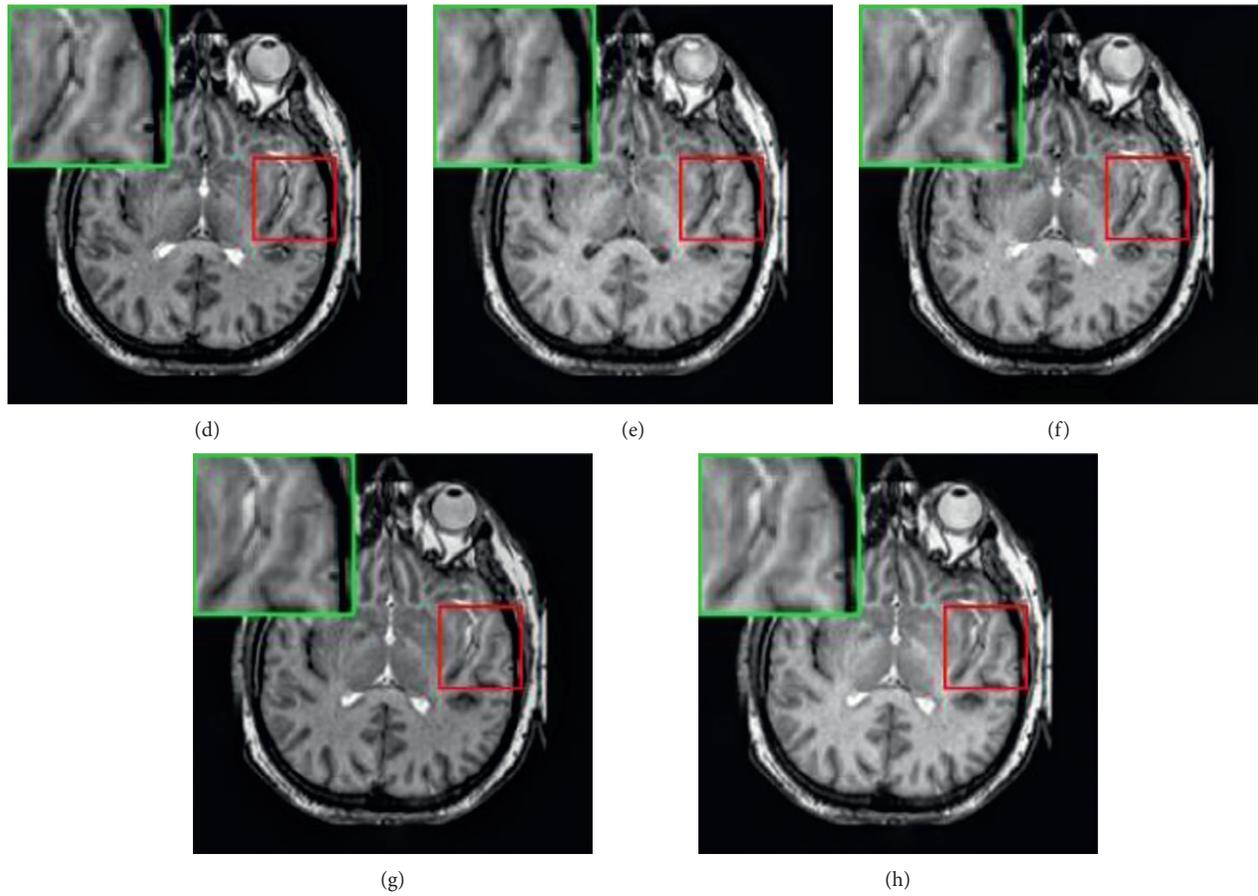


FIGURE 7: An example for the fusion results of the different methods. (c) SR. (d) BSR. (e) PAPCNN. (f) CSR. (g) CSMCA. (h) CTC SR.

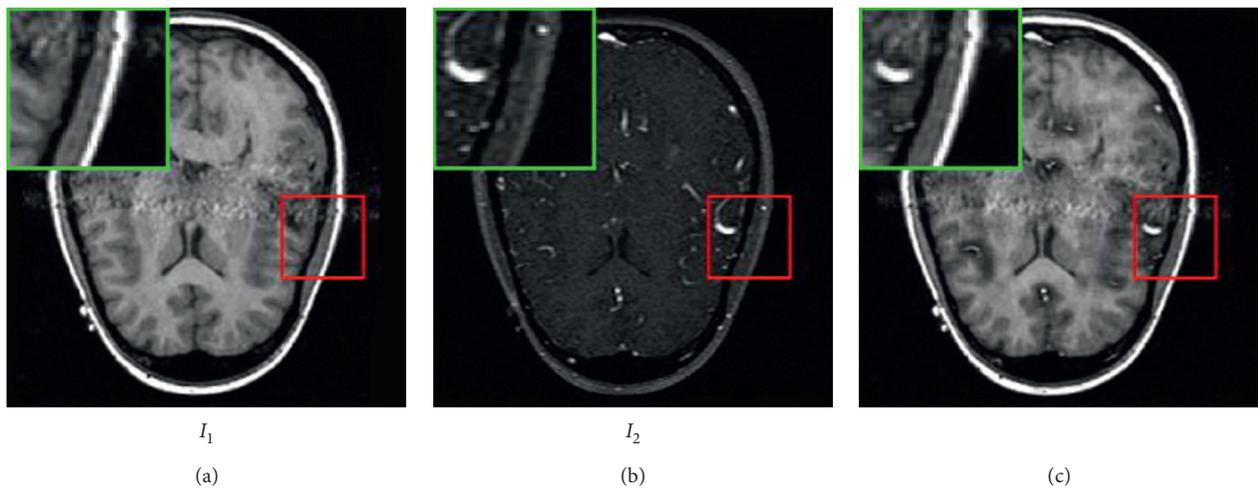


FIGURE 8: Continued.

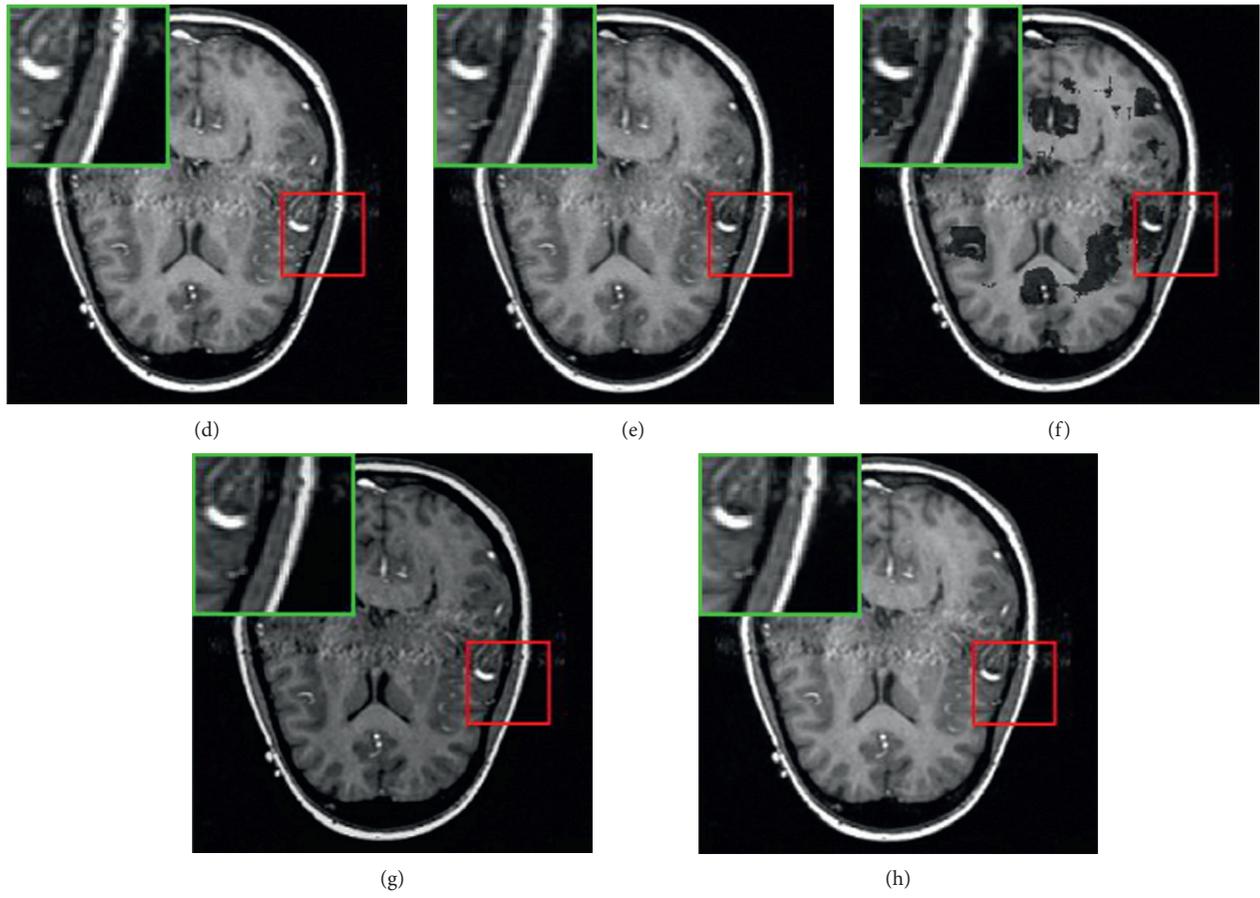


FIGURE 8: An example for the fusion results of the different methods. (c) SR. (d) BSR. (e) PAPCNN. (f) CSR. (g) CSMCA. (h) CTCRS.

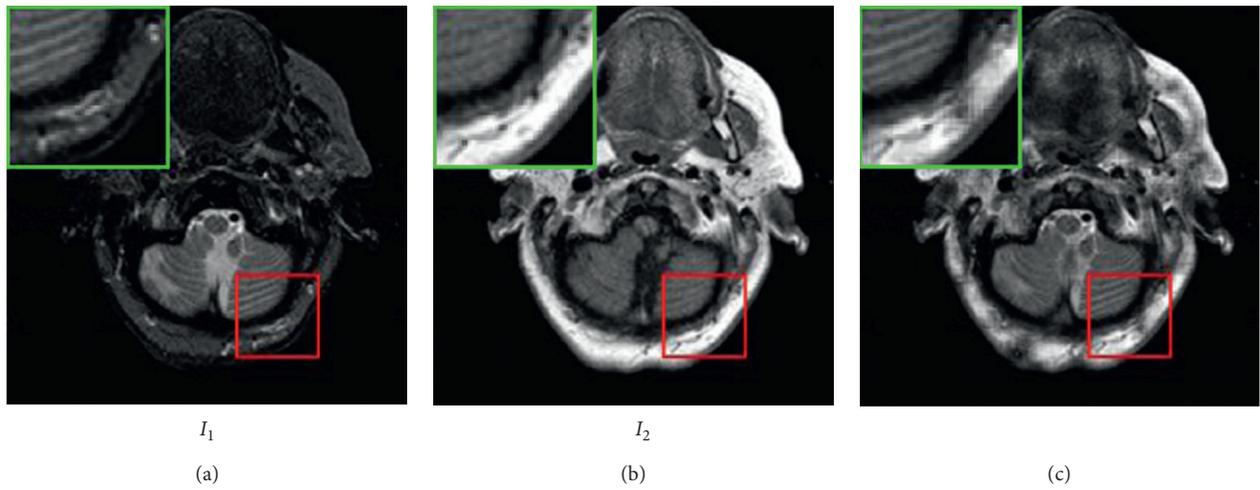


FIGURE 9: Continued.

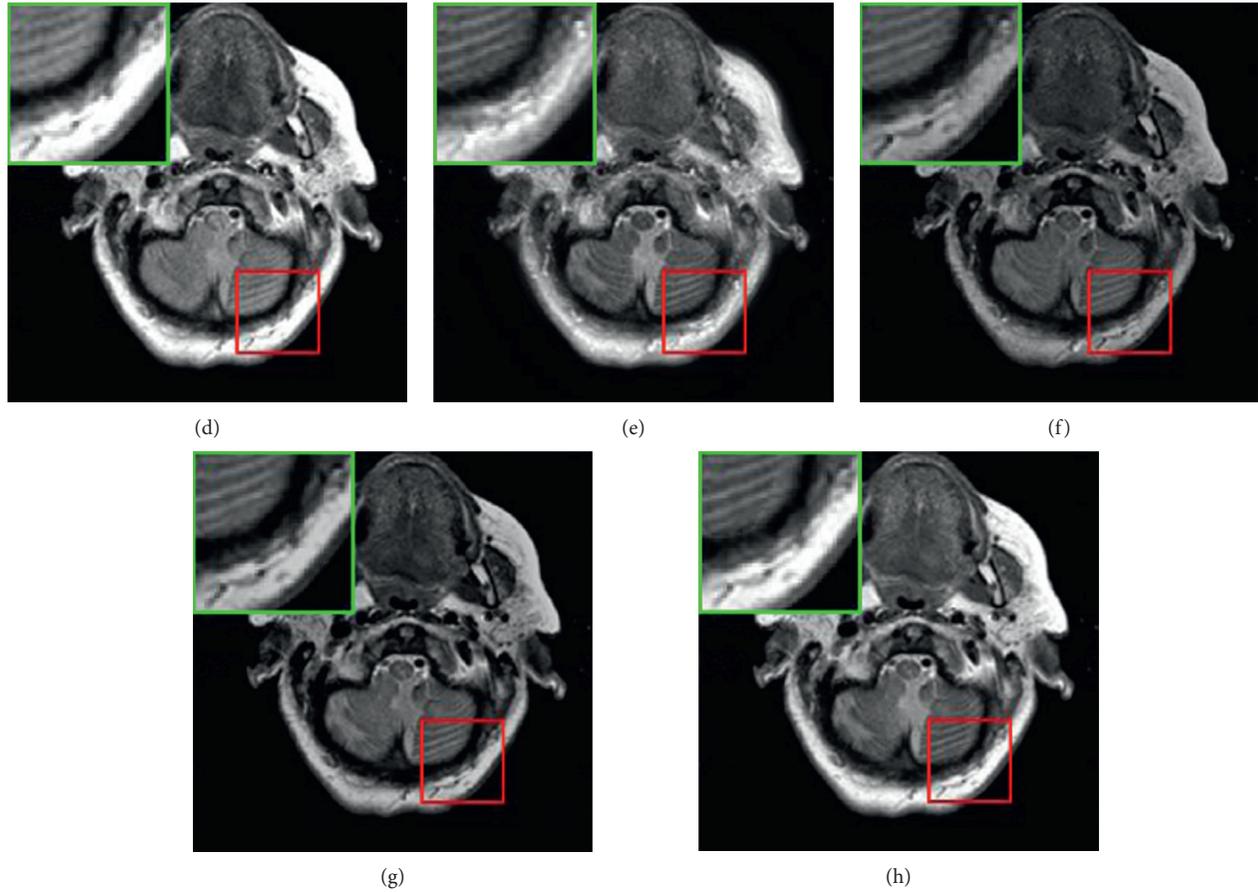


FIGURE 9: An example for the fusion results of the different methods. (c) SR. (d) BSR. (e) PAPCNN. (f) CSR. (g) CSMCA. (h) CTC SR.

TABLE 1: The average objective metrics of SR, BSR, PAPCNN, CSR, CSMCA, and CTC SR methods on pairs of medical images.

	SR	BSR	PAPCNN	CSR	CSMCA	CTCSR
EN	5.7277	5.7636	6.1723	6.1707	5.5939	5.9027
SD	65.3614	66.6342	73.3334	71.6315	62.1341	75.5059
MI	0.7729	0.7720	0.6135	0.6523	0.6439	0.7947
Q_G	0.5596	0.5767	0.5478	0.5928	0.5945	0.5987
Q_P	0.3732	0.40376	0.3849	0.4938	0.5301	0.5421

same visual effect. There are artificial defects in the result of CSR. The CSMCA method reduces the contrast of the fusion image. The details are seriously blurred in the SR method.

Figure 9 gives a set of H1 and H2 image fusion results. The BSR and CTC SR methods not only keep the brightness of the bone but also contain rich soft tissue information. They have good visual effect. There are artificial defects in the result of CSR. The CSMCA and CSR methods reduce the contrast of the fusion image. The result of the SR based-method lost a lot of details.

Table 1 lists the average objective metrics of different fusion methods on eight sets of CT and MR images. For each metric, the biggest value given in bold indicates the best results among all the methods. Overall, the proposed method shows the best performances on SD, MI, Q_G and Q_P . These metric values reflect the high robustness of the proposed method. It is further confirmed that the proposed method can achieve better fusion effect.

5. Conclusion

In this paper, a fusion method based on the cartoon-texture decomposition method and convolution sparse representation theory is proposed for medical images. The fusion rules of energy protection and feature extraction are used in the cartoon part, while the fusion method of convolution sparse representation is used in the texture part. Different fusion rules are selected in different feature parts, which can better represent image information and achieve better fusion effect. The experimental results show that the proposed algorithm is effective in terms of visual quality and objective metric values.

Data Availability

The experimental images used to support the findings of this study are composed of two parts. Part of the fusion images is

supplied by the database <http://www.imagefusion.org/>. Another part of the training dictionary filter images comes from the database <http://decsai.ugr.es/cvg/dbimagenes/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank first-class discipline construction in Ningxia Institutions of Higher Learning (Pedagogy) (Grant NXYLXK2017B11), the National Natural Science Foundation of China (Grants 61772389, 61972264, and 61971005), General Projects in Guyuan (Grant 2019GKGY041), and Key Research and Development Projects of Ningxia Autonomous Region (Talent Introduction Program) (2019BEB04021) for supporting our research work.

References

- [1] A. Mauro, "A generalised porous medium approach to study thermo-fluid dynamics in human eyes," *Medical & Biological Engineering & Computing Journal of the International Federation for Medical & Biological Engineering*, vol. 34, 2018.
- [2] A. Mauro, V. Romano, and P. Nithiarasu, "Suprachoroidal shunts for treatment of glaucoma," *International Journal of Numerical Methods for Heat & Fluid Flow*, vol. 28, no. 2, pp. 297–314, 2018.
- [3] Y. Liu, X. Chen, R. K. Wang, and X. Wang, "Deep learning for pixel-level image fusion: recent advances and future prospects," *Information Fusion*, vol. 42, pp. 158–173, 2018.
- [4] Z. Liu, K. Tsukada, K. Hanasaki, and Y. P. Ho, "Image fusion by using steerable pyramid," *Pattern Recognition Letters*, vol. 22, no. 9, pp. 929–939, 2001.
- [5] L. J. Chipman, T. M. Orr, and L. N. Graham, "Wavelets and image fusion," in *Proceedings of the International Conference on Image Processing*, New York, NY, USA, 1995.
- [6] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, and D. R. Bull, "Pixel- and region-based image fusion with complex wavelets," *Information Fusion*, vol. 8, no. 2, pp. 119–130, 2007.
- [7] Q. Canagarajah and B.-L. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Processing*, vol. 89, no. 7, pp. 1334–1346, 2009.
- [8] S. Li and B. Yang, "Hybrid multiresolution method for multisensor multimodal image fusion," *Sensors Journal, IEEE*, vol. 10, pp. 1519–1526, 2010.
- [9] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884–892, 2010.
- [10] L. Chang, X. Feng, X. Zhu et al., "CT and MRI image fusion based on multiscale decomposition method and hybrid approach," *IET Image Processing*, vol. 13, no. 1, pp. 83–88, 2018.
- [11] Y. Ming, L. Xiaoning, L. Yu et al., "Medical image fusion with parameter adaptive pulse coupled-neural network in non-subsampled shearlet transform domain," *IEEE Transactions on Instrumentation and Measurement*, vol. 34, pp. 1–16, 2018.
- [12] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.
- [13] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," *Proceedings of 20th International Conference on Information Fusion*, vol. 35, pp. 1–7, 2017.
- [14] Y. Liu, X. Chen, R. K. Ward et al., "Medical image fusion via convolutional sparsity based morphological component analysis," *IEEE Signal Processing Letters*, vol. 56, p. 1, 2019.
- [15] A. Azarang and H. G. Hasseemian, "A new pansharpening method using multiresolution analysis framework and deep neural networks," *Proceedings of 3rd International Conference on Pattern Recognition and Image Analysis*, vol. 56, pp. 1–5, 2017.
- [16] 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.
- [17] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion*, vol. 24, pp. 147–164, 2015.
- [18] 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 Recognition*, vol. 79, pp. 130–146, 2018.
- [19] M. Kim, D. K. Han, and H. Ko, "Joint patch clustering-based dictionary learning for multimodal image fusion," *Information Fusion*, vol. 27, pp. 198–214, 2016.
- [20] Q. Qiu, Z. Jiang, and R. Chellappa, "Joint patch clustering-based dictionary learning for multimodal image fusion," in *Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, 2011.
- [21] F. Siyahjani and G. Doretto, "Learning a context aware dictionary for sparse representation," *Lecture Notes in Computer Science Book Series*, pp. 228–241, Springer, Berlin/Heidelberg, Germany, 2013.
- [22] G. Qi, J. Wang, Q. Zhang, F. Zeng, and Z. Zhu, "An integrated dictionary-learning entropy-based medical image fusion framework," *Future Internet*, vol. 9, no. 4, p. 61, 2017.
- [23] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multimodality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.
- [24] A. Buades and J.-L. Lisani, "Directional filters for cartoon+texture image decomposition," *Image Processing on Line*, vol. 5, pp. 75–88, 2016.
- [25] R. Zhang, X. Feng, S. Wang, and L. Chang, "A sparse gradients field based image denoising algorithm via non-local means," *Acta Automatica Sinica*, vol. 14, no. 9, pp. 1542–1548, 2015.
- [26] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 301–315, 2016.
- [27] M. Morup and M. Schmidt, "Transformation invariant sparse coding," in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Process, MLSP*, New York, NY, USA, pp. 1–6, 2011.
- [28] 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.
- [29] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.
- [30] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. au, and W. Wu, "assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 94–109, 2012.