IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Multimodal Medical Image Fusion Techniques

Yogeshwaran R
Dept of IST
College of Engineering,Guindy
Chennai,India.

Geetha G
Dept of IST
College of Engineering,Guindy
Chennai,India.

***Abstract*— Multi-modal medical image fusion may be a well-known area of picture fusion research. Image fusion is the technique of creating one image from the pertinent data from several images taken of the same scene. The resulting merged image is more comprehensive and useful than any of the input images. Diagnostic accuracy depends on imaging technology. Medical image fusion research has gained popularity since the little information offered by single mode medical images cannot satisfy the demand for clinical diagnosis, which necessitates a substantial amount of information. Single-mode fusion and multimodal fusion are subcategories of medical picture fusion. The advantages and disadvantages of each medical technique, such as X-rays, CT scans, MRIs, nuclear medicine, and others, used to check the body's organs, vary. Then, a Non Subsampled Contourlet Transform (NSCT)-based fusion algorithm is used to fuse the Magnetic Resonance Imaging (MRI) and computed tomography (CT) scan images. Additionally, In order to improve the viability of denoising algorithms, In this research, a novel one-stage blind real picture method is proposed, using a modular architecture to denoising network RIDNet. A residual on the residual structure is used by us to Facilitate the flow of low-frequency data and use feature attention to take advantage of channel dependencies In order to increase the robustness, Denoising Convolutional Neural Network (DnCNN) is utilized.**

Keywords — Medical images, Multimodal fusion, Security, image denoising, nonsubsampled contourlet transform, RID Net, High and Low frequency ,Deep convolutional neural network.

1. INTRODUCTION

A unique process dubbed "image fusion" joins two or more existing images to create a single new image. A prominent field of picture fusion study may be multi-modal medical image fusion. Image fusion is a method of combining relevant data from several combining multiple images of the same scene into one. More complete and helpful than any of the input images is the final fused image. Object identification and recognition, medical diagnosis, satellite imaging for remote sensing, civilian and military surveillance, and navigation guidance are some of the uses for picture fusion. Imaging technology is essential for precise medical diagnosis. Additionally, the storage, processing, and ongoing availability of data from numerous sources makes cloud computing in healthcare a vital solution. The illicit copying, alteration, and forging of medical records, however, is a serious issue that has to be addressed [2]. Additionally, the need for copyright protection and identity theft prevention is growing every day [3, 4]. Data integrity must be protected from unauthorized users.

A low-level vision task called image denoising is crucial in a number of ways. First off, some noise corruption during image acquisition is unavoidable and can significantly lower the visual quality; as a result, removing for many computer vision and image analysis applications, removing noise from the collected image is a crucial step. Second, denoising is a special testing environment for comparing picture prior and optimization techniques from a Bayesian perspective.
from a particular angle. Additionally, a lot of activities involving picture restoration can be accomplished in the unrolled inference through strategies for separating variables into smaller denoising processes, It broadens the use of image denoising even more.

NSCT is especially advantageous for image fusion since it not only has outstanding local properties in the spatial and frequency domains but also enjoys multi-scale and translation invariance. Non-subsampled directional channel bank (NSDFB) and non-subsampled pyramid structure (NSP) are taken into account in its structure.
While NSDFB provides the multi-direction highlight for NSCT, NSP provides the multi-scale highlight. The arrangement of the image as well as the smoothed shape of the source image are described by the low-frequency sub-band. Even though the decomposition levels are periodically limited in this approach, the majority of the sign liveliness and a few nuanced elements of the initial picture are available in the low-recurrence sub-band. The specific details and rough estimates offered in the initial.

Multimodality medical image fusion based on machine learning methods cannot produce high-quality images. Deep learning fusion techniques are necessary to achieve this goal. These deep learning convolution models allow for improved characterization of medical pictures, improved handling of curved geometries, and improved production of fused details. One of the general advantages of deep learning systems is the improvement of the visual quality of the photographs and the reduction of noise and imperfections. 256*256 is the size of each image

2.RELATED WORKS

In this section, we present and discuss recent trends in the image denoising, image preprocessing and image fusing methods. In 2016 Du et al. introduced union Laplacian pyramid to complete the fusion of medical images. Some improved versions of DWT such as dual tree complex wavelet transform (DT-CWT) (Yu et al., 2016), non-subsampled rotated complex wavelet transform (NSRCxWT) (Chavan et al., 2017), discrete stationary wavelet transform (DSWT) (Ganasala and Prasad, 2020a; Chao et al., 2022) were presented to complete the fusion of medical images. Compared with DWT, these three new versions share both the redundancy feature and the shift-invariance property, which effectively avoid the Gibbs phenomenon in DWT. Similarly, in order to overcome the absence of shift-invariance in the original contourlet transform and shearlet transform, the corresponding improved versions namely non-subsampled contourlet transform (NSCT) and non-subsampled shearlet transform (NSST) were proposed successively. In comparison to the aforementioned transform domain-based methods, NSCT and NSST have both manifested competitive fusion performance due to their flexible structures. Zhu et al. (2019) combined NSCT, phase congruency and local Laplacian energy together to present a novel fusion method for multi-modality medical images. Liu X. et al. (2017), Liu et al. (2018) proposed two NSST-based methods to fuse the CT and MRI images.

In addition to spatial domain-based methods and transform domain-based methods, extensive work has also been conducted with soft computing-based methods dedicated to multimodal medical image fusion.

A great many representative models, including dictionary learning model (Zhu et al., 2016; Li et al., 2018), gray wolf optimization (Daniel, 2018), fuzzy theory (Yang et al., 2019), pulse coupled neural network (Liu X. et al., 2016; Xu et al., 2016), sparse representation (Liu and Wang, 2015; Liu Y. et al., 2016), total variation (Zhao and Lu, 2017), guided filter (Li et al., 2019; Zhang et al., 2021), genetic algorithm (Kavitha and Thyagharajan, 2017; Arif and Wang, 2020), compressed sensing (Ding et al., 2019), structure tensor (Du et al., 2020c), local extrema (Du et al., 2020b), Otsu's method (Du et al., 2020a) and so on, were successfully used to fuse the medical images.

Since the transform domain-based methods and soft computing-based methods have both manifested to be promising in the field of medical image fusion, some novel hybrid methods have also been presented in recent years. Jiang et al. (2018) combined interval type-2 fuzzy sets with NSST to complete the fusion task of multi-sensor images. Gao et al. (2021) proposed a fusion method based on particle swarm optimization optimized fuzzy logic in NSST domain. Asha et al. (2019) constructed a novel fusion scheme based on NSST and gray wolf optimization. Singh and Anand (2020) employed NSST to decompose the source images, and then used sparse representation and dictionary learning model to fuse the sub-images. Yin et al. (2019) and Zhang et al. (2020) each proposed a NSST-PCNN based fusion method for medical images. The guided filter was combined with NSST to deal with the issue of multi sensor image fusion (Ganasala and Prasad, 2020b). Zhu et al. (2022) combined the advantages of both spatial domain and transform domain methods to construct an efficient hybrid image fusion method. Besides, the collective view of the applicability and progress of information fusion techniques in medical imaging were reviewed respectively.

Recently, Schmidt et al. [53] introduced a cascade of shrinkage fields (CSF) which integrated half-quadratic optimization and random-fields. Shrinkage aims to suppress smaller values (noise values) and learn mappings discriminatively. The CSF assumes the data fidelity term to be quadratic and that it has a discrete Fourier transform based closed-form solution.

Currently, due to the popularity of convolutional neural networks (CNNs), image denoising algorithms [63, 64, 39,14, 53, 8] have achieved a performance boost. Notable denoising neural networks, DnCNN [63], and IrCNN [64] predict the residue present in the image instead of the denoised image as the input to the loss function is ground truth noise as compared to the original clean image. Both networks achieved better results despite having a simple architecture where repeated blocks of convolutional, batch normalization and Re LU activations are used. Furthermore, IrCNN and DnCNN [63] are dependent on blindly predicted noise i.e. without taking into account the underlying structures and textures of the noisy image.

The algorithms [64, 63,35] benefitted from the modeling capacity of CNNs and have shown the ability to learn a single-blind denoising model; however, the denoising performance is limited, and the results are not satisfactory on real photographs. Generally speaking, real-noisy image denoising is a two-step process: the first involves noise estimation while the second addresses non-blind denoising. Noise clinic (NC) [38] estimates the noise model dependent on signal and frequency followed by denoising the image using non-local Bayes (NLB). In comparison, Zhang et al. [65] proposed a non-blind Gaussian denoising network, termed FFDNet that can produce satisfying results on some of the real noisy images; however, it requires manual intervention to select high noise-level.

Although the results of TRND are favorable, the model requires a significant amount of data to learn the parameters and influence functions as well as overall fine-tuning,hyper-parameter determination, and stage-wise training.Similarly, non-local color net (NLNet) [39] was motivated by non-local self-similar (NSS) priors which employ non-local self-similarity coupled with discriminative learning.NLNet improved upon the traditional methods; but, it lags in performance compared to most of the CNNs [64, 63] due to the adaptaton of NSS priors, as it is unable to find the analogs for all the patches in the image.

3.PROPOSED WORK

3.1 Medical Image Fusion

Utilizing images from many sources, Multimodal Medical Image Fusion (MMIF) X-rays, computed tomography (CT), single photon emission computed tomography (SPECT), ultrasound, magnetic resonance imaging (MRI), infrared and ultraviolet, positron emission tomography, and others like as emission tomography (PET). The location, size, and appearance of the lesion as well as the morphological and structural changes it has caused in nearby tissues can all be seen in images from MRI, X-ray, CT, and US. In the medical field, each imaging modality provides specific information and characteristics. The numerous medical imaging methods utilised for disease detection and diagnosis in the human body span the whole electromagnetic (EM) spectrum. Each imaging technique has unique characteristics and operates at

a certain wavelength and frequency. An object scatters, reflects, or absorbs EM waves. Multi-modality medical picture fusion commonly uses transform domain based techniques. In contrast to spatial domain approaches, transform domain approaches first transform source images into specific coefficients. After combining the coefficients, and each combined coefficient is then reverse-transformed to produce a combined image. Recent years have seen a significant increase in the use of multi-scale transform (MST) and sparse representation-based picture fusion techniques.

In modern health care systems, where clinical imaging is essential, computed tomography (CT) and magnetic resonance imaging (MRI) are two of the most frequently used imaging modalities. Radiologists are now able to study the human body and generate a number of patterns that can be used for clinical analysis. These images reveal anatomical statistics. However, there is still a significant issue with the capacity to extract relevant functional information from a single image. Multimodal image fusion is necessary in this case because it enhances anatomical and functional information by fusing the complementary information of images from many modalities using simulation. A CT scan can show the spatial distribution of a certain physical quantity, while an MRI, or magnetic resonance imaging, may create a four-dimensional

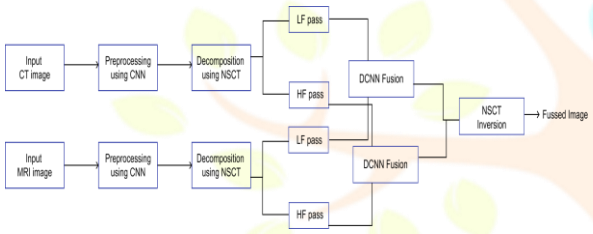


Fig 3.1.1 flow diagram

3.2 Data preprocessing

CNN has become a powerful resource for understanding medical images. Researchers have successfully applied CNNs to a range of medical image interpretation applications. Image denoising is a crucial low-level vision activity that has various uses. First off, during the capture of photos, noise contamination is inevitable and can seriously impair visual quality. The elimination of noise from the captured image is a vital step for many computer vision and image analysis applications. CNN models performed better when denoising because to their modelling ability, network design, and training.

3.3 CNN Denoiser

The input picture was first processed by 24 sequential connected convolutional units (SCCU). SCCU is made up of complex-valued (CV) BN, complex-valued (CV) ReLU, and complex-valued (CV) convolutional layer. A 64-convolutional kernel was used in the network. The remaining block was distributed to the middle 18 units. A convolution/deconvolution layer with a stride of 2 was built to improve computational efficiency**.**

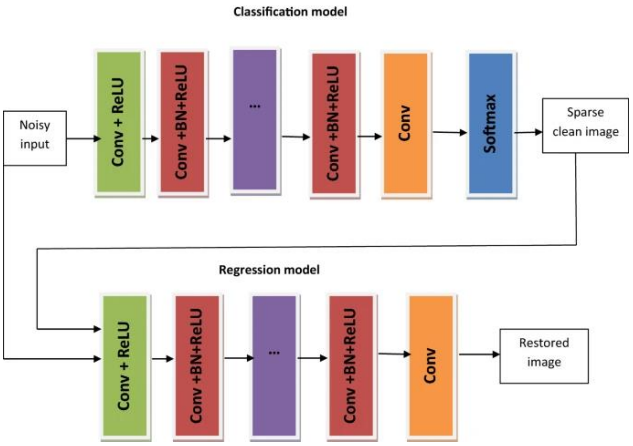


Fig 3.2.1 flow diagram

3.4 Convolutional layer

The cornerstone of a convolutional neural network is the CONV layer. The K learnable filters (also known as "kernels") that make up the CONV layer parameters are all almost square and have widths and heights. The depth is the number of CNN input channels in the image. The depth for volumes further down the network will depend on how many filters were employed in the layer before. The width and height of the input volume are used to convolve each of the K filters. By simply visualising each of our K kernels sliding across the input region, computing an element-wise multiplication, adding the results, and then doing so, the output value may be stored in a two-dimensional activation map.

3.4 Batch Normalization

Instead of using the raw data itself, the batch normalisation technique is used to normalise the layers of a neural network. Mini-batches are employed in place of the whole data set. Learning is facilitated by accelerating training and using greater learning rates. are used, as their name suggests, to normalise the activations of a particular input volume before transferring it to the network layer below. It has been demonstrated that utilising batch normalisation significantly reduces the number of epochs required to train a neural network. Batch normalisation also has the advantage of helping to "stabilise" training, allowing for a larger variety of learning rates and regularisation strengths. Your life will be easier by using batch normalisation to reduce the volatility of learning rate and regularisation.

3.5 Decomposition using NSCT

NSCT is especially advantageous for image fusion since it not only has outstanding local properties in the spatial and frequency domains but also enjoys multi-scale and translation invariance. Non-subsampled directional channel bank (NSDFB) and non-subsampled pyramid structure (NSP) are taken into account in its structure.

While NSDFB provides the multi-direction highlight for NSCT, NSP provides the multi-scale highlight. The arrangement of the image as well as the smoothed shape of the source image are described by the low-frequency sub-band. Even though the decomposition levels are periodically limited in this approach, the majority of the sign liveliness and a few nuanced elements of the initial picture are available

in the low-recurrence sub-band. The specific details and rough estimates offered in the initial.

The below equation explains it:

$$a_L^A(m,n) = \sum_{i=-1}^{i=1}\sum_{j=-1}^{j=1} {C_L^A}^2(m+i,n+j)\log\left({C_L^A}^2(m+i,n+j)\right)/9$$

Where $C^A{}_L(m,n)$ represents the source image A's location (m,n) in the low-frequency sub-band coefficient. When $C^B{}_L(m,n)$ at position (m,n) is supplied, similar to source image B, it represents the activity of the low-frequency sub-band coefficient.

$$a_L^B(m,n) = \sum_{i=-1}^{i=1}\sum_{j=-1}^{j=1} {C_L^B}^2(m+i,n+j)\log\left({C_L^B}^2(m+i,n+j)\right)/9$$

In order to obtain the decision map of fusion, use the maximum combination technique, which involves choosing the maximum coefficient using the following activity measure:

$$d_i(m,n) = \begin{cases} 1 & if\, a_L^A(m,n) \geq a_L^B(m,n) \\ 0 & if\, a_L^A(m,n) < a_L^B(m,n) \end{cases}$$

Then, using the decision map (df), the low-frequency sub-band coefficients are described by ($C^L{}_F(m,n)$) and calculated.

$$C_L^F(m,n) = \begin{cases} C_L^A(m,n) & if\, d_f(m,n) = 1 \\ C_L^B(m,n) & if\, d_f(m,n) = 0 \end{cases}$$

Edges, shapes, and question boundaries are examples of high-frequency sub-bands that convey detail information and maximum coefficient retrieval. The combination technique for high-recurrence sub-bands is most frequently used to find the most extreme coefficient with the absolute highest value.

$$a_{d,k}^A(m,n) = \mathrm{WSML}\left[C_{d,k}^A(m,n)\right]$$

The dth level and Kth directional sub-band coefficient of image A at position (m,n) are represented as $C^A{}_{d,k}(m,n)$.

$$a_{d,k}^B(m,n) = \mathrm{WSML}\left[C_{d,k}^B(m,n)\right]$$

Where $C^B{}_{d,k}$ (m,n) is used to describe the picture B's dth level and Kth directional sub-band coefficient at the area (m,n). When choosing the most extreme action coefficient plan, the underlying combination choice guide is found.

$$d_i(m,n) = \begin{cases} 1 & \text{if } a_{d,k}^A(m,n) \geq a_{d,k}^B(m,n) \\ 0 & \text{if } a_{d,k}^A(m,n) < a_{d,k}^B(m,n) \end{cases}$$

$$C_{d,k}^F(m,n) = \begin{cases} C_{d,k}^A(m,n) & \text{if } d_f(m,n) = 1 \\ C_{d,k}^B(m,n) & \text{if } d_f(m,n) = 0 \end{cases}$$

As mentioned in the low-frequency fusion rule, the final fusion decision map (df) is currently being discovered through consistency verification. Here, final fusion decision map is used to calculate the fused of high-frequency sub-band coefficients.

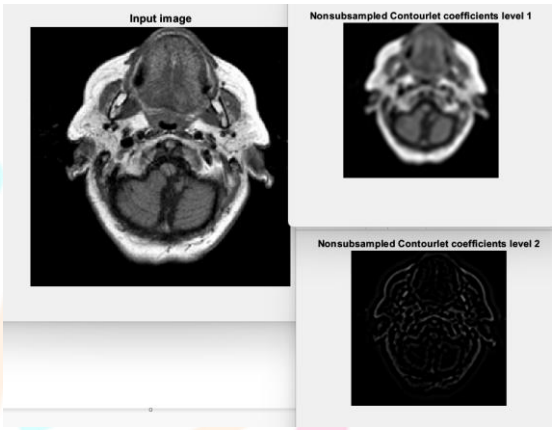


Fig 3.5.1 CT Image

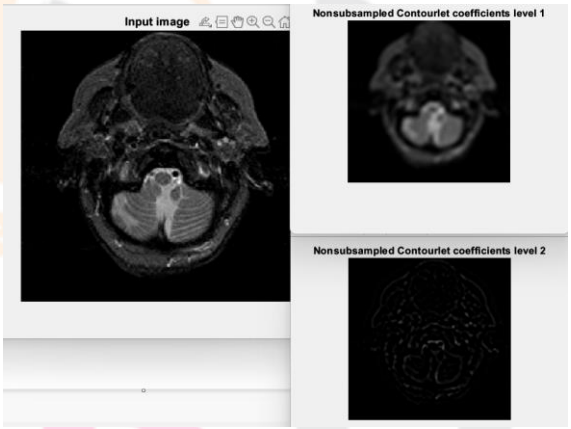


Fig 3.5.2 MRI Image

3.6 DCNN Fusion

Multimodality medical image fusion based on machine learning methods cannot produce high-quality images. Deep learning fusion approaches are needed to achieve this goal. These deep learning convolution models allow for improved characterization of medical pictures, improved handling of curved geometries, and improved production of fused details. One of the general advantages of deep learning methods is the improvement of the visual quality of the photographs and the reduction of noise and artefacts. 256*256 is the size of each image.

4.DATASET USED

4.1 Sample Dataset 1:

CT images were used as the sample dataset 1 in this suggested system. The brain's CT scan clearly reveals the skull bone and other hard tissues, but the soft tissues, such as the membranes that encircle the brain, are less obvious.

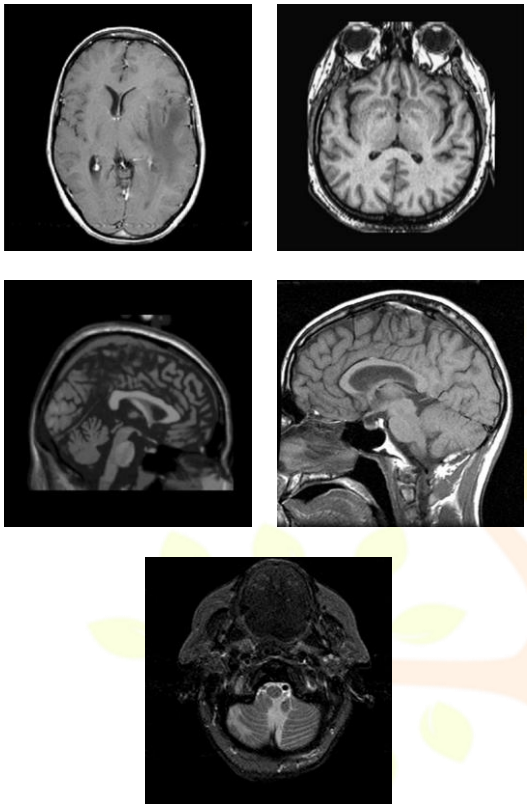
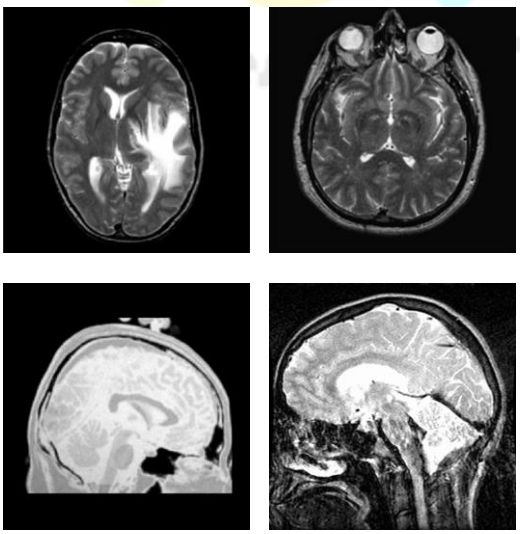


Fig 4.1.1 CT Images

4.2 Sample Dataset 2:

This dataset includes a set of similar MRI pictures in which the soft tissue, such as the membranes encasing the brain, can be seen clearly but the hard tissue, such as the skull's bones, cannot.



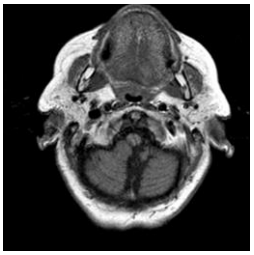


Fig 4.1.2 MRI Images

5.REFERENCES

1. Z. Zhang and R. S. Blum, "A categorization of multiscale decomposition- based image fusion schemes with a performance study for a digital camera application" Proceedings of the IEEE,vol. 87, no. 8, pp. 1315-1326, 1999.
2. W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," Pattern Recognit. Lett., vol. 28, no. 4, pp. 493–500,(2007)
3. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian Denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process. (2017)
4. J. Du, W. Li, K. Lu, B. Xiao, An overview of multi-modal medical image fusion, Neurocomputing (2016)
5. J.M. Dolly, A.K. Nisa, A survey on different multimodal medical image fusion techniques and methods, in: Proc. 1st Int. Conf. Innov. Inf. Commun. Technol, ICIICT 2019, (2019)
6. Goyal, S., Singh, V., Rani, A., and Yadav, N. (2022). Multimodal image fusion and denoising in NSCT domain using CNN and FOTGV.
7. He, C. T., Liu, Q. X., Li, H. L., and Wang, H. X. (2010). Multimodal medical image fusion based on IHS and PCA.
8. Hermessi, H., Mourali, O., and Zagrouba, E. (2021). Multimodal medical image fusion review: theoretical background and recent advances.

9. Hossny, M., Nahavandi, S., and Creighton, D. (2008). Comments on 'information measure for performance of image fusion.

10. Jiang, Q., Jin, X., Hou, J. Y., Lee, S., and Yao, S. W. (2018). Multi-sensor image fusion based on interval type-2 fuzzy sets and regional features in nonsubsampled shearlet transform domain.

11. Kavitha, S., and Thyagharajan, K. K. (2017). Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation.

12. Kumar, B. K. S. (2015). Image fusion based on pixel significance using cross bilateral filter.

13. Li, W. S., Jia, L. H., and Du, J. (2019). Multi-modal sensor medical image fusion based on multiple salient features with guided image filter.

14. Liu, X. B., Mei, W. B., and Du, H. Q. (2016). Multimodality medical image fusion algorithm based on gradient minimization smoothing filter and pulse coupled neural network.

15. Liu, X. B., Mei, W. B., and Du, H. Q. (2017). Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion.

16. Liu, Y., Chen, X., Ward, R., and Wang, Z. J. (2016). Image fusion with convolutional sparse representation.

17. Liu, Y., and Wang, Z. F. (2015). Simultaneous image fusion and denosing with adaptive sparse representation.

18. Shi, Z. H., Zhang, C. W., Ye, D., Qin, P. L., Zhou, R., Lei, L., et al. (2022). MMI-Fuse: multimodal brain image fusion with multiattention module.

19. Xu, H., and Ma, J. Y. (2021). EMFusion: an unsupervised enhanced medical image fusion network.

20. Yang, Y., Wu, J. H., Huang, S. Y., Fang, Y. M., Lin, P., Que, Y., et al. (2019). Multimodal medical image fusion based on fuzzy discrimination with structural patch decomposition.