Multimodal medical image fusion algorithm in the era of big data

Wei Tan[1] · Prayag Tiwari[2] · Hari Mohan Pandey[3] · Catarina Moreira[4] · Amit Kumar Jaiswal[5]

Abstract

In image-based medical decision-making, different modalities of medical images of a given organ of a patient are captured. Each of these images will represent a modality that will render the examined organ differently, leading to different observations of a given phenomenon (such as stroke). The accurate analysis of each of these modalities promotes the detection of more appropriate medical decisions. Multimodal medical imaging is a research field that consists in the development of robust algorithms that can enable the fusion of image information acquired by different sets of modalities. In this paper, a novel multimodal medical image fusion algorithm is proposed for a wide range of medical diagnostic problems. It is based on the application of a boundary measured pulse-coupled neural network fusion strategy and an energy attribute fusion strategy in a non-subsampled shearlet transform domain. Our algorithm was validated in dataset with modalities of several diseases, namely glioma, Alzheimer's, and metastatic bronchogenic carcinoma, which contain more than 100 image pairs. Qualitative and quantitative evaluation verifies that the proposed algorithm outperforms most of the current algorithms, providing important ideas for medical diagnosis.

✉ Hari Mohan Pandey
  pandeyh@edgehill.ac.uk

  Wei Tan
  wtan@stu.xidian.edu.cn

  Prayag Tiwari
  prayag.tiwari@dei.unipd.it

  Catarina Moreira
  catarina.pintomoreira@qut.edu.au

  Amit Kumar Jaiswal
  amitkumar.jaiswal@beds.ac.uk

1   School of Physics and Optoelectronic Engineering, Xidian University, Xi'an, China

2   Department of Information Engineering, University of Padova, Padua, Italy

3   Department of Computer Science, Edge Hill University, Ormskirk L39 4QP, UK

4   School of Information Systems, Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia

5   Institute for Research in Applicable Computing, University of Bedfordshire, Luton, UK

1 Introduction

Multimodal medical imaging is a research field that has been getting increasing attention in the scientific community in the last few years, specially due to its significance in medical diagnosis, computer vision, and internet of things [3, 5, 15, 20, 28, 31, 32, 35]. Defined as the simultaneous production of signals belonging to different medical imaging techniques, one of the biggest challenges in this research field is how to combine (or fuse) in an effective and optimal way multimodal medical imaging sensors, such as positron emission tomography (PET), single-photon emission computed tomography (SPECT), and magnetic resonance imaging (MRI). This image fusion process comprises many techniques and research areas, ranging from image processing techniques, computer vision to pattern recognition, with the goal of promoting more accurate medical diagnosis and more effective medical decision-making [8, 10, 18, 26, 45].

1.1 Current challenges in multimodal image fusion

Image fusion can usually be divided into three levels: pixel-level, feature-level, and decision level [21, 31, 42–44, 47]. Since the aim is to fuse pixel information from source images, medical image fusion belongs to the pixel-level.

Multi-scale transform (MST) method is one of the most famous categories [40]. Commonly, the MST fusion methods consist of three steps. First, the source images are transformed into MST domain. Then, the parameters in different scales merged in light of a specific fusion strategy. Finally, the fused image is reconstructed through the corresponding inverse transform. The MST methods mainly contain the Laplacian pyramid (LP) [6], the wavelet transform (WT) [27, 34], the non-subsampled contourlet transform (NSCT) [49], and the non-subsampled shearlet transform (NSST) [4, 23, 38]. However, if the MST method performs without other fusion measures, some unexpected block effect may appear [39].

To overcome this disadvantage, some fusion measures are applied in the MST method. For instance, spatial frequency (SF), local variance (LV), the energy of image gradient (EIG) and sum-modified-Laplacian (SML) are commonly used as fusion measures [17, 41]. However, most of these measures are acquired in the spatial domain or low-order gradient domain, which means the fusion map may not be always precise. This imprecision may lead to blocking artefacts.

Except for traditional MST methods, the edge-preserving filtering (EPF)-based MST decomposition method are also commonly used. In the EPF-MST methods, Gaussian filtering and EPF are used to decompose the input image into two scale-layers and one base layer. Then, three layers are fused based on suitable fusion strategies. Finally, the fused image is reproduced by a reconstruction algorithm. The EPF-MST methods contain bilateral filtering (BF)-based [51], curvature filtering (CF)-based [40], and co-occurrence filtering (CoF)-based [37] methods.

1.2 A pulse-coupled neural network model for medical image fusion

To overcome this challenge, a method called pulse-coupled neural network (PCNN) has been proposed in the literature [46]. This method was initially proposed to emulate the underlying mechanisms of a cat's visual cortex and became later an essential method in image processing [29]. Kong et al. presented an SF modulated PCNN fusion strategy in NSST domain with the solution of infrared and visible image fusion [19]. Inspired by this kind of fusion measure

modulated by the PCNN model, one interesting research path would be a solution to a new measure to modulate PCNN in the medical image fusion field.

To further improve the fusion quality of medical images, we propose a medical image fusion method based on boundary measure modulated by a pulse-coupled neural network in the non-subsampled shearlet domain. Firstly, the source images are transformed into the NSST domain with low-frequency bands and high-frequency bands. Then, the low-frequency bands are merged through an energy attribute-based fusion strategy, and the high-frequency bands are merged through a boundary measure modulated PCNN strategy. Finally, the fused image is reconstructed by combining the inverse NSST. We evaluate the proposed algorithm by comparing its performance with several existing methods using both a quantitative and qualitative evaluation. Experimental results demonstrate that the proposed method performs better than most of the existing fusion methods.

1.3 Contribution

The main contributions of the proposed research article are the following:

1. A medical image fusion framework based on boundary measured PCNN in NSST domain, which can complete the fusion task effectively;
2. The application of a boundary measured PCNN model for high-frequency bands. In this method, the gradient information of the image can be easily extracted, and the size of the structure can be changed to adapt to the scale of structure;
3. The application of an energy attribute-based fusion strategy to low-frequency bands.

Experiments conducted in this research paper suggest that the proposed boundary measured PCNN-NSST achieves the best performance in most cases in qualitative and quantitative when compared to other state-of-the-art image fusion techniques.

1.4 Organization

The rest of this paper is organized as follows. In Sect. 2, it is presented the most significant works in the image fusion domain. In Sect. 3, the proposed fusion method BM-PCNN-NSST is described. In Sect. 4, it is presented the set of experiments that were performed to evaluate the proposed algorithm. Finally, in Sect. 5, the main conclusions of this research work are presented.

2 Related work

In this section, we present an overview of the most significant image fusion algorithms in the literature, namely the non-subsampled shearlet transform (Sect. 2.1), the multi-scale morphological gradient (Sect. 2.2), and the pulse-coupled neural network (Sect. 2.3).

2.1 Non-subsampled shearlet transform

The non-sampled shearlet transform is an image fusion method, originally proposed by Easley [13]. It consists in combining the non-subsampled pyramid transform with different shearing filters, and it has the characteristics of multi-scale and multi-directionality. The non-subsampled pyramid transform makes it invariant, which is superior than the LP, and WT methods. Additionally, since the size of the shearing filter is smaller than the directional filter, NSST can represent smaller scales, which makes it better than NSCT.

Given the superiority of its underlying functions, NSST performs better than most commonly used MST. It is therefore widely used in the field of image denoising [36] and image fusion [22].

The NSST model can be described as follows. For the case, $n = 2$, the shearlet function is satisfied

$$\Omega_{AB}(\psi) = \left\{ \psi_{i,j,k}(x) = |\det A|^{j/2} \psi\left(B^l A^j x - k\right); j, l \in Z^2 \right\} \tag{1}$$

where $\psi \in L^2(R^2)$, both A and B are invertible matrices with size 2×2, and $|\det B| = 1$. For instance, A and B can be represented as

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{2}$$

In this situation, the tiling of the frequency plane of NSST is shown in Fig. 1 It can be seen that (a) represents the decomposition, and (b) represents the size of the frequency support of the shearlet element $\psi_{i,l,k}$.

For convenience, two related functions are used to represent the NSST and the inverse NSST

$$\{L_n, H_n\} = \text{nsst\_de}(I_{\text{in}}) \tag{3}$$

$$I_{\text{re}} = \text{nsst\_re}(L_n, H_n) \tag{4}$$

where $\text{nsst\_de}(\cdot)$ represents the NSST decomposition function for the input image I_{in}, and $\text{nsst\_re}(\cdot)$ represents the NSST reconstruction steps for the reconstructed image I_{re}. The parameters L_n and H_n represent low-frequency sub-bands and high-frequency sub-bands, respectively.

2.2 Multi-scale morphological gradient

Multi-scale morphological gradient (MSMG) is an effective operator which extracts gradient information from an image in order to indicate the contrast intensity in the close neighborhood of a pixel in the image. For this reason, MSMG is a method that is highly efficient and used in edge detection and image segmentation. In image fusion, MSMG has been used as a type of focus measure in multi-focus image fusion [50]. The specific details of MSMG are as follows.

A multi-scale structuring element is defined as

$$\text{SE}_j = \underbrace{\text{SE}_1 \oplus \text{SE}_1 \oplus \cdots \oplus \text{SE}_1}_{j}, \quad j \in \{1, 2, \ldots, N\} \tag{5}$$

where SE_1 denotes a basic structure element, and t represents the number of scales.

The gradient feature G_t can be represented by the morphological gradient operators from the image f.

$$G_t(x, y) = f(x, y) \oplus \text{SE}_t - f(x, y) \odot \text{SE}_t \tag{6}$$

where $\oplus$ and $\odot$ denote the morphological dilation and erosion operators, respectively. (x, y) denotes the pixel coordinate.

From the multi-scale structuring element and the gradient feature, then one can obtain the MSMG by computing the weighted sum of gradients over all scales.

$$M(x, y) = \sum_{t=1}^{N} w_t \cdot G_t(x, y) \tag{7}$$

where w_t represents the weight of gradient in t-th scale, and it can be represented as

$$w_t = \frac{1}{2t + 1} \tag{8}$$

Figure 2 shows an example of MSMG. One can see that the boundary information of the images has been well extracted, which demonstrates the effectiveness of the boundary measure.
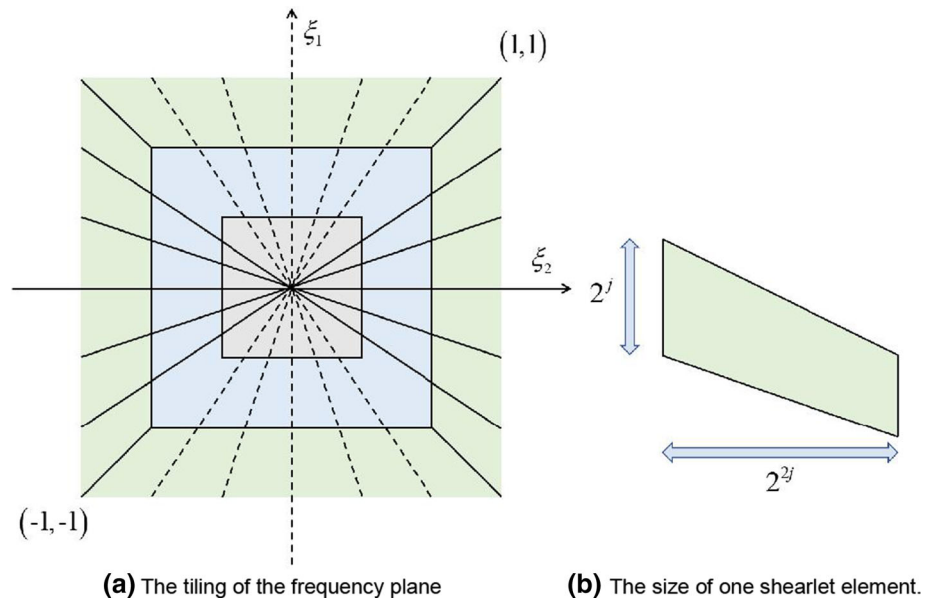
2.3 Pulse-coupled neural network

As the third-generation artificial neural network, PCNN has achieved great success in the image fusion field. A PCNN model often contains three parts: the receptive field, the modulation field and the pulse generator. The expressions of a simplified dual-channel PCNN model can be defined as
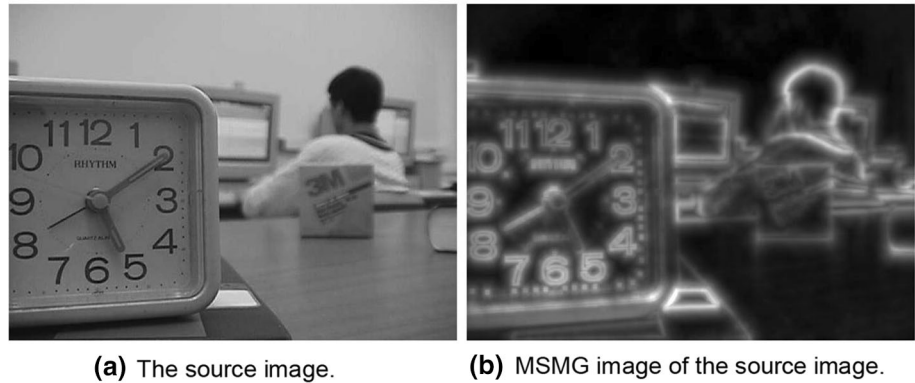
$$F_{ij}^1(k) = S_{ij}^1(k) \tag{9}$$

$$F_{ij}^2(k) = S_{ij}^2(k) \tag{10}$$

Fig. 1 The structure of the frequency tiling

(a) The tiling of the frequency plane   **(b)** The size of one shearlet element.

Fig. 2 An example of MSMG

(a) The source image.   **(b)** MSMG image of the source image.

$$L_{ij}(k) = \begin{cases} 1, & \text{if } \sum_{r,t \in S} Y_{rt}(k-1) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$U_{ij}(k) = \max\left\{ F^1_{ij}(k)\left(1 + \beta^1_{ij}L_{ij}(k)\right), F^2_{ij}(k)\left(1 + \beta^2_{ij}L_{ij}(k)\right) \right\} \quad (12)$$

$$Y_{ij}(k) = \begin{cases} 1, & \text{if } U_{ij}(k) \geq \theta_{ij}(k-1) \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

$$\theta_{ij}(k) = \theta_{ij}(k-1) - \Delta + V_\theta Y_{ij}(k) \quad (14)$$

$$T_{ij} = \begin{cases} k, & \text{if } U_{ij}(k) \geq \theta_{ij}(k-1) \\ T_{ij}(k-1), & \text{otherwise} \end{cases} \quad (15)$$

As is shown in Fig. 3, S^1_{ij} and S^2_{ij} denote the pixel value of two input images at point (i,j) in the neural network; L_{ij} represents the linking parameter; β^1_{ij} and β^2_{ij} denote the linking strength; F^1_{ij} and F^2_{ij} represent the feedback of inputs. U_{ij} is the output of the dual-channel. θ_{ij} is the threshold of step function, d_e is the declining extent of the threshold, V_θ decides the threshold of the active neurons,

and T_{ij} is the parameter to determine the number of iterations. $Y_{ij}(k)$ is the k-th output of PCNN.

3 A Bounded measured PCNN in NSST domain algorithm

In this section, we present the proposed algorithm for multimodal medical image fusion: a bounded measured PCNN approach in the NSST domain (BM-PCNN-NSST). The framework of the proposed algorithm is illustrated in Fig. 4. The fusion algorithm consists in four parts: the NSST decomposition, the low-frequency fusion, the high-frequency fusion, and the NSST reconstruction.

The algorithm starts with a pseudocolor image source A which contains three-bands (PET/SPECT image). The first step is to apply an intensity-hue-saturation (IHS) transform in A, which will result in a pair containing the intensity image I_A and a source image B. After performing the fusion

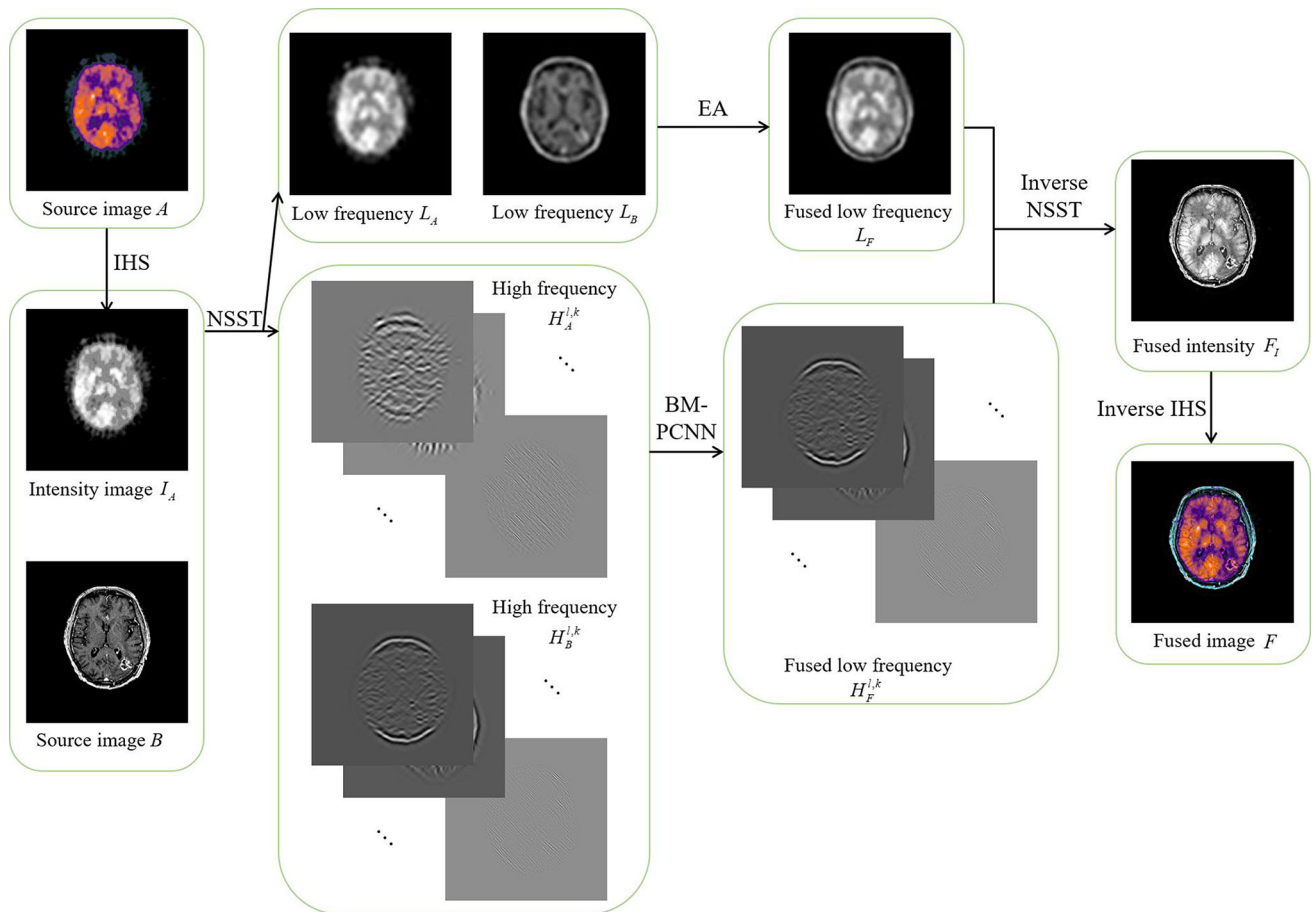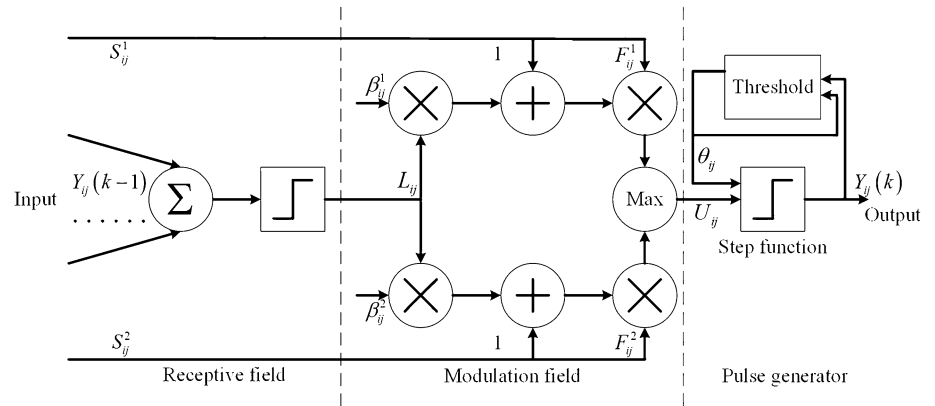**Fig. 3** Classical PCNN model

Fig. 4 Framework of the proposed algorithm

of this image pair, an inverse IHS transform is applied in order to obtain the final fused image.

3.1 NSST decomposition

An N-level NSST decomposition is performed on images I_A and B to acquire the decomposition bands L_A, $H_A^{l,k}$ and L_B, $H_B^{l,k}$ based on Eq (3), where L denotes low-frequency

sub-bands and $H^{l,k}$ represents high-frequency sub-bands at level l with direction k.

$$\left\{ L_A, H_A^{l,k} \right\} = \text{nsst\_de}(A) \tag{16}$$

$$\left\{ L_B, H_B^{l,k} \right\} = \text{nsst\_de}(B) \tag{17}$$

Fig. 5 One set of glioma disease MRI and PET image fusion results. **a** MRI; **b** PET; **c** CNN; **d** CSMCA; **e** LLF-IOI; **f** NFA; **g** NSST-PAPCNN; **h** PC-LLE-NSCT; **i** PSF; **j** Proposed
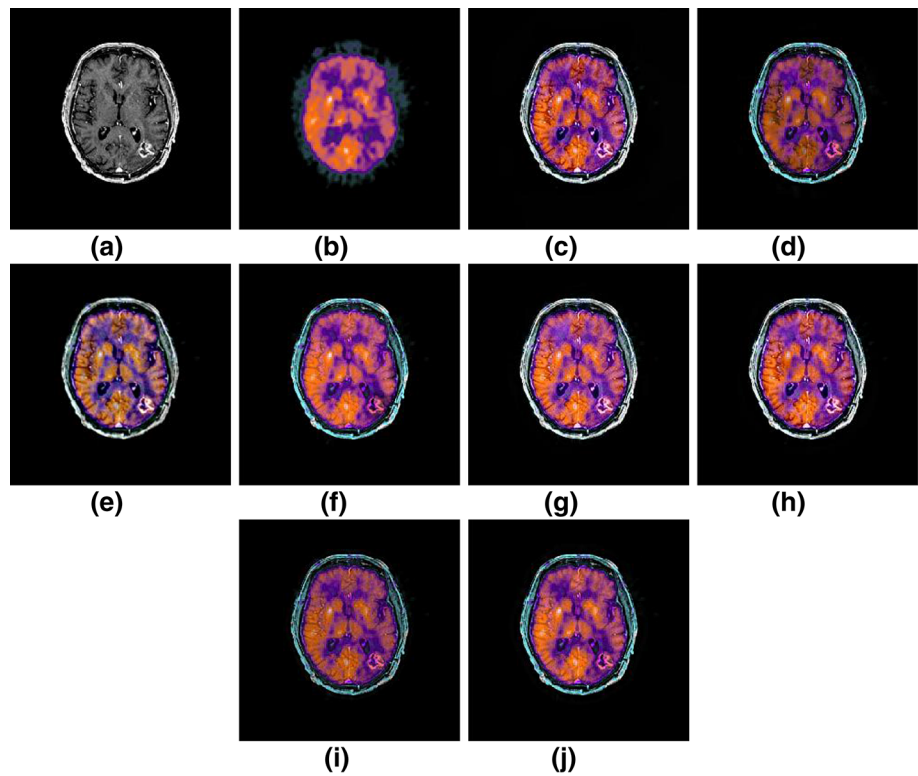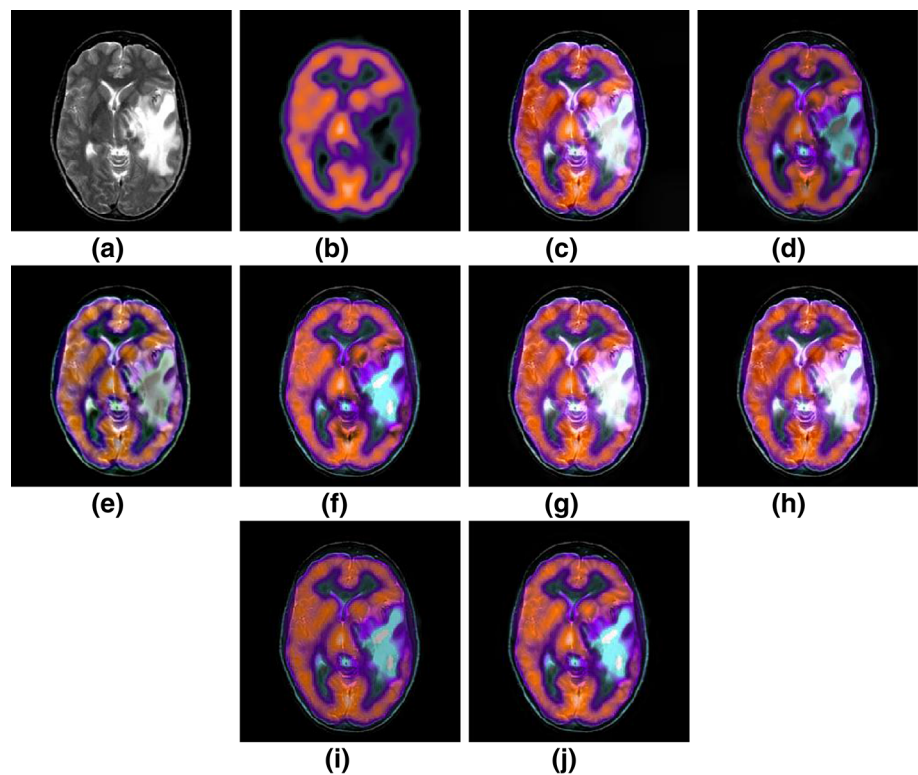

Fig. 6 One set of glioma disease MRI and SPECT image fusion results. **a** MRI; **b** SPECT; **c** CNN; **d** CSMCA; **e** LLF-IOI; **f** NFA; **g** NSST-PAPCNN; **h** PC-LLE-NSCT; **i** PSF; **j** Proposed
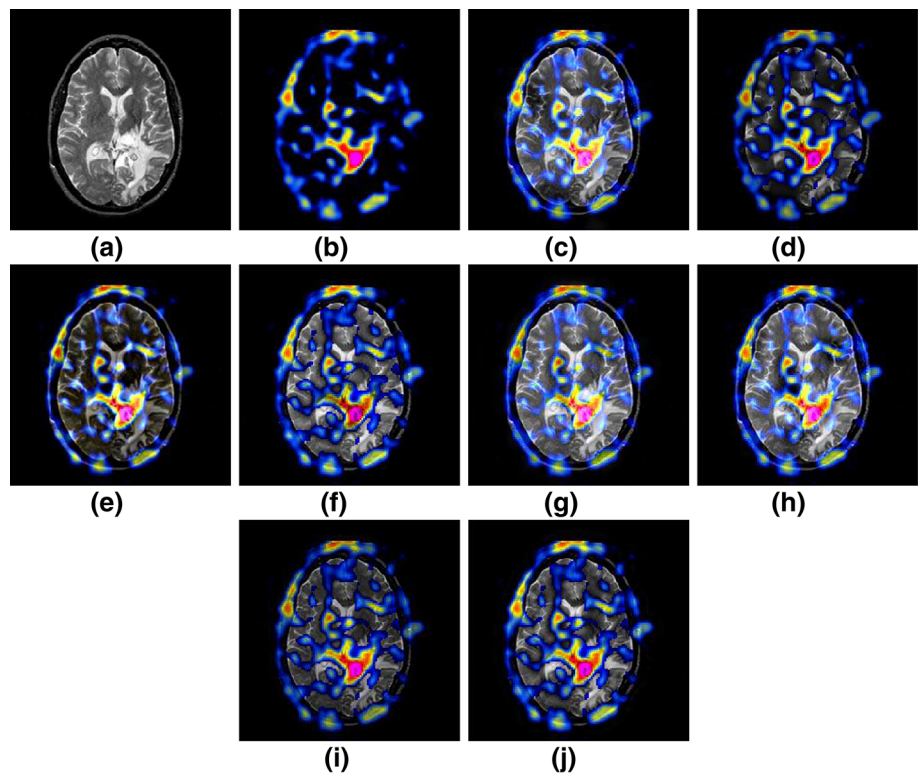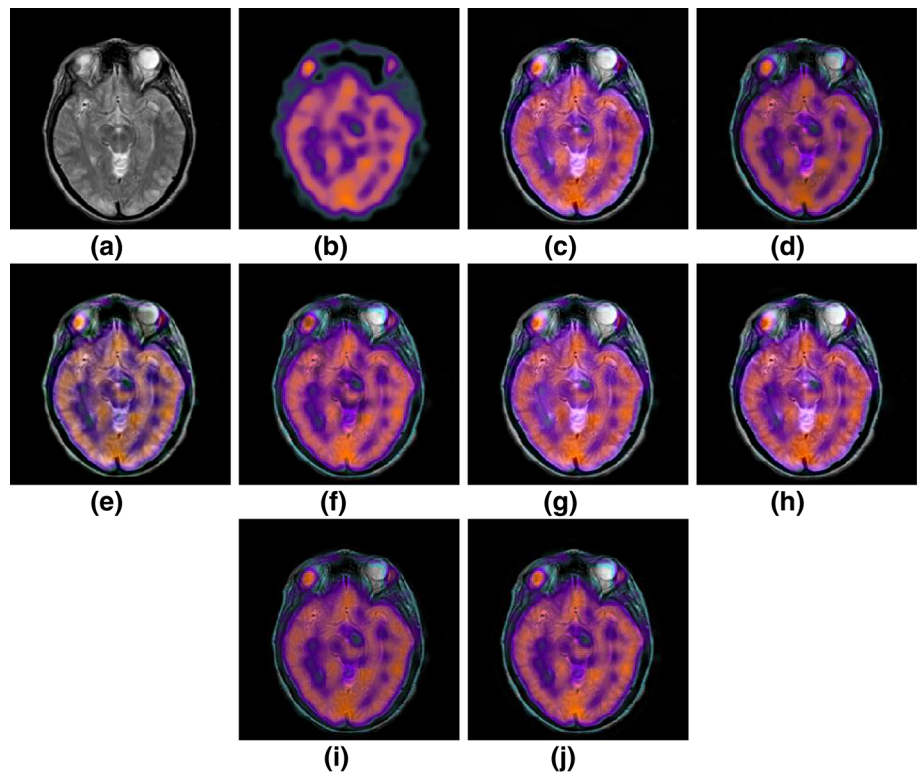
Fig. 7 One set of mild
Alzheimer's disease MRI and
PET image fusion results.
a MRI; **b** PET; **c** CNN;
d CSMCA; **e** LLF-IOI; **f** NFA;
g NSST-PAPCNN; **h** PC-LLE-
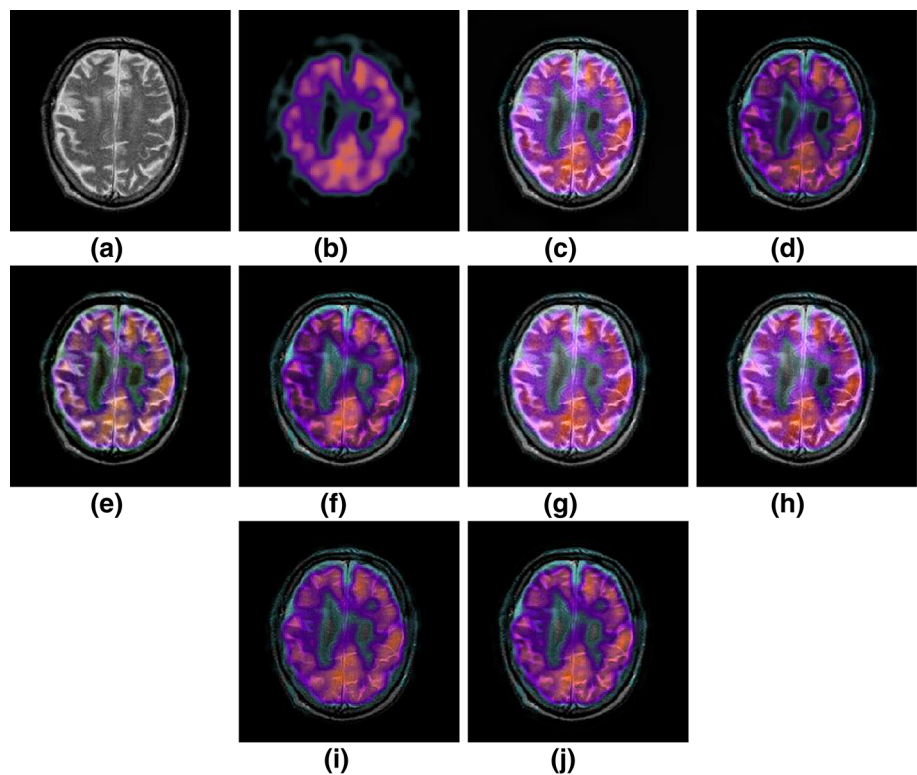NSCT; **i** PSF; **j** Proposed
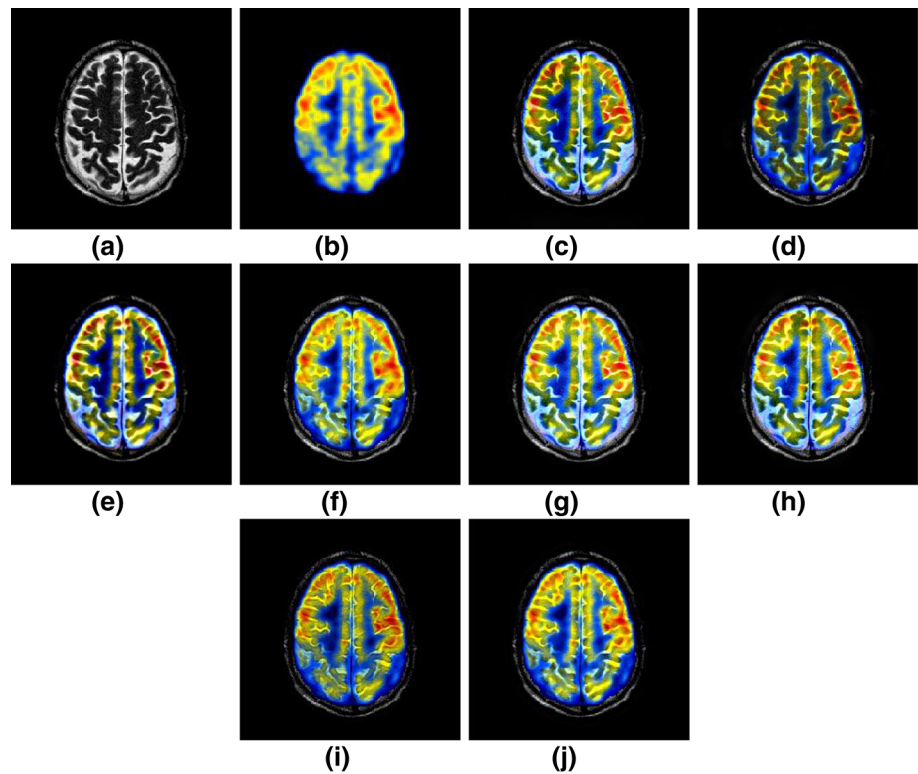
Fig. 8 One set of metastatic
bronchogenic carcinoma MRI
and SPECT image fusion
results. **a** MRI; **b** SPECT;
c CNN; **d** CSMCA; **e** LLF-IOI;
(f) NFA; **g** NSST-PAPCNN;
h PC-LLE-NSCT; **i** PSF;
j Proposed

Fig. 9 One set of hypertensive encephalopathy MRI and SPECT image fusion results. **a** MRI; **b** SPECT; **c** CNN; **d** CSMCA; **e** LLF-IOI; **f** NFA; **g** NSST-PAPCNN; **h** PC-LLE-NSCT; **i** PSF; **j** Proposed
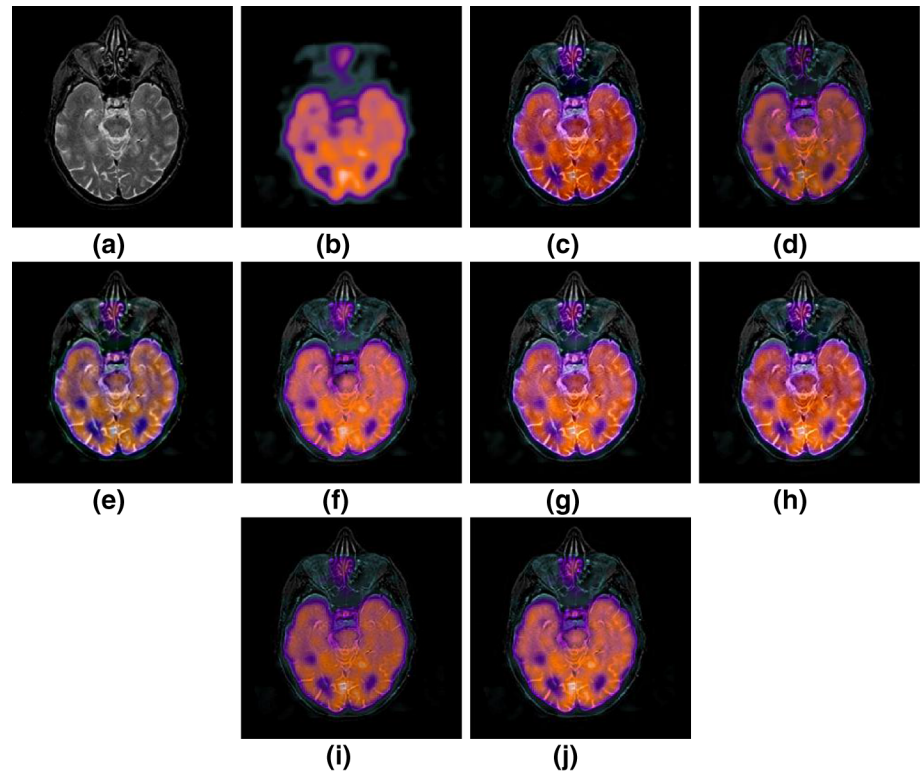

Fig. 10 One set of motor neuron disease MRI and SPECT image fusion results. **a** MRI; **b** SPECT; **c** CNN; **d** CSMCA; **e** LLF-IOI; **f** NFA; **g** NSST-PAPCNN; **h** PC-LLE-NSCT; **i** PSF; **j** Proposed

Fig. 11 One set of normal aging MRI and SPECT image fusion results. **a** MRI; **b** SPECT; **c** CNN; **d** CSMCA; **e** LLF-IOI; (f) NFA; **g** NSST-PAPCNN; **h** PC-LLE-NSCT; **i** PSF; **j** Proposed

Table 1 Mean quality of glioma disease (30 image pairs of MRI-PET)

	EN	SD	NMI	SS	VIF
CNN	**3.4365**	54.9628	0.7288	0.7149	**0.4732**
CSMCA	3.4309	**55.2631**	*0.7291*	*0.7265*	0.4673
LLF-IOI	3.3015	50.0933	0.6648	0.6827	0.4274
NFA	3.3876	53.5912	0.6951	0.7104	0.4633
NSST-PAPCNN	3.3981	54.1147	0.6964	0.7018	0.4430
PC-LLE-NSCT	3.4058	54.3635	0.7009	0.7087	0.4613
PSF	3.2162	50.6057	0.6726	0.6895	0.4284
Proposed	*3.4316*	*55.1167*	**0.7321**	**0.7289**	*0.4704*

Table 2 Mean quality of glioma disease (13 image pairs of MRI-SPECT)

	EN	SD	NMI	SS	VIF
CNN	4.0332	**60.1184**	*0.7198*	0.4215	0.2813
CSMCA	**4.1269**	59.9960	0.7134	0.4316	*0.2894*
LLF-IOI	3.7614	56.9402	0.6867	0.4192	0.2590
NFA	3.9186	58.2348	0.6719	0.3932	0.2418
NSST-PAPCNN	4.0627	58.6948	0.6879	0.4183	0.2635
PC-LLE-NSCT	4.0558	59.6137	0.6942	*0.4339*	0.2759
PSF	4.0186	57.0918	0.6751	0.4140	0.2667
Proposed	*4.1227*	*60.0368*	**0.7229**	**0.4481**	**0.2970**

Table 3 Mean quality of mild Alzheimer's disease (10 image pairs of MRI-PET)

	EN	SD	NMI	SS	VIF
CNN	3.4810	67.2903	0.8468	0.6332	**0.2894**
CSMCA	3.5207	*68.0381*	**0.8537**	*0.6356*	0.2618
LLF-IOI	3.5094	67.7324	0.8471	0.6271	0.2579
NFA	3.4356	65.3818	0.8326	0.6117	0.2360
NSST-PAPCNN	**3.5614**	67.9053	0.8441	0.6327	0.2617
PC-LLE-NSCT	3.4709	65.4932	0.8204	0.6268	0.2781
PSF	3.3790	66.9214	0.8306	0.6208	0.2615
Proposed	*3.5530*	**68.2571**	*0.8524*	**0.6447**	*0.2887*

Table 4 Mean quality of metastatic bronchogenic carcinoma (11 image pairs of MRI-SPECT)

	EN	SD	NMI	SS	VIF
CNN	4.5680	*65.0373*	0.9416	**0.6261**	0.4427
CSMCA	4.5361	64.1775	0.9328	0.6208	0.4358
LLF-IOI	4.4303	60.1125	0.9154	0.6027	0.4247
NFA	4.6005	64.5230	0.9428	0.6207	*0.4461*
NSST-PAPCNN	**4.6179**	64.6127	0.9314	0.6178	0.4408
PC-LLE-NSCT	4.5868	64.1052	*0.9470*	0.6206	0.4391
PSF	4.4291	63.1609	0.9162	0.6014	0.4108
Proposed	*4.6160*	**65.0717**	**0.9511**	*0.6239*	**0.4508**

Table 5 Mean quality of hypertensive encephalopathy (10 image pairs of MRI-SPECT)

	EN	SD	NMI	SS	VIF
CNN	4.6866	**59.0718**	0.9365	0.6651	0.4612
CSMCA	**4.8576**	56.3227	0.9419	*0.6734*	0.4590
LLF-IOI	4.4054	53.8491	0.9056	0.6417	0.4334
NFA	4.5782	54.2814	0.9146	0.6558	0.4487
NSST-PAPCNN	4.6930	56.7419	*0.9468*	0.6718	**0.4651**
PC-LLE-NSCT	4.6026	55.9921	0.9367	0.6648	0.4587
PSF	4.5124	54.6234	0.9263	0.6486	0.4475
Proposed	*4.8521*	*58.8718*	**0.9594**	**0.6974**	*0.4628*

Table 6 Mean quality of motor neuron disease (11 image pairs of MRI-SPECT)

	EN	SD	NMI	SS	VIF
CNN	4.3413	48.6742	**0.8965**	0.6538	*0.3761*
CSMCA	4.4647	49.6958	0.8697	**0.6712**	0.3746
LLF-IOI	4.1829	45.4876	0.8307	0.6352	0.3458
NFA	4.4535	46.3916	0.8615	0.6457	0.3548
NSST-PAPCNN	*4.5036*	48.2961	0.8365	0.6592	0.3719
PC-LLE-NSCT	4.4608	**50.8289**	0.8756	0.6437	0.3629
PSF	4.3062	46.3109	0.8417	0.6316	0.3418
Proposed	**4.5085**	*50.7957*	*0.8904*	*0.6672*	**0.3820**

Table 7 Mean quality of normal aging (16 image pairs of MRI-SPECT)

	EN	SD	NMI	SS	VIF
CNN	4.2768	50.2328	0.9027	0.5843	**0.4748**
CSMCA	4.1545	*51.6948*	*0.9209*	0.5964	0.4602
LLF-IOI	4.0341	48.6309	0.8981	0.5612	0.4473
NFA	4.1948	49.1654	0.9035	0.5741	0.4589
NSST-PAPCNN	*4.2876*	51.5940	**0.9227**	*0.6093*	0.4618
PC-LLE-NSCT	4.1349	50.2568	0.9141	0.6024	0.4538
PSF	4.0612	49.5587	0.8861	0.5714	0.4429
Proposed	**4.3342**	**52.1836**	0.9174	**0.6180**	*0.4731*

3.2 Low-frequency fusion

The low-frequency sub-band contains most information of the source images (texture structure and background). In this paper, an energy attribute (EA) fusion strategy is presented in the low-frequency fusion. This EA fusion strategy is divided into three steps:

1. The intrinsic property values of the low-frequency sub-band are computed as

$$IP_A = \mu_A + Me_A \tag{18}$$

$$IP_B = \mu_B + Me_B \tag{19}$$

where μ and Me represent the mean value and the median value of L_A and L_B, respectively.

2. The EA function E_A and E_B are calculated by

$$E_A(x,y) = \exp(\alpha|L_A(x,y) - IP_A|) \tag{20}$$

$$E_B(x,y) = \exp(\alpha|L_B(x,y) - IP_B|) \tag{21}$$

where $\exp(\alpha|L_A(x, y - IP_A)|)$ represents the exponential operator, and α denotes the modulation parameter.

2. The fused low-frequency sub-band is obtained by a weighted mean

$$L_F(x,y) = \frac{E_A(x,y) \times L_A(x,y) + E_B(x,y) \times L_B(x,y)}{E_A(x,y) + E_B(x,y)} \tag{22}$$

3.3 High-frequency fusion

While low-frequency sub-band contains most information about the source images (such as background and texture), high-frequency sub-bands contain more information about details in images (for example, pixel-level information). Since in the PCNN model one pixel corresponds to one neuron, it is suitable to use PCNN in high-frequency sub-bands. In addition, modulating PCNN with MSMG can increase the spatial correlation in the image. Therefore, the MSMG operator can be used to adjust the linking strength between β_{ij}^1 and β_{ij}^2

$$\beta_{ij}^A = M_A \tag{23}$$
$$\beta_{ij}^B = M_B \tag{24}$$

where M_A and M_B are computed by Eq. (7).

The high-frequency sub-bands are merged based on this MSMG-PCNN model until all neurons are activated (equal to 1). The fused high-frequency sub-bands can be obtained by

$$H_F^{l,k}(x,y) = \begin{cases} H_A^{l,k}(x,y), & \text{if } T_{xy,A} \geq T_{xy,B}; \\ H_B^{l,k}(x,y), & \text{otherwise.} \end{cases} \tag{25}$$

where $T_{xy,A}$ and $T_{xy,B}$ can be computed using Eq (15).

3.4 NSST reconstruction

The fused image F is reconstructed by L_F and $H_F^{l,k}$ through the inverse NSST according to Eq (4)

$$F = nsst\_re\left(L_F, H_F^{l,k}\right) \tag{26}$$

4 Experiments

To validate the proposed algorithm, a set of experiments was made using three datasets representing different diseases: (1) glioma, (2) mild Alzheimer's, and (3) hypertensive encephalopathy. The proposed algorithm was compared with seven state-of-the-art image fusion methods. Qualitative and quantitative analyses were made to assess its performance. The code of the paper is made available.[1]

4.1 Datasets

To verify the proposed algorithm, more than 100 pairs of multimodal medical images were used, including 30 image pairs of MRI-PET and 13 image pairs of MRI-SPECT of glioma disease, 10 image pairs of MRI-PET of mild Alzheimer's disease, 11 image pairs of MRI-SPECT of Metastatic bronchogenic carcinoma, 10 image pairs of MRI-SPECT of hypertensive encephalopathy, 11 image pairs of MRI-SPECT of motor neuron disease, and 16 image pairs of MRI-SPECT of normal aging. All the image pairs can be downloaded from the Whole Brain Atlas dataset [1]. All the pairs have been perfectly registered, and the size of all images is 256×256.

4.2 Comparison methods

The proposed BM-PCNN-NSST algorithm is compared with seven state-of-the-art fusion methods. There methods are the convolutional neural network (CNN) [[24], [53]], the convolutional sparsity-based morphological component analysis (CSMCA) [25], the information of interest in local Laplacian filtering domain (LLF-IOI) [11], the neuro-fuzzy approach (NFA) [9], the parameter-adaptive PCNN in NSST domain (NSST-PAPCNN) [48], the phase congruency and local Laplacian energy in NSCT domain (PC-LLE-NSCT) [52], and the parallel saliency features (PSF) [12]. These methods are recently proposed fusion methods. The parameters that we used in our experiments are the same as in their papers.

4.3 Parameter settings

In the proposed BM-PCNN-NSST algorithm, the following parameters were used:

- the NSST decomposition level N is set to 4;
- the number of directions in each level is set to 16,16,8,8;
- the modulation parameter is set to 4;
- the scales number of MSMG operator t is set to 3.

4.4 Evaluation metrics

To analyze the performance of the proposed algorithm in a quantitative way, we evaluate the different fusion methods using five metrics: entropy (EN), standard deviation (SD), normalized mutual information (NMI) [14], Piella's structure similarity (SS) [33], and visual information fidelity (VIF) [16]. In general, both SD and EN can measure the amount of information of the fused image. NMI evaluates the amount of information transferred from the source images to fused image. SS mainly evaluates the structure similarity between source images and fused image. VIF evaluates the visual information fidelity between the source images and fused image. More detailed information about these evaluation metrics can be found on the references related to each fusion method.

4.5 Experimental results

The results of medical image fusion cannot be completely dependent on visual effects evaluation. As long as the feature information is not lost, the medical diagnosis will not be misjudged because of this, and the visual effect will be acceptable. Therefore, in this paper, each disease demonstrates a set of experimental results, which is shown in Figs. 5, 6, 7, 8, 9, 10, and 11. Different methods have different visual effects, but the feature information does not seem to be lost. Therefore, objective evaluation indicators are needed for a further quantitative evaluation.

The mean value of each metrics of different fusion methods is listed in Tables 1, 2, 3, 4, 5, 6 and 7. Each column represents the same metrics for different methods. The highest value is shown in bold, while the second highest in italic. It can be seen that the proposed method performs the best in half of the cases. Even if it is not the highest value in one column, it is still the second highest value, except in the NMI of the normal aging case.

---

[1]

5 Conclusion

In this paper, a multimodal medical image fusion algorithm is proposed based on boundary measured PCNN and EA fusion strategies in NSST domain. The main advantage of the proposed algorithm is that the two fusion strategies are suitable for different scales. Decomposing images into different scales with NSST can give full play to the advantages of the two fusion strategies. Meanwhile, as an excellent decomposition method, NSST can well blend the differences of multimodal medical images. The performance of the proposed algorithm has been verified in public datasets, which represents it has reached state-of-the-art level. One of the important outcomes of this paper is reported in "Appendix," which showed the experimental performance of different values of α and t. One can see that when $\alpha = 4$ and $t = 3$ the performance is the best in most cases. Since the deep learning technology has been widely used, in the future research, we will focus on the deep learning method in multimodal medical image fusion [2, 7, 30].

Compliance with ethical standards

Appendix

Figures 12, 13, 14, 15, 16, 17, and 18 show the experimental performance of different α and t. One can see that when $\alpha = 4$, $t = 3$, the performance is the best in most cases.
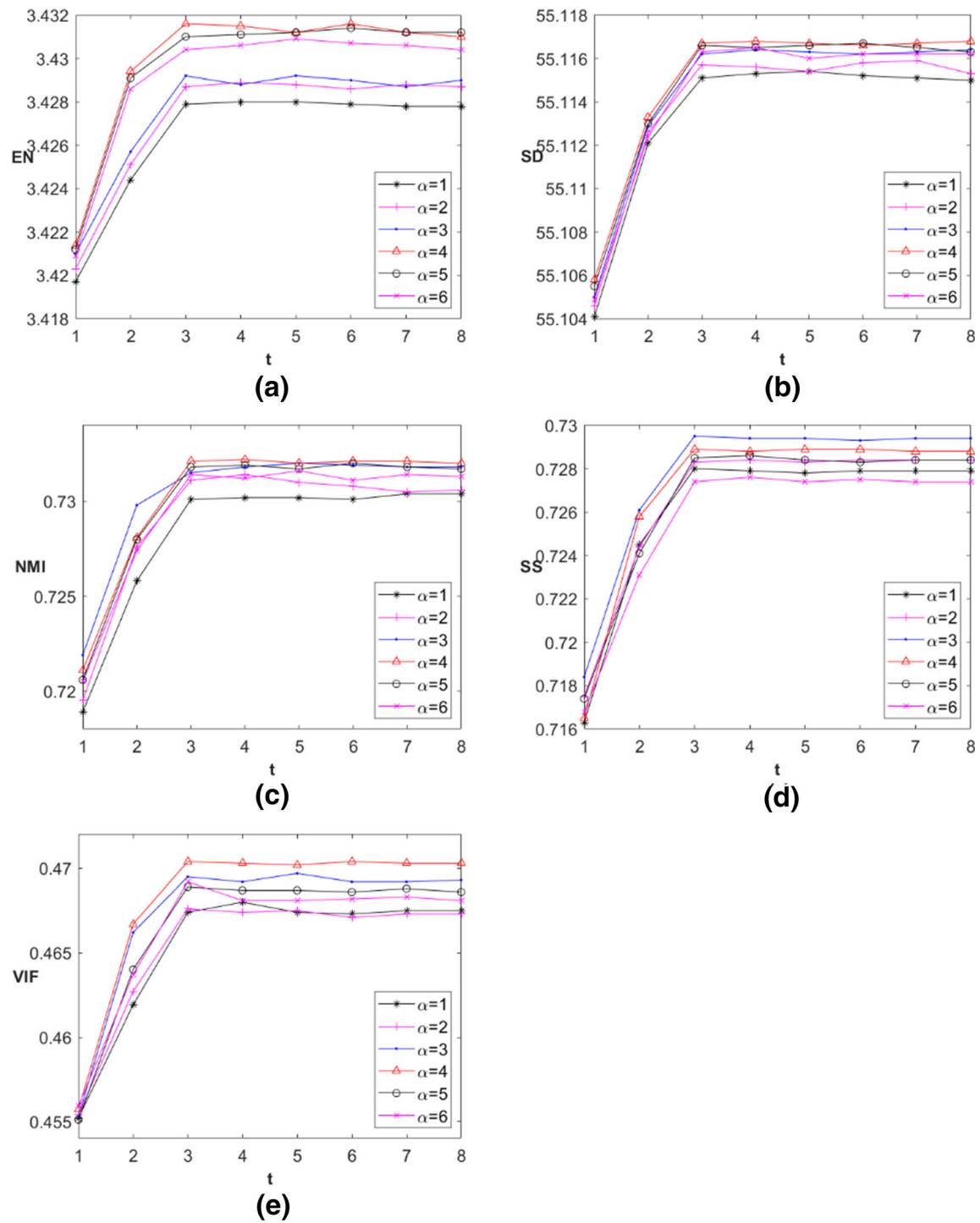
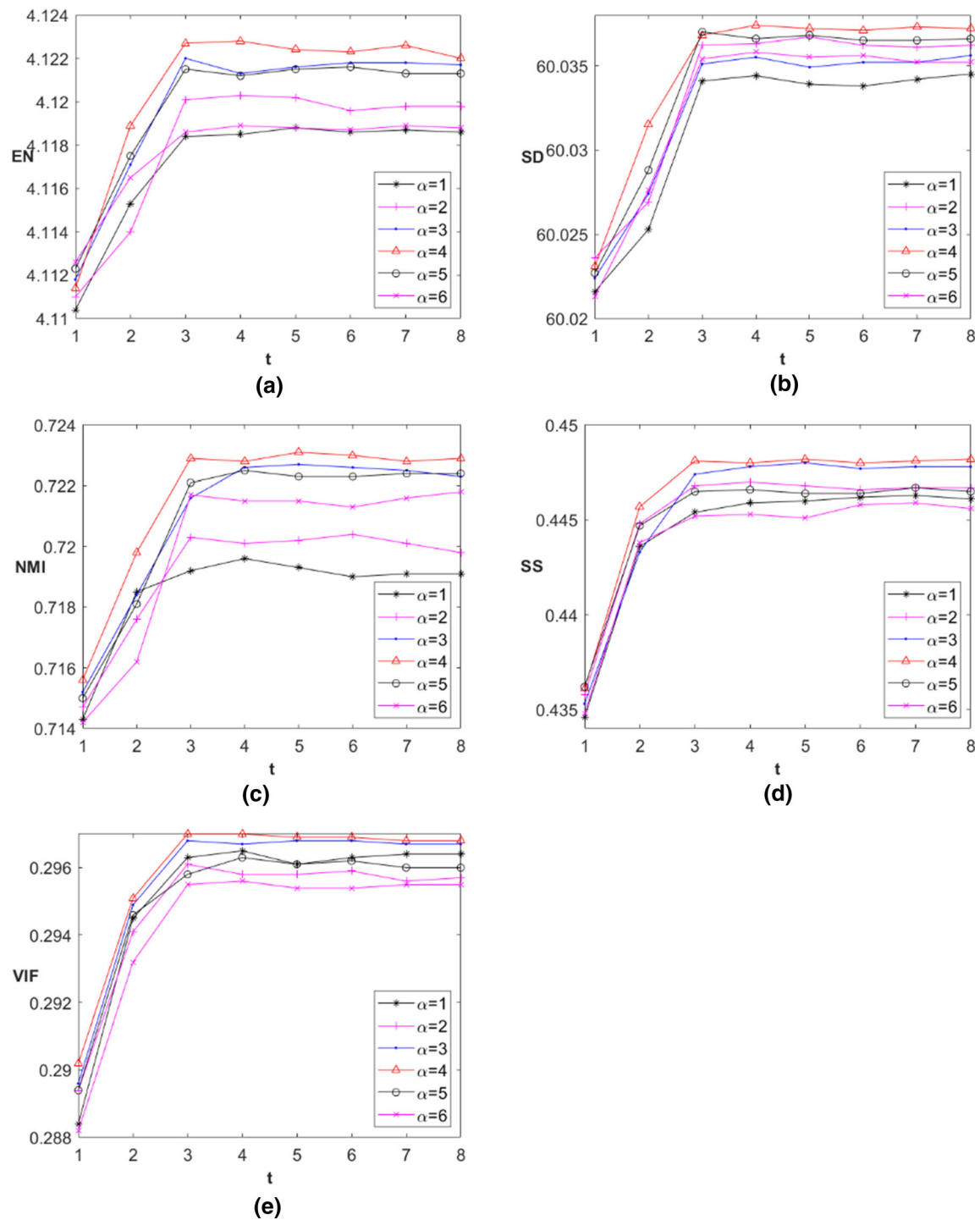**Fig. 12** Glioma disease (MRI-PET). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF

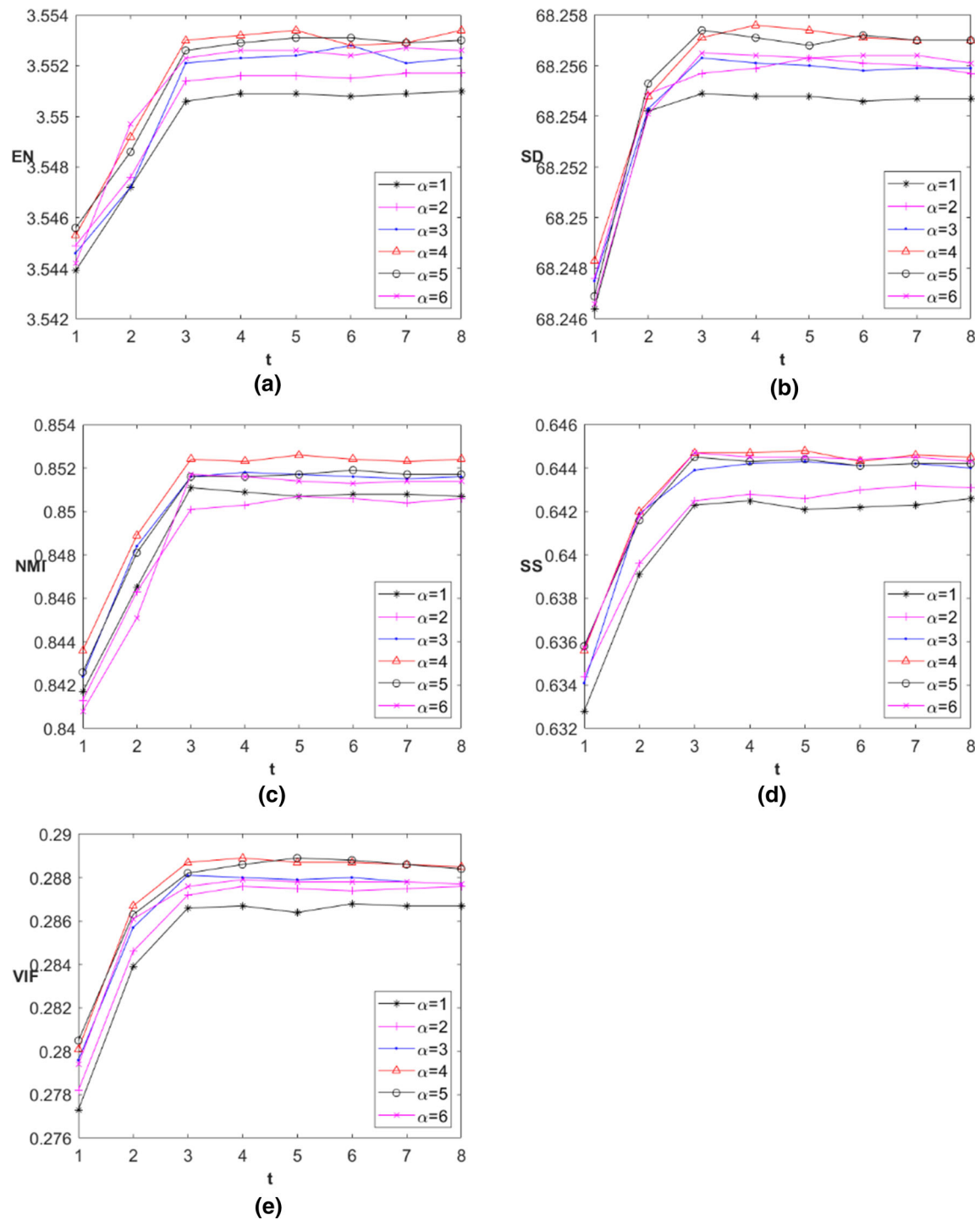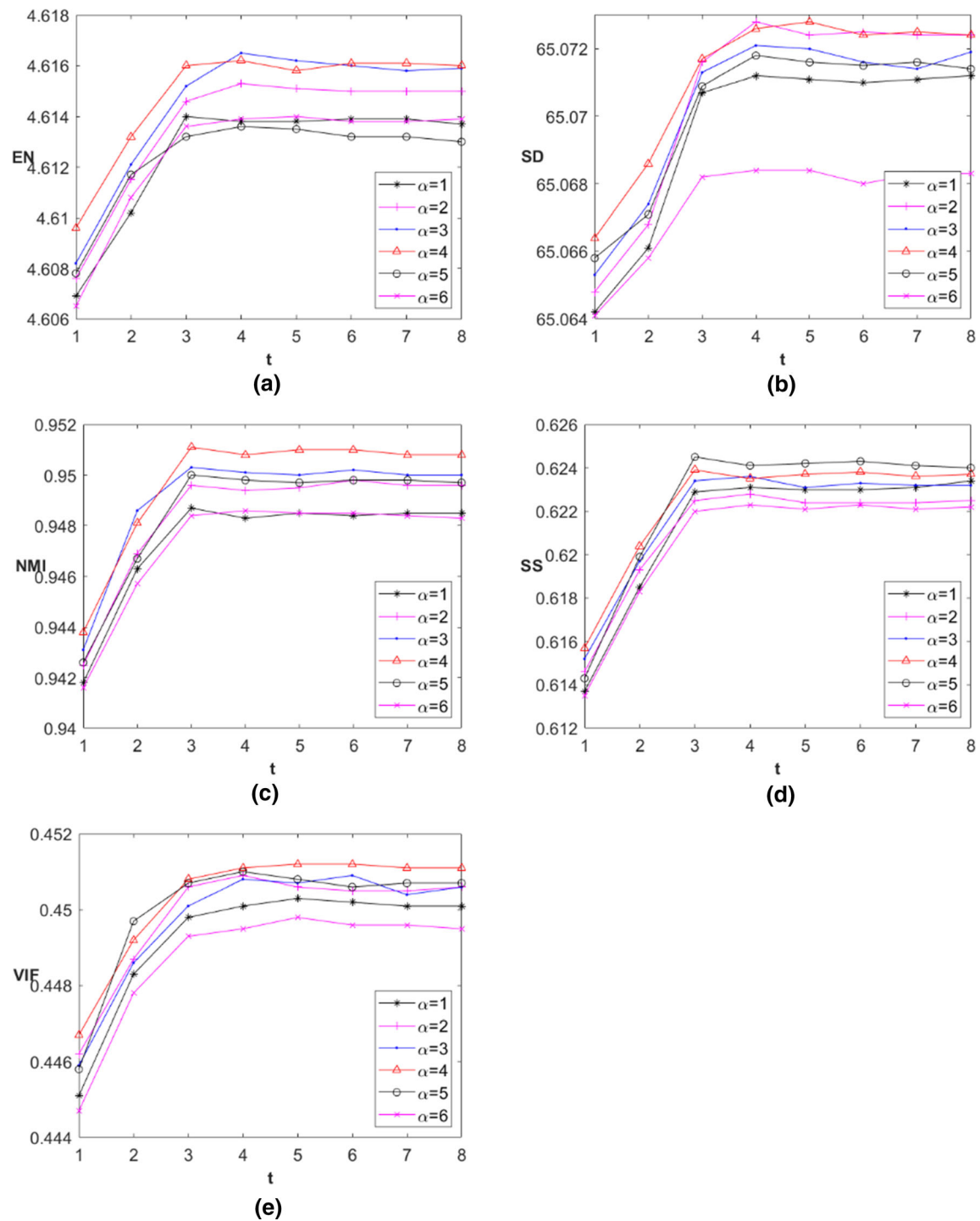Fig. 13 Glioma disease (MRI-SPECT). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF
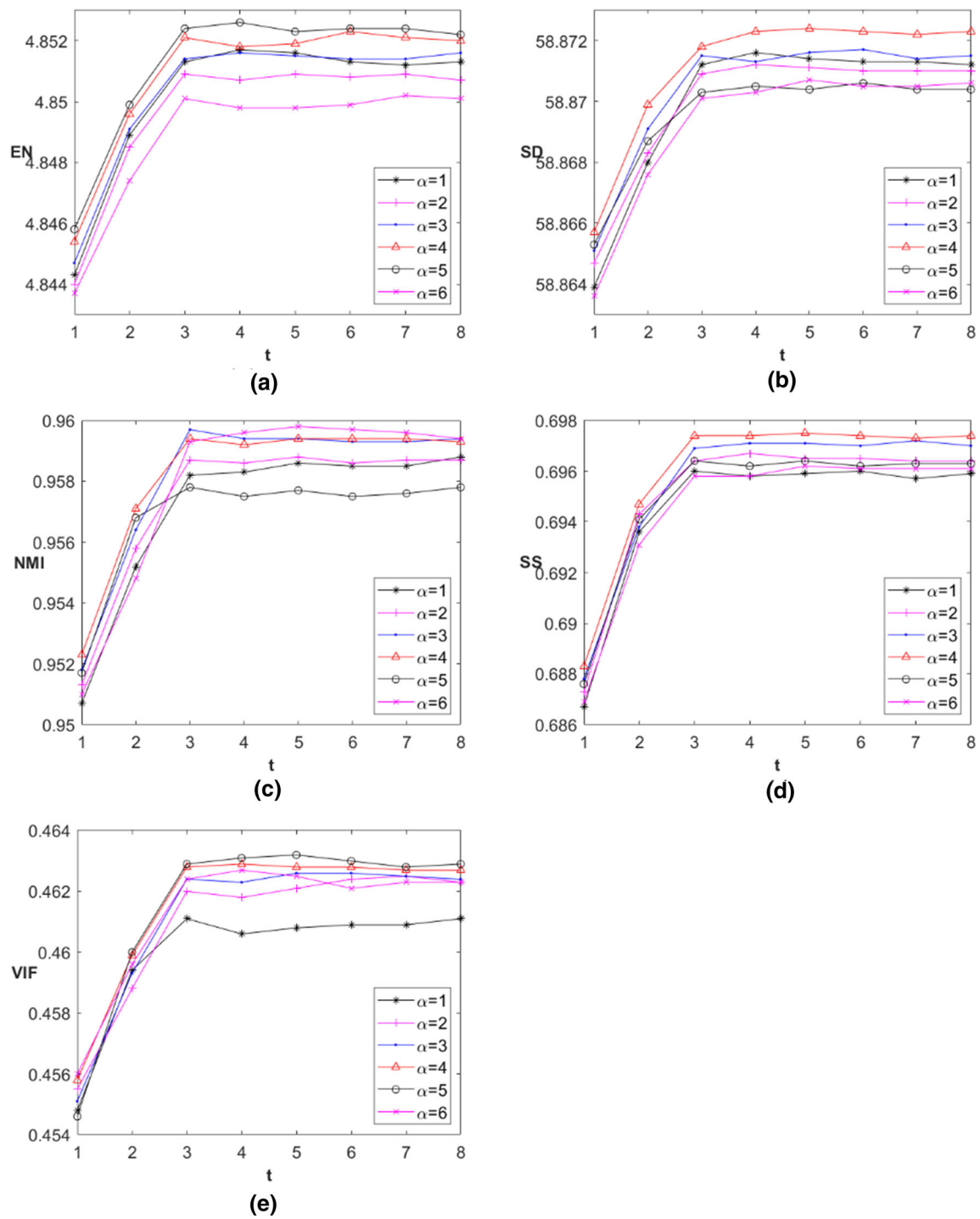
Fig. 14 Mild Alzheimer's disease (MRI-PET). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF

Fig. 15 Metastatic bronchogenic carcinoma (MRI-SPECT). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; (e) t–VIF

Fig. 16 Hypertensive encephalopathy (MRI-SPECT). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF
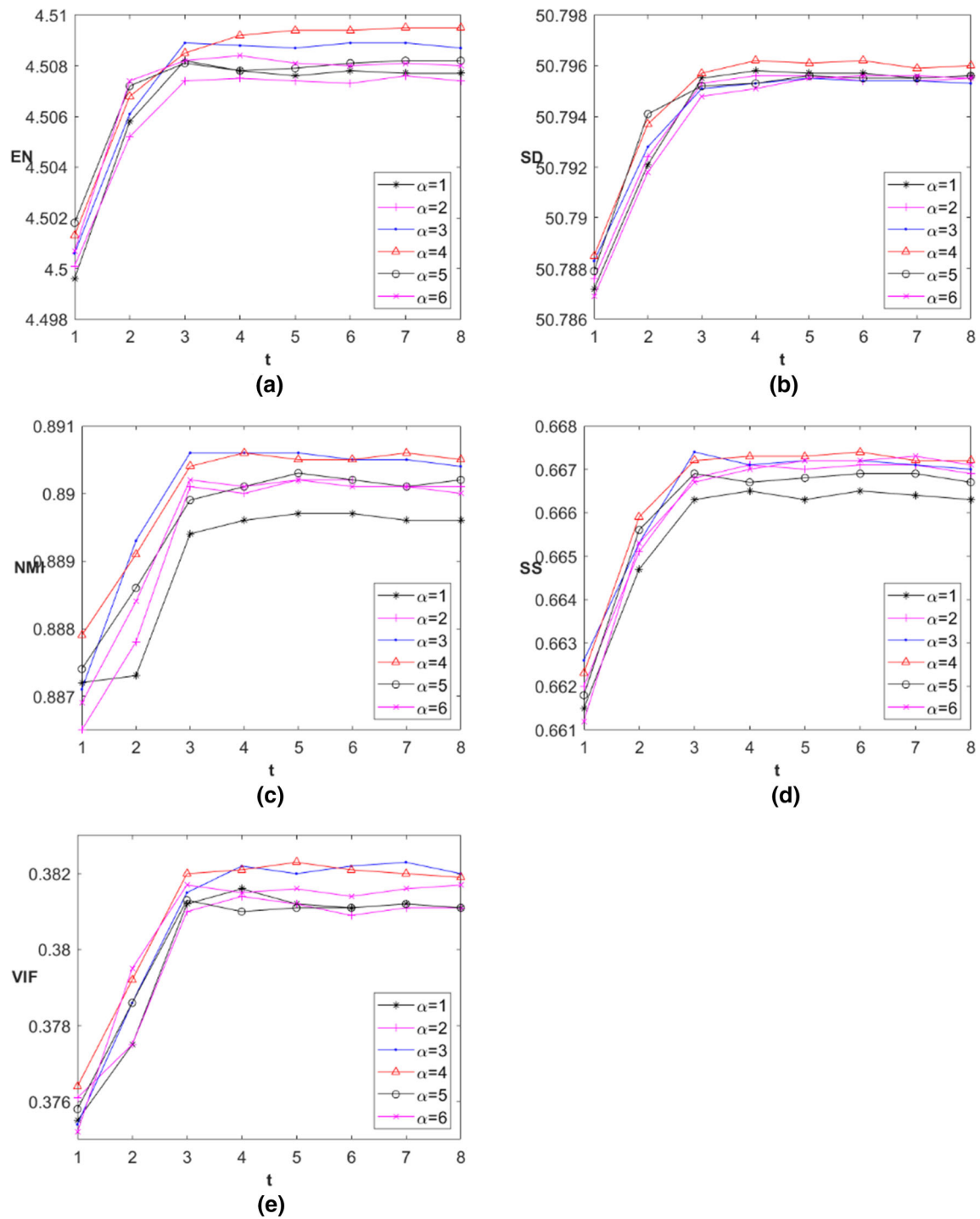
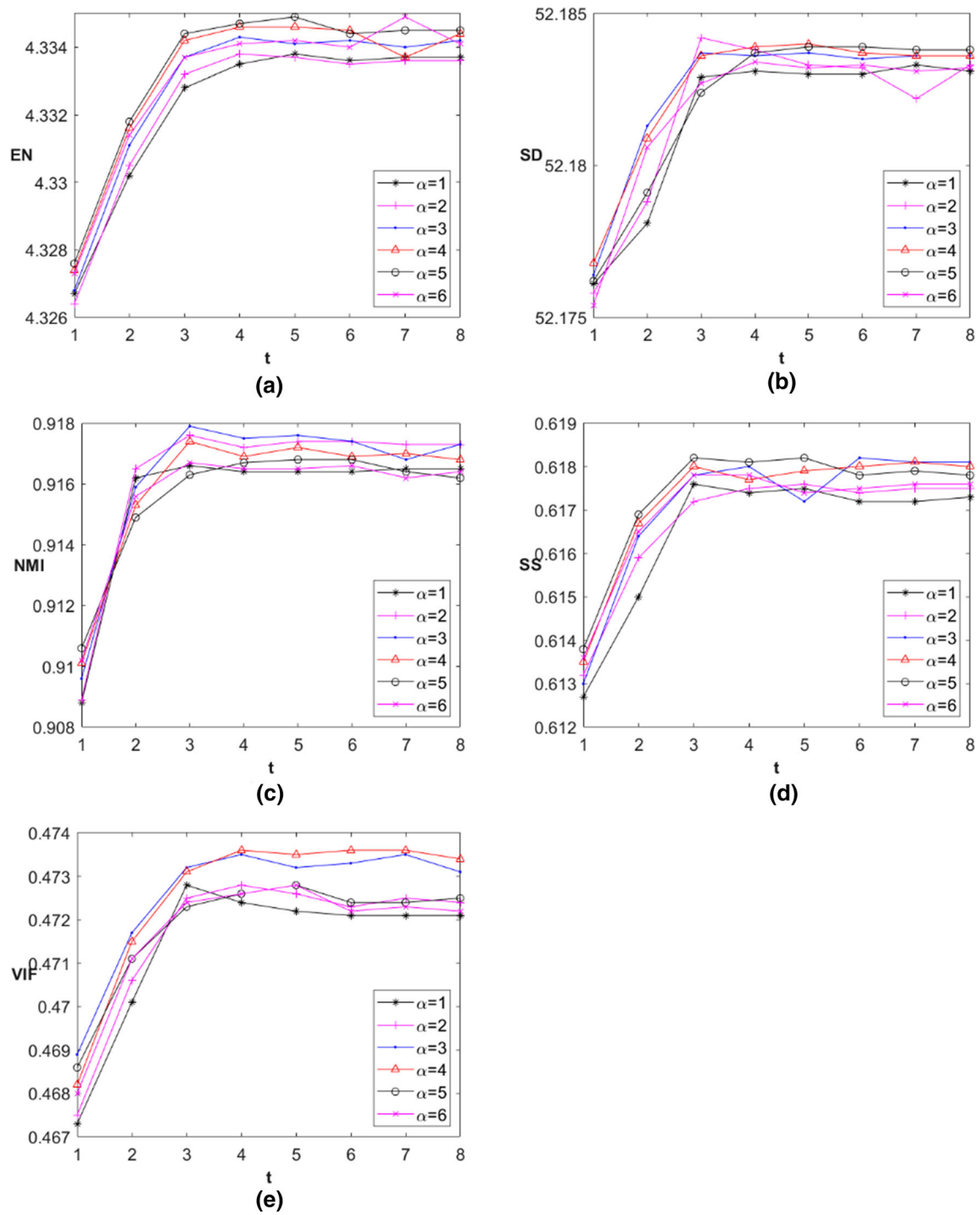Fig. 17 Motor neuron disease (MRI-SPECT). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF

Fig. 18 Normal aging (MRI-SPECT). **a** t–EN; **b** t–SD; **c** t–NMI; **d** t–SS; **e** t–VIF

References

1. Whole brain atlas. http://www.med.harvard.edu/AANLIB/
2. Ahmed I, Din S, Jeon G, Piccialli F (2019) Exploring deep learning models for overhead view multiple object detection. IEEE Int Things J. https://doi.org/10.1109/JIOT.2019.2951365
3. Amato F, Moscato V, Picariello A, Piccialli F, Sperl G (2018) Centrality in heterogeneous social networks for lurkers detection: an approach based on hypergraphs. Concurr Comput Pract Exp 30(3):e4188
4. Asha C, Lal S, Gurupur VP, Saxena PP (2019) Multi-modal medical image fusion with adaptive weighted combination of nsst bands using chaotic grey wolf optimization. IEEE Access 7:40782–40796
5. Bebortta S, Senapati D, Rajput NK, Singh AK, Rathi VK, Pandey HM, Jaiswal AK, Qian J, Tiwari P (2020) Evidence of power-law behavior in cognitive IoT applications. Neural Comput Appl pp 1–13
6. Burt P, Adelson E (1983) The Laplacian pyramid as a compact image code. IEEE Trans Commun 31(4):532–540
7. Casolla G, Cuomo S, Di Cola VS, Piccialli F (2020) Exploring unsupervised learning techniques for the internet of things. IEEE Trans Ind Inform 16(4):2621–2628. https://doi.org/10.1109/TII.2019.2941142
8. Chouhan V, Singh SK, Khamparia A, Gupta D, Moreira C, Damasevicius R, de Albuquerque VHC (2020) A novel transfer learning based approach for pneumonia detection in chest X-ray images. Appl Sci 10(2):559
9. Das S, Kundu MK (2013) A neuro-fuzzy approach for medical image fusion. IEEE Trans Biomed Eng 60(12):3347–3353
10. Du J, Li W, Lu K, Xiao B (2016) An overview of multi-modal medical image fusion. Neurocomputing 215:3–20
11. Du J, Li W, Xiao B (2017) Anatomical-functional image fusion by information of interest in local Laplacian filtering domain. IEEE Trans Image Process 26(12):5855–5866
12. Du J, Li W, Xiao B (2018) Fusion of anatomical and functional images using parallel saliency features. Inf Sci 430:567–576
13. Easley G, Labate D, Lim WQ (2008) Sparse directional image representations using the discrete shearlet transform. Appl Comput Harmon Anal 25(1):25–46
14. Estevez PA, Tesmer M, Perez CA, Zurada JM (2009) Normalized mutual information feature selection. IEEE Trans Neural Netw 20(2):189–201
15. Gochhayat SP, Kaliyar P, Conti M, Prasath V, Gupta D, Khanna A (2019) LISA: lightweight context-aware IoT service architecture. J Clean Prod 212:1345–1356
16. Han Y, Cai Y, Cao Y, Xu X (2013) A new image fusion performance metric based on visual information fidelity. Inf Fusion 14(2):127–135
17. Huang W, Jing Z (2007) Evaluation of focus measures in multi-focus image fusion. Pattern Recognition Lett 28(4):493–500
18. Jaiswal AK, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJ (2019) Identifying pneumonia in chest x-rays: a deep learning approach. Measurement 145:511–518
19. Kong W, Zhang L, Lei Y (2014) Novel fusion method for visible light and infrared images based on NSST-SF-PCNN. Infrared Phys Technol 65:103–112
20. Kumar S, Tiwari P, Zymbler M (2019) Internet of things is a revolutionary approach for future technology enhancement: a review. J Big Data 6(1):111
21. Li S, Kang X, Fang L, Hu J, Yin H (2017) Pixel-level image fusion: a survey of the state of the art. Inf Fusion 33:100–112
22. Liu S, Wang J, Lu Y, Li H, Zhao J, Zhu Z (2019) Multi-focus image fusion based on adaptive dual-channel spiking cortical model in non-subsampled shearlet domain. IEEE Access 7:56367–56388
23. Liu X, Mei W, Du H (2018) Multi-modality medical image fusion based on image decomposition framework and nonsubsampled shearlet transform. Biomed Signal Process Control 40:343–350
24. Liu Y, Chen X, Cheng J, Peng H (2017) A medical image fusion method based on convolutional neural networks. In: 2017 20th international conference on information fusion (Fusion), pp 1–7. IEEE
25. Liu Y, Chen X, Ward RK, Wang ZJ (2019) Medical image fusion via convolutional sparsity based morphological component analysis. IEEE Signal Process Lett 26(3):485–489
26. Mallick PK, Ryu SH, Satapathy SK, Mishra S, Nguyen GN (2019) Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network. IEEE Access 7:46278–46287
27. Nair RR, Singh T (2019) Multi-sensor medical image fusion using pyramid-based dwt: a multi-resolution approach. IET Image Proc 13(9):1447–1459
28. Piccialli F, Bessis N, Jung JJ (2020) Data science challenges in industry 4.0. IEEE Trans Ind Inform
29. Piccialli F, Casolla G, Cuomo S, Giampaolo F, di Cola VS (2020) Decision making in iot environment through unsupervised learning. IEEE Intell Syst 35(1):27–35. https://doi.org/10.1109/MIS.2019.2944783
30. Piccialli F, Cuomo S, di Cola VS, Casolla G (2019) A machine learning approach for iot cultural data. J Ambient Intell Human Comput pp 1–12
31. Piccialli F, Cuomo S, Giampaolo F, Casolla G, di Cola VS (2020) Path prediction in iot systems through Markov chain algorithm. Fut Gen Comput Syst
32. Piccialli F, Yoshimura Y, Benedusi P, Ratti C, Cuomo S (2020) Lessons learned from longitudinal modeling of mobile-equipped visitors in a complex museum. Neural Comput Appl 32:7785–7801. https://doi.org/10.1007/s00521-019-04099-8
33. Piella G, Heijmans H (2003) A new quality metric for image fusion. In: Proceedings 2003 international conference on image processing (Cat No 03CH37429), vol 3, pp III-173. IEEE
34. Polinati S, Dhuli R (2020) Multimodal medical image fusion using empirical wavelet de-composition and local energy maxima. Optik 205:163947
35. Qian J, Tiwari P, Gochhayat SP, Pandey HM (2020) A noble double dictionary based ecg compression technique for ioth. IEEE Intern Things J
36. Rong S, Zhou H, Zhao D, Cheng K, Qian K, Qin H (2018) Infrared x pattern noise reduction method based on shearlet transform. Infrared Phys Technol 91:243–249
37. Tan W, Xiang P, Zhang J, Zhou H, Qin H (2020) Remote sensing image fusion via boundary measured dual-channel pcnn in multi-scale morphological gradient domain. IEEE Access 8:42540–42549
38. Tan W, Zhang J, Xiang P, Zhou H, Thitøn W (2020) Infrared and visible image fusion via nsst and pcnn in multiscale morphological gradient domain. In: Optics, photonics and digital technologies for imaging applications VI, vol 11353, p 113531E. International society for optics and photonics
39. Tan W, Zhou H, Rong S, Qian K, Yu Y (2018) Fusion of multi-focus images via a Gaussian curvature filter and synthetic focusing degree criterion. Appl Opt 57(35):10092–10101
40. Tan W, Zhou H, Song J, Li H, Yu Y, Du J (2019) Infrared and visible image perceptive fusion through multi-level Gaussian curvature filtering image decomposition. Appl Opt 58(12):3064–3073
41. Tan W, Zhou Hx, Yu Y, Du J, Qin H, Ma Z, Zheng R (2017) Multi-focus image fusion using spatial frequency and discrete

wavelet transform. In: AOPC 2017: Optical sensing and imaging technology and applications, vol 10462, p 104624 K. International society for optics and photonics

42. Tiwari P, Melucci M (2018) Towards a quantum-inspired framework for binary classification. In: Proceedings of the 27th ACM international conference on information and knowledge management, pp 1815–1818

43. Tiwari P, Melucci M (2019) Binary classifier inspired by quantum theory. Proc AAAI Conf Artif Intell 33:10051–10052

44. Tiwari P, Melucci M (2019) Towards a quantum-inspired binary classifier. IEEE Access 7:42354–42372

45. Tiwari P, Qian J, Li Q, Wang B, Gupta D, Khanna A, Rodrigues JJ, de Al-buquerque VHC (2018) Detection of subtype blood cells using deep learning. Cogn Syst Res 52:1036–1044

46. Wang Z, Wang S, Guo L (2018) Novel multi-focus image fusion based on PCNN and random walks. Neural Comput Appl 29(11):1101–1114

47. Yin H (2018) Tensor sparse representation for 3-D medical image fusion using weighted average rule. IEEE Trans Biomed Eng 65(11):2622–2633

48. Yin M, Liu X, Liu Y, Chen X (2018) Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain. IEEE Trans Instrum Meas 68(1):49–64

49. Zhang Q, Guo BL (2009) Multifocus image fusion using the nonsubsampled contourlet transform. Signal Process 89(7):1334–1346

50. Zhang Y, Bai X, Wang T (2017) Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure. Inf Fusion 35:81–101

51. Zhou Z, Wang B, Li S, Dong M (2016) Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters. Inf Fusion 30:15–26

52. Zhu Z, Zheng M, Qi G, Wang D, Xiang Y (2019) A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain. IEEE Access 7:20811–20824

53. Pandey HM, Windridge D (2019) A comprehensive classification of deep learning libraries. In: Third international congress on information and communication technology. Springer, Singapore