Journal of Bionic Engineering (2021) 18:1452–1462
https://doi.org/10.1007/s42235-021-00100-0

RESEARCH ARTICLE

An Improved Pigeon-Inspired Optimization for Multi-focus Noisy Image Fusion

Yingda Lyu[1,4] · Yunqi Zhang[2] · Haipeng Chen[3,4]

Received: 6 July 2021 / Revised: 16 August 2021 / Accepted: 17 August 2021 / Published online: 12 November 2021

Abstract
Image fusion technology is the basis of computer vision task, but information is easily affected by noise during transmission. In this paper, an Improved Pigeon-Inspired Optimization (IPIO) is proposed, and used for multi-focus noisy image fusion by combining with the boundary handling of the convolutional sparse representation. By two-scale image decomposition, the input image is decomposed into base layer and detail layer. For the base layer, IPIO algorithm is used to obtain the optimized weights for fusion, whose value range is gained by fusing the edge information. Besides, the global information entropy is used as the fitness index of the IPIO, which has high efficiency especially for discrete optimization problems. For the detail layer, the fusion of its coefficients is completed by performing boundary processing when solving the convolution sparse representation in the frequency domain. The sum of the above base and detail layers is as the final fused image. Experimental results show that the proposed algorithm has a better fusion effect compared with the recent algorithms.

Keywords Improved pigeon-inspired optimization · Convolutional sparse representation · Noisy image fusion · Bionic algorithm

1 Introduction

In the process of taking photos, due to the movement of objects or cameras, different focal lengths will make the pictures blurry and poor registration, and it is difficult to focus on all objects in the picture under different depths of field [1]. These factors seriously reduce the fusion quality of multi-focus images and make it difficult to obtain complete information on all levels. Therefore, multi-focus image fusion can generate a multi-focus image using a set of images acquired with different focus. In the process of image acquisition, image fusion will result in noise accumulation, which will make the image quality descend after fusion. This makes multi-focus noise image fusion a very challenging task.

At present, most image fusion methods focus on pixel-level fusion, which mainly include transform domain-based algorithms and spatial domain-based algorithms. Among them, methods based on multi-scale transform are widely discussed. They decompose the source image into multiple scales, mainly including base layer and detail layer, extract sub-band coefficients of different scales, and then fuse transform coefficients according to different fusion rules, such as wavelet transform [2], gradient pyramid [3], contrast pyramid [4], etc.. Since more spatial features can be extracted on different scales, image fusion can be effectively performed. However, there are still difficulties in their strong sensitivity to sensor noise.

The spatial domain-based methods directly fuse the input image to the pixel information, and they are classified as pixel-based, block-based, and region-based methods [5–7]. However, the images fused by these methods usually have problems such as loss of detail and texture information, reducing contrast, artifact and insufficient sharpness, which make them face huge challenges.

✉ Haipeng Chen
chenhp@jlu.edu.cn

1 Public Computer Education and Research Center, Jilin University, Changchun 130012, China

2 College of Software, Jilin University, Changchun 130012, China

3 College of Computer Science and Technology, Jilin University, Changchun 130012, China

4 Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

Therefore, for overcoming the above shortcomings, many scholars have conducted different degrees of research based on the image fusion strategy of sparse representation (SR) [8–11]. Experiments have proved that compared with the MST-based algorithms, the SR-based method has splendid performance [12, 13]. Nevertheless, most SR-based methods always produce blocking effects. Therefore, a Convolutional Sparse Representation (CSR) can be used instead of SR, which processes the whole image rather than local image blocks. CSR is essentially an image representation model of invariant displacement, which significantly improves the quality of fusion in areas of misalignment. However, the main drawback of the convolutional sparse coding method is the high computational complexity. Generally, CSR is a nonconvex problem and many existing methods provide little to no guarantees for global convergence. Finding a solution to the convolutional sparse coding problem in a reasonable amount of time is not only challenging, but it is also difficult to find a good local minimum. The substantial computational cost of convolutional sparse coding problems has recently been shown to be significantly reduced by solving them in the frequency domain, but the periodic boundary conditions imposed by this approach can produce boundary artifacts. Therefore, it is also crucial to propose a method to avoid these effects. In recent years, Deep learning (DL) has achieved great success in many computer vision and image processing problems. The application of deep learning technology in the field of pixel-level multi-source image fusion has become an active topic. Liu et al. first applied CNN technology in the field of multi-source image fusion. Specifically, a classification CNN based on the Siamese architecture is designed to learn the clarity of the relative position in each source image [14]. Amin-Naji et al. used ensemble Learning technology to improve the classification accuracy of the CNN model [15]. Tang et al. proposed a pixel-based CNN model, which divides the pixels in each source image into focused pixels, defocused pixels, and unknown pixels [16].

Turning image fusion problems into optimization problems has attracted attention in recent years. Optimization methods are used to optimize the parameters in the above methods, and the fusion weight is the most common optimized parameter. Ref. [17] used genetic algorithms to find the best fusion weights in the wavelet domain and the eigenspace domain. Raghavendra et al. [18] used an improved particle swarm optimization to obtain the wavelet domain weights, both of which have certain effects, but they also expose the shortcomings of low efficiency and accuracy. In recent years, this type of method has been used in more fields. Zhao et al. [19] used the information entropy and gradient constraint of the fusion image as fitness evaluation indicators, and used the gradient descent method to optimize the fusion weight.

Although the fusion images of most optimized fusion algorithms can obtain better results in objective evaluation indicators, they assign too many weights to the irrelevant parts of the image. Fortunately, many efficient bio-inspired global search algorithms [20–24] are used to tackle this issue. These novel bio-inspired optimization algorithms usually have the feature of multiple structures. Hence, they complete the exploration and development tasks at different stages of the algorithm iteration. A large number of variants based on bio-inspired optimization algorithms have been proposed and applied in various fields. For example, Liu et al. proposed an improved PSO algorithm to solve the problems of multilevel thresholding [25]. Uros Mlakar et al. made improvements to the DE algorithm [26]. Gai et al. proposed an Improved Quantum-Behaved Particle Swarm Optimization (IQPSO) to solve the problems of image fusion [27]. Wang et al. used a cooperative pigeon-inspired optimization algorithm with dynamic distance threshold (CPIOD) to obtain higher performance [28]. The Pigeon-Inspired Optimization (PIO) algorithm has been shown to have advantages over traditional algorithms in standard function optimization [22].

In addition, when the source image is processed, noise will accumulate, resulting in a decrease in the quality of the fused image. Therefore, it matters that the image fusion can resist noise environment [29, 30]. However, the above method is mainly for the edge statistics of wavelet coefficients, and the noise reduction effect is not good.

To overcome the above shortcomings, we propose an improved PIO(IPIO) algorithm, and it is used to get the optimized weights for multi-focus noisy image fusion. The main contributions are as follows:

1. A novel improved PIO (IPIO) framework is proposed, and accordingly an effective image fusion method is presented for multi-focus noisy image. In the base layer, IOPO can get the optimized weights. In the detail layer, the detailed information can be handled well.
2. We provide a more accurate range for weights by fusing edge features, and obtain better optimization results through pigeon-inspired optimization, using global information entropy as the fitness index of pigeon-inspired optimization. It can effectively display important areas, and can achieve decent results in global entropy and local entropy. Besides, the detailed information can be transferred to the fusion image.
3. The proposed method introduces the boundary handling for convolutional sparse representation into multi-focus noisy image fusion, which is essentially an image representation model with invariable displacement, it can significantly improve the fusion quality of mismatched regions. Simultaneously, boundary handling is used to avoid possible boundary artifacts.

2 PIO Algorithm

Studies have shown that pigeons can easily return home through three guidance tools: geomagnetic field information, sun altitude information and landmark information. Inspired by this phenomenon, Duan et al. proposed a pigeon flock optimization algorithm. It has been verified that the algorithm can effectively solve problems such as numerical design and parameter optimization [31].

2.1 Map and Compass Operator

The PIO algorithm runs through two operators: the map and compass operator and the landmark operator. V_i represent speed of the pigeon i in the iteration round of t. The following equation gives the updated process [24].

$$V_i(t) = V_i(t-1) \times e^{-Rt} + r * (X_g - X_i(t-1)). \tag{1}$$

In the formula, R is the factor of the map and compass operator, X_g is the best solution of the current pigeons in all locations, r is the size of the window. X_i represent the position of the pigeon i in the iteration round of t.

$$X_i(t) = X_i(t-1) + V_i(t). \tag{2}$$

Figure 1 shows the map and compass operator. The flying pigeon will adjust its flying direction according to the position of the best pigeon. The first part of Eq. (1) represents the current direction of the pigeon, and the second part represents the process of the pigeon following the best pigeon (the current best solution).

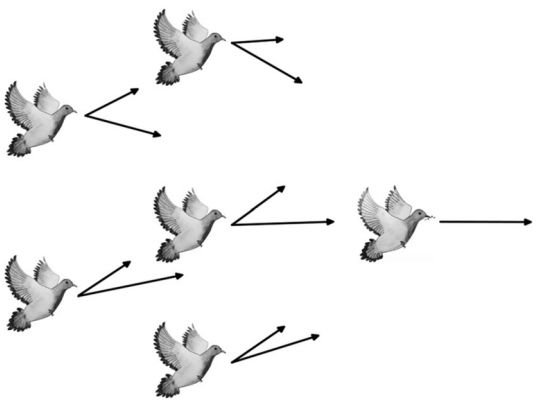

Fig. 1 Compass and map operator

2.2 Landmark Operator

In the landmark operator, all the pigeons will be sorted according to their fitness. The first half of the pigeons in the ranking will calculate the position of the center pigeon according to Eq. (3). This position is regarded as a landmark, and the remaining half of the pigeons will update their position based on this landmark, as shown in Eq. (4).

$$X_c(t) = \frac{\sum_{N_p} X_i(t)\, fitness\,(X_i(t))}{\sum_{N_p} fitness\,(X_i(t))}, \tag{3}$$

$$X_i(t) = X_i(t-1) + r_2(X_c(t) - X_i(t-1)), \tag{4}$$

$$N_p = \frac{N_p(t-1)}{2}, \tag{5}$$

where X_c is the position of the center pigeon (landmark), X_i is the current position of all the pigeons, fitness reflects the quality of the pigeon individual, and $N_p(t)$ is the number representing the pigeons, r_2 is the size of the window.

Figure 2 is a schematic diagram of the landmark operator. In the process of algorithm simulation, the pigeons with low fitness are considered to be unfamiliar with the landmarks, and they must follow the pigeons with high fitness. The number of pigeons in the circle is half of the number of pigeons calculated according to formula (5).

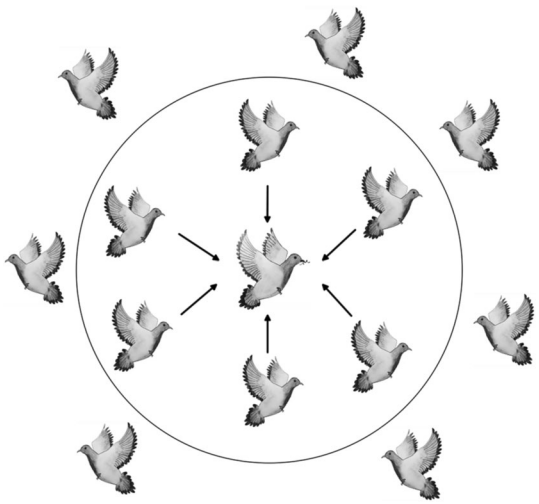

Fig. 2 Landmark operator

3 The Improved Pigeon-Inspired Optimization (IPIO) Algorithm

This paper considers a new bio-inspired swarm intellect optimizer PIO to obtain $\vec{X}_{best}$.

First, the pixel weight range is determined with the help of edge information, and then the fitness function is determined under the fusion framework and the pigeon-inspired optimization is used to obtain the pixel weight.

3.1 Determination of the Weight Range

Most optimization algorithms do not consider the range of optimization weights. Too much weight is assigned to irrelevant parts of the image, resulting in too much useless information, making the target not be accurately expressed. Thus, we use the fused edge information to determine the range of weight.

Many researchers have integrated edge information into image fusion [21]. And it is effective to use edge information to determine the range of weight w_{ij}. So Canny algorithm is used to extract the edge binary image E_{A} and E_{B}, and obtain complementary edges $\overline{\overline{E_{\mathrm{A}}}}$, $\overline{\overline{E_{\mathrm{B}}}}$ using Eq. (6):

$$\begin{cases} \overline{\overline{E_A}}(i,j) = E_A(i,j) \oplus E_B(i,j) \wedge E_A(i,j) \\ \overline{\overline{E_B}}(i,j) = E_B(i,j) \oplus E_B(i,j) \wedge E_A(i,j) \end{cases}, \tag{6}$$

where $\oplus$ indicates XOR operation; $\wedge$ indicates AND operation. $\overline{\overline{E_{\mathrm{A}}}}$, $\overline{\overline{E_{\mathrm{B}}}}$ can effectively represent the unique and effective detail information of the image. This paper considers that giving higher weights to the corresponding position of the weight matrix can obtain a better fused image, w_{ij}^{lp}, w_{ij}^{up}, respectively, indicate the upper and lower limits of the weight w_{ij} at the position (i,j), that is w_{ij}^{up}, w_{ij}^{lp} satisfy Eq. (7):

$$\begin{cases} w_{ij}^{\mathrm{up}} = \theta^{up} + \lambda\overline{\overline{E_{\mathrm{A}}}}(i,j) - \lambda\overline{\overline{E_{\mathrm{B}}}}(i,j) \\ w_{ij}^{lp} = \theta^{lp} - \lambda\overline{\overline{E_{\mathrm{A}}}}(i,j) + \lambda\overline{\overline{E_{\mathrm{B}}}}(i,j) \end{cases}, \tag{7}$$

where θ^{up}, θ^{lp} and λ are all parameters.

3.2 Design of the Loss Function

The source image is decomposed into $\mathbf{I}_k^b$ and $\mathbf{I}_f^d$, and the final fused image I_f is the fusion of $\mathbf{I}_k^b$ and $\mathbf{I}_f^d$. The result depends on the pixel weight $\vec{X}$. Therefore, this paper hopes to obtain a better fused image with appropriate pixel weight $\vec{X}$. However, there is no optimal fusion image for the image fusion problem, so an objective index is used to evaluate the pros and cons of the fusion result. Information Entropy (IE) is a commonly used objective index [21], such as the Eq. (8):

$$\mathrm{IE(F)} = -\sum_{i=0}^{L-1} p_F(g)\log_2 p_F(g), \tag{8}$$

g represents the total number of gray levels, where $g = 256$,$p_F = \{p_F(1), \ldots, p_F(g)\}$ represents the probability of each gray level in **F**. A higher **IE** means that the image has more information, so the fusion quality can be evaluated by **IE**. Regarding **IE** as the fitness, the function of fitness **Y** with respect to the pixel weight $\vec{X}$ can be obtained by Eqs. (7) and (8), as in Eq. (9):

$$\mathrm{Y} = \mathrm{Y}\left(\vec{X}\right) = \mathrm{IE} = \mathrm{IE}\left(\mathrm{F}\left(\vec{X}\right)\right). \tag{9}$$

Then the optimal pixel weight $\vec{X}_{best}$ can be expressed as Eq. (10):

$$\vec{X}_{\mathrm{best}} = \arg\max_{\vec{X}} \mathrm{Y}(\vec{X}). \tag{10}$$

4 Method

To obtain excellent fusion for multi-focus noisy images, we present a multi-focus noise image fusion method based on the improved pigeon-inspired optimization bionic algorithm. The proposed fusion algorithm framework is shown in Fig. 3. The detailed fusion algorithm includes the following four steps.

Step 1 Two-scale image decomposition: the source image is decomposed into base layer $\mathbf{I}_k^b$ and detail layer $\mathbf{I}_k^d$ using the discrete gradient operator.

Step 2 Base layer fusion: the IPIO method is used to fuse the base layer.

Step 3 Detail layer fusion: the detail layer is fused by the boundary handling for CSR method.

Step 4 Two-scale image reconstruction: the fused base layer $\mathbf{I}_k^b$ and the detail layer $\mathbf{I}_k^d$ are through the superposition to obtain the final fused image.

4.1 Image Decomposition

First, the source image $\mathbf{I}_k$ is decomposed into base layer $\mathbf{I}_k^b$ and detail layer $\mathbf{I}_k^d$ by the discrete gradient operator. $\mathbf{I}_k^b$ is acquired by the following optimization process:

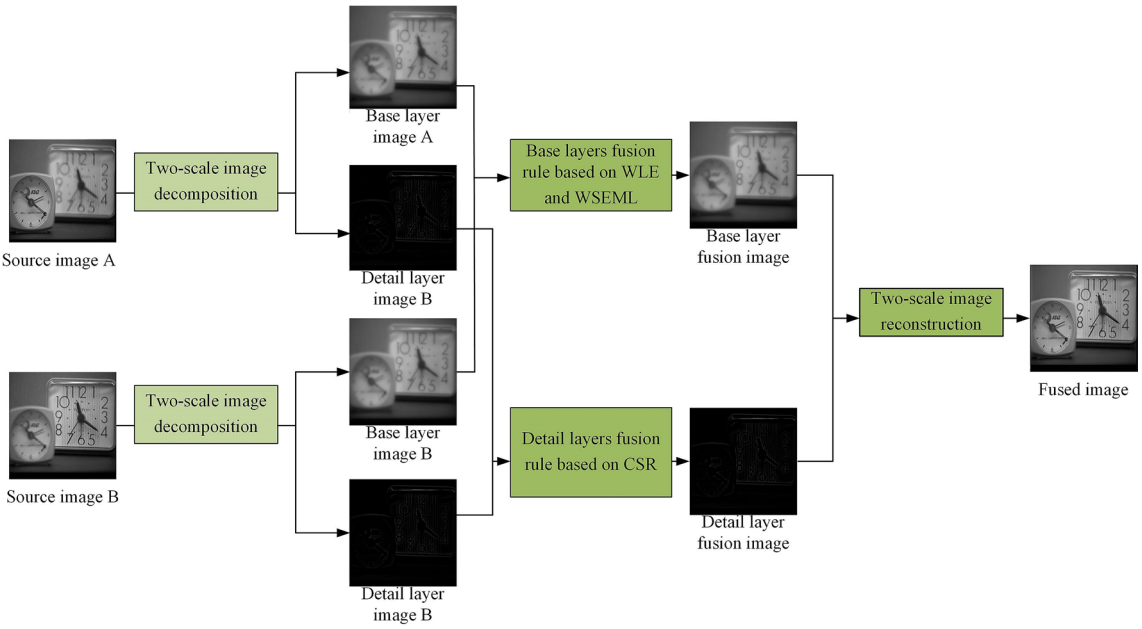

Fig. 3 The algorithm flow

$$\arg\min_{I_k^b} ||\boldsymbol{I}_k - \boldsymbol{I}_k^b||_F^2 + \eta\left(||g_x \times \boldsymbol{I}_k^b||_F^2 + ||g_y \times \boldsymbol{I}_k^b||_F^2\right). \quad (11)$$

It is a common Tikhonov regularization problem, which can be solved by fast Fourier transformation, where $g_x = [-1\ \ 1]$ and $g_y = [-1\ \ 1]^T$ are the row and column gradient filters. The value of the regularization parameter η is 5.

$$\arg\min_{I_k^b} ||\boldsymbol{I}_k - \boldsymbol{I}_k^b||_F^2 + \eta\left(||G_X\boldsymbol{I}_k^b||_F^2 + ||G_Y\boldsymbol{I}_k^b||_F^2\right). \quad (12)$$

The formula (12) is a problem of quadratic optimization. The standard approach is the least squares linear regression. However, if no $\mathbf{I}_k^b$ satisfies (12) or more than one $\mathbf{I}_k^b$ do, that is, the solution is not unique. This type of convex optimization problem can be converted into a dual problem, all of its local optimal solutions are global optimal solutions, and because the derivative of 0 is a necessary condition for the function to take the extreme value, the solution of this type of problem can directly set the partial derivative be 0 to obtain the global optimal solution. Therefore, to solve the above optimization problems, it should solve the partial derivative of $\mathbf{I}_k^b$ set to 0:

$$\boldsymbol{I}_k^b = \frac{\boldsymbol{I}_k}{1 + \eta\left(G_X^2 + G_Y^2\right)}. \quad (13)$$

The inverse frequency Fourier transform is performed by (13) to obtain the base layer component $\mathbf{I}_k^b$, and the detail layer component $\mathbf{I}_k^d$ can be obtained by subtracting the base layer component from the source image $\mathbf{I}_k$.

$$\boldsymbol{I}_k^d = \boldsymbol{I}_k - \boldsymbol{I}_k^b, \quad (14)$$

4.2 Base Layer Fusion Rule Based on IPIO

The fusion strategy for base layer bands also has a significant impact on the final fusion quality. The conventional averaging-based low-frequency fusion rule tends to cause the loss of energy in the fused image. To address this issue, we use the fusion edge feature to provide a more accurate range for the weight, and use the global information entropy as the fitness index of the pigeon-inspired optimization, and obtain a better optimization effect.

This paper uses pixel weight $\{w_{ij}\}$ for the base layer to perform fusion according to the following formula (15), w_{ij} represents the fusion weight on the pixel (i,j), which is used to calculate the value of $\mathbf{I}_f^b$, and $\{w_{ij}\}$ will be expressed as the pixel weight $\vec{X}$ according to Eq. (16). The use of pixel weight facilitates the use of optimization algorithms to improve the fusion effect, and the optimized weights can

effectively enhance the image quality and are suitable for real-time calculation:

$$\mathbf{I}_f^b(i,j) = w_{ij}A_h(i,j) + \big(1 - w_{ij}\big)B_h(i,j), \tag{15}$$

$$\begin{cases} X^d = w_{ij}, d = i + j(i-1) \\ \vec{X} = \left\{X^d, d = 1, \ldots, i \times j\right\} \end{cases}. \tag{16}$$

4.3 Detail Layers Fusion Rule Based on Boundary Handling CSR

The detail layer reflects features such as details and textures of the image, for each detail layer I_d^k, its sparse coefficient maps C_k^m, $m \in \{1, \cdots, M\}$ are obtained by solving the CSR model with the method in [32]:

$$\arg\min_{\{\mathbf{C}_{k,m}\}} \frac{1}{2}||\sum_{m=1}^{M} \mathbf{d}_m \times \mathbf{C}_{k,m} - \mathbf{I}_k^d||_2^2 + \lambda \sum_{m=1}^{M} ||\mathbf{C}_{k,m}||_1. \tag{17}$$

Since this method is solved in the DFT domain, it implicitly imposing periodic boundary conditions. One possible solution is to mask the boundary of the fill estimation through a boundary spatial mask [33].

$$\begin{aligned} &\arg\min_{C,y_0,y_1} \tfrac{1}{2}||Wy_1 - \mathrm{I}_k^d||_2^2 + \lambda||y_0||_1 \\ &\quad s.t. \begin{pmatrix} \mathrm{C} \\ \mathrm{DC} \end{pmatrix} - \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = 0 \end{aligned}. \tag{18}$$

Defining

$$A = \begin{pmatrix} \mathrm{I} \\ \mathrm{D} \end{pmatrix} \; y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \; u = \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}. \tag{19}$$

The corresponding ADMM iterations are

$$\mathrm{C}^{(j+1)} = \arg\min_{\mathrm{C}} \frac{\rho}{2}||AC - y^{(j)} + u^{(j)}||_2^2, \tag{20}$$

$$y^{(j+1)} = \arg\min_{y} \frac{1}{2}||Wy_1 - \mathrm{I}_k^d||_2^2 + \lambda||y_0||_1 + \frac{\rho}{2}||AC^{(j+1)} - y + u^{(j)}||_2^2, \tag{21}$$

$$u^{(j+1)} = u^{(j)} + AC^{(j+1)} - y^{(j+1)}. \tag{22}$$

The functional minimized in (20) can be expanded as

$$\begin{aligned} &\frac{\rho}{2}||AC - y + u||_2^2 = \frac{\rho}{2}||\begin{pmatrix} \mathrm{C} \\ \mathrm{DC} \end{pmatrix} - \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} + \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}||_2^2 \\ &= \frac{\rho}{2}||\mathrm{DC} - \big(y_1 - u_1\big)||_2^2 + \frac{\rho}{2}||\mathrm{C} - \big(y_0 - u_0\big)||_2^2. \end{aligned} \tag{23}$$

The functional minimized in (21) can be expanded as

$$\frac{1}{2}||Wy_1 - \mathrm{I}_k^d||_2^2 + \lambda||y_0||_1 + \frac{\rho}{2}||y_1 - \big(\mathrm{DC} + u_1\big)||_2^2 + \frac{\rho}{2}||y_0 - \big(\mathrm{C} + u_0\big)||_2^2. \tag{24}$$

Since the y_0 and y_1 components of y are decoupled, minimization with respect to y can be achieved by the independent minimizations.

$$y_0^{(j+1)} = \arg\min_{y_0} \lambda||y_0||_1 + \frac{\rho}{2}||y_0 - \big(\mathrm{C} + u_0\big)||_2^2, \tag{25}$$

$$y_1^{(j+1)} = \arg\min_{y_1} \frac{1}{2}||Wy_1 - \mathrm{I}_k^d||_2^2 + \frac{\rho}{2}||y_1 - \big(\mathrm{DC} + u_1\big)||_2^2, \tag{26}$$

The solution to (25) is just soft thresholding, and the solution to (26) is given by

$$\big(W^T W + \rho \mathrm{I}\big)y_1 = W^T \mathrm{I}_\mathrm{k}^d + \rho\big(\mathrm{DC} + u_1\big). \tag{27}$$

Then, we obtain $\mathbf{C}_{k,1:M}(x,y)$, $\mathbf{C}_{k,1:M}(x,y)$ is a vector with M-dimensional, representing the position of $\mathbf{C}_{k,m}$ in the spatial domain (x,y), and $||\mathbf{C}_{k,1:M}(x,y)||_1$ indicates the activity level of the source image. Thus, the activity level map $\mathbf{A}_k(x,y)$ can be generated by

$$\mathbf{A}_k(x,y) = ||\mathbf{C}_{k,1:M}(x,y)||_1. \tag{28}$$

To enhance the noise robustness of the algorithm, a window-based averaging strategy is applied in $\mathbf{A}_k(x,y)$ to achieve the final activity level map:

$$\bar{A}_k(x,y) = \frac{\sum_{p=-r}^{r}\sum_{q=-r}^{r} A_k(x+p, y+q)}{(2r+1)^2} \tag{29}$$

where r is the size of the window, and $r = 9$. Then the fused coefficient maps are obtained through the "choose-max" strategy:

$$\mathbf{C}_{f,1:M}(x,y) = \mathbf{C}_{k^*,1:M}(x,y), \; k^* = \arg\max_k \left(\overline{\mathbf{A}}_k(x,y)\right). \tag{30}$$

Accordingly, the detail layer can be fused as

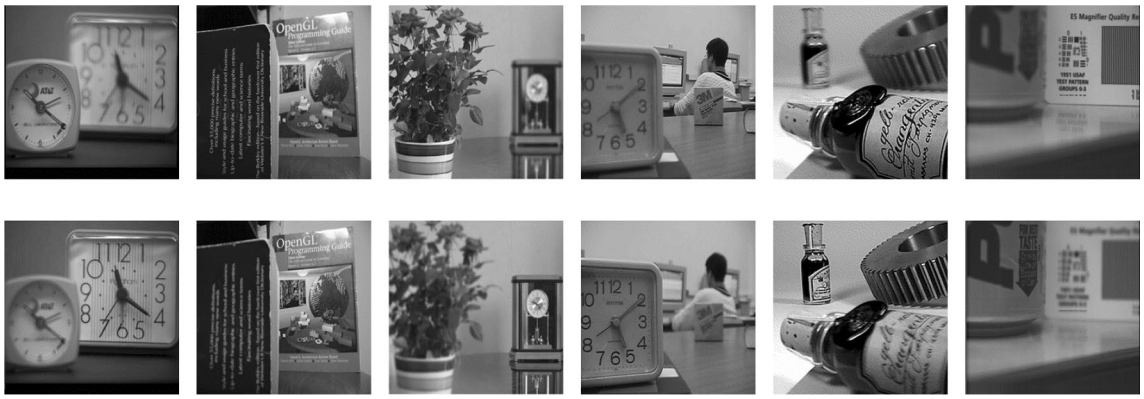

Fig. 4 A portion of multi-focus images used in our experiments

$$\mathbf{I}_f^d = \sum_{m=1}^{M} \mathbf{d}_m \times \mathbf{C}_{f,m}. \tag{31}$$

Finally, through the coefficient $\mathbf{I}_f^b$ of the base layer and $\mathbf{I}_f^d$ of the detail layer, the fused image can be reconstructed as (32)

$$\mathbf{I}_f = \mathbf{I}_f^d + \mathbf{I}_f^b. \tag{32}$$

5 Experimental Results and Comparison

To show the performance of the method, we use the widely used 14-pair multi-focus images as the testing dataset, a portion is as shown in Fig. 4.

5.1 Measurement

The following four objective metrics are selected to verify the performance of our algorithm:

The phase congruency based fusion metric (Q_P) [34], the mutual information-based fusion metric (MI) [35], the edge similarity measure ($Q^{AB/F}$) [36], and the fusion measure of image distortion (Q_0) [37]. The larger the value, the better the performance.

5.2 Comparison Algorithm and Parameter Settings

We select the following five fusion algorithms ASR [38], NSCT [39], CS-MCA [40], NSST-PAPCNN [41], DWT-SR [10],which are more advanced in recent years, to conduct comparative experiments.

The parameter settings of the proposed fusion method are: convolution sparse representation: $\lambda = 0.01$, and a dictionary filter of $12 \times 12 \times 36$ is selected.

We use the commonly used multi-focus image database for experiments, all images are grayscale images, and their size is 256×256.

5.3 Comparative Analysis of Experimental Results

For the testing images, Gaussian noise ($\mu = 0, \sigma = 0.0005$, $0.001, 0.005, 0.01$), salt and pepper noise (with density of 0.01 and 0.02) and Poisson noise are added, respectively.

In Fig. 5, the fusion results obtained by NSST-PAPCNN and ASR are relatively low, lose a lot of detail and cannot provide clear visual effect. The detailed information above the number "11" on the right dial is relatively fuzzy, and the black dot on the right is not clear. By CS-MCA, the collar of the character is relatively blurred, and there are artifacts in the upper left corner of the clock dial. By DWT-SR, the scale on the upper right of the dial is not clear, and there are artifacts on the face. By NSCT, the letter part of the clock has no detailed information and the border of the human head is blurred. And the above fusion results also have some spatial discontinuities. In contrast, the fusion image generated by our algorithm has higher definition and the details are clearly distinguishable.

In Fig. 6, the NSST-PAPCNN and ASR method scenes are relatively blurry, and the fusion result has block effects. The detail information on the lower left part of

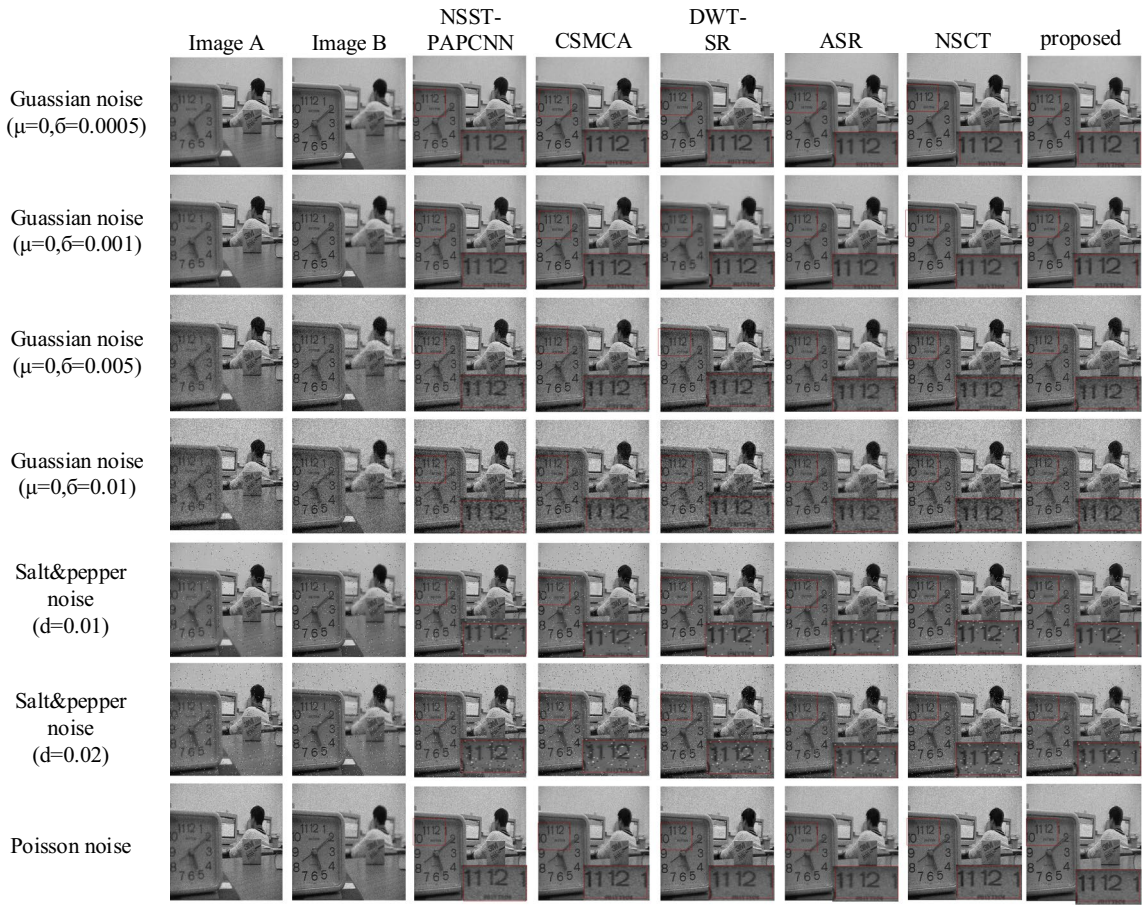

Fig. 5 The first pair of image fusion results

the large bottle body is not obvious and there are artifacts on the inner edge of the gear, indicating that there are misclassified pixels. The fusion results of NSCT and DWT-SR all have spatial discontinuities, which affect the overall fusion effect. The pendant on the large bottle of the DWT-SR method has artifacts and the upper edge is not clear.

In the NSCT method, the edge of the table in the lower left corner is blurred and has artifacts. In the fusion results obtained by the CS-MCA method and the method in this paper, the scene detail information is retained better and the overall effect is better. However, in CS-MCA, there are still some details that are not completely extracted, the upper right part of the gear and the pendant on the large bottle are both blurred. In contrast, the fusion result of our method has higher contrast and is more clearly distinguishable.

To further illustrate the robustness of the algorithm, the objective evaluation indicators are analyzed. Statistical results in Table 1 are the average values of all methods in the same dataset and among the maximum is bolded.

The data in Table 1 show that DWT-SR and ASR are low in Q_P and $Q^{AB/F}$, indicating that they cannot well retain the structural information and edge information of the source image. And the NSST-PAPCNN, DWT-SR, and NSCT algorithms are also low in Q_0. The above results indicate that there are defects in the reservation of its details, which is consistent with subjective vision. The NSST-PAPCNN, DWT-SR, and NSCT algorithms

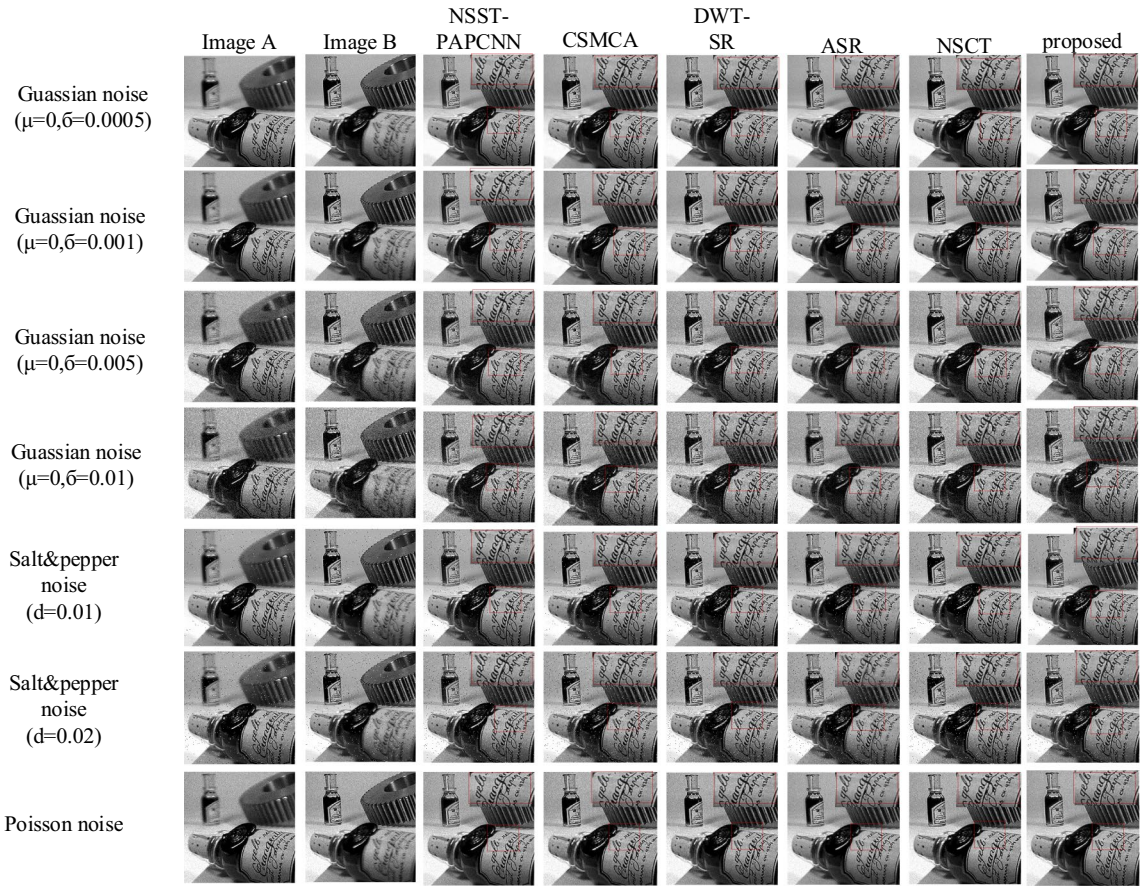

Fig. 6 The second pair of image fusion results

are also low in MI, which indicates that the fused images retain less the original information of the source image, resulting in poor fusion effect. Our method performs better than other methods in distortion and detail preservation. Because its detail structure is clearer, and more rich detail information is preserved. It can be seen from Table 1 that except for salt and pepper noise, our method has higher value for each index, and the quality of fusion is better.

6 Conclusion

In this study, an improved pigeon-inspired optimization bionic algorithm is proposed for multi-focus noisy image fusion. First, by two-scale image decomposition, the source image containing noise is decomposed into base layer and detail layer. The detail layer is fused by boundary handling for convolution sparse representation to complete the fusion of the detail layer coefficients. Based on the convolutional sparse representation for image fusion greatly reduces the time complexity by solving in the frequency domain, and at the same time avoids possible boundary artifacts through effective boundary handling. The base layer image is fused with optimized weight obtained through pigeon-inspired optimization, and the range of the optimized weight is determined by complementary edges, which can effectively strengthen target information and edge information. Finally, the final fusion image is generated by the superposing the base layer and the detail layer. The experimental results show that our method has better performance in fusion effect and noise robustness.

Table 1 Comparison of objective evaluation indexes

Noise	–	Indicators	NSST-PAPCNN	CS-MCA	DWT-SR	ASR	NSCT	Proposed
Gaussian noise	$\sigma=0.0005$	Q_0	0.8166 ± 0.0173	0.8178 ± 0.0125	0.8150 ± 0.0159	0.8172 ± 0.0164	0.8173 ± 0.0139	**0.8182** ± 0.0114
		MI	3.4565 ± 0.6155	3.6595 ± 0.2662	3.2613 ± 0.3489	3.5876 ± 0.5759	3.5908 ± 0.1684	**3.6349** ± 0.1244
		$Q^{AB/F}$	0.6161 ± 0.0246	0.6507 ± 0.0505	0.6080 ± 0.0476	0.6342 ± 0.0488	0.6494 ± 0.0868	**0.6510** ± 0.0196
		Q_P	0.7464 ± 0.0802	**0.8133** ± 0.0311	0.7338 ± 0.0592	0.7809 ± 0.0936	0.8042 ± 0.0815	0.7843 ± 0.0104
	$\sigma=0.001$	Q_0	0.8145 ± 0.0168	0.8152 ± 0.0132	0.8127 ± 0.0182	0.8146 ± 0.0161	0.8146 ± 0.0109	**0.8155** ± 0.0107
		MI	3.2091 ± 0.2463	3.3624 ± 0.4029	2.9637 ± 0.3410	3.2906 ± 0.8387	3.2721 ± 0.3098	**3.3747** ± 0.1380
		$Q^{AB/F}$	0.5774 ± 0.0175	0.6165 ± 0.0966	0.5752 ± 0.0920	0.5874 ± 0.0660	0.6113 ± 0.0145	**0.6192** ± 0.0114
		Q_P	0.7169 ± 0.0451	**0.7780** ± 0.0418	0.7038 ± 0.0367	0.7413 ± 0.0798	0.7607 ± 0.0380	0.7775 ± 0.0395
	$\sigma=0.005$	Q_0	0.8095 ± 0.0117	0.8096 ± 0.0111	0.8077 ± 0.0180	0.8095 ± 0.0113	0.8091 ± 0.0126	**0.8101** ± 0.0109
		MI	2.5160 ± 0.5374	2.5767 ± 0.8081	2.1734 ± 0.6230	2.5735 ± 0.6953	2.4831 ± 0.7493	**2.6366** ± 0.2919
		$Q^{AB/F}$	0.4631 ± 0.0306	0.5072 ± 0.0416	0.4337 ± 0.0264	0.4568 ± 0.0912	0.4954 ± 0.0318	**0.5203** ± 0.0270
		Q_P	0.6126 ± 0.0839	0.6822 ± 0.0770	0.5732 ± 0.0538	0.6135 ± 0.0576	0.6778 ± 0.0922	**0.6945** ± 0.0351
	$\sigma=0.01$	Q_0	0.8078 ± 0.0131	0.8077 ± 0.0172	0.8062 ± 0.0115	0.8079 ± 0.0163	0.8073 ± 0.0122	**0.8083** ± 0.0112
		MI	2.2353 ± 0.4767	2.2499 ± 0.4442	1.8530 ± 0.5823	2.2837 ± 0.7408	2.1624 ± 0.5169	**2.3456** ± 0.1229
		$Q^{AB/F}$	0.4984 ± 0.0797	0.4526 ± 0.0337	0.3763 ± 0.0562	0.4046 ± 0.0950	0.4439 ± 0.0464	**0.4993** ± 0.0626
		Q_P	0.6419 ± 0.0492	0.6079 ± 0.0249	0.5030 ± 0.0836	0.5478 ± 0.0462	0.6165 ± 0.0542	**0.6452** ± 0.0231
Salt and pepper noise	$d=0.01$	Q_0	0.8223 ± 0.0149	0.8242 ± 0.0170	0.8204 ± 0.0157	**0.8276** ± 0.0182	0.8251 ± 0.0167	0.8242 ± 0.0132
		MI	3.7950 ± 0.0187	4.0639 ± 0.0718	3.5202 ± 0.0502	**4.4484** ± 0.0642	4.1567 ± 0.0843	4.1814 ± 0.0139
		$Q^{AB/F}$	0.6583 ± 0.0222	0.6683 ± 0.0168	0.6274 ± 0.0540	0.6957 ± 0.0219	**0.7034** ± 0.0252	0.6673 ± 0.0105
		Q_P	0.7019 ± 0.0265	**0.7348** ± 0.0376	0.6344 ± 0.0585	0.7331 ± 0.0882	0.7309 ± 0.0404	0.7274 ± 0.0232
	$d=0.02$	Q_0	0.8199 ± 0.0163	0.8214 ± 0.0171	0.8173 ± 0.0140	0.7671 ± 0.0186	0.8221 ± 0.0149	**0.8225** ± 0.0132
		MI	3.4970 ± 0.3806	3.7191 ± 0.7754	3.0654 ± 0.8239	**4.3275** ± 0.2774	3.8078 ± 0.8431	3.8136 ± 0.2821
		$Q^{AB/F}$	0.6246 ± 0.0948	0.6188 ± 0.0684	0.5645 ± 0.0432	0.6544 ± 0.0558	**0.6775** ± 0.0689	0.6209 ± 0.0154
		Q_P	0.6434 ± 0.0997	**0.6677** ± 0.0460	0.5522 ± 0.0622	0.6664 ± 0.0211	0.6601 ± 0.0910	0.6647 ± 0.0114
Poisson noise	–	Q_0	0.7994 ± 0.0131	0.8145 ± 0.0112	0.8115 ± 0.0164	0.8122 ± 0.0153	0.8131 ± 0.0123	**0.8146** ± 0.0032
		MI	3.0399 ± 0.3065	3.1765 ± 0.3373	3.0678 ± 0.8431	3.1214 ± 0.7369	3.0929 ± 0.7232	**3.2365** ± 0.2853
		$Q^{AB/F}$	0.5455 ± 0.0961	0.5892 ± 0.0759	0.5320 ± 0.0663	0.5504 ± 0.0560	0.5819 ± 0.0720	**0.5978** ± 0.0618
		Q_P	0.6852 ± 0.0170	0.7523 ± 0.0334	0.6620 ± 0.0595	0.6986 ± 0.0266	0.7464 ± 0.0432	**0.7528** ± 0.0265

Acknowledgements This work is supported in part by National Key Research and Development Program of China (2018YFB0804202, 2018YFB0804203); Regional Joint Fund of NSFC (U19A2057); National Natural Science Foundation of China (61876070); Jilin Province Science and Technology Development Plan Project(20190303134SF).

Declaration

Conflict of interest The Authors declare that they have no conflict of interest.

References

1. Xu, Y. D., Sun, B. B., Yan, X. A., Hu, J. Z., & Chen, M. L. (2020). Multi-focus image fusion using learning based matting with sum of the gaussian-based modified laplacian. *Digital Signal Processing, 106*, 102821.
2. Ramlal, S. D., Sachdeva, J., Ahuja, C., & Khandelwal, N. (2019). An improved multimodal medical image fusion scheme based on hybrid combination of nonsubsampled contourlet transform and stationary wavelet transform. *International Journal of Imaging Systems and Technology, 29*, 146–160.
3. Fu, J., Li, W. S., Du, J., & Xiao, B. (2020). Multimodal medical image fusion via laplacian pyramid and convolutional neural network reconstruction with local gradient energy strategy. *Computers in Biology and Medicine, 126*, 104048.
4. Wang, K. P., Zheng, M. Y., Wei, H. Y., & Qi, G. Q. (2020). Multi-Modality medical image fusion using convolutional neural network and contrast pyramid. *Sensors, 20*, 2169.
5. Gai, D., Shen, X. J., Chen, H. P., & Su, P. X. (2020). Multi-focus image fusion method based on two stage of convolutional neural network. *Signal Processing, 176*, 107681.
6. Guo, L. Q., Cao, X., & Liu, L. (2020). Dual-tree biquaternion wavelet transform and its application to color image fusion. *Signal Processing, 171*, 107513.
7. Zhang, L. X., Zeng, G. P., & Wei, J. J. (2019). Adaptive region-segmentation multi-focus image fusion based on differential

evolution. *International Journal of Pattern Recognition and Artificial Intelligence, 33*, 1954010.
8. Ma, X. L., Hu, S. H., Liu, S. Q., Fang, J., & Xu, S. W. (2019). Multi-focus image fusion based on joint sparse representation and optimum theory. *Signal Processing-image Communication, 78*, 125–134.
9. Zhang, Q., Li, G. H., Cao, Y. F., & Han, J. G. (2020). Multi-focus image fusion based on non-negative sparse representation and patch-level consistency rectification. *Pattern Recognition, 104*, 107325.
10. Liu, Y., Liu, S. P., & Wang, Z. F. (2015). A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion, 24*, 147–164.
11. Gai, D., Shen, X. J., Cheng, H., & Cheng, H. P. (2019). Medical image fusion via PCNN based on edge preservation and improved sparse representation in NSST domain. *IEEE Access, 7*, 85413–85429.
12. Zhang, Q., Wang, F., Luo, Y. J., & Han, J. G. (2020). Exploring a unified low rank representation for multi-focus image fusion. *Pattern Recognition, 113*, 107752.
13. Maqsood, S., & Javed, U. (2020). Multi-modal medical image fusion based on two-scale image decomposition and sparse representation. *Biomedical Signal Processing and Control, 57*, 101810.
14. Liu, Y., Chen, X., Peng, H., & Wang, Z. F. (2017). Multi-focus image fusion with a deep convolutional neural network. *Information Fusion, 36*, 191–207.
15. Amin-Naji, M., Aghagolzadeh, A., & Ezoji, M. (2019). Ensemble of CNN for multi-focus image fusion. *Information Fusion, 51*, 201–214.
16. Tang, H., Xiao, B., Li, W. S., & Wang, G. Y. (2018). Pixel convolutional neural network for multi-focus image fusion. *Information Sciences, 433*, 125–141.
17. Bebis, G., Gyaourova, A., Singh, S., & Pavlidis, I. (2006). Face recognition by fusing thermal infrared and visible imagery. *Image and Vision Computing, 24*, 727–742.
18. Raghavendra, R., Dorizzi, B., Rao, A., & Kumar, G. H. (2011). Particle swarm optimization based fusion of near infrared and visible images for improved face verification. *Pattern Recognition, 44*, 401–411.
19. Zhao, J. F., Cui, G. M., Gong, X. L., Zang, Y., Tao, S. Y., & Wang, D. D. (2017). Fusion of visible and infrared images using global entropy and gradient constrained regularization. *Infrared Physics and Technology, 81*, 201–209.
20. Lawrence, T., Zhang, L., Lim, C. P., & Phillips, E. J. (2021). Particle swarm optimization for automatically evolving convolutional neural networks for image classification. *IEEE Access, 9*, 14369–14386.
21. Zhang, Y. J., Chen, H. J., Yang, L., Liu, K., Li, F., Bai, C., Wu, H. T., & Yao, J. F. (2020). A proportional genetic algorithm for image reconstruction of static electrical impedance tomography. *IEEE Sensors Journal, 20*, 15026–15033.
22. Duan, H. B., & Qiao, P. X. (2014). Pigeon-inspired optimization: A new swarm intelligence optimizer for air robot path planning. *International Journal of Intelligent Computing and Cybernetics, 7*, 24–37.
23. Sandhya, G., Kande, G. B., & Savithri, T. S. (2020). Segmentation of magnetic resonance brain images using the advanced ant colony optimization technique. *Journal of Biomimetics Biomaterials and Biomedical Engineering, 44*, 37–49.
24. Yan, X. H., Zhang, Y. J., Zhang, D. J., & Hou, N. (2020). Multimodal image registration using histogram of oriented gradient distance and data-driven grey wolf optimizer. *Neurocomputing, 392*, 108–120.
25. Liu, Y., Mu, C. H., Kou, W. D., & Liu, J. (2015). Modified particle swarm optimization-based multilevel thresholding for image segmentation. *Soft Computing, 19*, 1311–1327.
26. Mlakar, U., Potocnik, B., & Brest, J. (2016). A hybrid differential evolution for optimal multilevel image thresholding. *Expert Systems with Applications, 65*, 221–232.
27. Gai, D., Shen, X. J., Chen, H. P., Xie, Z. Y., & Su, P. X. (2020). Medical image fusion using the PCNN based on IQPSO in NSST domain. *IET Image Processing, 14*, 1870–1880.
28. Wang, Y., Zhang, G. B., & Zhang, X. F. (2019). Multilevel image thresholding using tsallis entropy and cooperative pigeon-inspired optimization bionic algorithm. *Journal of Bionic Engineering*, *16*, 954–964.
29. Loza, A., Bull, D., Canagarajh, N., & Achim, A. (2010). Non-Gaussian model-based fusion of noisy images in the wavelet domain. *Computer Vision and Image Understanding, 114*, 54–65.
30. Yin, S. F., Cao, L. C., Ling, Y. S., & Jin, G. F. (2011). Fusion of noisy infrared and visible images based on anisotropic bivariate shrinkage. *Infrared Physics and Technology, 54*, 13–20.
31. Bo, Z., & Duan, H. B. (2017). Three-dimensional path planning for uninhabited combat aerial vehicle based on predator-prey pigeon-inspired optimization in dynamic environment. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 14*, 97–107.
32. Wohlberg, B. (2016). Boundary handling for convolutional sparse representations. *Proceedings of the IEEE International Conference on Image Processing*, Phoenix, United States, 1833–1837
33. Heide, F., Heidrich, W., Wetzstein, G. (2015). Fast and flexible convolutional sparse coding. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, United States, 5135–5143
34. Zhao, J. Y., Laganiere, R., & Liu, Z. (2007). Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement. *International Journal of Innovative Computing Information and Control, 3*, 1433–1447.
35. Haghighat, M. B. A., Aghagolzadeh, A., & Seydarabi, H. (2011). A non-reference image fusion metric based on mutual information of image features. *Computers and Electrical Engineering, 37*, 744–756.
36. Xydeas, C. S., & Petrovic, V. (2000). Objective image fusion performance measure. *Electronics Letters, 36*, 308–309.
37. Wang, Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Processing Letters., 9*, 81–84.
38. Liu, Y., & Wang, Z. F. (2015). Simultaneous image fusion and denoising with adaptive sparse representation. *IET Image Processing, 9*, 347–357.
39. Zhang, Q., & Guo, B. L. (2009). Multifocus image fusion using the nonsubsampled contourlet transform. *Signal processing*, *89*, 1334–1346.
40. Liu, Y., Chen, X., Ward, R. K., & Wang, Z. J. (2019). Medical image fusion via convolutional sparsity based morphological component analysis. *IEEE Signal Processing Letters, 26*, 485–489.
41. Yin, M., Liu, X. N., Liu, Y., & Chen, X. (2019). Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain. *Instrumentation and Measurement IEEE Transactions, 68*, 49–64.