Future Generation Computer Systems 116 (2021) 200–208

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Multiple improved residual networks for medical image super-resolution

Defu Qiu, Lixin Zheng *, Jianqing Zhu, Detian Huang

Engineering Research Centre in Industrial Intellectual Techniques and Systems of Fujian Providence College of Engineering, Huaqiao University, Quanzhou, China

ARTICLE INFO

Article history:
Received 23 March 2020
Received in revised form 25 October 2020
Accepted 3 November 2020
Available online 5 November 2020

Keywords:
Intelligent medical image processing
Super-resolution reconstruction
Deep learning
Improved residual network
Residual block

ABSTRACT

The rapid development of deep learning has resulted in great breakthroughs in image super-resolution reconstruction technology in medical imaging modalities. The application of artificial intelligence to medical image processing has been the focus of scholars both domestically and internationally in recent years. Due to the fast super-resolution convolutional neural network (FSRCNN) algorithm has fewer convolutional layers and lacks the correlation between the feature information of adjacent convolutional layers, it is difficult to be used to extract deep information of an image, and the super-resolution rate of the image reconstruction effect is not good. To solve this problem, we propose the multiple improved residual network (MIRN) super-resolution reconstruction method. First, MIRN designs the residual blocks connected by multi-level skips to build multiple improved residual block (MIRB) modules. A deep residual network with multi-level skip connection is used to solve the lack of correlation between the characteristic information of adjacent convolutional layers. Then, the stochastic gradient descent method (SGD) is used to train a deep residual network connected by multi-level jumpers with an adjustable learning rate strategy to obtain a super-resolution reconstruction model of the network. Finally, the low-resolution image is input in the MIRN super-resolution reconstruction model, and the residual block obtains the predicted residual eigenvalues and then combines the residual image and the low-resolution image into a high-resolution image. Most quantitative and qualitative evaluations on benchmark datasets demonstrate that the proposed model can better reconstruct the details and textures of images and avoid the over-smoothing of medical images after iteration, and the performance of the proposed algorithm is revealed to be better than that of existing state-of-the-art methods.

1. Introduction

During medical imaging, transmission, conversion, storage, etc., images obtained by manual techniques are usually degraded due to the influences of adverse factors, such as the medical imaging equipment and external environment; degradation refers to the loss of useful information in images. Image super-resolution (SR) technology reconstructs low-resolution (LR) images into high-resolution (HR) images, and is widely applied in the fields of medical imaging, satellite imaging, facial recognition, and surveillance [1–5], among others. At present, image super-resolution methods can be categorized into three types, as follows.

Interpolation-based methods [6] include nearest-neighbor interpolation (bilinear) [7] and bicubic interpolation (bicubic) [8]. The nearest-neighbor interpolation method [7] is prone to suffer from the block effect and the jagged-edge effect because the interpolation rule is too simple. The bicubic interpolation method [8] is slightly more complex than the nearest-neighbor interpolation method, but the smoothness of the edges of images is improved. The common advantage of these two methods is that the algorithms are simple and easy to implement, and the calculation speed is fast. However, only the local pixel points are used to calculate the pixel values of the interpolation points; therefore, the high-frequency details of an image cannot be repaired, and the resulting super-resolution image has limited sharpness.

Reconstruction-based methods [9] can be further divided into two categories, namely the frequency-domain and space-domain methods. Frequency-domain methods usually improve the resolution of an image by eliminating the phenomenon of spectral aliasing, and the airspace method has a strong ability to constrain airspace a priori. The representative reconstruction-based methods are the Projection onto Convex Sets method (POCS) [10–12],

* Correspondence to: Huaqiao University, Chenghua North Road, Fengze District, Quanzhou, Fujian, 362021, China.
E-mail addresses: defu.qiu@hqu.edu.cn (D. Qiu), zlx@hqu.edu.cn (L. Zheng), jqzhu@hqu.edu.cn (J. Zhu), huangdetian@hqu.edu.cn (D. Huang).

https://doi.org/10.1016/j.future.2020.11.001
0167-739X/

the Iterative Back-Projection method (IBP) [13,14], the Application of Maximum a Posteriori method (MAP) [15–17], and the hybrid MAP/POCS method [18–20].

Learning-based methods [21] have been the focus of research in recent years. A learning model is obtained by learning a large number of HR images, and then, when processing LR medical images, the prior knowledge in the model is fully combined with the LR image to obtain a reconstructed image with rich high-frequency information. Of the existing learning-based methods, the method based on neighbor embeddings [22–24] and the method of sparse representation [25] can achieve better super-resolution performance. The single-image super-resolution (SISR) method based on neighbor embeddings uses the geometric similarity between HR and LR image blocks to understand their regression or mapping relationship. In the SISR method based on sparse representation, the LR image block and the corresponding HR image block are respectively represented as corresponding dictionaries by using the sparse coding theory, and the mapping relationship between dictionaries is established by machine learning methods. The mapping relationship generates an HR image of a given single-frame image. The establishment of models and the selection of training sets are the core problems of super-resolution reconstruction algorithms based on learning. Yang et al. [26,27] assumed that the input LR image blocks can be expressed sparsely and linearly by the training sample set. Wang et al. [28] proposed the learning of a more effective over-complete dictionary from the training sample set, which is suitable for fast multi-level, super-resolution image reconstruction tasks.

In recent years, the use of deep learning has achieved remarkable research results in the field of image processing. In many tasks, features obtained by deep learning have been proven to have stronger representation capabilities than features constructed by traditional methods [29]. Dong et al. [30] proposed the super-resolution using a convolutional neural network (SRCNN) algorithm, and applied it to the field of image super-resolution reconstruction. This network has a simple structure and an excellent super-resolution reconstruction effect, but its convolution kernel is large, and traditional bicubic interpolation is used for upsampling; these structural characteristics substantially affect the network's operating speed. Subsequently, Dong et al. [31] proposed an accelerated super-resolution convolutional neural network (FSRCNN) based on SRCNN. The structure of FSRCNN is deeper than that of SRCNN, and it is replaced by deconvolution bicubic interpolation. FSRCNN is significantly faster than SRCNN, and the super-resolution effect of the image is also improved. While FSRCNN has succeeded in the super-resolution reconstruction of images, its number of convolutional layers is small, and the feature information of adjacent convolutional layers lacks correlation; therefore, it is difficult to extract deep information of the image, resulting in poor super-resolution reconstruction.

Based on SRCNN, a super-resolution reconstruction method based on a very deep network (VDSR) [32] was borrowed from the VGG [33] network structure. This method takes an image preprocessed by interpolation and increases the network depth to 20 layers; this network structure is different from that of image classification. The super-resolution reconstruction method based on the very deep network establishes a mapping relationship model between LR and HR images [34]. However, the increase in the network depth also increases the amount of calculation, it is difficult to find a satisfactory threshold during training, and the method is time consuming. Kim et al. [35] proposed a deeply-recursive convolutional network (DRCN) for image super-resolution. The DRCN algorithm applies a recursive neural network structure to super-resolution reconstruction, and the use of residual learning deepens the number of network layers, increases the receptive field of the network, and improves reconstruction performance. Tai et al. [36] proposed a method of image super-resolution via a deep recursive residual network (DRRN) that uses both local and global residual learning. Each residual unit in this model has a common input, and each contains 2 convolutional layers. In each recursive block in DRRN, the convolutional layer parameters at the same position in the residual unit are shared.

Although the application of the deep neural network structure can obtain more image information and achieve better results in super-resolution reconstruction technology, to extract the deep-level information of the image, the deep neural network is constructed by simply stacking convolutional layers. This makes it difficult to train deep convolutional neural networks, and problems such as gradient disappearance and gradient explosion can occur. Additionally, the convolutional layers of the deep neural network only perform convolution calculations on the feature vectors of the previous convolutional layer, and do not use the correlation between the feature information of the two adjacent convolutional layers; thus, the context information of the image region is reconstructed with a lack of relevance.

2. Methods

In this paper, a multiple improved residual network (MIRN) method for super-resolution is proposed. First, multiple improved residual blocks (MIRBs) are designed. The MIRN consists of eight MIRBs and upsampling modules. The MIRB extracts LR image features with three improved residual blocks (IRBs); it is composed of 3 sub-residual blocks and β skip connections, each sub-residual block is composed of a convolution layer and β skip connections, and the upsampling module is composed of a sub-pixel convolution layer [37]. Then, the stochastic gradient descent (SGD) method is used to train the MIRN with an adjustable learning rate strategy to obtain a super-resolution reconstruction model of the network. Finally, the proposed model is applied to the reconstruction of LR images of a test set; it is found that the reconstructed images have not only better visual effects, but also higher peak signal-to-noise ratios (PSNRs) [38] and average structural similarity (SSIM) [39].

2.1. Residual learning

He et al. [40] first proposed a residual network in 2016. The residual network solves the problem of model degeneration when constructing a deep neural network, and is composed of multiple residual blocks connected end-to-end. The structural diagram of the residual block is presented in Fig. 1.

The residual block function is defined as:

$$y = F(x, \{W_i\}) + x, \tag{1}$$

where y is the output vector of the residual block, x is the input vector of the residual block, W_i represents the weight of the layer i, $F(x)$ represents the output of the residual block before the activation function of the second layer, and the function $F(x, \{W_i\})$ is the residual mapping to be learned. The residual block presented in Fig. 1 includes two convolutional layers; thus, the residual function is $F = W_2(W_1(x))$, in which the bias is omitted to simplify the representation. The input vector x is connected to the residual function $F(x)$ through a skip connection [41] to obtain the residual block function y as follows:

$$y = F(x, \{W_i\}) + x. \tag{2}$$

In the residual block, the sizes of the input vector x and the residual function $F(x)$ must be equal.

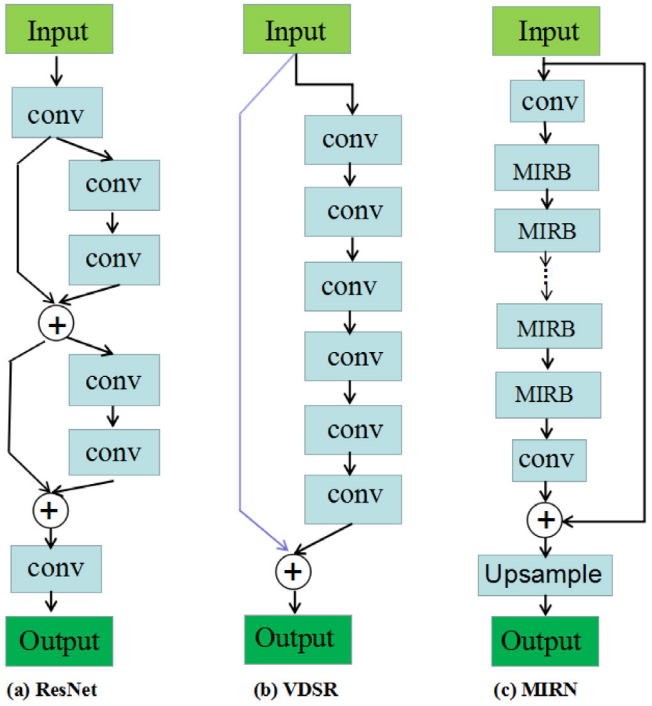

Fig. 1. Structural comparison of MIRN, ResNet [40], and VDSR [32].

Assuming that the convolutional layer behind the deep neural network is an identity-mapping function represented by $H(x)$, then the training of the deep neural network is simplified to learn an identity-mapping function. If it is difficult to directly fit the convolutional layer to an identity-mapping function $H(x) = x$, but the deep neural network is designed as $H(x) = F(x) + x$, then the training of the deep neural network can be converted into the learning of residual function $F(x) = H(x) - x$; as long as $F(x) = 0$, it is equivalent to fitting the identity-mapping function $H(x) = x$. Under the same calculation conditions, it is easier for the previous convolutional layer to fit the residual function $F(x) = H(x) - x$ than to fit the identity function $H(x) = x$.

2.2. Multiple improved residual blocks

The convolutional layers of the deep neural network only perform convolution calculations on the feature vectors of the previous convolutional layer, and do not use the correlation between the feature information of two adjacent convolutional layers; this results in a lack of correlation of the context information of the image region. To solve this problem, the MIRB is designed, the structure of which is shown in Fig. 2.

Because the smart-network design is a concatenated convolutional network, the efficient sub-pixel convolution layer acts as the output layer of the super-resolution network. These network sections are described in detail in the subsequent section.

In Fig. 2, a convolution layer and λ skip connections are connected to form a sub-residual block, and 2 sub-residual blocks and β skip connections are connected to form an IRB. Assuming that the input of the residual block connected by the multi-level skip connection is x, the output obtained by the first sub-residual block is y_1, the output obtained by the second sub-residual block is y_2, and the output obtained by the multiple improved residual modules is y_3, then the following is output:

$$y_1 = W_1(x) + \lambda x, \tag{3}$$

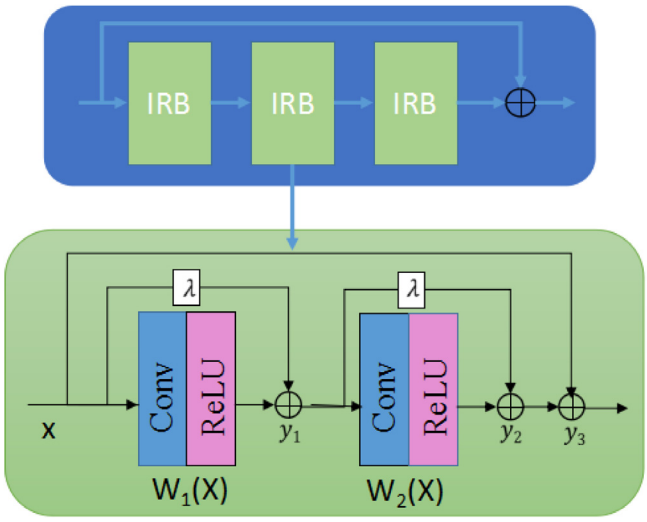

Fig. 2. Structural framework of the MIRB.

$$\begin{aligned} y_2 &= W_2(y_1) + \lambda y_1 \\ &= W_2(W_1(x) + \lambda x) + \lambda W_1(x) + \lambda^2 x, \end{aligned} \tag{4}$$

$$\begin{aligned} y_3 &= y_2 + \beta x \\ &= W_2(W_1(x) + \lambda x) + \lambda W_1(x) + (\lambda^2 + \beta)x. \end{aligned} \tag{5}$$

From the output y_3, it can be determined that when the input x is connected through a multi-level skip connection, not only the vector $W_2(W_{\ 1}(x))$, but also the output value λy_1, of the first sub-residual block is obtained. Therefore, the residual block connected by the multi-level skip connection can extract the relevant information of the feature vector of the convolution layer inside the residual block.

The structural diagram of the residual blocks of two adjacent multi-level skip connections is shown in Fig. 3, in which the output y_4 of the third sub-residual block is:

$$\begin{aligned} y_4 &= W_3(y_3) + \lambda y_3 \\ &= W_3(y_2 + \beta x) + \lambda(W_2(y_1)) + \lambda^2 W_1(x). \\ &+ (\lambda^3 + \lambda\beta)x \end{aligned} \tag{6}$$

The output y_5 of the fourth sub-residual block is:

$$\begin{aligned} y_5 &= W_4(y_4) + \lambda y_4 \\ &= W_4(y_4) + \lambda(W_3(y_2 + \beta x)) + \lambda^2 W_2(y_1). \\ &+ \lambda^3 W_1(x) + (\lambda^4 + \lambda^2\beta)x \end{aligned} \tag{7}$$

Then, the output y_6 of the second multi-level skip connection of the residual block is:

$$\begin{aligned} y_6 &= y_5 + \lambda y_3 \\ &= W_4(y_4) + \lambda(W_3(y_2 + \beta x)) + (\lambda^2 + \beta)(W_2(y_1)). \\ &+ (\lambda^3 + \lambda\beta)W_1(x) + (\lambda^2 + \beta)^2 x \end{aligned} \tag{8}$$

It can be determined from the output y_6 that each convolution layer output value of the residual block connected by two adjacent multi-level skip convolves the feature vector of each previous convolution layer, making full use of the adjacent residual information of the feature vector of the convolution layer inside the difference block; thus, the residual block can extract more feature information.

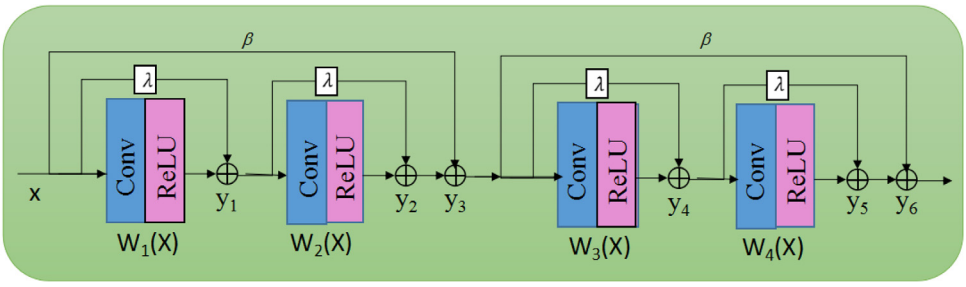

Fig. 3. Residual block structural diagram of two adjacent multi-level skip connections.

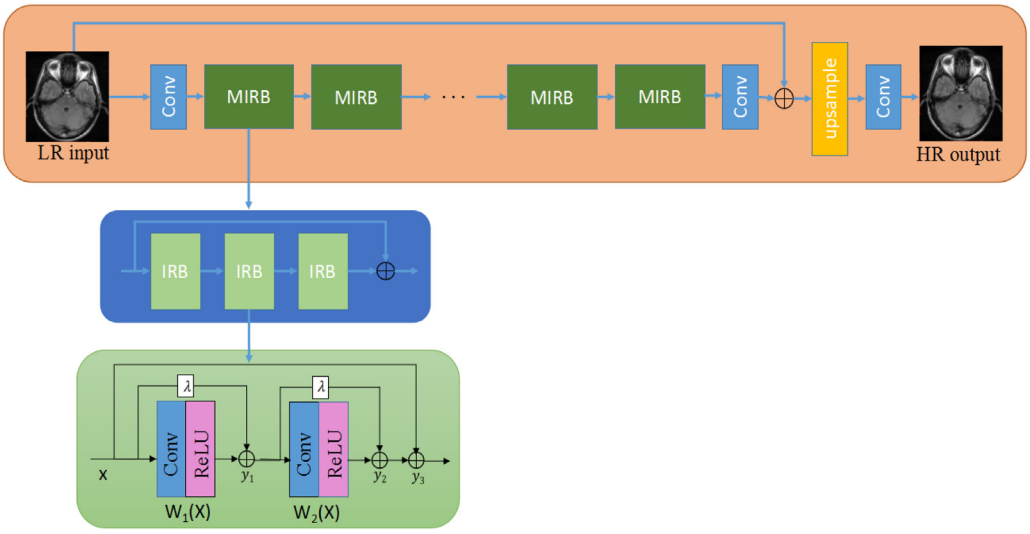

Fig. 4. Structural framework of MIRN for medical image processing.

2.3. Multiple improved residual network

To extract the deep information of a medical image and avoid the problems of gradient disappearance and gradient explosion when training the network, the MIRN is designed. The network has a total of 47 convolutional layers, consisting of 7 MIRBs (42 convolutional layers), intermediate layers, input layers, output layers, and sub-pixel convolutional layers (2 convolutional layers). Additionally, it consists of end-to-end connections, and the network structure is illustrated in Fig. 4. An LR image is input into the network, the predicted residual eigenvalues are obtained through the MIRB, and the residual block eigenvalues and LR image values are then combined into an HR image value.

3. Experimental results

3.1. Experimental settings

The experimental environment was an Ubuntu 18.04.02 LTS operating system. The deep-learning framework used included PyTorch 1.0, CUDA Toolkit 9.0, cuDNN 7.0, Python 3.6, and an NVIDIA GeForce TITAN XP GPU. The entire network was divided into two components, namely the training and testing components. In the training component, the learning rate was set to 10^{-4}, the weight decay in the iterative process was 0, the momentum was set to 0.9, and the batch size in the training dataset was 128. In the testing component, the batch size was set to 32. The Euclidean loss function [25] was employed, which enables the computation of the loss between the predicted value and the tag value.

This method uses the mean-square error (MSE) function [42] as the loss function to estimate the network parameter A. The MSE function is as follows:

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n} \|G(Y_i, \theta) - X_i\|^2, \tag{9}$$

where the parameter θ is estimated by minimizing the loss of the reconstructed image and the corresponding true HR image, $G(Y_i, \theta)$ is the reconstructed image, X_i is the HR image sample, Y_i is the corresponding LR image sample, and n is the number of training samples.

The proposed method uses the SGD algorithm to optimize the training of the MIRN, and the variable for the stochastic sample is "*sample*". The expression of the updated parameter θ of the SGD algorithm is as follows:

$$\nabla L(\theta_t^l) = \frac{\partial L}{\partial \theta_t^l}, \tag{10}$$

$$\Delta_{t+1} = \mu\Delta_t - \eta\nabla L(\theta_t^l), \tag{11}$$

$$\theta_{t+1}^l = \theta_t^l + \Delta_{t+1}, \tag{12}$$

where t is the number of iterations, l is the number of convolution layers, μ is the weight of the previous iteration and $\mu \in [0,1]$, and η is the learning rate.

The activation function used in this method is ReLu [43]. During the calculation process of a deep residual network connected by multi-level skip connections, the use of many parameters will increase the complexity of network calculation, resulting in a slow network training speed; therefore, the same residual function is used in this paper.

$$F(x) = \max(0, W_i^* x + B_i), \tag{13}$$

where i represents the number of layers, W_i is a $64 \times 3 \times 3 \times 64$ filter, and B_i is a 64-dimensional deviation. The filter size of the input convolution layer of the network is $1 \times 3 \times 3 \times 64$, and the filter size of the output convolution layer of the network is $64 \times 3 \times 3 \times 1$. To ensure that the feature map after the convolution calculation maintains the same size as the feature map of the input convolution layer, the moving step size of all convolutional layers of the network was set to 1, and the padding was set to 1. The skip connection parameters of the network residual block were $\lambda = 0.1$ and $\beta = 0.1$.

3.2. Dataset and protocol

The public data sets BSD500 and T91 [32] were used in the experiment, and the two training sets have a total of 591 images. Because the depth model usually benefits from large amounts of data, 591 images are not sufficient to push the model to its best performance. Thus, to make full use of the data set, MATLAB was used to expand the data of the BSD500 and T91 training set images. Two methods, namely scaling and rotation, were used to increase the data. Each image was scaled by the ratios of 0.7, 0.8, and 0.9. Additionally, each image was respectively rotated by 90°, 180°, and 270°. The original images from BSD500 and T91 were first sampled 4 times with a Gaussian filter ($\delta = 1.5$) and bicubic interpolation to obtain an LR image. The LR training image was then cut into a group of 96×96 HR image blocks with a step size of 12, and 9456 images were finally obtained. For the initialization of the parameters in the network, they are randomly generated from a Gaussian distribution with a zero mean and a standard deviation of 0.001, and the obtained high-resolution image block is first blurred as much as possible using a Gaussian kernel with a size of 7×7 and a standard deviation of 1.6 to simulate a naturally blurred image.

3.3. Evaluation standard

The public datasets Set 5 [44], Set 14 [45], and Urban100 [46] were chosen for testing, and the PSNR [38] and SSIM [39] were selected as reference evaluation indexes for the evaluation of image quality. The superiority of the proposed MIRN medical image super-resolution reconstruction algorithm was then evaluated based on the publicly available data set IDI (I Do Imaging). The IDI website offers free and open-source medical imaging software, and it includes nearly 300 categorized, ranked, and searchable software projects [34].

In general, a higher PSNR indicates a higher-resolution image, and is generally defined by the MSE:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2. \tag{14}$$

The PSNR is expressed as:

$$PSNR = 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right), \tag{15}$$

where I and K represent two images of size $m \times n$, and the unit of the specified PSNR is dB; the smaller the value, the larger the distortion, and vice versa.

The SSIM is expressed as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_1^2 + \mu_2^2 + c_1)(\sigma_1^2 + \sigma_2^2 + c_2)}, \tag{16}$$

where μ_x represents the average value of x, μ_y represents the average value of y, σ_x represents the variance of x, σ_y represents the variance of y, σ_{xy} represents the covariance of x and y, and the range of the prescribed structural similarity is $(-1, 0]$.

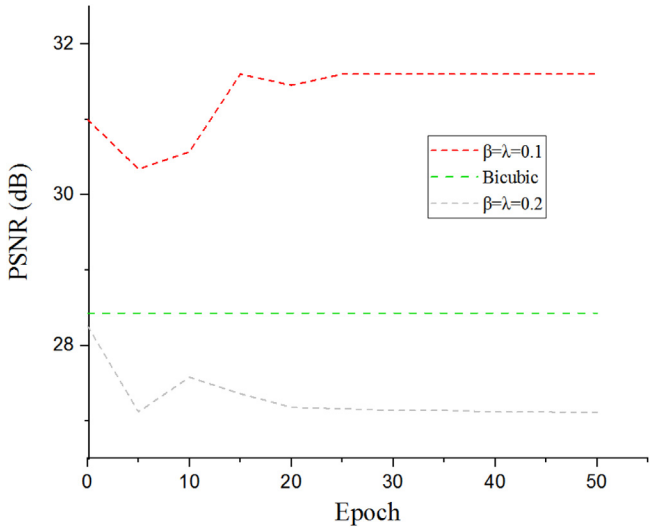

Fig. 5. PSNR curves measured with different skip coefficients.

3.4. Model analysis

In the proposed MIRN, the parameters are initialized to a Gaussian distribution with a mean of 0 and a standard deviation of 0.001, and the initialization of the bias is set to 0. The network was trained for a total of 200 epochs. To speed up the training process, an adjustable learning rate strategy was used to train the network. The initial learning rate was set to 0.1, and the learning rate decreased to 0.1 times the original learning rate every 10 epochs. When the learning rate decreased to 0.0001, it was then maintained at 0.0001 and the batch size was set to 128. Fig. 5 presents the curve of the PSNR of the network super-resolution model with the skip coefficients $\beta = \lambda = 0.1$ and $\beta = \lambda = 0.2$, respectively, on the Set 5 test set when the amplification factor was 2. The reconstruction effect of the network reconstruction model using the skip coefficient $\beta = \lambda = 0.1$ was significantly better than that using the bicubic algorithm and the skip coefficient $\beta = \lambda = 0.2$.

Fig. 6 presents the reconstruction performance versus the number of network parameters of CNN-based SR methods. By sharing parameters and using recursive layers, the proposed MIRN has about 76% fewer parameters than the EDSR [47] and 6% more than the DRCN. Although the proposed model has a smaller footprint, it was found to achieve state-of-the-art performance as compared to the CNN-based methods. As compared to the SRCNN [30] and FSRCNN [31], the proposed MIRN presented an improvement of about 0.89 to 1.2 dB on the challenging Set 5 dataset for $2 \times$ SR. Therefore, the proposed MIRN strikes a balance between reconstruction accuracy and execution time.

3.5. Comparison with state-of-the-art methods

To objectively compare the advantages and disadvantages of the algorithms, the Bicubic [8], A+ [23], SCN [48], SRCNN [30], FSRCNN [31], VDSR [32], DRCN [35], LapSRN [49], and DRRN [36] methods were experimentally compared with the proposed algorithm. Table 1 respectively present the performances of the algorithms on the Set 5, Set 14, and Urban100 test sets with respective amplification factors of 2, 3, and 4. As presented in Table 1, the average PSNR and SSIM values obtained by the MIRN algorithm were greatly improved when the amplification factors were 2, 3, and 4. The state-of-the-art SR algorithms were quantitatively evaluated by the average PSNR/SSIM for the scale factors

Table 1
Comparison of experimental results with an amplification factor of 2, 3, and 4.

Algorithm	Scale	Set 5		Set 14		Urban 100	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic [7]	2×	33.69	0.931	30.25	0.870	26.88	0.841
A+ [23]	2×	36.60	0.955	32.32	0.906	29.25	0.895
SCN [47]	2×	36.58	0.954	32.35	0.905	29.52	0.897
SRCNN [30]	2×	36.72	0.955	32.51	0.908	29.53	0.896
FSRCNN [31]	2×	37.05	0.956	32.66	0.909	29.88	0.902
VDSR [32]	2×	37.53	0.959	33.05	0.913	30.77	0.914
DRCN [35]	2×	37.63	0.959	33.06	0.912	30.76	0.914
LapSRN [48]	2×	37.52	0.959	33.08	0.913	30.41	0.910
DRRN [36]	2×	37.74	0.959	33.23	0.914	31.23	0.919
MIRN	2×	37.88	0.960	33.72	0.919	31.27	0.919
Bicubic [7]	3×	30.41	0.869	27.79	0.775	24.46	0.735
A+ [23]	3×	32.62	0.909	29.15	0.820	26.05	0.799
SCN [47]	3×	32.62	0.908	29.16	0.818	26.21	0.801
SRCNN [30]	3×	32.78	0.909	29.32	0.823	26.25	0.801
FSRCNN [31]	3×	33.18	0.914	29.37	0.824	26.43	0.808
VDSR [32]	3×	33.67	0.921	29.78	0.832	27.14	0.829
DRCN [35]	3×	33.83	0.922	29.77	0.932	27.15	0.828
LapSRN [48]	3×	33.82	0.922	29.87	0.832	27.07	0.828
DRRN [36]	3×	34.03	0.924	29.96	0.835	27.53	0.764
MIRN	3×	34.04	0.924	30.14	0.843	27.24	0.832
Bicubic [7]	4×	28.43	0.811	26.22	0.715	23.14	0.658
A+ [23]	4×	30.32	0.860	27.34	0.751	24.34	0.721
SCN [47]	4×	30.41	0.863	27.39	0.751	24.52	0.726
SRCNN [30]	4×	30.50	0.863	27.52	0.753	24.53	0.725
FSRCNN [31]	4×	30.72	0.866	27.61	0.755	24.62	0.728
VDSR [32]	4×	31.35	0.883	28.02	0.768	25.18	0.754
DRCN [35]	4×	31.54	0.884	28.03	0.768	25.14	0.752
LapSRN [48]	4×	31.54	0.885	28.19	0.772	25.21	0.756
DRRN [36]	4×	31.68	0.888	28.21	0.772	25.44	0.764
MIRN	4×	31.61	0.886	28.49	0.784	25.27	0.759

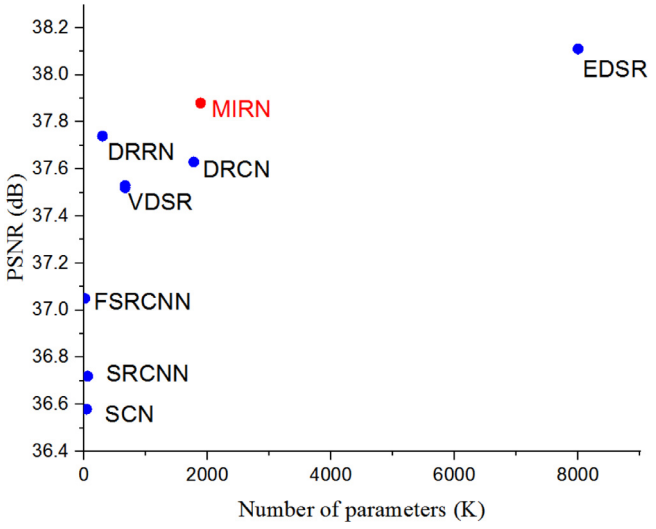

Fig. 6. Number of network parameters versus performance.

of 2×, 3×, and 4×. Red text indicates the best performance, and green text indicates the second-best performance.

Fig. 7 presents visual comparisons of the performances of the investigated algorithms on the Head, Brain, and Knee datasets at 4 × SR. The proposed method was found to accurately reconstruct parallel straight lines, grid patterns, and text from magnetic resonance imaging (MRI) images.

4. Discussion

The advantages and disadvantages of the proposed method and similar existing state-of-the-art algorithms were experimentally compared, and the test results of multiple algorithms on the same image were analyzed to judge whether the reconstruction quality was good or bad. As presented in Table 1, on the Set 5 test set, the average PSNR value obtained by the proposed method was improved by 0.35 dB and 4.19 dB as compared with those of the VDSR [32] and Bicubic algorithms, respectively. Additionally, the average PSNR and SSIM values obtained with the amplification factors of 2, 3, and 4 were greatly improved. Fig. 7 presents the effects of the proposed and existing state-of-the-art algorithms on the test set with an amplification factor of 4. Fig. 7(A), (B), and (C) are respectively the reconstruction results on the Head, Brain, and Knee datasets. As is evident from Fig. 7(A), the head scan images reconstructed by the DRRN [36], DRCN [35], FSRCNN [31], and VDSR [32] algorithms are significantly clearer than those reconstructed using the traditional SCN and SRCNN algorithms. By comparing the DRRN [36], DRCN [35], FSRCNN [31], and VDSR [32] algorithms, the reconstructed head scan images do not appear to be significantly different in terms of clarity; nevertheless, the lesion area of each reconstructed head scan image was enlarged for clearer observation. It is evident that the lesion areas of the head scan images reconstructed using the SRCNN [30] and SCN [48] algorithms are blurred, and the details are not clear; additionally, the lesion area of the head scan image reconstructed using the VDSR [32] algorithm is clear, but the blood vessels are blurred. In contrast, the lesion area of the head scan image reconstructed using the proposed MIRN method is clear, and the details of the blood vessels are evident. In addition, the enlarged knee image texture details and the brain image details shown in the lower-right corners of the medical images reconstructed by MIRN in Fig. 7(B) and (C) are clearer than those reconstructed by the other algorithms.

By comparing the PSNR and SSIM test values in Table 1, the PSNR and SSIM values obtained using the proposed MIRN method on the Set 5 and Set 14 test sets under the amplification factors of 2, 3, and 4 were higher than those obtained using the existing state-of-the-art algorithms. Additionally, although the

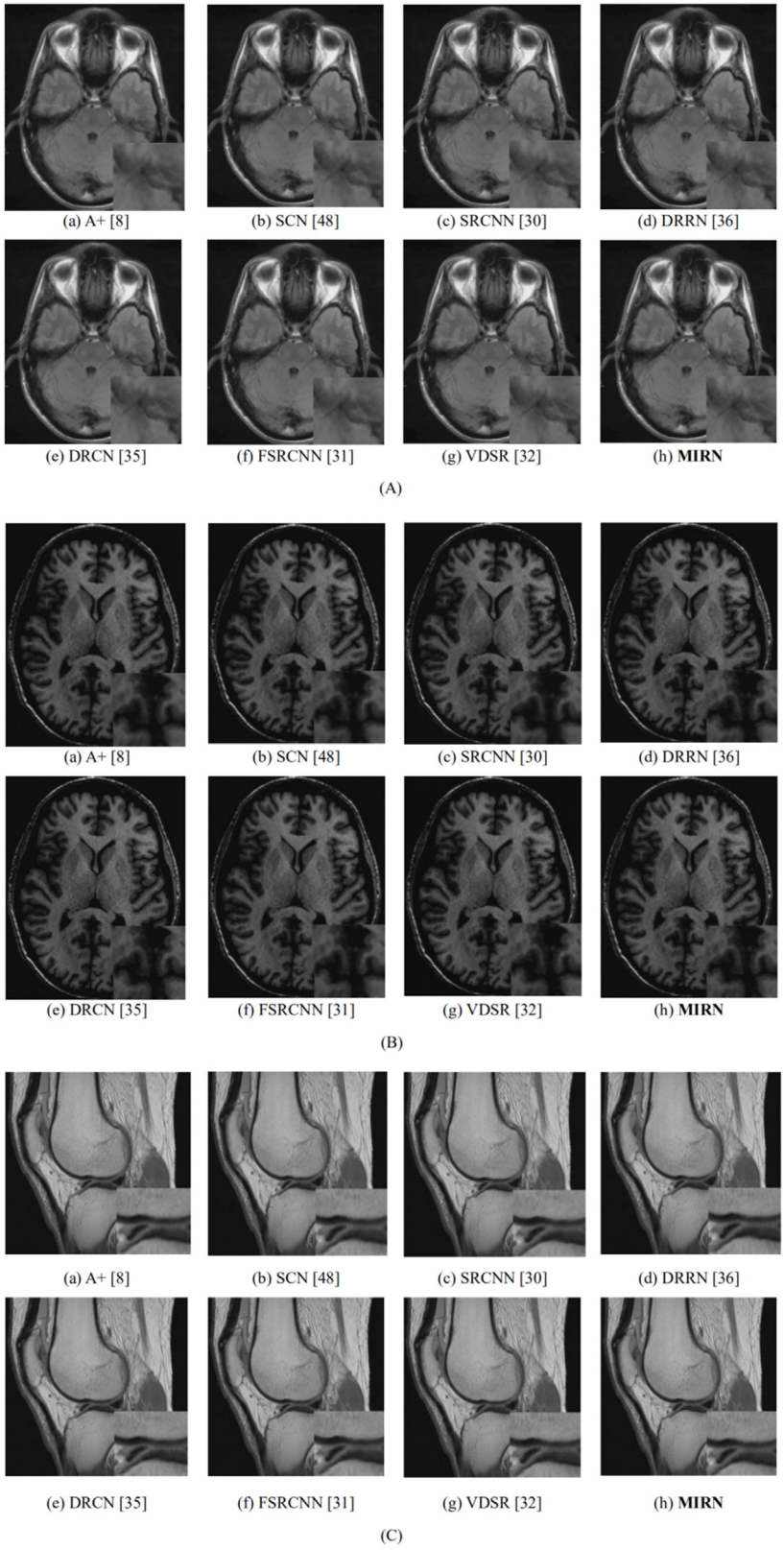

Fig. 7. Qualitative comparison of the proposed model with other state-of-the-art methods at 4× super-resolution based on MRI images.

PSNR values obtained using the proposed MIRN method on the Urban100 test set with the amplification factors of 3 and 4 were lower than those obtained using the DRRN [36] algorithm, the PSNR and SSIM values obtained by the proposed method with the amplification factor of 2 were higher than those obtained by the existing state-of-the-art algorithms. As can be seen from Fig. 7, the reconstructed images obtained using the proposed method were not only clearer, but also had more detailed information,

more delicate image textures, more realism, and better visual effects. Therefore, the proposed MIRN method is superior to the investigated existing state-of-the-art algorithms.

5. Conclusion

To solve the problem of the lack of correlation between magnetic resonance image feature information, this paper proposed a multiple improved residual network super-resolution reconstruction method to extract the deep information of medical images. The network structure of this method is primarily composed of 7 multiple improved residual blocks with end-to-end connections. A low-resolution image is input into the network, and a predicted residual image is obtained via the improved residual blocks connected by multi-level skip. Finally, the residual image and the low-resolution image are combined into a high-resolution image. The proposed method was compared with the state-of-the-art Bicubic [8], A+ [23], SCN [48], SRCNN [30], FSRCNN [31], VDSR [32], DRCN [35], LapSRN [49], and DRRN algorithms [36], and it was found that the proposed method displayed better performance, and its reconstructed images had better visual effects. The findings fully illustrate the effectiveness of the improved algorithm proposed in this paper. In future work, better optimization strategies will be sought, the width of the network structure will be expanded, and the accuracy of super-resolution image reconstruction will be further improved. The research model implemented in this paper can be used as a guide for improving medical scan resolution in the future, and can support the improvement of medical image diagnosis.

CRediT authorship contribution statement

Defu Qiu: Carried out the experiments, The statistical analysis and manuscript preparation. **Lixin Zheng:** Experimental design, Analysis and interpretation of data, Provided final approval of the version of the submitted manuscript. **Jianqing Zhu:** Responsible for data processing and revises the article. **Detian Huang:** Participated in the analysis of the IDI dataset.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61976098 and 61901183, in part by the Science and Technology Bureau of Quanzhou, China under Grant No. 2017G027, and in part by the Science and Technology Bureau of Xiamen, China under Grant No. 3502Z20173045. The authors would like to thank the Fujian Provincial Academic Engineering Research Centre in Industrial Intellectual Techniques and Systems and the Innovation Center of Engineering College of Huaqiao University for providing high-performance servers to support this research. All authors read and approved the final manuscript.

References

[1] S. Peled, Y. Yeshurun, Super resolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging, Magn. Reson. Med. Off. J. Soc. Magn. Reson. Med. 45 (1) (2001) 29.
[2] Y. Xu, L. Peng, G.Y. Li, Multi modal registration of structural features and mutual information of medical image, Future Gener. Comput. Syst. 93 (2019) 499–505.
[3] M.W. Thornton, P.M. Atkinson, D.A. Holland, Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping, Int. J. Remote Sens. 27 (3) (2006) 473–491.
[4] B.K. Gunturk, R.A.U. Batu, Y. Altunbasak, et al., Eigen-face-domain super-resolution for face recognition, IEEE Trans. Image Process. 12 (5) (2003) 597–606.
[5] L. Zhang, H. Zhang, H. Shen, et al., A super-resolution reconstruction algorithm for surveillance images, Signal Process. 90 (3) (2010) 848–859.
[6] H. Chavezroman, V. Gonzalezhuitron, A. Hernandez, V.I. Ponomaryov, Super resolution imaging via sparse interpolation in wavelet domain with implementation in DSP, and GPU, in: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Springer International Publishing, 2014.
[7] H. Su, J. Zhou, Z.H. Zhang, A summary of super-resolution image reconstruction methods, Acta Automat. Sinica 39 (8) (2013) 1202–1213.
[8] R. Keys, Cubic convolution interpolation for digital image processing, IEEE Trans. Acoust. Speech Signal Process. 29 (6) (1981) 1153–1160.
[9] D.Y. Ge, X.F. Yao, Z.T. Lian, Binocular vision calibration and 3D reconstruction with an orthogonal learning neural network, Multimedia Tools Appl. 75 (23) (2016) 15635–15650.
[10] Q. Yuan, S.X. Jin, Improved sequence image super-resolution reconstruction method, J. Comput. Appl. 29 (12) (2009) 3310–3313.
[11] H. Ji, Fermuller, Wavelet-based super-resolution reconstruction: theory and algorithm, in: European Conference on Computer Vision, 2006, pp. 295–307.
[12] H. Ji, Fermuller, Robust wavelet-based super-resolution reconstruction: theory and algorithm, IEEE Trans. Pattern Anal. Mach. Intell. 31 (4) (2009) 649–660.
[13] H. Stark, P. Oskoui, High-resolution image recovery from image-plane arrays, using convex projections, J. Opt. Soc. Amer. A 6 (11) (1989) 1715–1726.
[14] N. Nguyen, P. Milanfar, A wavelet-based interpolation-restoration method for super resolution (wavelet super resolution), Circuits Systems Signal Process. 19 (4) (2000) 321–338.
[15] V. Patanavijit, S. Jitapunkul, An iterative super-resolution reconstruction of image sequences using fast affine block-based registration with BTV regularization, in: Proceedings of 2006 IEEE Asia Pacific Conference on Circuits and Systems, Singapore, 2006, pp. 1717–1720.
[16] X. Li, M.T. Orchard, New edge directed interpolation, IEEE Trans. Image Process. 10 (10) (2001) 1521–1527.
[17] X. Zhang, X. Wu, Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation, IEEE Trans. Image Process. 17 (6) (2008) 887–896.
[18] M. Irani, S. Peleg, Improving resolution by image registration, CVGIP: Graph. Models Image Process. 53 (3) (1991) 231–239.
[19] M. Haris, Shakhnarovich, N. Ukita, Deep back-projection networks for super-resolution, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 1664–1673.
[20] E.S. Meinel, Origins of linear and nonlinear recursive restoration algorithms, J. Opt. Soc. Amer. A 3 (6) (1986) 787–799.
[21] N. Ai, J. Peng, X. Zhu, X. Feng, Single image super-resolution by combining self-learning and example-based learning methods, Multimedia Tools Appl. 75 (11) (2016) 6647–6662.
[22] R. Timofte, V. De, L.V. Gool, Anchored neighborhood regression for fast example-based super-resolution, in: Proceedings of the 2013 IEEE International Conference on Computer Vision, IEEE, Piscataway, NJ, 2013, pp. 1920–1927.
[23] R. Timofte, V.D. Smet, L.V. Gool, A +: adjusted anchored neighborhood regression for fast super-resolution, in: Proceedings of the 2014 Asian Conference on Computer Vision, Springer, Berlin, 2014, pp. 111–126.
[24] C.Y. Yang, M.H. Yang, Fast direct super-resolution by simple functions, in: Proceedings of the 2013 IEEE International Conference on Computer Vision, IEEE Computer Society, Washington, DC, 2013, pp. 561–568.
[25] R. Zeyde, M. Elad, M. Protter, On single image scale-up using sparse-representations, in: International Conference on Curves and Surfaces, 2010, pp. 3–13.
[26] J. Yang, J. Wright, T. Huang, et al., Image super-resolution as sparse representation of raw image patches, in: CVPR 2008: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, Washington, DC, 2008, pp. 1–8.
[27] J. Yang, J. Wright, T.S. Huang, et al., Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.
[28] J. Wang, S. Zhu, Y. Gong, Resolution enhancement based on learning the sparse association of image patches, Pattern Recognit. Lett. 31 (1) (2010) 1–10.
[29] A. Krizhevsky, I. Sutskever, G.E. Hinton, Image net classification with deep convolutional neural networks, in: Proceedings of the 2012 International Conference on Neural Information Processing Systems, Curran Associates Inc, West Chester, OH, 2012, pp. 1097–1105.

[30] C. Dong, C.L. Chen, K. He, et al., Learning a deep convolutional network for image super-resolution, in: ECCV 2014: Proceedings of the 2014 European Conference on Computer Vision, Springer, Berlin, 2014, pp. 184–199.
[31] C. Dong, C.L. Chen, X. Tang, Accelerating the super-resolution convolutional neural network, in: ECCV 2016: Proceedings of the 2016 European Conference on Computer Vision, Springer, Berlin, 2016, pp. 391–407.
[32] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in: Computer Vision and Pattern Recognition, IEEE, 2016, pp. 1646–1654.
[33] S. Karen, Andrew Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, ArXiv preprint arXiv:1409.1556.
[34] D. Qiu, S. Zhang, L. Zheng, Y. Liu, J. Zhu, Super-resolution reconstruction method of knee MRI based on deep learning, Comput. Methods Programs Biomed. (2019).
[35] J. Kim, J. Kwon Lee, K. Mu Lee, Deeply-recursive convolutional network for image super-resolution, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1637–1645.
[36] Y. Tai, J. Yang, X. Liu, Image super-resolution via deep recursive residual network, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 3147–3155.
[37] W. Shi, J. Caballero, F. Huszar, J. Totz, A. Aitken, R. Bishop, D. Rueckert, Z. Wang, Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1–3.
[38] F. Yuan, L. Huang, Y. Yao, An improvedPSNR algorithm for objective video quality evaluation, in: 2007 Chinese Control Conference, Hunan, China, 2007, pp. 376–380.
[39] Z. Wang, A.C. Bovik, H.R. Sheikh, et al., Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[40] K. He, X. Zhang, S. Ren, et al., Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016, pp. 770–778.
[41] J. Johnson, A. Alahi, F.F. Li, Perceptual losses for real-time style transfer and super-resolution, in: Proceedings of the 2016European Conference on Computer Vision, Springer, Berlin, 2016, pp. 694–711.
[42] A. Bruhn, J. Weickert, C. Schnorr, Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods, Int. J. Comput. Vis. (2005) 211–231.
[43] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 2010 International Conference on International Conference on Machine Learning, Omnipress, Madison, WI, 2010, pp. 807–814.
[44] M. Bevilacqua, A. Roumy, C. Guillemot, et al., Low complexity single image super-resolution based on nonnegative neighbor embedding, in: Proceedings of the 2012 British Machine Vision Conference, Vol. 135, BMVC Press, [S. l.], 2012, pp. 1–10.
[45] R. Zeyde, M. Elad, R.M. Protte, On single image scale-up using sparse-representations, in: Proceedings of the 2010 International Conference on Curves and Surfaces, Springer, Berlin, 2010, pp. 711–730.
[46] J. Huang, A. Singh, N. Ahuja, Single image super resolution using transformed self-exemplars, in: CVPR, 2015.
[47] B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, Enhanced deep residual networks for single image super-resolution, in: CVPRW, 2017.
[48] Z. Wang, D. Liu, J. Yang, W. Han, T. Huang, Deep networks for image super-resolution with sparse prior, in: IEEE International Conference on Computer Vision, 2015, pp. 1–15.
[49] W. Lai, J. Huang, N. Ahuja, M. Yang, Deep laplacian pyramid networks for fast and accurate super-resolution, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2–15.

Defu Qiu was born in Linshu County, Shandong Province, China in 1992. He is currently pursuing a MS. degree in Computer Science and Technology from Huaqiao University of Engineering. Since 2017, he has been a Research Assistant with the College of Engineering, Huaqiao University. As of now, his research interest includes machine learning and the application of deep learning in medical image super-resolution reconstruction.

Lixin Zheng was born 19 April 1967, his research is mainly focus on image recognition and machine vision technology. Now he is the dean and professor of Engineering College of Huaqiao University, vice president of Fujian automation society, vice president of Xiamen automation society and director of Fujian power supply society.

Dr. Zheng graduated from the Department of electronic engineering, Huaqiao University and received a bachelor's degree in 1987. In 1990, he was graduated from the Department of Mechanical Engineering in the field of testing and automation control and received a Master's degree in 1990. In 1997, he went to Japan for further study. In 2002, he received his Ph.D. in predictive control at Tianjin University. In 2004, he won the backbone of the outstanding young and middle-aged teachers of Huaqiao University, and in 2006, was awarded the title of "Outstanding Teacher of Huaqiao University". In 2007, he was supported by the Fujian excellent talent support program.

Jianqing Zhu received the B.S. degree in communication engineering and the M.S. degree in communication and information system from the School of Information Science and Engineering, Huaqiao University, Xiamen, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2015. He is currently an Assistant Professor with the College of Engineering, Huaqiao University, Quanzhou, China. His current research interests include computer vision and pattern recognition, with a focus on image and video analysis, particularly person re-identification, object detection, and video surveillance.

Detian Huang received the B.S. degree in electronic information engineering from Xiamen University, Xiamen, China, in 2008, and the Ph.D. degree in circuits and systems from the University of Chinese Academy of Sciences, Beijing, in 2013. He is currently an Assistant Professor with the College of Engineering, Huaqiao University, Quanzhou, China. His research interests include computer vision, target tracking, image restoration, and machine learning.