Received: 19 April 2022 | Revised: 25 June 2022 | Accepted: 9 January 2023

IET Image Processing

DOI: 10.1049/ipr2.12742

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Medical image blind super-resolution based on improved degradation process

Dangguo Shao[1,2] | Li Qin[1] | Yan Xiang[1,2] | Lei Ma[1] | Hui Xu[3]

[1]Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, People's Republic of China

[2]Yunnan Provincial Key Laboratory of Artificial Intelligence, Kunming, People's Republic of China

[3]First Affiliated Hospital of Kunming Medical University, Kunming, People's Republic of China

Correspondence
Dangguo Shao, Faculty of Information Engineering and Automation, Kunming University of Science and Technology, No. 727, Jingming South Road, Chenggong District, Kunming City, Yunnan Province 644500, People's Republic of China.
Email: huntersdg@126.com

Funding information
National Science Foundation of China, Grant/Award Number: 11773012

Abstract

Clinical diagnosis has high requirements for the resolution of medical images, but most existing medical images super- resolution (SR) methods are performed under a known or specific degradation kernel. However, the difference between the actual degradations and their assumed degradation kernels results in a severe performance drop for the advanced SR methods in real applications. This paper proposes a medical image blind super-resolution model (Med-BSR) based on an improved degradation process to handle this issue. The model makes each of the degradation factors in medical image blind SR, such as blur, noise, and downsampling, more complex and practical. Specifically, the authors use the random select/combine strategy to randomly arrange and combine the type and order of each degradation factor, which significantly expands the degradation space. The authors also improved the loss function of the primary enhanced super-resolution generative adversarial networks (ESRGAN) network. The extensive experimental results demonstrate that the authors' designed model can accurately restore the natural degradation process, which can reconstruct high-quality SR medical images. It also has a good generalization ability to realistic images simultaneously.

1 | INTRODUCTION

Medical image has gained importance as a diagnostic, therapeutic, and regenerative medicine tool [1–3]. The most commonly used imaging techniques include magnetic resonance imaging (MRI) [4, 5], computed tomography (CT) [6, 7], and positron emission computed tomography [8, 9]. Although high-resolution (HR) medical images are of great help in clinical diagnosis, their resolution is limited by factors such as the precision of the machining, the performance of the detector material, and the radiation dose that the patient can withstand. Thus, it is valuable to use super-resolution methods to improve the resolution of medical images.

Nowadays, the learning-based methods [10–17] are the dominant approach to medical image super-resolution (SR) task. Its essence is to obtain a mapping relationship between HR and low resolution (LR) images [12]. However, this is a typical ill-posed problem, because the mapping relationship between the LR and HR images is one-to-many. An LR image can be degraded from different HR images. Therefore, an accurate degradation model is crucial for medical image SR tasks.

The non-blind SR methods [16, 25] usually only consider the degradation of bicubic downsampling. While these networks are simple and easy to implement, their performance drops significantly when dealing with arbitrary images, such as medical images with complex degradation types. As shown in Figure 1, if a predefined non-blind SR model is used on arbitrary LR images of unknown degradation type, this leads to a large domain gap, which in turn produces poor reconstruction results. To avoid the domain gap, some works try to research blind SR methods [21–26] with unknown degradation kernel. These methods introduce noise or blur kernel estimation modules. Compared with the traditional non-blind SR methods, this model is more practical. However, it cannot accurately predict the real noise level and blur degree, and the performance deterioration is more serious when the input image is MRI or CT image.

To overcome the above problems, this paper proposes a medical image blind super-resolution model (Med-BSR) based on

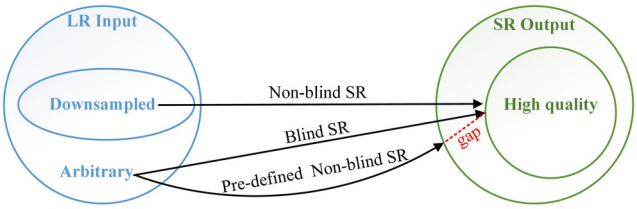

FIGURE 1 The main distinction between non-blind SR and blind SR methods. SR, super-resolution.

the improved degradation process, which can achieve a better medical image SR performance. Where the image degradation parameters are unknown, Med-BSR introduces all types of degradation factors and aligns them as closely as possible with the real degradation. Then we use the random select/combine strategy to combine the types and order of blur, noise, and downsampling. The degradation space is expanded, and the matching degree between the model and the real degradation process is further improved. Based on the enhanced super-resolution generative adversarial networks (ESRGAN) model, we used the enhanced visual geometry group (VGG) loss function to optimize the SR effect of medical images. The main contributions of our Med-BSR are as follows:

1. To restore the real degradation process of medical images, we take into account the diversity of degradation factors. Our model introduces two kinds of blur, one kind of noise, and three kinds of downsampling.
2. Use the random select/combine strategy to arrange and combine degradation factors to enlarge degradation space.
3. The VGG perceptual loss function also been adopted with the aim of increasing the quality of the SR medical images.

2 | RELATED WORK

2.1 | Non-blind SR

These methods are carried out on the assumption that the degradation kernel is known. Dong et al. designed a three-layer convolutional network (SRCNN) containing feature extraction, non-linear mapping, and image reconstruction functions, which pioneered the application of deep learning in the field of image SR [12]. SRCNN achieves better reconstruction results than traditional methods, and subsequently convolutional neural networks are widely used for image SR tasks. Enhanced deep residual networks for single image super-resolution (EDSR) [18] expands the model scale by removing unnecessary modules from traditional residual networks and solves the network degradation problem in SRCNN, but the complex network structure is inconvenient to train. To maintain the image texture details, the generative adversarial network (GAN) is introduced into the image SR training task. SRGAN [19] consists of a generator network and a discriminator network. The former can generate images with both high- and low-frequency information, and the latter is used to judge the effects of the generated images. Enhanced super-resolution generative adversarial networks (ESRGAN) [20] introduces a novel Residual-in-Residual Dense Block (RRDB) and uses a pre-activation feature expression to constrain the perceptual loss function. A series of GAN-based SR methods effectively improve the visual effect of reconstructed images. However, the degradation kernel of the above methods are all known or specific (e.g. bicubic downsampling), which is still different from the real degradation process of the LR image.

2.2 | Blind SR

The degradation kernel of this method is unknown. Kligler et al. [22] first estimated the degradation parameters of the input LR image, and then used the non-blind SR methods [27] to obtain the HR image. It can generate kernels based on the characteristics of the input image, but the non-blind SR methods are insensitive to changes in the degraded kernels resulting in overly smooth SR images. To solve this problem, Gu et al. [23] proposed an iterative kernel correction approach (IKC), which utilizes the intermediate SR reconstruction results to correct the blur kernel and repeats for several times until convergence. Luo et al. [21] proposed a deep alternating network (DAN) consisting of kernel estimator and HR image restorer module. They aim to jointly estimate blur kernel and HR images, improving the sensory quality of the generated images to some extent. However, these methods do not take into account all noise types and still suffer from inconsistency with the true degradation model. As such, they cannot be used as an accurate priori information to constrain the solution space.

3 | METHOD

3.1 | Problem formulation

In general, the traditional non-blind SR degradation models [16] assume that only downsampling of the HR image is required to obtain the LR image, expressed mathematically as follows:

$$I^{LR} = (I^{HR})\downarrow_S \tag{1}$$

where I^{LR} and I^{HR} denote respectively the LR image and the HR image, $\downarrow_s$ represents downsampling operation, and s is the amplification factor. In practical training, bicubic interpolation is the most commonly used downsampling method, so the deformation of Equation (1) is

$$I^{LR} = (I^{HR})\downarrow_s^{bicubic} \tag{2}$$

The blind SR methods [24, 30] argue that the process of degradation from a HR image to an LR image is influenced by a variety of factors. The above two formulas cannot be totally generalized. Mathematically, the complete degradation model is

$$I^{HR} = (k \otimes I^{LR})\downarrow_s + n \tag{3}$$

where $\otimes$ is the convolution operation, k denotes the blur kernel, and n represents the noise. Blur, noise, and downsampling are three key elements in the HR image degradation model. The degradation of a HR image requires these three steps: convolution with a blur kernel, the addition of a certain level of noise, and a corresponding downsampling operation. In order to follow the real degradation process of medical images, our approach is based on Equation (3) and delves into the three key degradation factors.

3.2 | Proposed method

3.2.1 | Blur

Whether it is a natural image or a medical image, there are varying degrees of blurring. We use two Gaussian blur kernels [23], one is an isotropic Gaussian blur kernel, denoted as B_{iso}, and the other is an anisotropic Gaussian blur kernel, denoted as B_{aniso}. Two Gaussian blurring kernels are chosen because it is possible to model blurring not only from the HR space but also from the LR space. On the one hand, the blurring operation is performed on the HR medical images first, so that more image detail information can be retained in the downsampling operation. On the other hand, as some original medical images are not clear, it can expand the space of blur degradation and provide more options for subsequent degradation operations.

For details, the values of the two blur kernel sizes range from (7×7, 9×9, …, 21×21). The scale factor s determines the width of the isotropic Gaussian kernel. When the scale factor is 2, the width is (0.1–2.4), and when the magnification scale is 4, the width is (0.1–2.8). The rotation angle of the anisotropic Gaussian blur kernel is (0–π). When the scale factor is 2 and 4, the length of each axis is (0.5–6) and (0.5–8), respectively. We chose all values in equal proportions for the experiments to avoid randomness results.

3.2.2 | Noise

Noise is inevitable, all medical images contain some noise more or less. It may come from the process of image acquisition or the image signal transmission process. MRI images have Rician noise, Gaussian noise, and other noise, and the noise within the CT image projection data approximately follows a Gaussian distribution [29]. To facilitate training, we combine the common noise types of the two images and use Gaussian noise to unify the noise model, denoted as N_g . The mathematical model is as follows:

$$y = x + n \tag{4}$$

where x and y represent respectively the clear and noise images, n denotes additive Gaussian noise. We employ a 3D zero-mean Gaussian noise model [31]. Here $n \sim N(0, \mathbf{\Sigma})$, $\mathbf{\Sigma}$ is the covariance matrix. When $\mathbf{\Sigma} = \sigma^2\mathbf{1}$ adding greyscale additive Gaussian noise; when $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, additive Gaussian noise is adopted; in other cases, generalized Gaussian noise is added. Here $\mathbf{1}$ and $\mathbf{I}$ are the 1×1 and 3×3 identity matrixes, respectively. The selection probabilities of the above three types of Gaussian noise are 2/5, 2/5, and 1/5, respectively; as for the value of σ , randomly selected from (1/255, 2/255, …, 25/255).

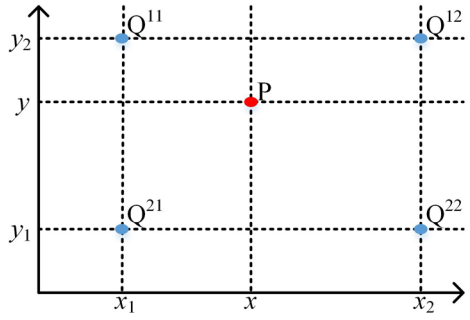

FIGURE 2 The schematic diagram of bilinear interpolation. Suppose we know the value of the function f at the four points: $Q^{11} = (x_1, y_2)$, $Q^{12} = (x_2, y_2)$, $Q^{21} = (x_1, y_1)$, $Q^{22} = (x_2, y_1)$.The value of the unknown function f at point $p = (x, y)$ can be obtained by doing two linear interpolations in the x and y directions, respectively

3.2.3 | Downsampling

Downsampling is used to reduce the size of the medical image. The nearest neighbour, bilinear, and bicubic interpolation [32] are the most common downsampling methods. Due to the many-to-one relationship between the HR and LR images, artificially LR images can be obtained if the downsampling method does not match the actual degradation of image. Such lossy images have a detrimental effect on subsequent experiments. To avoid this error and improve the match to the true degradation, we introduce the following three downsampling methods.

Nearest-neighbour interpolation ($D_s^{nearest}$). Selecting the nearest pixel as the pixel value of the unknown point is the most direct downsampling method without a calculation. The generated images are discontinuous in greyscale and have obvious jagged edges.

Bilinear interpolation ($D_s^{bilinear}$). As depicted in Figure 2, the linear interpolation [33] is performed twice for the four adjacent pixel points along the plane coordinate axis to obtain the pixel value of the unknown point P. The output images do not have greyscale discontinuities and look smoother.

Bicubic interpolation ($D_s^{bicubic}$). Sixteen adjacent pixels are interpolated cubically along the direction of the spatial coordinate axis to calculate the pixel of the unknown point. It considers the influence of four direct adjacent points and takes the greyscale change rate of each adjacent point into consideration. It is a widely used downsampling method with the best performance but an extensive computation. According to the approximate bicubic kernel estimation with different magnification scale factors, the bicubic interpolation method selects

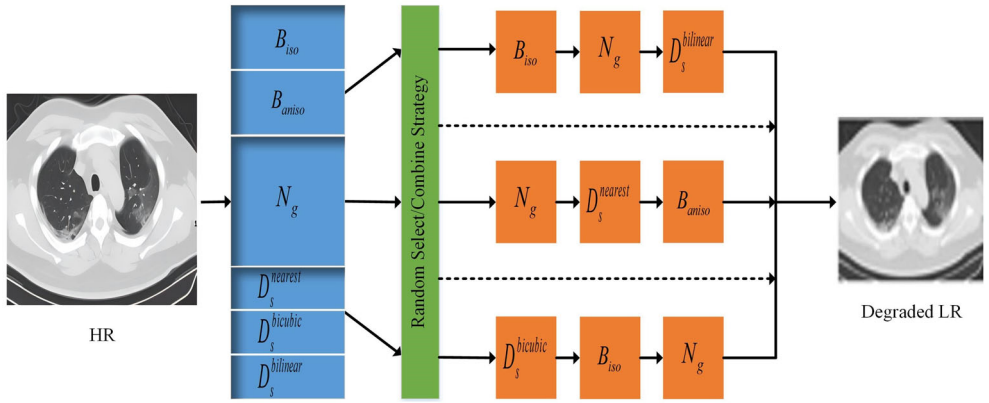

FIGURE 3 Schematic diagram of the improved degradation model

the upper left pixel block for each different $s \times s$ pixel block by default [21]. When the s is 2 or 4, the degraded LR images are shifted to the upper left by 0.5 or 1.5 pixels, respectively. Therefore, we use the two-dimensional (2D) linear grid interpolation method for correction. When s is 2, the Gaussian kernel is moved 0.5 pixels to the lower-right; if s is 4, the Gaussian kernel is moved to the lower-right by 1.5 pixels. Performing the bicubic interpolation operation after the above correction.

The above three methods introduce different degrees of blurring in practice, which is not mentioned in Section 3.2.1 because they are coupled during the actual operation. We randomly choose one of these three interpolation methods to change the size of the input MRI and CT images, and the probability of each method being chosen is 1/3.

3.3 | Random select/combine strategy

Combining the above degradation settings, we reformulate Equation (3,5) as

$$\begin{aligned} I^{HR} &= (K^* \otimes I^{LR}) \downarrow_s^* + N_g \\ K^* &\subset \{B_{iso}, B_{aniso}\} \\ \downarrow_s^* &\subset \left\{D_s^{nearest}, D_s^{bicubic}, D_s^{bilinear}\right\} \end{aligned} \tag{5}$$

From this it is known that there are many kinds of blur, noise, and downsampling in practical applications, and their order is also different in the actual degradation process. An LR image can be degraded from an HR image by blur, Gaussian noise, bicubic interpolation downsampling or bilinear interpolation downsampling, blur, Gaussian noise. The previous image degradation models are convenient but too simple. It only considers single degradation types and cannot completely include all degradation types of real images. Therefore, we adopt a random select/combine strategy to arrange and combine the types and order of these degradation factors. It substantially enlarges the degradation space, which can improve the matching degree between the designed and actual degradation models.

As shown in Figure 3, we first randomly select the types of blur, noise, and downsampling, then combine the three selected degradation factors in a random order. In this way, 36 degradation methods are formed, and the training can no longer be limited to the previous single degradation model. A model closest to the actual degradation process can be selected through training. After the designed degradation operation in this paper, we can be obtain the aligned LR/HR medical image pairs. Medical images are inconvenient to obtain due to the privacy of patients, and there are fewer paired datasets. It is possible to effectively solve the problem of unpaired image datasets by using this method. At the same time, the random select/combine strategy can combine more degradation types, so the proposed degradation model can restore the uncommon degradation process in natural scenes. It is suitable not only for medical images but also for natural images, which improves the generalization ability of our proposed model.

3.4 | Loss function and network structure

The mean squared error (MSE) loss [19] is an optimization scheme for obtaining higher peak signal-to-noise ratio (PSNR) in generative adversarial SR networks. It reflects the mean square error of the original image and the reconstructed image in terms of pixels. The loss is smaller when the reconstructed image is more similar to the original image. Equation (6) is the mathematical definition of MSE loss.

$$L_{MSE}^{SR} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} \left(I_{x,y}^{HR} - G(I_{x,y}^{LR})\right)^2 \tag{6}$$

where I^{LR} represents the LR image of size $H \times W \times C$, I^{HR} is the HR image of size $rH \times rW \times rC$,G stands for generator network, and x, y denote the pixel coordinates. We can get a relatively high PSNR value by calculating the mean square pixel error, but the output image is very blurred due to insufficient high-frequency information. Therefore, we use the VGG perceptual loss to avoid this problem [27]. As shown in

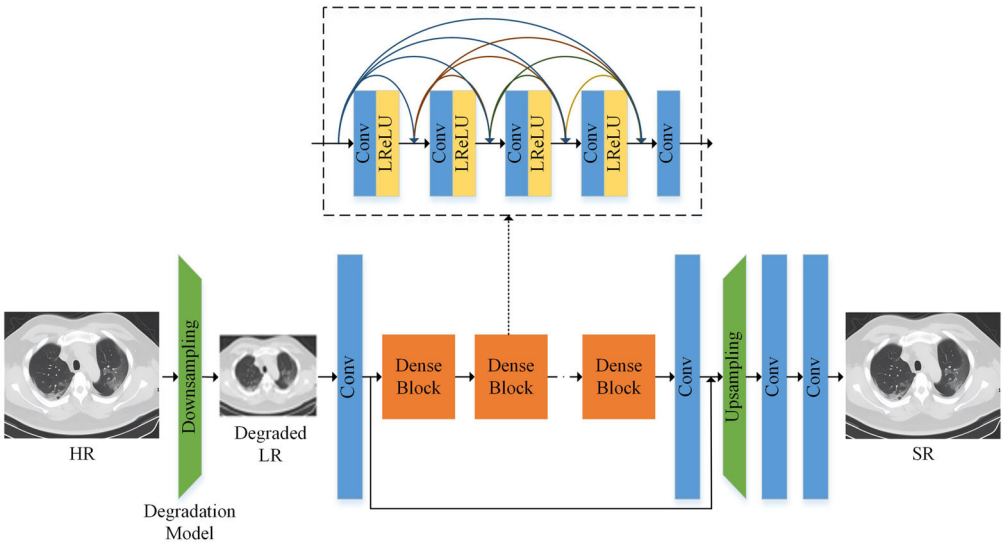

FIGURE 4 The structure of Med-BSR. Med-BSR, medical image blind super-resolution.

Equation (7), by calculating the Euclidean distance between the features of high- and low-resolution images, it can effectively maintain the perception quality of the reconstructed images.

$$LR_{VGG}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} \left(\phi_{i,j}\left(I^{HR}\right) - \phi_{i,j}\left(G\left(I^{LR}\right)\right)\right)^2 \tag{7}$$

where $\phi_{i,j}(x)$ represents the activation value obtained by the jth convolution before the ith maxpooling layer within the fine-tuned VGG network, $G(I^{LR})$ denotes the reconstructed image, $W_{i,j}$ and $H_{i,j}$ describe the respective feature maps size in the VGG network.

Therefore, we train two models in this paper. One is an objective evaluation model, which uses the MSE loss function to increase the objective evaluation value of the reconstructed image. The other is a sensory evaluation model using a combined loss function, which optimizes the L1, VGG perceptual and PatchGAN loss [25] with a loss combination coefficient of 1, 1, and 0.1. In practical applications, we prefer to use the VGG perceptual loss that can obtain a better visual experience.

The designed degradation model is added before the basic ESRGAN structure to form a medical image Med-BSR as demonstrated in Figure 4. The Dense Block (DB) can improve the image reconstruction quality by obtaining richer detailed information. The part inside the dotted frame is its internal structure. DB removes the most commonly used batch normalization layer [35], reducing computational complexity but not degrading network performance. We first degrade the input medical image through the model designed in Section 3.2. The generated LR image and the original HR image form an aligned HR/LR image pair input to the GAN. Finally, the reconstructed SR image is obtained after an upsampling layer and two convolutional layers.

4 | EXPERIMENTS

4.1 | Experimental setup

4.1.1 | Data and trainging details

We chose two different types of medical image datasets to test our model. One is a public dataset of brain tumour MRI images provided by the International Association for Medical Image Computing and Computer-Assisted Intervention [36], denoted as Dataset1. We selected 3000 MRI images with a transparent texture of 240 × 240 pixels, 2500 as the training set, and 500 as the test set. The other group is a lung CT image dataset containing 216 COVID-19 cases, marked as Dataset 2. We selected 300 images with 287 × 202 pixels as the training set and 50 images with 287 × 202 pixels as the test set.

We used the Adam optimizer to train our model, where the first-order momentum term $\beta_1 = 0.9$ and the second-order momentum term $\beta_2 = 0.99$. Training is performed in mini-batches with a batch_size of 32. Our model was trained for 200 epochs and the initial learning rate is set to 1×10^{-5}, then decreased to half every 1×10^5 iterations. The experiments are carried out on a computer equipped with two NVIDIA GeForce RTX3090 graphics processing unit by using the PyTorch1.6 deep learning framework.

4.1.2 | Evaluation metric

To objectively and directly compare the differences between different SR methods, this paper uses the PSNR and structural similarity (SSIM) [37] as evaluation metrics, which are the most

commonly used in the medical image SR field. PSNR is an image evaluation metric based on the sensitivity of inter pixel errors. The higher the value, the lower the distortion of the super-resolution image. The SSIM compares the similarity of two images in terms of structure, contrast, and brightness. The value range of SSIM is 0 to 1. The more significant the SSIM value, the better the super-resolution image quality. Here are their definitions

$$R_{PSNR} = 10\lg\left(\frac{255^2}{R_{MSE}}\right) \tag{8}$$

where R_{MSE} is the mean square error between the original and the SR image. The lower the MSE, the higher the PSNR.

$$R_{MSE} = \frac{1}{W \times H}\sum_{i=1}^{W}\sum_{y=1}^{H}\left(I_{i,j}^{HR} - I_{i,j}^{SR}\right)^2 \tag{9}$$

where I^{SR} represents the reconstructed SR image, $I_{i,j}$ is the pixel value of point (i, j).

$$R_{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{10}$$

where μ_x and μ_y represent the mean value of the input images x and y, respectively; σ_x and σ_y are the standard deviation of x and y, respectively; σ_{xy} is the covariance of the image x and image y; C_1 and C_2 are constants; we set them to be 0.01 and 0.03, respectively.

4.2 | Compared with SOTA

In order to verify the effect of our Med-BSR model, we carry out extensively comparative experiments with some state-of-the-art algorithms, such as Bicubic, SRCNN [12], fast super-resolution convolutional neural network (FSRCNN) [13], SRGAN [19], ESRGAN [20], DSRGAN [34], EDSR [18], SRMD [28], and IKC [23]. These methods are representative in the task of medical image SR and have strong correlation with our proposed method. SRCNN and FSRCNN adopt simple convolution neural networks.

SRGAN, ESRGAN, and DSRGAN are SR methods based on GAN. EDSR utilizes the residual structure without BN layers. SRMD and IKC are blind SR methods proposed in recent years, the former feeds both noise and blur to the network and the latter focuses on predicting blur kernels. As mentioned in Section 3.4, we designed two models based on ESRGAN. One is Med-BSR1 for objective evaluation and the other is Med-BSR2 aiming at perceptual quality. We set each network parameter optimal under the same experimental configuration to ensure the validity of the comparative experiments.

Table 1 shows the comparison results between our algorithm and other algorithms on Dataset1 and Dataset2. The experimental results with scale factors of ×2 and ×4 are given respectively. Whether MRI or CT images are input, the Med-BSR1 model exhibits a significant improvement. Compared with the traditional interpolation method Bicubic, the improvement of Med-BSR1 is particularly obvious. Because it has more parameters to learn, which is able to better capture the mapping of HR and LR images. At the same time, it can be found that the objective evaluation value of the three GAN-based methods SRGAN, ESRGAN, DSRGAN, and our designed Med-BSR2 is not very high. This is because these methods are not trained based on PSNR and SSIM but on perceptual loss and focus primarily on the recovery of detail in the image.

To demonstrate the perceptual quality intuitively, we selected some texture-rich images from the MRI and CT reconstructed images for partial magnification, and the magnified position has been marked with a white or black box. Figures 5, 6, 7, and 8 show the visual comparison of our algorithm with some other state-of-the-arts when the magnification factor is ×2 and ×4. As shown, the Med-BSR2 method exhibits an excellent visual effect, recovering edge texture and details very well. In terms of overall quality, Bicubic and SRCNN are of poor quality, which may be due to an insufficient network depth. Three GAN-based methods, SRGAN, ESRGAN, and DSRGAN, all have good

TABLE 1 The average objective evaluation of different SR algorithms on Dataset1 and Dataset2. The best result is in bold and the second best result is underlined.

		Dataset1		Dataset2	
Method	**Scale**	**PSNR**	**SSIM**	**PSNR**	**SSIM**
Bicubic	×2	33.3956	0.8814	29.3061	0.7937
SRCNN [12]	×2	35.4572	0.9081	31.5942	0.8089
FSRCNN [13]	×2	35.8735	0.9143	31.7834	0.8127
SRGAN [19]	×2	35.0125	0.9156	30.5438	0.8043
ESRGAN [20]	×2	34.6189	0.8966	29.8953	0.7968
DSRGAN [34]	×2	34.8357	0.9032	30.3596	0.8058
EDSR [18]	×2	36.8936	0.9352	32.7362	0.8268
SRMD [28]	×2	36.5402	0.9251	32.4017	0.8169
IKC [23]	×2	37.0 824	0.9 416	33.1747	0.8 221
Med-BSR1 (ours)	×2	**37.5794**	**0.9501**	**33.2375**	**0.8341**
Med-BSR2 (ours)	×2	36.8096	0.9335	32.6592	0.8194
Bicubic	×4	31.6437	0.8651	26.8745	0.7180
SRCNN [12]	×4	33.5248	0.8874	28.1047	0.7489
FSRCNN [13]	×4	33.7904	0.8903	28.4872	0.7517
SRGAN [19]	×4	32.8575	0.8846	27.7626	0.7354
ESRGAN [20]	×4	32.2951	0.8716	26.2904	0.7205
DSRGAN [34]	×4	32.5076	0.8787	27.5591	0.7276
EDSR [18]	×4	34.9335	0.9072	29.2023	0.7695
SRMD [28]	×4	34.3841	0.8942	29.1794	0.7586
IKC [23]	×4	34.8947	0.9 057	30.3217	0.77 62
Med-BSR1 (ours)	×4	**35.0916**	**0.9223**	**30.7985**	**0.7855**
Med-BSR2 (ours)	×4	34.7381	0.9017	29.6734	0.7724

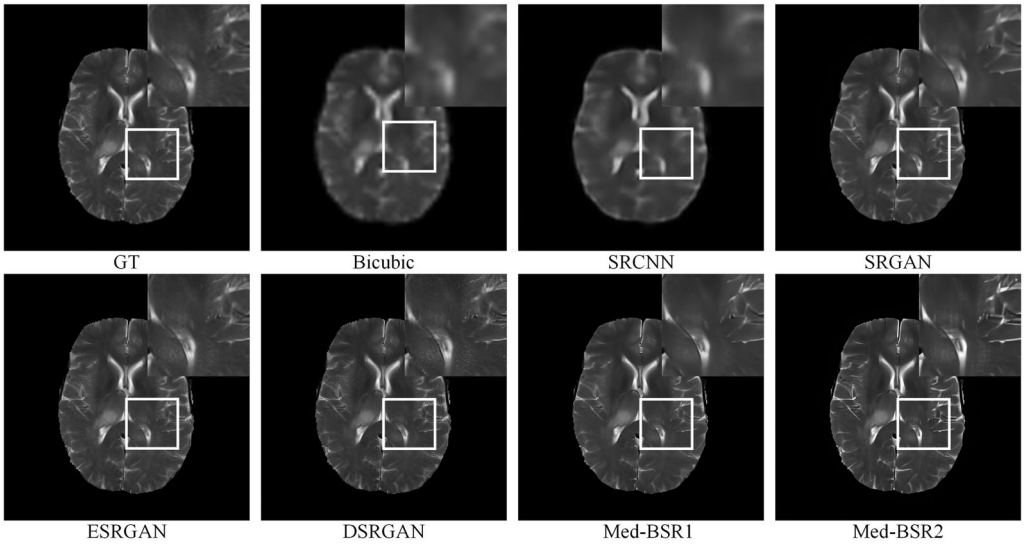

FIGURE 5 Comparison of ×2 reconstruction results of brain MRI images by different SR methods. MRI, magnetic resonance imaging.

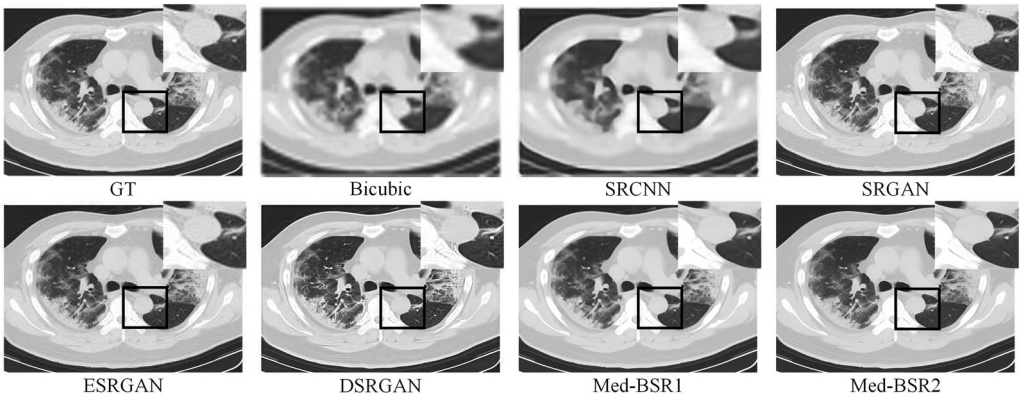

FIGURE 6 Comparison of ×2 reconstruction results of lung CT images by different SR methods. CT, computed tomography.

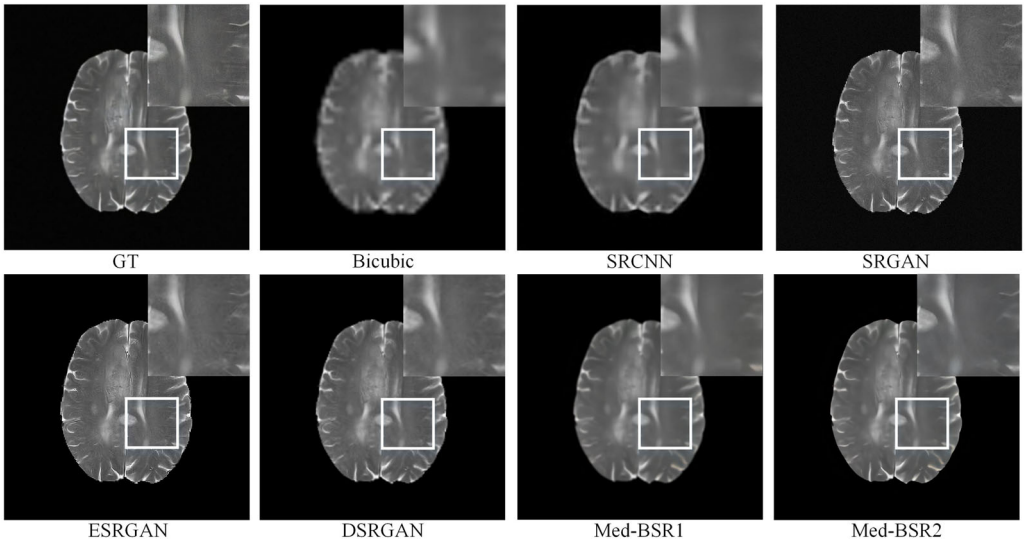

FIGURE 7 Comparison of ×4 reconstruction results of brain MRI images by different SR methods

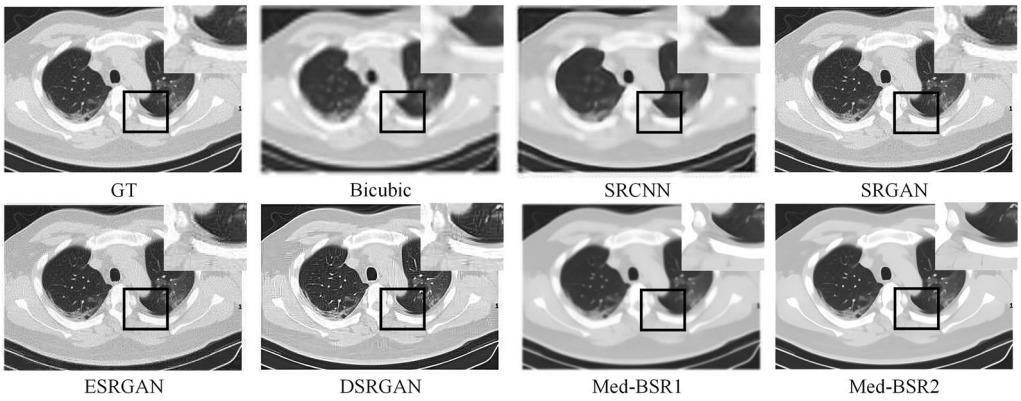

FIGURE 8 Comparison of ×4 reconstruction results of lung CT images by different SR methods

visual effects. The images reconstructed by our model Med-BSR 2 have precise contour and a pleasing visual experience. As for the details, take the 4× lung CT image shown in Figure 8 as an example. The SR images of Bicubic and SRCNN have many fictional artefacts that make people look very unpleasing. Although the images reconstructed by SRGAN and ESRGAN are less blurry, there are problems such as details loss and edge sharpening. The HR image reconstructed by DSRGAN is relatively more transparent but lacks texture information. Med-BSR2 outperformed the previous methods, and the obtained images have reasonable sharpness and rich texture, which is closest to the actual image.

Those above results illustrate that the model proposed by us can restore the actual degradation process of the HR image more accurately and is able to increase the objective index and perceptual quality of the reconstructed image significantly.

4.3 | Ablation study

We separately removed some components from the complete model and developed some variants of our degradation model Med-BSR for a better examination of the influences of each factor.

1. The degradation model A does not consider blur and the type of noise and downsampling remain the same.
2. The degradation model B does not introduce noise.
3. The degradation model C does not use multiple downsampling methods, and the other two degradation factors remain unchanged. We need to train with paired data, but Dataset1 and Dateset2 are both unpaired images. Moreover, paired datasets are rare in real life. So we must use the downsample operation to get aligned HR/LR image pairs, this section cannot remove the downsampling operation completely, and the degradation model C only uses the most commonly used bicubic interpolation method.

The training environment of the above three model variants is the same as that of the original model Med-BSR2. Table 2 shows the objective evaluation results, and the overall visual comparison is shown in Figures 9 and 10. Combined with the graph content, it can be seen that the objective and sensory performance of SR medical images decreases severely when the degradation model does not contain blur or noise. This phenomenon is more pronounced when the input is a CT image. The reconstruction results of the degradation model C with single downsampling are also not very satisfactory. Finally, our proposed Med-BSR2 with a triple degradation factor achieves the best reconstruction results, producing clear and detailed images. Thus, we can conclude that blur, noise, and downsampling can significantly impact the quality of medical super-resolution images, and any lack of one will deteriorate the reconstruction effect. Our designed degradation model contains rich degradation types and is beneficial for medical image SR tasks.

4.4 | Generalization analysis

To further test the generalization ability of the our designed Med-BSR model, we also selected 20 images of 160 × 160 pixels (denoted as Dataset3) from the natural image super-resolution dataset [26] for the network generalization test. The SR reconstruction results are shown in Table 3 and Figure 11. We can see from Table 3 that the Med-BSR1 model obtained the highest PSNR and SSIM values and achieved different degrees of improvement compared with other SR methods. As can be seen from the comparison of reconstruction effects in Figure 11, the SR natural image obtained by Med-BSR2 has accurate details, uniform brightness without blurring artefacts, which is the closest to the real image. This indicates that our Med-BSR model has good generalization ability, which is not only suitable for the MRI and CT images, but also has good performance in natural image SR tasks. Although we did not simulate the degradation process of natural images, the designed degradation model contains rich degradation forms. It can simulate the natural degradation process with a high probability and reconstruct SR natural images with an excellent objective and perceptual performance.

TABLE 2 The ablation study in our experiments. We used the Med-BSR2 for training, its relatively low objective performance and high sensory quality better reflect the accuracy and validity of the ablation experiment, where '√' means the factor is active and '×' indicates to eliminate it. The best result is in bold and the second best result is underlined.

Method	Blur	Noise	Multiple Downsampling	PSNR/SSIM	
				Dataset1	Dataset2
A	×	√	√	32.5237/0.8753	24.4083/0.7012
B	√	×	√	32.2045/0.8710	24.6958/0.7081
C	√	√	×	3 3.6144/ 0.888 2	2 6.1628/ 0.7 245
Ours	√	√	√	**34.1451/0.9025**	**28.4537/0.7784**

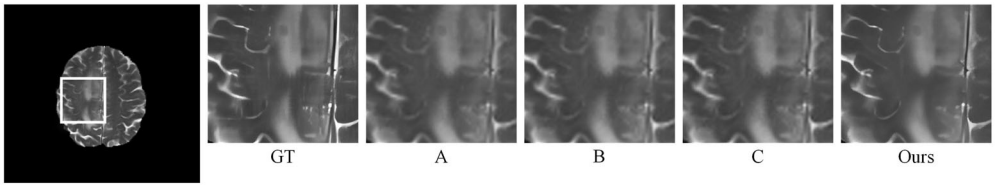

FIGURE 9 The SR visual comparisons of brain MRI image

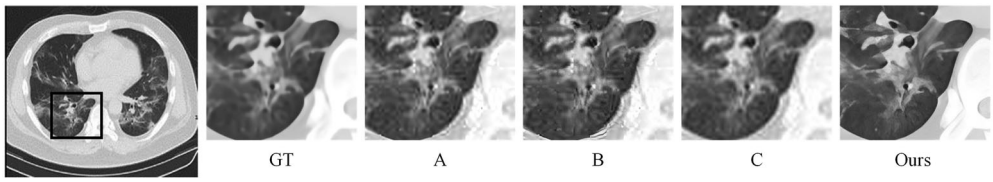

FIGURE 10 The SR visual comparisons of lung CT image

5 | CONCLUSION

In this paper, we propose a blind SR model based on an improved degradation process for medical images. Different from the existing SR methods, the proposed degradation model takes blur, noise, and downsampling as input. Specifically, the random select/combine strategy is used to arrange and combine each degradation factor, which greatly expands the degradation space. We also introduced VGG loss to further improve the overall perceived quality of SR images. Extensive experiments on brain MRI and lung CT images demonstrate the superiority of our designed Med-BSR model over the state-of-the-arts in

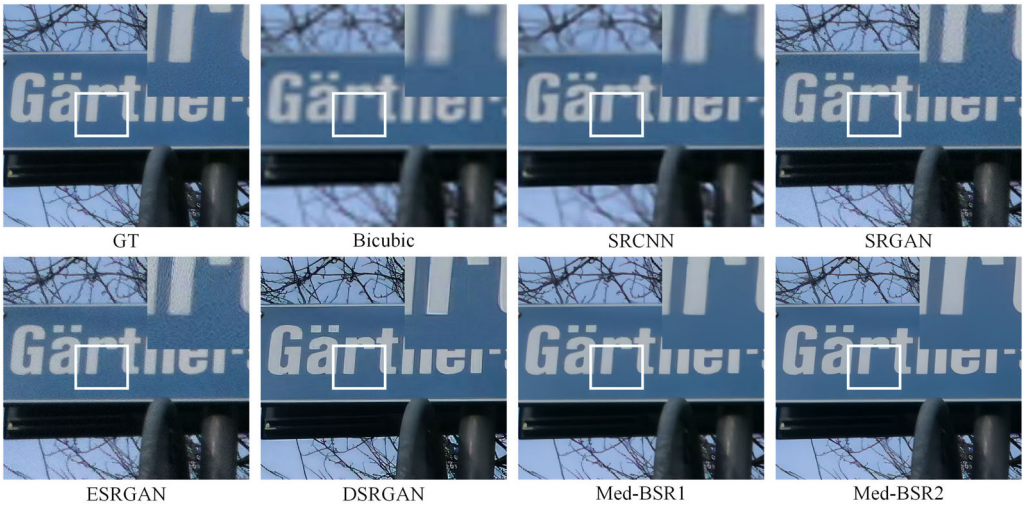

FIGURE 11 Comparison of ×4 reconstruction results of natural images by different SR methods

TABLE 3 The average objective evaluation of different SR algorithms on Datasets3. The best result is in bold and the second best result is underlined.

	×2		×4	
Method	**PSNR**	**SSIM**	**PSNR**	**SSIM**
Bicubic	30.0894	0.8554	27.5761	0.8346
SRCNN [12]	32.0926	0.8986	29.4573	0.8658
FSRCNN [13]	32.3407	0.8757	29.7832	0.8712
SRGAN [19]	31.7483	0.8721	28.9685	0.8654
ESRGAN [20]	31.1457	0.8647	28.0438	0.8527
DSRGAN [34]	31.4583	0.8695	28.4783	0.8635
EDSR [18]	33.7512	0.8915	30.6295	0.8836
SRMD [28]	32.5793	0.8823	29.9861	0.8772
IKC [23]	3 3.8916	0.9 032	3 1.6265	0.8 938
Med-BSR1 (ours)	**34.1278**	**0.9067**	**31.8654**	**0.8985**
Med-BSR2 (ours)	32.6729	0.8853	30.4071	0.8814

terms of objective evaluation and visual quality. In addition, the experimental results on natural images show that the Med-BSR model also has good generalization performance on different types of datasets. To sum up, the proposed method can restore the actual degradation process of medical image and provides a feasible solution for the complex degradation types in the actual scene.

AUTHOR CONTRIBUTIONS

Dangguo Shao: Methodology (lead); Writing—review and editing (equal).

ACKNOWLEDGEMENTS

We thank the Yunnan Provincial Key Laboratory of Artificial Intelligence and the First Affiliated Hospital of Kunming Medical University for providing equipment and medical image datasets. This work was supported by the National Science Foundation of China (Grant no. 11773012).

CONFLICT OF INTEREST

The authors declared that they have no conflicts of interest to this manuscript.

Li Qin: Software (lead); Writing—original and draft preparation (supporting).

Yan Xiang: Writing—review and editing (equal).

Lei Ma: Writing—review and editing (equal).

Hui Xu: Resources (equal); Conceptualization (supporting).

FUNDING INFORMATION

This work was supported by the National Science Foundation of China (Grant no. 11773012).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Dangguo Shao https://orcid.org/0000-0002-1222-9545

REFERENCES

1. Kavanagh, J., Liu, G., Menezes, R., et al.: Importance of long-term low-dose ct follow-up after negative findings at previous lung cancer screening. Radiology 289(1), 218–224 (2018)
2. Li, Y., Sixou, B., Peyrin, F.: A review of the deep learning methods for medical images super resolution problems. IRBM 42(2), 120–133 (2021)
3. Zhou, S.K., Greenspan, H., Davatzikos, C., et al.: A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises. Proc. IEEE 109(5), 820–838 (2021)
4. Zhang, H., Shinomiya, Y., Yoshida, S.: 3d mri reconstruction based on 2d generative adversarial network super-resolution. Sensors 21(9), 2978 (2021)
5. Liang, Z., He, X., Teng, Q., et al.: 3d MRI image super-resolution for brain combining rigid and large diffeomorphic registration. IET Image Proc. 11(12), 1291–1301 (2017)
6. Georgescu, M.I., Ionescu, R.T., Verga, N.: Convolutional neural networks with intermediate loss for 3d super-resolution of CT and MRI scans. IEEE Access 8, 49112–49124 (2020)
7. Liu, H., Liu, J., Hou, S., et al.: Perception consistency ultrasound image super-resolution via self-supervised cyclegan. Neural Comput. Appl. 1–11 (2021)
8. Reader, A.J., Corda, G., Mehranian, A., et al.: Deep learning for pet image reconstruction. IEEE Trans. Radiat. Plasma Med. Sci. 5(1), 1–25 (2020)
9. Song, T.A., Chowdhury, S.R., Yang, F., et al.: Pet image super-resolution using generative adversarial networks. Neural Netw. 125, 83–91 (2020)
10. Wang, Z., Chen, J., Hoi, S.C.: Deep learning for image super-resolution: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 43(10), 3365–3387 (2020)
11. Cai, Q., Li, J., Li, H., et al.: TDPN: Texture and detail-preserving network for single image super-resolution. IEEE Trans. Image Process. 31, 2375–2389 (2022)
12. Dong, C., Loy, C.C., He, K., et al.: Learning a deep convolutional network for image super-resolution. In: The European Conference on Computer Vision, Zurich, Switzerland, pp. 184–199 (2014)
13. Dong, C., Loy, C.C., Tang, X.: Accelerating the super-resolution convolutional neural network. In: The European Conference on Computer Vision. Amsterdam, The Netherlands, pp. 391–407 (2016)
14. Kim, J., Lee, J.K., Lee, K.M.: Accurate image super-resolution using very deep convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1646–1654 (2016)
15. Mahapatra, D., Bozorgtabar, B., Garnavi, R.: Image super-resolution using progressive generative adversarial networks for medical image analysis. Comput. Med. Imag. Graph. 71, 30–39 (2019)
16. Zhang, Y., Tian, Y., Kong, Y., et al.: Residual dense network for image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 2472–2481 (2018)
17. Li, H., Cen, Y., Liu, Y., et al.: Different input resolutions and arbitrary output resolution: A meta learning-based deep framework for infrared and visible image fusion. IEEE Trans. Image Process. 30, 4070–4083 (2021)
18. Lim, B., Son, S., Kim, H., et al.: Enhanced deep residual networks for single image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 136–144 (2017)
19. Ledig, C., Theis, L., Huszár, F., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 4681–4690 (2017)
20. Wang, X., Yu, K., Wu, S., et al.: Esrgan: Enhanced super-resolution generative adversarial networks. In: The European Conference on Computer Vision, Munich, Germany, (2018)
21. Huang, Y., Li, S., Wang, L., et al.: Unfolding the alternating optimization for blind super-resolution. Adv. Neural Inf. Process. Syst. 33, 5632–5643 (2020)
22. Bell Kligler, S., Shocher, A., Irani, M.: Blind super-resolution kernel estimation using an internal-gan. Adv. Neural Inf. Process. Syst. 32 (2019)

23. Gu, J., Lu, H., Zuo, W., et al.: Blind super-resolution with iterative kernel correction. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 1604–1613 (2019)
24. Wang, L., Wang, Y., Dong, X., et al.: Unsupervised degradation representation learning for blind super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 10581–10590 (2021)
25. Qiao, J., Song, H., Zhang, K., et al.: Image super-resolution using conditional generative adversarial network. IET Image Proc. 13(14), 2673–2679 (2019)
26. Zhang, K., Liang, J., Van Gool, L., et al.: Designing a practical degradation model for deep blind image super-resolution. In: IEEE/CVF International Conference on Computer Vision, Montreal, Canda, pp. 4791–4800 (2021)
27. Shocher, A., Cohen, N., Irani, M.: "Zero-shot" super-resolution using deep internal learning. In: IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 3118–3126 (2018)
28. Zhang, K., Zuo, W., Zhang, L.: Learning a single convolutional super-resolution network for multiple degradations. In: IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 3262–3271 (2018)
29. Yadav, R.B., Srivastava, S., Srivastava, R.: A partial differential equation-based general framework adapted to Rayleigh's, Rician's and Gaussian's distributed noise for restoration and enhancement of magnetic resonance image. J. Med. Phys. 41(4), 254–265 (2016)
30. Lin, H., Fan, J., Zhang, Y., et al.: Generative adversarial image super-resolution network for multiple degradations. IET Image Proc. 14(17), 4520–4527 (2020)
31. Nam, S., Hwang, Y., Matsushita, Y., et al.: A holistic approach to cross-channel image noise modeling and its application to image denoising. In: IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 1683–1691 (2016)
32. Ayhan, B., Kwan, C.: Mastcam image resolution enhancement with application to disparity map generation for stereo images with different resolutions. Sensors 19(16), 3526 (2019)
33. Haledyan, D., Amirany, A., Jafari, K., et al.: Low-cost implementation of bilinear and bicubic image interpolation for real-time image super-resolution. In: IEEE Global Humanitarian Technology Conference, Seattle, WA, USA, pp. 1–5 (2020)
34. Fritsche, M., Gu, S., Timofte, R.: Frequency separation for real-world super-resolution. In: IEEE/CVF International Conference on Computer Vision Workshop, Seoul, Repubic of Korea, pp. 3599–3608 (2019)
35. Nah, S., Hyun Kim, T., Mu Lee, K.: Deep multi-scale convolutional neural network for dynamic scene deblurring. In: IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 3883–3891 (2017)
36. Xiang, Y., Liu, H., Wang, S., et al.: Segmentation method of multiple sclerosis lesions based on 3D-CNN networks. IET Image Proc. 14(9), 1806–1812 (2020)
37. Wang, Z., Bovik, A.C., Sheikh, H.R., et al.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13(4), 600–612 (2004)

How to cite this article: Shao, D., Qin, L., Xiang, Y., Ma, L., Xu, H.: Medical image blind super-resolution based on improved degradation process. IET Image Process. 17, 1615–1625 (2023).
https://doi.org/10.1049/ipr2.12742