Signal Processing: Image Communication 75 (2019) 1–10

Contents lists available at ScienceDirect

Signal Processing: Image Communication

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

Towards efficient medical lesion image super-resolution based on deep residual networks☆

Sheng Ren [a,b], Deepak Kumar Jain [c], Kehua Guo [a,*], Tao Xu [a], Tao Chi [d]

[a] School of Computer Science and Engineering, Central South University, Changsha, China
[b] School of Computer and Electrical Engineering, Hunan University of Arts and Science, Changde, China
[c] Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing, China
[d] Key Laboratory of Fisheries Information, Ministry of Agriculture, Shanghai Ocean University, Shanghai, China

ARTICLE INFO

Keywords:
Super-resolution
Deep residual networks
Medical diagnosis

ABSTRACT

Super-resolution reconstruction helps doctors clearly observe the details of medical lesion images and increases the likelihood of the disease being diagnosed and cured. In this paper, we propose an efficient medical lesion image super-resolution method based on deep residual networks. First, a multi-scale super-resolution reconstruction model based on a deep residual network is trained. Second, an easy-to-use interface is designed. Third, a multi-scale super-resolution reconstruction model is used to reconstruct different types of medical lesion images with different scales and calculate their peak signal-to-noise ratios and structural similarity index values. The experimental results show that the proposed super-resolution reconstruction method achieves superior performance over the other methods compared in this work.

1. Introduction

In recent years, an increasing number of new technologies have been applied in the field of healthcare, such as heterogeneous big data processing, cyber–physical–social services, body area networks and wireless sensor networks with mobile sinks [1–4]. The new feature representation learning methods, such as deep learning, have yielded encouraging results [5]. With the development of deep learning, convolutional neural networks (CNNs) have been proven to be a powerful tool for computer vision [6]. A CNN can greatly reduce the number of parameters in the neural network and simultaneously reduce the complexity of the neural network model [7]. Based on a CNN, the classification of dermatological images generated by confocal laser scanning microscopy is almost as accurate as that performed by individual dermatologists [8]. In the field of super-resolution (SR), CNNs have significantly improved the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) index [9].

Medical imaging (X-ray imaging, CT imaging, magnetic resonance imaging, ultrasound imaging, nuclear medicine imaging, etc.) is widely used for medical examinations in varying departments such as internal medicine, surgery, ophthalmology, otolaryngology, and orthopaedics [10]. Medical imaging is not only an important diagnostic tool for doctors but also the basis for artificial intelligence medical assistant diagnosis systems. For example, Nature published a cancer-related research paper from Stanford University that used deep learning algorithms to classify skin cancer in photographs [11]. The medical imaging-assisted diagnostic systems developed by iFlytek, Tencent MI-AIS, and Ali Health can assist in diagnosing in specific medical fields (CT imaging of lungs, early screening of oesophageal cancer). These systems can reduce the workload of doctors and improve the accuracy of examinations. Whether it is the skin cancer classification algorithm based on medical images proposed by Stanford University or the medical imaging-assisted diagnostic systems developed by Tencent, Ali and iFlytek Company, the resolution of medical images must be very high. However, for some medical imaging techniques (MR images, CT images, etc.) it is difficult to obtain high-resolution images due to the imaging environment, limitations of physical imaging systems and quality constraints. Doctors are limited by low-resolution images and cannot clearly observe the details of lesions in medical images, which is not conducive to the accurate diagnosis and treatment of diseases. Reconstructing SR images from low-resolution (LR) medical images can effectively help doctors observe the details of medical lesion images and increase the possibilities of disease confirmation and cure. Therefore, SR reconstruction is of great significance to medical image processing.

SR reconstruction is an effective method for improving image resolution, and the current research is mainly focused on single image

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.image.2019.03.008.

* Corresponding author.
E-mail address: guokehua@csu.edu.cn (K. Guo).

https://doi.org/10.1016/j.image.2019.03.008
Received 21 October 2018; Received in revised form 11 March 2019; Accepted 13 March 2019
Available online 18 March 2019

super-resolution (SISR). Early SR methods were mainly based on image interpolation methods such as zero-order interpolation, bilinear interpolation and bicubic interpolation. These methods are simple and efficient but do not perform well in the non-uniform grey areas of an image and produce blurred images. Before CNNs were applied to SR reconstruction, the most advanced SISR methods were mostly based on examples, such as the sparse-coding based SR method (SCSR) [12]. These SR methods require much pre/post-processing and focus on learning and optimizing dictionaries or alternative ways of modelling them. However, the rest of the steps in the process have rarely been optimized or considered in a unified optimization framework [13]. For the first time, super-resolution convolutional neural networks (SRCNNs) were applied to CNNs in the field of SR reconstruction in the work of [14], and SRCNNs surpassed the example-based methods and demonstrated the power of CNNs in the field of SR. However, SRCNNs rely on the context of small image regions, and training converges too slowly. Therefore, the quality of the SRCNN reconstructed images is not sufficiently high, and details such as edge contour are not clear enough. The edge contour details of images are the key to medical lesion image SR reconstruction, which is an important basis for doctors' diagnosis. Some SR methods proposed in recent years, such as super-resolution using very deep convolutional networks (VDSR) and deeply-recursive convolutional network (DRCN), are mainly used in the reconstruction of ordinary images but are not often used in the reconstruction of medical lesion images [9,15].

To solve the above problems, we propose a medical lesion image SR method based on deep residual networks (MDRN). An MDRN can construct a deeper level to obtain larger receptive fields and extract more features. The introduction of residual learning in the MDRN enables faster convergence and thus higher execution efficiency. The SR images produced by an MDRN have clear outline details, and the PSNRs and SSIM index values are greatly improved compared with other methods. The innovations of MDRNs are as follows: (1) The application of a deep residual network in the SR reconstruction of medical images not only improves the quality of reconstruction but also improves the efficiency of reconstruction. (2) On the basis of single-scale models, a smaller and more stable multi-scale model is constructed, and a simple and easy-to-use SR reconstruction interface is designed for doctors.

The rest of this paper is organized as follows: Section 2 explains the related works on SR reconstruction, Section 3 gives an overview of the work process, Section 4 describes the actual realization process for medical lesion SR images based on deep residual networks, and Section 5 provides experimental results. Section 6 concludes our work and states our plans for future work.

2. Related works

In recent years, SR methods have been widely used in medical image reconstruction. To obtain high-resolution images from noisy LR medical images, D.H. Trinh et al. proposed an example-based SR and de-noising method for medical images [16], which used high- and low-resolution matching databases to model sparse representations of high- and low-resolution images using non-negative quadratic programming and estimated high-resolution images from a single noise LR image. This method requires a large amount of preprocessing, such as preparing standard images (good quality, taken of the same organ as the given LR image, with the same medical imaging modality) and constructing databases between HR and LR patch pairs. Due to the lack of back propagation optimization, the quality of SR reconstruction is not sufficient. To eliminate the negative effects of low-resolution CT images, reference [17] proposed an SR reconstruction method for medical images based on adaptive quadtree decomposition, which used a K-means clustering algorithm to determine the clustering centre to construct the mapping between LR image blocks and high-resolution image blocks and then reconstructed high-resolution images with a mapping function. The time complexity of this method is high, and it is does not easily converge. Selecting an appropriate K value is also difficult. Because the result of the K-means algorithm is not necessarily globally optimal, the reconstruction quality of this method is not high enough. Chang Liu et al. proposed a medical image SR technique via a minimum error regression model selection using a random forest algorithm [18]. The training samples are clustered iteratively, and then a series of typical regression models are trained. Clustered samples are used to train a random forest model, which then chooses the best regression model for reconstruction. This method requires training a series of regression models that consume a large amount of storage space. Random forest algorithms may not produce good choices for small data sets or data with fewer features, and over-fitting may occur for regression problems with high noise levels.

In addition to the special medical image SR methods mentioned above, there are also some general SR methods, such as example-based methods (SCSR) and CNN-based methods (SRCNN, VDSR, etc.) [19]. At present, most advanced SR methods are based on CNNs. Because of the widespread adoption and development of processing power and big data technologies [20], the effectiveness of deep convolutional neural networks has significantly improved the quality of SR reconstruction. The Chinese University of Hong Kong first proposed an SR reconstruction method based on deep learning named SRCNN [13]. An SRCNN has a three-layer structure, which is used for the extraction of LR image patches, nonlinear mapping and SR image reconstruction. The shallow structure of an SRCNN enables it to rely on the context of small image regions. The interpolated image is used as an input into the network, which results in a large amount of computations and a slow convergence rate. In addition, an SRCNN constructs only a single-scale model, which cannot meet the needs of multi-scale SR reconstruction. J. Kim et al. improved SRCNNs and proposed a very deep convolution neural network SR method called VDSR [15]. For the first time, a residual network was used to train a 20-layer SR reconstruction model, which has a larger receptive field, faster convergence speed and multi-scale amplification, and PSNRs and SSIM index values are also better than those achieved with an SRCNN. VDSR uses interpolated images as input, which increases computational complexity and reduces execution efficiency. The low-frequency part of the SR image is only processed by a one-time interpolation method, and the quality of reconstruction is not high enough. The high-frequency part of the SR image is processed by a 20-layer residual network; however, the depth is not sufficient and the reconstruction effect of some detailed features is not good. The residual network improves SR. Based on the residual network (ResNet), C. Ledig et al. proposed a photorealistic single image SR using a generative adversarial network [19]. The core of the SR method based on residual network (SRResNet) is a deep residual network (DRN) composed of 16 residual blocks, each of which contains 2 convolutional layers, 2 batch normalization (BN) layers and 1 parametric rectified linear unit (PReLU) layer. The DRN processes the residual image, and the output layer adds the residual image and the LR image processed by the interpolation method to generate the SR image, which optimizes the performance of the SR image. This method directly uses the original ResNet structure, but the BN layer in ResNet is suitable for image classification and not for SR reconstruction [21]. Since batch normalization layers normalize the features, the range flexibility is eliminated from networks. On the other hand, the BN layer consumes a large amount of computer capacity and reduces the efficiency of reconstruction.

In this paper, we apply a deep residual network to the SR of medical lesion images. MDRNs improved performance by removing the BN layer in the residual block. The SR model includes 80 residual blocks, and each layer uses 64 filters to extract features to reconstruct the residual images. This deep structure enables the network to learn enough features to achieve high-quality SR reconstruction. To reduce computational complexity and improve the speed of medical image SR reconstruction, we do not use an interpolation method to preprocess LR images and instead directly input LR images into the model. This

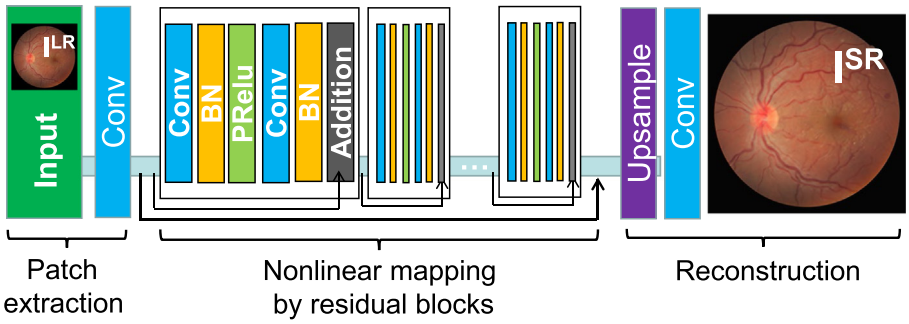

Fig. 1. The architecture of SRResNet.

approach can reduce a large number of computations without reducing the capacity of the model [14]. Based on the single scale model, MDRN can reconstruct multi-scale SR images by adding different magnification preprocessing and upsampling modules before and after the residual network. Details such as contours and edges of images reconstructed by the MDRN are clearly visible. The PSNR and SSIM are also improved compared with previous methods. The MDRN reconstructed SR medical image details are clearer, which is of great significance to doctors who diagnose diseases via imaging.

3. Preliminary

3.1. Deep residual networks

SR reconstruction is based on the input LR image, I^{LR}, which is used to estimate its SR image, I^{SR}. The image I^{LR} is an LR image corresponding to the high-resolution image I^{HR}. I^{HR} is used only during the training phase. The residual image is defined as $I^R = I^{HR} - I^{LR}$. Given a training set $\{I_i^{LR}, I_i^R\}_{i=1}^N$ with N as the number of images in the training set, our goal is to learn the model f to predict $\hat{y} = f(x)$ values, where $\hat{y}$ is the estimation of the target I^R image. We minimize the loss function L to predict the residual image.

$$L = \frac{1}{N}\sum_{i=1}^{N}\left|I_i^R - f(x)\right|. \tag{1}$$

SRResNet applies a residual network to SR reconstruction and achieves good results. Its network structure is shown in Fig. 1. An enhanced deep super-resolution network (EDSR) model improves performance by stacking more layers and increasing the number of filters, but it cannot increase the number of layers and filters indefinitely because too many levels and filters can cause the training process to become unstable [22]. The EDSR benchmark model has 16 layers, with 64 filters in each layer. The EDSR final model has 32 layers and 256 filters in each layer [22]. The final model structure is shown in Fig. 2. The MDRN model has 80 layers and 64 filters in each layer. Because the BN layer will normalize the colour distribution of the image, destroying the original contrast information of the image, the MDRN residual block removes the BN layer, which is not conducive to SR reconstruction. The MDRN model with simplified residual blocks not only saves approximately 40% of the computing resources but also improves SR reconstruction by stacking more residual blocks and increasing the number of filters.

The EDSR benchmark model occupies approximately $O(16 \times 64)$ of memory, and the final model occupies approximately $O(32 \times 256)$ of memory [22]. The MDRN occupies approximately $O(80{\times}64)$ of memory. To reduce computational complexity, the MDRN model uses LR medical lesion images as inputs. The MDRN model has 80 layers, and the deep residual network structure allows the model to have a larger receptive field to achieve high-quality SR reconstruction. Residual learning enables the network to converge faster and achieve high-efficiency SR reconstruction of medical images while maintaining depth.

3.2. Work process overview

Medical images of lesions, such as those obtained by gastroscopy, CT, and MR, are an important basis for doctors to diagnose diseases. The amount of medical image data is growing rapidly year by year. Some medical lesion images have a lower resolution due to factors such as the imaging environment, the limitations of the physical imaging systems, and quality constraints. To make an accurate diagnosis, doctors often need to magnify the image to observe the details of the lesion. If the I^{LR} is magnified directly, as shown in Fig. 3(b), the medical image will appear blurry, and the edge contours are not recognizable, which is not conducive to formulating a diagnosis. SR reconstruction can reconstruct the I^{SR}, with clear edges while enlarging it as well, as shown in Fig. 3(c). This level of image clarity is important for doctors to appropriately evaluate a condition and diagnose the disease.

The SR reconstruction process of medical lesion images is summarized in Fig. 4. To simplify the process for doctors, we encapsulate the SR reconstruction model and design an easy-to-use interface with Python. Doctors only need to operate a mouse to achieve SR reconstruction and see clearly enlarged medical lesion images.

The process for medical image SR reconstruction is divided into three steps. In the first step, the doctor selects an image that requires SR reconstruction and sets the magnification scale. The medical conditions of patients are typically complex, and some images may have lower resolution due to factors such as an inadequate imaging environment. To accurately determine the patient's condition, it is often necessary to magnify medical lesion images to observe their details. In the first step, based on their expertise, doctors select a low-resolution images, I^{LR}, requiring SR reconstruction. In the second step, the low-resolution image I^{LR} is input into the MDRN model for SR reconstruction. First, the I^{LR} is processed by a preprocessing module with different magnification scales, and then a residual image, I^R, is generated by the DRN. Finally, the I^{LR} and I^R are combined, and the SR image I^{SR} is sampled and output. In the third step, the doctors view or save the SR image. The image I^{SR} can be viewed directly by the doctors or saved in a specified folder.

4. Enhanced deep residual networks for medical lesion image SR

4.1. The role of ResNet in medical lesion image SR

In SR, a deeper network level often means a better reconstruction result. Deep convolution neural networks tend to cause gradients to disappear or explode with an increase in network depth. Szegedy proposed the BN structure, which can maintain a stable gradient during backpropagation that will not be too large or too small. However, increasing the BN layer still cannot make the network converge rapidly, and when the convolutional neural network reaches a certain depth, the accuracy becomes saturated and declines rapidly. K. He proposed a residual network (ResNet) in 2015 [23], which solved the problems of slow convergence and declining accuracy. The SR model based on ResNet also performs better than the traditional SR model.

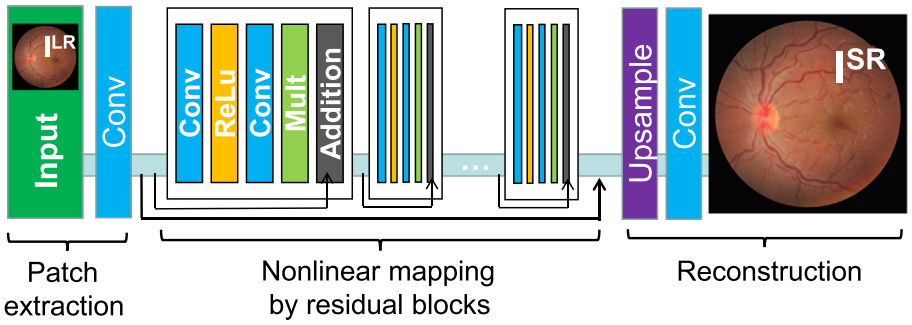

Fig. 2. The architecture of the final EDSR model.

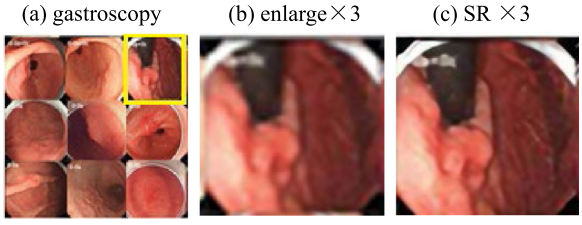

Fig. 3. A gastroscopy image at ×3 magnification and ×3 SR-reconstructed.

ResNet uses shortcut connections to connect a low-resolution input image, I^{LR}, directly to the SR output image I^{SR}. As shown in Fig. 5, the structure of the ResNet residual block is divided into two layers [23], which are defined as:

$$I^{SR} = F\left(I^{LR}, \{w_i\}\right) + I^{LR}. \tag{2}$$

Here, I^{SR} and I^{LR} are the output and input vectors of the layers considered. The function $F\left(I^{LR}, \{w_i\}\right)$ represents the residual mapping to be learned [23]. After passing image I^{LR}, each module learns only the residual function $F\left(I^{LR}\right)$, and the network is stable and learns easily. With the increase in network depth, optimizing the residual function $F\left(I^{LR}\right) = H\left(I^{LR}\right) - I^{LR}$ is simplified into optimizing the complex I^{SR} nonlinear mapping $H\left(I^{LR}\right)$.

ResNet has two architectures that are designed for shallow networks (ResNet34) and deep networks (ResNet50/101/152) [23]. In the shallow network residual block structure, two layers are the same size (3×3×64). To reduce the calculations and parameters, the residual block structure of the deep network is changed into three layers, and the intermediate layer size is consistent with that of the shallow residual block. The convolution kernel of 1×1 is used in the first and second layers. ResNet can be constructed with a deeper network structure, which is very suitable for SR reconstruction of medical lesion images. A deep residual network can learn more detailed features and reconstruct SR images with more details, such as edge contours, which are more representative. Residual learning greatly reduces the computational complexity of SR and improves the reconstruction efficiency while deepening the network.

4.2. SR of medical lesion images

In this paper, we use the PSNR and SSIM index to evaluate the performance of SR reconstruction. Compared with SRCNNs, MDRNs not only improve the quality of SR reconstruction by deepening the network but also accelerate the speed of reconstruction.

4.2.1. SR evaluation criteria

The PSNR is an image quality evaluation method based on the error sensitivity between corresponding pixels. It is the most common and widely used objective evaluation measure of images [24,25]. X is defined as I^{SR}, and Y represents the ground truth image I^{HR}.

$$MSE = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - Y(i,j))^2, \tag{3}$$

$$PSNR = 10 \log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right). \tag{4}$$

The MSE is the mean square error between the high-resolution image I^{HR} and the SR image, I^{SR}. H and W are the height and width of the image, and n is the number of bits per pixel, which is generally 8, indicating that the grey level of the pixel is 256. The PSNR is measured in dB units, and the larger the value is, the smaller the distortion.

The SSIM index is used to measure the similarity of two images. It measures image similarity from three aspects: brightness, contrast and structure.

$$l(X,Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2\mu_Y^2 + C_1} \quad c(X,Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2\sigma_Y^2 + C_2} \quad s(X,Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3}, \tag{5}$$

$$SSIM(X,Y) = l(X,Y) \cdot c(X,Y) \cdot s(X,Y). \tag{6}$$

μ_X and μ_Y represent the mean values of image X and Y, respectively, and are used to estimate brightness. σ_X and σ_Y represent the variance in images X and Y, respectively, and are used to estimate contrast. σ_{XY} represents the covariance of images X and Y as a measure of the degree of structural similarity. C_1, C_2, and C_3 are constants. The SSIM index has become an important performance index and is widely used in SR reconstruction because of its excellent performance in objective image evaluation. Its value range is [−1,1]; the greater the value is, the smaller the distortion.

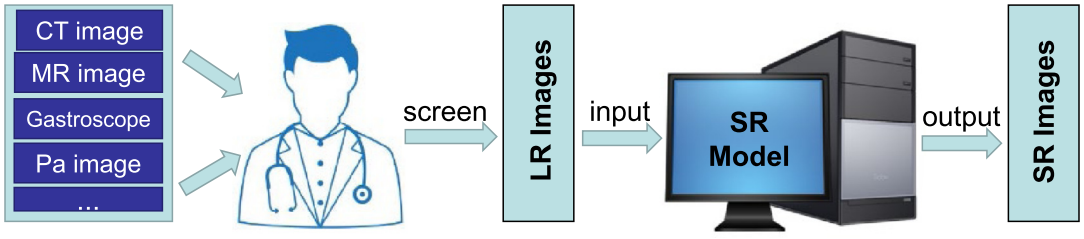

Fig. 4. Medical lesion image SR reconstruction process overview.

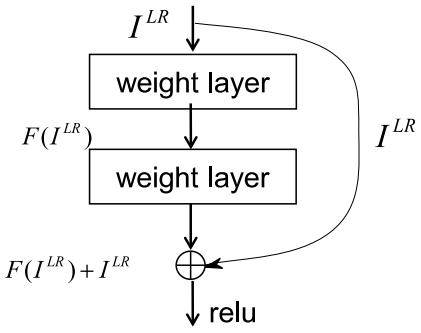

Fig. 5. Residual building block.

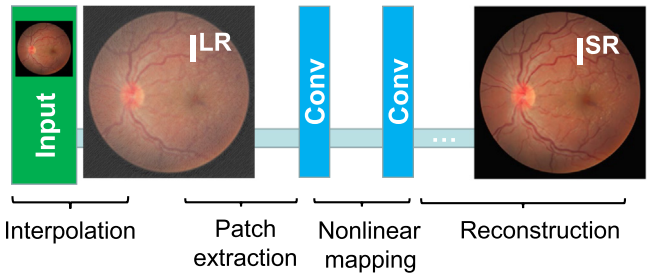

Fig. 6. The architecture of SRCNN.

4.2.2. Comparison of SR methods for medical lesion images

The SRCNN does not use ResNet; the input image I^{LR} is transferred layer by layer through the convolutional neural network. The network structure of the SRCNN is shown in Fig. 6. Dong et al. deepened the network; after a week of training, no better results were obtained. As shown in Fig. 6, the SRCNN has only three layers. This shallow structure makes the receptive field small and the model cannot learn enough reconstruction features, so the quality of the reconstructed SRCNN image is not high. On the other hand, the transmission of interpolated images in the network results in a large amount of computational effort and low execution efficiency.

ResNet can solve the problem of convergence and accuracy decline while deepening the network, so that the SR model based on ResNet can have a deeper level. The VDSR model residual block has 20 layers, and each layer includes a convolution layer and an ReLU layer [15]. The VDSR is approximately 0.86 dB higher than the PSNR produced by the SRCNN on the Set5 dataset (scale factor×4). The MDRN contains 80 layers, and each residual layer includes 2 convolutional layers and an ReLU layer [22]. The network structure of the MDRN is shown in Fig. 7. The MDRN produces a PSNR that is approximately 2 dB higher than that of the SRCNN on the Set5 dataset (scale factor×4). This figure also shows that depth of the MDRN is much higher than that of the SRCNN. The deep network structure has a larger receptive field and can learn abundant reconstruction features to improve the quality of SR reconstruction. Residual learning can effectively reduce the computational complexity and improve the reconstruction speed.

Before the low-resolution image I^{LR} is input into the network, the SRCNN uses an interpolation method (bicubic) to upsample it. This method is simple and easy to use and can effectively improve the PSNR and the SSIM index value of the SR image I^{SR}, but bicubic preprocessing requires many computations, which is not conducive to improving the efficiency of SR reconstruction. The SRCNN does not use ResNet, and the algorithm based on interpolation preprocessing can be divided into three steps: patch extraction and representation; nonlinear mapping; and reconstruction [13], as shown below. Formally, layer is expressed as an operation F, where W and B represent the filters and biases respectively.

Table 1
Average PSNR/SSIM for scale factors ×2, ×3, ×4 on datasets Set5, Set14, and Urban 100.

Dataset	Scale	Bicubic		A+		SRCNN		VDSR		MDRN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	×2	33.65	0.929	36.54	0.954	36.66	0.954	37.53	0.959	38.11	0.960
	×3	30.36	0.868	32.58	0.909	32.75	0.909	33.66	0.921	34.66	0.928
	×4	28.42	0.810	30.29	0.860	30.49	0.863	31.35	0.883	32.49	0.897
Set14	×2	30.29	0.869	32.33	0.906	32.35	0.905	32.98	0.912	33.86	0.920
	×3	27.55	0.774	29.13	0.819	29.28	0.821	29.77	0.831	30.42	0.845
	×4	26.01	0.703	27.39	0.750	27.53	0.750	28.01	0.769	28.72	0.786
Urban 100	×2	26.88	0.840	29.21	0.894	29.51	0.895	30.76	0.914	32.84	0.935
	×3	24.45	0.744	26.03	0.797	26.24	0.799	27.14	0.828	28.77	0.866
	×4	23.15	0.658	24.33	0.719	24.52	0.723	25.18	0.753	26.66	0.804

The MDRN directly inputs the low-resolution image I^{LR} and enlarges it according to the scale factor by adding an upsampling module at the end of the model. This method can improve the efficiency of SR reconstruction [22]. The MDRN uses ResNet because the residual values are mostly zeros or a very small value, so the computational burden is relatively small, and the network can stack layers deeper to learn more detailed features and improve the quality of SR reconstruction. The SR algorithm based on a deep residual network is shown below.

In Algorithm 1, we set the convolution kernels used in the three convolution layers to be divided into 9×9, 1×1 and 5×5. The number of output features of the first two layers is 64 and 32, respectively. In Algorithm 2, we set $t = 80$ and use 80 residual blocks to generate high-frequency information from the images. Compared with the SRCNN without a residual network, the MDRN can stack more layers to learn more detailed features, and the reconstructed SR image edge outlines are clearer. We tested the effectiveness of the DRN for SR reconstruction using several different data sets. As shown in Table 1, the PSNR and SSIM index values of the DRN based SR methods (VDSR and MDRN) are higher than those of other methods (Bicubic, A+, and SRCNN), and the MDRN has the highest PSNR and SSIM values.

The MDRN has 80 residual blocks, so the number of parameters is much higher than those in the SRCNN. Although the MDRN has deeper layers and more parameters, its execution efficiency is higher than that of the SRCNN due to the small number of residual calculations. In conclusion, compared with other methods, the MDRN has a higher SR reconstruction quality, faster reconstruction speed and offers multi-scale SR reconstruction, which is of great significance to doctors who diagnose diseases via medical imaging.

4.3. Multi-scale model for SR reconstruction of medical lesion images

Single-scale SR models need to train and store different scale-related models for different scale factors. At present, the SR training scale factors are the ×2, ×3, and ×4 models. It is impractical to train and store a model for all possible scenarios. The multi-scale SR model uses a single network model to achieve multi-scale SR reconstruction, which not only simplifies the training complexity but also saves memory space. SR reconstructions with the same scaling factor using a single-scale model and a multi-scale model are saved in the same data set [22], and the performance comparison is shown in Fig. 8.

The PSNR and SSIM index values of the single-scale model and the multi-scale model are presented in Fig. 8. The average difference between the PSNRs of the two models is 0.015, and the SSIM index average difference is 0.0027, indicating that the SR reconstruction performance of the two models is similar. To ensure the stability of the SR model, save storage space and adapt to the multi-scale amplification requirements of medical lesion images, our final model uses a multi-scale SR model.

Bicubic interpolation is applied to I^{LR} to upsample it before input into the network, and the SR model can be expanded during the training phase to address multi-scale SR reconstruction. The SR model

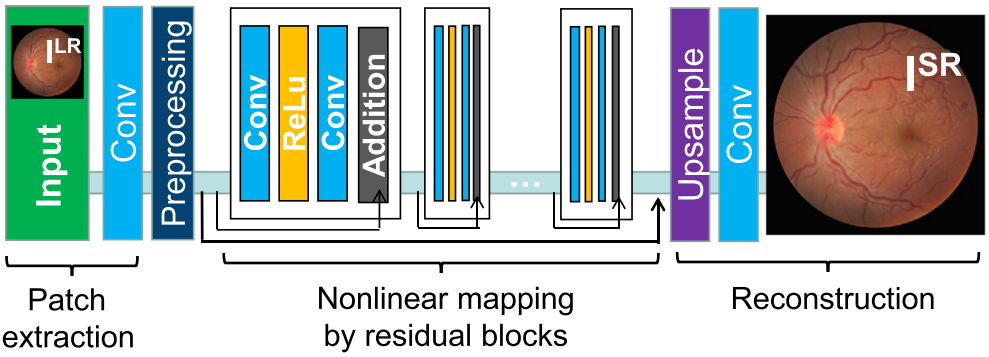

Fig. 7. The architecture of MDRN.

Algorithm 1. An SRCNN algorithm based on interpolation preprocessing

1	**input** I^{LR} # input LR image
2	**output** I^{SR} # output SR image
3	**define** I^{IT} **as** interpolation results
4	**require** $S \in (2, 3, 4)$ # scale factors
5	I^{IT} =upsample(I^{LR}, S, bicubic)
6	$F_1(I^{IT}) = \max(0,\ W_1 * I^{IT} + B_1)$ # patch extraction
7	$F_2(I^{IT}) = \max(0,\ W_2 * F_1(I^{IT}) + B_2)$ # nonlinear mapping
8	$F_3(I^{IT}) = W_3 * F_2(I^{IT}) + B_3)$ # reconstruction
9	$I^{SR} \leftarrow F_3(I^{IT})$
10	**return** I^{SR}

of the training scale factor ×2 can handle the SR reconstruction well with a × 2 scale in the test phase, but the results of ×3 and ×4 scales are not good. Expanding the training scale and training the scale factors of ×2, ×3, and ×4 in the test phase, the model can address the SR reconstruction of ×2, ×3, and ×4 scales well. Compared with the sum of multiple single-scale models with the same scale factor, the parameters of the multi-scale model are reduced by 29%, and the capacity is smaller and more stable [15].

The SR model of the I^{LR} input network cannot realize the multi-scale model directly by expanding the training scale. A single-scale SR model is the basis of realizing a multi-scale model. To train the multi-scale model, we first train the single-scale models ×2, ×3 and ×4 and use the relationships between the models to achieve multi-scale SR reconstruction. SR reconstruction at different scales is interrelated. In the SR model with a ×4 training scale, the training converges faster when the parameters are initialized by the pre-trained ×2 scale network. Using this correlation and adding different magnification preprocessing modules in front of the single-scale model to reduce the differences among input images with different magnifications [21], the upsampling modules with different magnifications are placed in parallel behind the MDRN, and the training scale is expanded to realize the multi-scale SR model.

There are many types of medical lesion images, so it is not appropriate to build an SR model for each image type. In this paper, different types of LR medical lesion images are processed into two parts: one consists of the regions where the grey value changes smoothly (low-frequency information), and the other consists of the regions where the grey value changes sharply (high-frequency information). The multi-scale model is used to reconstruct the SR medical image. The low-frequency information of the image is reconstructed by the interpolation method or by placing a sampling module at the end of the model, and the high-frequency information is reconstructed by the DRN. The two are combined to output the I^{SR}. The MDRN-based medical image SR reconstruction process is shown in Fig. 9.

The focus of SR reconstruction for medical imaging is to use an MDRN to reconstruct the high-frequency information of the image so that details such as the contour edges of the SR image are more clearly visible, enabling doctors to make an appropriate diagnosis.

5. Experimental results

At present, SR training data are usually 2D images; only a few methods utilize 3D data since the computational complexity of 3D data is very high [26]. The SRCNN uses a very large dataset, ImageNet, as the training dataset, and the VDSR uses a training dataset consisting of 291 pictures [27]. Medical lesion imaging involves patient privacy. Therefore, enhancing the security of the system to protect the privacy of

Algorithm 2. Residual algorithm without interpolation preprocessing

input I^{LR} # input LR image

output I^{SR} # output SR image

define residual block **as** $\Re$

define I^{TP} **as** intermediate results in residual blocks

require $S \in (2, 3, 4)$ . . # scale factors

$I^{TP} \leftarrow I^{LR}$

for i =1, . . .t **do** . # t is the number of residual blocks

$\Re_i \leftarrow \Re(F(I^{TP}, \{w_i\}))$ # execution residual block

$I^{TP} \leftarrow sum(I^{TP} + \Re_i)$ # elementwise sum

end for

$I^{SR} \leftarrow$ upsample(I^{LR}+$\Re_t$, S)

Return I^{SR};

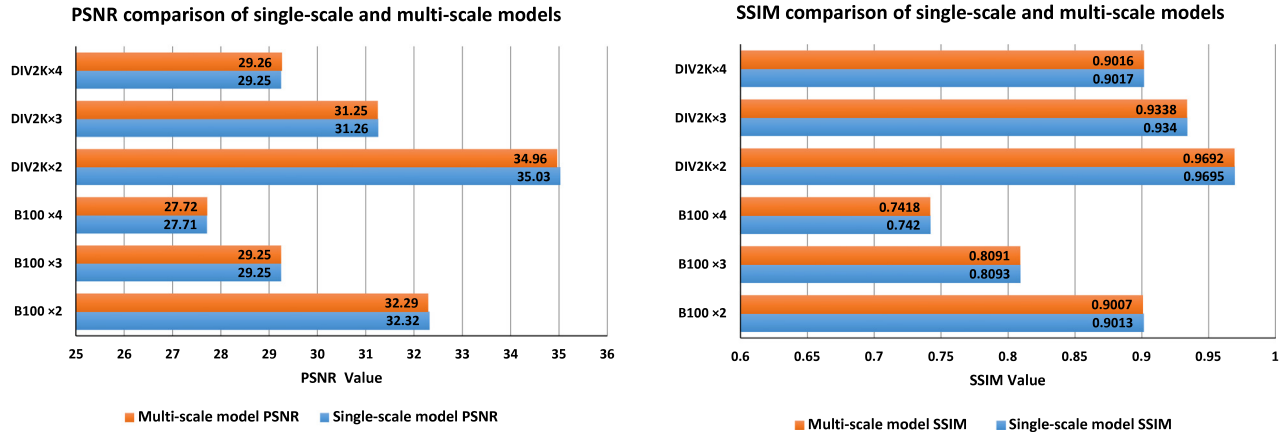

Fig. 8. Comparison between single-scale model and multi-scale model PSNR and SSIM.

sensitive data is of great importance [28,29]. In this paper, the DIV2K data as well as 200 other medical lesion images are used as training datasets [30]. DIV2K is a public dataset, and the 200 medical images are private data obtained from hospitals. The DIV2K dataset is the latest 2K resolution dataset proposed for image restoration and includes 800 training images, 100 verification images and 100 test images. We used four different types of medical lesion images: retinopathy, gastroscopy, chest X-ray and demodicosis, as test datasets. All baseline methods were obtained from the author's MATLAB implementation and run on Intel CPU 3.20 GHz, 8 GB memory machines. The MDRN is implemented using the PyTorch platform and runs on NVIDIA TITAN Xp GPUs.

5.1. MDRN-based SR reconstruction of multiple medical lesion images

Three types of medical lesion images, retinopathy, gastroscopy and chest X-ray, were reconstructed on a ×4 scale SR model based on the MDRN. The input I^{LR} is shown in Figs. 10–12, but the size is too small (120 x 120) to observe the details of the lesion and make a diagnosis. In SR reconstruction, the three types of medical lesion images are processed according to low-frequency and high-frequency information to eliminate the differences between the different medical lesion images. Figs. 10(b), 11(b) and 12(b) show the effect of ×4 scale upsampling. The low-frequency information of the image is sufficiently clear, but the edge contour is not. In Figs. 10(c), 11(c) and 12(c), the high-frequency information images are reconstructed by the MDRN. Detailed information such as the image edge contour is processed by the MDRN to generate a better effect. Figs. 10(d), 11(d) and 12(d) show the integration of low and high-frequency information in the image to the output I^{SR}.

As shown in Figs. 10–12, if the LR medical image is magnified directly, the image becomes blurred, and the details of the retinopathy,

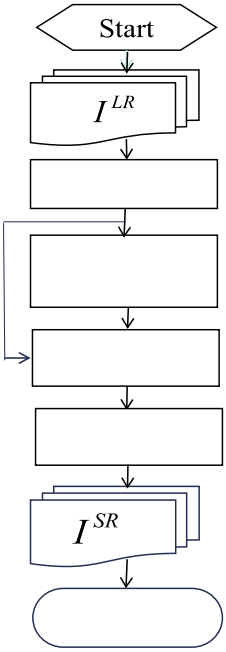

Fig. 9. The reconstruction process of medical image SR based on an MDRN.

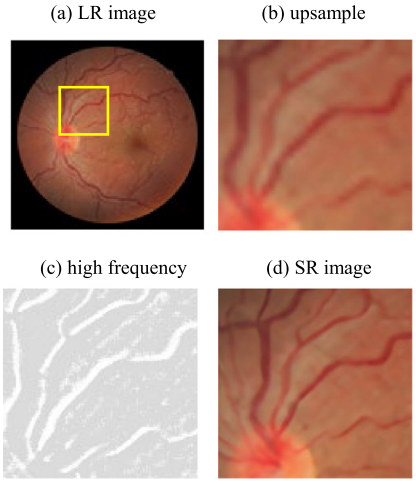

Fig. 10. A ×4 SR reconstruction of retinopathy.

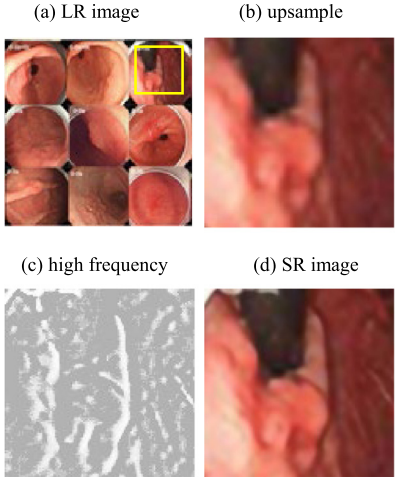

Fig. 11. A ×4 SR reconstruction of gastroscopy.

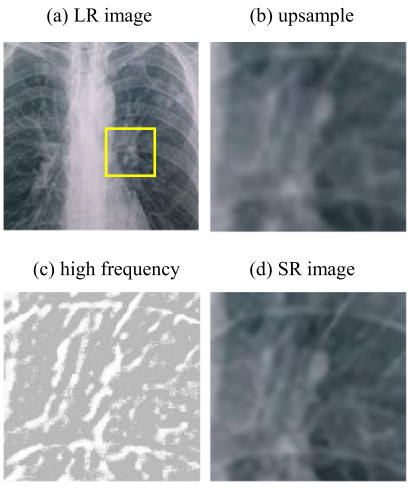

Fig. 12. A ×4 SR reconstruction of chest X-ray.

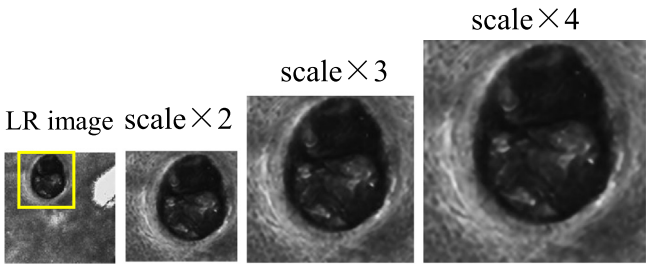

Fig. 13. Scale ×2 ×3 ×4 SR reconstruction of demodicosis.

gastroscopy and chest X-ray image cannot be recognized. Early symptoms of some serious diseases (e.g., gastric cancer) are not obvious. If the images are not sufficiently clear, doctors may not be able to identify the lesions in a timely manner, resulting in serious consequences. The MDRN reconstructed SR images, edge contours and other details are clearly visible, and some small lesions can also be observed in the images, which is of great significance for doctors who diagnose diseases via medical imaging.

5.2. SR reconstruction of multiple-scale medical lesion images

Doctors need a variety of magnification scales when viewing image data. As discussed in 4.3 of this paper, the multi-scale SR model has fewer parameters and uses less storage space, which is more suitable for medical image SR reconstruction. We used the MDRN to reconstruct the ×2, ×3, and ×4 scales of SR medical lesion images of demodicosis and retinopathy. As shown in Fig. 13, the medical image of demodicosis cannot be used for diagnosis because it is too small to determine whether there are any demodicosis in the hair follicle. The antennae and bodies of demodicosis are clearly visible after SR reconstruction, which is very important for the diagnosis. The size of the retinopathy image input in Fig. 14 is too small to clearly observe the details of the lesion. The details of the retinopathy after the SR reconstruction are clearly visible, which can effectively help the doctor make an appropriate diagnosis.

The PSNR and SSIM indexes are important performance indicators for SR reconstruction. As shown in Table 2, we calculated the PSNR and SSIM index values of retinopathy and demodicosis images using the MDRN at ×2, ×3, and ×4 scales.

The PSNR of the I^{SR} of retinopathy at the ×2 scale is 43.56 dB and the SSIM index is 0.9776, which are very close to the values for the I^{HR}. The lowest group of PSNR and SSIM index values are for the

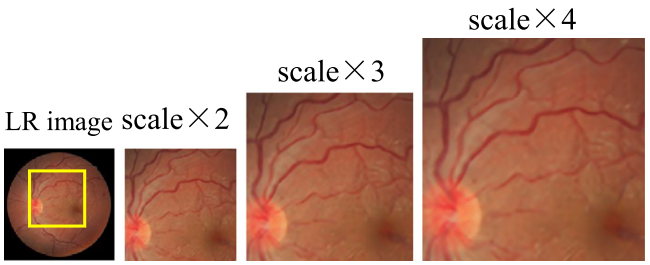

Fig. 14. Scale ×2 ×3 ×4 SR reconstruction of retinopathy.

Table 2
MDRN-based, PSNR(dB) and SSIM values of SR reconstruction at different scales.

Types of medical lesion images	Scale ×2		Scale ×3		Scale ×4	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Demodicosis	32.76	0.8889	29.95	0.7680	28.18	0.6528
Retinopathy	43.56	0.9776	41.16	0.9595	39.76	0.9451

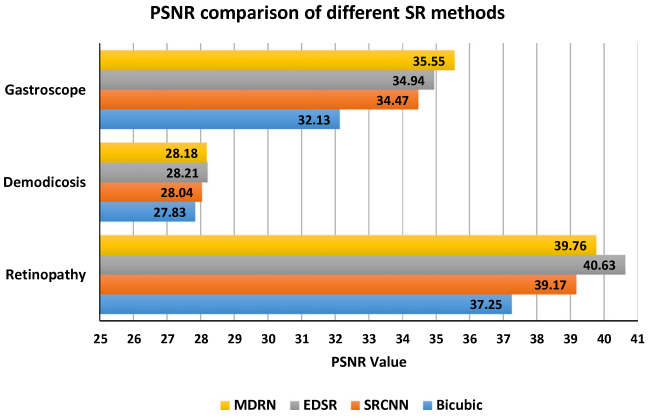

Fig. 15. PSNR comparison of different SR methods.

demodicosis ×4 scale SR image, I^{SR} (28.18 dB, 0.6528) This is because of the limitations of the instrument. The image of the demodicosis consists mostly of low-frequency information, and the image contrast is also low. In the SR image, I^{SR}, the demodicosis have a clear visual outline, which can provide strong support for doctors to diagnose the disease.

5.3. Comparison of SR methods for medical lesion images

The medical lesion image effects of different SR models are also different. We chose bicubic interpolation as the benchmark to compare the MDRN, EDSR and SRCNN. A demodicosis, retinopathy and gastroscopic image were selected from the dataset to perform ×4 scale SR reconstruction using different models. To detect the SR reconstruction effects of different models in three medical lesion images more intuitively, we calculated the PSNR and SSIM index of each I^{SR}. The performance comparison of the different models, using bar charts with values, is shown in Figs. 15 and 16.

As shown in Figs. 15 and 16, the performance of the SR methods based on the DRN (EDSR and the MDRN) is better than that of the SRCNN. The DRN can maintain a high level of execution efficiency while constructing a deep network. The deep network structure can learn an abundant amount of high-frequency information details and make a high-quality reconstructed image. The EDSR and MDRN have similar performance levels because both methods are based on deep residual networks, but the MDRN can also reconstruct multi-scale SR images. In the SR reconstruction of the gastroscopy image, the MDRN had the best performance, and the PSNR and SSIM index values were the highest

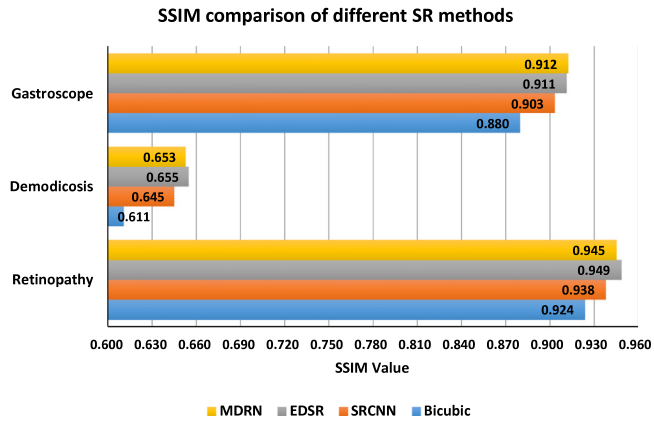

Fig. 16. SSIM comparison of different SR methods.

(35.55 dB, 0.912). The EDSR performed best in SR reconstruction of the retinopathy and demodicosis images; the PSNR and SSIM index values were (40.63 dB, 0.949) and (28.21 dB, 0.655), respectively. The retinopathy SR images obtained values of PSNR = 40.63 dB and SSIM index = 0.949, which are very close to the actual high-resolution image, showing the superiority of its performance.

6. Conclusion

Medical lesion imaging is an important basis for doctors to diagnose diseases. The resolution of some medical lesion images is not high due to issues such as the imaging environment, limitations of physical imaging systems and quality limitations. If one directly zooms in on an LR image, I^{LR}, the image will appear blurry, and the lesion will not be recognized. SR reconstruction can reconstruct the I^{LR} at different scales (usually ×2, ×3, ×4) and obtain the I^{SR} with clear details and contours. It can help doctors identify the details of the lesions clearly, thus improving the possibility of an accurate diagnosis and appropriate treatment of the disease. Therefore, medical lesion image SR reconstruction has a strong practical significance. In this paper, a medical lesion image SR reconstruction method based on a DRN is proposed. A single-scale model and a multi-scale model are trained using PyTorch to reconstruct an SR medical lesion image, I^{SR}. To simplify the doctor's operation, Python is used to design an easy-to-use operation interface. The doctor performs only simple mouse operations to reconstruct multiple medical lesion images at different scales. Several SR reconstruction models are used to reconstruct different types of medical lesion images. The experimental results show that the PSNRs and SSIM index values based on an MDRN are superior to other models, and in the reconstructed SR image I^{SR}, edge contours and other details are visibly clearer. Our next step is to study the medical image SR reconstruction on a large scale (×8, ×16). At the same time, we will explore the feasibility of SR reconstruction based on an MDRN in clinical practice.

Acknowledgements

This work is supported by the Natural Science Foundation of China (61672535, 61502540), National Science Foundation of Hunan Province, China (Nos. 2019JJ20025 and 2019JJ40406).

References

[1] K. Guo, Z. Liang, Y. Tang, et al., SOR: an optimized semantic ontology retrieval algorithm for heterogeneous multimedia big data, J. Comput. Sci. 28 (2018) 455–465.
[2] X. Wang, L.T. Yang, X. Xie, J. Jin, M. Jamal Dee, A cloud-edge computing framework for cyber-physical-social services, IEEE Commun. Mag. 55 (11) (2017) 80–85.

[3] W. Tang, K. Zhang, J. Ren, Y. Zhang, X. (Sherman) Shen, Flexible and efficient authenticated key agreement scheme for BANs based on physiological features, IEEE Trans. Mob. Comput. (2018) http://dx.doi.org/10.1109/TMC.2018.2848644.
[4] J. Wang, J. Cao, S. Ji, J. Hyuk Park, Energy efficient cluster-based dynamic routes adjustment approach for wireless sensor networks with mobile sinks, J. Supercomput. 73 (7) (2017) 3277–3290.
[5] W. Yang, Z. Wang, B. Zhang, Recognition using adaptive local ternary patterns method, Neurocomputing 213 (2016) 183–190.
[6] Y. Tu, Y. Lin, J. Wang, J. Kim, Semi-supervised learning with generative adversarial networks on digital signal modulation classification, Comput. Mater. Contin. 55 (2) (2018) 243–254.
[7] W. Fang, F. Zhang, Victor S. Sheng, Y. Ding, A method for improving CNN-based image recognition using DCGAN, Comput. Mater. Contin. 57 (1) (2018) 167–178.
[8] K. Guo, T. Li, R. Huang, et al., DDA: A deep neural network-based cognitive system for IoT-aided dermatosis discrimination, Ad. Hochsch. Netw. 80 (2018) 95–103.
[9] J. Kim, J. Kwon Lee, K.M. Lee, Deeply-recursive convolutional network for image super-resolution.
[10] K. Guo, Y. He, X. Kui, et al., LLTO: towards efficient lesion localization based on template occlusion strategy in intelligent diagnosis, Pattern Recognit. Lett. 116 (2018) 225–232.
[11] A. Esteva, B. Kuprel, RA Novoa, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[12] J. Yang, J. Wright, T. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.
[13] C. Dong, C.C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, in: ECCV, 2014, pp. 184–199.
[14] C. Dong, C.C. Loy, X. Tang, Accelerating the superresolution convolutional neural network, in: ECCV, 2016, pp. 391–407.
[15] J. Kim, J. Kwon Lee, Lee. K.M., Accurate image superresolution using very deep convolutional networks.
[16] D.H. Trinh, M. Luong, F. Dibos, J.M. Rocchisani, C.D. Pham, Novel example-based method for super-resolution and denoising of medical images, IEEE Trans. Image Process. 23 (2014) 1882–1895.
[17] J. Song, H. Liua, K. Deng, C. Zhang, Super resolution reconstruction of medical image based on adaptive quad-tree decomposition, J. Comput. Methods Sci. Eng. 17 (2017) 411–422.
[18] C. Liu, X. Wu, X. Yu, Y. Tang, J. Zhang, J. Zhou, Medical image super-resolution via minimum error regression model selection using random forest, Sustainable Cities Soc. 42 (2018) 1–12.
[19] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., Photo-realistic single image super-resolution using a generative adversarial network.
[20] X. Wang, L.T. Yang, L. Kuang, X. Liu, Q. Zhang, M. Jamal Deen, A tensor-based big-data-driven routing recommendation approach for heterogeneous networks, IEEE Netw. 33 (2019) 64–69.
[21] D. Zeng, Y. Dai, F. Li, R. Simon Sherratt, J. Wang, Adversarial learning for distant supervised relation extraction, Comput. Mater. Contin. 55 (1) (2018) 121–136.
[22] B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, Enhanced Deep Residual Networks for Single Image Super-Resolution.
[23] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[24] A. Hore, D. Ziou, Image quality metrics: PSNR vs SSIM, in: International Conference on Pattern Recognition, IEEE, 2010, pp. 2366–2369.
[25] S.S. Channappayya, A.C. Bovik, R.W. Heath, C. Caramanis, Rate bounds on ssim index of quantized image DCT coefficients, IEEE Trans. Image Process. 17 (2008) 1624–1639.
[26] W. Yang, C. Sun, W. Zheng, K. Ricanek, Gender classification using 3D statistical models, Multimedia Tools Appl. 76 (3) (2017) 4491–4503.
[27] S. Schulter, C. Leistner, H. Bischof, Fast and accurate image upscaling with super-resolution forests.
[28] J. Ren, H. Guo, C. Xu, Y. Zhang, Serving at the edge: A scalable IoT architecture based on transparent computing, IEEE Netw. 31 (5) (2017) 96–105.
[29] C. Yin, J. Xi, R. Sun, J. Wang, Location privacy protection based on differential privacy strategy for big data in industrial internet of things, IEEE Trans. Ind. Inf. 14 (8) (2018) 3628–3636.
[30] R. Timofte, E. Agustsson, L. Van Gool, M.H. Yang, L. Zhang, et al., Ntire 2017 challenge on single image superresolution: Methods and results.