

## Research Article

# Image Super-Resolution Using Lightweight Multiscale Residual Dense Network

Shilin Li,<sup>1</sup> Ming Zhao,<sup>1</sup> Zhengyun Fang,<sup>2</sup> Yafei Zhang ,<sup>3,4</sup> and Hongjie Li<sup>1</sup>

<sup>1</sup>Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming 650217, China

<sup>2</sup>College of Land Resource Engineering, Kunming University of Science and Technology, Kunming 650500, China

<sup>3</sup>College of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China

<sup>4</sup>Key Laboratory of Artificial Intelligence in Yunnan Province, Kunming University of Science and Technology, Kunming 650500, China

Correspondence should be addressed to Yafei Zhang; zyfeimail@163.com

Received 12 June 2020; Revised 31 August 2020; Accepted 20 September 2020; Published 13 October 2020

Academic Editor: Wonho Jhe

Copyright © 2020 Shilin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The current super-resolution methods cannot fully exploit the global and local information of the original low-resolution image, resulting in loss of some information. In order to solve the problem, we propose a multiscale residual dense network (MRDN) for image super-resolution. This network is constructed based on the residual dense network. It can integrate the multiscale information of the image and avoid losing too much information in the deep level of the network, while extracting more information under different receptive fields. In addition, in order to reduce the redundancy of the network parameters of MRDN, we further develop a lightweight parameter method and deploy it at different scales. This method can not only reduce the redundancy of network parameters but also enhance the nonlinear mapping ability of the network at different scales. Thus, it can better learn and fit the feature information of the original image and recover the satisfactory super-resolution image. Extensive experiments are conducted, which demonstrate the effectiveness of the proposed method.

## 1. Introduction

The purpose of image super-resolution reconstruction is to recover the corresponding high-resolution image from one or more low-resolution images. At present, the technology has been successfully applied to such fields as remote sensing, medical imaging, image compression, video monitoring, and military. As an important image processing technology, image super-resolution has received wide attention from researchers, and many effective super-resolution methods have been proposed [1].

The existing image super-resolution methods mainly include the following: interpolation-based methods [2], reconstruction-based methods [1,3,4], and learning-based methods [5–7]. Interpolation-based methods include bilinear interpolation, nearest neighbor interpolation, bicubic interpolation, and edge information based image interpolation [2]. These methods have lower computational

complexity, but the high-frequency detail information of the original image is often lost, resulting in poor visual quality. Reconstruction-based methods are based on the assumption that the low-resolution image is generated from the high-resolution image with downsampling, deformation, translation, and noise. Combined with prior knowledge, the high-resolution image can be recovered by optimization. Although reconstruction-based methods can retain more image details, the reconstruction effect is susceptible to the parameters of the image degradation model and the regularization model. Meanwhile, when the resolution is improved more than the scale factor of 4×, the reconstruction effect is usually not ideal.

Compared with the other two types of methods, learning-based methods can introduce more high-frequency information and have better robustness to noise, so they have become a research hot spot in recent years. The basic idea of learning-based methods is to establish a relationship between

low-and high-resolution images through learning and then use this relationship to guide high-resolution image reconstruction. Freeman et al. [5] first established this relationship based on Markov random field, but this method required large amounts of time to construct training sets, and image reconstruction was also time-consuming. Chang and Yeung [6] proposed a super-resolution reconstruction method based on neighborhood embedding, assuming that the high- and low-resolution image blocks have the same geometry manifold. Due to the excellent performance of sparse representation in computer vision tasks [8–13], Yang et al. [7] proposed a super-resolution reconstruction method based on sparse representation, which assumes that high- and low-resolution images have the same sparse coding coefficients.

In recent years, deep learning has attracted wide attention because it displays excellent performance in image super-resolution. A series of effective super-resolution algorithms based on deep learning have been proposed [14–20]. The deep learning-based methods usually use a single-scale network to extract image features. Under the constraint of minimizing the loss function, the network can extract the texture details from the low-resolution image to restore the high-resolution image. However, some feature information at different scales will be lost when extracting features through the single-scale network, resulting in unsatisfactory image quality of super-resolution.

To solve the above problem, we propose a multiscale feature extraction model for single image super-resolution. The model can extract features at different scales and in different receptive fields, which improves the efficiency of feature extraction considerably while reducing the depth of the network. Therefore, our method can not only make the model lighter and improve its training efficiency but also effectively avoid the significant degradation of the quality of the reconstructed image. Specifically, we design a multiscale residual dense network to extract the feature information on different scales, and propose to integrate the features of each layer to realize the sharing of multiscale information. Thus, the method can receive more information from different receptive fields, which is very beneficial to avoid information loss [21]. In addition, as we know, convolutional neural network (CNN) introduces adaptive parameters by adding a fully connected layer to the network, resulting in an increase in parameter size. Inspired by the idea of lightweight network [22], we propose to add lightweight parameters at each scale. Thus, the method ensures the quality of reconstruction without obvious degradation. Moreover, it reduces the parameter scale of the model, enhances the nonlinear mapping ability of the network, and improves the efficiency of the algorithm. The major contributions and innovations of our work are summarized as follows:

- (i) In order to obtain high-resolution image from the corresponding low-resolution image, we design a multiscale residual dense network. The network can extract the features of different scales and improve the efficiency of the algorithm, while keeping the reconstruction quality from degrading significantly.
- (ii) In multiscale reconstruction, a lightweight parameter learning method is developed and added to

each scale to enhance the nonlinear mapping ability of the network. Different from the existing residual dense networks, we do not introduce the common fully connected layer after the feature extraction at various scales but use a lightweight method to generate multiscale parameters.

- (iii) The proposed method not only uses multiple scales to extract feature information under different receptive fields but also adopts dilated convolution to increase the area of receptive fields. It can extract more receptive field features under different scales with the same number of parameters, so that the recovered high-resolution image retains more feature information.

The remainder of this paper is organized as follows: some related work is briefly reviewed in Section 2. Section 3 describes the proposed network structure. The experimental results and analysis are presented in Section 4. Finally, we conclude our work in Section 5.

## 2. Related Work

In recent years, researchers have witnessed the impressive performance of deep learning in single image super-resolution. In particular, Dong et al. [23] first applied deep learning to image super-resolution and proposed super-resolution convolutional neural network (SRCNN), which achieved clear reconstruction results. In order to solve the problem of a large number of parameters in SRCNN, researchers have proposed various solutions. Dong et al. [24] improved SRCNN and proposed fast super-resolution convolutional neural networks (FSRCNNs). This method first performs convolution and extracts features at the low-resolution stage, and then generates super-resolution images with upsampling at the end of the network. Kim et al. a deeply-recursive convolutional network (DRCN) [14] and a very deep convolutional network (VDSR) [25] for image super-resolution. Based on the idea of residual, VDSR reduces the training parameters by adding a skip connection to learn the residual parameters instead of all parameters. Zhang et al. [16] proposed an image super-resolution reconstruction method based on residual dense network (RDN). This method well solves the problems of gradient vanishing and low convergence efficiency in training and the problem of progressive loss of image information during the convolution process. Ledig et al. [17] proposed a super-resolution generative adversarial network (SRGAN). This method is based on adversarial learning and uses the adversarial training of generators and discriminators to generate texture details consistent with the distribution of natural images. The above methods all extract image features at a single scale and perform super-resolution. However, they ignore the complementary information of image features at different scales, leading to the loss of some information of the source image. This will hinder the reconstruction of detailed high-resolution images.

In order to solve the above problems, researchers have proposed the concept and model of multiscale convolutional neural networks in recent years [18, 26, 27]. The multiscale

methods use convolution kernels of different scales to extract features on different scale layers of the image, and then fuse them, thus alleviating the loss of image feature information and improving the quality of super-resolution. Specifically, Hu et al. [27] proposed a multiscale convolutional neural network, which can effectively extract feature information under different receptive fields. It performs feature fusion of different scale layers after each feature extraction module and then extracts the residuals between the fusion information of adjacent modules. Gao and Zhuang [26] developed a multiscale super-resolution method based on the deep neural network and showed the advantages of the multiscale residual dense network in feature extraction compared with the single-scale network. In addition, the enhanced deep super-resolution network [18] also utilizes multiscale residual blocks to eliminate gradient disappearance and gradient explosion. In the meanwhile, it adds a fixed hyperparameter to the multiscale network to enhance the network's ability, so that it can fit feature information at each scale and remove the batch normalization layer, thereby reducing the scale of parameters.

The existing multiscale models can solve the problem that the single-scale model cannot extract enough feature information because of too few branches. However, there are still some issues to consider: (1) by increasing the number of branches of the network, the models can enhance the ability of the network to extract complementary feature information. However, most of them do not consider the multiscale basis to further enhance extraction capability. (2) Most of the existing models input the multiscale information from the previous multiscale feature extraction block directly into the later feature extraction block after fusion, which is likely to cause the problem of gradient vanishing. To solve this problem, a multiscale dense residual network is proposed. In this method, the convolution kernel with different receptive fields is set up at different scales to integrate the advantages of multiscale, and the feature information of different receptive fields is also extracted. In this paper, the residual feature is directly input into the subsequent layer instead of multiscale fusion feature, which speeds up the convergence rate of the network. Moreover, dilated convolution is used to expand the receptive field without changing the parameters of the convolution kernel, which reduces the scale of parameters compared with the direct use of different sizes of convolution kernel.

### 3. Our Approach

We propose a super-resolution model based on MRDN. The model consists of three modules: shallow feature extraction, multiscale deep feature extraction, and reconstruction. In the shallow feature extraction module, a  $3 \times 3$  convolution layer is used for shallow feature extraction. Let  $F_0$  denote the output of shallow feature extraction block, which also is the input of the multiscale deep feature extraction module. In the multiscale deep feature extraction module, we have  $N$  multiscale residual blocks (MRBs). Each MRB contains  $M$  multiscale fusion layers (MFLs). Each MFL contains three different scale feature extraction branches. Each scale

contains a dilated convolution with a kernel size of 3, and the dilated ratio is 1, 3, and 5, respectively. In order to prevent grid effect, a convolution kernel of the same size as the dilated ratio is added before the dilated convolution of each scale. Each MRB is linked in the form of dense connection to ensure that the extracted feature information of each layer is not lost. Meanwhile, the information flows quickly to the convolution layer behind, accelerating the convergence speed. In the reconstruction module, upsampling is performed by  $3 \times 3$  deconvolution to generate a high-resolution feature map after reducing the dimension by a  $1 \times 1$  convolution. The specific network structure is shown in Figure 1.

**3.1. Multiscale Deep Feature Extraction Module.** The existing deep network cannot extract the feature information under different receptive fields. To solve this problem, we propose a multiscale residual block (MRB), which can extract the feature information under different receptive fields and produce better texture details of the recovered high-resolution image. The MRB is composed of three parts: feature extraction layer, multiscale fusion layers, and residual learning. The proposed multiscale residual network structure is shown in Figure 2.

Let  $F_{n-1}$  and  $F_n$  denote the input and output of the  $n$ -th MRB, respectively. In the feature extraction layer, features are extracted, and the feature map generated by this layer is directly transmitted to the end to obtain the residual, which accelerates network convergence and prevents gradients from disappearing. The formula is as follows:

$$FS_n = \text{relu}(\text{sconv}(F_{n-1}, k = 3)), \quad (1)$$

where  $FS_n$  is the output of convolution with a scale of  $k = 3$ ,  $\text{relu}$  is the ReLU activation function, and  $\text{sconv}$  is an equal-size convolution; that is, the input feature map is the same size as the output feature map by adding 0, and  $k = 3$  is the convolution kernel size.

$FS_n$  is used as input to multiscale fusion layers. Each MRB consists of  $M$  MFLs with a scale of 3, and the output of the former MFL is the input of the latter MFL. We use dilated convolution for feature extraction, which can expand the receptive field of the feature map without changing the parameters. By this method, the feature information of different receptive fields at different scales is extracted. Moreover, in the multiscale fusion layer, we propose a lightweight parameter method to simulate the channel attention mechanism. Instead of using the fully connected layer to generate adaptive parameters such as channel attention, the proposed method uses the lightweight parameters. The lightweight parameters are actually learnable tensors, which can be generated by built-in function. The tensors are introduced into the MRDN and become the trainable parameters. As shown in Figure 2, the lightweight parameters such as  $p1$ ,  $p2$ , and  $p3$  can be initialized to 1. In the feature fusion stage, the concatenation and  $1 \times 1$  convolution are used to fuse the feature maps extracted from three-scale feature extraction branches. The formula for the multiscale feature fusion layer is as follows:

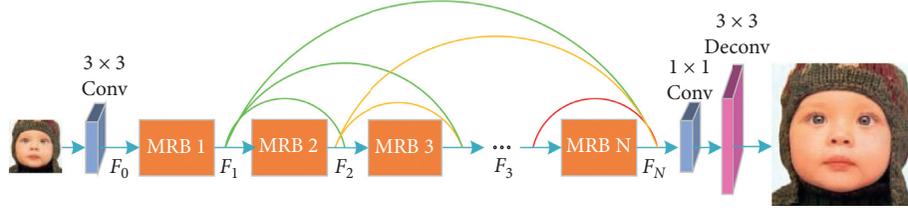


FIGURE 1: General structure of super-resolution reconstruction network.

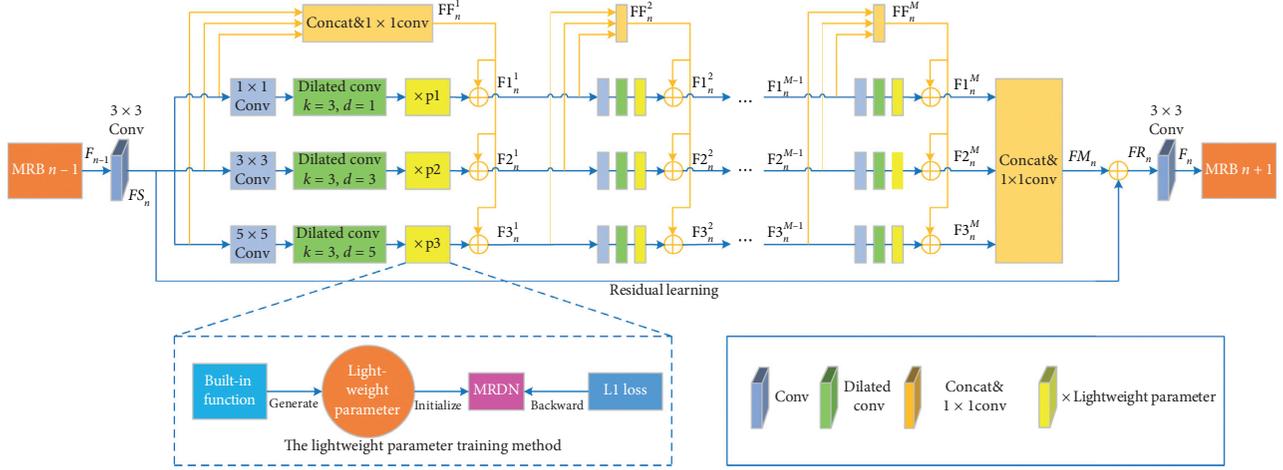


FIGURE 2: Structure of the multiscale residual block.

$$\begin{aligned}
 F_1^m &= \text{sconv}(\text{sconv}(F_1^{m-1}, k=1), k=3, d=1) \times p1 + FF_n^m, \\
 F_2^m &= \text{sconv}(\text{sconv}(F_2^{m-1}, k=3), k=3, d=3) \times p2 + FF_n^m, \\
 F_3^m &= \text{sconv}(\text{sconv}(F_3^{m-1}, k=5), k=3, d=5) \times p3 + FF_n^m,
 \end{aligned} \quad (2)$$

where  $F_1^m$ ,  $F_2^m$ , and  $F_3^m$  are three outputs of the  $m$ -th MFL in  $n$ -th MRB and  $F_1^{m-1}$ ,  $F_2^{m-1}$ , and  $F_3^{m-1}$  are its inputs.  $F_1^0 = F_2^0 = F_3^0 = F_n^0$  when  $m = 1$ .  $k$  is the convolution kernel size, and  $d$  is the dilation rate, which is used to control the size of the dilated convolution. In the process of dilated convolution, part of the feature image information may not be convoluted as the dilated convolution complement 0 increases the receptive fields, resulting in information loss. Therefore, the common convolution of the same size as the dilated convolution is performed before the dilated convolution, which can effectively eliminate the mesh effect of the dilated convolution and a large number of parameters.  $FF_n^m$  denotes the fused feature map, and  $FF_n^m = \text{sconv}(\text{concat}(F_1^{m-1}, F_2^{m-1}, F_3^{m-1}), k=1)$ .  $\text{concat}$  is used to concatenate three sets of feature maps extracted from different scales in the channel dimension.

At the end of multiscale fusion layers, the feature information extracted from the  $M$ -th MFL is fused by concatenation and  $1 \times 1$  convolution:

$$FM_n = \text{sconv}(\text{concat}(F_1^M, F_2^M, F_3^M), k=1), \quad (3)$$

where  $F_1^M$ ,  $F_2^M$ , and  $F_3^M$  are the output of three scales in the last MFL.  $\text{concat}$  is concatenating operation.

In the multiscale residual block, the residual learning is utilized to further improve the feature map by

$$FR_n = FS_n + FM_n, \quad (4)$$

where  $FR_n$  denotes the feature after residual learning.

At the end of the multiscale residual block, the  $3 \times 3$  convolution is used to further extract the feature, and the final output of the  $n$ -th MRB can be formulated as

$$F_n = \text{sconv}(FR_n, k=3). \quad (5)$$

**3.2. Loss Function.** In this paper, we use  $L1$  loss defined in equation (6) to make the reconstructed super-resolution image  $I^{\text{SR}}$  approximate to the real high-resolution image  $I^{\text{HR}}$ :

$$L1(I^{\text{HR}}, I^{\text{SR}}) = \|I^{\text{HR}} - I^{\text{SR}}\|_1. \quad (6)$$

## 4. Experiments

**4.1. Training Image Preprocessing.** In order for the training model to reconstruct the power and daily pictures well, we use the power images and DIV2K dataset (<https://data.vision.ee.ethz.ch/cvl/DIV2K/>) to train the model, respectively. The model can realize image super-resolution with scale factors of  $2\times$  and  $4\times$ . First, the DIV2K and power data sets are randomly cropped to construct high-resolution label sets. Then, the low-resolution training sets are obtained by downsampling the high-resolution images. When training the network model with magnification factor 2, the image

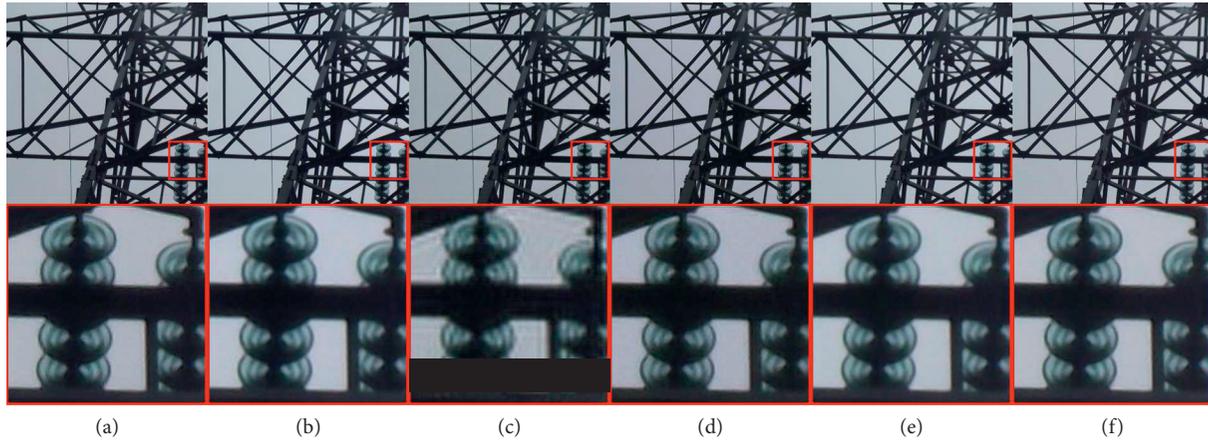


FIGURE 3: Super-resolution results of the power image with magnification factor 2: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

size of high-resolution label sets is  $128 \times 128$ , and that of low-resolution training sets is  $64 \times 64$ . When training the network model with magnification factor 4, the image size of high-resolution label sets is  $256 \times 256$ , and that of low-resolution training sets is  $64 \times 64$ .

**4.2. Implementation Details.** In the experiment, the training network contains 8 MRBs, each containing 3 MFLs. Each MFL has 3 feature extraction branches with different receptive fields.  $L1$  loss is used to make the reconstructed super-resolution image approximate to the real high-resolution image. The learning rate in our method is set to be 1 and remains unchanged during the iteration. The parameters of the whole MRDN are trained by back propagation until the model converges. In addition, we set epoch = 2200 when training on DIV2K, and epoch = 1200 when training on the power image set.

In the testing process, a group of high-resolution power images and three groups of high-resolution natural images in DIV2K dataset are selected to verify the effectiveness of the proposed method. The low-resolution images are obtained by downsampling the selected high-resolution images with scale factors of  $2\times$  and  $4\times$ . In addition, the proposed method is compared with SRCNN [23], FSRCNN [24], SRGAN [17], and RDN [16]. We use peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) to evaluate the quality of super-resolution reconstruction results objectively. PSNR can measure the mean square error between the reconstructed image and the original one. A larger PSNR means less image distortion and higher quality of the reconstructed image. SSIM can measure the structural similarity between the reconstructed image and the original one. The greater SSIM, the closer the reconstructed image is to the original.

**4.3. Comparisons with the State-of-the-Arts.** For the power image, the results of super-resolution reconstruction with scale factors of  $2\times$  and  $4\times$  using different methods are shown in Figures 3 and 4. From the perspective of visual effect, the

results of FSRCNN have artifacts, and the reconstructed images are blurred. The  $2\times$  super-resolution results of SRCNN, SRGAN, and RDN have better visual effect, but the  $4\times$  results are blur and have jagged edges. The  $2\times$  and  $4\times$  super-resolution reconstruction results of the proposed method have good visual effects. This is because the method can well integrate the feature information under different receptive fields and generate a reconstructed result with rich high-frequency information such as edge details and textures. To further compare the performance of different methods, quantitative assessments are presented in Table 1. As can be seen from these results, the proposed method achieves the highest values of PSNR and SSIM on the power image. Compared with RDN, our method improves PSNR and SSIM by 1.86 dB and 0.01, respectively, for magnification factor 2, and by 4.13 dB and 0.03, respectively, for magnification factor 4. Therefore, in terms of subjective and objective evaluation, the multiscale fusion model proposed in this paper is superior to the single-scale single receptive field models, such as SRCNN, FSRCNN, SRGAN, and RDN. Thus, the effectiveness of our method is verified.

In the second experiment, the “baby,” “mountain,” and “girl” images from DIV2K are used to further validate the effectiveness of the proposed method. The results of  $2\times$  and  $4\times$  super-resolution reconstruction of different methods are shown in Figures 5–10. In terms of visual effect, the reconstructed image of SRCNN displays jagged edges, and SRGAN produces details which are not consistent with the real texture. The results produced by FSRCNN lose more high-frequency details and have poor visual effect. The  $2\times$  results of RDN have better visual effects, but its  $4\times$  results lose details. Based on these results, we can see that our method can better fit the real high-resolution image and enhance the brightness of the reconstructed high-resolution image while recovering its texture and removing its sharp edge. Table 1 shows the quantitative evaluation of these reconstructed results. From these data, it can be seen that the PSNR value of the  $4\times$  super-resolution result of the “mountain” image by RDN is slightly higher than that of our method, and our method outperforms

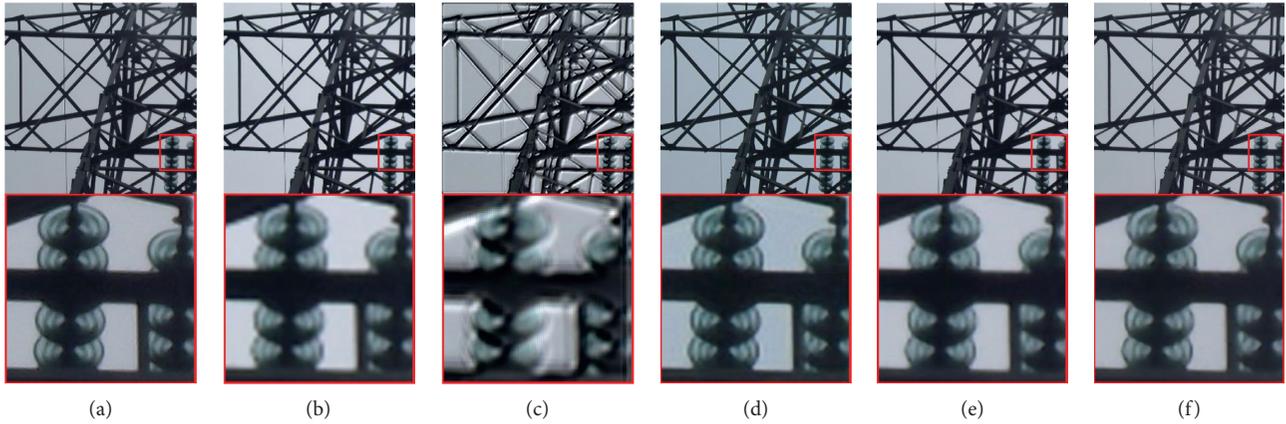


FIGURE 4: Super-resolution results of the power image with magnification factor 4: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

TABLE 1: Evaluation index (PSNR/SSIM) of super-resolution results by different methods.

Scale	Test images	SRCNN	FSRCNN	SRGAN	RDN	MRDN (ours)
2×	Power	28.61/0.97	16.37/0.56	27.76/0.86	29.40/0.97	<b>31.26/0.98</b>
	Baby	27.63/0.92	23.46/0.69	32.42/0.93	32.03/0.94	<b>34.00/0.95</b>
	Mountain	24.94/0.80	17.32/0.45	20.88/0.66	26.15/0.83	<b>26.39/0.86</b>
	Girl	28.94/0.74	25.67/0.62	26.88/0.66	30.64/0.79	<b>30.75/0.80</b>
4×	Power	21.92/0.87	27.88/0.85	30.93/0.91	27.04/0.92	<b>31.17/0.95</b>
	Baby	25.17/0.80	24.09/0.65	29.36/ <b>0.83</b>	27.75/0.81	<b>29.70/0.83</b>
	Mountain	21.96/0.57	20.47/0.40	18.23/0.48	<b>22.34/0.60</b>	<b>22.31/0.61</b>
	Girl	27.45/0.66	19.46/0.51	24.97/0.58	26.95/0.63	<b>28.70/0.68</b>

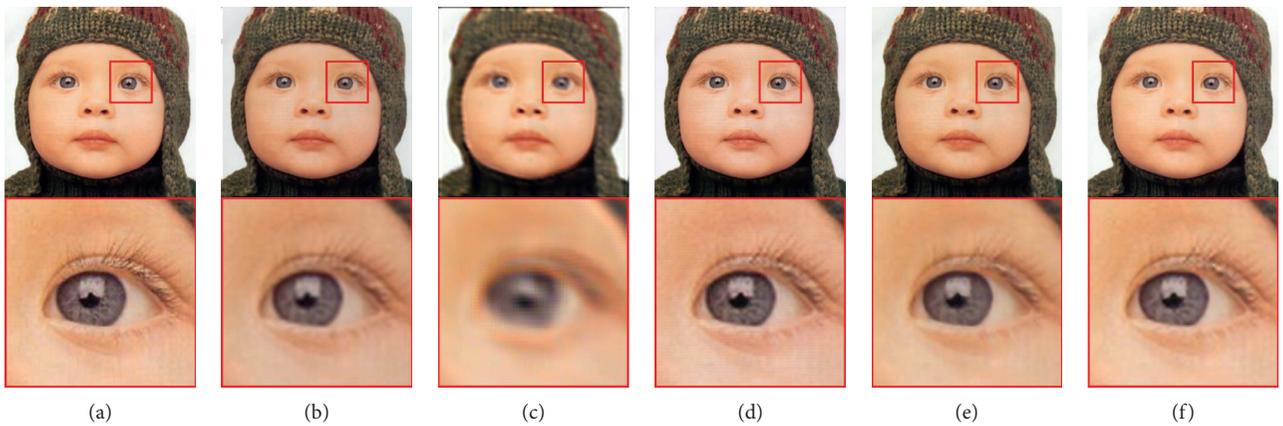


FIGURE 5: Super-resolution results of the “baby” image with magnification factor 2: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

the others in other metrics. Therefore, the effectiveness and superiority of our method are verified from the visual effect and quantitative evaluation.

#### 4.4. Model Validity Analysis

**4.4.1. Lightweight Parameter Validity Analysis.** In order to verify the effect of the proposed lightweight parameter

method, we replace the lightweight part of the proposed method with the traditional nonlightweight method and make a comparison to prove the validity of our method. In the process, we train the model on the power image dataset and reconstruct the low-resolution power image with a scale factor of 4. The PSNR and SSIM of the results are shown in Table 2, and the visual results are illustrated in Figure 11.

From these results, it can be seen that our method achieves improvement of PSNR which demonstrates the

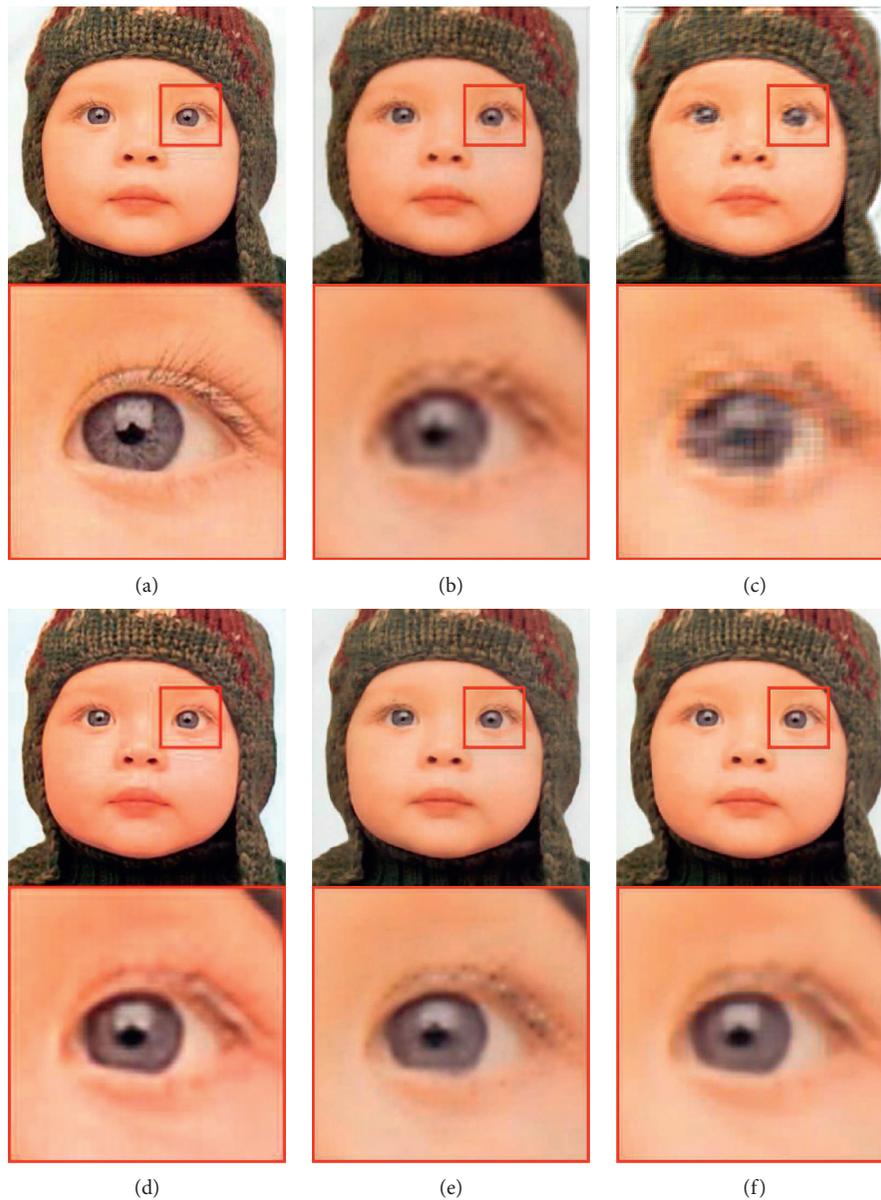


FIGURE 6: Super-resolution results of the “baby” image with magnification factor 4: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

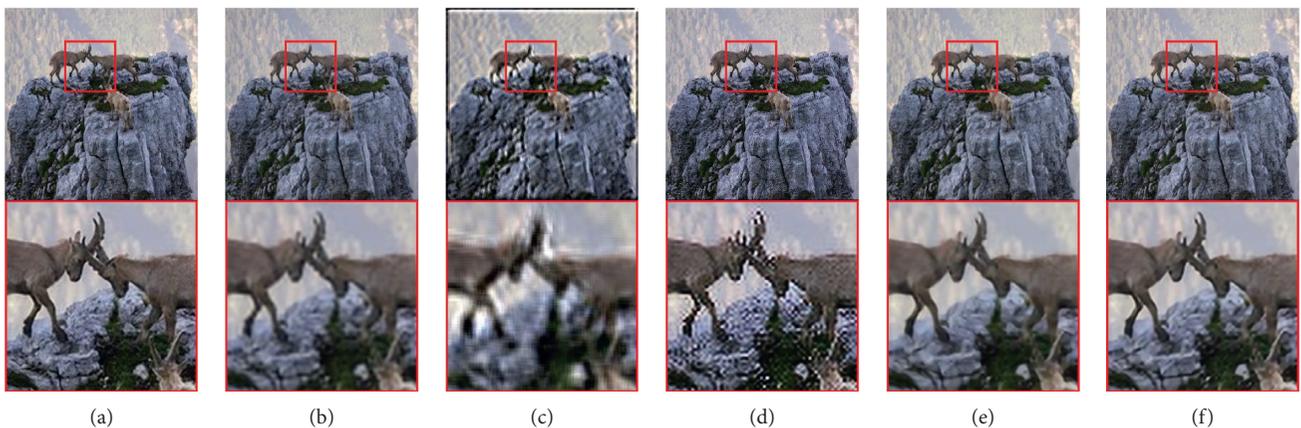


FIGURE 7: Super-resolution results of the “mountain” image with magnification factor 2: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

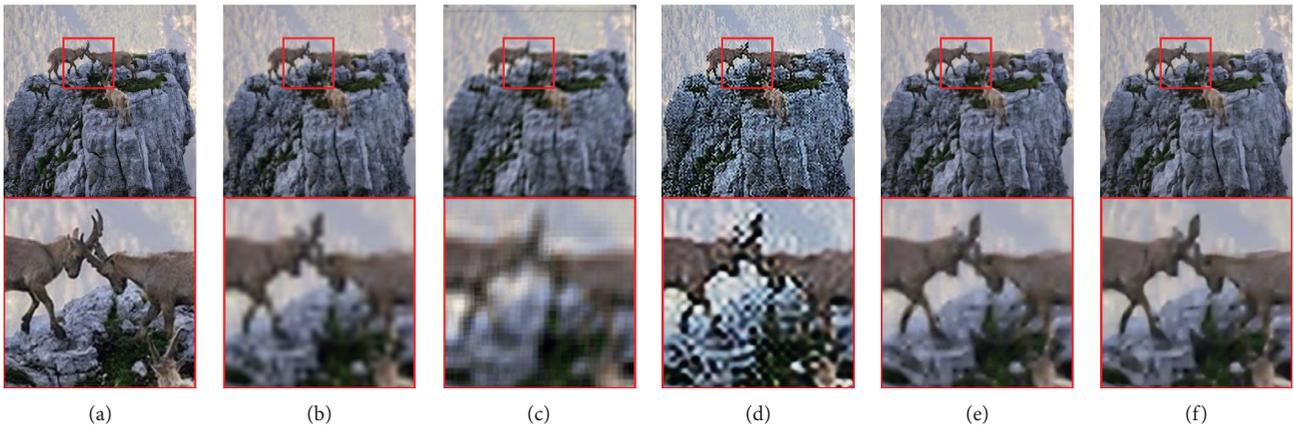


FIGURE 8: Super-resolution results of the “mountain” image with magnification factor 4: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

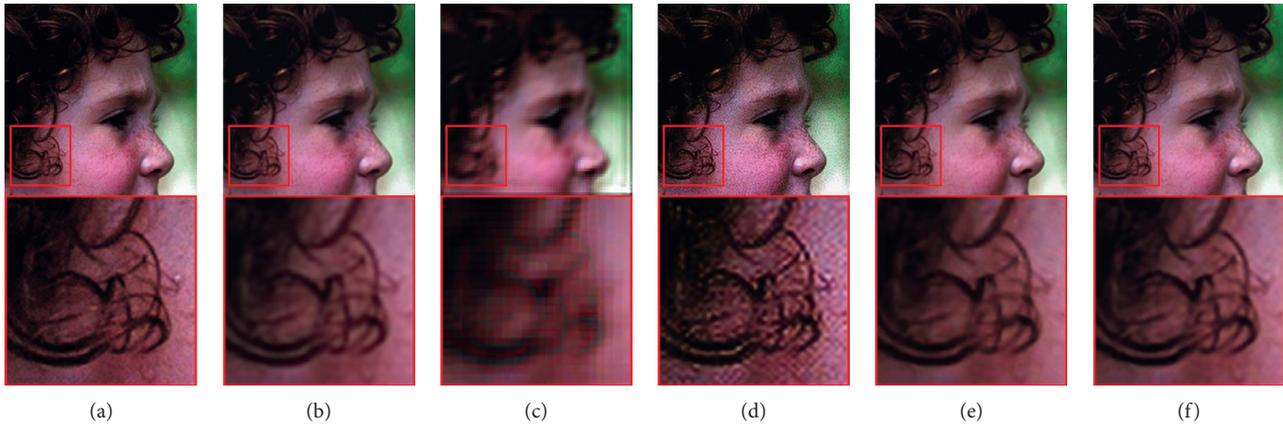


FIGURE 9: Super-resolution results of “girl” image with magnification factor 2: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

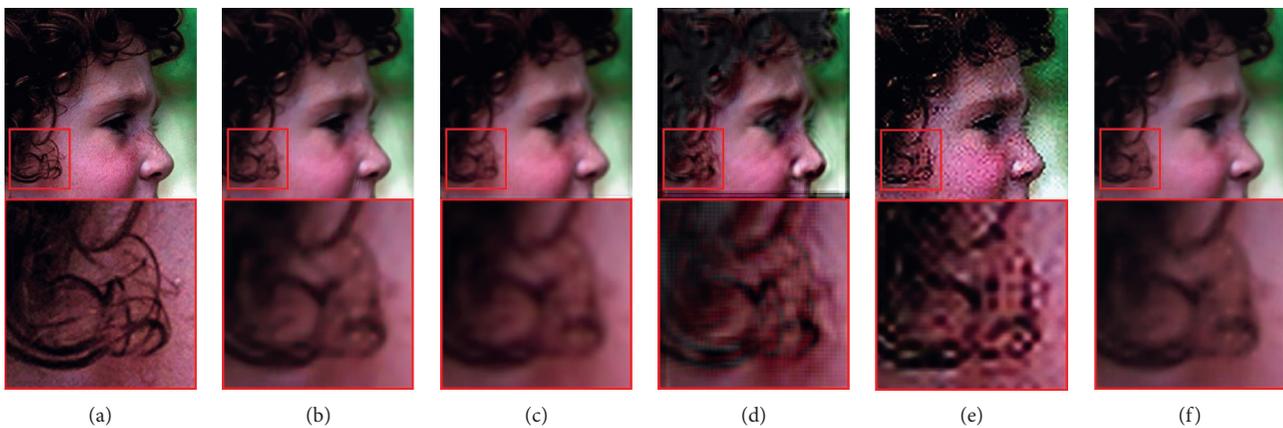


FIGURE 10: Super-resolution results of the “girl” image with magnification factor 4: (a) original high-resolution image; (b) result reconstructed by SRCNN; (c) result reconstructed by FSRCNN; (d) result reconstructed by SRGAN; (e) result reconstructed by RDN; (f) result reconstructed by our method.

TABLE 2: Performance analysis of super-resolution reconstruction results by the lightweight method.

Methods	PSNR	SSIM	PSNR	SSIM
Nonlightweight method	27.95 (Figure 11(a))	0.91 (Figure 11(a))	25.56 (Figure 11(c))	0.91 (Figure 11(c))
Our method	28.05 (Figure 11(b))	0.91 (Figure 11(b))	26.81 (Figure 11(d))	0.91 (Figure 11(d))

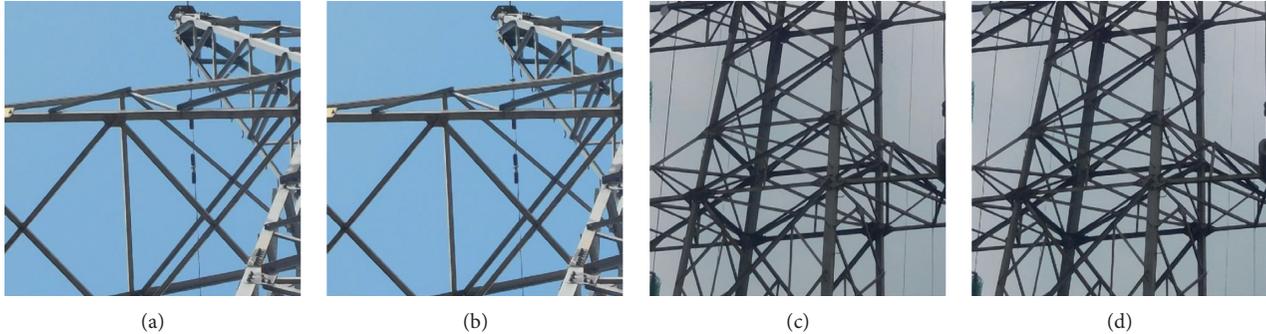
FIGURE 11: 4 $\times$  super-resolution results of networks with and without lightweight parameters: (a) tower top without lightweight parameters; (b) tower top with lightweight parameters; (c) tower body without lightweight parameters; (d) tower body with lightweight parameters.

TABLE 3: Performance analysis of the lightweight method.

Methods	PSNR	SSIM	Time (h)
Nonlightweight method	27.49	0.90	5.0
Our method	28.05	0.91	4.5

TABLE 4: Performance comparison of our method with different scales.

Scale layer	Training time (h)	PSNR	SSIM
Two scales	4	22.15	0.90
Three scales	4.5	28.05	0.91
Four scales	9.5	26.21	0.90

effectiveness of adding lightweight parameters to the multiscale networks. The reason is mainly because this design can further enhance the ability of nonlinear mapping at each scale, thus improving the quality of the generated high-resolution images. Although some parameters are added in the training, the scale of the parameters is greatly reduced compared with the traditional method of using the fully connected layer. More importantly, while increasing the nonlinear mapping ability, the complexity of the algorithm is not increased greatly.

**4.4.2. Effectiveness Analysis of Lightweight Model.** In order to verify the validity of the lightweight model, we replace the lightweight parameter with the traditional channel attention mechanism (the nonlightweight method) to compare their efficiency. For the fairness of comparison,

both methods use power dataset to train the model and the low-resolution images are generated by downsampling their corresponding high-resolution version. Table 3 shows that the proposed multiscale lightweight method improves the training efficiency by 0.5 h under the same 1200 epoch, which proves the effectiveness of the proposed algorithm. In addition, as can be seen from Table 3, the proposed lightweight model improves the objective quality compared with the traditional approach. This shows that the algorithm not only reduces the complexity of the model training and improves the training efficiency of the algorithm but also keeps the quality of the results from degrading.

**4.4.3. Discussion of Multiscale Selection.** In the above experiments, our method is compared with the single-scale network, which demonstrates its validity and superiority. However, in the multiscale selection, it is verified whether the three-scale selection is optimal. To this end, we analyze the quality of the reconstructed image at different scales and the training time of the algorithm. In this process, the power image dataset is used as the training set to train the model, and the reconstructed high-resolution results are shown in Table 4. As can be seen from these results, the reconstruction performance of this algorithm is the best, and the training efficiency of the algorithm has not been greatly reduced in the three-scale feature learning. On the contrary, under the four-scale condition, the quality of image reconstruction decreases, and the training efficiency of the algorithm degrades from 4.5 h to 9.5 h. This shows that the feature extraction at three scales is an appropriate choice.

## 5. Conclusion

This paper proposes a new lightweight dense residual network based on multiscale analysis for single image super-resolution. The method can avoid the problem of losing feature information when extracting features by the single-scale network. The power image dataset and the natural scene image dataset (DIV2K) are used to train the network separately, and the power test images and the natural test images are employed to verify the effectiveness and performance of the model. The experiments demonstrate the validity of our algorithm. After analysis, the following conclusions are drawn: (1) compared with the single-scale model, the multiscale residual dense network proposed in this paper can extract the feature information of different scale layers and different sizes of receptive fields, which is very conducive to the extraction of image feature information. (2) Lightweight parameters reduce the redundancy of the algorithm effectively while enhancing the nonlinear mapping of the network, and the experiments indicate that using three scales to construct the network model produces the best performance.

## Data Availability

The power images used to support the findings of this study are supplied by the Electric Power Research Institute of Yunnan Power Grid Co., Ltd., under license and so cannot be made freely available. The DIV2K dataset is available at <https://data.vision.ee.ethz.ch/cvl/DIV2K/>.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was funded by the Science and Technology Project of Yunnan Power Grid Co., Ltd. (no. YNKJXM20190729) and the Yunnan Natural Science Foundation (no. 2017FB094).

## References

- [1] C. Xiao, J. Yu, and K. Su, "Gibbs artifact reduction for pocs super-resolution image reconstruction," *Frontiers of Computer Science in China*, vol. 2, no. 1, pp. 87–93, 2008.
- [2] L. Zhang and X. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2226–2238, 2006.
- [3] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam, and F. E. A. EL-SAMIE, "Wavelet fusion: a tool to break the limits on lmmse image super-resolution," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 4, no. 1, pp. 105–118, 2006.
- [4] H. Ji and C. Fermuller, "Robust wavelet-based super-resolution reconstruction: theory and algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp. 649–660, 2008.
- [5] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25–47, 2000.
- [6] D.-Y. Chang and Y. X. Yeung, "Super-resolution through neighbor embedding," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. I, Washington, DC, USA, July 2004.
- [7] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Anchorage, AK, USA, June 2008.
- [8] H. Li, Y. Wang, Z. Yang, R. Wang, X. Li, and D. Tao, "Discriminative dictionary learning-based multiple component decomposition for detail-preserving noisy image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1082–1102, 2020.
- [9] H. Li, S. Yan, Z. Yu, and D. Tao, "Attribute-identity embedding and self-supervised learning for scalable person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, 2019.
- [10] Y. Zhang, M. Yang, N. Li, and Z. Yu, "Analysis-synthesis dictionary pair learning and patch saliency measure for image fusion," *Signal Processing*, vol. 167, p. 107327, 2020.
- [11] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, "Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning," *Pattern Recognition*, vol. 79, pp. 130–146, 2018.
- [12] M. Xie, Z. Zhou, and Y. Zhang, "Joint framework for image fusion and super-resolution via multicomponent analysis and residual compensation," *IEEE Access*, vol. 7, pp. 174092–174107, 2019.
- [13] H. Li, X. He, Z. Yu, and J. Luo, "Noise-robust image fusion with low-rank sparse decomposition guided by external patch prior," *Information Sciences*, vol. 523, pp. 10–24, 2020.
- [14] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Las Vegas, NV, USA, pp. 1637–1645, June 2016.
- [15] Y. Qiu, R. Wang, D. Tao, and J. Cheng, "Embedded block residual network: a recursive restoration model for single-image super-resolution," in *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Korea (South), pp. 4179–4188, November 2019.
- [16] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, UT, USA, pp. 2472–2481, June 2018.
- [17] C. Ledig, L. Theis, F. Huszar, J. Caballero et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Honolulu, HI, USA, pp. 105–114, July 2017.
- [18] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Honolulu, HI, USA, pp. 1132–1140, July 2017.
- [19] K. Zeng, J. Yu, R. Wang, C. Li, and D. Tao, "Coupled deep autoencoder for single image super-resolution," *IEEE Transactions on Cybernetics*, vol. 47, no. 1, pp. 27–37, 2017.
- [20] S. Maeda, "Unpaired image super-resolution using pseudo-supervision," in *Proceedings of the 2020 IEEE/CVF Conference*

- on *Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, WA, USA, pp. 288–297, November 2020.
- [21] R. Wang, M. Gong, and D. Tao, “Receptive field size versus model depth for single image super-resolution,” *IEEE Transactions on Image Processing*, vol. 29, no. 1, pp. 1669–1682, 2020.
  - [22] C. Wang, Z. Li, and J. Shi, “Lightweight image super-resolution with adaptive weighted learning network,” Ithaca, NY, USA.
  - [23] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.
  - [24] C. Dong, C. C. Loy, and X. Tang, “Accelerating the super-resolution convolutional neural network,” in *Proceedings of the Computer Vision -ECCV 2016*, Springer, Vegas, NV, USA, pp. 391–407, June 2016.
  - [25] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1646–1654, IEEE, Las Vegas, NV, USA, June 2016.
  - [26] S. Gao and X. Zhuang, “Multi-scale deep neural networks for real image super-resolution,” in *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Seoul, Korea, November 2019.
  - [27] Y. Hu, X. Gao, J. Li et al., “Single image super-resolution via cascaded multi-scale cross network,” Ithaca, NY, USA.