

Image Super-Resolution Reconstruction Based on Dual Residual Network

Zhe Wang¹, Liguang Zhang^{1,2,*}, Tong Shuai³, Shuo Liang³ and Sizhao Li^{1,4}

¹Harbin Engineering University, Harbin, 150001, China

²Xidian University, Xi'an, 710000, China

³The 54th Research Institute of CETC, Shijiazhuang, 050000, China

⁴The Chinese University of Hong Kong, Hong Kong, 999077, China

*Corresponding Author: Liguang Zhang. Email: zhangliguo@hrbeu.edu.cn

Received: 26 January 2022; Accepted: 09 March 2022

Abstract: Research shows that deep learning algorithms can effectively improve a single image's super-resolution quality. However, if the algorithm is solely focused on increasing network depth and the desired result is not achieved, difficulties in the training process are more likely to arise. Simultaneously, the function space that can be transferred from a low-resolution image to a high-resolution image is enormous, making finding a satisfactory solution difficult. In this paper, we propose a deep learning method for single image super-resolution. The MDRN network framework uses multi-scale residual blocks and dual learning to fully acquire features in low-resolution images. Finally, these features will be sent to the image reconstruction module to restore high-quality images. The function space is constrained by the closed-loop formed by dual learning, which provides additional supervision for the super-resolution reconstruction of the image. The up-sampling process includes residual blocks with short-hop connections, so that the network focuses on learning high-frequency information, and strives to reconstruct images with richer feature details. The experimental results of $\times 4$ and $\times 8$ super-resolution reconstruction of the image show that the quality of the reconstructed image with this method is better than some existing experimental results of image super-resolution reconstruction in subjective visual effects and objective evaluation indicators.

Keywords: Super-resolution; convolution neural network; residual learning; dual learning

1 Introduction

Single-image super-resolution reconstruction (SISR) and multi-image super-resolution reconstruction are the two types of image super-resolution reconstruction (SR). This study focuses on the super-resolution reconstruction of a single image. SISR technology aims to transform low-resolution (LR) images into high-resolution (HR) counterparts [1,2]. It is widely used in the fields of medical imaging [3], image compression [4], remote sensing [5], security [6], and underwater target recognition



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[7]. High-quality images require not only sufficient pixel density; but also rich detailed information. Super-resolution reconstruction is an ill-conditioned inverse problem. The main challenge faced by many researchers is that LR images can correspond to multiple HR images, and it is very difficult to restore HR images with vivid and rich details. According to studies, image super-resolution reconstruction has always been a difficulty in the realm of computer vision. Because of the rapid rise of deep learning technology, it has quickly become a major tool for current SISR approach research [2,7–10].

The difficulty of trying to rebuild HR images from LR images is essentially an ill-conditioned task, according to research. This is because the LR image might correlate to numerous HR images, and the available mapping space is enormous, making finding the correct relationship extremely challenging. Furthermore, when high magnification is necessary, the reconstructed image's features are frequently inadequate. Collecting useful context information from LR images is critical for reconstructing the high-frequency details of HR, as stated in the previous section of the research [11].

Deeper convolutional neural networks have been effectively employed to recreate single image super-resolution in recent years [12–14], and the structure of CNN has been continuously enhanced [15]. The benefit of utilizing a deep network is that it can use a bigger receptive field to extract more contextual information from the LR image to predict the information in the HR image. However, simply raising the network depth will cause the features to fade away throughout the transmission process, making it difficult to recover the image's details.

This study aims to contribute to this growing area of research by exploring residual networks and regression networks. We proposed the MDRN network framework. To construct the down-sampling section of the network in this framework, we coupled the utilization of multi-scale residual blocks (MSRB) [11]. At various sizes, it may be utilized to extract image information. By integrating local multi-scale features and global features during feature extraction, it may maximize the use of LR image features and effectively tackle the problem of feature disappearance during transmission. Furthermore, the introduced 1×1 kernel convolutional layer can achieve global feature fusion.

The dual regression method is used in the reconstruction module to reduce the possible spatial function of the LR data by introducing a regression scheme with additional constraints [16]. The newly learned double regression mapping, in addition to the LR to HR image mapping, forms a closed-loop for the down-sampling kernel, providing additional supervision for reconstructing the LR image, significantly reducing the mapping space and enabling greater feature retention.

The model is trained using the DIV2K data set, with no weight initialization or additional training approaches. Our model has produced good results based on the ensuing experimental data. The following is a summary of our contributions:

- A multi-scale residual block is utilized in the down-sampling process to adaptively detect image features and realize feature fusion of different scales. The residual channel attention block can be used to fully understand the image features throughout the up-sampling process. When the two modules are combined, an excellent super-resolution effect can be achieved.
- The function mapping from the LR image to the HR image can be built into a constrained closed loop using the dual learning framework, and the LR image reconstructed based on this can be utilized to improve the performance of the SR model.

2 Related Work

2.1 Single Image Super-Resolution

Image super-resolution is an improper inverse technique, and the input LR image has several equivalent outputs. A variety of image super-resolution algorithms have been developed to handle this inverse problem, including interpolation-based methods [17], reconstruction-based methods [18], and learning-based methods [11,16,19–22]. The interpolation-based method has the advantage of low complexity, but the contour edge of the reconstructed image is blurred, the details are not clear enough, and the accuracy is somewhat lacking. In different contexts, reconstruction-based approaches will have different performance discrepancies. The deep learning method requires more training data than the classic interpolation method, but it provides great denoising and augmentation results. It learns the difference in complexity between low-resolution and high-resolution images using a deep neural network's powerful nonlinear reflection ability. Richer image features and texture information can be restored via the mapping relationship.

The deep learning-based SISR approach can learn the mapping relationship between the end-to-end LR image and the HR image directly. Super-Resolution Convolutional Neural Network (SRCNN) [22], a SISR approach based on convolutional neural networks, was first proposed by Dong et al. SRCNN employs only three convolutional layers and learns the nonlinear mapping between LR and HR image from start to finish [23]. Kim et al. presented a deeper network SISR approach based on the global residual learning method [24]. It has been discovered that images of various magnifications can be combined for training purposes, allowing the model to solve the SR problem of images of various magnifications. Tai et al. devised Deep Recursive Residual Network (DRRN), which combines the local residual learning and global residual learning of the multi-path mode, as well as the SR method of multi-weight recursive learning, to successfully restrict the increase of network parameters while expanding the network depth [25]. Through the sharing of parameters between residual units, the performance of VDSR and DRCN [26] is improved. The input image contains rich low-frequency and high-frequency information. The existing CNN-based SR network does not treat each channel of this information differently, which limits the expressive ability of the network. Therefore, Zhang et al. [11] proposed a 400-layer residual network Residual Channel Attention Networks (RCAN) with a channel attention mechanism, which uses the RIR structure to ease the burden of information transmission, and can learn coarse-grained residual information to stabilize the training process. To obtain reconstructed images with different magnification factors, multi-scale images based on residual networks can be done. Lai et al. [27] proposed the SISR method LapSRN of the Laplacian Pyramid Super-Resolution Network, which gradually up-sampling and predicts residuals, and can simultaneously complete HR image reconstruction of multiple sizes. The Multi-Scale Residual Network (MSRN) method proposed by Li et al. [11] is an image SR network that can use multi-scale hierarchical features to zoom in at any scale. It uses convolutional layers of different receptive fields in the residual block to extract different scales. Feature information further improves the performance.

The above method uses a lightweight network. The depth of the network and the number of parameters are important factors that affect the performance of SISR. Some heavyweight networks have also achieved good results. Lim et al. [28] proposed a heavyweight Enhanced Deep Super-Resolution Network (EDSR) method, which removed the normalization module and superimposed the residual block. The Residual Dense Network (RDN) method proposed by Zhang et al. [23] combines residual structure; and dense structure; and makes full use of the hierarchical feature information of LR images. High-resolution reconstruction of the LR images obtained from different degradation models has achieved good reconstruction results. Liu et al. [29] proposed the Residual

Feature Aggregation Network (RFANet) method, which uses a spatial attention module with larger receptive fields and smaller parameters in the residual block. By filtering the feature information, and fusing the features extracted by the residual branch of each residual block, the image reconstruction quality is improved.

Most of the previous techniques do not fully utilize the information contained in each convolutional layer. Local convolutional layers cannot directly access the following layers, even though gate units in storage blocks are offered as a way to regulate short-term memory. It's also difficult to say that the memory block can fully utilize all layers' information.

2.2 Residual Learning

Is it possible to lower the training error by adding a new layer to the neural network model after enough training? In theory, the previous model's solution space is just a subspace of the new model's solution space, and the newly added layer can be trained as an identity mapping. The more layers appear to be simpler to lower the training error since the new model may be able to find a better solution to fit the training data set. However, it has been discovered that when too many layers are added to the training process, the error does not reduce but rather grows. Moreover, the depth of the network has a great influence on the SISR. In the neural network, the local residual block is used to simplify the training of the network.

Currently, many feature extraction blocks have been proposed thus far. The easy linking of various characteristics at different scales, as shown in Fig. 1, will lead the local features to be underutilized. (Fig. 1a) The goal is to make network training easier so that it can produce more competitive results. When dense blocks are introduced (Fig. 1b), both residual blocks and dense blocks employ the same size convolution kernel, increasing the computational cost of dense blocks [30].

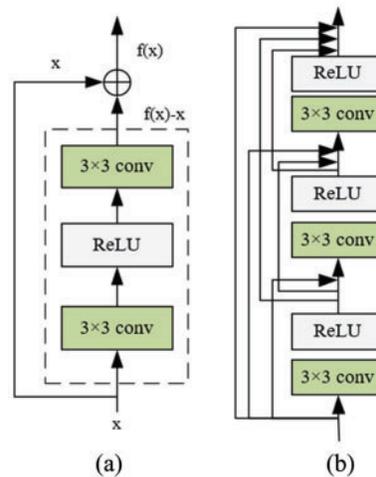


Figure 1: (a) Residual block. (b) Dense block

The deep residual channel attention network's concept is to build a deep network with residual in residual (RIR) structure, which is made up of numerous residual groups connected by long-hop connections [21]. There are several residual blocks with short jump connections in each residual group. The entire RIR structure uses multiple hop connections to allow low-frequency information to bypass the network; so that the backbone network only learns high-frequency information. The channel

attention mechanism (CA) adaptively adjusts the weights of different channels by considering the interdependence between channels.

2.3 Dual Learning

In real life, artificial intelligence tasks that have meaning and practical value often appear in pairs. Dual learning involves at least two learning tasks. These two tasks are formed into a closed-loop feedback system, and the feedback information is used to improve the two machine learning models in the dual-task.

The dual learning approach consists of an original model and a dual model, which can simultaneously train two opposed mappings to improve language translation performance. The key is that the model of the original task can provide feedback to the model of the dual-task. The parameters of the original model and the dual model are shared, which means that the model can have fewer parameters. Compared to standard supervised learning, the data will be more fully utilized. Due to the sharing of parameters, the complexity of the two models is reduced, so there will be better generalization capabilities. CycleGAN [31] and DualGAN [32] are two recent examples of this technology being utilized to conduct image translation without paired training data. The loop consistency loss is proposed to help decrease the distribution difference and avoid the mode collapse problem of the GAN approach. These solutions, however, cannot be applied directly to conventional SR situations. As a result, a closed-loop can be utilized to minimize the number of SR functions that are conceivable.

3 Proposed Method

The main purpose of this paper is to use the low-resolution image to rebuild the super-resolution image. The primary structure of the network and its significant components are mostly introduced in this part. The multi-scale dual residual network (MDRN) is made up of two main components: the feature extraction module and the image reconstruction module. Fig. 2 depicts the broad structure of our concept, which will be detailed in the next chapters.

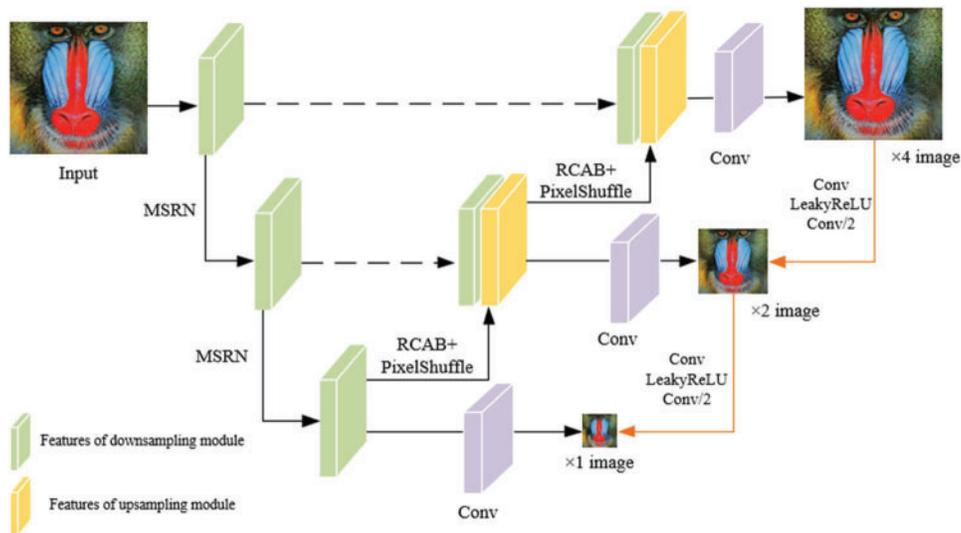


Figure 2: MDRN frame structure

3.1 Feature Extraction Module

Both the down-sampling module and the up-sampling module contain $\log_2(s)$ basic blocks, where s is the scale factor. It is equivalent to using 2 blocks for 4 times magnification and 3 blocks for 8 times magnification. The feature extraction module includes the convolutional layer, the LeakyReLU layer, and the residual layer for feature extraction of the input image. Here, the residual structure is used in the hope that more detailed features can be extracted and put into the dual model for supervision. As shown in Fig. 3, it is the main structure of our feature extraction module, the content of which is described in detail below.

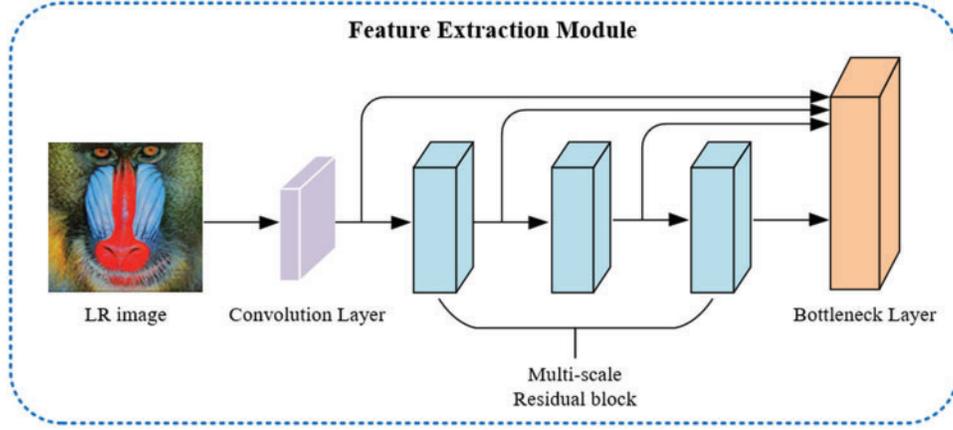


Figure 3: The structure of the feature extraction module

We use multi-scale residual blocks to extract features and design a two-bypass network to learn image features of multiple sizes throughout the down-sampling process. Convolution kernels are used differently by different bypasses. Image features of various scales can be detected by transferring information between the two bypasses.

$$D_1 = \sigma(w_{3 \times 3}^1 * M_{n-1} + b^1) \quad (1)$$

$$T_1 = \sigma(w_{5 \times 5}^1 * M_{n-1} + b^1) \quad (2)$$

$$D_2 = \sigma(w_{3 \times 3}^2 * [D_1, T_1] + b^2) \quad (3)$$

$$T_2 = \sigma(w_{5 \times 5}^2 * [T_1, D_1] + b^2) \quad (4)$$

$$D' = w_{1 \times 1}^3 * [D_2, T_2] + b^3 \quad (5)$$

where w and b represent weights and deviations, $w_{3 \times 3}^1$ represents the use of a 3×3 convolution kernel in the first layer, and so on. $\sigma(x) = \max(0, x)$ represents the ReLU function, $[D_1, T_1]$ represents the series operation.

For image feature extraction, we use the multi-scale residual block described in [11]. If the number of feature maps is denoted by the letter M . M feature maps will output from the first convolutional layer, while $2M$ feature maps are output from the second convolutional layer. The 1×1 convolutional layer receives these feature maps and can make the number of input and output feature maps the same.

For each MSRB we use residual learning, which can be described as:

$$M_n = D' + M_{n-1} \quad (6)$$

Among them, M_n and M_{n-1} represent MSRB input and output, respectively. The operation $D + M_{n-1}$ is performed through shortcut connection and element-wise addition.

3.2 Image Reconstruction Module

The image reconstruction module is a dual network that creates down-sampled LR images from super-resolution images to give extra supervision to the model. The emphasis is on learning the down-sampling procedure, which is made up of two convolutional layers and a LeakyReLU layer. As seen in Fig. 4, the main body employs a dual learning framework that employs the channel attention module. The up-sampling section incorporates B residual channel attention blocks with small jumps, which can boost the model's capacity. Short-hop connections can filter unwanted low-frequency information.

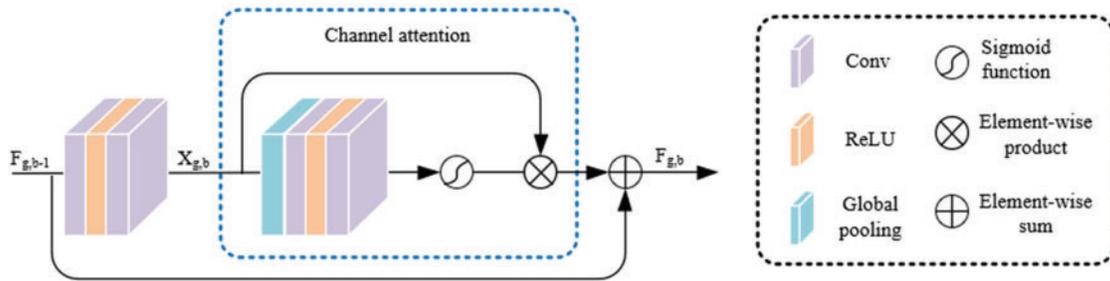


Figure 4: The structure of the up-sampling module

The interdependence between feature channels is used to construct a channel attention mechanism to have the network pay attention to additional information features. There are two main issues to consider. One is that the information in the LR space has a lot of low-frequency information and a lot of high-frequency information. Details such as edges and textures in the area are common in high-frequency components. The other is that each filter in the convolution layer operates with a local receptive field. As a result, outside of the immediate area, the output of convolution cannot utilize context information. The channel approach is used to capture the channel's reliance on aggregated data using global average pooling, and a sigmoid function mechanism is implemented.

$$s = f(W_u \delta(W_c z)) \quad (7)$$

Among them, sigmoid gating and ReLU functions are denoted by $f(\cdot)$ and $\delta(\cdot)$, respectively. W_c is a Conv layer weight set that functions as the channel reduction and reduction ratio r .

We employed the PixelShuffle method and the residual channel attention block in the up-sampling process. The feature of each channel is adaptively rescaled by modeling the dependency between feature channels using the channel attention method. During the up-sampling process, PixelShuffle exploits multi-channel recombination of the convolution kernel to get high-resolution feature maps from low-resolution feature maps. Learning dual regression mapping estimates the down-sampling kernel to reconstruct low-resolution images to give extra supervision for the network. With this additional limitation, the available mapping space from low resolution to high resolution may be constrained.

4 Experiment

4.1 Implementation Details

The training dataset we use is DIV2K, and the input of the network is obtained by down-sampling the original images in the data set by bicubic interpolation. The ADAM optimizer with the parameters set to $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$ uses the weight normalization processing in the training model, the learning rate is set to 10^{-4} , and a total of 1000 iteration cycles are trained. And the loss function is the L1 loss function. Use the Set5 standard test dataset and the 801st to 900th images in the DIV2K dataset for comparison experiments. For testing, use Set14, BSDS100, Urban100, and Manga109. Consider the average peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) to evaluate the experimental outcomes. PyTorch was used to create the model, which was then trained on an NVIDIA Quadro RTX 5000.

4.2 Quantitative Analysis

On the one hand, the evaluation index of image super-resolution reconstruction quality can be judged as subjective evaluation by human eyes, which is more one-sided. On the other hand, PSNR and SSIM are used as objective evaluation indexes. The PSNR is an evaluation metric that is sensitive to the mistake of reconstructed pixels and ignores the human eye’s visual features. The SSIM is assessed using the three dimensions of brightness, contrast, and structure, with a focus on people’s subjective feelings. Combining the two evaluation indicators can more objectively evaluate the quality of the reconstructed image. For $\times 4$ and $\times 8$ SR, the algorithm in this paper is compared with 11 SR methods. As shown in [Tab. 1](#), in the comparison algorithm, Bicubic is a traditional bicubic interpolation algorithm, SRCNN [22], ESPCN [15] and FSRCNN [33] are shallow linear networks based on CNN, VDSR [34] is a deep linear network, DRCN [26] is a recursive network, and LapSRN [27] is a progressive reconstruction Network, EDSR is the residual network. It can be seen from [Tab. 1](#) that our model performs better than other models on different magnification factors and test data sets.

Table 1: Quantitative evaluation of different SR algorithms on $\times 4$ and $\times 8$

Algorithm	Scale	Set5 PSNR/ SSIM	Set14 PSNR/ SSIM	BSDS100 PSNR/ SSIM	Urban100 PSNR/ SSIM	Manga109 PSNR/ SSIM
Bicubic	$\times 4$	28.43/0.8022	26.10/0.6936	25.97/0.6517	23.14/0.6599	24.91/0.7826
SRGAN [35]	$\times 4$	29.46/0.8380	26.60/0.7180	25.74/0.6660	24.50/0.7360	27.79/0.8560
A+ [21]	$\times 4$	30.33/0.8565	27.44/0.7450	26.83/0.6999	24.34/0.7211	27.03/0.8439
SRCNN [22]	$\times 4$	30.50/0.8573	27.62/0.7453	26.91/0.6994	24.53/0.7236	27.66/0.8505
ESPCN [15]	$\times 4$	30.66/0.8646	27.71/0.7562	26.98/0.7124	24.60/0.7360	27.70/0.8560
FSRCNN [33]	$\times 4$	30.73/0.8601	27.71/0.7488	26.98/0.7029	24.62/0.7272	27.90/0.8517
VDSR [34]	$\times 4$	31.36/0.8796	28.11/0.7624	27.29/0.7167	25.18/0.7543	28.83/0.8809
LapSRN [27]	$\times 4$	31.54/0.8811	28.19/0.7635	27.32/0.7162	25.21/0.7564	29.09/0.8845
DRCN [26]	$\times 4$	31.56/0.8810	28.15/0.7627	27.24/0.7150	25.15/0.7530	28.98/0.8816
DBPN [36]	$\times 4$	32.42/0.8970	28.75/0.7860	27.67/0.7390	26.38/0.7940	30.90/0.9130
EDSR [28]	$\times 4$	32.46/0.8968	28.80/0.7876	27.71/0.7420	26.64/0.8033	31.02/0.9148
MDRN (Our)	$\times 4$	32.57/0.9000	28.82/0.7986	27.73/0.7436	26.62/0.8145	31.55/0.9193
Bicubic	$\times 8$	24.40/0.6045	23.19/0.5110	23.67/0.4808	20.74/0.4841	21.46/0.6138

(Continued)

Table 1: Continued

Algorithm	Scale	Set5 PSNR/ SSIM	Set14 PSNR/ SSIM	BSDS100 PSNR/ SSIM	Urban100 PSNR/ SSIM	Manga109 PSNR/ SSIM
SRGAN [35]	×8	23.04/0.6260	21.57/0.4950	21.78/0.4422	19.64/0.4685	20.42/0.6253
A+ [21]	×8	25.53/0.6548	23.99/0.5535	24.21/0.5156	21.37/0.5193	22.39/0.6454
SRCNN [22]	×8	25.34/0.6471	23.86/0.5443	24.14/0.5043	21.29/0.5133	22.46/0.6606
ESPCN [15]	×8	25.75/0.6738	24.21/0.5109	24.37/0.5277	21.59/0.5420	22.83/0.6715
FSRCNN [33]	×8	25.42/0.6440	23.94/0.5482	24.21/0.5112	21.32/0.5090	22.39/0.6357
VDSR [34]	×8	25.73/0.6743	23.20/0.5110	24.34/0.5169	21.48/0.5289	22.73/0.6688
LapSRN [27]	×8	26.15/0.7028	24.45/0.5792	24.54/0.5293	21.81/0.5555	23.39/0.7068
DRCN [26]	×8	25.93/0.6743	24.25/0.5510	24.49/0.5168	21.71/0.5289	23.20/0.6686
DBPN [36]	×8	27.25/0.7860	25.14/0.6490	24.90/0.6020	22.72/0.6310	25.14/0.7980
EDSR [28]	×8	27.03/0.7740	24.91/0.6420	24.81/0.5985	22.51/0.6221	24.69/0.7841
MDRN (Our)	×8	27.21/0.7915	25.06/0.6516	24.78/0.6045	22.77/0.6426	25.11/0.7985

4.3 Qualitative Analysis

The seven algorithms of Bicubic, A+ [21], SRCNN [22], FSRCNN [33], VDSR [34], LapSRN [27], and DRCN [26] are selected for qualitative comparison with the algorithm of this paper. Figs. 5 to 8 show the super-resolution visual effects of the 4 standard test sets under ×4 magnification. We compared the “baby” image in the Set5 standard test data set, the “baboon” image in the Set14 standard test set, the “76053” image in the BSDS100 test set, and the “img001” image in the Urban100 test set. Manga109 is a comic data set. Since this training set does not contain any comic images, the renderings are not displayed for comparison. It can be seen from a subjective vision that the local details in the reconstruction results of other methods are seriously blurred and distorted, and the reconstruction details of our model are better and clearer. For example, in Fig. 8 in the “image001” image, our reconstruction result is significantly clearer than other models, which is very close to the original image.

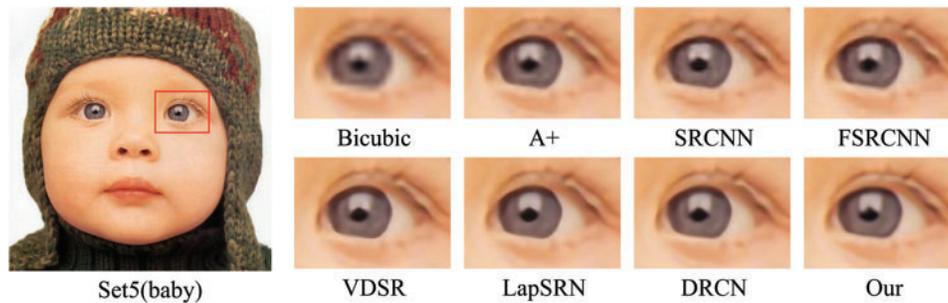


Figure 5: Visualized results of (baby) ×4 SR on Set5

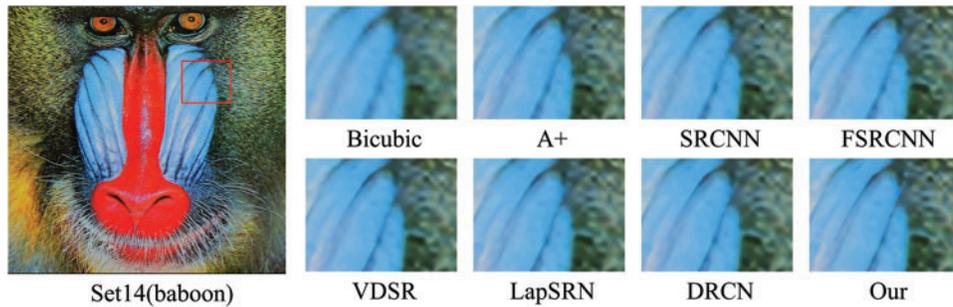


Figure 6: Visualized results of (baboon) $\times 4$ SR on Set14

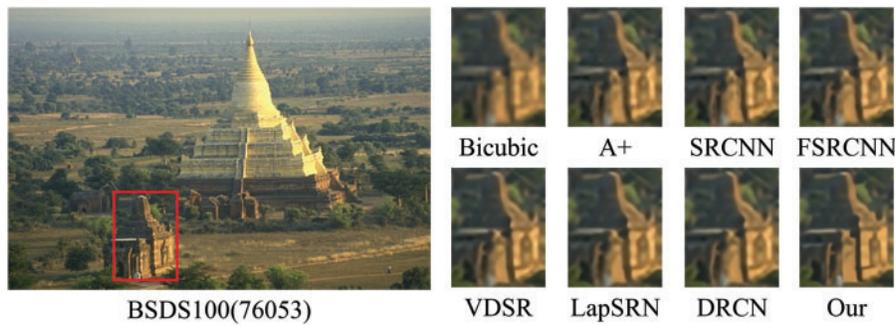


Figure 7: Visualized results of (76053) $\times 4$ SR on BSDS100

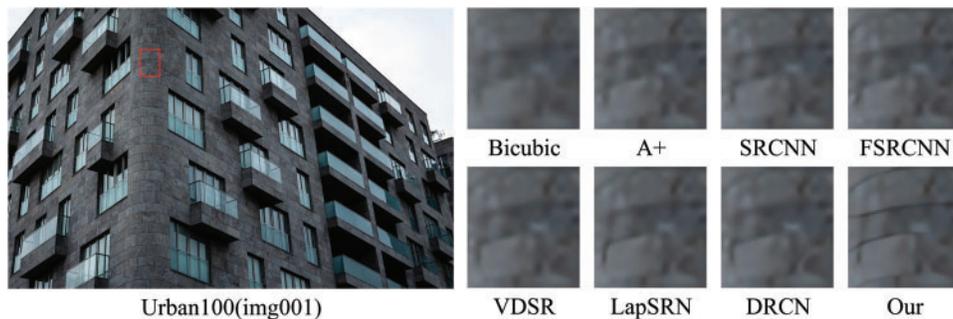


Figure 8: Visualized results of (img001) $\times 4$ SR on Urban100

5 Conclusion

In this paper, we propose an MDRN framework for single-image super-resolution, which uses residual blocks to fully acquire image details to obtain accurate SR images. Dual learning is used to improve the model's performance by introducing additional constraints for the reconstruction of low-resolution images, reducing the possible mapping space from low-resolution images to high-resolution images, and reducing the possible mapping space from low-resolution images to high-resolution images. From the results of the experimental analysis, our method has achieved good results. Despite these promising results, questions remain. I look forward to more ways to solve such problems in the future.

Funding Statement: Our research is funded by National Key R&D Program of China (2021YFC3320302) and Network threat depth analysis software (KY10800210013).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] N. Zhang, Y. C. Wang, X. Zhang and D. D. Xu, “A review of single image super-resolution based on deep learning,” *Acta Automatica Sinica*, vol. 46, no. 12, pp. 2479–2499, 2020. <https://doi.org/10.16383/j.aas.c190031>.
- [2] Y. F. Zhang, Y. Liu, C. Jiang and X. Cheng, “A curriculum learning approach for single image super resolution,” *Acta Automatica Sinica*, vol. 46, no. 2, pp. 274–282, 2017. <https://doi.org/10.3970/cmc.2017.053.133>.
- [3] C. Y. You, G. Li, Y. Zhang, X. L. Zhang, H. M. Shan *et al.*, “CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE),” *IEEE Transactions on Medical Imaging*, vol. 39, no. 1, pp. 188–203, 2020. <https://doi.org/10.1109/TMI.2019.2922960>.
- [4] Y. Tan, J. Cai, S. Zhang, W. Zhong and L. Ye, “Image compression algorithms based on super-resolution reconstruction technology,” in *2019 IEEE 4th Int. Conf. on Image, Vision and Computing (ICIVC)*, Xiamen, China, pp. 162–166, 2019.
- [5] D. W. Zhou, L. J. Zhao, R. Duan and X. L. Chai, “Image super-resolution based on recursive residual networks,” *Acta Automatica Sinica*, vol. 45, no. 6, pp. 1157–1165, 2019. <https://doi.org/10.16383/j.aas.c180334>.
- [6] W. W. W. Zou and P. C. Yuen, “Very low resolution face recognition problem,” *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 327–340, 2012. <https://doi.org/10.1109/TIP.2011.2162423>.
- [7] A. K. Cherian and E. Poovammal, “A novel alphasrgan for underwater image super resolution,” *Computers, Materials & Continua*, vol. 69, no. 2, pp. 1537–1552, 2021.
- [8] J. Zhang, Z. Wang, Y. Zheng and G. Zhang, “Design of network cascade structure for image super-resolution,” *Journal of New Media*, vol. 3, no. 1, pp. 29–39, 2021.
- [9] G. Zhang, Y. Ge, Z. Dong, H. Wang, Y. Zheng *et al.*, “Deep high-resolution representation learning for cross-resolution person re-identification,” *IEEE Transactions on Image Processing*, vol. 30, pp. 8913–8925, 2021.
- [10] X. Sun, X. G. Li, J. F. Li and L. Zhuo, “Review on deep learning based image super-resolution restoration algorithms,” *Acta Automatica Sinica*, vol. 43, no. 5, pp. 697–709, 2017. <https://doi.org/10.16383/j.aas.2017.c160629>.
- [11] J. Li, F. Fang, K. Mei and G. Zhang, “Multi-scale residual network for image super-resolution,” in *Proc. of the European Conf. on Computer Vision (ECCV)*, Munich, Germany, pp. 517–532, 2018.
- [12] E. Agustsson and R. Timofte, “NTIRE 2017 challenge on single image super-resolution: Methods and results,” in *2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, pp. 126–135, 2017.
- [13] K. He, X. Zhang, S. Ren and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770–778, 2016.
- [14] W. El-Shafai, A. M. Ali, E. M. El-Rabaie, N. F. Soliman, A. D. Algarni *et al.*, “Automated covid-19 detection based on single-image super-resolution and cnn models,” *Computers, Materials & Continua*, vol. 70, no. 1, pp. 1141–1157, 2022.
- [15] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken *et al.*, “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 1874–1883, 2016.
- [16] Y. Guo, J. Chen, J. D. Wang, Q. Chen, J. Z. Cao *et al.*, “Closed-loop matters: Dual regression networks for single image super-resolution,” in *2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 5406–5415, 2020.

- [17] 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. <https://doi.org/10.1109/TIP.2006.877407>.
- [18] K. Zhang, X. Gao, D. Tao and X. Li, "Single image super-resolution with non-local means and steering kernel regression," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, 2012. <https://doi.org/10.1109/TIP.2012.2208977>.
- [19] R. Timofte, V. De and L. V. Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *2013 IEEE Int. Conf. on Computer Vision*, Sydney, NSW, Australia, pp. 1920–1927, 2013.
- [20] R. Timofte, V. De Smet and L. Van Gool, "A+: Adjusted anchored neighborhood regression for fast super-resolution," in *Asian Conf. on Computer Vision*, Singapore, pp. 111–126, 2014.
- [21] T. Peleg and M. Elad, "A statistical prediction model based on sparse representations for single image super-resolution," *IEEE Transactions on Image Processing*, vol. 23, no. 6, pp. 2569–2582, 2014. <https://doi.org/10.1109/TIP.2014.2305844>.
- [22] C. Dong, C. C. Loy, K. He and X. Tang, "Learning a deep convolutional network for image super-resolution," in *European Conference on Computer Vision*, Zurich, Switzerland, pp. 184–199, 2014.
- [23] Y. Zhang, Y. Tian, Y. Kong, B. Zhong and Y. Fu, "Residual dense network for image super-resolution," in *2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 2472–2481, 2018.
- [24] T. Ying, J. Yang and X. M. Liu, "Image super-resolution via deep recursive residual network," in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 3147–3155, 2017.
- [25] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong *et al.*, "Image super-resolution using very deep residual channel attention networks," in *Proc. of the European Conf. on Computer Vision (ECCV)*, Munich, Germany, pp. 286–301, 2018.
- [26] J. Kim, J. K. Lee and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 1637–1645, 2016.
- [27] W. S. Lai, J. B. Huang, N. Ahuja and M. H. Yang, "Deep laplacian pyramid networks for fast and accurate super-resolution," in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 5835–5843, 2017.
- [28] B. Lim, S. Son, H. Kim, S. Nah and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, pp. 1132–1140, 2017.
- [29] J. Liu, W. Zhang, Y. Tang, J. Tang and G. Wu, "Residual feature aggregation network for image super-resolution," in *2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 2356–2365, 2020.
- [30] J. Zhou, J. Liu, J. Li, M. Huang, J. Cheng *et al.*, "Mixed attention densely residual network for single image super-resolution," *Computer Systems Science and Engineering*, vol. 39, no. 1, pp. 133–146, 2021.
- [31] J. Y. Zhu, T. Park, P. Isola and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *2017 IEEE Int. Conf. on Computer Vision (ICCV)*, Venice, Italy, pp. 2242–2251, 2017.
- [32] Z. Yi, H. Zhang, P. Tan and M. Gong, "DualGAN: Unsupervised dual learning for image-to-image translation," in *2017 IEEE Int. Conf. on Computer Vision (ICCV)*, Venice, Italy, pp. 2868–2876, 2017.
- [33] C. Dong, C. C. Loy and X. Tang, "Accelerating the super-resolution convolutional neural network," in *European Conf. on Computer Vision*, Amsterdam, The Netherlands, pp. 391–407, 2016.
- [34] J. Kim, J. K. Lee and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 1646–1654, 2016.

- [35] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 105–114, 2017.
- [36] M. Haris, G. Shakhnarovich and N. Ukita, “Deep back-projection networks for super-resolution,” in *2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 1664–1673, 2018.