International Journal of Computational Intelligence Systems
Vol. **12(2)**, 2019, *pp.* 1592–1601
DOI: https://doi.org/10.2991/ijcis.d.191209.001; ISSN: 1875-6891; eISSN: 1875-6883
https://www.atlantis-press.com/journals/ijcis/

An Advanced Deep Residual Dense Network (DRDN) Approach for Image Super-Resolution

Wang Wei[1,*], Jiang Yongbin[1], Luo Yanhong[2], Li Ji[1], Wang Xin[1], Zhang Tong[1,*]

[1] *School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China*
[2] *Hunan Children's Hospital, Changsha 410000, China*

ARTICLE INFO

Article History
Received 07 Aug 2019
Accepted 02 Dec 2019

Keywords
Deep residual dense network (DRDN)
Single image super-resolution
Fusion reconstruction
Residual dense connection
Multi-hop connection

ABSTRACT

In recent years, more and more attention has been paid to single image super-resolution reconstruction (SISR) by using deep learning networks. These networks have achieved good reconstruction results, but how to make better use of the feature information in the image, how to improve the network convergence speed, and so on still need further study. According to the above problems, a novel deep residual dense network (DRDN) is proposed in this paper. In detail, DRDN uses the residual-dense structure for local feature fusion, and finally carries out global residual fusion reconstruction. Residual-dense connection can make full use of the features of low-resolution images from shallow to deep layers, and provide more low-resolution image information for super-resolution reconstruction. Multi-hop connection can make errors spread to each layer of the network more quickly, which can alleviate the problem of difficult training caused by deepening network to a certain extent. The experiments show that DRDN not only ensure good training stability and successfully converge but also has less computing cost and higher reconstruction efficiency.

1. INTRODUCTION

Single image super-resolution reconstruction (SISR) is to reconstruct a corresponding high-resolution (HR) image based on a low-resolution (LR) image by a certain algorithm. The super-resolution reconstruction, implemented at the algorithm level, can help to break the limitations of some imaging devices. Therefore, it is widely used in medicine imaging [1], satellite images, security monitoring [2], and so on.

SISR is an ill-posed inverse problem. This indicates that there are many different solutions to reconstruct the HR image from the corresponding LR image. The early super-resolution reconstruction methods are mainly interpolation methods. Although this kind of method is relatively simple, the reconstruction effect is not so ideal. Other methods for image super-resolution reconstruction mainly include the method by using prior information of the image [3,4], internal patch recurrence method [5,6], and traditional learning-based reconstruction methods [7–9]. In recent years, with deep learning network showing strong learning ability, more deep learning effective reconstruction method has been widely used to solve the ill-posed problem of image super-resolution reconstruction.

Dong *et al.* firstly applied the deep learning network to realize image super-resolution reconstruction (SRCNN) [10]. SRCNN constructed a three-layer convolutional neural network for the mapping of LR image to HR image. Its reconstruction effect was significantly improved compared with the traditional learning algorithm.

With the success of deep neural network in ImageNet [11,12], the application of deep neural network in image super-resolution reconstruction has become an important research content. Accurate Image Super-Resolution (VDSR) [13] deepened the depth of convolutional neural network to 20 layers. In order to ensure the effective convergence of deep neural network in training, VDSR introduced global residual connection and gradient clipping technology. Inspired by the residual network model, deeply-recursive convolutional network (DRCN) [14] and deep recursive residual network (DRRN) [15] were proposed one after another. In addition to deepening the network, efficient sub-pixel convolutional neural network (ESPCN) [16] convolved directly on LR images, and finally used sub-pixel convolution layer to realize up-sampling process. Densely connected convolutional networks (DenseNet) [17] maximized the transmission of feature information between layers by densely connecting and the dense connection made full use of the feature maps obtained by each convolutional layer. Inspired by DenseNet, Tong *et al.* [18] firstly introduced dense network to realize super-resolution image reconstruction.

Besides deepening the network, many researchers also proposed other structures to get better reconstruction results. Dong *et al.* [19] proposed a compact hourglass-shape CNN structure, namely FSRCNN, for faster and better SR, and re-designed the SRCNN structure mainly in three aspects. Wei-Sheng *et al.* [20] proposed the Laplacian Pyramid Super-Resolution Network (LapSRN). They

* *Corresponding authors. Email: wangwei@csust.edu.cn, 1214304762@qq.com*

trained LapSRN with deep supervision using a robust Charbonnier loss function and achieved high-quality reconstruction. Namhyuk *et al.* [21] designed an architecture that implements a cascading mechanism upon a residual network, and also presented variant models of the proposed cascading residual network (CARN) to further improve efficiency. Xiangxiang *et al.* [22] proposed a fast, accurate, and lightweight super-resolution (FALSR) method to get the balance between the restoration capacity and the simplicity of models.

Although the above methods have achieved good reconstruction results, with the deepening of network layers, the network parameters will increase dramatically and the network itself becomes more difficult to train and converge. On the other hand, many useful features obtained in each layer are often neglected in above networks. Moreover, information of each layer is often fused to a certain middle layer, rather than directly used for the final reconstruction, which will affect the utilization of layer's feature information.

Aiming at the above problems, this paper proposes a deep residual dense network (DRDN), which can effectively utilize feature information in and between the layers by connecting dense blocks (DB) to deepen the network. The main contributions of our paper are as follows:

1. Residual dense connection. Dense connection can make full use of local hierarchical information and provide more hierarchical feature information for final reconstruction. At the same time, the residual learning of each dense connected network block through the skip connection can alleviate the problems such as the training difficulty when deepening the network.
2. Fusion reconstruction. When deepening the network, the intermediate reconstruction results of each DB can be obtained by introducing residual skip connection to each DB, and the intermediate results of each DB can be fused and reconstructed, so as to obtain the final reconstruction results. At the same time, Multi-hop connection can make errors spread to each layer of the network more quickly, which can alleviate the problem of difficult training caused by deepening the network.

2. DRDN

Based on the comparison with DRCN and DRRN, this section first introduces the network structure of DRDN, and then analyzes the local feature fusion approach and the global residual fusion reconstruction approach in detail. In the analysis, we focus on the improvement of DRDN compared with DRCN and DRRN and why DRDN is superior to them.

2.1. Architecture of DRDN

Different from the cascade convolutional layer unit in DRCN and DRRN, DRDN adopts DB [17] (shown in Figure 1) as the basic block of network architecture. DRDN combines the ways of deepening the network in DRCN and DRRN, and connects each block in the same way through identical skip connection. Then after adding each basic block and identity connection, the outputs are taken as the intermediate results in reconstruction. Finally, all intermediate results are fused through the reconstruction layer (convolution kernel is 1×1) to obtain HR reconstructed images. The network structure of DRCN, DRRN, and DRDN is shown in Figure 2.

DRDN mainly consists of three parts: shallow feature extraction, residual-dense network, and fusion reconstruction network. In the part of shallow feature extraction, two convolution layers with the convolution kernel size of 3×3 and 1×1 respectively are used to extract shallow features. Specifically, we use a multiple dense connected convolution layer (the convolution kernel is 3×3) and a feature contraction layer with the convolution kernel size of 1×1 as the basic block (also called DB) of the network. For residual learning, an identical skip connection is introduced between each DB and the shallow feature extraction layer. Multiple DBs are connected in the same way to deepen the network continuously. In the fusion reconstruction network, the results of each residual-DB are taken as intermediate results, and the final reconstruction results are obtained by weighting and summing these intermediate results.

2.2. Local Feature Fusion

In the reconstruction, we expect to make use of more LR image information, which needs the network reach a certain depth to obtain better reconstruction effects. When deepening the network, instead of taking two ordinary convolutional layers as the basic network block in DRRN, densely connected convolutional network is taken as the basic block in DRDN, as shown in Figure 3. In Figure 3, $f_0\,(LR)$ and $f_0'\,(LR)$ represent the outputs of the shallow feature extraction layers in DRRN and DRDN respectively. U_i and U_i' represent the outputs of the ith blocks in DRRN and DRDN

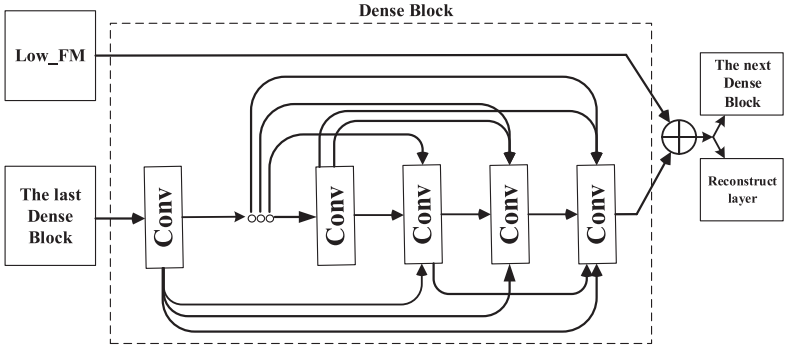
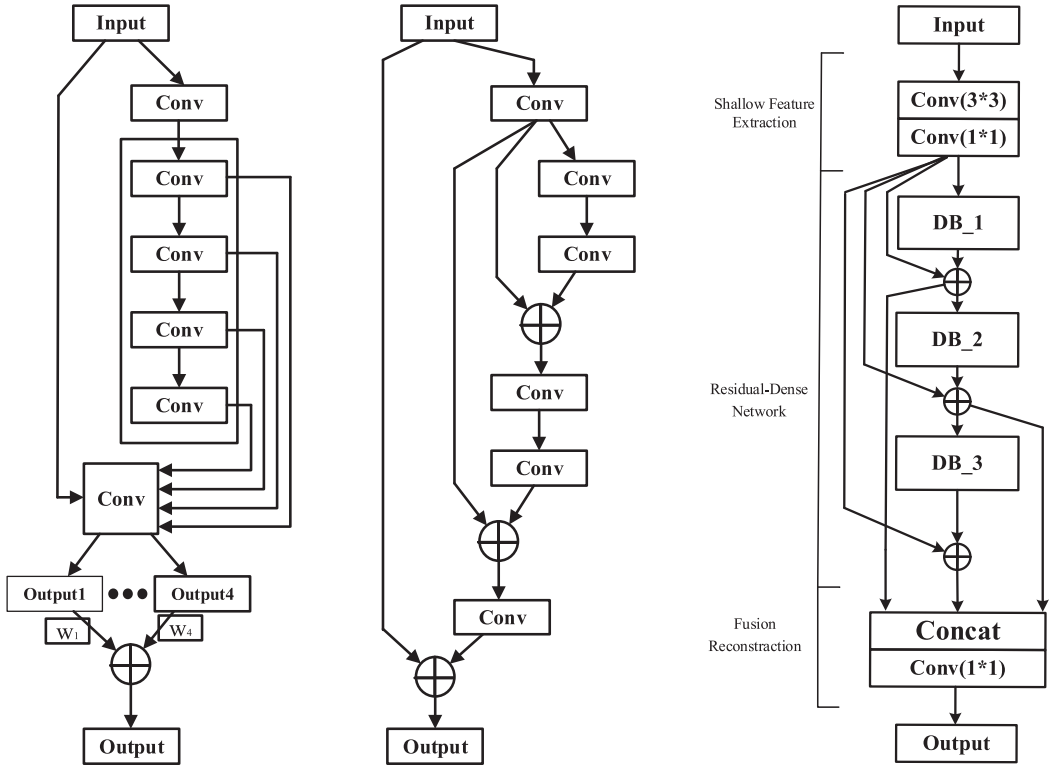
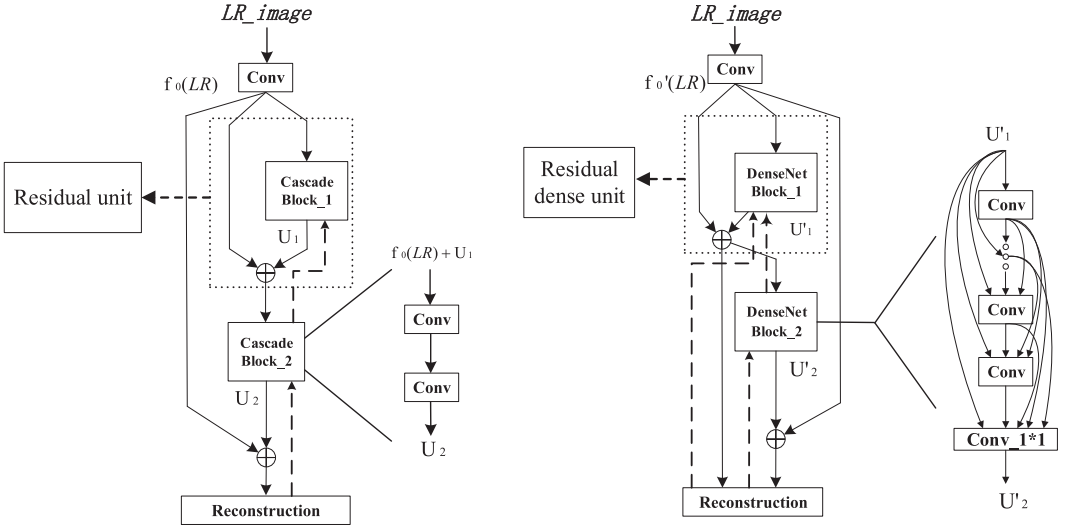


Figure 1 | Dense Block (DB).



(a) DRCN (b) DRRN (c) DRDN(Ours)

Figure 2 | The diagram of network architecture.



（a）DRRN （b）DRDN

Figure 3 | Simplified net architecture of deep recursive residual network (DRRN) and deep residual dense network (DRDN).

respectively. $H_i()$ and $H_i'()$ represent the input–output function relationships of the ith blocks in DRRN and DRDN respectively.

DBs are connected by multi-layer densely connected convolution layer and convolution layer with kernel size of 1×1, so that the DB can make full use of the feature information of each convolutional layer. Then, the intermediate result can be obtained via a convolution kernel with size of 1×1 from the feature information of the convolutional layer. Similar to DRRN, both the outputs of the above block in DRDN and the shallow feature are used as the inputs of the next block. The formula is expressed as:

$$U_{i+1}' = H_{i+1}'\left(f_0'(LR) + U_i'\right) \tag{1}$$

In addition, for the diversity of feature information, different from DRRN adopting block weight sharing, DRDN trains the weight parameters in each block independently.

2.3. Global Residual Fusion Reconstruction

As shown in Figure 3, in DRRN, only the output of the last residual block is directly used to reconstruct the HR image, and the formula is expressed as:

$$HR = Rec\left(U_{last} + f_0(LR)\right) \tag{2}$$

where LR and HR represent the input LR image and the reconstructed HR image respectively. $Rec()$ is the output function of the last convolutional layer. In DRDN, by introducing identity connection between the shallow feature extraction layer and each block, the outputs of each DB and the shallow feature extraction layer can be added to obtain an intermediate reconstruction result. All intermediate results are taken as inputs of the final fusion reconstruction layer, and the final fusion reconstruction layer is a convolution layer with kernel size of 1×1, which is equivalent to weighted sum of all intermediate results.

$$HR = \sum_{i=1}^{n} w_i\left(U_i' + f_0'(LR)\right) \tag{3}$$

where n is number of the blocks in the entire network, and w_i is the corresponding weight which obtained by training. Through the introduction of multi-path identical skip connection, the intermediate reconstruction results are fused to obtain the final result, which is the fusion and utilization of the global feature information of the deep network. In this way, local and global hierarchical feature information can be fully utilized in network reconstruction.

At the same time, by these skip connections, the error in the deeper layer can be propagated to the shadow convolution layer faster and more directly. As shown in Figure 3, the dashed lines indicate the error backward propagation path. In DRRN, it can be seen that the reconstruction error needs to pass through all the blocks to reach the previous block. This is the main reason for gradient vanishing problem in deep networks. However, for DRDN, the identical skip connection can make the reconstruction error directly propagated to these blocks, which can alleviate the problems of gradient vanishing to a certain extent and accelerate the training. By using residual learning between DBs, the whole network can simultaneously complete local hierarchical features fusion and global features fusion, both of which provide rich hierarchical features information for final reconstruction.

3. EXPERIMENTS AND RESULTS ANALYSIS

3.1. Training

As in VDSR [13] and DRRN [15], MSE is taken as the loss function in training. Given a train set $\{x_i, y_i\}_{i=1}^{N}$, the specific expression of loss function is:

$$L(\Theta) = \frac{1}{2N}\sum_{i=1}^{N} \|y_i - HR(x_i)\|^2 \tag{4}$$

where N is the number of training patches, x_i is the ith LR image, y_i is the corresponding HR label image, Θ represents the parameter sets, and $HR(x_i)$ represents the reconstructed image. The method proposed in Reference [23] is used to initialize the network parameters and optimize the loss function. The training is based on Caffe framework, and the mini-batch size is 64. The initial value of learning rate is 0.0001, which is reduced to half of the previous after 200,000 iterations.

3.2. Experiments Settings

As in VDSR, DRCN, and DRRN, the experiments also use the same data set containing 291 images [13–15] as the training data set, Set14, Urban100, BSD100, and other commonly public data sets are adopted as the test data set. Some graphic examples of the train dataset are shown in Figure 4.

In the stage of data preparation, the training data is artificially expanded by rotating and flipping the 291 images in 4 directions respectively, as shown in Figure 5(a). Data preprocessing is required before training, and the specific operation is as follows:

1. The RGB images are converted into YCbCr images, and only Y-channel is used for training and testing, as well as in SRCNN [10], DRCN [14], and DRRN [15].
2. Scale transformation and degradation processing are carried out on the images obtained in step (1) to obtain the LR images and the corresponding tag images with different magnification multiples (×2, ×3, ×4), as shown in Figure 5(b).
3. The images obtained in step (2) are cropped into image blocks with a size of 40×40 for network training, which is shown in Figure 5(c). These LR images with different magnification scales and label image pairs are used to train the same network model, so that the network model is suitable for reconstruction with different magnification scales.

So, the dataset containing 291 images is rotated and flipped to $291 \times 8 = 2328$ images. Among them, 2164 are training images and 164 are test images. After the processing in step (2), we can get $2164 \times 3 = 6492$ training images and 492 test images. After the tailoring in step (3), we can get about 648,000 image patches for training.

The specific operations of image reconstruction are as follows:

1. The size of the LR image is adjusted to 320×240, and then it is cut to 40×40 image patches.

Figure 4 | Examples of train data sets.

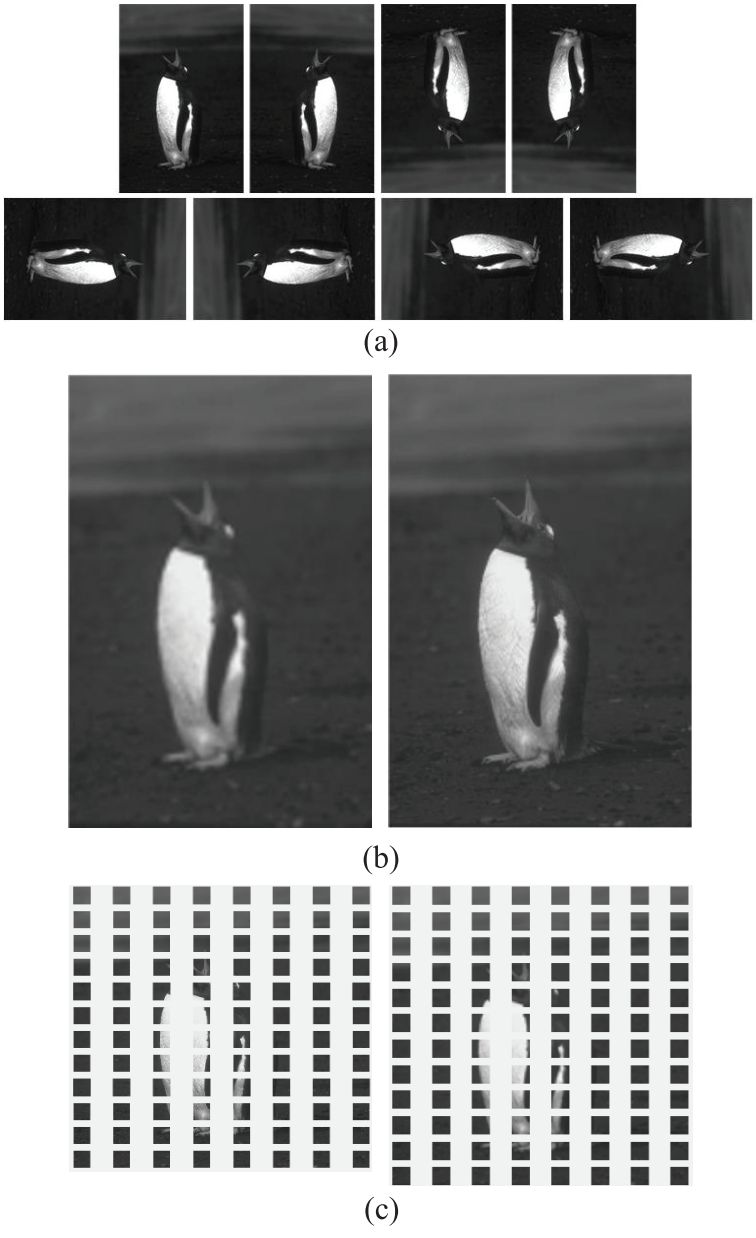


Figure 5 | (a) Training data expansion diagram, (b) Low-resolution image (left) and high-resolution image (right), and (c) Low resolution image patches (40 × 40) and label image patches (40 × 40).

2. DRDN is used to reconstruct image patches, and the reconstructed image patches are spliced according to the original locations to get the reconstructed HR images.

For experimental comparison, we trained the two networks (DRRN and DRDN) in the same hardware and software environment. The specific hardware and software environment configuration are shown in Table 1. Specially, when training DRRN, we carried out the training completely according to the network structure and parameter settings in Reference [15].

Table 1 | Experimental environment configuration.

Configuration	Parameter
OS	Ubuntu 16.04
CPU	Intel i7 3.30GHz
GPU	GTX1080Ti (11G)
RAM	16G/DDR3/2.10GHz
cuDNN	CuDNN 7.0
CUDA	CUDA9.0
Frame	Caffe

Table 2 | Network parameters and reconstruction time.

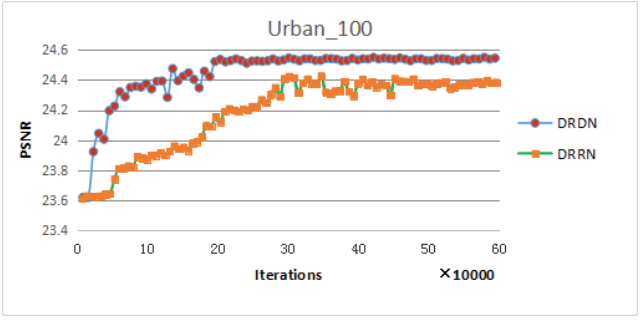
	DRRN	DRDN
Network layers	52	50
Convolution kernels size	51 (3 × 3) + 1 (1 × 1)	49 (3 × 3) + 1 (1 × 1)
Reconstruction time (Set14)	3.942 s	2.824 s

DRDN = deep residual dense network; DRRN = deep recursive residual network.

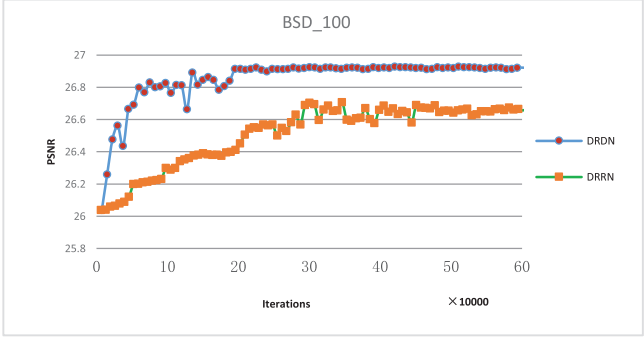
3.3. Experiments Analysis

1. **Comparison of network parameters and reconstruction time**
Table 2 shows the comparison of the two network parameters. DRRN deepens the network to 52 layers by introducing recursive units, and ours DRDN makes the network depth reach 50 layers by stacking multiple DBs. Meanwhile, the size of convolution kernel of each layer in the two networks is mainly 3 × 3. Under the same hardware and software environment, the average reconstruction time of the two networks on Set14 is as follows: DRRN takes 3.942s and DRDN takes 2.824s. This reflects that DRDN has less computing cost compared with DRRN.

2. **Network robustness and convergence speed**
Through repeated training of DRRN and DRDN, we find that under the same training environment, the training of DRRN is unstable, and often difficult to converge. There is only about one successfully converge among 4 times training. By contrast, DRDN can ensure good training stability and successfully converge among 4 training. Figure 6 shows the curves of peak signal to noise ratio (PSNR) with different iterations on Urban100 and BSD100 datasets. It can be clearly found that DRDN has a significantly faster convergence rate than that of DRRN. DRDN tends to converge after about 200,000 iterations, while DRRN gradually converges after about 400,000 iterations. This fully demonstrates that the introduction of multi-path reconstruction of DRDN with skip connection is conducive to accelerating network convergence. In addition, the experimental



(a)



(b)

Figure 6 | (a) The curves of peak signal to noise ratio (PSNR) values and iterations on Urban100, and (b) the curves of PSNR values and iterations on BSD100.

results also show that DRDN achieves better reconstruction effects.

3. **Fusion reconstruction**
Figure 7 shows the intermediate results obtained by each residual-DB, and the final HR image is obtained through fusion reconstruction. In Figure 7, the above images from left to right are 5 intermediate results, obtained from the 1–5 residue-intensive blocks, which named result_1- result_5, respectively. As can be seen from the figure, the result_1 obtained by the first residue-intensive block at the lower level is still fuzzy. With the stacking of network blocks, the feature information is more abundant, and the intermediate result obtained by the residual-DB at the deeper level is clearer. In the end, the residual-DBs of different depths can improve the feature information of different levels for reconstruction, so that the final reconstructed image has clearer and richer information.

4. **Visual effect comparison of reconstructed images**
Figure 8 shows the reconstructed images of VDSR, DRRN, and DRDN. Their PSNRs are shown in Table 3. In Figure 8, there are 3 test images. The images from left to right are the LR image, the reconstruction result of VDSR, the reconstruction result of DRRN, and the reconstruction result of DRDN, respectively. Because the LR image in Figure 8(b) contains more image details and textures, the results in Figure 8(b) reconstructed by 3 methods are better than those in Figure 8(a) and 8(c). Compared with other methods, DRDN achieves the best reconstruction effect, both in visual effect and PSNR value. This indicates that the rich feature information extracted by DNDN is helpful to improve the reconstruction effect.
Figure 9 shows detail parts corresponding to the images of Figure 8. It can be seen from the figure that the edge details of the reconstructed image obtained by DRDN are clearer and the lines are smoother.

In addition to the above experiments, we also test different approaches under our own experimental conditions, including SRCNN [10], FSRCNN [19], DRCN [14], LapSRN [20], DRRN [15], CARN [21], FALSR [22], and DRDN. The experimental results are shown in Table 4. PSNR is the average value of 200,000 iterations after the network is stable. The data sets used in the experiment are BSD100 and Urban100. As can be seen from Table 4, our method has achieved better results compared with other methods. On BSD100, the effect of DRDN is a little worse than that of FALSR, but better than those of other methods.

4. CONCLUSIONS

In this paper, a new network structure named DRDN is proposed for image super-resolution. In DRDN, dense connection can make full use of local hierarchical information and provide more hierarchical feature information. Each DB can get an intermediate result, and the final reconstructed image can be obtained by weighted summation of all intermediate results. Reconstruction is the full fusion and direct utilization of the network's global hierarchical feature information. Furthermore, Multi-hop connection can make error spread to each layer of the network more quickly, which can alleviate the problem of difficult training caused by deep network to a certain extent. The experimental results show that both the reconstruction visual effect and PSNR of DRDN are better than those of CDSR and DRRN, and the details are clearer. Compared with the other 7 methods in Table 4, our DRDN approach also achieves good reconstruction results.

In the experiments, we found that different hardware/software and training methods have great impacts on the experimental results. Therefore, our future research is mainly about the adaptability of the methods to the environment and the further optimization of the network structure.

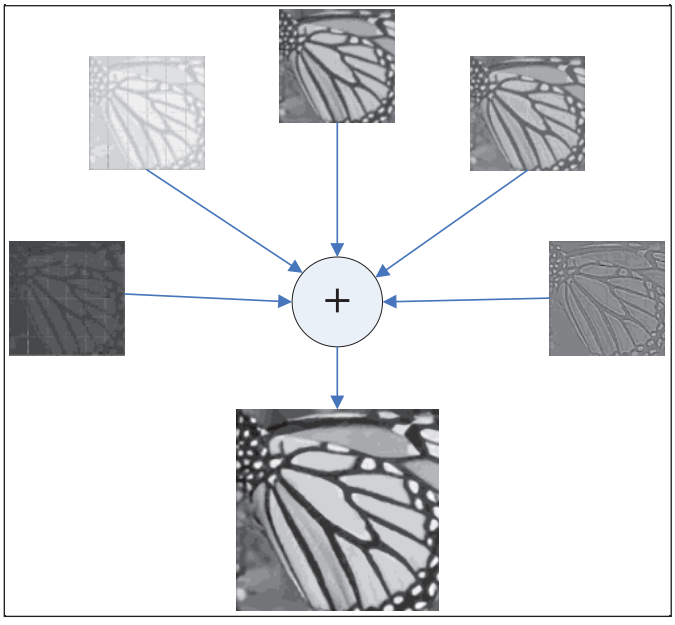


Figure 7 | Fusion reconstruction result.

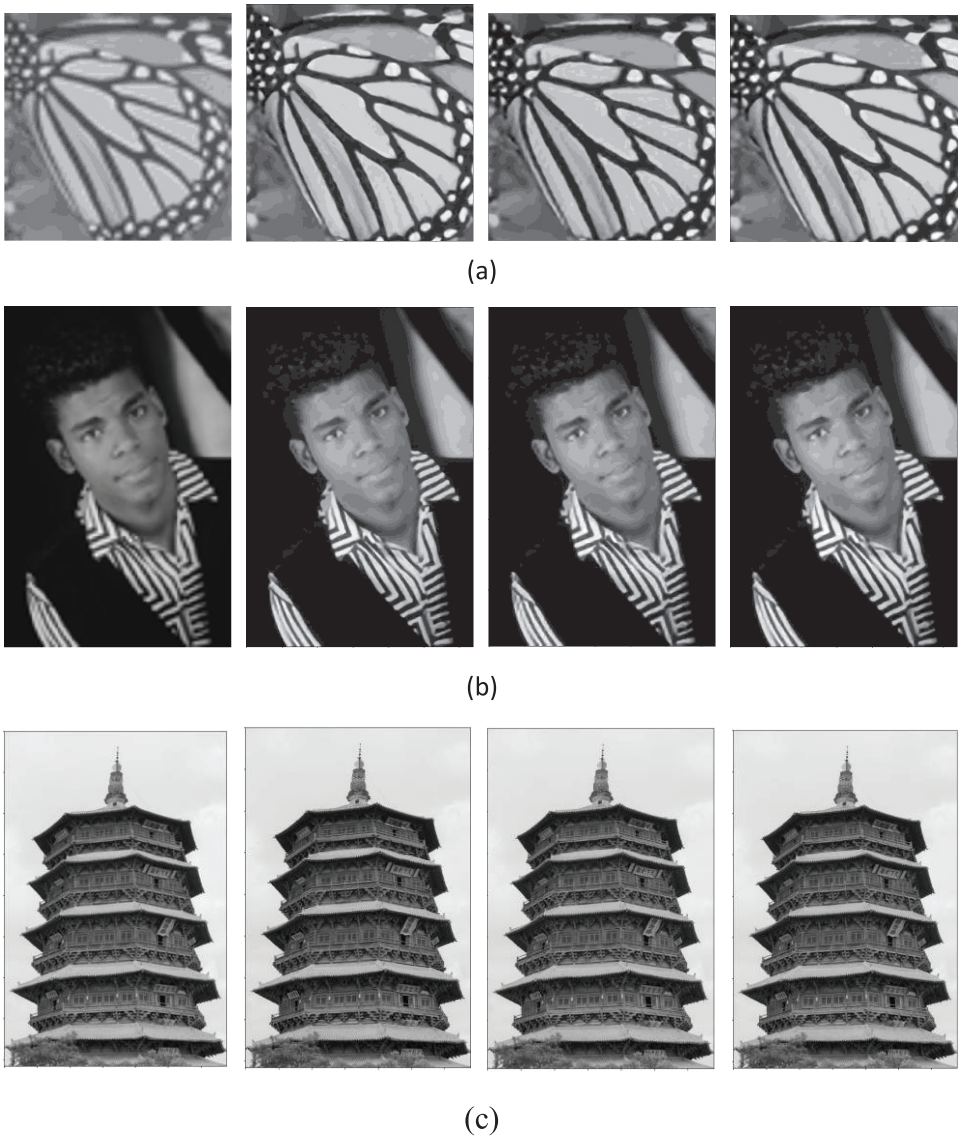


Figure 8 | Visualization of reconstruction results.

Table 3 | PSNR of the reconstructed images.

	VDSR	DRRN	DRDN
(a)	26.5500	25.8549	26.6885
(b)	31.3138	30.4452	31.6819
(c)	30.3441	30.3072	30.4184

DRDN = deep residual dense network; DRRN = deep recursive residual network; PSNR = peak signal to noise ratio.

Table 4 | PSNR of different networks.

Approach	Scale	BSD100	Urban100
SRCNN [10]	4	21.12	20.13
FSRCNN [19]	4	21.14	20.21
DRCN [14]	4	22.23	21.08
LapSRN [20]	4	23.01	21.17
DRRN [15]	4	26.67	24.38
CARN [21]	4	26.72	24.52
FALSR [22]	4	26.86	24.55
DRDN(ours)	4	26.92	24.53

CARN = cascading residual network; DRDN = deep residual dense network; DRRN = deep recursive residual network; FALSR = fast, accurate and lightweight super-resolution; LapSRN = Laplacian Pyramid Super-Resolution Network; PSNR = peak signal to noise ratio.

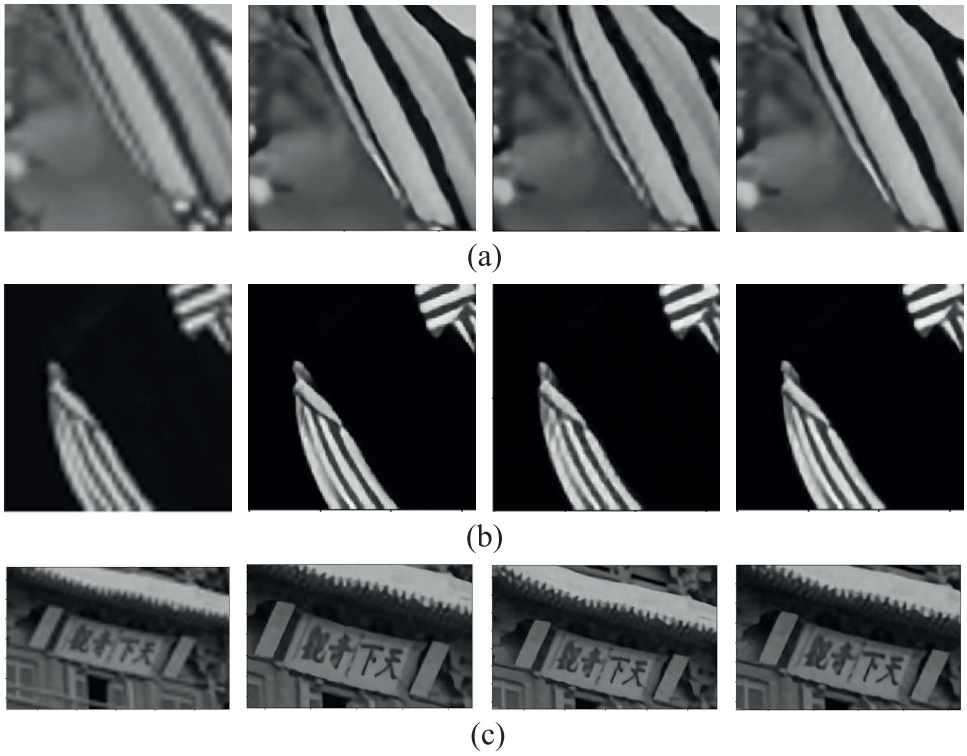


Figure 9 | Details corresponding to the images in Figure 8.

CONFLICT OF INTEREST

Authors have no conflict of interest to declare.

Funding Statement

National Defense Pre-Research Foundation of China (7301506); National Natural Science Foundation of China (61070040); Education Department of Hunan Province (17C0043); Hunan Provincial Natural Science Fund (2019JJ80105).

REFERENCES

[1] W. Shi, J. Caballero, C. Ledig, *et al.*, Cardiac image super-resolution with global correspondence using multi-atlas PatchMatch, in International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Nagoya, 2013, pp. 9–16.

[2] W.W.W. Zou, P.C. Yuen, Very low resolution face recognition problem, IEEE Trans. Image Process. 21 (2012), 327–340.

[3] J. Sun, Z. Xu, *et al.*, Image super-resolution using gradient profile prior, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Anchorage, 2008.

[4] K.I. Kim, Y. Kwon, Single-image super-resolution using sparse regression and natural image prior, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010), 1127–1133.

[5] D. Glasner, S. Bagon, M. Irani, Super-resolution from a single image, in IEEE 12th International Conference on Computer Vision (ICCV), IEEE, Kyoto, 2009.

[6] J.B. Huang, A. Singh, N. Ahuja, Single image super-resolution from transformed self-exemplars, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, 2015.

[7] E. Perez-Pellitero, J. Salvador, J. Ruiz-Hidalgo, *et al.*, PSyCo: manifold span reduction for super resolution, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, 2016.

[8] J. Salvador, E. Perez-Pellitero, Naive bayes super-resolution forest, in 2015 IEEE International Conference on Computer Vision (ICCV), IEEE, Santiago, 2015.

[9] S. Schulter, C. Leistner, H. Bischof, Fast and accurate image upscaling with super-resolution forests, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, 2015.

[10] C. Dong, C.C. Loy, K. He, *et al.*, Image super-resolution using deep convolutional networks, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2014), 295–307.

[11] W. Wei, Y. Yujing, W. Xin, *et al.*, The development of convolution neural network and its application in image classification: a survey, Opt. Eng. 58 (2019), 040901.

[12] O. Russakovsky, J. Deng, H. Su, *et al.*, ImageNet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (2014), 211–252.

[13] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, 2016.

[14] J. Kim, J.K. Lee, K.M. Lee, Deeply-recursive convolutional network for image super-resolution, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, 2016.

[15] Y. Tai, J. Yang, X. Liu, Image super-resolution via deep recursive residual network, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, 2017.

[16] W. Shi, J. Caballero, F. Huszár, *et al*., Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, 2016.
[17] G. Huang, Z. Liu, L. Maaten, *et al*., Densely connected convolutional networks, in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, 2017.
[18] T. Tong, G. Li, X. Liu, *et al*., Image super-resolution using dense skip connections, in 2017 IEEE International Conference on Computer Vision (ICCV), IEEE, Venice, 2017.
[19] C. Dong, C.C. Loy, X. Tang, Accelerating the super-resolution convolutional neural network, in ECCV, Amsterdam, 2016.
[20] L. Wei-Sheng, H. Jia-Bin, A. Narendra, *et al*., Deep laplacian pyramid networks for fast and accurate super-resolution, in IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[21] A. Namhyuk, K. Byungkon, S. Kyung-Ah, Fast, accurate, and, lightweight super resolution with cascading residual network, arXiv: 1803.08664, 2018.
[22] C. Xiangxiang, Z. Bo, M. Hailong, *et al*., Fast, accurate and lightweight super-resolution with neural architecture search, arXiv: 1901.07261, 2019.
[23] D. Kingma, J. Ba, Adam: a method for stochastic optimization, in International Conference on Learning Representations, 2014.