


Super-Resolution Reconstruction of Remote Sensing Images Based on Symmetric Local Fusion Blocks

Xinqiang Wang, Tianjin Sino-German University of Applied Sciences, China*

 <https://orcid.org/0000-0002-8230-482X>

Wenhuan Lu, Tianjin University, China

ABSTRACT

In view of the rich information and strong autocorrelation of remote sensing images, a super-resolution reconstruction algorithm based on symmetric local fusion blocks is proposed using a convolutional neural network based on local fusion blocks, which improves the effect of high-frequency information reconstruction. By setting local fusion in the residual block, the problem of insufficient high-frequency feature extraction is alleviated, and the reconstruction accuracy of remote sensing images of deep networks is improved. To improve the utilization of global features and reduce the computational complexity of the network, a residual method is used to set the symmetric jump connection between the local fusion blocks to form the symmetry between them. Experimental results show that the reconstruction results of 2-, 3-, and 4-fold sampling factors on the UC Merced and nwpu-resisc45 remote sensing datasets are better than those of comparison algorithms in image clarity and edge sharpness, and the reconstruction results are better in objective evaluation and subjective vision.

KEYWORDS

Convolutional Neural Network, Global Symmetric Connection, Local Fusion, Remote Sensing Image, Residual-Dense Network, SR-SLFB, Super-Resolution Reconstruction, Symmetric Residual

1. INTRODUCTION

Image super-resolution reconstruction aims to reconstruct high-resolution images with richer details from single or multiple low-resolution images. It has broad application prospects in public security surveillance, medical image imaging, satellite remote sensing, and other fields. In the field of satellite remote sensing, high-resolution (HR) images with rich texture details are particularly important in target detection and recognition (Ding et al., 2016; Wu et al., 2015; Cui et al., 2018), disaster monitoring (Shamsolmoali et al., 2019), land cover classification (Michael et al., 2018), and the processing and

DOI: 10.4018/IJISP.319019

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

analysis of remote sensing images. Remote sensing image is huge and widely used, including some confidential information and special features. The application in urban public security mainly includes traffic system, emergency system, public security emergencies, etc. It is one of the important means to obtain information in urban construction. In remote sensing image acquisition, it is impossible to obtain high-resolution images due to the limitation of imaging conditions in the detection imaging system. Many researchers use algorithms to obtain higher-resolution images. The traditional super-resolution reconstruction algorithm has limited feature extraction and expression ability, resulting in a poor image effect. Therefore, the efficient reconstruction of high-quality, high-resolution images to meet the application requirements in satellite remote sensing is a concern.

2. BACKGROUND

ISR reconstruction algorithms include difference-based (Bätz, Eichenseer et al., 2015; Wei, 2016), reconstruction-based (Papayan & Elad, 2015; Xu et al., 2013), and learning-based (Lai et al., 2017; Wang et al., 2022) methods. While simple and effective, the classical difference-based method relies heavily on prior knowledge of natural images, resulting in obvious jagged edges of reconstructed images and limited restoration of detailed information. The reconstruction-based approach utilizes powerful image priori, such as non-local self-similar (Dong et al., 2013), sparse (Yang et al., 2010) and denoising (Zhang et al., 2017) priori, which can flexibly rebuild an HR image of relatively high quality. However, due to limited prior knowledge and time consumption of the optimization process, the restoration effect of high-frequency details of images is poor. Most methods to reconstruct remote sensing images are based on this method, and their performance is limited.

There is great interest in deep learning technology with strong characteristic learning and expression ability. With its end-to-end training mode and superior performance, it is seeing increased practical application. Dong et al. (Dong et al., 2014) of the Chinese University of Hong Kong first applied deep learning to ISR reconstruction, adopting a three-layer super-resolution convolutional neural network (SRCNN) to reconstruct an image's super-resolution. An algorithm introduced residual learning to improve the depth and performance of the network (Kim et al., 2016a). Dong et al. (Dong et al., 2016) improved SRCNN and reduced the computational complexity by taking the original low-resolution image as the input and adopting a deconvolution layer at the end to complete the image reconstruction of different sampling factors. Problems of gradient disappearance and network degradation become more obvious as the network becomes deeper (Kim et al., 2016b). Recently several algorithms were proposed using a symmetric residual convolution neural network (CNN) for image super-resolution reconstruction (Liu et al., 2019; Wang et al., 2022; Liu et al., 2021). Wang et al. (Wang et al., 2022) applied methods of symmetric short- and long-hop connection, respectively, within and outside the block were adopted to mitigate gradient disappearance and network degradation caused by the depth of the network. Wang et al. tried evaluate deep-learning-based infrared-visible images fusion method based on encoder-decoder architecture. Wang et al. proposed a degradation and super-resolution attention model (D-SRA) using unsupervised machine learning to super-resolution reconstruct high resolution (HR) time-resolved fluid images from coarse data. The deep network is thought to have a good effect in extracting image features, which can be improved by increasing the network depth. However, some critical problems still exist in the extraction of deep features from remote sensing images:

1. Remote sensing images have richer features than ordinary images, so it is impossible to extract deep detail features only by improving the network depth.
2. Only increasing network depth may lead to gradient disappearance and network degradation during network training, resulting in the decline of network extraction ability.
3. The number of parameters also increases, requiring increased network computation and training difficulty, longer reconstruction time images, and lower reconstruction efficiency.

Considering these deficiencies, a CNN model based on symmetric local fusion blocks is proposed in this paper. The combination of a skip connection and residual structure achieves fully integrated information in the block, improving the extraction capability of high-frequency information of remote sensing images and accelerating convergence. A symmetric residual connection mode promotes the effective integration of global information. The original low-resolution remote sensing image is used as the input of the model, reducing computational complexity and improving image reconstruction efficiency.

3. FEATURE FUSION STRUCTURE

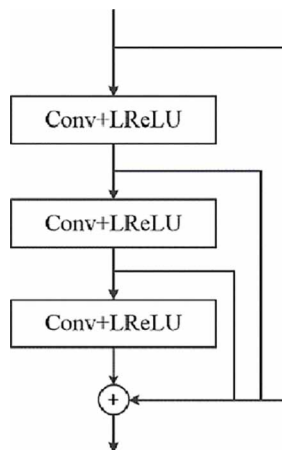
A residual-dense network (RDN) based on residual-dense blocks is proposed to solve the problem of insufficient feature extraction of a deep CNN(Liu et al., 2019). To make full use of the image features extracted from each residual dense block, the RDN adopts a feature fusion structure among the blocks to effectively fuse and adaptively learn the global features extracted from the network. This algorithm demonstrates that the feature fusion structure can effectively improve the ability of network feature extraction and greatly reduce the number of network parameters. Another method for image super-resolution reconstruction (HDRN)(Jiang et al., 2020), also fully demonstrates that the local fusion method is not only beneficial for the reconstruction of ordinary images, but also applicable to the reconstruction of remote sensing images. In view of the advantages of a feature fusion structure for deep networks, we use feature fusion and a local fusion structure to effectively integrate and utilize internal network features. The local fusion structure is shown in Figure 1.

The output of each convolutional layer is directly connected to the end of the last layer to realize the effective integration of feature information extracted from each layer; as the input of the next convolutional layer, it promotes the fluidity of shallow and deep characteristic information. With the increase of network depth, the local fusion structure among convolutional layers can quickly complete the update of shallow network parameters, thus mitigating the phenomenon of gradient disappearance caused by network depth.

4. SUPER-RESOLUTION RECONSTRUCTION ALGORITHM OF REMOTE SENSING IMAGE BASED ON SYMMETRIC LOCAL FUSION BLOCK

To improve the reconstruction quality of remote sensing images, make full use of spatial information, and effectively reduce the computational burden during network training, a super-resolution

Figure 1.
Local Fusion Structure



reconstruction algorithm (SR-SLFB) for remote sensing images based on a symmetric local fusion block is proposed. The network model consists of five local fusion blocks and 38 convolutional layers, primarily for feature extraction, feature fusion, and image reconstruction. The network structure of the algorithm is shown in Figure 2, where x and y are the input and output, respectively, of the network.

The SR-SLFB network has internal fusion connections and external symmetric connections. The information transmission of each convolutional layer is realized by a skip connection. These multiple local fusion blocks are formed to fully utilize the features in the block. Outside the network, symmetric connections are established between the front and back local fusion blocks to realize information sharing. This symmetric connection shares features extracted from the front end to the back end, effectively reducing information loss caused by gradient disappearance.

4.1. Shallow Feature Extraction Module

Given the low resolution and the large amount of information contained in remote sensing images, we take the original low-resolution remote-sensing images as the input of the network to reduce the amount of computation and learning. Feature extraction adopts two convolutional layers with a 3×3 convolution kernel to extract features from the original low-resolution input image x . The extracted shallow feature can be expressed as

$$B_0 = f(x), \quad (1)$$

where f is a feature-extraction function.

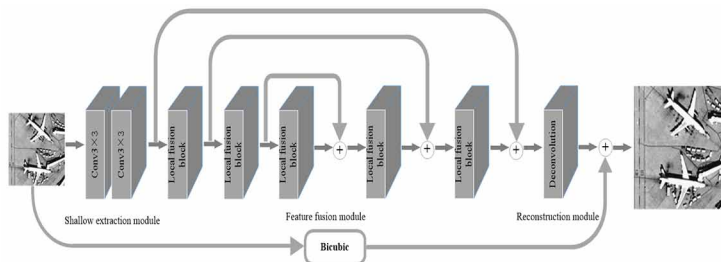
The network uses two convolutional layers to extract shallow features in the low-resolution space, which act as input data for later local fusion blocks. The activation functions in the network adopt leaky rectified linear unit (LReLU) functions, and the negative slope is set to 0.05, for stability and fast convergence.

4.2. Feature Fusion Module

4.2.1. Global Symmetric Connection

Feature fusion is the core of the network. During image reconstruction, gradient disappearance and network degradation become more obvious with increasing network depth. At the same time, super-resolution reconstruction relies heavily on the feature information of low-resolution images. To improve feature extraction without increasing computational complexity, feature fusion is divided into five local fusion blocks with identical structures and global symmetry among blocks, so that shallow and deep feature information are mapped uniformly and feature sharing is realized. The output of the m th local fusion block can be expressed as

Figure 2.
Framework of CNN based on locally fused symmetric blocks



$$\mathbf{B}_m = F_m \mathbf{B}_{m-1} \quad m = 1, 2, \dots, n, \quad (2)$$

where F_m is the function corresponding to the m th local fusion block, and \mathbf{B}_{m-1} is its input.

4.2.2. Local Fusion Block

With the local fusion block, image feature enhancement is realized through convolutional layers with 6-layer convolution kernels of size 3×3 , and feature enhancement uses a residual jump connection to fully integrate extracted image features to fully utilize features of different levels. A convolutional layer with a 1×1 kernel is used to complete image feature compression, and key features are extracted through the convolutional layer with one layer of feature compression, which greatly reduces the computational complexity of network training. The local fusion block is shown in Figure 3.

There are 64 convolution kernels in each convolutional layer, with LReLU activation. The output of the current local fusion block is

$$\mathbf{B}_{m-1} = D\left(\mathbf{B}_{m-2} + \left(\mathbf{M}_0^{m-1}, \mathbf{M}_1^{m-1}, \mathbf{M}_2^{m-1}, \dots, \mathbf{M}_N^{m-1}\right)\right), \quad (3)$$

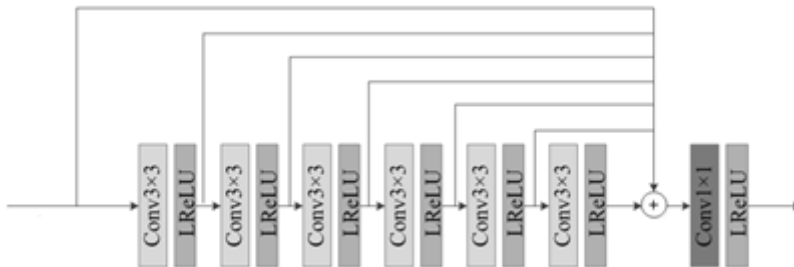
$$\mathbf{M}_i^{m-1} = C\left(\mathbf{M}_{i-1}^{m-1}\right), \quad (4)$$

where \mathbf{B}_{m-1} is both the output of the current local fusion block and the input of the next local fusion block; \mathbf{B}_{m-2} is the input of the current local fusion block; \mathbf{M}_0^{m-1} is the output of the convolutional layer of the first layer in the $(m-1)$ th local fusion block; D is a dimension reduction operation; \mathbf{M}_i^{m-1} ($i = 1, 2, \dots, 5$) is the output of the i th convolutional layer in the $(m-1)$ th local fusion block; and C is a convolution operation.

4.3. Image Reconstruction Module

The traditional reconstruction algorithm based on a CNN preprocesses the original low-resolution image by interpolation before input to the network, which increases computational complexity and destroys image context information, leading to unsatisfactory reconstruction. Due to the great similarity between the original low-resolution image and the high-resolution image, we take the former as the network input, image upsampling is completed through a transposed convolutional layer without an activation function, and the effective features output from the final local fusion block are combined to generate residual images for high-resolution image reconstruction, which further improves image reconstruction quality. The result is

Figure 3.
Local Fusion Block



$$\mathbf{y} = R\left(F_n\left(\mathbf{B}_{n-1}\right)\right) + U(\mathbf{x}), \quad (5)$$

where R is a reconstruction operation and U is a bi-cubic interpolation operation.

5. EXPERIMENTAL ANALYSIS

To verify the effectiveness of the proposed algorithm, ablation experiments were carried out on the NWPU-RESISC45 and UC Merced remote sensing datasets. The proposed algorithm was compared to bicubic, SRCNN (Dong et al., 2014), FSRCNN (Dong et al., 2016), and SymRCN (Liu et al., 2019). The peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were adopted to measure the quality of image reconstruction. These are defined as

$$PNSR = 10 \cdot \log \frac{M \times N}{Y - Y^{*2}}, \quad (6)$$

$$SSIM = \frac{(2\mu_Y \mu_{Y^*} + C_1)(2\sigma_{YY^*} + C_2)}{(\mu_Y^2 + \mu_{Y^*}^2 + C_1)(\sigma_Y^2 + \sigma_{Y^*}^2 + C_2)}, \quad (7)$$

where Y and Y^* are high-resolution images and reconstructed images, respectively; $M \times N$ are the size of the image; μ_Y and μ_{Y^*} respectively represent the average gray values of the original LR and reconstructed image; σ_Y and σ_{Y^*} respectively represent the variance of the original and reconstructed images; σ_{YY^*} is the covariance of the original and reconstructed images; C_1 and C_2 are constants.

In order to avoid the situation where the denominator is 0, $C_1 = (k1 * L)^2$, $C_2 = (K2 * L)^2$, generally $K1 = 0.01$, $K2 = 0.03$, $L = 255$. The values of constants C_1 and C_2 are 6.5025 and 58.5225 respectively. The larger the value of PSNR and SSIM, the better the effect of the reconstructed image, and the closer it is to the original image.

The hardware environment was a computer with an Intel Xeon CPU E5-1650 v4 @ 3.6 GHz \times 12 processor, Tesla K20c GPU, and 64 GB memory. The software environment was a Linux operating system, MATLAB R2016a, Caffe deep learning framework, and CUDA Toolkit 8.0 development package.

In the experiment, the slope of the linear rectification function with leakage was set to 0.05. To initialize the network, the method proposed in document (Jichang et al., 2019) was used for weight initialization, and the offset was set to zero. Adam was used for network optimization. The initial learning rate was set to 10^{-4} , and the training ended at 600,000 iterations.

5.1 Dataset Setup

The UC Merced (He et al., 2015) and NWPU-RESISC45 (Yang et al., 2010) datasets were used as training sets for the model. To demonstrate the effectiveness of the proposed method, a total of 600 images were randomly selected from the two datasets, and rotated 90, 180, and 270 degrees, turned horizontally, and reduced by factors of 0.9, 0.8, 0.7, and 0.6, for a total of 24000 images for training data. Similarly, a total of 51 images were randomly selected as test data.

Bicubic interpolation was used to downsample original high-resolution images with a factor m ($m=2,3,4$), which generates the corresponding low-resolution image, then crop the obtained low-resolution image into sub-images of $h \times h$ size. Original high-resolution images were similarly

cropped into $mh \times mh$ sub-images. Since the over-cropped sub-image cannot obtain enough image information when the sampling factor becomes large. To make full use of the information in images, low- and high-resolution images were cropped to $29^2 / 57^2$, $15^2 / 43^2$, and $11^2 / 41^2$, respectively, according to m ($m=2, 3, 4$) times the sampling size.

5.2 Analysis of SR-SLFB Network Structure

The proposed local fusion structure, as the core of the SR-SLFB network, can adaptively aggregate effective image features and improve the expression ability of the deep network. The depth of the network is improved by setting up several local fusion blocks, which can extract the deep features of remote sensing images and improve network feature extraction. To analyze the network structure of the SR-SLFB algorithm and show that the local fusion method can effectively reduce the gradient disappearance caused by the deep network, the average PSNR value of the UC Merced dataset was compared to that of two-fold sampling factor reconstruction using in-block partial fusion and non-partial fusion. As can be seen from Figure 4, after local fusion, the average PSNR of the network during training was significantly improved compared to that without local fusion. Moreover, with the increase of training iteration times, network convergence was faster and reconstruction performance was better after using the local fusion method.

Table 1 shows the average runtime with and without local fusion. Not using local fusion will take longer to run, while using local fusion will significantly improve the runtime. In the test procedure, using local fusion is 2.3ms faster than do not use local fusion.

5.3 Comparison of Proposed Algorithm to Other Algorithms

To demonstrate the effectiveness of SR-SLFB, downsampling operations of 2-, 3-, and 4-fold factors for the UC Merced and NWPU-RESISC45 remote sensing datasets were performed, and SR-SLFB algorithm was compared to the bicubic, SRCNN, FSRCNN, and Sym-RCN algorithms. For more convincing comparisons, the number of iterations of algorithms was set to 600,000, and training was performed again.

Tables 2 and 3 show the comparative results of average PSNR and SSIM, respectively, on the UC Merced and NWPU-RESISC45 datasets for different reconstruction algorithms with sampling factors

Figure 4.
Quantitative evaluation (PSNR) on UC Merced dataset

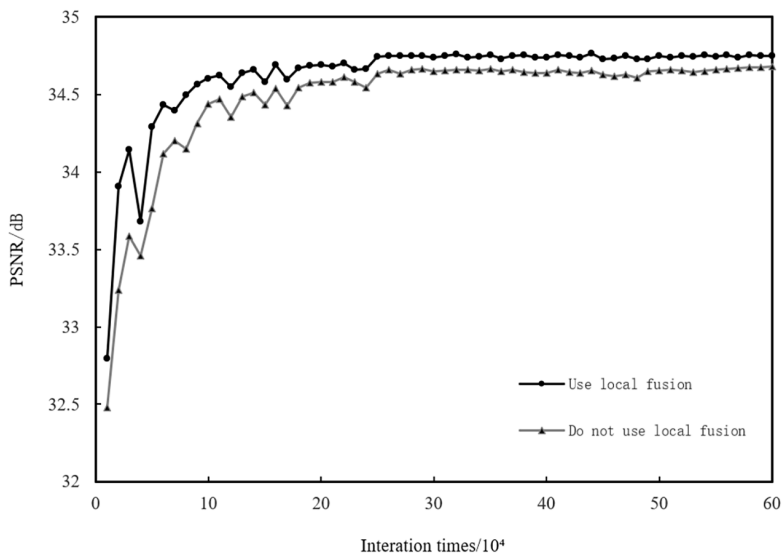


Table 1.
The average runtime with and without local fusion

Method	SR-SLFB
Not using local fusion	30.8ms
Using local fusion	28.5ms

of 2, 3, and 4. SR-SLFB performed best, with the highest average PSNR and SSIM values. For the UC Merced dataset, compared to the other four algorithms, the PSNR and SSIM values of SR-SLFB increased by 3.10db and 0.0797, 1.67db and 0.0378, 1.12db and 0.0210, and 0.22db and 0.0046, respectively. Due to the low resolution of the original NWPU-RESISC45 dataset, its reconstruction performance was slightly lower than that of the UC Merced dataset. The PSNR and SSIM values of the SR-SLFB algorithm were 2.99db and 0.0647, 1.57db and 0.0341, 1.08db and 0.0195, and 0.16db and 0.0021, respectively, higher than those of the other four algorithms.

There are three reasons for the high-image quality reconstructed by SR-SLFB and its high PSNR and SSIM values.

1. The algorithm's local fusion structure achieves a high degree of combination of features at different levels in images and enhances the flow of information in the network. The algorithm effectively reduces the phenomenon of network gradient disappearance by setting skip connections among convolutional layers in the local fusion block.

Table 2.
Average PSNR for sampling factors 2, 3, and 4 on UC Merced and NWPU-RESISC45 datasets

Dataset	Sampling factor	Bicubic (PSNR)	SRCNN (PSNR)	FSRCNN (PSNR)	SymRCN (PSNR)	SR-SLFB (PSNR)
UC Merced	×2	30.81	32.41	33.04	34.55	34.77
	×3	27.76	29.25	29.69	30.47	30.74
	×4	26.02	27.20	27.78	28.19	28.36
NWPU-RESISC45	×2	30.78	32.37	32.90	34.48	34.50
	×3	27.62	29.16	29.57	30.31	30.48
	×4	26.15	27.28	27.82	28.26	28.54

Table 3.
Average SSIM for sampling factors 2, 3, and 4 on UC Merced and NWPU-RESISC45 datasets

Dataset	Sampling factor	Bicubic (SSIM)	SRCNN (SSIM)	FSRCNN (SSIM)	SymRCN (SSIM)	SR-SLFB (SSIM)
UC Merced	×2	0.8633	0.8981	0.9021	0.9223	0.9290
	×3	0.7512	0.8083	0.8245	0.8392	0.8435
	×4	0.6627	0.6953	0.7253	0.7397	0.7423
NWPU-RESISC45	×2	0.8630	0.8934	0.8967	0.9102	0.9138
	×3	0.7674	0.7992	0.8256	0.8362	0.8379
	×4	0.6993	0.7287	0.7428	0.7709	0.7719

2. The symmetric connection mode realizes the symmetric sharing of the global features extracted from the network, and establishes an information-sharing bridge between the local fusion blocks, realizes the effective complementation of image features, and further improves reconstruction performance.
3. The algorithm increases network depth by adding local fusion blocks, enhances the network's ability to extract deep-level features of the image, makes full use of high-frequency information in the image context, and enhances network reconstruction ability.

Table 4 compares the average test time of different reconstruction algorithms on the UC Merced and NWPU-RESISC45 datasets when the sampling factor is 2, 3, and 4.

It can be seen from Table 4 that the average test time of SR-SLFB was less than those of the other algorithms, and the average time of SRCNN was the longest. When the sampling factor was 2, the average test time of SR-SLFB for image reconstruction between the two datasets was 0.031 s, and those of FSRCNN and SymRCN were 0.877s and 0.045 s, respectively. SR-SLFB was 0.846 s faster than FSRCNN and 0.014 s faster than SymRCN. When the sampling factor was 4, the average test time of SR-SLFB for image reconstruction between the two datasets was 0.0295s, and the average test times of FSRCNN and SymRCN were 0.7925s and 0.0315s, respectively. SR-SLFB was 0.763s faster than FSRCNN and 0.002s faster than SymRCN. Through the analysis, it can be seen that SR-SLFB took less time to reconstruct the image, and was more efficient at image reconstruction.

SR-SLFB was relatively fast at reconstructing images because it adopts the residual structure and applies it to the local fusion block, which effectively reduces the number of network parameters, alleviates the complexity of network training, and increases the speed of network convergence and reconstruction. SR-SLFB sets a small convolution kernel of 3×3 , which effectively reduces the number of network parameters, increases learning ability, and further improves network reconstruction speed.

Figures 5 to 10 show the effects of the four algorithms after the reconstruction of 2-, 3-, and 4-fold sampling factors on the UC Merced and NWPU-RESisc45 remote sensing datasets respectively. It is shown that the texture of remote sensing images reconstructed by SR-SLFB for different sampling factors was clearer and more detailed, and the sawtooth and artifacts were reduced. Because the proposed symmetric local fusion block effectively extracts and fuses image features, the reconstructed remote sensing image has a clearer edge, is closer to the original image, and has a better visual effect.

6. CONCLUSION

A method of super-resolution reconstruction of remote sensing images based on a symmetric local fusion block was proposed in view of characteristics such as difficulty in super-resolution reconstruction, strong autocorrelation, and rich texture details. The network model consisted of 38 convolutional layers. The core part of the network model consisted of five local fusion blocks, each

Table 4.
Average test time for sampling factors 2, 3, and 4 on UC Merced and NWPU-RESISC45 datasets

Dataset	Scale	Bicubic	SRCNN	FSRCNN	SymRCN	SR-SLFB
UC Merced	$\times 2$	-	5.162	0.812	0.055	0.036
	$\times 3$	-	5.075	0.745	0.028	0.024
	$\times 4$	-	5.137	0.729	0.022	0.020
NWPU-RESISC45	$\times 2$	-	4.252	0.942	0.035	0.026
	$\times 3$	-	4.136	0.878	0.043	0.037
	$\times 4$	-	3.893	0.856	0.041	0.039

Figure 5.
Results of super-resolution on agricultural15 with 2-fold sampling factors

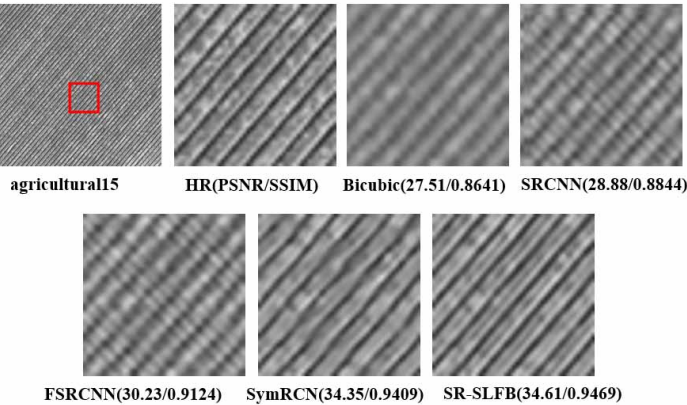


Figure 6.
Results of super-resolution on runway_095 with 2-fold sampling factors

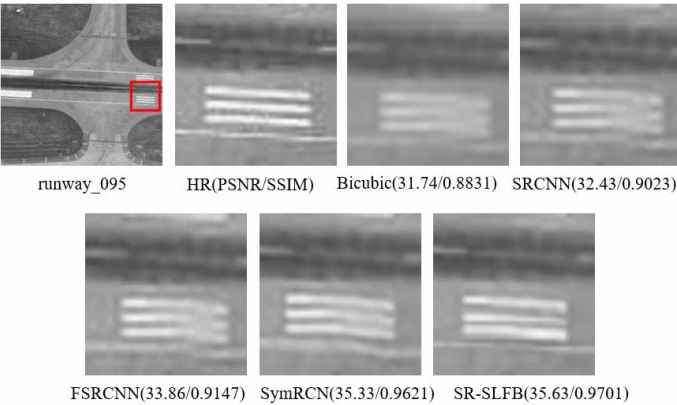


Figure 7.
Results of super-resolution on airplane 40 with 3-fold sampling factors

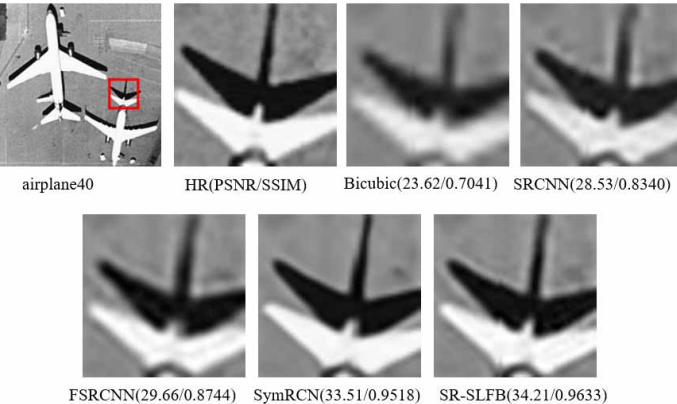


Figure 8.
Results of super-resolution on roundabout_169 with 3-fold sampling factors

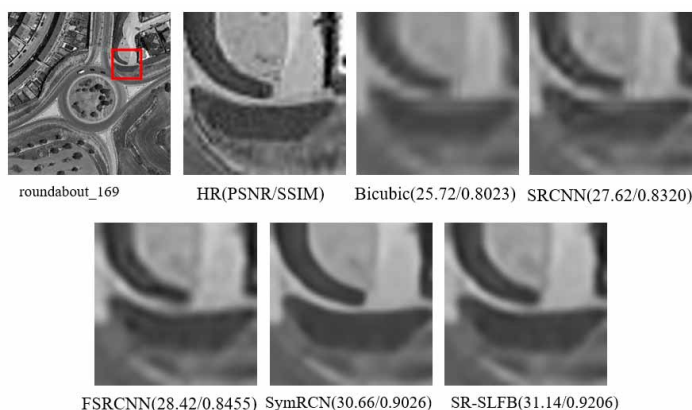


Figure 9.
Result of super-resolution on baseball diamond03 with 4-fold sampling factors

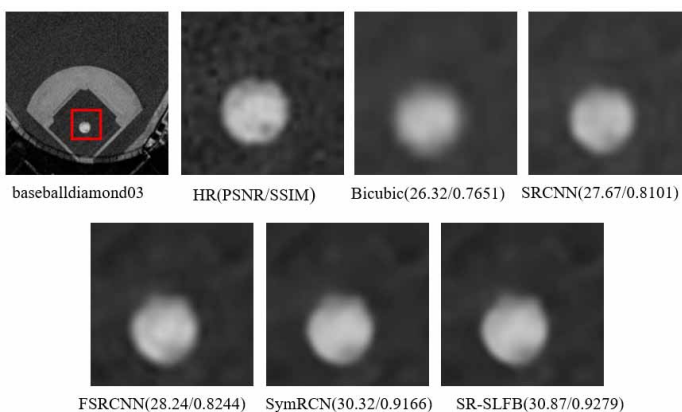
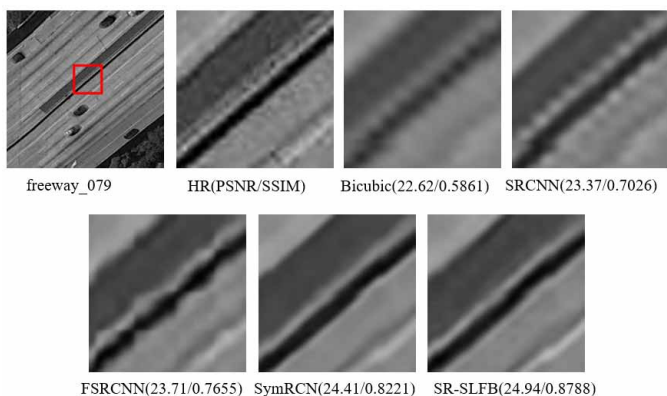


Figure 10.
Results of super-resolution on freeway_079 with 4-fold sampling factors



containing six convolutional layers with the same structure and one compression unit. Through the connection mode of symmetric local fusion, the effective integration of multi-level features was realized, which improved the convergence speed of the network and effectively diminished the phenomena of gradient disappearance and network degradation. In the next step, we will consider training image super-resolution reconstruction models suitable for different scenes by learning different degradation models.

DATA AVAILABILITY

The data used to support the findings of this study are available from the corresponding author upon request.

CONFLICTS OF INTEREST

Regarding the publication of this article, the authors declare that they have no conflicts of interest.

FUNDING STATEMENT

This work was supported by the National Natural Science Foundation of China, grant number 61876131.

REFERENCES

- Bätz, M., Eichenseer, A., Seiler, J., Jonscher, M., & Kaup, A. (2015). *Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches*. In *2015 IEEE international conference on image processing. ICIP*.
- Cui, Z., Dang, S., Cao, Z., Wang, S., & Liu, N. (2018). SAR target recognition in large scene images via region-based convolutional neural networks. *Remote Sensing*, *10*(5), 776. doi:10.3390/rs10050776
- Ding, J., Chen, B., Liu, H., & Huang, M. (2016). Convolutional Neural Network With Data Augmentation for SAR Target Recognition. *IEEE Geoscience and Remote Sensing Letters*, 364–368. doi:10.1109/LGRS.2015.2513754
- Dong, C., Loy, C. C., He, K., & Tang, X. (2014). *Learning a deep convolutional network for image super-resolution*. Springer International Publishing.
- Dong, C., Loy, C. C., & Tang, X. (2016). *Accelerating the super-resolution convolutional neural network*. Springer International Publishing.
- Dong, W., Lei, Z., & Shi, G. (2013). Nonlocally Centralized Sparse Representation for Image Restoration. *IEEE Transactions on Image Processing*, *22*(4).
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*, 1026-1034.
- Jiang, K., Wang, Z., Yi, P., & Jiang, J. (2020). Hierarchical Dense Recursive Network for Image Super-Resolution. *Pattern Recognition*, *107*, 107475.
- Lai, W. S., Huang, J. B., Ahuja, N., & Yang, M. H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 624-632.
- Liu, S. D., Wang, X. M., & Zhang, Y. (2019). An image super-resolution reconstruction method of symmetrical residual CNN. *Journal of Xi'an University of Electronic Science and Technology*, *46*(05), 15–2.
- Liu, Z., Huang, J., Zhu, C., Peng, X., & Du, X. (2021). Residual attention network using multi-channel dense connections for image super-resolution. *Applied Intelligence*, *51*, 85–99.
- Michael, K., & Arnt-Børre, S. (2018). Urban Land Cover Classification With Missing Data Modalities Using Deep Convolutional Neural Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(6), 1758–1768. doi:10.1109/JSTARS.2018.2834961
- Papayan, V., & Elad, M. (2015). Multi-Scale Patch-Based Image Restoration. *IEEE Transactions on Image Processing*, *25*(1), 249-261.
- Shamsolmoali, P., Zareapoor, M., Jain, D. K., Jain, V. K., & Yang, J. (2019). Deep convolution network for surveillance records super-resolution. *Multimedia Tools and Applications*, *78*(17), 23815–23829. doi:10.1007/s11042-018-5915-7
- Wang, B., Zou, Y., Zhang, L., Li, Y., Chen, Q., & Zuo, C. (2022). Multimodal super-resolution reconstruction of infrared and visible images via deep learning. *Optics and Lasers in Engineering*, *156*, 107078.
- Wang, Y., Li, X., Nan, F., Liu, F., Li, H., Wang, H., & Qian, Y. (2022). Image super-resolution reconstruction based on generative adversarial network model with feedback and attention mechanisms. *Multimedia Tools and Applications*, *81*(5), 6633–6652.
- Wang, Z., Li, X., Liu, L., Wu, X., Hao, P., Zhang, X., & He, F. (2022). Deep-learning-based super-resolution reconstruction of high-speed imaging in fluids. *Physics of Fluids*, *34*(3), 037107.
- Wei, D. (2016). Image super-resolution reconstruction using the high-order derivative interpolation associated with fractional filter functions. *IET Signal Processing*, *10*(9), 1052–1061.
- Wu, H., Zhang, H., Zhang, J., & Xu, F. (2015, September). Fast aircraft detection in satellite images based on convolutional neural networks. *2015 IEEE International Conference on Image Processing (ICIP)*, 4210-4214. doi:10.1109/ICIP.2015.7351599

Xu, H., Zhai, G., & Yang, X. (2013). Single Image Super-resolution With Detail Enhancement Based on Local Fractal Analysis of Gradient. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(10), 1740–1754.

Yang, J. W. J., Huang, T., & Ma, Y. (2010). Image Super-Resolution Via Sparse Representation. *IEEE Transactions on Image Processing*, 19(11), 2861–2873.

Yang, Y., & Newsam, S. (2010). Bag-of-visual-words and spatial extensions for land-use classification. *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, 270-279.

Zhang, K., Zuo, W., Gu, S., & Zhang, L. (2017). Learning deep CNN denoiser prior for image restoration. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3929-3938.

Xinqiang Wang received the M.S. degree from Nankai University, Tianjin, China, in 2009. He is currently pursuing the Ph.D. degree with the School of Software, Tianjin University, Tianjin, China. His research interests include intelligent human computer interaction system and visualization.

Wenhuan Lu is a professor at College of Intelligence and Computing in Tianjin University, P. R. China. She received her Ph.D. Degree at Japan Advanced Institute of Science and Technology (JAIST), Japan and she has also worked as postdoctoral researcher at JAIST; afterwards, she joined in Tianjin University, P. R. China. She is working on intelligent human computer interaction system and visualization. She is currently the Principal Investigator for several national grants in China, including National High-Technology R&D Program of China, National Natural Science Foundation of China, and Scientific Research Foundation from Ministry of Education of China, etc. She has also (co) authored more than 60 academic articles.