An Improved U-Net Model for Astronomical Images Denoising

1st Jin Qi
School of Internet of Things
Nanjing University of Posts and Telecommunications
NanJing, China
qijin@njupt.edu.cn

2nd Mengwei Chen
School of Internet of Things
Nanjing University of Posts and Telecommunications
NanJing, China
1221077031@njupt.edu.cn

3rd Zhaofei Wu
School of Internet of Things
Nanjing University of Posts and Telecommunications
NanJing, China
2829342079@qq.com

4th Can Su
School of Internet of Things
Nanjing University of Posts and Telecommunications
NanJing, China
335212080@qq.com

5th Xiaoliang Yao
School of Computer Science
Nanjing University of Posts and Telecommunications
NanJing, China
1392803263@qq.com

6th Chaojun Mei
School of Computer Science
Nanjing University of Posts and Telecommunications
NanJing, China
1221045717@njupt.edu.cn

2022 China Automation Congress (CAC) | 978-1-6654-6533-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/CAC57257.2022.10054887

***Abstract*—To explore and understand the universe, it is inevitable to analyze astronomical images. But additional and incidental noise information will be added during the acquisition process. These noise will interfere with the subsequent image analysis. Hence it is necessary to denoise the astronomical images before analyzing. Astronomical image resolution is large and the star signal is small. Meanwhile,the existing denoising methods are difficult to recover the feature information of stars. To enhance the ability of detail feature extraction while denoising, we propose an improved U-Net model. First, we add the residual connections inside the network to break the symmetry of the network. Residual connections are useful to information transmission, and avoid gradient explosion and overfitting. Meanwhile, residual connections can improve the network representation ability. Secondly, we use the encoding and decoding parts of Receptive Field Block(RFB) connection model. RFB increases the receptive field to strengthen the capability of multi-scale fea-ture extraction. RFB also recover more feature information of stars. Finally, we used an improved upsampling method to avoid generating checkerboard pattern of artifacts. We use the dataset from the Hubble Space Telescope(HST) Archives to verify the experiment. The experimental results verify that the proposed model performs well in the task of astronomical image denoising. Our method can better remove the existing noise of astronomical images and recover the feature in-formation of stars.**

Keywords—Astronomical image denoise, U-Net, Residual connection, Multi-scale feature extraction

I. INTRODUCTION

The research of astronomical images are very important for deep space exploration. But it is inevitable that astronomical images will be disturbed by various noises in the process of shoot and transmission. In daily photography, the optical signal received by the camera is much larger than the noise. Therefore we can ignore these noise. But in the Astronomical photography, the target is very faint object in the universe. The optical signal may be as strong as the noise signal. Astronomical images contain a lot of detailed information about the universe. The existence of noise makes it difficult to analyze the images. We can improving instruments and methods of shoot to reduce noise. But it is almost impossible to eliminate noise completely. Hence it is necessary to denoise the astronomical images before analyzing. Worden et al. first proposed to use the designed threshold to denoise astronomical images. Applying linear or nonlinear filters can achieve further image enhancement to improve the noise level. Zhu et al. proposed the use of median filter for image denoising. Early image denoising work includes sparse coding, Block Matching 3D Filter (BM3D) [1] etc. All methods rely he-avily on image prior. Traditional methods can perform image denoising better. But these algorithms have high complexity and limited application scope. Image denoising task has made a new breakthrough in the era of deep learning. Initially, Multilayer Perception (MLP) was applied to the denoising task. The denoising effect was comparable to that of the traditional method BM3D. Later, more network structures were proposed for denoising tasks, including blind and non-blind denoising methods. Chunhua Shen et al. proposed a complete convolutional codecs framework RED-Net [2] for blind denoising. The framework uses jump connection to acce-lerate convergence and strengthen the recovery of details. Kai Zhang et al. Proposed blind Denoising Convolutional Neural Network DnCNN [3].This network makes residual learning and Batch Normal-ization (BN) be valued in the field of denoising. FFD-Net [4], as a non-blind denoising method,establishes the relationship between noise level and denoising model in the form of graph. Subsequently, ShiGuo et al. proposed a non-blind denoising network CBD-Net [5]. The network embed a noise estimation sub-network with asymmetric learning. The sub-network can suppress the underestimation of noise level. In recent years, models from high-level visual tasks such as image segmentation and object detection have been applied to denoising tasks. Olaf Ronneberger et al. [6] proposed to encode and decode U-Net network for medical image segmentation. Subsequently, it is also used in the field of image denoising. MIR-Net [7] proposed a general network architecture. The network can extract multi-scale feature information for image enhancement tasks such as denoising and super resolution . These denoising models based on convolutional neural networks have better performance than traditional denoising models. NB-Net [8] proposed non-local attention module, learning base generation and subspace projection to maintain the local structure of the input signal naturally. NAF-Net [9] further simplified the baseline by deleting or replacing nonlinear activation functions. The network proposed a network of no more than linear activation.

Above traditional denoising method and deep learning denoising algorithm can achieve the denoising task and recover the original image with a relatively clear. Due to the large astronomical images resolution and the characteristics of few star signals. These methods can not recover the feature information of stars well while denoising. To enhance the ability of detail feature extraction while denoising, we propose an improved U-Net model. The experimental results verify that the proposed model performs well in the task of astronomical image denoising. The network can better remove the existing noise of astronomical images and recover the feature information of stars.

In summary, our contributions are as follows:

•To enhance the reusability of features and information, we add residual connections within the network. The connections break the symmetry of the inherent network to improve the representation capability of the network. It also avoid the problems of gradient explosion and gradient disappearance.

• we integrated the RFB [10], used the encoding and decoding part to increase the receptive field. RFB can recover more star feature information by strengthening multi-scale feature extraction ability.

•To avoid checkerboard patterns of artifacts [11] caused by the uneven overlap feature of transpose convolution, we use an improved upsampling method.

II. Our Method

A. The original U-Net model

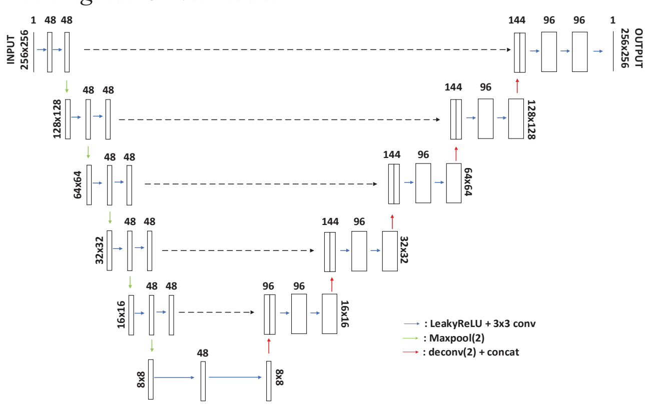


Fig. 1. Original U-Net network structure.

The model proposed is an improved model based on U-Net network structure. U-Net network structure is shown in Figure 1. The network structure is symmetric. The encoder and decoder correspond to the left and right parts of the picture respectively. The encoder part is a feature extraction network consisting of a 3*3 convolution, a down-sampling operation and a LeakyReLU activation function. The encoder part uses the max-pooling operation for down-sampling and reduces feature maps resolution. The decoder part is a feature fusion network consisting of 3*3 convolution, an upsampling operation, a LeakyReLU activation function and feature splicing operation. The decoder part uses the transposed convolution operation for upsampling to recover feature images resolution . It also uses the concatenate operation for upsampling to concatenate the encoding images.

B. The improved U-Net model

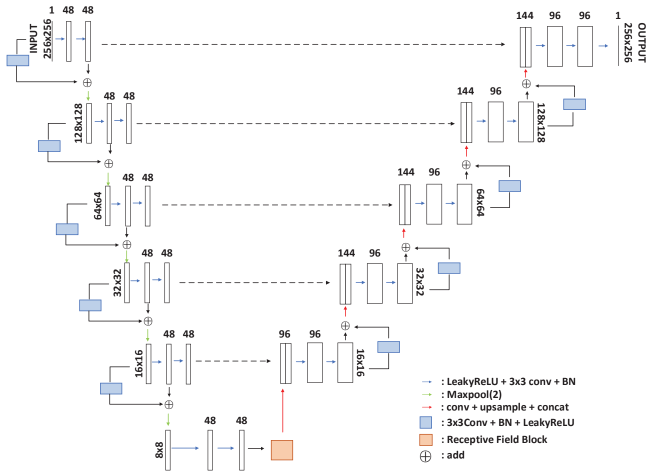


Fig. 2. Improved U-Net network structure.

a. Deep residual connection

The increase of network depth can improve the performance. The network will be difficult to train and degradation while increasing the depth of network. Therefore, training deeper networks is more challenging. He et al. [12] proposed Residual Module, which solved the problem of degradation and promoted the training process. ResUNet, a neural network proposed in medical image segmentation. The network combines the idea of ResNet and uses residual units to promote the better completion of image segment-ation tasks. Residual connection is used to break the symmetry of the network because it is difficult to recover the detailed feature information of astronomical image de-noising. Residual connection is useful for information trans-mission and improves the representation ability of the network. With the deepening of network layers, the use of residual connection can also avoid the emergence of gradient explosion and gradient disappearance. Similarly, network degradation will be avoided and convergence speed will be accelerated. We can add residual connections in U-Net network to improve network performance according to these advantages.

b. Receptive Field Block

To simulate the receptive field of human vision and reinforce the feature extraction ability of the network, Liu et al. proposed RFB [10]. RFB borrows ideas from Inception [13] networks in structure, mainly adding dilated convolution on the basis of inception to increase the receptive field. In the

RFB structure, the rates of the dilated convolution layer are different. Finally, concat-enate feature fusion is performed on convolution layer outputs of different rates and sizes. The RFB structure increases the receptive field and reinforce the ability of multi-scale feature information extraction. These advantages are expected to enhance the capability of feature extraction in astronomical denoising tasks. Therefore, we add RFB to the bottom of U-Net as the connection between encoding part and decoding part. RFB is used to capture multi-scale information and recover more star features while denoising.

c. Upsampling

The original U-Net model uses transpose convolution to perform upsampling in the decoding part. It should be noted that the transpose convolution will cause uneven overlap if the nuclear size can not be divided by the stride. In the two-dimensional case, the uneven overlaps of different axes are superimposed to produce checkerboard pattern of artifacts [11] with different intensities. Nowadays, multi-layer trans-pose convolution is widely used to generate images. It is theoretically possible to eliminate the effects of the illusion. However, the superposition of convolution layers after transformation will still produce illusions of different scales. Limiting filters can be used to avoid serious artifacts, which conversely can lead to a reduction in the capacity of model. Considering this situation, we use upsampling plus convo-lution to replace transpose convolution to achieve better performance in the improved model. The detailed procedure is as follows: the image is first upsampled and performs the operation of 5*5 convolution and 3*3 convolution with residual connection.

C. Loss Function

In this paper, the input of the model is the noise observation value $y = x + v$. Gaussian noise is used to simulate the real noise. The L1 loss and MSE loss function are used to learn the potential clean images. The L1 loss function is used for the first 40 training rounds and the MSE loss function for the last 60 rounds.

$$L1 = \frac{1}{N}\sum_{i=1}^{N} |x_i - x_i'| \quad (1)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N} |x_i - x_i'|^2 \quad (2)$$

Where x represents the ground-truth image, and x' represents the output of the network.

III. Experiments

A. Dataset

The Hubble Space Telescope(HST or Hubble) is a telescope that orbits around the Earth and is above the Earth's atmosphere. Since its 1990 launched, HST has become the most important instrument in the history of astronomy. We used the dataset obtained from HST archives. Each image was captured by the UVIS detector in the Wide Field Camera 3(WFC3). WFC3, a detector with 4000*4000 pixels, has s pair of charge-coupled devices (CCDs), each with 2051*4096 pixels and with a 31-pixel gap between them. We selected 174 images containing celestial bodies of various scales, we use 140 images as the training set and 34 as the validation set.

B. Implementation Details

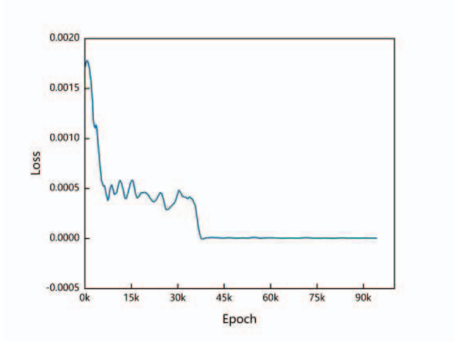
Our model is implemented with PyTorch. The improved U-Net is an end-to-end trainable network that do not need pre-training. We cut the images for training into a size of 256*256 pixels, resulting in totally 26545 images.Consi-dering that astronomical images often contain large areas without signal, we then filtered out such invalid images and finally ended up with a total of 15076 for training, each with 256*256 pixels. Here, we use the Adam optimizer during the experimental optimization process. The batch size is 16, the initial learning rate is 0.0003. Our training iteration executes 100 times, in which the learning rate is halved every 20 times, and the weight is saved as well as the validation is performed every 10 times. All experiments are conducted on servers with Python 3.8.1, PyTorch 1.8.0, and Nvidia RTX 3090 GPU.

We set Gaussian noise level σ as 15, 25, 50 to explore the influences of different degrees of noise on the ex-perimental results.

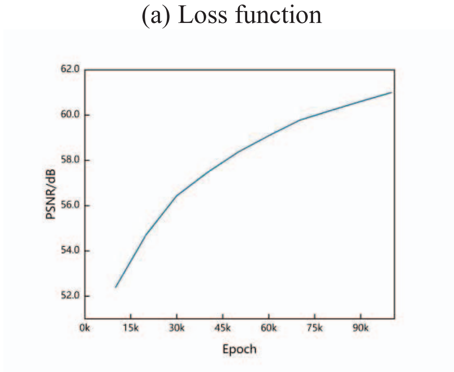
C. Create New Image

After the input of astronomical images, we apply a mosaic approach, whose sliding window size is set to 256*256, to recreate the full image size. In practical, our network runs iteratively until all the images are traversed. In each iteration, it will 1) put the current window into the network and run, and then 2) move 32 pixels into the network again. Average the pixels in the area where the windows overlap.

D. Image Denoising Results



(a) Loss function



(b) PSNR value

Fig. 3. Loss function and PSNR curves under gaussian noise $\sigma = 25$.

Figure 3 shows the loss function curves of the proposed denoising method at noise level 25 and the PSNR curves of the test set. Table 1 shows the average PSNR and SSIM values of our denoising model(in the last column), the classical denoising method BM3D [1] and other deep learning methods on the test data set under different noise levels. It also shows the quantitative results of denoising the test set images beyond different methods. Table 1 shows both the PSNR and SSIM values of the traditional denoising method BM3D [1] at different noise levels are smaller than the several deep learning based denoising models. Therefore, the denoising method based on deep learning is more effective. Motivated by this observation, we propose our improved model. Firstly, in view of the degradation of deep networks, residual connections within networks are added to enhance re-presentation learning ability, and the problems of gradient disappearance and gradient explosion are solved. Then, to recover more star features, we integrate RFB [10] to enhance the network ability of feature representation and multi-scale feature extraction. Finally, checkerboard pattern of artifacts appeared when the kernel size could not be divided by the stride, for which an improved upsampling method is used. Table 1 shows the algorithm proposed in this paper is superior to NAFNet [9], CDLNet [14], SCUNet [15], RIDNet [16] and the traditional denoising method BM3D. The comparison of experimental results is shown in Figure 4 which indicates the superiority of our method.

TABLE 1. QUANTITATIVE COMPARISON OF DENOISING RESULTS OF DIFFERENT ALGORITHMS AT DIFFERENT NOISE LEVELS（PSNR（DB）/SSIM）

Dataset	Noise Level	NAFNet[9]	CDLNet[14]	SCUNet[15]	RIDNet[16]	DnCNN[3]	BM3D[1]	Ours
Astro-Image	15	60.81/0.998	62.48/0.996	**63.17**/**0.998**	63.06/0.997	59.95/0.991	58.12/0.972	**63.17**/0.997
	25	56.64/0.996	60.35/0.996	60.86/0.996	60.90/0.996	55.91/0.978	54.72/0.926	**60.91**/**0.996**
	50	54.28/0.993	56.83/0.994	56.59/0.993	57.01/**0.995**	51.78/0.956	50.63/0.914	**57.26**/0.994

T

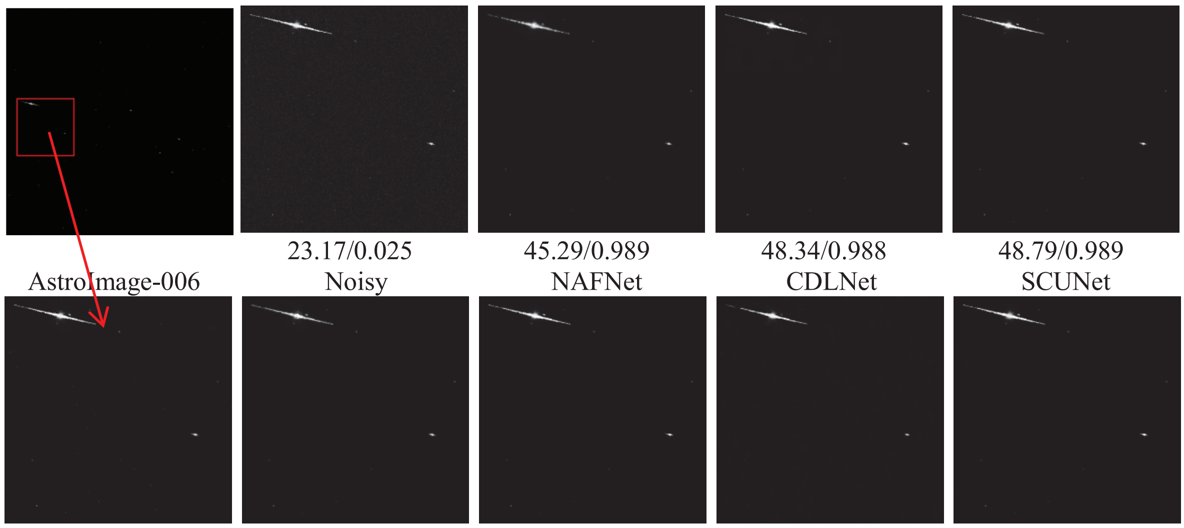


Fig. 4. Visual comparison of our method with other methods on the sixth test set image in the case of Gauss σ = 25 Quantitative PSNR(dB)/SSIM results are listed below the image.

E. Ablation Study

Taking noise level $\sigma = 25$ as an example, three main components of the improved model are examined: a) residual connection, b)RFB module, and c) the effect of the improved upsampling on the model. The network symmetry is forcibly broken by the residual connections within it, which promotes information transmission, improves the representation ability of the network, and avoids gradient explosion and overfitting problems in the meanwhile. The improved upsampling method avoids the checkerboard pattern of artifacts. The encoding and decoding parts of RFB connection model are used to increase receptive fields, hence to enhance multi-scale feature extraction ability and recover more star feature information. On the basis of verifying the effect of the original U-Net model, we successively added three components for verification. The experimental results are shown in Table 2. As seen, all three components optimize the model to varying degrees.

TABLE2. INFLUENCE OF EACH COMPONENT ON THE MODEL

	residual connection	RFB	Improved upsampling	PSNR(dB)/SSIM
U-Net				58.68/0.997
	√			59.53/0.996
	√		√	59.79/0.996
	√	√	√	60.91/0.996

IV. CONCLUSIONS

In this paper, we propose an improved U-Net model for astronomical image denoising. Firstly, motivated by the challenge that traditional deep network degrades with the deepening of network layers, we propose to add residual connections within the network. The residual connections improve representation learning ability and avoid gradient explosion and gradient disappearance. Then, we integrate the encoding and decoding parts in the RFB connection model to simulate the receptive field size in human vision. RFB increases the receptive field to strengthen the ability of multi-scale feature extraction. RFB also recover more feature information of stars. Finally, we use an improved up-sampling method to avoid the artifacts checkerboard pattern produced by transpose convolution.

ACKNOWLEDGMENT

This paper was supported by China Postdoctoral Science Foundation(2019M651923), Natural Science Foundation of Jiangsu Province of China under Grant BK20191381.

REFERENCES

[1] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," in IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2080-2095, Aug. 2007, doi: 10.1109/TIP.2007.901238.

[2] Mao X, Shen C, Yang Y B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections[J]. Advances in neural information processing systems, 2016, 29.

[3] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, July 2017, doi: 10.1109/TIP.2017.2662206.

[4] K. Zhang, W. Zuo and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising," in IEEE Transactions on Image Processing, vol. 27, no. 9, pp. 4608-4622, Sept. 2018, doi: 10.1109/TIP.2018.2839891.

[5] S. Guo, Z. Yan, K. Zhang, W. Zuo and L. Zhang, "Toward Convolutional Blind Denoising of Real Photographs," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 1712-1722, doi: 10.1109/CVPR.2019.00181.

[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, Cham, 2015: 234-241..

[7] S. W. Zamir et al., "Learning Enriched Features for Fast Image Restoration and Enhancement," in IEEE Transactions on Pattern Analysis and Machine Intelligence, doi: 10.1109/TPAMI.2022.3167175.

[8] S. Cheng, Y. Wang, H. Huang, D. Liu, H. Fan and S. Liu, "NBNet: Noise Basis Learning for Image Denoising with Subspace Projection," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 4894-4904, doi: 10.1109/CVPR46437.2021.00486.

[9] Chen L, Chu X, Zhang X, et al. Simple baselines for image restoration[J]. arXiv preprint arXiv:2204.04676, 2022.

[10] Liu S, Huang D. Receptive field block net for accurate and fast object detection[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 385-400.

[11] Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts[J]. Distill, 2016, 1(10): e3.

[12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 770-778.

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.

[14] N. Janjušević, A. Khalilian-Gourtani and Y. Wang, "CDLNet: Noise-Adaptive Convolutional Dictionary Learning Network for Blind Denoising and Demosaicing," in IEEE Open Journal of Signal Processing, vol. 3, pp. 196-211, 2022, doi: 10.1109/OJSP.2022.3172842.

[15] Guo X, O'Neill W C, Vey B, et al. SCU - Net: A deep learning method for segmentation and quantification of breast arterial calcifications on mammograms[J]. Medical physics, 2021, 48(10): 5851-5861.

[16] S. Zhuo, Z. Jin, W. Zou and X. Li, "RIDNet: Recursive Information Distillation Network for Color Image Denoising," 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019, pp. 3896-3903, doi: 10.1109/ICCVW.2019.00483.