Image Enhancement and Improvement Algorithm Based on Esrgan Singal Frame Remote Sensing Image

Zhiwei Yang 1, *, Yunyan Wang 2

1 Key Laboratory of High Efficiency Utilization and Energy Storage Operation Control of Hubei Province, Hubei University of Technology, Wuhan 430068, China
2 School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan430068, China

*Corresponding author: zhiweiyang@hbut.edu.cn

Abstract. Traditional spline interpolation algorithm for reconstruction of visual effects are not good, based on Super Resolution against Network (Super - Resolution Generative Adversarial Network, SRGAN) edge character to deal with such problems as imperfect, using a generated based on the enhanced Super Resolution against Network to improve the Resolution of ordinary optical remote sensing images, the first Network generated by high Resolution data sets to train (G network), then lower Resolution test data model test, The test results and real results are put into the discrimination network (D network) to get the adversarial loss, and then the generated network is modified according to the adversarial loss. The superiority of the network is improved by introducing dense residuals to SRGAN, modifying the judgment object of the discriminator to be relatively real, and using the eigenvalue before activation to improve the perceived loss. The desert, farmland, forest and mountain data were tested on AID data set, and the algorithm in this paper could obtain the recomposition of the real image more closely. Compared with SRResNet and SRGAN algorithms, PSNR improved by about 4.0db and SSIM improved by about 0.14. This method improves the feature comprehensiveness by increasing the network fineness degree, and USES the modified perception loss to get the brightness closer to the real image, which is beneficial to improve the quality of single frame remote sensing image.

Keywords: Remote sensing image Confrontation network Super resolution Reconstruction error.

1. Introduction
Remote sensing images are often used in natural disaster monitoring, environmental monitoring, ground feature analysis, resource exploration and other important fields. Therefore, obtaining remote sensing data with higher resolution has become an important link of image interpretation. The commonly used acquisition means is to completely rely on high-resolution satellite shooting, and then filter and cut the data through software. With the progress of image processing methods, interpolation [1] and statistics [2] are used to reconstruct low-resolution images into higher-resolution images, which has become another way to obtain high-resolution remote sensing data sets. Remote sensing images are easily
affected by weather, day and night. Aiming at the problem of unclear images caused by the above problems, image Super Resolution (SR) reconstruction is an important means to improve image clarity. It plays an important role in the image preprocessing stage of remote sensing ground object classification [3], remote sensing target detection [4], remote sensing image coding [5] and other fields. With the emergence of antagonistic network in deep learning, image reconstruction technology has achieved a great breakthrough.

Remote sensing images have strict requirements on resolution, edge features and spectral information, and traditional machine learning has great limitations when processing remote sensing data. Therefore, deep learning soon becomes an important tool for remote sensing data processing, and supervised learning gradually comes into the eyes of researchers. Chao Dong et al. [6] first constructed an end-to-end deep convolutional network through spline interpolation to train low-resolution images. The test found that deep network was more effective in image reconstruction, but the construction of deep network was not sliming enough, the features were not comprehensive enough, and the design of loss function was relatively simple. It was not until 2016 that Ledig et al. [7] used the generated adverse network for the first time to improve the image resolution and adopted the Perceptual Loss Function. Although the reconstructed Loss Function could more effectively correct the generated network, the problem of incomplete feature extraction from the network was still not solved. Nan Fangzhe et al. [8] have analyzed Convolutional neural networks (CNN), deep residual networks and Recurrent neural networks in deep learning, (RNN), generation advesion network, dense convolutional neural network and the relationship between single frame image super-resolution processing, the hierarchical feature learning algorithm model is more conducive to high resolution image reconstruction. Bao Xiaoan et al. [9] proposed to take the improved generative adconfronntation network as a model, extract image features through the combination of coarse-grained main content and fine-grained detail edges, and realize image reconstruction by using a relatively complete feature space. However, the convergence speed of the network did not reach the ideal speed. Su Jianmin et al. [10] used a loss function based on the discriminant reconstruction error to accelerate the convergence of the network. Although the admissive network has a good reconstruction effect on image reconstruction, there is still room for improvement. Zhang Ning et al. [11] summarized the development of single-frame image super-resolution algorithms in recent years, and classified them into the SR model of standard convolutional neural network, the SR model based on residual network structure, and the SR model based on generated adveravtion network structure. Generated confrontation network can obtain more complete detailed information through confrontation training, so it becomes an important means of image super-resolution reconstruction. In view of the blurring of remote sensing images caused by precipitous terrain, incomplete hardware equipment and harsh image acquisition environment, traditional interpolation algorithm and deep convolutional neural network can improve the resolution to a certain extent. When the degree of image restoration is required to be higher and the precision is more rigorous, Choosing a more superior loss function and network model becomes the key to improve the clarity of remote sensing. In this paper, an improved SRGAN network is proposed to reconstruct remote sensing data sets. Firstly, the residual network is used to replace the normalized layer of the network, the judgment object of the discriminant is modified to be relative authenticity, the eigenvalue before activation is used to improve the perceived loss, and then the AID high-resolution remote sensing data is used to train and generate the network. The test data set is imported into the preservation model to get the high-resolution data, and the antagonistic loss is obtained by comparing the obtained high-resolution data with the real data through the discriminant network. The antagonistic loss is used to modify the generated network until the relatively superior high-resolution data is obtained. Compared with the traditional interpolation algorithm, statistical algorithm and deep convolutional network algorithm, the deep generation network in this paper has higher fineness, more comprehensive feature space and more reasonable discriminant function to ensure the convergence speed of loss.
2. Related Work

2.1. Dense block
Original SRGAN network of convolution firstly by 3 x 3 piece of convolution convolution and normalized layer, selection of parameters of linear rectifier function (PReLU) function activated, only increased the parameters of the very small amounts, makes the characteristic information more rich, the two get original feature vector and feature vector vector convolution process combined with the feature of guarantee the information complete. As shown in figure 1, residual convolution is by residual convolution blocks to replace all of the normalized layers, each layer of normalized replacement for 3 x3 convolution kernels convolution and PReLU activation layer, increase the depth and complexity of the network, every characteristics after the convolution has made full use of, generating network edge character processing is improved. The original convolution block and residual-density block are shown in Fig. 1.

![Original convolution block and residual dense block](image)

2.2. SRGAN Generation Network
The original SRGAN network inputs the low-resolution remote sensing data set into the generation network for training. First, it passes through a 9×9 convolution layer and connects five original convolution blocks to obtain a complete underlying feature space. Then, it adopts twice double upsampling and PReLU activation, and finally connects a 9×9 convolution layer to restore high-resolution remote sensing data. In ESRGAN network, the normalization process is completely removed and replaced with 5 residual dense blocks (PRDBs) to generate a new network. The rest of the network is generated by SRGAN network module. As shown in Equations (1) and (2), it can be seen that the size...
of the upsampling process is determined by the size of the convolution kernel and the filling step size \( m \).

\[
O = \left( \frac{i + 2p - k}{s} \right) + 1 \quad (1)
\]

\[
Out = (In - 1) * m + k \quad (2)
\]

In the formula, \( S \) is the step size, \( k \) is the size of the convolution kernel, \( p \) is the filling of the same dimension, \( I \) and \( O \) are the input and output of the convolution operation, \( In \) is the input size, \( k \) is the size of the convolution kernel, \( m \) is the filling step size, and \( Out \) is the output size. SRGaN generates the network, as shown in Figure 2.

![SRGAN generation network](image)

**Fig.2** SRGAN generation network

### 2.3. SRGaN discriminant network

SR remote sensing images output by the SRGN network are generated as input data, as shown in Figure 3. The middle layer adopts the basic network structure of VGG network, with convolution cores of 64, 128, 256 and 512 layers respectively, and the step sizes are set as 1 and 2 respectively. Finally, the full connection layer is connected. The cost function of Equations (3), (4) and (5) is used to compare SR remote sensing data and HR real data to calculate the loss.

\[
loss = loss_x + 10^3 loss_{Gen} \quad (3)
\]

\[
loss_x = \frac{1}{W_{ij}H_{ij}} \sum_{i,j} \sum_{x,y} \left( \hat{O}^{HR}(x, y) - \theta G(I^{HR}(x, y)) \right)^2 \quad (4)
\]

\[
loss_{Gen} = \sum_{n=1}^{N} - \log D(G(I^{HR})) \quad (5)
\]

Where, \( D \) is the probability that the original image belongs to the real resolution image, \( G \) is the probability that the reconstructed high-resolution image belongs to the real resolution image, \( W_{ij} \) and \( H_{ij} \) represent the pixel parameters on the x-scale and y-scale of the image, \( \theta \) is the deviation parameter, \( I^{LR} \) is the original low-resolution data, and \( I^{HR} \) is the real high-resolution data. The discriminator of the original SRGAN network is \( D(x) = \sigma(C(x)) \), where \( \sigma \) is a sigmoid function and \( C(x) \) represents a non-transformation discriminator. Improved discriminant for \( D(x_r, x_f) = \sigma(C(x_r) - IE_{x_f}[C(x_f)]) \), including \( IE_{x_f} \) said the false data to average operation, \( x_r, x_f \) respectively against the loss function of the two. The loss function of the improved SRGAN network is shown in Formula (6) and (7) as follows:

\[
E_{loss} = loss + \lambda L_G^{Ra} + \eta L_i
\]

\[
L_i = IE_{x} \parallel G(x_{i} - y) \parallel \quad (7)
\]
Where $\lambda$ and $\eta$ are the coefficients to balance different loss terms, loss is the result obtained by Formula 1-3, $I_E$, represents the operation of false data to average value, $L_G$ represents the average loss of G-net in the training process and the real value, $G(x_i)$ represents the evaluation and recovery image, and $y$ represents the real image. SRGAN discrimination network, as shown in Figure 3.

Fig.3 SRGAN discriminant network

2.4. ESRGAN super resolution algorithm

In order to obtain high-resolution remote sensing data sets more quickly, this algorithm uses stacked convolution, PReLU activation function, and upsampling to form a generated network. It uses the convolution layer of VGG19 network model, the weight of activation layer, and the sigmoid function to get the probability of generating remote sensing images. Finally, the generated network is adjusted according to the discriminant. The specific steps are as follows:

Step1: GID high-resolution remote sensing data is cut into data with a resolution of 480×480 according to the step size of 400 as the label of the generated network. Then, double-triple subsampling is used to obtain low-resolution remote sensing data as the training set, and the training data and label data are saved in the corresponding folder respectively.

Step2: The residual convolution block, the PReLU activation function layer and the double upper sampling layer were used to build the generated network, and the basic VGG network model, the full connection layer and Sigmoid were used to form the discriminant network.

Step3: Set discriminator function and network loss function, adjust learning rate and iteration number.

Step4: Import the training data, labels and VGG19 official model into the generation network and discriminant network for training, and save the training weights and training losses of G network and D network.

Step5: Use the saved G network weights to predict the predicted data, and calculate the mean error, peak signal-to-noise ratio and structural similarity. G network in ESRGAN, as shown in Figure 4.
3. Experimental results and analysis

3.1. Experimental data and environment

In this paper, remote sensing images adopted GID data set as training data and part of AID data set as test data, as shown in Figure 5. GID data included 150 pieces of five categories, including building area, farmland, forest, grassland and water area, and the resolution of each category was 7200×6800. As shown in Figure 6, AID is a new large aerial image data set. By collecting sample images from Google Earth images, the data set has 10,000 images in 30 categories, and the resolution of each category is 480×480. The test data includes 5 images of desert, farmland, forest and mountains. 150 pieces of high-resolution remote sensing data from the GID data set were cut into 480×480 data with a step size of 400 as training data, and 4 pieces of data from each category of desert, farmland, forest and mountain were selected as test data in the AID data set. In the experimental environment of Win10 system, in order to quickly build the antagonistic network, the current popular deep learning framework TensorFlow is adopted, and the auxiliary software is Matlab2018 to process the picture data. In the experiment, in order to improve the training speed of the experimental network, the computing mode of a single 8G capacity GPU was adopted, NVIDIA1080 card was used in the hardware, and the GUP accelerated mode was used. GID data, as shown in Figure 5. AID data, as shown in Figure 6
3.2. Experimental parameter setting

In order to better evaluate the validity of the experimental results, mean error, peak signal-to-noise ratio and structural similarity were set as evaluation indexes.

Mean-square error, MSE:

$$MSE = \frac{1}{hw} \sum_{i=1}^{h} \sum_{j=1}^{w} \left( x(i,j,k) - y(i,j,k) \right)^2$$  \hspace{1cm} (8)

Peak signal to noise ratio, PSNR:

$$PSNR = 10 \times \log \left( \frac{\left( 2^n - 1 \right)^2}{MSE} \right)$$  \hspace{1cm} (9)

Structural similarity, SSIM:

$$SSIM = l(x, y) \cdot c(x, y) \cdot s(x, y)$$  \hspace{1cm} (10)

Among them, $k$, $w$ and $h$ respectively represent the number of channels and the length and width of the image; $x(i,j,k)$ and $y(i,j,k)$ respectively represent the pixel coordinate information under a single pass; $l(x, y)$, $c(x, y)$ and $s(x, y)$ respectively represent the brightness, contrast and similarity of the image; $x$ and $y$ in the formula represent the transverse and longitudinal coordinate space, which are universally used in image operation. The larger the PSNR and SSIM values are, the smaller the image distortion is and the better the effect is.

3.3. Experimental results

In this paper, the data of building area, farmland, forest, grassland and water area of GID data were firstly cut into a data set of 480×480, and the SRRESNET, SRGAN and ESRGAN networks were trained respectively. Then, the data of desert, farmland, forest and mountain range were converted into 480×480 data, and the training model was tested for four types of data. The regions of interest obtained from the experimental results are shown in Figure 7 and Figure 8. Compared with the real high-resolution beach data, the results of SRRESNET are more fuzzy, while the results of SRGAN are clearer but the brightness is dim. The algorithm in this paper not only has high resolution, but can see the abrupt dune effect in the beach, and the brightness can be restored to the real standard. In the data of farmland, forest and mountains, the brightness of images of the three algorithms are not very different, and the boundary line between farmland and green space of the algorithm in this paper is clearer. The PSNR parameters of the proposed algorithm are up to 28.54dB, and the SSIM parameters are up to 0.74, which are improved compared with the parameters of other algorithms, fully demonstrating the effectiveness of the proposed algorithm and the ability to reconstruct super-resolution images that are more consistent with real images intuitively and microscopically. The key to reconstruct the image is to process the low-resolution image from the texture detail, and convert the original smooth image into the data with larger texture area, which can more clearly distinguish the sand dunes in the beach, the boundary of the...
farmland, the road in the forest, and the path in the mountain. The results of superresolution reconstruction of various methods are shown in Figure 7.

![Super-resolution reconstruction results of various methods](image)

**Fig. 7** Super-resolution reconstruction results of various methods

### 3.4. Experimental Analysis

According to ROI super resolution reconstruction structure analysis, although the three algorithms can improve the clarity of low-resolution data to a certain extent, the effect of SRRESNET is the worst, but compared with the low-resolution original data, the clarity of the road is higher, but the texture of the junction line of the two scenes is difficult to distinguish, and the texture of the trees is smooth. The main reason is that the deep network is relatively simple and the feature dimension is too large to be processed in detail. In view of the large difference of the characteristic matrix after convolution, the SRGAN network reduces the difference of the characteristic matrix data by adding the normalization layer after each convolution layer, and improves the sharpness of the experimental results. Obviously, the texture features of trees are better described, but the image brightness in the beach image is difficult to be reconstructed into the target brightness. The ESRGAN network adopts the method of dense residual network and network interpolation to balance the image quality without introducing artifacts and ensure the brightness of reconstructed remote sensing image. Evaluation parameters of various algorithms are shown in Table 1.

**Table 1. Various algorithm evaluation parameters**

| algorithm | SRResNet | SRGAN | ESRGAN |
|-----------|----------|-------|--------|
| MSE       | 221.51   | 276.42| 225.23 |
| PSNR      | 21.69    | 24.65 | 28.54  |
| SSIM      | 0.59     | 0.60  | 0.74   |
The result of ROI super-resolution reconstruction is shown in Figure 8

![Figure 8 ROI super-resolution reconstruction results](image)

Compared with other remote sensing super resolution algorithms, the classification algorithm in this paper has the following advantages: 1) Compared with other algorithms, this algorithm has a certain degree of improvement in the resolution of desert, farmland, forest and mountains. 2) In this algorithm, mean error, peak signal-to-noise ratio and structured similarity are used to measure the effectiveness of this experiment. 3) The selection of experimental data is diverse, and some categories are carefully cut and tested, which proves that the algorithm has the mobility and robustness.

4. Conclusion
This paper proposes a modified SRGAN network to reconstruct high resolution remote sensing data, and compared with the original SRGAN network structure, the use of residual error to replace the network of normalized layer, modify the judgment criterion of object is relative truth, using the characteristic value to improve the sensory loss before activation, and then use the AID cut high resolution remote sensing data generated to train the network, Adversarial loss is obtained by comparing the obtained high-resolution data with the real data through the discriminant network. Adversarial loss is used to modify the generated network until the relatively superior high-resolution data is obtained. This algorithm improves the resolution of desert, farmland, forest and mountains to a certain extent, which not only guarantees the reconstruction of image category, texture and region, but also guarantees the authenticity of image brightness. However, the algorithm has some limitations, and overfitting can be avoided to some extent by optimizing the model parameters. This method is not limited to the enhancement of ground object data. For the data of aircraft, tanks and ships that are difficult to be captured by satellites, this algorithm can be applied to reconstruct the data and strengthen military strategic monitoring and deployment.
References

[1] DAI J, WU Q Y. Super-resolution Optimization Algorithm of Edge Preserving Interpolation Based on Single Image [J]. Computer Applications, 2018, 38(S1):191-193.

[2] LIU Y. Research on Super-resolution Calculation Method Based on Edge Direction [J]. Fujian Computer, 2017, 33(09):91-92+113.

[3] ZHANG Z X. Research on Hyperspectral Image Super-resolution Restoration Algorithm Based on Ground Object Category [D]. Beijing University of Technology, 2017.

[4] ZHANG W G. Vision Detection of Circular Hole Position and Pose Based on Image Super-resolution Reconstruction [D]. Zhejiang University, 2019.

[5] HUANG T Y, SUN T T, ZHOU Z H, et al. Based on adaptive coupling half dictionary learning super-resolution image reconstruction [J/OL]. Computer application research: 1-6 [2019-05-31]. https://doi.org/10.19734/j.issn.1001-3695.2018.11.0852.

[6] Dong C, Loy C C, He K, et al. 2014. Image Super-Resolution Using Deep Convolutional Networks [J]. Transactions on Pattern Analysis & Machine Intelligence, 38(2):295-307.

[7] Ledig C, Theis L, Huszár F, et al. 2016. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [J].

[8] NAN F Z, QIAN Y R, XING Y N, et al. Based on the deep study of single image super-resolution reconstruction research review [J/OL]. Computer application research: 1-7 [2019-05-31]. HTTP: // http://kns.cnki.net/kcms/detail/51.1196.TP.20181219.1953.003.html.

[9] BAO X A, GAO C B, ZHANG N, et al. Image super-resolution method based on generated against networks [J/OL]. Journal of zhejiang university of technology (natural science edition) : 1-10 [2019-05-31]. HTTP: // http://kns.cnki.net/kcms/detail/33.1338.TP.20190322.0834.004.htm ml.

[10] SU J M, YANG L X. Single frame based on the generated against network remote sensing image super-resolution [J/OL]. Computer engineering and application: 1-9 [2019-05-31]. http://kns.cnki.net/kcms/detail/11.2127.tp.20181128.1730.054.html.

[11] ZHANG N, WANG Y C, ZHANG X, et al. Based on the deep study of single image super-resolution reconstruction research progress [J/OL]. Acta automatica: 1-21 [2019-10-28]. https://doi.org/10.16383/j.aas.c190031.

[12] W. Xu, G. Xu, Y. Wang, X. Sun, D. Lin and Y. Wu, "High Quality Remote Sensing Image Super-Resolution Using Deep Memory Connected Network," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 8889-8892. doi: 10.1109/IGARSS.2018.8518855

[13] W. Ma, Z. Pan, J. Guo and B. Lei, "Super-Resolution of Remote Sensing Images Based on Transferred Generative Adversarial Network," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 1148-1151. doi: 10.1109/IGARSS.2018.8517442

[14] Q. Yuan, L. Yan, J. Li and L. Zhang, "Remote sensing image super-resolution via regional spatially adaptive total variation model," 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, 2014, pp. 3073-3076. doi: 10.1109/IGARSS.2014.6947126

[15] J. Fu, Y. Liu and F. Li, "Single Frame Super Resolution with Convolutional Neural Network for Remote Sensing Imagery," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018, pp. 8014-8017. doi: 10.1109/IGARSS.2018.8518584

[16] CHEN C, LIU M M, LIU B, et al. Image super-resolution reconstruction algorithm based on residual networks [J/OL]. Computer engineering and application: 1-8 [2019-05-31]. HTTP: // http://kns.cnki.net/kcms/detail/11.2127.tp.20190527.1720.009.html.

[17] Lu W Y, Wang J H. Science of Surveying and Mapping, 2016, 41(12):53-58+69.

[18] LI C M, DENG K Z, SUN J Y, et al. Science of Surveying and Mapping, 2018, 43(10):82-89.

[19] WU Q, TIAN Y, ZHOU C P, et al. Current situation and development of super resolution research of remote sensing image [J]. Science of Surveying and Mapping, 2008, 33(06):66-69+15