Research on Image Super Resolution Reconstruction Based on Deep Learning

Zhiwen Chen, Qiong Hao(✉), and Liwen Liu

Wuhan Railway Vocational College of Technology, Wuhan 430205, China
wruqhao@163.com

Abstract. To enhance the precision and clarity of graphic and image depictions, we propose a super-resolution image reconstruction method driven by the power of deep learning. This method initiates by obtaining the reconstruction object from graphics and images, subsequently simulating their degradation process. The preprocessing of initial images is accomplished via registration and expansion, setting a solid foundation for the subsequent stages. Deep learning algorithms are employed to interrogate and dissect the inherent features of the graphics and images. Subsequently, a lineup of techniques including feature fusion and bilinear interpolation are deployed to gain super-resolution reconstruction results of the graphics and images. Upon examining and juxtaposing our deep learning-based method with conventional techniques, we discerned a noticeable advantage of the former. Intriguingly, the resolution deviation within the image reconstruction results derived via our idealized strategy has been remarkably minimized. Concurrently, peak signal-to-noise ratio and structural similarity attributes have been substantially augmented. This unique confluence of improvements as embodied in our approach places it squarely as a potential game-changer in the domain of super-resolution image reconstruction.

Keywords: Deep Learning · Image Reconstruction · Super-Resolution Image

1 Introduction

A key index to evaluate the resolution of an image acquisition system is the spatial resolution of the image. Super resolution reconstruction is the process of obtaining the highest quality image from one or more low resolution images through signal processing and image processing methods. Super resolution reconstruction is widely used in image coding, image processing, high-definition television, image synthesis, face recognition and monitoring, medical diagnosis and other fields. Due to camera cost, limited bandwidth, limited storage space, limited computing power and other reasons, image resolution is often compressed [1]. An intuitive example is that due to the limitation of network bandwidth, compressed low resolution images are often transmitted in the network. Super resolution reconstruction method has become an important part of the network receiver.

L. Yun et al. (Eds.): ADHIP 2023, LNICST 548, pp. 435–450, 2024.
https://doi.org/10.1007/978-3-031-50546-1_29

In the context of super-resolution reconstruction, the pursuit of increased spatial resolution relies on acquiring multiple low-resolution images of the same scene. These images must not only be undersampled and aliased but also possess sub-pixel displacement relative to one another. When the displacement is by integer pixels, the information provided on the same low-resolution grid sampling points is identical across all images. Consequently, integrating these images for reconstruction would be futile, as no new information about the scene would be gained. However, when there is sub-pixel displacement between low-resolution images, along with aliasing, each image no longer represents a simple transformation or duplication of another. Each low-resolution image captures a specific recording outcome for the same scene. In this scenario, the new scene information contained within each low-resolution image can be harnessed for super-resolution image reconstruction. Current reconstruction methods predominantly focus on identifying easily extractable new features and extracting image details for training purposes. This approach, unfortunately, exhibits several drawbacks, including heavy reliance on predefined features, high computational complexity, limited robustness, and fixed input image requirements. These limitations ultimately hinder the quality of super-resolution reconstruction for graphical images. Acknowledging these limitations, there is a need for more advanced approaches that address these challenges. This research aims to overcome the shortcomings by proposing a novel reconstruction method that leverages deep learning techniques. By utilizing deep learning algorithms, the proposed method aims to enhance feature extraction capabilities, reduce computational complexity, improve robustness, and flexibly adapt to varying input image conditions. The objective is to ultimately enhance the quality and fidelity of super-resolution reconstruction for graphic images.

The proposed approach contributes to the broader field of super-resolution reconstruction by addressing key limitations of existing methods. By combining the power of deep learning with the unique characteristics of sub-pixel displacement and aliasing, this research demonstrates the potential to significantly improve the quality and robustness of reconstructed images. The outcomes of this study will provide valuable insights and advancements towards more effective super-resolution reconstruction techniques, benefiting fields such as medical imaging, remote sensing, and computer vision.

Deep learning inSearch technology, data mining, machine learningMT, natural language processing, Multimedia learning, voice, recommendation and personalization technology, as well as other related fields have made many achievements. In order to improve the final image reconstruction quality, the deep learning algorithm is applied to optimize the image super-resolution reconstruction method.

2 Design of Super-Resolution Reconstruction Method for Graphics and Images

Graphic image reconstruction is a technology to obtain the shape information of three-dimensional objects through digital processing from the external measured data of objects. With the support of deep learning algorithm, the operation process of optimizing the design of super-resolution reconstruction method for graphics and images is shown in Fig. 1.

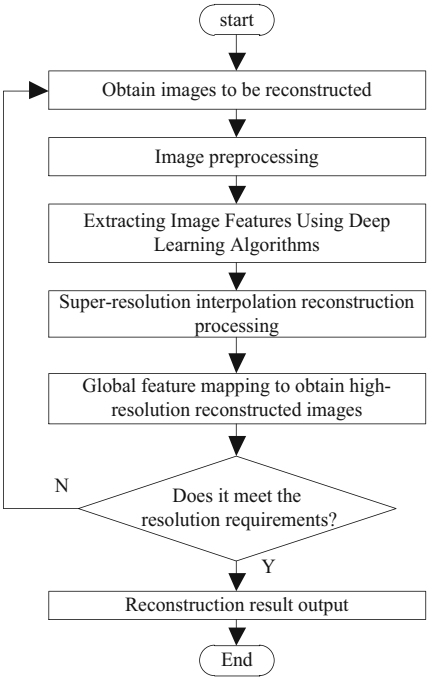


Fig. 1. Flow chart of super-resolution reconstruction of graphic image

The optimized super-resolution reconstruction method takes graphics as the basic content of the image, and obtains the super-resolution reconstruction results of graphics and images through feature fusion, feature reconstruction and other steps.

2.1 Acquire Graphic Image Reconstruction Object

Let the graphic image be sampled separately on the two coordinate systems x and y k_1 and k_2. Second, the Dirac distribution function is used to represent the sampling function of the grid image:

$$f(x, y) = \sum_{k_1=1}^{M} \sum_{k_2=1}^{N} \chi(x - k_1 \Delta x, y - k_2 \Delta y) \tag{1}$$

where M and N Respectively represents the length and width of the image to be reconstructed, $\chi(x, y)$ Is Dirac distribution function, Δx and Δy The sampling interval [2] on the corresponding real coordinate axes x and y. Where Dirac distribution function $\chi(x, y)$. The expression of is as follows:

$$\chi(\mathrm{x}, \mathrm{y}) = \begin{cases} 1, x = y = 1 \\ 0, \text{ else} \end{cases} \tag{2}$$

The Dirac distribution function expressed by formula 2 meets the following conditions:

$$\int_{-\infty-\infty}^{\infty}\int_{-\infty}^{\infty}\chi(\mathrm{x},\mathrm{y})=1 \tag{3}$$

Substitute the solution result of formula 2 into formula 1 to get the sampling result of graphic image.

2.2 Simulate the Degradation Process of Graphics and Images

In the process of natural image imaging, the original natural scene and target image are a continuous signal. The imaging system samples the input continuous signal to obtain the observed digital image. Because the original image will be affected by motion and deformation, blur, downsampling, noise and other factors, the actually acquired low resolution image will be degraded to a large extent compared with the original high resolution image [3]. Usually, the linear shift invariant or linear shift variant model is used to model the fuzzy interference. Downsampling is the discretization and quantization of the input analog continuous signal. The commonly used downsampling strategies are average sampling or ideal sampling, and generally assume average sampling;The quantization process is to convert analog signals into digital signals. In addition, the image imaging process will also be polluted by noise, resulting in errors, mainly including sensor, sampling and quantization, and model errors. The degradation model of graphics and images is shown in Fig. 2.

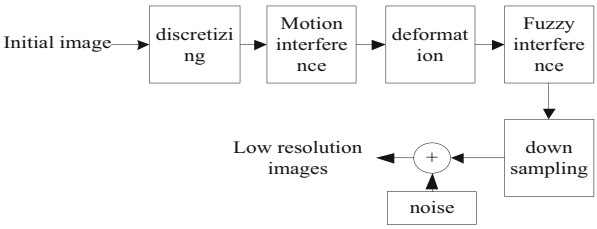


Fig. 2. Graph Image Degradation Model

According to the representation principle in Fig. 2, it can be concluded that the degradation result of graphics and images is:

$$g(x,y)=D\cdot B\cdot H_{atm}M\cdot f(x,y)+n_{\text{additive}} \tag{4}$$

Among $f(x,y)$ is the sample object of the initially collected graphics and images, D, B, H_{atm} and M the corresponding real down sampling matrix expression, optical ambiguity matrix expression, atmospheric disturbance matrix and deformation matrix expression, n_{additive} Is the additive noise in the spatial domain, and the final result is obtained $g(x,y)$ is the low resolution image to be reconstructed.

2.3 Initial Graphic Image Pre-processing

Before super-resolution reconstruction of graphic images, the initial image needs to be preprocessed first. The preprocessing steps mainly include registration and expansion. Image registration is the estimation of image motion information. Image registration is based on the scene, and it needs to use each point to estimate [4]. That is, it is necessary to find a highly accurate point-to-point matching model between input images: suppose two different images are taken for the same scene, and the corresponding points in the other image can be found for one point in one image, which are exactly the same points of the real scene. If the image registration is not accurate, then the information used in the image reconstruction is not correct, can not express the information of the real image, and can not make the image add more details than the original. Therefore, motion estimation is a very critical step in the super-resolution reconstruction algorithm, and if the estimation accuracy cannot reach the sub-pixel level, it will greatly affect the quality of the reconstructed image. The specific registration process in the graphic image preprocessing can be quantified as:

$$g'(x, y) = h_{\text{Registration}}\left[g(x, y)\right] \tag{5}$$

Among $h_{\text{Registration}}()$ represents the registration transformation function, which includes three parts: rigid body transformation registration, affine transformation registration and projection transformation registration. Rigid body transformation registration refers to the coordinate transformation of the three relations of translation, rotation and mirror image, that is, the image frames must be from the same perspective and taken by the same sensor. The rigid body transformation registration function can be expressed as:

$$h_{\text{Registration}}^{\text{rigid body}} = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \end{bmatrix} + \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{6}$$

Variables in Formula 6 ε_x and ε_y Is the offset in the horizontal and vertical directions respectively, θ is the rotation angle. Affine transformation registration is a coordinate transformation of translation, rotation, scaling and mirroring, which is slightly more general than rigid body transformation model and increases the ability to handle scaling transformation [5]. The projection transformation registration can complete the linear transformation of two-dimensional and three-dimensional planes. The affine transformation registration and projection transformation registration functions are as follows:

$$\begin{cases} h_{\text{Registration}}^{\text{affine}} = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ h_{\text{Registration}}^{\text{projection}} = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{cases} \tag{7}$$

among λ_{ij} Is a nonsingular matrix element. The registration result of the initial graphic image can be obtained by weighted fusion of Formula 6 and Formula 7 and substituting

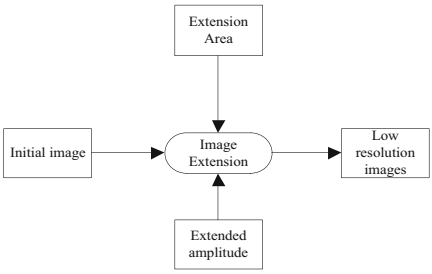


Fig. 3. Schematic diagram of graphic image expansion processing

them into Formula 5. In addition, the expansion principle of graphics and images is shown in Fig. 3.

According to the principle shown in Fig. 3, the result of image expansion processing is:

$$g_{extend}(x, y) = g(x, y) \times \kappa_{extend} \tag{8}$$

Parameters in the above formula κ_{extend} Is the image expansion coefficient. All pixels in the initial graphic image are processed according to the above process to obtain the preprocessing result of the initial graphic image.

2.4 Using Deep Learning Algorithm to Extract Image Features

In order to ensure the quality of image reconstruction, convolutional neural network algorithm is used for image reconstruction. The basic structure of the network is shown in Fig. 4.

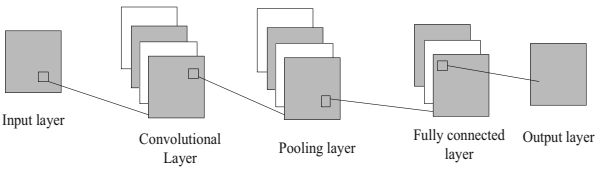


Fig. 4. Structure of deep convolution neural network

As can be seen from Fig. 4, the deep convolution neural network is composed of convolution layer, pooling layer, full connection layer, etc. The convolution layer is the most important part of the convolution neural network, and its main function is to extract various features in the image. Input a feature map. The convolution core in the convolution layer and the feature map are convolutioned and added with the offset weight value. Repeat this process until the whole map is traversed. A new feature map is obtained and used as the input of the next network layer [6]. The convolution process of the convolution layer is shown in Fig. 5.

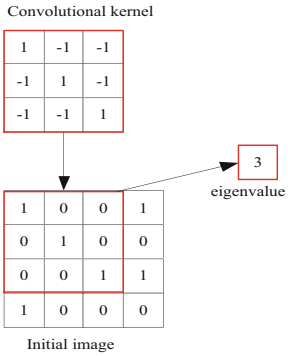


Fig. 5. Schematic diagram of convolution process of convolution layer

For a training process, more than one feature needs to be extracted. The convolution process of a convolution kernel can be regarded as a feature extraction. In order to get more features, we need to use multiple convolution kernels. Each convolution kernel is equivalent to a feature. The feature map of the previous layer in the convolution layer is convolved with the learnable kernel function, and the feature map of this layer is output after the activation function. A trainable parameter will be added after convolution b At this time, the output result of the convolution layer is:

$$x^l_{\text{convolution}} = f_{\text{activation}}\left(\sum_{i \in M_j} x_i^{l-1} * u_{ij}^l + b^l_{\text{convolution}}\right) \tag{9}$$

among $x^{l-1}_{\text{convolution}}$ by $l-1$ Image input value of layer, u^l_{ij} Is the convolutional kernel, $f_{\text{activation}}()$ Is the activation function, and its expression is:

$$f_{\text{activation}}(x) = \frac{1}{1+e^{-x}} \tag{10}$$

Each output will give a deviation b However, for some special output characteristic graphs, the input characteristic graphs will be convolved with some special kernel functions. The pooling layer can reduce the dimension of the feature map processed by the convolution layer. After being processed by pooling layer, the size of feature map tends to become smaller, which is also the result of reducing redundant information of feature map in pooling layer. The pooling layer reduces the size of the image and correspondingly reduces the number of feature parameters, which also speeds up the convergence of the model and improves the efficiency of network training [7]. Most importantly, the pooling layer can keep the most important features of the image from being lost after processing the feature map, that is, feature invariance. The pooling layer only removes some irrelevant information, while the important feature information is still not broken and can be expressed. The specific pooling process is shown in Fig. 6.

According to the above process, the output result of pooling layer is:

$$x^l_{\text{pooling}} = f_{\text{activation}}\left(\beta^l_{\text{pooling}} \text{down}\left(x_j^{l-1}\right) + b^l_{\text{pooling}}\right) \tag{11}$$

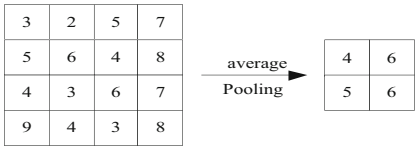


Fig. 6. Diagram of average pooling process of pooling layer

Each output characteristic graph has a multiple deviation value β^l_{pooling} and an additional deviation value b^l_{pooling}. Among, down$\left(x_j^{l-1}\right)$ is a down sampling function, which determines the pooling method. The features of the input image are obtained through the convolution layer and pooling layer, and then can be input into any classifier with differentiable weights. Generally speaking, the full connection layer is equivalent to the hidden layer of a multi-layer perceptron, which is usually located at the last part of the network model. As a classifier, the classification results are obtained through the feature vectors obtained in the convolution layer and pooling layer. Its principle is similar to that of the artificial neural network [8]. In the actual image feature extraction process, in addition to forward propagation, backpropagation is also required to improve the quality of feature extraction. In the backpropagation process of deep learning, first assign a random value in the interval (-1,1) to the network's convolution kernel weight parameter, randomly select the training set, including k input samples and the expected output corresponding to the input samples, and mark it as x_{in} and E_{out}. Through this process, the input and output of each neuron in the hidden layer are calculated and the partial derivatives of the error function in each layer to each neuron in the output layer are computed based on the actual output and expected output of the last layer of the network. The connection weight value is then adjusted using the derivatives of each neuron in the output layer and the output of each neuron in the hidden layer. Similarly, the derivatives of each neuron in the hidden layer and the input of each neuron in the input layer are used to modify the connection weight value. The global error is evaluated using the error formula to determine if it meets the requirements. If not, the next learning sample and its expected output are selected, and the process returns to the third step to commence the next round of learning. This continues until the error reaches the preset precision or the maximum number of learning times is reached [9]. Through the forward and backward propagation of the convolutional neural network, the extraction results of graphic image features are obtained, in which the extraction results of graphic image texture features are:

$$\tau_{\text{texture,infra-red}} = \frac{P(i,j|d,\varphi)}{\sum_i \sum_j P(i,j|d,\varphi)} \tag{15}$$

Among d and φ respectively represents the distance and direction vector between the target pixel and the image center, $P(i,j|d,\varphi)$ is the gray level co-occurrence matrix. Similarly, we can get the extraction results of other features such as color, shape and so on.

2.5 Realize Super-Resolution Reconstruction of Graphics and Images

The image super-resolution reconstruction is carried out through bilinear interpolation. The basic idea of bilinear interpolation is as follows: the output pixel value at the interpolation point after interpolation is its size in the input image 2×2 is the average value of four vertex samples in the field, and the output pixel interpolates the gray value of four pixels around the corresponding position of the input image in both horizontal and vertical directions [10]. The pixel value of any point in the square is obtained by interpolation, and the bilinear interpolation result is:

$$f_{\text{interpolation}}(x, y) = \alpha x + \psi y + \gamma xy + \zeta \tag{16}$$

Equation 16 represents a hyperbolic paraboloid that will be fitted through four known points. Coefficient α, ψ, γ and ζ. It is selected by the known pixel values of four vertices. The final image super-resolution reconstruction results are:

$$F = \max\left(0, \tau * f_{\text{interpolation}}(x, y)\right) \tag{17}$$

Repeat the above process to obtain the super-resolution reconstruction results of graphics images.

3 Experimental Analysis of Image Super-Resolution Reconstruction Quality Testing

To evaluate the efficacy of the deep learning-based image super-resolution reconstruction method optimized for enhanced performance, we conducted quality test experiments employing a contrast testing approach. The experiment aimed to assess the reconstruction quality by selecting multiple sets of graphic images that met the resolution requirements as experimental samples. Low-resolution images were generated through the process of degradation and downsampling, and served as inputs for the super-resolution reconstruction algorithm under investigation. Employing the configured development environment, we implemented the optimized reconstruction method and supplied the prepared image samples to obtain the corresponding reconstruction results. To measure the quality of the reconstructed images, we employed multiple quality indicators. Furthermore, we compared the performance of our optimized design method with traditional reconstruction results to highlight the advantages and efficacy of our approach.

3.1 Configuration and Reconstruction Method Development Environment

To fulfill the operational requirements of deep learning algorithms and the super-resolution reconstruction method for graphics and images, a comprehensive experimental development environment is established encompassing both hardware and software components. The hardware environment comprises two distinct components: a high-performance server employed in processing the neural network training process, and a low-configuration personal computer dedicated to software UI design and neural network testing. The neural network undergoes training and testing phases, which are facilitated

by Ubuntu 16.04, a 64-bit operating system providing essential support. Anaconda serves as the package management tool for efficient virtual environment management, while PyCharm acts as the integrated development environment (IDE) platform to streamline the development and coding processes. Software design, including UI aspects, is conducted utilizing PyQt5, PyQt5 tools, and other relevant packages on the Windows 10 operating system. The UI design phase is executed within the QtDesigner application software on macOS Catalina. Developing the image super-resolution reconstruction method relies on the PyTorch framework, celebrated as one of the most prominent frameworks among deep learning researchers. The framework's dynamic graph computation capabilities facilitate seamless model building and training, while also affording users the option to leverage GPU acceleration for enhanced computational performance.

3.2 Prepare Graphic Image Samples

In this experiment, a graphic image data set is used as the training set and test set of the reconstruction method. The data set contains 2000 training data and 200 test data. The resolution of the initial graphic image samples is 72 dpi. The initial preparation of some graphic image samples is shown in Fig. 7.

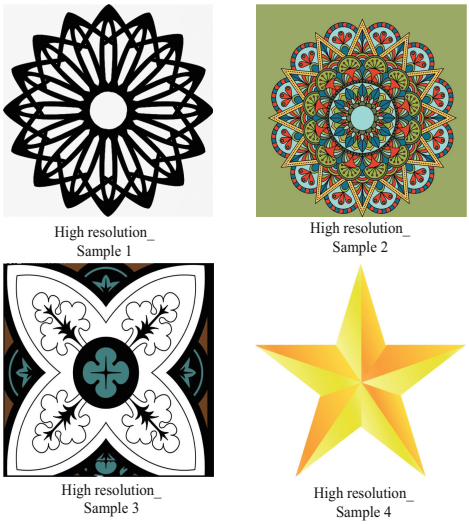


Fig. 7. Schematic Diagram of Initial Sample of Graphic Image

The OpenCV Bicubic interpolation method is used to preprocess the data set for 4-fold down sampling to generate a low resolution image, that is, a graphic image to be reconstructed. The processing result is shown in Fig. 8.

Similarly, the preparation results of all reconstructed object samples in the experiment can be obtained.

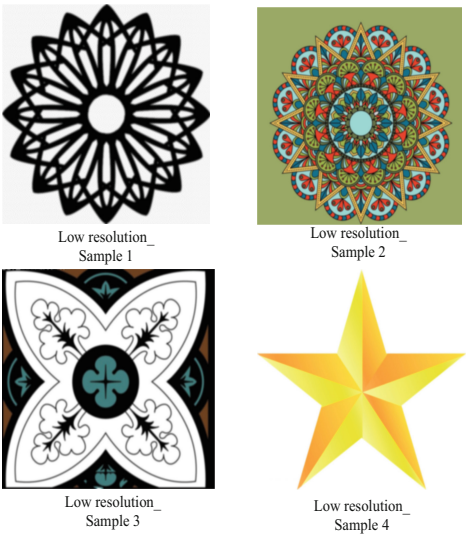


Fig. 8. Schematic diagram of image samples to be reconstructed

3.3 Input Operating Parameters of Deep Learning Algorithm

The convolutional neural network algorithm in the deep learning algorithm is selected as the technical support for the optimization design method, and the operation parameters of the algorithm need to be set from both the structure and implementation aspects (Table 1).

Table 1. Convolutional Neural Network Table

Convolutional Layer	Convolutional kernel size	Convolutional Kernel Number
1	9 * 9	64
2	1 * 1	32
3	5 * 5	1
4	9 * 9	64
5	1 * 1	32
6	5 * 5	1

The activation function is sigmoid, and the step size of the network is set to 1. The learning rate is set to 10-4, the learning rate is reduced to half of the original 250000 iterations, a total of 1000000 iterations, and the image batch size is set to 64.

3.4 Describe the Reconstruction Quality Test Experiment Process

Under the configured experimental environment, the development of the super-resolution reconstruction method of graphics and images based on deep learning for optimal design is realized, and the set operating parameters of the deep learning algorithm are input into the running program of the reconstruction method. In order to reflect the advantages of optimization design method in reconstruction quality, the traditional super-resolution reconstruction method based on maximum a posteriori estimation and super-resolution reconstruction method based on multiple sparse representation are set as the two comparison methods of the experiment. The comparison method is developed under the same experimental environment. In the actual operation process, the comparison method is not allowed to call the deep learning algorithm,The parallel switching mode is adopted to ensure the complementary influence between reconstruction methods. Select a reconstruction object from the prepared graphic image samples, and obtain the corresponding reconstruction results by running the reconstruction method. Figure 8 shows the reconstruction output results of some graphic images obtained by the optimization design method (Fig. 9).

According to the above method, the reconstruction result output by the traditional reconstruction method can be obtained, and the resolution information of the reconstructed image can be recorded.

3.5 Setting Super Resolution Reconstruction Quality Test Indicators

Resolution deviation, peak signal to noise ratio and structural similarity are set as test indicators to measure the quality of super-resolution reconstruction of graphic images. The numerical results of resolution deviation indicators are:

$$\Delta\eta = \eta_{\text{target}} - \eta_{\text{rebuild}} \tag{18}$$

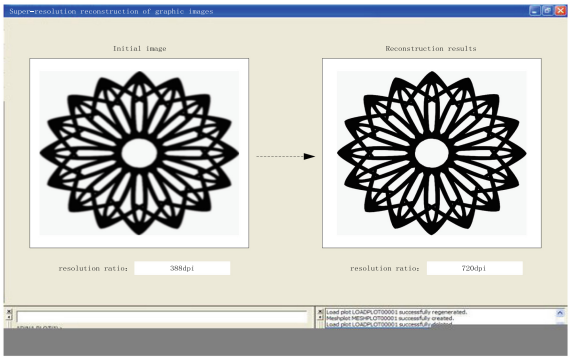
among η_{target} and η_{rebuild} They are reconstruction target resolution and actual reconstruction resolution. In addition, the test results of peak signal to noise ratio index and structural similarity index are as follows:

$$\begin{cases} \sigma = \dfrac{N(\text{image})}{N(\text{noise})} \\ \mu = \displaystyle\sum_{i=1}^{n_{\text{pixel}}} \left(\left| x_{\text{original}}(i) - x_{\text{rebuild}}(i) \right| + \left| y_{\text{original}}(i) - y_{\text{rebuild}}(i) \right| \right) \end{cases} \tag{19}$$

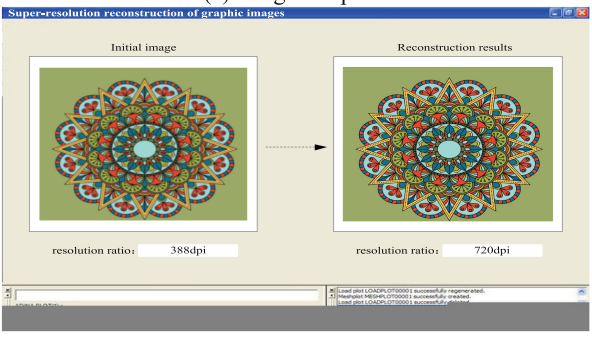
Among, $N(\text{image})$ and $N(\text{noise})$ It respectively represents the number of effective pixels and noise points in the reconstruction result image, $\left(x_{\text{original}}(i), y_{\text{original}}(i)\right)$ It is the No i Pixel value of pixels, $(x_{\text{rebuild}}(i), y_{\text{rebuild}}(i))$ Is the pixel value of the reconstructed image.

3.6 Experimental Results of Image Super-Resolution Reconstruction Quality Test

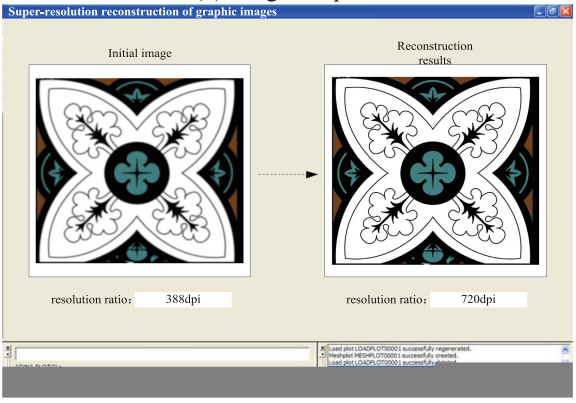
Through the statistics of relevant data, the test results of the reconstructed image resolution deviation index are obtained, as shown in Table 2.



(a) Image sample 1



(b) Image sample 2



(c) Image sample 3

Fig. 9. Super resolution reconstruction results of graphics and images

By substituting the data in Table 2 into Formula 18, it is calculated that the average resolution deviation of the two comparison methods is 112.5 dpi and 50 dpi respectively, and the average value of the reconstructed image resolution deviation of the graphic image super-resolution reconstruction method based on deep learning optimized design is 2.5 dpi. In addition, the peak signal to noise ratio and structure similarity test results

Table 2. Reconstructed image resolution deviation test data table

Image sample number	Reconstruction target resolution/dpi	Super-resolution reconstruction method based on maximum a posteriori estimation	Super-resolution reconstruction method based on multiple sparse representations	Image super-resolution reconstruction method based on deep learning
1	720	650	680	720
2	720	660	650	715
3	720	540	680	715
4	720	550	680	720
5	720	600	660	720
6	720	650	650	720
7	720	660	680	720
8	720	550	680	710

of the reconstructed image are obtained through the calculation of Formula 19, as shown in Fig. 10.

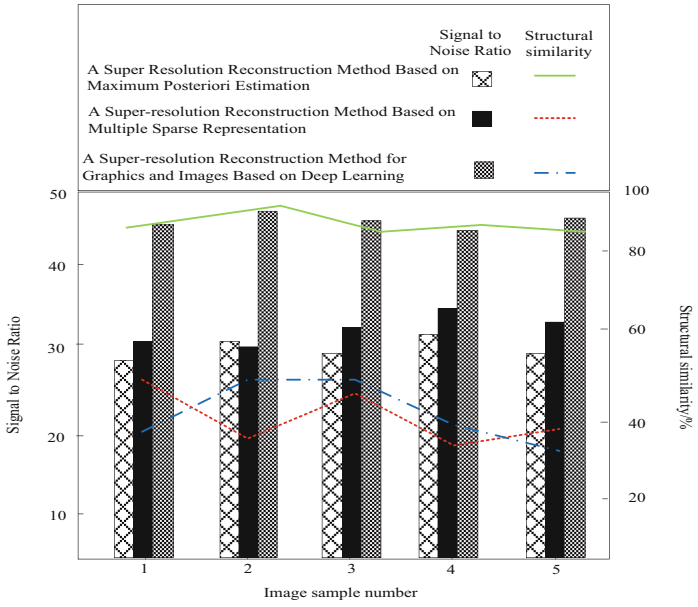


Fig. 10. Image SNR and Structure Similarity Test Results

From Fig. 10, it can be intuitively seen that for the image sample numbered 2, the maximum posterior estimation method has a signal-to-noise ratio of 31 and a structural similarity of 90.2%; The signal-to-noise ratio of the multiple sparse representation method is 29.2, and the structural similarity is 38.6%; The signal-to-noise ratio of the deep learning method is 46.8, and the structural similarity is 48.5%; Compared with two traditional super-resolution reconstruction methods, the optimized design method has higher peak signal-to-noise ratio and structural similarity.

4 Conclusion

In recent years, the utilization of image super-resolution reconstruction has seen a significant rise across various domains. For instance, in gaming, somatosensory peripheral devices enable the capture of players' body movements, enhancing human-computer interaction. By employing super-resolution reconstruction technology to improve the quality of depth data, the accuracy of posture recognition can be substantially elevated, greatly enhancing players' overall gaming experience. Super-resolution reconstruction techniques also find application in 3D reconstruction. By enhancing the density and accuracy of point cloud data obtained from depth cameras, a more realistic surface model of 3D objects can be achieved. This advancement promotes the utilization of 3D reconstruction in diverse domains like biomedicine, video surveillance, criminal case analysis, augmented reality, and more. In the field of unmanned driving, depth information plays a crucial role in determining the 3D position of the vehicle. Implementing super-resolution reconstruction on depth images enables unmanned vehicles to achieve higher positioning accuracy. Simultaneously, it enhances the vehicle's capability for accurate environmental description and obstacle avoidance operations. The development and maturation of depth image super-resolution reconstruction technologies have significantly propelled their widespread adoption in the realm of computer vision. Consequently, this progress has presented new and heightened requirements for future research endeavors. By leveraging deep learning algorithms, the reconstruction quality of graphics and images can be substantially improved. This advancement holds immense research significance and practical value. In summary, deep learning-based super-resolution reconstruction methods for graphic images demonstrate immense potential across various applications, including image enhancement, video compression, and medical image processing. Nevertheless, further research is necessary to enhance the method's robustness and generalization ability, while also addressing challenges such as computational speed and model size.

References

1. Kholil, M., Ismanto, I., Fu'Ad, M.N.: 3D reconstruction using structure from motion (SFM) algorithm and multi view stereo (MVS) based on computer vision. IOP Conference Series: Materials Science and Eng. **1073**(1), 12066–12072 (2021)
2. Lee, H., Chon, B.H., Ahn, H.K.: Rapid misalignment correction method in reflective fourier ptychographic microscopy for full field of view reconstruction. Opt. Lasers Eng. **16**(5), 138–145 (2021)

3. Inam, O., Qureshi, M., Laraib, Z., et al.: GPU accelerated Cartesian GRAPPA reconstruction using CUDA. J. Magn. Reson. **337**(21), 107175–107186 (2022)
4. Zhang, J., Xu, T., Zhang, Y., et al.: Multiplex Fourier ptychographic reconstruction with model-based neural network for Internet of Things. Ad Hoc Netw. **111**(22), 102350–102359 (2021)
5. Shi, Q., Hui, W., Huang, K., et al.: Under-sampling reconstruction with total variational optimization for Fourier ptychographic microscopy. Optics Communications **10**, 126986–126993 (2021)
6. Zhang, J., Tao, X., Yang, L., et al.: The integration of neural network and physical reconstruction model for Fourier ptychographic microscopy. Optics Communications **504**(22), 127470–127483 (2022)
7. Pan, B., Betcke, M.M., Arridge, S.R., et al.: Photoacoustic reconstruction using sparsity in curvelet frame: image versus data domain. IEEE Trans. Computational Imaging **26**(9), 8–15 (2021)
8. Zhang, Y., Zhang, Z., Li, T.: Research on image super-resolution reconstruction based on deep learning. J. Phys. Conf. Ser. **1802**(4), 42034–42043 (2021)
9. Jia, R., Wang, X.: Research on super-resolution reconstruction algorithm of image based on generative adversarial network. J. Phys. Conf. Ser. **1944**(1), 12014–12019 (2021)
10. Agarwal, V., Chitkariya, P., Miglani, A., et al.: Deep learning-based image processing for analyzing combustion behavior of gel fuel droplets. Smart Electrical and Mechanical Syst. **16**(22), 65–85 (2022)
11. Meng, Z., Zhang, J., Qiu, J.: Multi supervised loss function smoothing image super-resolution reconstruction. Chinese J. Image Graphics **27**(10), 2972–2983 (2022)
12. Ni, R., Zhou, L.: Face image super-resolution reconstruction method based on Convolutional neural network. Computer and Digital Eng. **50**(01), 195–200 (2022)
13. Ge, P., You, Y.: Super-resolution reconstruction of light field images based on sparse representation. Progress in Laser and Optoelectronics **59**(02), 94–100 (2022)
14. Liang, M., Wang, H., Zhang, Y., Li, J.: Image super-resolution reconstruction method based on accelerated residual network. Computer Appl. **41**(05), 1438–1444 (2021)
15. Wang, H.Y., Zhang, K.X., Guan, W.Z.: Single image super-resolution reconstruction method based on dense Inception. Computer Appl. **41**(12), 3666–3671 (2021)