Design Studies and Intelligence Engineering
L.C. Jain et al. (Eds.)

doi:10.3233/FAIA231514

Research on the Method of Artistic Image Restoration Based on Artificial Intelligence

Jian WANG[1]
Changsha University of science and technology, School of Design Art, Changsha, 410114, China

Abstract: In order to improve the current inpainting algorithm for natural art images, an artistic image inpainting model is proposed by integrating artificial intelligence and embedding multi-scale attention expansion convolution. According to the uniqueness of artistic images, the network structure of DMFB is improved, and its repair model combines two plug-and-play optimization modules: extended convolution block and coordinate attention mechanism. The extended convolution module is used to capture multi-scale context information, and the coordinate attention mechanism is used to improve the remote migration ability of features of the repaired network. The combination of the two makes the repair results meet the visual visibility and semantic rationality. The experimental results show that ablation experiments are carried out to evaluate the effectiveness of each module in this model. In this paper, the hybrid extended convolution and coordinated attention mechanism are used to train the network, and the effectiveness of MADC is verified. When the expansion rate changes to 2, 4, 6 and 8, the method in this paper will lead to moderately distorted structure and fuzzy texture. On the contrary, there is no coordination attention. In the branch of mechanism, the output of this method shows texture defects and discontinuities. By using these two branches, the method in this paper has achieved good results in structure and texture. The quantitative evaluation of branches without coordinated attention mechanism is given when the expansion rate is fixed at 2, 4, 6 and 8. Conclusion: The results restored by this method are more consistent with the real image in human senses, and the objective evaluation of PSNR, SSIM and MSE has also been improved to some extent.

Keywords: image inpainting, deep learning, attention mechanism, artistic image

1 Introduction

Culture and art play a very important role in the development of human civilization. Painting is an important cultural and artistic form. It is an important part of Chinese culture, rooted in the soil of national culture, and reflects a wide range of real-life content through beautiful artistic forms, thus reflecting the cultural outlook and aesthetic taste of all ethnic groups. It is a unique and important way for human beings to observe and express the world. For thousands of years, China has produced a large number of paintings. The study of these paintings is an important means for people to understand the development history of human history, culture, art and science and technology, thus further promoting the development of human civilization. These paintings span

[1] Corresponding Author.

thousands of years and contain rich cultural, artistic, scientific and historical values. However, due to natural disasters (earthquakes) and natural weathering, as well as more and more economic activities of human beings, some paintings are more or less damaged or missing, which seriously affects the appreciation, cultural creativity and cultural communication based on these paintings. Therefore, how to use the latest scientific and technological means to digitally repair these damaged or missing artistic images is of great significance.

With the development of computer collection technology and hardware and the states emphasis on the protection of cultural heritage, with the support of relevant state departments, the relevant cultural protection departments have carried out a large number of high-precision digital collection projects of painting works (artistic images), such as typical projects such as the painting series of past dynasties, the complete works of Song paintings and the complete works of Yuan paintings, and accumulated a huge amount of artistic image data [1]. These digitized data provide a data base for intelligent analysis and processing of artistic images. At the same time, computer artificial intelligence technology, especially deep learning technology, has made a major breakthrough, such as the accuracy of image classification based on deep learning has reached 96%, and the accuracy of face recognition has reached more than ninety-seven percent. The development of computer artificial intelligence technology provides technical feasibility for digital restoration of artistic images. In addition, at present, Chinas per capita GDP has exceeded \$10,000. According to international experience, when the per capita GDP reaches \$10,000, the demand for cultural and entertainment consumption will explode. Therefore, the restoration of artistic images is of great significance to the current cultural industry in China [2].

2 Literature review

The image inpainting algorithm based on PDE was first put forward by the concept of image inpainting technology [3]. Anantarisirichai, N. et al. put forward a fast marching algorithm based on BSCB, which can completely repair the image by determining the direction of the isochromatic line in the area to be repaired and progressing layer by layer [4]. Zhang, Z. and others put forward an algorithm to keep the direction of the isochromatic line by using Euler-Lagrange equation and anisotropic diffusion, which is called total variation (TV) [5].

Traditional image inpainting usually uses the redundancy of the image itself, that is, the information of the known area is used to complete the unknown area. The approximate method is to search the pixels around the unknown area and select a pixel with higher reliability to complete it. However, due to the low efficiency of this method, it is difficult to be competent for the task of artistic image restoration. With the vigorous development of deep learning, many researchers began to apply deep learning to the task of artistic image restoration. In the task of repairing old photos, two encoders are trained to project old photos and high-definition photos into hidden space, and a hidden space mapping module is designed to transform the feature distribution of old photos into the feature distribution of high-definition photos, so as to realize the repair of old photos [6]. Focusing on the task of Dunhuang frescoes restoration, the researchers put forward a method of Dunhuang frescoes restoration based on structure guidance. They used the representation of color in high-dimensional feature space to improve the quality of regional color restoration, making the overall color of the restored Dunhuang frescoes relatively smooth [7]. In addition, by constructing two restoration paths, namely, the path to predict the prior distribution of missing areas and the path to generate conditions, good results have also been achieved in the natural scene image restoration task [8]. With the rapid popularization of the new generation of artificial intelligence science and technology and products, the application scope of artificial intelligence is more and more extensive, which can be applied to medical impact analysis, investment analysis, agricultural robots, self-driving cars, virtual reality and other fields. The structure of artificial intelligence industry can be divided into application layer, platform interface layer, technology layer and foundation layer. At present, the technological innovation of most scientific researchers mainly focuses on the technical layer and the basic layer, while the project innovation of enterprises mainly focuses on the application layer. However, the platform interface layer, which integrates algorithms and provides interfaces for the application layer, has relatively little work. This part of the work is of great significance for promoting the development of artificial intelligence industry and the transformation of scientific research results. If the algorithm is limited to the laboratory and cannot be transformed into a tool to facilitate peoples lives, then the practical significance of this work is quite limited. Through the intervention of the platform interface layer, many developers who have no experience in artificial intelligence can simply and quickly use mature artificial intelligence algorithms for application development, thus promoting the popularization and application of artificial intelligence technology. At present, more and more large enterprises around the world have begun to build their own artificial intelligence platforms. In addition to the purpose of profit, I also hope to expand the companys own popularity in this way to attract more outstanding talents.

In recent years, the image inpainting method based on deep learning has shown good inpainting results in the task of inpainting large missing areas. The existing methods are mainly aimed at natural images, but the restoration effects of structural distortion, inconsistent lines and blurred textures are often output on artistic images. The main reason is that artistic images are generally unstructured data, which have complex line drawing, rich colors and abstract content. Therefore, in order to solve the above problems, this paper proposes an artistic image restoration method embedded with multi-scale attention expansion convolution.

3 research methods

In order to make the natural image inpainting model better applied in artistic images, this paper proposes an artistic image inpainting model embedded with multi-scale attention expansion convolution. The model structure is improved on the basis of generating countermeasure network. The generator is an encoding-decoding network, and the discriminator has two branches-global discriminator and local discriminator. The generator generates natural and reasonable restoration images, and the discriminator performs network countermeasure training to assist the generator to generate more credible restoration images. The proposed network structure and loss function will be described in detail below.

3.1 Generation Network

The generated network adopts a coding-decoding structure, in which the characteristic coding network shown in Table 1 consists of a series of 3×3 convolution layers and multiple attention expansion convolution. In this network, a 5×5 convolution kernel is used in the first convolution layer, which can fully extract the potential information features of the input image [9,10]. Then, a multi-scale attention expansion convolution module (MADC) is embedded behind the convolution layer. This module combines the advantages of extended convolution and attention mechanism, which can not only expand the receptive field, but also effectively extract the interesting features of the image. In addition, the convolution layer with 3×3 convolution kernel and the combination module of multiple attention expansion convolution are used continuously in this paper, which makes more image feature details remain in the encoding process and makes the decoding process generate images with consistent content and reasonable semantics.

Table 1. Coding network structure

Layer number	**Layer type**	**Convolution kernel size**	**dilatation coefficient**	**step**	**Output size**	**Number of output channels**
1	Convolution layer	5*5	2	1	256*256	64
2	Convolution layer	3*3	1	2	128*128	128
3	MADC					
4	Convolution layer	3*3	1	1	128*128	128
5	MADC					
6	Convolution layer	3*3	1	2	64*64	256
7	MADC					
8	Convolution layer	3*3	1	3	32*32	512
9	MADC					

At the same time, Table 2 gives the detailed architecture of the decoding network. The decoding process consists of a 3×3 convolution layer, three consecutive up-sampling layers and a 3×3 convolution layer, in which the last two up-sampling layers use ReLU activation function to output information, and the last convolution layer uses Tanh activation function to output information [11]. A series of up-sampling and convolution layers continuously map the features extracted from the coding network to the real image, thus generating a credible restoration image.

Table 2. Decoding network structure

Layer number	**Layer type**	**Convolution kernel size**	**dilatation coefficient**	**step**	**Output size**	**Number of output channels**
1	Layer type	3*3	1	1	32*32	512
2	Convolution layer	3*3	1	2	64*64	256
3	Up-sampling	3*3	1	2	128*128	128
4	Up-sampling	3*3	1	2	256*256	64
5	Up-sampling	3*3	1	1	256*256	3
6	Convolution layer	3*3	1	2	64*64	256

3.2 Multi-scale attention expansion convolution

Multi-scale attention expansion convolution (MADC) is a plug-and-play optimization module, which combines the mixed expansion convolution block with the coordination attention layer, as shown in Figure 1. It uses extended convolution to expand the receptive field without increasing the amount of calculation, and at the same time uses attention mechanism to retrieve interested information and suppress useless information [12]. The following two parts will be introduced in detail.

(1) Extended convolution block

The extended convolution is an improvement on the standard convolution, which is a special extended convolution with an expansion rate of 1, while the extended convolution with an expansion rate greater than 1 can expand the receptive field and extract background information on different scales, and at the same time, it does not need to introduce additional computational overhead in the feature extraction process. By expanding the acceptance domain of the network, expanding convolution is also very important for the image inpainting task, because it does not need to downsample or increase the number of model parameters [13,14]. Firstly, the serial layer with increased expansion rate is used to expand the receptive field of output neurons. However, the use of large expansion rate information may only benefit some large areas, but it may have disadvantages for small areas. Here, this paper uses the idea of DMFB, which consists of four consecutive convolution layers, and the expansion rates in DMFB are 1, 2, 4 and 8 respectively. However, the convolution kernel of the expanded convolution is not continuous, and not all pixels participate in the calculation, so adopting the expansion rates of 1, 2, 4 and 8 will lose the continuity of information and cause the grid effect. When using multiple expansion convolution, it is necessary to design the expansion coefficient of each convolution kernel so that it can just cover the underlying features [15]. Using different expansion coefficients of 1, 3, 5 and 7 can ensure that the receptive field is equally large and more information can be used, otherwise grid effect will occur, and the following ablation experiments prove the effectiveness of expansion coefficients of 1, 3, 5 and 7. Therefore, this paper changes the expansion rate to 1, 3, 5 and 7. Specifically, the first convolution layer on the left in MADC reduces the number of channels of input features to 64, which is used to reduce parameters, and then sends these processed features to four branches, and extracts multi-scale features through expanded convolution with different expansion factors, which is recorded as $x_i (i = 1,2,3,4)$. except x_1, each x_i Has a corresponding 3 x 3 convolution, using $K_i(\cdot)$ Express. By cumulative addition, various sparse multi-scale features are combined to obtain dense multi-scale features. This article uses y_i express $K_i(\cdot)$ The output of. The specific combination mode is as follows (1):

$$y_i = \begin{cases} x_i, & i = 1 \\ K_i(x_{i-1} + x_i), & i = 2 \\ K_i(y_{i-1} + x_i), & 2 < i \leq 4 \end{cases} \tag{1}$$

The next step is to simply use 1 x 1 convolution to fuse the connected features. In a word, the extended convolution layer increases the receptive field of general extended convolution, and its parameters are less than those of general convolution kernel.

(2) Coordinate attention mechanism

Coordinate attention mechanism uses more effective methods to capture location space information and channel relations, so as to enhance the feature representation ability of mobile networks. In the coordinate attention mechanism, two-dimensional channel attention is decomposed into two independent one-dimensional feature codes, and both coding processes perform feature aggregation operations along the spatial direction, as shown in Figure 1. One of the one-dimensional feature coding processes can capture the long-distance dependence between features along the spatial direction, and the other can retain the accurate location information of features. Then, the feature maps extracted from the two branches are fused and encoded into a set of joint feature maps with both direction perception and accurate position information. The complementary application of feature maps to the input feature map can improve the representation of the feature map of interest [16]. It plays an important role in image restoration network. Therefore, coordinate attention is inserted into the network of this paper to guide image restoration. In order to verify the effectiveness of coordinate attention, this paper uses Grad-CAM as a visualization tool to visualize it. Compared with the two attention mechanisms of CABM and coordinate attention, coordinate attention mechanism can emphasize the relative position of the most interesting background area in human senses, and the color of the background block of interest is brighter and more obvious.

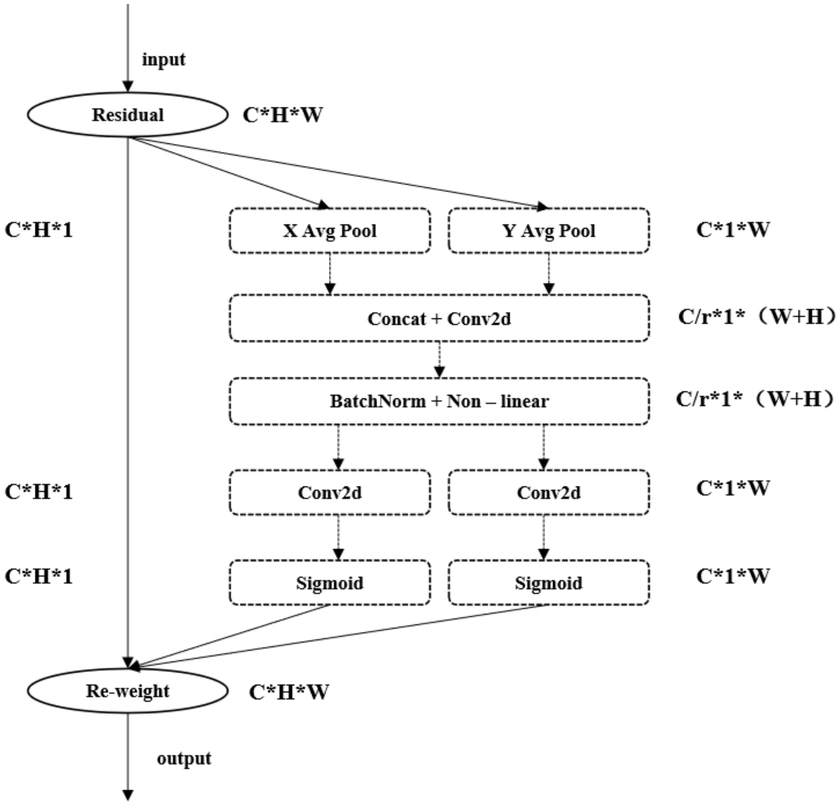

Figure 1. Coordinate Attention Model

3.3 Authentication Network

In order to make the image repaired by the generator network consistent with the original image structure and texture, this paper uses RaGAN to pursue a more realistic generated image. The identification network consists of a global discriminator and a local discriminator, which ensures the consistency between local and global contents. The global discriminator consists of six convolution layers, with a convolution kernel size of 5, an expansion coefficient of 2 and a step of 2. The local discriminator consists of five convolution layers, with a convolution kernel size of 5, an expansion coefficient of 2 and a step of 2. Except for the last layer, Leaky ReLu with a slope of 0.2 is used, as shown in Table 3. In this paper, VGG is used to extract the feature space of the picture, and the discriminator is used to calculate the anti-loss on this feature space, which can better help the generator understand the image information. In addition, the network adopts spectral normalization to realize the training of network stability, thus making the generated image more realistic and reasonable [17,18].

Table 3 Identification network structure

Branch type	Layer type	Convolution kernel size	dilatation coefficient	step	Output size	Number of output channels
Global discriminator	1	Convolution layer	5*5	2	128*128	64
	2	Convolution layer	5*5	2	64*64	128
	3	Convolution layer	5*5	2	32*32	256
	4	Convolution layer	5*5	2	16*16	512
	5	Convolution layer	5*5	2	8*8	512
	6	Convolution layer	5*5	2	4*4	512
Local discriminator	1	Convolution layer	5*5	2	64*64	64
	2	Convolution layer	5*5	2	32*32	128
	3	Convolution layer	5*5	2	16*16	256
	4	Convolution layer	5*5	2	8*8	512
	5	Convolution layer	5*5	2	4*4	512

3.4 Loss function

In order to improve the training efficiency of the generator, this paper adopts four loss functions, including reconstruction loss, perception loss, style loss and confrontation loss, which are introduced as follows.

(1) Reconstruction loss: Reconstruction loss is very useful for the regional convolution filter to learn and generate meaningful content in different regions, so this paper uses L1 reconstruction loss to ensure that the restored image is consistent with the real image, and its calculation formula is as follows (2):

$$L_{\text{rec}} = \left\| I_{\text{out}} - I_{\text{gt}} \right\|_1 \tag{2}$$

In the above formula (2),I_{out} Its an image of network repair, I_{gt} Is a real image;

(2) Perceptual loss: The perceptual loss formula is used to extract the advanced semantic features between the real image and the inpainted image, and to perceive the quality of the inpainted image. In this paper, the perceptual loss function that has been defined on the VGG-16 network and pre-trained on ImageNet is used, as shown in the following formula (3)]:

$$L_{\text{prec}} = E\left[\sum_i \frac{1}{N_i} \left\| \Phi_i(I_{\text{out}}) - \Phi_i(I_{\text{gt}}) \right\|_1\right] \tag{3}$$

In the above formula (3), Φ_i It is the first in the VGG16 network diagram. i The activation diagram of the layer, in this experiment. Φ_i Corresponding to ReLu1_1, ReLu2_1, ReLu3_1, ReLu4_1 and ReLu5_1.

(3) Style loss: The convolution layer transposed from the decoder will bring artifacts similar to chessboard. In order to alleviate this influence, this paper introduces style loss, which is widely used in image inpainting and style transfer tasks, and is also an effective

tool to fight against "checkerboard" artifacts. The given size is $C_j \times H_j \times W_j$, its calculation formula is as follows (4):

$$L_{\text{style}} = E_j\left[\left\| G_j^{\Phi}(I_{\text{out}}) - G_j^{\Phi}(I_{gt})\right\|_1\right] \tag{4}$$

(4) Relative average LS countermeasure loss: Antagonistic loss is the catalyst to repair the missing area. The discriminator in this paper adopts relative average LS countermeasure loss and uses local and global discriminators. Both local discriminator and global discriminator use spectral normalization to realize stable training. For generators, antagonistic losses are defined as:

$$L_{adv} = -E_{xr}\left[\log\left(1 - D_{ra}(\mathrm{x_r}, x_f)\right)\right] - E_{xf}\left[\log\left(1 - D_{ra}(\mathrm{x}_f, x_r)\right)\right] \tag{5}$$

In the above formula (5),$D_{ra}(\mathrm{x_r}, x_f) = \text{sigmoid}\left(C(x_r) - E_{\mathrm{xf}}[C(x_f)]\right)$ and $C(.)$ indicates that the last layer has no local or global discriminator with sigmoid function,$(\mathrm{x_r}, x_f)$They are real images and network-generated images.

(5) The total loss function is the following formula (6):

$$L_{\text{total}} = \lambda_{\mathrm{r}} L_{\mathrm{re}} + \lambda_{\mathrm{p}} L_{\text{prec}} + \lambda_{\mathrm{s}} L_{\text{style}} + \lambda_{\text{adv}} L_{\text{adv}} \tag{6}$$

among$\lambda_{\mathrm{r}}, \lambda_{\mathrm{p}}, \lambda_{\mathrm{s}}, \lambda_{\mathrm{adv}}$Is the balance parameter.

4 Result analysis

4.1 Experimental environment

In order to verify the superiority of this model in restoring artistic images, NVIDIARTX2080Ti graphics card is selected as the experimental platform, and Pytorch 1.10.2 architecture is used to build a restoration network with 32 GB of memory. In this paper, the model is qualitatively and quantitatively compared with DMFB, RFR, PRVS, CSA and MEDFE algorithms on art image data sets, and ablation experiments are designed to prove the effectiveness of the MADC optimization module in this model. The motives for choosing these methods as comparison algorithms are as follows: (1) Inspired by the work of DMFB, this paper has made some improvements in the network structure. In the DMFB literature, multiple DMFBs with expansion rates of 1, 2, 4 and 8 are used continuously in the repair network to obtain a larger and more effective receiving domain. In this method, the expansion rates are changed to 1, 3, 5 and 7, and the interest characteristics of the four combined branches are enhanced by combining the CA attention mechanism, so the DMFB and the DMFB are combined. (2)RFR module, like MADC, is also an image inpainting method based on attention mechanism. The difference is that the attention mechanism in RFR is knowledge consistency attention mechanism, while MADC is coordinate attention mechanism, which can verify the effectiveness of this attention mechanism. (3) In 3)PRVS, a series of visual structure reconstruction layers (VSR) are superimposed in the encoding and decoding stages, and the structural information is integrated into the reconstructed feature map, and the visual structural features are alternately reconstructed in a gradual way. Using structural diagram to guide the generation of the next feature diagram is an image inpainting method based on structural information. Artistic images usually have complex topological structures, so a masterpiece based on structural information is selected for comparison [19]. (4)CSA is an image inpainting based on coherent semantic attention mechanism, which is different from this paper. Comparing the two can verify the effectiveness of this papers attention mechanism; (5)MEDFE recombines the deep features of the encoder into structural features and the shallow features into texture features, and balances the output features in the channel domain and the spatial domain. Similar to the structure in this paper, it is a single-stage network, so MEDFE is compared with the experimental results in this paper.

4.2 Construction of artistic image data set

At present, most of the public data sets used in the field of image restoration are natural images. However, due to the important cultural value of artistic images to the nation, most of them are distributed in national museums, art galleries, cultural centers and other places, and some of them are in the hands of private collectors, so there are few authoritative public data sets for the restoration research and application of artistic images in deep learning. With the world paying more and more attention to the protection of cultural heritage, with the support of the relevant state departments, the relevant cultural protection departments have carried out a large number of high-precision digital collection projects of paintings (artistic images). Typical projects, such as the painting series of past dynasties, the complete works of Song paintings and the complete works of Yuan paintings, have accumulated a huge amount of artistic image data, and research teams in some universities have also collected and built databases based on artistic images. These digitized data provide a data base for intelligent analysis and processing of artistic images.

I have obtained the data set of China traditional landscape painting opened in Chinese-Landscape-Painting, which contains 2192 pieces. Besides Chinas landscape painting, I also collected the artistic images of famous foreign painters. Including Alfred Sisley, eugene boudin, claude monet, Pierre Auguste Renoir, Henri Rousseau and Vincent Van Gogh, there are 464, 549, 937, 583, 80 and 934 works, with a total of 3,574 works (oil painting data source is https://www. wikiart。 org/)。 At present, there are few open art image data sets. In order to enlarge the data set as much as possible to verify the effectiveness of the model, this paper cuts the traditional high-definition landscape painting in China into several images with rich contents. Finally, a total of 21,000 pictures were collected, all of which were 256x256x3. In order to ensure the randomness and uncertainty of the damaged area, the position of the mask in the image is random, and the size of the damaged area is also random.

4.3 Qualitative comparison

In this paper, the method is used to inpaint images with different scale masks. From the figure, it can be found that the inpainting results of images with mask ratio of 0-20% are no different from the real images in human senses, and the structure is consistent and the texture is reasonable. However, with the expansion of the damaged area, the semantic information of the image is missing too much, and there will be discontinuous lines, and the detailed texture is blurred and inconsistent with the real image.

In order to show the superiority of the art image inpainting method, the inpainting results of this method and five schemes, such as DMFB, RFR, PRVS, CSA and DEM Fe, are qualitatively evaluated. As a result of RFR, PRVS, CSA and DEMFE, there are obvious visual artifacts in the missing area, including image blur and distortion (shown in the white box). Among them, when there is face information with rich semantic information in artistic images, although DMFB produces more pleasing content, there is a lack of correlation between the cavity area and the background area in its restoration map, such as line discontinuity [20]. Compared with the above two methods, MADC method has a good visual effect, can effectively generate results with real texture, and has a high correlation between the repair area and the background area.

4.4 Quantitative comparison

In this paper, different mask ratios are used for numerical evaluation on artistic image data sets. During training, the data set is divided into training set, verification set and test set in the ratio of 8:1:1. 1000 pictures were randomly selected from the art image data set for testing. For the evaluation indicators, this paper follows the use of SSIM, MAE and PSNR evaluation indicators. The evaluation results are shown in Figure 2 and Table 4. Fig. 2 is a quantitative evaluation and comparison between the method in this paper and five methods: DMFB, RFR, PRVS, CSA and MEDFE. Table 4 compares the repair results of this method under different scale masks. From the two tables, it is found that the quality of image restoration in this paper is better than the existing methods in pixel level, structure level and perception level. At the same time, this method has a good effect on filling irregular holes under different scale masks.

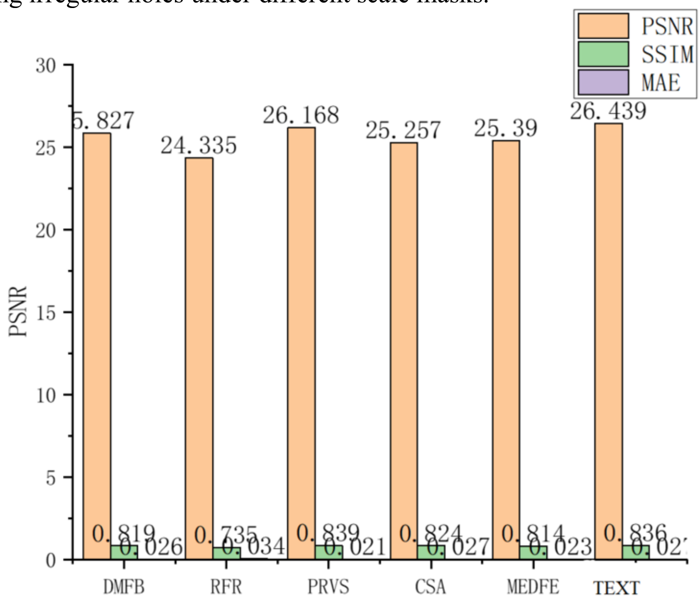

Figure 2. Quantitative evaluation results on art image data sets

Table 4. Quantitative evaluation results of different scale masks

Mask	10%-20%	20%-30%	30%-40%	40%-50%
PSNR	29.902	27.3 82	26.439	24.579
SSIM	0 .927	0.867	0.836	0.761
MAE	0 .011	0.021	0.027	0.032

4.5 Ablation experiment

Finally, this paper conducts ablation experiments to evaluate the effectiveness of each module in this model. In this paper, the hybrid extended convolution and coordinated attention mechanism are used to train the network, and the effectiveness of MADC is verified. When the expansion rate changes to 2, 4, 6 and 8, the method in this paper will lead to moderately distorted structure and fuzzy texture. On the contrary, there is no coordination attention. In the branch of mechanism, the output of this method shows texture defects and discontinuities. By using these two branches, the method in this paper has achieved good results in structure and texture. The quantitative evaluation of branches without coordinated attention mechanism is given when the expansion rate is fixed at 2, 4, 6 and 8.

5 conclusions

In this paper, an artistic image inpainting method embedded with multi-scale attention expansion convolution is proposed. In the coding stage of the network, there are five convolution layers, four of which are followed by an optimization module, and the optimization module consists of several attention expansion convolutions (MADC), and the missing content is roughly extracted by re-creating loss training. In the decoding stage, three upsamples and two convolution layers are used to generate a natural and reliable restored image. In the loss function, four different loss functions, namely reconstruction loss, style loss, perception loss and confrontation loss, are combined to train and generate the confrontation network. In the mixed expansion convolution, this paper fuses four expansion convolution blocks with expansion rates of 1, 3, 5 and 7 respectively, and obtains dense multi-scale features. At the same time, the coordination attention layer is inserted into the network to enhance the expression ability of interested features, improve the remote migration ability of features embedded in the network, and make the repair results meet the visual visibility and semantic rationality. Experiments verify the effectiveness of this method.

references

[1] Willemink, M. J., & Noël, P. B. (2019). The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence. European radiology, 29, 2185-2195.

[2] Jin, Z., Iqbal, M. Z., Bobkov, D., Zou, W., Li, X., & Steinbach, E. (2019). A flexible deep CNN framework for image restoration. IEEE Transactions on Multimedia, 22(4), 1055-1068.

[3] Koonjoo, N., Zhu, B., Bagnall, G. C., Bhutto, D., & Rosen, M. S. (2021). Boosting the signal-to-noise of low-field MRI with deep learning image reconstruction. Scientific reports, 11(1), 8248.

[4] Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. Artificial intelligence review, 1-68.

[5] Zha, Z., Yuan, X., Zhou, J., Zhu, C., & Wen, B. (2020). Image restoration via simultaneous nonlocal self-similarity priors. IEEE Transactions on Image Processing, 29, 8561-8576.

[6] Muckley, M. J., Riemenschneider, B., Radmanesh, A., Kim, S., Jeong, G., Ko, J., ... & Knoll, F. (2021). Results of the 2020 fastMRI challenge for machine learning MR image reconstruction. IEEE transactions on medical imaging, 40(9), 2306-2317.

[7] Knoll, F., Murrell, T., Sriram, A., Yakubova, N., Zbontar, J., Rabbat, M., ... & Recht, M. P. (2020). Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge. Magnetic resonance in medicine, 84(6), 3054-3070.

[8] He, W., Yao, Q., Li, C., Yokoya, N., Zhao, Q., Zhang, H., & Zhang, L. (2020). Non-local meets global: An iterative paradigm for hyperspectral image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(4), 2089-2107.

[9] Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2020). Residual dense network for image restoration. IEEE transactions on pattern analysis and machine intelligence, 43(7), 2480-2495.

[10] Arabi, H., AkhavanAllaf, A., Sanaat, A., Shiri, I., & Zaidi, H. (2021). The promise of artificial intelligence and deep learning in PET and SPECT imaging. Physica Medica, 83, 122-137.

[11] Racine, D., Becce, F., Viry, A., Monnin, P., Thomsen, B., Verdun, F. R., & Rotzinger, D. C. (2020). Task-based characterization of a deep learning image reconstruction and comparison with filtered back-projection and a partial model-based iterative reconstruction in abdominal CT: a phantom study. Physica Medica, 76, 28-37.

[12] Mei, Y., Fan, Y., Zhang, Y., Yu, J., Zhou, Y., Liu, D., ... & Shi, H. (2023). Pyramid Attention Network for Image Restoration. International Journal of Computer Vision, 1-19.

[13] Pan, X., Zhan, X., Dai, B., Lin, D., Loy, C. C., & Luo, P. (2021). Exploiting deep generative prior for versatile image restoration and manipulation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(11), 7474-7489.

[14] Pan, J., Dong, J., Liu, Y., Zhang, J., Ren, J., Tang, J., ... & Yang, M. H. (2020). Physics-based generative adversarial models for image restoration and beyond. IEEE transactions on pattern analysis and machine intelligence, 43(7), 2449-2462.

[15] Malik, J., Kiranyaz, S., & Gabbouj, M. (2021). Self-organized operational neural networks for severe image restoration problems. Neural Networks, 135, 201-211.

[16] Lu, J., Li, N., Zhang, S., Yu, Z., Zheng, H., & Zheng, B. (2019). Multi-scale adversarial network for underwater image restoration. Optics & Laser Technology, 110, 105-113.

[17] Hou, G., Pan, Z., Wang, G., Yang, H., & Duan, J. (2019). An efficient nonlocal variational method with application to underwater image restoration. Neurocomputing, 369, 106-121.

[18] Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., & Timofte, R. (2021). Plug-and-play image restoration with deep denoiser prior. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10), 6360-6376.

[19] Yu, K., Wang, X., Dong, C., Tang, X., & Loy, C. C. (2021). Path-restore: Learning network path selection for image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10), 7078-7092.

[20] He, W., Yao, Q., Li, C., Yokoya, N., Zhao, Q., Zhang, H., & Zhang, L. (2020). Non-local meets global: An iterative paradigm for hyperspectral image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(4), 2089-2107.