J. Meas. Sci. Instrum., 2025, 16(2): 195-204. http://jmsi.publish.founderss.cn/
DOI: 10.62756/jmsi.1674-8042.2025019 jmsi@nuc.edu.cn

Generative adversarial mural inpainting algorithm based on structural and texture hybrid enhancement

TAO Meifeng[1], CHEN Yong[1,2*], ZHAO Mengxue[1], ZHANG Jiaojiao[1]

1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, Lanzhou 730070, China

***Corresponding author:** CHEN Yong (edukeylab@126.com)

Received: October 5, 2023 **Revised:** November 22, 2023 **Accepted:** January 20, 2024

Abstract: For the existing deep learning image restoration methods, the joint guidance of structure and texture information is not considered, which leads to structural disorder and texture blur in the restoration results. A generative adversarial mural inpainting algorithm based on structural and texture hybrid enhancement was proposed. Firstly, the structure guidance branch composed of dynamic convolution cascade was constructed to improve the expression ability of structure features, and the structure information was used to guide the encoder coding to enhance the edge contour information of the coding feature map. Then, the multi-granularity feature extraction module was designed to obtain the texture features of texture guided branches, and the multi-scale texture information was used to guide the decoder to reconstruct and repair, so as to improve the texture consistency of murals. Finally, skip connection was used to promote the feature sharing of structure and texture features, and the spectral-normalized PatchGAN discriminator was used to complete the mural restoration. The digital restoration experiment results of real Dunhuang murals showed that the proposed method was better than the comparison algorithms in both subjective and objective evaluation, and the restoration results were clearer and more natural.

Key words: image processing; mural inpainting; structural and texture enhancement; dynamic convolution; multi-granularity feature extraction

0 Introduction

Dunhuang Mogao Grottoes is a splendid treasure of world culture and art. Its murals are world-famous and have high research value. However, due to factors such as harsh natural environment, man-made damage, and the fragility of painting materials, the murals have various degrees of peeling, cracks, molds, and other diseases. The application of digital restoration technology to the restoration and protection of ancient murals has become a current research hotspot[1].

Image restoration is a computer technology that uses the prior information of the image and the prior information of the existing intact area to estimate and fill the damaged area, so that the restoration results can meet the human visual perception. Image inpainting methods can be divided into two groups: traditional methods and deep learning based approaches. Traditional image restoration methods use pixel diffusion, sample patch matching, and sparse representation reconstruction, which can complete the repair of small damaged images[2-4].

With the advantages of deep learning in image feature learning, deep learning-based image restoration has gradually become the main research trend in the field of image restoration. Quan et al.[5] proposed a novel three-stage inpainting framework with local and global refinement. It improves the local and global consistency of the repair results. Lahiri et al.[6] proposed a prior guided GAN based semantic inpainting model. It used data-driven noise prior learning to improve the image repair efficiency. Wang et al.[7] proposed a dynamic selection network for image inpainting. It makes full use of the information in the known region and achieves good restoration results. Li et al.[8] proposed a recurrent feature reasoning for image inpainting, which was mainly constructed by a recurrent feature reasoning module and a knowledge consistent attention (KCA) module, enhancing the semantic information of the repair results.

However, these deep learning image inpainting algorithms do not consider the prior information such as structure and texture, which often leads to problems such as structural disorder and loss of texture details. Therefore, enhancing the ability to perceptualize the features of the structure and texture is of great significance for the results of image restoration. Li et al.[9] proposed a progressive image reconstruction method based on visual structure. It integrates the information about the structure information into the features of the image. However, the information about the texture is not constrained. Liao et al.[10] proposed an image inpainting algorithm guided by coherence priors of semantics and textures. It introduces semantic and texture to complete image restoration, but the structural information is not considered, resulting in the structural disorder of the repair results. Nazeri et al.[11] proposed the EdgeConnect image restoration model, which used edge images to guide the restoration and achieved good restoration results. However, the method ignores the importance of texture information, resulting in texture blur in the restoration results. Liu et al.[12] proposed an image inpainting method via a mutual encoder-decoder with feature equalizations. The joint constraint of structure texture is not considered in the repair process, and the repair results are prone to be inconsistent as a whole. Yang et al.[13] used image structure knowledge to repair, but did not pay attention to texture information. Kim et al.[14] applied two discriminators work in a complementary manner to learn both facial features and texture details. However, the attention to structural information is ignored.

To sum up, the existing deep learning image restoration methods lack the joint constraint guidance of the prior information of structure and texture on the restoration process when repairing murals, which leads to problems such as structural disorder and texture blur in the restoration results of murals. A deep learning model for mural restoration was proposed based on structure texture hybrid enhancement. Firstly, the structure feature was extracted by dynamic convolution, and the structure guided branch was used to guide the encoder to encode the mural image, so as to obtain rich edge contour information. Then, a multi-granularity feature extraction module was designed to obtain the multi-scale features of the texture image, which could be used as the guide information of the decoder to guide the generation of mural images, and improved the structural consistency and texture rationality of the repair results. Finally, skip connection was used to promote the feature fusion and sharing between encoder and decoder, and the damaged murals were repaired by spectral-normalized PatchGAN (SN-PatchGAN) discriminator. The restoration experiments of the real damaged Dunhuang murals showed that the proposed algorithm was superior to the comparison algorithms in both subjective and objective evaluation.

1 Generative adversarial networks

The generated adversarial networks (GAN) is shown in Fig. 1. It is mainly composed of a generator and a discriminator. The input of the generator is random noise, which is mainly used to generate data. The discriminator mainly discriminates between the generated data and the real data, and feeds back the judgment results to the generator. In image restoration, through the game between the generator and the discriminator, the ability of the generator to learn the distribution of real image samples and generate image samples is continuously improved, so as to achieve the purpose of learning image restoration. When the discriminator cannot distinguish between the real sample and the generated sample, that is, when the generator and the discriminator reach the game equilibrium, the image restoration task is completed.

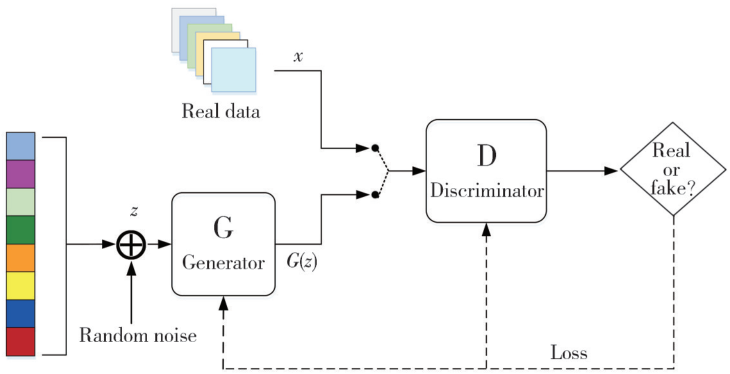

Fig. 1 Basic structural framework of GAN

The adversarial training process of GAN is expressed as

$$\min_G \max_D V(G,D)=E_{x\sim P_{\text{data}}(x)}\Big[\log\big(D(x)\big)\Big]+E_{z\sim P_z(z)}\Big[\log\Big(1-D\big(G(z)\big)\Big)\Big], \tag{1}$$

where z is random noise; x is the real samples; E is the expected value of the distribution function; $P_{\text{data}}(x)$ is the real data distribution; $P_z(z)$ is the random noise distribution. The generative adversarial networks maximizes $\log\left(D(x)\right)$ training discriminator and minimizes $\left[\log\left(1-D\left(G(z)\right)\right)\right]$ training generator, so that they can carry out the maximum and minimum iterative adversarial game until the generator and discriminator reach equilibrium.

2 Proposed method

2.1 Overall network architecture

The overall network architecture is shown in Fig. 2. The network model is based on the generative adversarial networks and consists of generator and discriminator. The generator includes: 1) encoder and structure guided branch; 2) decoder and texture guided branch. The discriminator adopts the spectral-normalized PatchGAN model.

When the model works, firstly, the mural image is encoded by the encoder, and the information of the structure map is extracted by the dynamic convolution cascade method to improve the expression ability of the structure features. The structure features are used as the guide information for the encoder to enhance the structure outline information of the encoded feature map. Then, the multi-granularity feature extraction module is designed to obtain the texture feature map, which is used as the guide information for the reconstruction of the decoder to improve the reconstruction accuracy of the repair results. At the same time, a jump connection is used between the encoder and the decoder to promote the consistency of the structure and texture of the repair results. Finally, the spectral-normalized PatchGAN discriminant model is used to judge whether the output image and the ground truth image are real or fake, so as to complete the restoration of the mural.

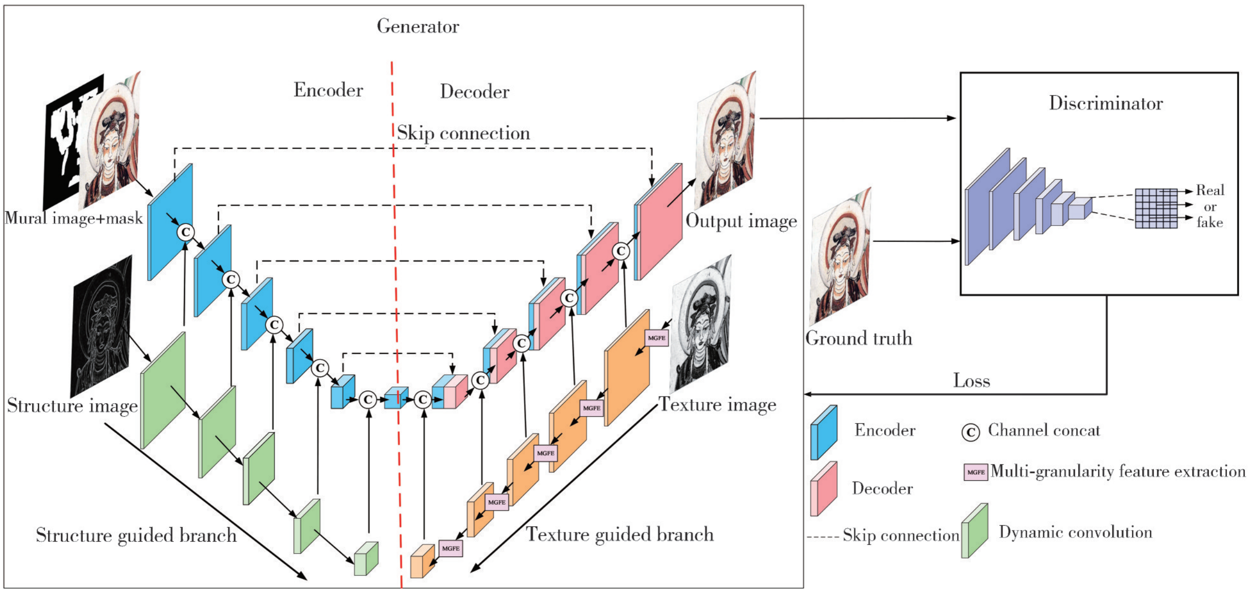

Fig. 2 Overall network architecture of proposed method

2.2 Generator

Due to the lack of prior information about the structure and texture of the damaged murals, the restoration results are prone to the problems of blurred boundaries or broken lines. In the image restoration process, it can obtain better restoration performance by using prior information such as structure or texture to guide image restoration[11,13]. Therefore, in order to enhance the guiding role of structure and texture information in restoration, it was proposed to use structure guidance branch and texture guidance branch to guide the mural features to be encoded and decoded in the generator.

2.2.1 Encoder and structure guided branch

The encoder mainly extracts the basic feature information of mural image through multi-layer convolution operation, which is expressed as

$$F_{\mathrm{E}}^{j}=H_{\downarrow}\left(\sigma\left(W_{\mathrm{E}}^{j}*F_{\mathrm{E}}^{j-1}\right)+b_{\mathrm{E}}^{j}\right), \tag{2}$$

where F_{E}^{j} is the extracted feature map of the jth layer of murals; $H_{\downarrow}(\cdot)$ is the $2\times$ down sampling operation; $*$ is the convolution operation; $\sigma(\cdot)$ is the ReLU activation function; W_{E}^{j} and b_{E}^{j} are the weights and bias of convolution, respectively.

In order to better extract the structural information of mural images, we use dynamic convolution to replace ordinary convolution in the structure guidance branch, and set multiple parallel convolution kernels into a dynamic kernel to improve the expression ability of features, realize the extraction of network depth features, and avoid the incomplete expression of single feature information[15]. The structural diagram of dynamic convolution is shown in Fig. 3, which mainly includes attention weight calculation and dynamic weight fusion.

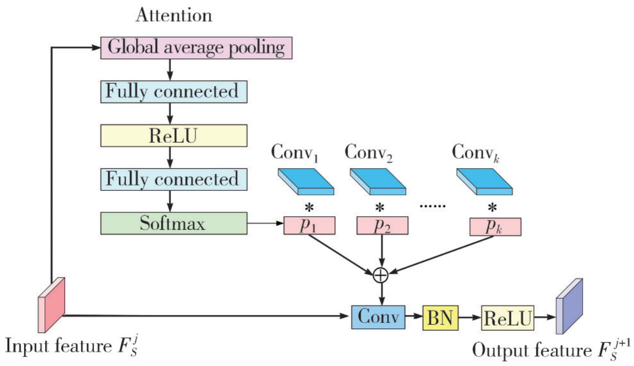

Fig. 3 Dynamic convolution

Firstly, global average pooling (GAP) is used to obtain the global information of the structural feature map, so as to overcome the problem that the ordinary convolution only works on the local receptive field and cannot use the global information of the feature. Then, two fully connected (FC) and ReLU activation functions are used for feature mapping, and the weight of each convolution kernel is dynamically adjusted according to the input information. Finally, the attention weights are obtained by Softmax normalization operation and are assigned to K convolution cores in this layer. The weights of each convolution core are linearly weighted and fused to obtain dynamic convolution. Dynamic convolution changes from a fixed convolution kernel to a convolution kernel dynamically adjusted according to the input, which significantly improves the ability of feature expression.

Feature extraction of structure image through dynamic convolution can be expressed by

$$F_{\mathrm{S}}^{j+1}=\sigma(\tilde{W}(F_{\mathrm{S}}^{j})F_{\mathrm{S}}^{j}+\tilde{b}(F_{\mathrm{S}}^{j})), \tag{3}$$

$$\tilde{W}(F_{\mathrm{S}}^{j})=\sum_{k=1}^{K}p_k(F_{\mathrm{S}}^{j})\tilde{W}_k, \tag{4}$$

$$\tilde{b}(F_{\mathrm{S}}^{j})=\sum_{k=1}^{K}p_k(F_{\mathrm{S}}^{j})\tilde{b}_k, \tag{5}$$

where F_{S}^{j} is the jth layer of the structure feature map; F_{S}^{j+1} is the $(j+1)$th layer of the structure feature map; $\tilde{W}(F_{\mathrm{S}}^{j})$ and $\tilde{b}(F_{\mathrm{S}}^{j})$ are the weights and bias of the dynamic convolution, respectively; p_k is the weight of the kth convolution kernel, which satisfies the constraint condition of $0\leqslant p_k(F_{\mathrm{S}}^{j})\leqslant 1$ and $\sum_{k=1}^{K}p_k(F_{\mathrm{S}}^{j})=1$.

After obtaining the mural feature map F_E^j and the structural feature map F_{S}^{j} of the ith layer, they are fused through channel concat, and the structural information is used to guide the mural features for coding, which is expressed as

$$F_{\mathrm{ES}}^{j}=\mathrm{Concat}(F_{\mathrm{E}}^{j},F_{\mathrm{S}}^{j}), \tag{6}$$

where F_{ES}^{j} is the jth layer coding feature map after fusion; Concat(•) is the channel concat operation.

2.2.2 Decoder and texture guided branch

After the structure guided branch guides the mural image to complete encoding, the encoded feature map is then input into the decoder. When designing the decoder, the structure and texture information of the image are interrelated to form the content of the image. If joint constraint guidance of the outline of the structure and the details of the texture of the image is not taken into account, the repair results are susceptible to problems such as content blur and structure distortion[12]. Therefore, in order to improve the accuracy and overall consistency of the restoration results, we propose a texture guided branch to guide the decoder to reconstruct and restore the image.

Input the coding feature map into the decoder, and obtain the decoding feature through the 2× up sampling operation. The calculation is expressed as

$$F_{\mathrm{D}}^{i}=H^{\uparrow}(\sigma(W_{\mathrm{D}}^{i}*F_{\mathrm{D}}^{i-1})+b_{\mathrm{D}}^{i}), \tag{7}$$

where F_{D}^{i} is the ith decoding feature map; $H^{\uparrow}(\cdot)$ is the 2× up sampling operation; $\sigma(\cdot)$ is the ReLU activation function; W_{D}^{i} and b_{D}^{i} are the weights and bias of convolution, respectively.

Texture image contains rich texture details. In order to obtain more fine-grained features of texture image and increase the range of receptive fields of each network layer, we used the Res2Net multi-granularity feature extraction module to group the feature channels, so as to achieve the purpose of feature extraction for different scales and different granularity of texture image. At the same time, the residual mapping was realized by introducing skip connection to promote the back propagation of gradient[16]. Res2Net multi-granularity feature extraction module is shown in Fig.4.

Firstly, the input feature map is divided into four subsets by the number of channels using the convolution operation of 1×1, and each feature has the same scale. Then, the corresponding output is obtained through the convolution kernel of size 3×3. The process of obtaining y_i is expressed as

$$y_i=\begin{cases}x_i, & i=1,\\ \mathrm{Conv}_{3\times 3}(x_i), & i=2,\\ \mathrm{Conv}_{3\times 3}(x_i+y_{i-1}), & i=3,4,\end{cases} \tag{8}$$

where $\mathrm{Conv}_{3\times 3}(\cdot)$ is the convolution operation with convolution kernel size of 3×3.

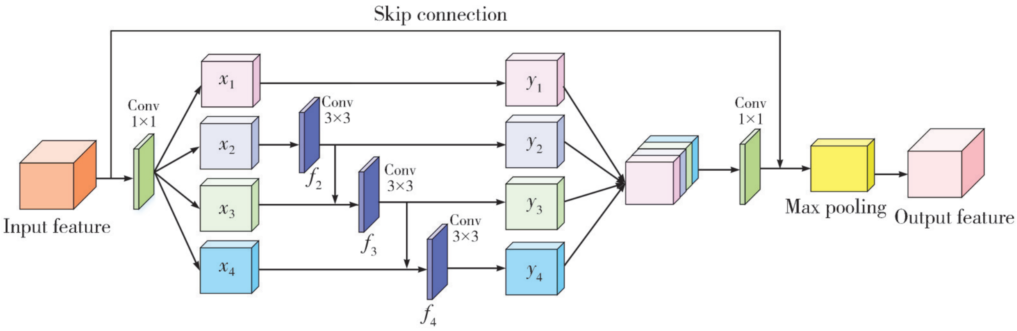

Fig. 4 multi-granularity feature extraction

Finally, the convolution operation of 1×1 is used to reduce the channel dimension of channel cascade y_i, and the extracted texture feature map is obtained by 2× down sampling through max pooling operation, which is expressed as

$$F_{\mathrm{T}}^{i}=H_{\downarrow}(\mathrm{Conv}_{1\times 1}(\mathrm{Concat}(y_i))), \tag{9}$$

where F_{T}^{i} is the extracted feature map of the ith layer of texture image; $\mathrm{Conv}_{1\times 1}(\bullet)$ is the convolution operation with convolution kernel size of 1×1; $\mathrm{Concat}(\bullet)$ is the channel concat operation; $H_{\downarrow}(\bullet)$ is the 2× down sampling through max pooling operation.

After obtaining the decoding feature map F_{D}^{i} and the texture feature map F_{T}^{i} of the ith layer, they are fused through channel concat, and the texture information is used to guide the mural reconstruction, which is expressed as

$$F_{\mathrm{DT}}^{i}=\mathrm{Concat}(F_{\mathrm{D}}^{i},F_{\mathrm{T}}^{i}), \tag{10}$$

where F_{DT}^{i} is the ith layer decoded feature map after fusion; $\mathrm{Concat}(\bullet)$ is the channel concat operation.

2.3 Discriminator

When designing the discriminator, in order to enhance the local semantic consistency and structural continuity of mural restoration results, spectral-normalized PatchGAN (SN-PatchGAN) discriminator was adopted[17,18].

The SN-PatchGAN discriminator is shown in Fig.5. The PatchGAN discriminator is composed of 6 convolution kernels with the size of 5×5 and the step size of 2. The feature of the input image is extracted through multi-level down sampling, and then the discriminator is used to judge whether each image block is real or fake. The output result is a two-dimensional matrix, in which each lattice point of the matrix corresponds to the receptive field of the image block in the original image, and the response mean value of all lattice points is taken as the real or fake probability of the generated image.

Since the SN-PatchGAN discriminator is used to extract features and judge whether each image block is real or fake, it can enhance the local texture details of the image and better distinguish the features of the generated image at different positions and semantic contents. At the same time, normalizing the weight gradient spectrum of PatchGAN discriminator can avoid the problem of overfitting in the network and make the training of the network more stable.

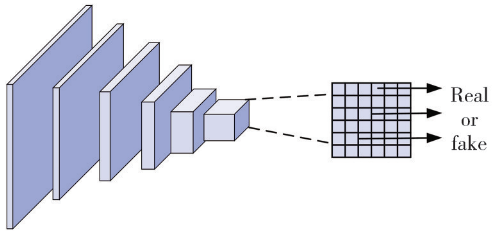

Fig. 5 SN-PatchGAN discriminator

2.4 Loss function

The joint loss of reconstruction loss, adversarial loss, and perceptual loss were used to train the model, so as to generate mural images with coherent structure and reasonable semantics[19].

Reconstruction loss L_{rec} is used to measure the pixel difference between the generated mural image and the ground truth image, expressed as

$$L_{\mathrm{rec}}=\left\| I^{\mathrm{out}}-I^{\mathrm{gt}}\right\|_1, \tag{11}$$

where I^{out} is the generated mural image; I^{gt} is the ground truth image; $\|\bullet\|_1$ is the l_1-norm.

Adversarial loss L_{adv} is used to make the generated image more realistic and natural. Hinge loss is used for network optimization adversarial training, which is expressed as

$$L_{\mathrm{adv}}=E_{I^{\mathrm{gt}}}[\sigma(D^{\mathrm{sn}}(I^{\mathrm{gt}}))]+E_{I^{\mathrm{out}}}[\sigma(1-D^{\mathrm{sn}}(G(I^{\mathrm{in}})))], \tag{12}$$

where D^{sn} is the SN-PatchGAN discriminator; I^{in} is the input image with holes; $\sigma(\bullet)$ is the ReLU activation function.

Perceptual loss is used to further improve the structural consistency of the generated image, which can be expressed as

$$L_{\mathrm{prec}}=E\left[\sum_i\left\|\phi_i(I_{\mathrm{out}})-\phi_i(I_{\mathrm{gt}})\right\|_1\right], \tag{13}$$

where $\phi_i(\cdot)$ denotes the activation map of the ith pooling layer from VGG-16 given the input image I^{in}.

In summary, the joint loss can be written as

$$L = \lambda_{rec} L_{rec} + \lambda_{adv} L_{adv} + \lambda_{prec} L_{prec}, \tag{14}$$

where λ_{rec}, λ_{adv}, and λ_{prec} are the corresponding weights of reconstruction loss, adversarial loss, and perceptual loss, respectively.

3 Experiments

In order to verify the effectiveness of the proposed method based on the self-made Dunhuang mural data set, the restoration experiments of Dunhuang murals with random damage, damage central in large area, and real damage were carried out and compared with other methods. The experimental hardware was configured with Intel (R) Core i7-10700k CPU, 32GB RAM and NVIDIA GeForce RTX 2060 SUPER. The comparison experimental environment was the same. Peak signal to noise ratio (PSNR) and structural similarity index measurement (SSIM) were used for quantitative comparison.

3.1 Experimental results of artificial addition of randomly damaged murals

Five Dunhuang murals were selected for the restoration experiment by adding random damage artificially. The results are shown in Fig. 6. Fig. 6(c) shows that the repair results of Ref.[8] have structural disorder and block effect, such as the structural disorder of the first mural and the fifth mural, the incomplete repair of facial features of the second and third murals, and the block effect of the fourth mural. Fig.6(d) shows that the repair results of Ref.[12] have obvious repair traces and broken lines. Fig. 6(e) shows the repair results of the proposed algorithm. Compared with the comparison methods, the proposed algorithm has the best repair results, with a higher degree of line fitting and clearer and natural texture details.

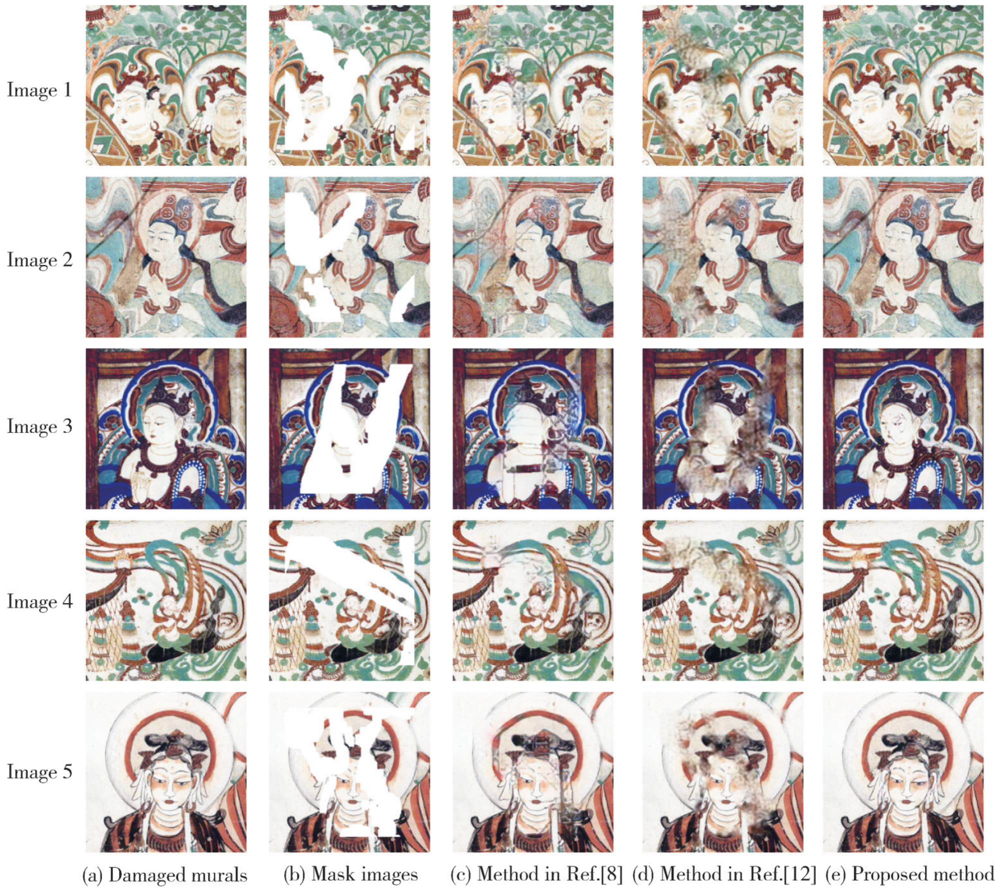

Fig. 6 Comparison of restoration results of artificially added random damaged murals

In order to further evaluate the repair results in Fig.6, PSNR and ssim objective quantitative indicators are used to compare the repair results[20,21]. The larger the PSNR value, the smaller the distortion of the repair results. The larger the SSIM value, the more consistent the restoration result with the original mural structure. The comparison results are shown in Table 1. It can be found that the PSNR and SSIM of the proposed algorithm are higher than those of the comparison algorithms, which indicates that the proposed algorithm has better repair performance.

Table 1 Analysis results of artificial random damage added by different algorithms

Image	Method in Ref.[8]		Method in Ref.[12]		Proposed method	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
1	21.141 5	0.807 7	22.153 4	0.818 1	25.556 5	0.874 4
2	23.401 7	0.804 6	23.936 1	0.811 3	27.898 3	0.893 9
3	18.169 6	0.749 7	18.708 3	0.737 3	25.350 3	0.872 1
4	20.744 9	0.827 2	21.327 4	0.817 2	26.397 4	0.900 8
5	20.424 5	0.770 2	20.667 8	0.775 6	26.316 5	0.889 3

3.2 Experimental results of repairing central damaged murals in large area

In order to further verify the repair effect of large-area damaged murals, five Dunhuang murals were selected for repair experiments. The results are shown in Fig.7. Fig.7(c) shows that the repair results of Ref.[8] have a repair block effect, and the main objects of the mural images have not been repaired. Fig.7(d) shows that the repair results in Ref.[12] have serious repair traces and ambiguous contents, and it is impossible to complete the repair of a large area of damaged murals in the center. Fig.7(e) shows that the repair results of the proposed method are better and more complete than those of the comparison algorithms.

In addition, the repair results in Fig. 7 are also evaluated objectively and quantitatively, as shown in Table 2. It can be found that the PSNR and SSIM of the proposed algorithm are higher than those of the comparison algorithms, that is, the proposed method can complete the repair of damaged murals in large areas and achieve good repair performance.

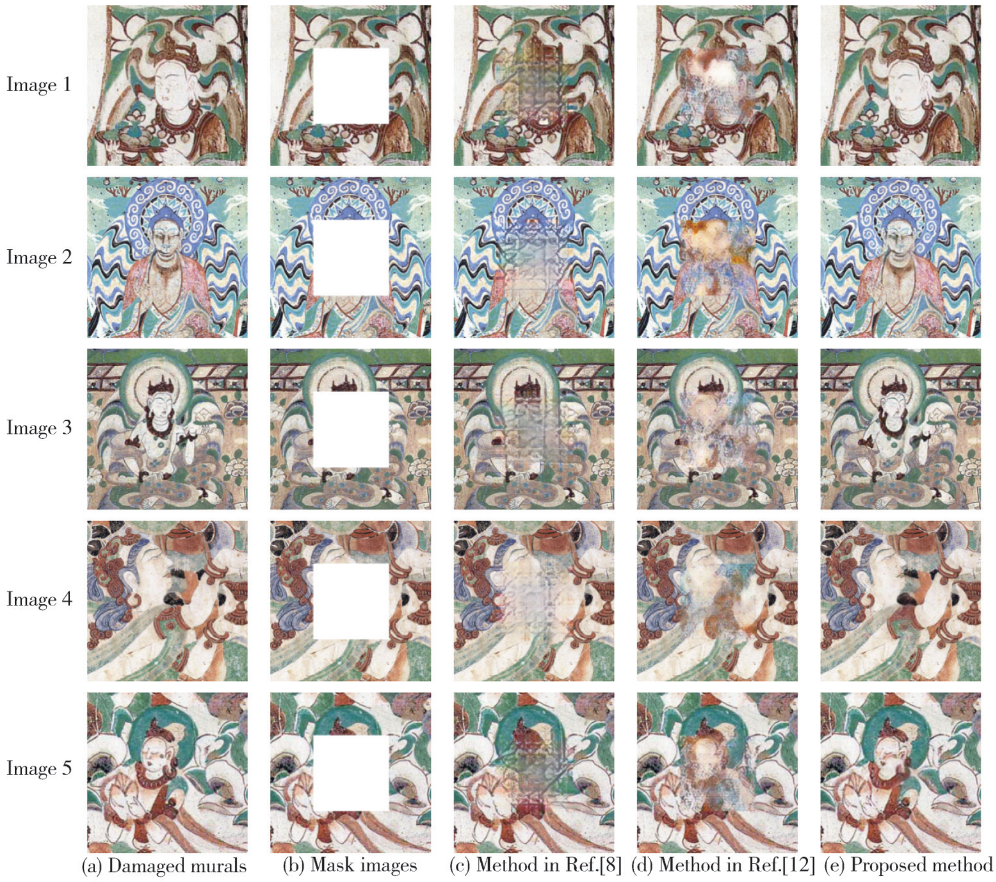

Fig. 7 Comparison of restoration results of damaged murals in center of artificially added areas

Table 2 Analysis results of artificially added large area center damage by different algorithms

Image	Method in Ref.[8]		Method in Ref.[12]		Proposed method	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
1	18.070 7	0.782 4	21.593 3	0.809 1	33.466 1	0.961 8
2	20.092 1	0.771 6	21.417 5	0.796 0	27.233 5	0.924 4
3	19.350 1	0.779 1	20.380 5	0.796 8	26.762 0	0.914 9
4	20.578 5	0.794 5	23.570 1	0.809 4	28.985 1	0.889 1
5	22.472 2	0.812 5	19.347 8	0.784 6	27.908 2	0.911 1

3.3 Experimental results of real damaged murals

In order to further illustrate the effectiveness of the proposed algorithm, five real damaged murals are used for repair experiments, as shown in Fig.8. As for the repair results of the first mural, it can be seen that there are repair errors in Ref.[8], and there are a large number of repair residues in Ref.[12], and the line fitting of the proposed algorithm is harmonious and natural. The repair results of the second mural can be seen that the repair results of Refs.[8] and [12] have structural disorder, and the texture consistency of the proposed algorithm is good, which makes the repair results more consistent with the human visual effect. From the repair results of the third mural, it can be seen that the headdress areas in Ref. [8] and Ref.[12] have structural repair discontinuities, and the repair results of the proposed algorithm have good structural continuity. From the repair results of the fourth mural, it can be seen that neither Ref.[8] nor Ref.[12] has completed the repair of the damaged area, and the proposed algorithm has achieved good structural texture consistency. From the restoration results of the fifth mural, it can be seen that Ref.[8] and Ref.[12] have obvious traces of restoration, and the lines have not been fitted, but the proposed algorithm fits the texture information well.

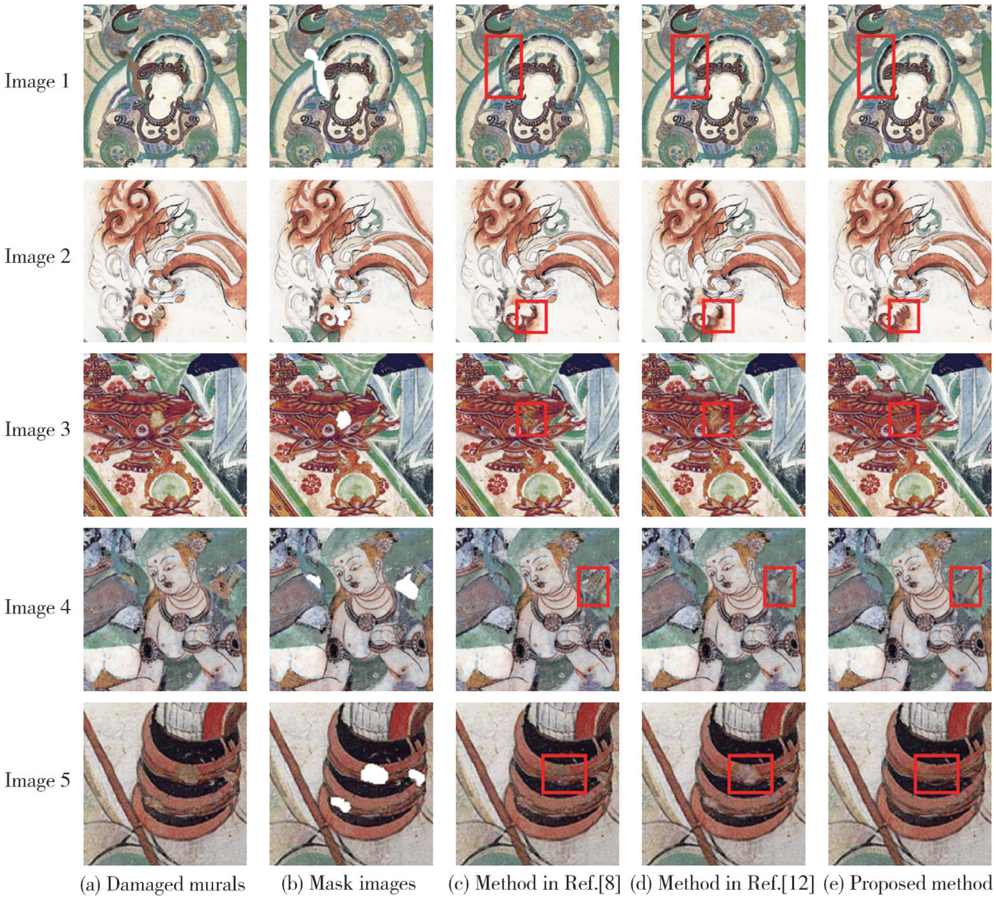

Fig. 8 Comparison of restoration results of real damaged murals

4 Conclusions

A mural inpainting generative adversarial networks model based on structural and texture hybrid enhancement was proposed. In the structure guided branch, the structure features were extracted by cascading dynamic convolution to guide coding, so as to improve the structure profile information in the coded feature map. Then a multi-granularity feature extraction module was designed to obtain the texture features of texture guided branches, guide the decoder to better complete the generation and reconstruction of murals, and use skip connection to enhance the sharing and complementarity of structure and texture information, so as to promote the structure and texture of mural images to be more consistent. The experimental results of Dunhuang murals showed that the proposed method had completed the repair of damaged murals well, and it was superior to the comparison algorithms in both subjective and objective evaluation.

Acknowledgement

This work was supported by Ministry of Education in China Project of Humanities and Social Sciences (No. 19YJC760012); Star of Innovation Project for Outstanding Graduate Students in Gansu Province (No.2022CXZX-546)

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] WANG H, LI Q Q, JIA S. A global and local feature weighted method for ancient murals inpainting. International Journal of Machine Learning and Cybernetics, 2020, 11(6): 1197-1216.

[2] SRIDEVI G, SRINIVAS KUMAR S. Image inpainting based on fractional-order nonlinear diffusion for image reconstruction. Circuits, Systems, and Signal Processing, 2019, 38(8): 3802-3817.

[3] CHEN Y T, ZHANG H P, LIU L W, et al. Research on image inpainting algorithm of improved total variation minimization method. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(5): 5555-5564.

[4] ZHA Z Y, YUAN X, WEN B H, et al. Image restoration using joint patch-group-based sparse representation. IEEE Transactions on Image Processing, 2020, 29: 7735-7750.

[5] QUAN W Z, ZHANG R S, ZHANG Y, et al. Image inpainting with local and global refinement. IEEE Transactions on Image Processing, 2022, 31: 2405-2420.

[6] LAHIRI A, JAIN A K, AGRAWAL S, et al. Prior guided GAN based semantic inpainting//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 13693-13702.

[7] WANG N, ZHANG Y P, ZHANG L F. Dynamic selection network for image inpainting. IEEE Transactions on Image Processing, 2021, 30: 1784-1798.

[8] LI J Y, WANG N, ZHANG L F, et al. Recurrent feature reasoning for image inpainting//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 7757-7765.

[9] LI J Y, HE F X, ZHANG L F, et al. Progressive reconstruction of visual structure for image inpainting//2019 IEEE/CVF International Conference on Computer Vision, October 27-November 2, 2019, Seoul, Korea. New York: IEEE, 2019: 5961-5970.

[10] LIAO L, XIAO J, WANG Z, et al. Image inpainting guided by coherence priors of semantics and textures//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 20-25, 2021, Nashville, TN, USA. New York: IEEE, 2021: 6535-6544.

[11] NAZERI K, NG E, JOSEPH T, et al. EdgeConnect: generative image inpainting with adversarial edge learning. 2019: 1901.00212. https://arxiv.org/abs/1901.00212v3.

[12] LIU H Y, JIANG B, SONG Y B, et al. Rethinking image inpainting *via* a mutual encoder-decoder with feature equalizations//Computer Vision – ECCV 2020: 16th European Conference, August 23 – 28, 2020, Glasgow, UK, Cham: Springer International Publishing, 2020: 725-741.

[13] YANG J, QI Z Q, SHI Y. Learning to incorporate structure knowledge for image inpainting. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12605-12612.

[14] KIM J, YANG J L, TONG X. Learning high-fidelity face texture completion without complete face texture//2021 IEEE/CVF International Conference on Computer Vision, October 10-17, 2021, Montreal, QC, Canada. New York: IEEE, 2021: 13970-13979.

[15] CHEN Y P, DAI X Y, LIU M C, et al. Dynamic convolution: attention over convolution kernels//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 13-19, 2020. Seattle, WA, USA. New York: IEEE, 2020: 11027-11036.

[16] WANG Z S, WU Y Y, WANG J Y, et al. Res2Fusion: infrared and visible image fusion based on dense Res2net and double nonlocal attention models. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 5005012.

[17] YU J H, LIN Z, YANG J M, et al. Free-form image inpainting with gated convolution//2019 IEEE/CVF International Conference on Computer Vision, October 27-November 2, 2019, Seoul, Korea. New York: IEEE, 2019: 4470-4479.

[18] LI H, HU L, ZHANG J. Irregular mask image inpainting based on progressive generative adversarial networks. The Imaging Science Journal, 2023, 71(3): 299-312.

[19] SONG J, YI H, XU W, et al. ESRGAN-DP: Enhanced super-resolution generative adversarial network with adaptive dual perceptual loss. Heliyon, 2023, 9(4): e15134.

[20] Al NAJJAr Y. Comparative analysis of image quality assessment metrics: MSE, PSNR, SSIM and FSIM. International Journal of Science and Research (IJSR), 2024, 13(3): 110-114.

[21] WANG H, SONG H X, ZHANG Z. Two - stream face image restoration algorithm based on texture and structure. Journal of Test and Measurement Technology, 2024, 38(3): 274-280.

结构与纹理混合增强的生成对抗壁画修复算法

陶美风[1], 陈　永[1,2*], 赵梦雪[1], 张娇娇[1]

1. 兰州交通大学 电子与信息工程学院,甘肃 兰州 730070;
2. 甘肃省人工智能与图形图像处理工程研究中心,甘肃 兰州 730070

摘　要: 现阶段的深度学习图像修复方法没有考虑结构和纹理信息的联合引导,导致修复结果易出现结构紊乱和纹理模糊等问题。本文提出了一种结构与纹理混合增强的生成对抗壁画修复算法。首先,构建动态卷积级联的结构引导分支,提高结构特征的表达能力,并利用结构信息指导编码器编码,增强编码特征图的边缘轮廓信息。然后,设计多粒度特征提取模块得到纹理引导分支的纹理特征,利用多尺度纹理信息引导解码器重构修复,提高壁画的纹理一致性。最后,采用跳跃连接促进结构和纹理的特征共享,并利用谱归一化马尔科夫判别器完成壁画修复。对真实敦煌壁画数字化修复实验结果表明,所提方法主客观评价均优于其他对比算法,修复结果较清晰自然。

关键词: 图像处理; 壁画修复; 结构与纹理增强; 动态卷积; 多粒度特征提取

引用格式: TAO Meifeng, CHEN Yong, ZHAO Mengxue, et al. Generative adversarial mural inpainting algorithm based on structural and texture hybrid enhancement. Journal of Measurement Science and Instrumentation, 2025, 16(2): 195-204. DOI: 10.62756/jmsi.1674-8042.2025019