

RESEARCH

Open Access



SeparaFill: Two generators connected mural image restoration based on generative adversarial network with skip connect

Chaohui Lv^{1,2*}, Zilu Li¹, Yinghua Shen¹, Jinghua Li¹ and Jin Zheng¹

Abstract

Mural is an important component of culture and art of Dunhuang in China. Unfortunately, these murals had been ruined or are being ruined by some diseases such as cracking, hollowing, falling off, mildew, dirt, and so on. Existing image restoration algorithms have problems such as incomplete repair and disharmonious texture during large-area repair, so the effect of mural image disease area repair is poor. Due to lack of a standard mural datasets, Dunhuang mural datasets are created in the paper. Meanwhile, our network architecture SeparaFill is proposed which connects two generators based on U-Net. Based on the characteristics of the painting, the contour line pixel area of the mural image is innovatively separated from the content pixel area. Firstly, the contour restoration generator network with skip connect and hierarchical residual blocks is employed to repair contour lines. Then, the color mural image is repaired by the content completion network with guide of the repaired contour. Full resolution branches and generator branches of the U type are exploited in content completion generators. Convolution layers of different kernel sizes are fused to improve the reusability of the underlying features. Finally, global and local discriminant networks are applied to determine whether the repaired mural image is authentic in terms of both the modified and unmodified areas. The proposed SeparaFill shows good performance in restoring the line structure of the damaged mural images and retaining the contour information of the mural images. Compared with existing restoration algorithms in mural real damage repair experiment, our algorithm increases the peak signal-to-noise ratio (PSNR) by an average of 1.1–4.3 dB and the structural similarity (SSIM) values were slightly improved. Experimental results reveal the good performance of the proposed model, which can contribute to digital restorations of ancient murals.

Keywords: SeparaFill, Contour restoration generator, The content completion network, Mural restoration

Introduction

Dunhuang murals retain the authentic works of famous artists over thousands of years, which have profound artistic, historical and cultural value. Due to the long-term impact of natural weathering and human factors, these murals have undergone different kinds of disease, such as cracking, falling off, hollowing, pulverization, fading, color changing, getting mildewed, smudging and

scratches, and so on. Therefore, there is an urgent need to restore the murals combined with the environment and painting materials. Meanwhile, the manual restoration work of Dunhuang murals is arduous and complex. It requires the joint efforts and support of multiple disciplines [1]. The development of image processing and deep learning technology have enabled digital restoration of mural images to become a research hot topic.

The research for digital restoration of mural images are divided into two main categories, including digital image restoration based on traditional algorithms and the restoration method based on deep learning. Cao et al. [2], Jiao et al. [3], and Wang Hong [4] improved the traditional

*Correspondence: llvch@cuc.edu.cn

¹ School of Information and Communication Engineering, Communication University of China, Beijing 100024, China
Full list of author information is available at the end of the article

algorithm to restore the mural image. While Wang et al. [5], Zhang et al. [6], and Wen et al. [7] proposed the improvement algorithms of the Generative Adversarial Networks (GAN) [8] based on the deep learning research. Currently, most of the researches are based on the public datasets of natural images in the field of deep learning images restoration, which can capture the potential statistical laws in natural images through the learning of massive data. Dunhuang murals are meticulous heavy color painting with extremely and complexly structures and obviously different from the natural images restored by self-similarity characteristics, while the structure of mural images is mainly determined by the line drawing. Due to the complex content of murals, it is also difficult to distinguish the foreground and background. So it is very difficult to restore the mural image well using the original GAN networks. As there is lack of public mural image datasets so far, it is essential to establish the mural datasets based on the Dunhuang mural album.

The mural images are mainly subject to damage from a wide variety of causes, resulting in some small scale of mural diseases and some large areas missing. These small missing areas such as cracking, small falling off blocks and pulverization of the mural, tend to appear in clusters and the areas intertwined with the complete areas. Those large missing areas such as scratch and large peeling, account for a significant portion of the information of the mural image, which seriously affects the integrity of the mural. The traditional image restoration algorithms are the ill-posed problems of filling in the missing pixels in digital images by interpolating from the prior information. Rudin et al. [9] proposed the Total Variation (TV), which used the partial differential equation based on the principle of thermal diffusion to carry out anisotropic diffusion to repair the damaged areas. Zhang et al. [10] replaced the integer order differential in the TV model with fractional order and further considered the image texture structure information to improve the accuracy of image restoration. The TV method achieves good effects on small missing areas, but it is prone to the problems of error diffusion and poor visual connectivity. The algorithm based on texture synthesis proposed by Criminisi et al. [11] considered the continuity of texture and can solve the problem of poor visual connectivity of partial differential equations to a certain extent. On the basis of the above algorithm, Guo et al. [12] and Cao et al. [2] improved the confidence of priority and the adaptive selection of template window size to optimize the effect of image restoration and avoided the diffusion phenomenon caused by the restoration.

With the continuous development of deep learning, deep neural networks have shown excellent ability in the prior knowledge learning of massive data [13, 14]. GAN

networks can generate non-existent images through the learning of image features and become a research hotspot in the field of image restoration. Yeh et al. [15] proposed a semantic image restoration model with depth generation mode by adopting the Deep Convolutional GAN (DCGAN) structure and using context-weighted loss to search for the closest encoding of the corrupted image, which has good performance on restoring images with the simple structure. Zhang et al. [16] used a four step incremental generation networks to restore the image under the square mask, but this restoration method cannot deal with the irregular mask. Liu et al. [17] proposed an image restoration method based on partial convolution (Pconv) by using partial convolution with mask instead of full convolution filling and it can repair the irregular holes of images. Guo et al. [18] designed the network generator as a network block with two parallel branches of low resolution and full resolution. Through the superposition of network blocks, the mask area was gradually cleaned and reduce to zero. Yu et al. [19], Nazeri et al. [20] and Zamir et al. [21] adopted two connected generators to restore the image. Yu et al. [22] further proposed gate convolution on the previously model and replaced the ordinary convolution in the generator to make the mask update learnable. Xiong et al. [23] proposed a three-stage image restoration model to distinguish the foreground and background, which was suitable for a single object or architectural image with clear outline and was not operable for images with complex lines and rich content for murals. Ronneberger et al. [24], Li et al. [25], Yang et al. [26], Jo et al. [27] and Liu et al. [28] introduced skip connections into the corresponding convolution layer of down sampling and up sampling based on the U-Net networks model to transfer the feature information extracted by the networks, so as to improve the utilization efficiency of the networks for low-level features and refine the texture of image restoration.

Traditional image repair methods can restore small damage, which is prone to blur in repair details. The image repair methods based on deep learning also have problems such as incomplete repair and disharmonious texture during large-area repair. In view of the above problems arising in the restoration of the murals, the painting characteristics of the Dunhuang murals are analyzed to improve the restoration network, so as to better repair the details of the murals. Two generators connected image restoration networks (SeparaFill) is proposed according to the characteristics of mural painting. The algorithm based on U-Net network [24] restore the contour lines of the image firstly and then restore the color areas inside the contour lines of the image. The improvements achieved by the proposed method are mainly reflected in the following aspects: (1)

skip connections are added to the contour restoration network for feature channel fusion to realize the reuse of low-level features and extract the high-level semantic features of the image by using Hierarchical Residual Networks(Res2Net) [29]; (2) an accumulation feature extraction mechanism is proposed to realize multi-level feature fusion under different resolutions in the content completion restoration network, and self-attention mechanism is introduced to restore image details. (3) Because of the small datasets, the Siamese Network idea of Meta learning is introduced to the discriminator network, and the contrast loss function commonly used in the Siamese Network is incorporated into the discriminator optimization. Compared with other algorithms, this method can achieve better performance. The proposed network architecture *SeparaFill* which connected two generators based on U-Net, separates the image repair work into contour line restoration and color block restoration, so that the image contour structure can be repaired carefully, and can also be better repaired when a large area is damaged.

Methods

It is observed that the murals are dominated by contour lines through the analysis of the painting characteristics of murals. Therefore, its pixels can be divided into contour line parts and color block parts separated by contour line for a mural image. The contour line parts of the mural image are composed of rich thin, narrow and continuous lines. Meanwhile, the contour line color is relatively single, which usually contains the dark colors such as black, brown and red. However, the color block parts contain rich color information. The size and shape depend on the contour line, but the gray scale inside the color block is continuous and the texture information is simple. With the above mural image characteristics, a two generators connected mural image restoration network based on U-Net network architecture is proposed. This method restores the mural image contour and its internal color blocks separately and reduces the difficulty of restoration. After training, the network can obtain the better restoration result compared with other algorithms.

The mural restoration network mainly consists of three parts: contour restoration generator network, content completion restoration generator network, global and local discriminator network.

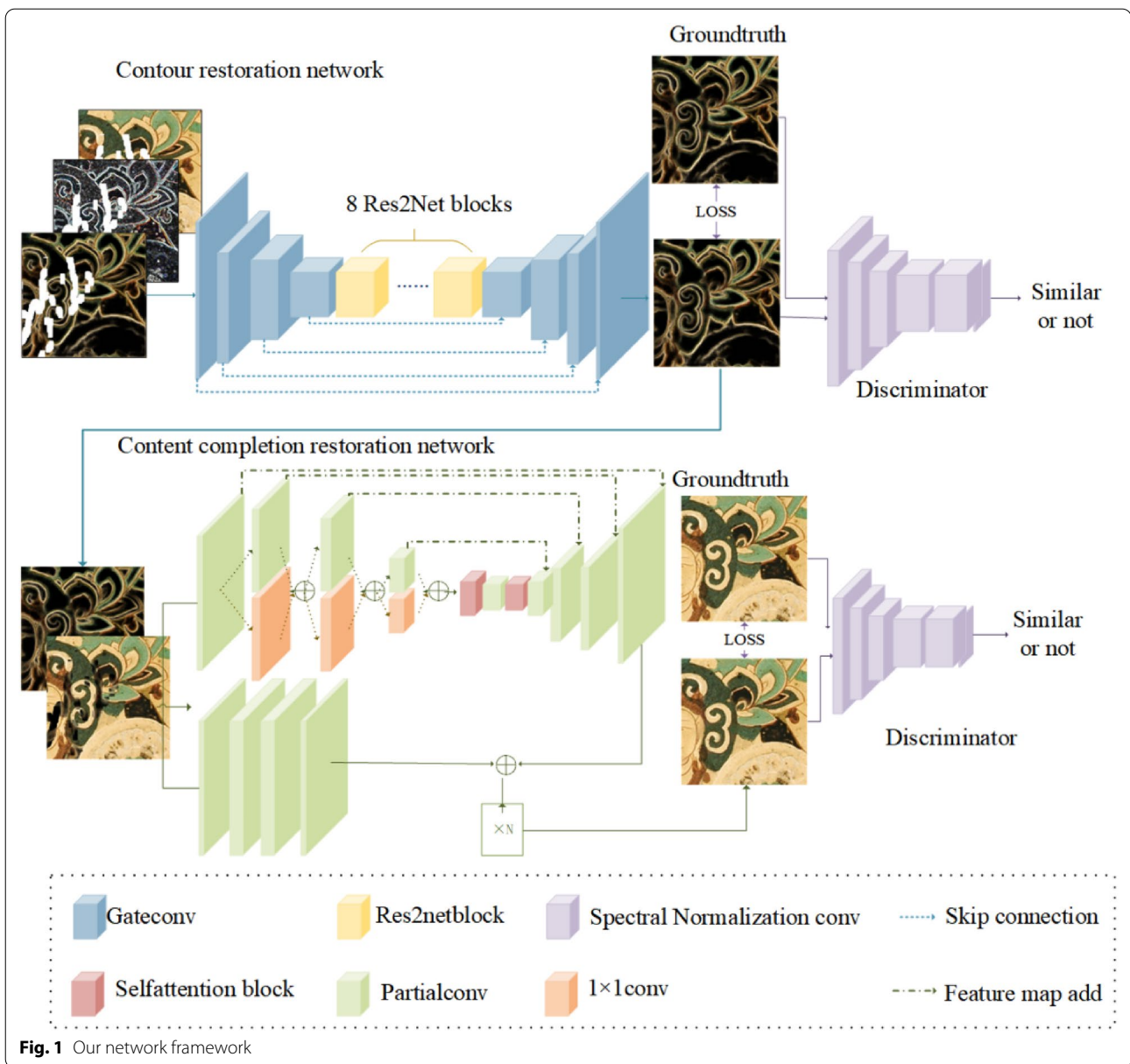
The damaged image is obtained by multiplying the image with the mask, Let I_{gt} be ground truth images, M be mask images, $x_{masked}^{(i)} = I_{gt} \odot M$ be the masked image, $sketchgray^{(i)}$ is obtained by the Holistically-nested Edge Detection (HED) algorithm [30], and the contour damaged image $sketch_{masked}^{(i)} = x_{masked}^{(i)} \odot sketchgray^{(i)} \odot M$.

Feed the damaged contour image and damaged image into the contour restoration generator network, at the same time, Let $sobel^{(i)}$ as the edge image of the damaged image via Sobel edge detection processing. Using the $x_{masked}^{(i)}$, $sketch_{masked}^{(i)}$ and $sobel^{(i)}$ as input of contour restoration generator. Fill the contour recovery map $sketch_g^{(i)}$ into the image's missing area, $x_{masked1}^{(i)} = sketch_g^{(i)} \odot (1 - M) + x_{masked}^{(i)}$, the large missing block is further divided into several small areas. The damaged contour image obtained by multiplying with the mask is input into the contour repair generator network. At the same time, the Sobel edge detection processing is conducted on the damaged image to acquire the edge image, which is input to the generator input end together with the damaged color image to assist the contour restoration. The recovered contour is filled into the image to be repaired to further split the damaged areas of the image. Meanwhile, the contour recovery map is also sent to the content completion restoration generator in the second stage to guide the restoration of the image. The discriminator network of contour restoration is as same as content completion restoration, which is composed of local discriminator and global discriminator. The network framework is shown in Fig. 1.

Contour restoration network

The generator of the contour restoration network is improved on the EdgeConnect model [20]. 9 channels are put into the network, which include the damaged contour image of RGB (red, green, blue) three channels, the damaged image and the edge feature map extracted by Sobel edge detection. The input of the damaged image with the edge image together provide rich information for network and guide the contour restoration. The network consists of a down sampling convolution block contained a 5×5 convolution layer and three 3×3 convolution layers, 8 Res2Net blocks and a up sampling recovery resolution convolution block. Skip connections are added to four up sampling convolution layers and down sampling convolution layers. Skip connections reuse the edge features of lower layers and retain more dimension information. So, the up sampling network of the generator can select between shallow features and deep features to enhance the robustness of the network. For the convolution in the shallow feature extraction network and the up-sampling convolution network, we use gate convolution to replace the ordinary convolution.

The block of Res2Net is used to replace the residual block structure in contour restoration network. Res2Net block modifies the structure of Residual Networks (ResNet) block [31] which is shown in Fig. 2. Firstly, the input features are passed through a layer of 1×1



convolution, further divide the output features equally according to the number of channels, and fuse the segmented features in different channel blocks. The expression formula is as follows:

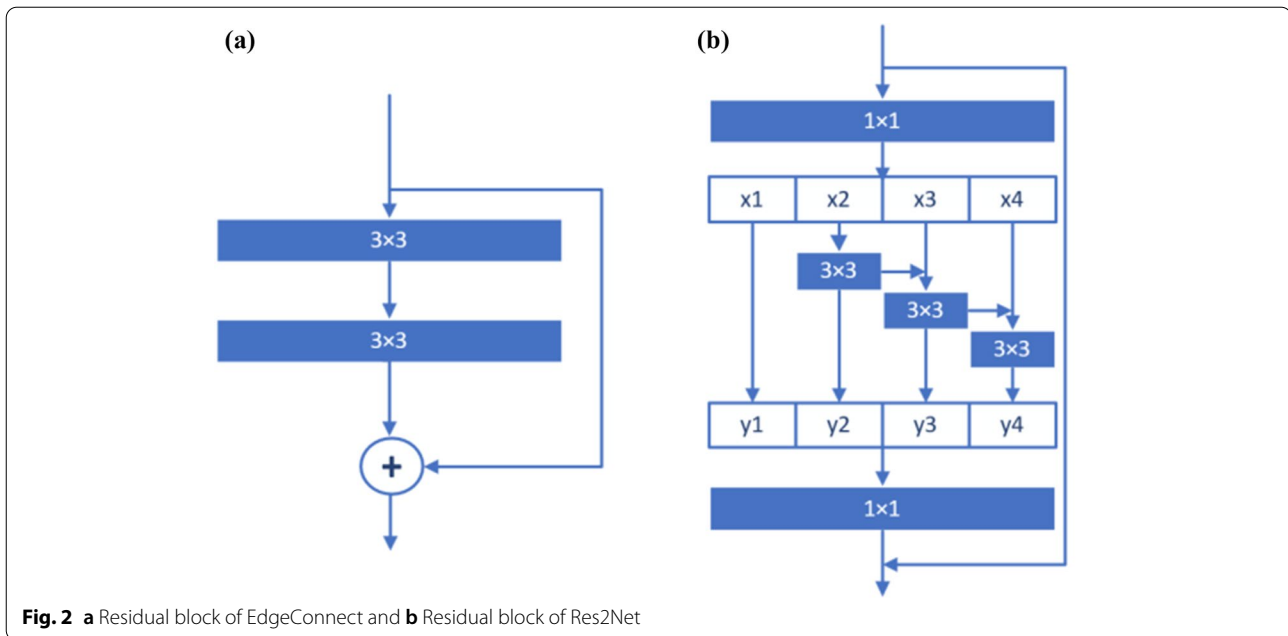
$$y_i = \begin{cases} x_i & i = 1; \\ K_i(x_i + y_{i-1}) & 1 < i \leq s \end{cases} \quad (1)$$

where x_i represents the number of equally divided blocks, s represents the number of equally divided blocks, y represents the output of convolution, and $K_i()$ represents 3×3 convolution. The residual structure of Res2Net retains the function of ResNet to avoid gradient disappearance and gradient explosion, and realizes channel

block multiplexing of the 3×3 convolution layer in the ResNet block. This multi-scale channel fusion of input features can make the ability of feature extraction stronger without add network parameters.

Content completion restoration network

The aim of the second stage is to complete the color blocks between the contour lines. The network inputs are the contour map generated in the first stage and the corresponding damaged image after repairing contour lines. Content completion restoration block include a U-shaped image repair branch network and a convolution branch without down sampling. Through the superposition of multiple modules,



the missing areas can be repaired finely. The U-shaped image repair branch network consists of the 4 down sampling convolution layers that the kernel is 3×3 with 2 step sizes, the feature extraction network with two self-attention blocks, and up sampling layers of 3×3 convolution corresponding to the down sampling layers. After adding feature fusion directly, the feature map of each dimension contains more features. This operation reduces network parameters and memory footprint, thus can provide space for module stacking. Furthermore, an accumulation feature extraction mechanism is proposed, the feature map output through each convolution layer is superimposed with the feature map output through the layer in front, so as to realize multi-level feature fusion under different resolutions. The implementation formula is as follows:

$$y^l = \frac{y^l}{2^l} + \frac{1}{2^l} \sum_{i=1}^l 2^{i-1} y^i, \quad (2)$$

where y^l represents the convolution fusion output of layer l , and y^i represents the fusion output of convolution layer i . It can be seen from the formula that the superposition fusion mechanism makes the features extracted by the fusion convolution layer accumulate the characteristics output of all the convolution layers in front. Since the feature fusion of direct addition requires the same size of the input feature map, the feature map with different resolution is matched by reducing the resolution through 1×1 convolution layer.

As the convolution kernel operates from the local region of the image and represents the local features,

the influence of the global features on the current region becomes very small with the deepening of the convolution network. Self-attention mechanism [32] can capture long-distance dependencies, namely pay attention to the global characteristics so as to enlarge receptive fields of the network. After the feature accumulation layer, the self-attention mechanism is employed to capture the overall features and detail features of the mural image. It can make the generated image more detailed in Fig. 3.

The other branch does not use down sampling in the process of sending and processing the input image information, and keeps the resolution of the original input information, so as to reuse the input information and help refine the texture of the image restoration.

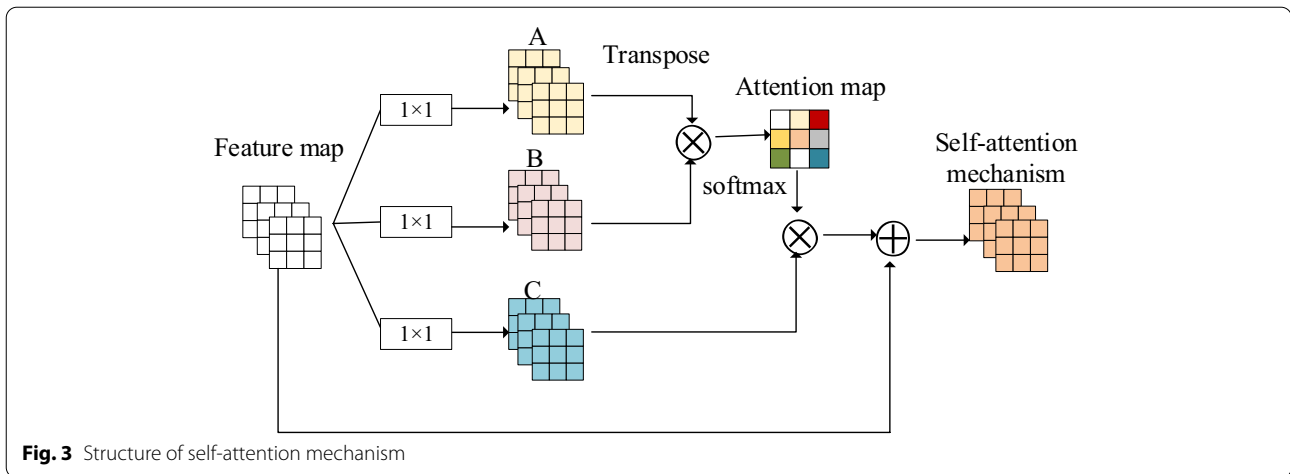
Dilated convolutions are used in the content completion restoration network. The dilation is set as a loop of 1, 2 and 5 to increase the perception domain of the convolution. Since the large damaged area has divided into small pieces after the contour lines have been repaired in the first stage, the difficulty of restoration becomes easier. Therefore, the partial convolutions with fewer parameters are exploited to update the mask and perform detailed restore through the superposition of modules.

Loss function

The loss function in the contour inpainting phase is expressed as:

$$L_{s_G} = \lambda_{adv} L_{adv} + \lambda_{rec} L_{rec} + \lambda_{FM} L_{FM}, \quad (3)$$

where L_{adv} is the adversarial loss based on the discriminator, L_{rec} is the L_1 reconstruction loss, L_{FM} is the feature



matching loss, and λ_{adv} , λ_{rec} and λ_{FM} are the weights of each loss respectively.

GAN obtains the optimal solution by optimizing the value function. The value function is expressed as:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{i \sim P_{out}(i)} [\log (1 - D(G(i)))] \tag{4}$$

where x represents the input data, $P_{data}(x)$ represents the distribution of the real data $P_{out}(i)$ represents the distribution of the image generated by the generator, D represents the discriminator, the probability that the output input is the real data, and G represents the generator, which outputs the generated image. The goal of the discriminator is to maximize the value function.

The reconstruction loss is used to constrain the image pixel level restoration, so as to optimize the detail restore ability of the contour. The reconstruction loss is expressed as follows:

$$L_{rec} = \|I_{recover} - I_{gt}\|_1 * \lambda_{rec} + \|I_{recover} \odot masks - I_{gt} \odot masks\|_1 * \lambda_{rec} \tag{5}$$

where $masks$ is the binary mask image, and \odot is the Hadamard product, used to calculate the global and local reconstruction losses for the generated image and the hole area under mask constraints respectively, λ_{rec} represents the weight value of the loss function.

The feature-matching loss is used to compare the feature maps in the intermediate layers of the discriminator. The feature-matching loss is expressed as follows:

$$L_{FM} = E \left[\sum_{i=1}^L \frac{1}{N_i} \|D_1^{(i)}(I_{gt}) - D_1^{(i)}(I_{recover})\|_1 \right] * \lambda_{FM} \tag{6}$$

where L is the number of convolution layers of the discriminator, N_i is the number of characteristic diagrams of the activation layer of layer i , and $D_1^{(i)}$ is the activation number of layer of the discriminator. λ_{FM} is the regularization parameter.

The content restoration network needs to restore the texture of the image and maintain the semantic consistency between the restored image and the ground truth image. The loss function consists of confrontation loss, reconstruction loss, perceptual loss [33] and structural similarity loss. The loss function is expressed as follows:

$$L_G = \lambda_{adv} L_{adv} + \lambda_{rec} L_{rec} + \lambda_{SSIM} L_{MS-SSIM} + \lambda_{style} L_{style} \tag{7}$$

The loss function and weight of the reconstruction loss L_{rec} and adversarial loss L_{adv} are the same as first part of contour restore. In order to better ensure that the texture and color of the image restoration area fit the original mural, and make the style of the whole restored image consistent, the perceptual loss function is introduced. The perceptual function is divided into content loss and style loss, compares the high-level abstract features through the VGG 19 pre-training model, the formula is as follows:

$$\ell_{feat}^{\varphi, j}(\hat{y}, y) = \frac{1}{C_j H_j W_j} \left\| \varphi_j(\hat{y}) - \varphi_j(y) \right\|_2^2 \tag{8}$$

where C_j , H_j and W_j represent the channel numbers, height and width of the characteristic graph respectively, j represents the j th layer of the network, and φ represents the output after convolution network processing. Content loss let the generated image obtain better visual effect, but large loss weight will produce texture to image

that does not conform to the original image, so it is necessary to reduce the weight of content loss in the later stage of training.

The multi-scale structure similarity loss function [34] is introduced. The combination of structural similarity loss and L1loss can balance the brightness and color of the image, thus making the restored image more detailed. The function expression is as follows:

$$L_{MS-SSIM}(P) = 1 - MS-SSIM(\tilde{p}), \quad (9)$$

where $MS-SSIM(\tilde{p})$ is SSIM calculation for images with different resolutions after scaling, which can obtain better results than simple SSIM loss.

Training and testing procedures

Limited by the small mural image datasets, the parameter of batch size of training will affect the training results. Therefore the parameter of batch size is set as 5, each of which batch has 3000 data, and the parameter of num_workers is set as 16, which is used to preload the batch data of the next iteration into memory.

The specific algorithm steps are as follows in Table 1.

Results and discussion

Data source

In this paper, some clear and well-preserved murals of characters in the Tang Dynasty from the electronic scanning edition of Dunhuang Mural Art (all 10 volumes),

Chinese Dunhuang murals (all 11 volumes) and Dunhuang Grottoes (all 26 volumes) were used. After cleaning the image data, eliminating duplicate mural images, 2175 original data images, consist of 172 images of the early Tang Dynasty, 271 images of the prosperous Tang Dynasty, 859 images of the middle Tang Dynasty, 743 images of the late Tang Dynasty and 130 images of the Five Dynasties were obtained, and all of them were 512×512 in size. A total of 300 images were randomly selected from the images of each dynasty as the test sets, and the remaining 1875 images as training sets. The datasets are expanded by mirror operation, and then divided each original mural image into four small sub-images (256×256) by horizontal and vertical segmentation. After that, a training dataset with 15,000 images and a test dataset with 2400 images were obtained and the size of each image was $256 \times 256 \times 3$, which contained rich mural feature elements such as the clothing, texture, face and decoration. The same number of images was randomly selected from the mask datasets to correspond to the mural images one by one. The size of each mask image was $256 \times 256 \times 1$. Mask datasets imitated irregular damages such as cracking, falling off, pulverization, getting mildewed, smudging and scratches in mural diseases. The damaged RGB image was acquired by multiplying the mural image and the mask image.

Experimental environment

To verify the effectiveness of the proposed method, tests on mural image restoration were carried out. The

Table 1 Pseudocode of algorithm

Initialize the parameters of generator G, discriminators D1 and D2, iterations = n, batch_size = m, and the super parameter k = 1 representing the steps of discriminator

for number of training iterations do

for k steps do

- The m real images $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$ were taken randomly from $P_{data}(x)$
- The m mask maps $\{M_1, M_2, \dots, M_m\}$ were taken in M_{mask} paired with the images
- via the HED networks to obtain contour $sketchgray^{(i)}$ from $x^{(i)}$
- $sketch^{(i)} = x^{(i)} \odot sketchgray^{(i)}$
- $x_{masked}^{(i)} = x^{(i)} \odot M$, $sketch_{masked}^{(i)} = sketch^{(i)} \odot M$
- $sketch_g^{(i)} = sketch^{(i)} \odot M + G_1(sketch^{(i)}, M) \odot (1 - M)$
- $x_{masked1}^{(i)} = sketch_g^{(i)} \odot (1 - M) + x_{masked}^{(i)}$
- Image inpainting $x_g^{(i)} = x_{masked1}^{(i)} \odot M + G_2(x_{masked1}^{(i)}, M) \odot (1 - M)$
- Calculate the restoration loss L_{dS}, L_d based on the input and restore samples
- Update the parameters of discriminators D1 and D2 through Adam optimization algorithm

end for

■ Get m samples randomly from $P_{data}(x)$ and M_{mask} , constructing damaged images $x_{masked}^{(i)}$, $sketch_{masked}^{(i)}$, calculate the loss L_{dS}, L_d based on the input samples and restored samples generated by the network

- Update the parameters of discriminators D1 and D2 through Adam optimization algorithm

end for

hardware environment in this experiment mainly consists of an Intel Xeon e5-2620 V4 @2.1GHZ with 128 GB memory and four Nvidia GeForce GTX 1080 Ti graphics cards with 11 GB memory. The software environment includes the JetBrains Pycharm compiler, running on a Windows 10 system. The software was written in python 3.8, and Pytorch was used as the framework for complete mural image restoration.

Restoration of the randomly damaged murals

In order to verify the effectiveness of the mural image restoration model proposed in this paper, our network model are compared with Pconv [17], EdgeConnect [20], FRRN [18], RN [35] and RFR [36] networks on the test image datasets established in the paper. The mask images were randomly selected from the test mask data set to cover the test images, and the mural images of artificial damage were obtained. These images were then sent to the trained network to obtain the prediction image. For the image restoration results of different network models, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used as the quantitative evaluation indices. The objective comparison of repair results is shown in Tables 2, 3. The running environment of all networks above is the same as our algorithm.

As can be seen from Fig. 4, the restoration part of the algorithms in RN [35], EdgeConnect [20] and RFR [36] is relatively fuzzy. When the damaged area contains structural information, the original contour of the image cannot be well restored with algorithms Pconv [17] and FRRN [18]. In this case, our network can not only obtain the prediction map with high PSNR and SSIM values,

but also express the visual connectivity and structural consistency of the image well. With the increase of the degree of damage, the PSNR and SSIM of all 5 compared algorithms decrease significantly, problems such as incomplete inpainting, and image pixel error diffusion have appeared. However our algorithm can still maintain the contour integrity and continuity of mural image, such as ribbon, sphere, hat wing, shoulder of the characters in the painting, and restore its complete structure with less effective information of the image.

Restoration of the authentic damaged murals

For the real damaged Dunhuang mural images, it is need to label and mark the real damaged area artificially, and the ground truth for these damaged historical relics disappeared many years ago. Therefore, 181 images with disease were selected from the test set images and labeled the disease areas manually. These images are used to test the repair effect of the algorithm on real disease images. Our method is compared with the algorithm in Refs. [20] and [18]. The objective evaluation value of all the test image repair results was averaged, and the results are shown in Figs. 5 and 6. It can be seen that the peak signal-to-noise ratio (PNSR) of our proposed algorithm are 4.3 dB and 1.1 dB higher than the EdgeConnect and FRRN algorithms respectively, and the structural similarity are 3.4% and 0.9% higher than the above algorithm respectively.

To further analyze the algorithm’s availability, some samples were selected for sample-by-sample analysis.

Figure 7 gives five samples with small damage areas. The objective comparison of repair results is shown in

Table 2 Comparison of SSIM value of each algorithm

Sample	Pconv	EdgeConnect	FRRN	RN	RFR	SeparaFill (ours)
1	0.9556	0.9187	0.9769	0.9322	0.9354	0.9788
2	0.9347	0.9113	0.9699	0.9434	0.9483	0.9677
3	0.7801	0.7904	0.9310	0.8774	0.8453	0.9234
4	0.7982	0.7509	0.8999	0.8108	0.8355	0.8826
5	0.8166	0.8135	0.9118	0.8291	0.8623	0.9067

Table 3 Comparison of PSNR value of each algorithm

Sample	Pconv	EdgeConnect	FRRN	RN	RFR	SeparaFill pl (ours)
1	31.73	24.86	33.78	26.92	24.57	33.86
2	31.54	26.16	33.69	30.57	30.45	34.25
3	30.54	24.58	32.71	28.66	27.52	33.36
4	22.36	19.69	25.96	22.39	22.86	25.33
5	24.22	23.15	27.48	22.67	24.66	27.02

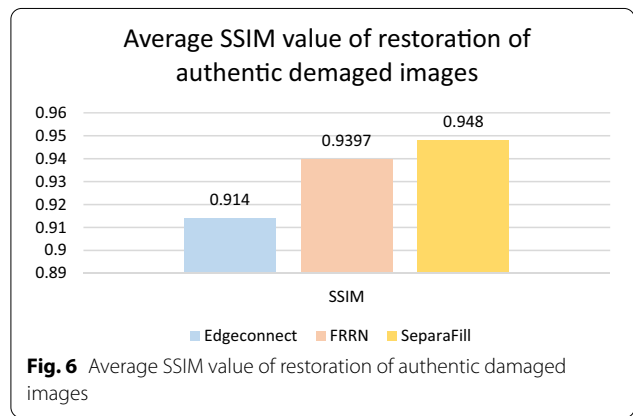
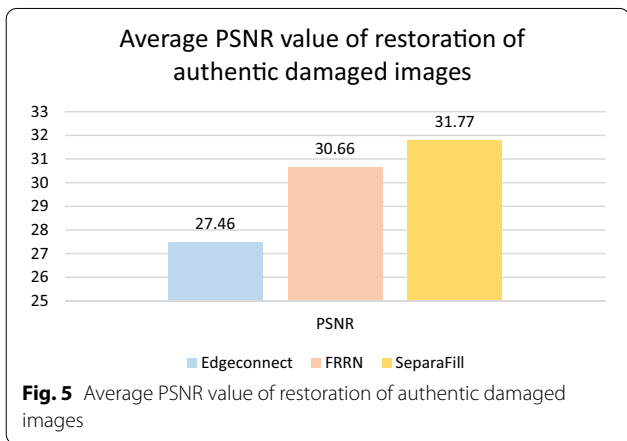
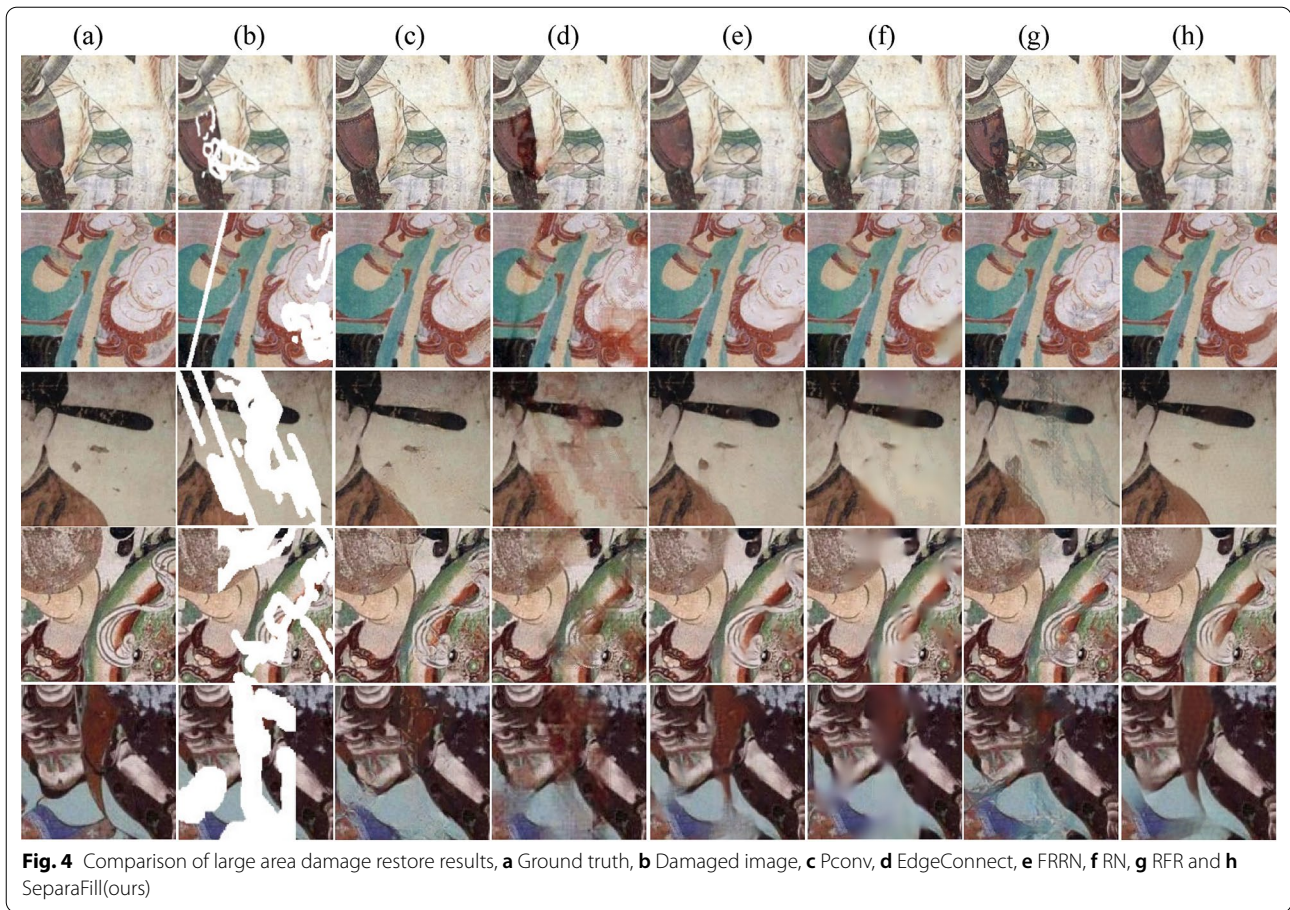


Table 4. It can be seen that the proposed method has significant advantages in objective evaluation in PSNR and SSIM indexes. For example, sample 2 is a figure’s neck image, the PSNR and SSIM values of our algorithm are 34.53 and 0.9714 respectively, while the EdgeConnect algorithm are 30.86 and 0.9540, FRRN algorithm are 33.72 and 0.9678 respectively.

As can be seen from Fig. 7, the predicted image quality of our algorithm is high if damaged area is small. The proposed algorithm has a good and stable restoration effect on distributed discrete damages, small pulverization, scratches, cracks and other diseases. Our method can keep the fill area consistent with the style and texture of the original image, and effectively remove the impact of this kind of disease on the mural image. The contour

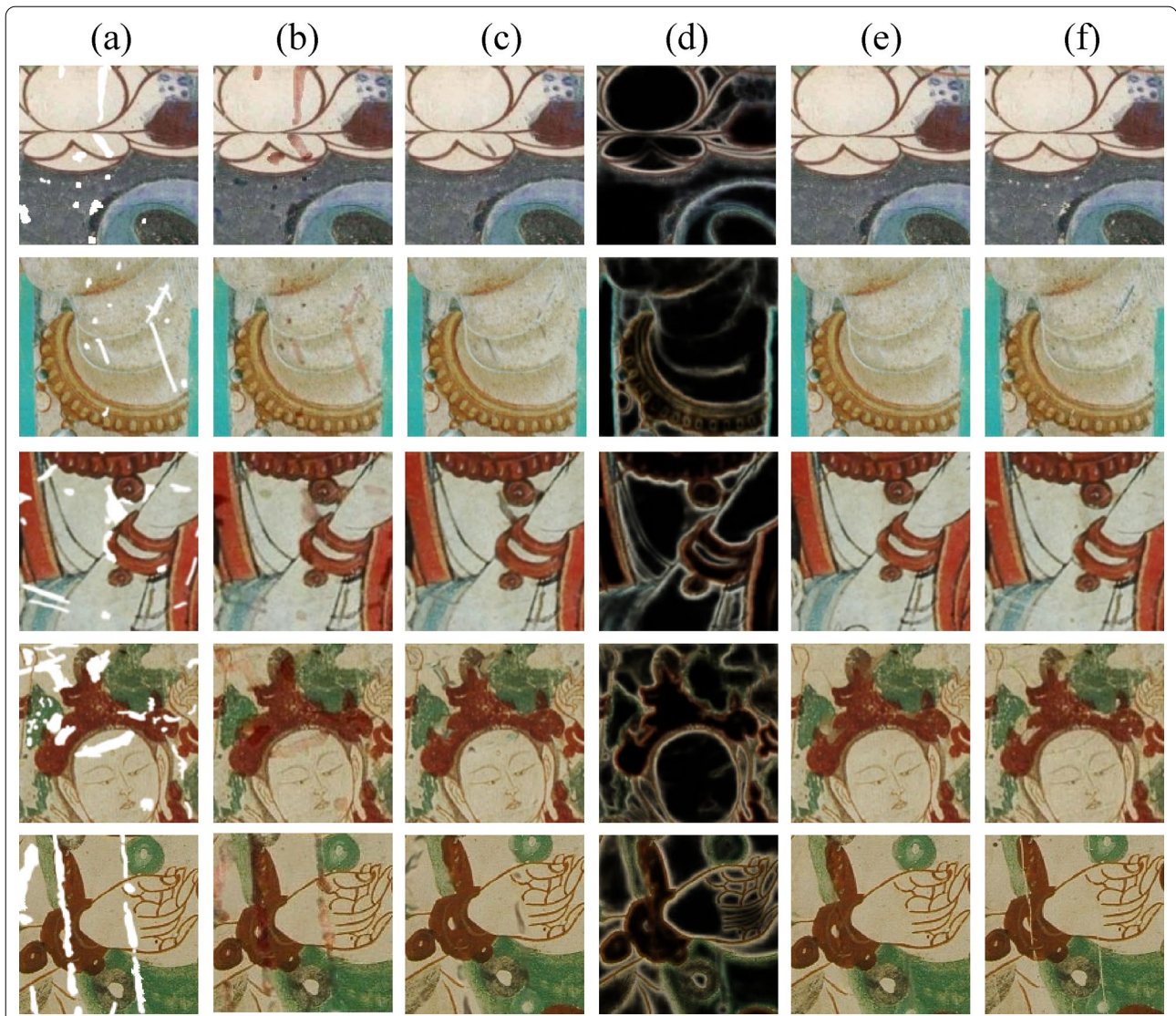


Fig. 7 Comparison of small area real damage inpainting, **a** masked image, **b** EdgeConnect, **c** FRRN, **d** restored contour(ours), **e** SeparaFill(ours), **f** ground truth

Table 4 Comparison of SSIM value and PSNR value of each algorithm with large area damage restoration

Sample	EdgeConnect		FRRN		SeparaFill (ours)	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
1	25.19	0.8874	29.86	0.9457	30.65	0.9512
2	30.86	0.9540	33.72	0.9678	34.53	0.9714
3	25.69	0.8946	29.77	0.9566	31.76	0.9613
4	23.03	0.8327	26.98	0.9059	28.51	0.9206
5	26.11	0.8687	30.23	0.9342	32.36	0.9533

line is repaired firstly, and then the color block area in the contour is contented, which has strong robustness. By increasing the mask area, the influence of discolored

pixels in the neighborhood of the crack area can be reduced, so that the damaged mural can be effectively repaired.

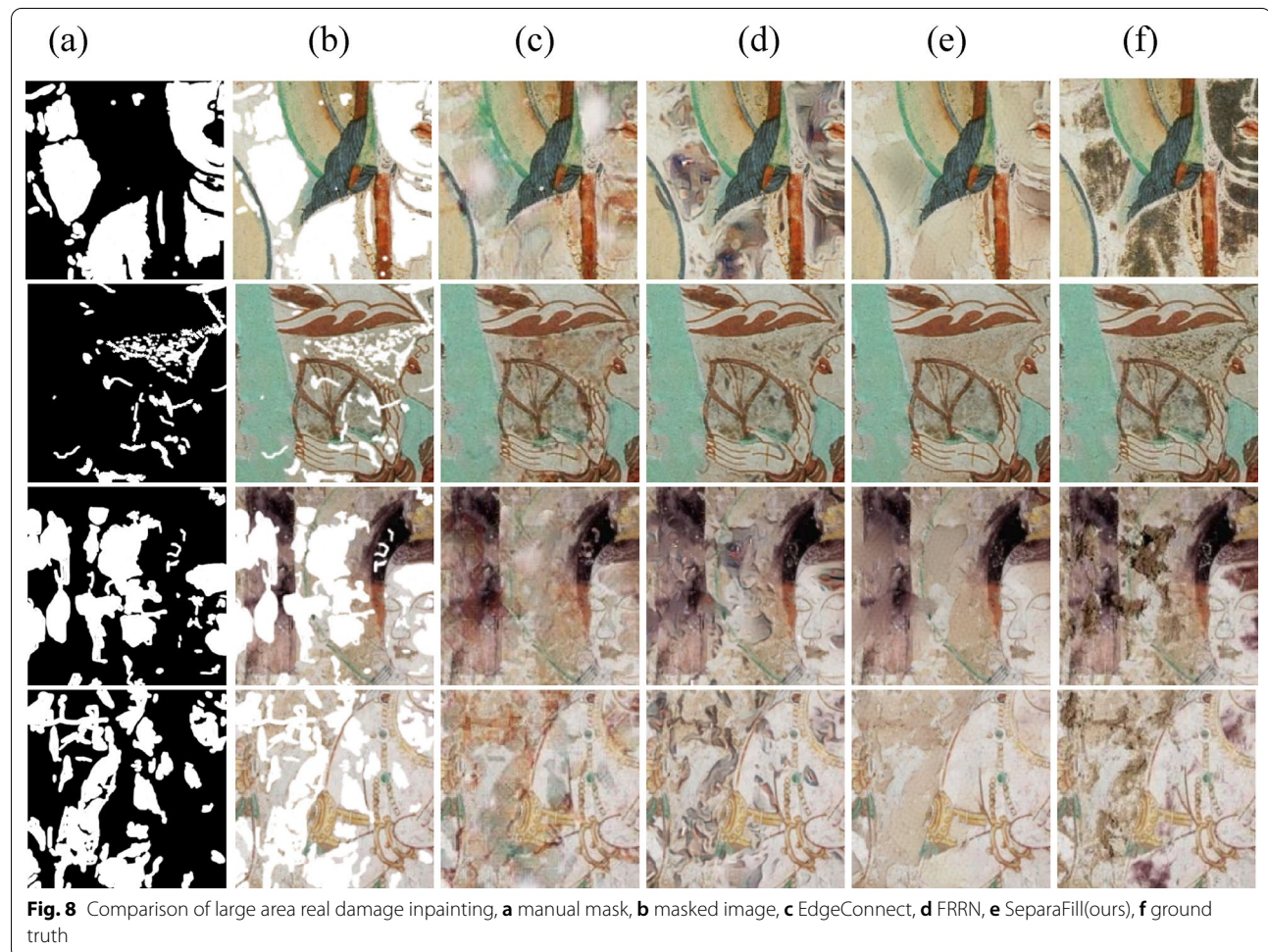
Figure 8 gives four samples with large damaged areas, including diseases of the large-area pulverization, falling off, fading and other diseases of murals. For these diseases, it is difficult to obtain realistic inpainting results because there is less effective information and it is very difficult to label the mask completely. As can be seen from Fig. 8, compared with the other algorithms, the proposed algorithm has obvious advantages in eliminating the influence of dense mildew on the image, completing the damaged wall, and coloring the faded area. The style of the repaired area and the background area tends to be consistent. Our method can avoid the unrepaired area in the FRRN algorithm and the disharmonious texture in the EdgeConnect algorithm. It can reduce the impact of the damaged area and restore the "original appearance" of murals as much as possible when the mask completely covers the damaged area.

Conclusions

According to the painting characteristics of Dunhuang murals, a two generators connected image restoration networks (SeparaFill) based on U-Net is proposed. An

accumulation feature extraction mechanism is presented to reuse the low-level features of images effectively. Firstly, the contour restoration of the image is achieved, and then the completed image contour is used to guide the restoration of color mural image. Experimental results indicate that our algorithm is very effective and show outstanding inpainting performance for both the objective comparisons and the visual performance. Compared with recent algorithms, the algorithm proposed in this study maintains a high PSNR and SSIM index of the repaired mural image. Our algorithm is more consistent with human subjective vision in large-scale damaged image restoration and complex texture restoration.

However, the contour line extracted from mural image is not very distinct because the mural image is extracted directly on the HED network trained by the mural datasets. When extracting the contour line, there is no improvement of network and neither modify the parameters of network. At the same time, the image source comes from the electronic scanning version of the



mural album, which has greatly influence on the quality of the mural image in the process of printing and scanning. The pixel blocks at the edge of the damaged area are badly jagged, so it is very difficult to cover them one by one during the manual mask labels. Due to the above problems, the inpainting of some local details of line structures is not very ideal. In the future, we will search for high-quality mural images to create datasets, and the contour extraction algorithm will be improved to obtain the distinctly contour of mural images. For large damaged mural images, the interactive manual assistance will be provided to restore line details and labeled more accurately to reduce the adverse effects caused by inaccurate calibration.

Abbreviations

GAN: Generative Adversarial Network; ResNet: Residual Networks; Res2Net: Hierarchical Residual Networks; VGG: Visual geometry group network.

Acknowledgements

None.

Author contributions

All the authors contributed to the current work. LCH and LZL devised the research plan and led the writing of the article. LQH and ZJ arranged the data of experiment. SYH and LCH supervised the entire process and provided constructive advice. All authors read and approved the final manuscript.

Funding

None.

Availability of data and materials

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Information and Communication Engineering, Communication University of China, Beijing 100024, China. ²Key Laboratory of Acoustic Visual Technology and Intelligent Control System, Ministry of Culture and Tourism, Beijing 100024, China.

Received: 9 May 2022 Accepted: 14 August 2022

Published online: 31 August 2022

References

- Fan JS. For the long-term survival of Dunhuang—exploration of the protection of Dunhuang Grottoes. *Dunhuang Res.* 2004;111:35–9 (in Chinese).
- Cao JF, Li YF, Zhang Q, et al. Restoration of an ancient temple mural by a local search algorithm of an adaptive sample block. *Herit Sci.* 2019;7:39.
- Jiao LJ, Wang WJ, Li BJ, et al. Wutai mountain mural inpainting based on improved block matching algorithm. *J Comput Aid Design Comput Graph.* 2019;31:119–25 (in Chinese with an English abstract).
- Wang H. Inpainting of Potlala Palace murals based on sparse representation. In: Proceedings of the international conference on biomedical engineering & informatics. IEEE; 2015. p. 737–741
- Wang N, Wang W, Hu W, et al. Thinka mural inpainting based on multi-scale adaptive partial convolution and stroke-like mask. *IEEE Trans Image Process.* 2021;30:3720–33.
- Yan Z, Li X, Li M, et al. Shift-net: image inpainting via deep feature rearrangement. In: Proceedings of the European conference on computer vision (ECCV). 2018. doi:https://doi.org/10.1007/978-3-030-01264-9_1
- Wen LL, Xu D, Zhang X, et al. Some irregular partial repair method of ancient mural based on generative model. *J Graphics.* 2019;5:925–31 (in Chinese with an English abstract).
- Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. In: Proceedings of the international conference on neural information processing systems (NIPS);2014. p. 2672–80.
- Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D.* 1992;60(1–4):259–68. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- Zhang G, Yan Y. Image repair of a fractional TV model combining the texture structure. *J Image Graphics.* 2019;24:5700–13 (in Chinese with an English abstract).
- Criminisi A, Pérez P, Object TK. Object Removal by Exemplar-Based Inpainting. *IEEE conference on computer vision and pattern recognition (CVPR 2003)*. IEEE. 2003;2003:16–22.
- Guo Q, Li J. Damaged Image Restoration Based on Improved Criminisi Algorithm. In: International Conference on Computer Network, Electronic and Automation; 2019. p. 31–35
- Lv C, Lan H, Yu Y, et al. Objective Evaluation Method of Broadcasting Vocal Timbre Based on Feature Selection. *Wirel Commun Mob Comput.* 2022;2022:17.
- Yan M, Li S, Chan C, et al. Mobility prediction using a weighted Markov model based on mobile user classification. *Sensors.* 2021;21(1740):2021.
- Yeh R. A, Chen C, Lim T. Semantic Image Inpainting with Deep Generative Models. In: Proceedings of the 30th IEEE conference on computer vision and pattern recognition (CVPR). IEEE; 2017. p. 6882–90.
- Zhang H, Hu Z, Luo C, et al. Semantic Image Inpainting with Progressive Generative Networks. In: Proceedings of the 26th ACM international conference on multimedia (MM '18). 2018. p. 1939–47.
- Liu G, Reda F. A, Shih K. J, et al. Image Inpainting for Irregular Holes Using Partial Convolutions. In: Proceedings of the European conference on computer vision (ECCV).2018. p. 89–105.
- Guo Z, Chen Z, Yu T, et al. Progressive Image Inpainting with Full-Resolution Residual Network. In: Proceedings of the 27th ACM international conference on multimedia (MM '19). 2019. p. 2496–2504. doi:<https://doi.org/10.1145/3343031.3351022>.
- Yu J, Lin Z, Yang J, et al. Generative Image Inpainting with Contextual Attention. In: Proceedings of the 31th IEEE conference on computer vision and pattern recognition (CVPR). IEEE; 2018. p. 5505–5514.
- Nazeri K, Ng E, Joseph T, et al. EdgeConnect: Structure Guided Image Inpainting using Edge Prediction. In: Proceedings of the international conference on computer vision workshop (ICCVW). IEEE; 2019. p. 3265–74.
- Zamir SW, Arora A, Khan S, et al. Multi-Stage Progressive Image Restoration. In: Proceedings of the conference on computer vision and pattern recognition (CVPR). IEEE; 2021. p. 14816–26.
- Yu J, Lin Z, Yang J, et al. Free-Form Image Inpainting with Gated Convolution. In: Proceedings of the international conference on computer vision (ICCV). IEEE; 2019. p.4470–9.
- Xiong W, Yu J, Lin Z, et al. Foreground-Aware Image Inpainting. In: Proceedings of the conference on computer vision and pattern recognition (CVPR). IEEE; 2019. p. 5833–5841.
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Proceedings of the medical image computing and computer-assisted intervention (MICCAI). 2015. p. 234–241. doi:https://doi.org/10.1007/978-3-319-24574-4_28.
- Li H, Wang W, Guo L, Zhou L. Deep feature rearrangement image repair algorithm based on a dual-transfer network. *J Huazhong Univ Sci Technol.* 2021;49:774–81 (in Chinese with an English abstract).
- Yang H, Yu Y. Image repair using channel attention with hierarchical residual networks. *J Comput Aid Design Comput Graph.* 2021;33:5671–81 (in Chinese with an English abstract).
- Jo Y, Park J. SC-FEGAN: Face Editing Generative Adversarial Network with User's Sketch and Color. In: Proceedings of the international conference on computer vision (ICCV). IEEE; 2019. p. 1745–1753.

28. Liu H, Jiang B, Song Y, et al. Rethinking Image Inpainting via a Mutual Encoder-Decoder with Feature Equalizations. In: Proceedings of the European conference on computer vision (ECCV). 2020. p. 725–741.
29. Gao S, Cheng MM, Zhao K, et al. Res2Net: A New Multi-scale Backbone Architecture. *IEEE*. 2021;43(2):652–62.
30. Xie S, Tu Z. Holistically-Nested Edge Detection. In: Proceedings of the international conference on computer vision (ICCV). *IEEE*; 2015. p. 1395–403. doi:<https://doi.org/10.1109/ICCV.2015.164>.
31. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. In: Proceedings of the conference on computer vision and pattern recognition (CVPR). *IEEE*; 2016. p. 770–8.
32. Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. In: Proceedings of the 31st international conference on neural information processing systems (NIPS'17). 2017. p. 6000–10.
33. Johnson J, Alahi A, Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In: Proceedings of the European conference on computer vision (ECCV). 2016. p. 694–711. doi: https://doi.org/10.1007/978-3-319-46475-6_43.
34. Zhao H, Gallo O, Frosio I, et al. Loss functions for neural networks for image processing. *Computer Sci*. 2015. <https://doi.org/10.48550/arXiv.1511.08861>.
35. Yu T, Guo Z, Jin X, et al. Region Normalization for Image Inpainting. In: the 34th AAAI Conference. 2019. p. 12733–40.
36. Li J, Wang N, Zhang L, et al. Recurrent Feature Reasoning for Image Inpainting. In: Proceedings of the conference on computer vision and pattern recognition (CVPR). 2020. p. 7757–65.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
