

RESEARCH ARTICLE

Open Access



# Ancient mural restoration based on a modified generative adversarial network

Jianfang Cao<sup>1,2\*</sup> , Zibang Zhang<sup>1</sup>, Aidi Zhao<sup>1</sup>, Hongyan Cui<sup>1</sup> and Qi Zhang<sup>1</sup>

## Abstract

How to effectively protect ancient murals has become an urgent and important problem. Digital image processing developments have made it possible to repair damaged murals to a certain extent. This study proposes a consistency-enhanced generative adversarial network (GAN) model to repair missing mural areas. First, the convolutional layer from a fully convolutional network (FCN) is used to extract deep image features; then, through deconvolution, the features are mapped to the size of the original image and the repaired image is output, thereby completing the regenerative network. Next, 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. In adversarial learning, the generative and discriminant network models are optimized to better complete the mural repair. The network introduces a dilated convolution that increases the convolution kernel’s receptive field. Each network convolutional layer joins in the batch standardization (BN) process to accelerate network convergence and increase the number of network layers and adopts a residual module to avoid the vanishing gradient problem and further optimizing the network. Compared with existing mural restoration algorithms, the proposed algorithm increases the peak signal-to-noise ratio (PSNR) by an average of 6–8 dB and increases the structural similarity (SSIM) index by 0.08–0.12. From a visual perspective, this algorithm successfully complements mural images with complex textures and large missing areas; thus, it may contribute to digital restorations of ancient murals.

**Keywords:** Generative adversarial network, Fully convolutional network, Residual module, Mural restoration

## Introduction

Ancient murals are important cultural treasures that record information about religious and cultural characteristics, people’s living conditions and major events in various historical periods, which constitute important references for the study of ancient history. Beyond that, the vivid images and graceful lines of ancient murals have intrinsic artistic value. However, due to the primitive drawing techniques and the fragile materials used in their creation, these murals have undergone constant change throughout their long history. Coupled with the destruction wrought by natural and artificial factors, murals are subject to damage from a wide variety of causes, resulting

in disruption, flaking, blistering and cracking [1]. Therefore, there is an urgent need to protect and digitally restore ancient murals.

Digital restoration of ancient murals has made some progress in recent years. Based on two cases by Mario Sironi and Edmondo Bacci in Venice, Izzo et al. [2] conducted a thorough study of materials and Italian mural painting techniques of different ages to understand their protection needs and formulate sustainable conservation plans. Sakr et al. [3] carried out a study on how *Streptomyces* have affected the colors of ancient Egyptian tomb murals and explored protective solutions, which laid the foundation for the digital protection of ancient murals. Abdel-Haliem et al. [4] isolated and identified *Streptomyces* as the primary cause of the discoloration of tomb murals in ancient Egypt, which provided a new idea for *Streptomyces* elimination. Regarding the reinforcement of murals in the Mogao Grottoes as well as the validity of

\*Correspondence: caojianfangcn@163.com

<sup>2</sup> Department of Computer Science & Technology, Xinzhou Teachers University, No. 10 Heping West Street, Xinzhou 034000, China  
Full list of author information is available at the end of the article

the reinforcement, Li et al. [5] conducted a site inspection to determine whether previously restored murals had undergone additional deterioration. Focusing on ancient murals from aesthetic, biological and chemical perspectives, the abovementioned studies reported the need to protect ancient murals and provided ideas for using intelligent computer information processing technology to help with their repair. In China, digital restoration of ancient murals is still new, although research has made some progress. Jiao et al. [6] proposed an improved block matching algorithm based on the traditional Criminisi algorithm. The principle is to find a suitable match in the image to be repaired to complete the damaged area. This algorithm has a good repair effect on textures, edges and smooth parts, but it does not perform well on complex textures or large areas of loss. Ren et al. [7] proposed a wavelet texture description algorithm as an improvement of the Telea algorithm. The aim of the algorithm is to gradually complete the mural image based on the diffusion of pixels at the edge of the missing area of the mural. This algorithm has a good effect on mural images with narrow and long missing areas, and the repair time is short. The method needs only a broken mural image, but the completion of large missing areas is poor. The basic idea of Cao et al. [8] is similar to that of Jiao et al. [6]. The difference is that by optimizing the local search strategy, the efficiency of searching for matching blocks is improved, so the completion of broken murals is completed faster. However, the shortcomings of Jiao et al.'s method have not been solved completely. Wu et al. [9] improved the information diffusion mode of the curvature-driven diffusions (CDD) algorithm; they replaced orthogonal diffusion with cross diffusion instead of using a specific orientation to adapt to the missing area, providing a more reasonable filling method [9]. However, their approach proved unsatisfactory to restore the textures of large missing areas. Li et al. [10] used the sample-based restoration algorithm to repair calibrated mud-spotted areas. Despite the realization of automatic labeling and virtual restoration, their approach is greatly limited by damaged areas.

With the development of computer vision and image processing technology, image restoration technology has become an increasingly important aspect and has developed rapidly. Image restoration algorithms are generally divided into traditional methods and deep learning methods. The traditional algorithms are mainly based on the partial differential equations proposed by Bertalmio et al. [11] and on the sample blocks method proposed by Crinisi et al. [12]. Iteration completion is a method based on partial differential equations, whose principle is to diffuse the information outside the missing area into

the missing area in stages. Each iteration step propagates pixel information from the image along an isoillumination line to the missing area. This method achieves good effects on small missing areas, such as slits and areas with words; however, its effect on larger missing areas is poor. The basic idea of the sample-block-based algorithm is to find an appropriate sample block from the image area to fill the missing area of the image. Although this method can be used to complete large loss areas, it produces a satisfactory result only when similar content occurs elsewhere in the undamaged areas of the image itself or in the database; by itself, this method cannot generate new content. Moreover, it is inefficient at searching for matching sample blocks. In recent years, image restoration methods based on deep learning have also developed rapidly. Pathak et al. [13] first used the neural network (NN) method in 2016 [mainly convolutional neural networks (CNNs) and generative adversarial networks (GANs)] to propose an unsupervised visual feature-learning algorithm for context-based pixel prediction; this proposal laid the foundation for subsequent methods. Yang et al. [14] proposed a multiscale neural patch synthesis method based on jointly optimizing the image content and texture constraints, which greatly improved the effect of high-resolution image restoration. Liu et al. [15] introduced local convolution to repair arbitrary, noncentral and irregular regions. Combined with the traditional patch synthesis method, Yu et al. [16] achieved a satisfactory repair effect by diffusing the texture information around the repair area. Yan et al. [17] proposed ShiftNet, which can guide the shift of coder features in known regions by calculating the decoder features of the missing regions, thereby completing image restoration. Zhang et al. [18] proposed an end-to-end progressive generation network approach that divides image restoration into several parts. A long short-term memory (LSTM) recurrent NN is used to integrate these parts, which then gradually reduce the size of the repair area until repair is complete. Chen et al. [19] proposed a progressive restoration algorithm based on a GAN that performs low-resolution image restoration followed by gradual refinement until the high-resolution image is restored.

However, the abovementioned traditional image restoration methods still have many shortcomings. For instance, when no similar content in the image database corresponds to the missing areas of a mural, the restoration effect of these methods is poor, and they cannot provide a satisfactory solution when a texture deficiency exists in a large missing area. Moreover, applying deep learning directly to mural restoration faces problems such as difficulty in extracting features, long network model training times and inconsistent global integrity of the repaired image.

To solve these problems, this study proposes a new consistency-enhanced GAN algorithm and applies it to the restoration of ancient mural images. The improvements achieved by the proposed algorithm are mainly reflected in the following aspects: (1) the framework of the proposed algorithm is based on a GAN, which consists of a generative network and a discriminant network; mean squared error (MSE) and adversarial loss functions are introduced to optimize these networks in two stages; (2) the generative network is based on a fully convolutional network (FCN), which takes the damaged image as input and outputs the repaired image; (3) cavity convolution is used instead of a pooling operation to optimize the network structure and reduce image information loss; and (4) the discriminant network uses global and local discriminant networks to jointly optimize the network model and determine whether the input sample image is an authentic mural image. This approach enhances the consistency between the global and local effects of the repaired image.

## Methodology

### Theoretical background

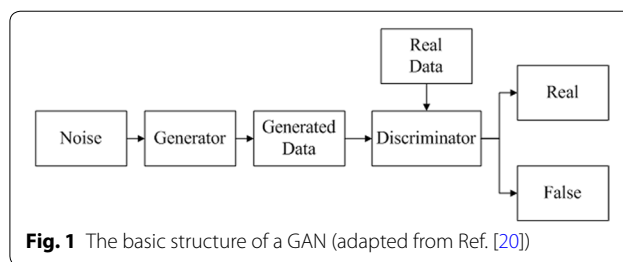
#### GAN

A GAN [20] is an unsupervised learning algorithm that includes generators and discriminators. Inspired by game theory, GANs regard image generation as a confrontation and a game between generators and discriminators. The purpose of the generative network is to generate synthetic data from noise data (such as data from a normal or Gaussian, distribution) as a false sample. The input to the discriminant network includes both the output data of the generative network and real input data (i.e., both false samples and true samples). Generally, a probability value is output by an NN—when the probability value exceeds 0.5, the input sample is true, and when it is below 0.5, the input sample is false. The generative network attempts to produce increasingly realistic data, while the corresponding discriminant network attempts to distinguish the real data from the synthetic data more accurately. The two networks learn from each other by competing, and they continually challenge each other, causing the data generated by the generative network to become closer to the real data. As the network learns in this circular manner, the generated samples become increasingly more similar to the real data.

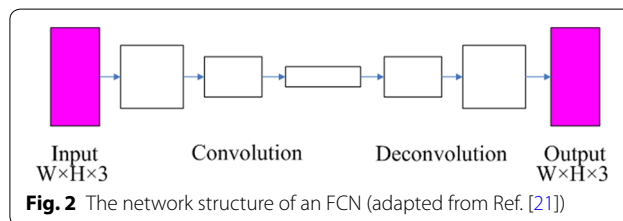
The basic GAN structure is shown in Fig. 1.

#### FCN

The FCN was proposed by Berkeley et al. [21] in 2015 for image semantic segmentation. FCNs can classify images at the pixel level. The traditional CNN connects the fully connected layers placed after the convolutional



**Fig. 1** The basic structure of a GAN (adapted from Ref. [20])



**Fig. 2** The network structure of an FCN (adapted from Ref. [21])

layers to obtain a fixed-length eigenvector for classifying the images. Unlike a traditional CNN, images of any size can be input into an FCN, which restores the deep image information obtained by the convolutional layer to the same size and spatial characteristics as the input image through deconvolution to finally achieve pixel-by-pixel classification. The network structure of an FCN is shown in Fig. 2.

#### Batch normalization

In the traditional deep NN training process, which involves optimizing and updating the parameters in each network layer, the input data distribution in each layer is often quite different from the distribution in the initial input. The network continuously adapts to the new data distribution, which leads to problems such as long training times.

Similar to the principle of preprocessing data before inputting them into the NN, batch normalization (BN) is performed before each activation function of the convolutional layers of the NN to achieve a stable data distribution in each layer. The concept of BN was first proposed by Sergey Ioffe et al. in 2015 [22] to solve the internal covariate shift problem. The basic idea of BN is to fix the input distribution of the hidden-layer nodes in each network.

Applying BN ensures a higher learning rate during the training process, which speeds up network convergence and reduces the network model training time. Moreover, BN transforms the input data distribution of the activation function into a Gaussian distribution, which is helpful for solving vanishing and exploding gradient problems.

### Modified GAN-based ancient mural restoration

Based on the characteristics of ancient murals and image restoration algorithms, this study designs a new consistency-enhanced GAN algorithm. This algorithm focuses on three main aspects: designing and optimizing the network structure, selecting the loss function and planning the training and testing process.

A GAN has the ability to generate new data according to the characteristics of existing data. Therefore, this paper selects a data set that satisfies the experimental requirements. The generator attempts to repair the to-be-completed area of a mural. The neural network learns the composition rules and the connections between the pixels inside the mural image. Then, the network model is gradually optimized until the capability to generate pixels that meet the requirements of the completion area is obtained.

GANs avoid the need to design complex loss functions. In the field of image restoration, designing a loss function that can perfectly express the quality of a completed image is a difficult problem, and a generator that generates an adversarial network can avoid this difficulties. The generator uses the original image as a correct indicator, which is sufficient to generate an image similar to the original image. This is especially true in the field of mural image restoration. Ancient Chinese mural images involve other knowledge, such as art and culture, and care must be taken in the quality evaluation of mural image completion. We cannot simply use ordinary loss functions and GANs to address this problem.

The algorithm in this paper uses two discriminators to solve the problem of poor consistency between the overall effect and the local effect of mural image restoration to better complete the mural image restoration.

### Network structure design

The mural image restoration network consists of generators and discriminators. Generators are responsible for outputting a repaired mural image, and discriminators are responsible for judging whether the images output by the generators are “true” or “false” compared with the authentic mural image. The learning rate of the network is  $1e-4$ , the optimizer is AdamOptimizer, the batch size is 16, the last layer after the convolution in the generation network uses the tanh activation function, and the remaining layers use the relu activation function. The relu activation function is used after discriminating the convolutional layers in the network.

**1. Generative network design** The generative network uses an FCN framework and follows an encoder-decoder structure. It treats the task of mural image comple-

tion as one of changes in the pixel information in only a masked area, leaving the image information in other areas unchanged. The FCN extracts image feature information using the convolutional layers, and deconvolution restores the feature information into an image of the same size as the original image. The final restored image is output by restoring the pixel values of the nonmasked areas and retaining the image information assigned to the restored masked areas. The details of the generative network are summarized in Table 1.

Basic CNN knowledge predicts that increasing the number of network layers will result in the extraction of more feature information pairs; thus, increasing the number of network layers results in more abundant mural image features. Layers 3–7 are convolutional layers. The residual module avoids the vanishing and exploding gradient problems to some extent by means of the residual learning method.

The current generative network employs cavity convolution during image feature extraction. In a traditional FCN, the input image undergoes convolution, pooling and deconvolution, and finally outputs an image matching the original size. During this process, the reduction in the size of the original image followed by image size augmentation causes a partial loss of image information; however, retaining as much information from the original image as possible is desirable to ensure the integrity of the restored image. Cavity convolution enables each of

**Table 1 Detailed information of the generative network**

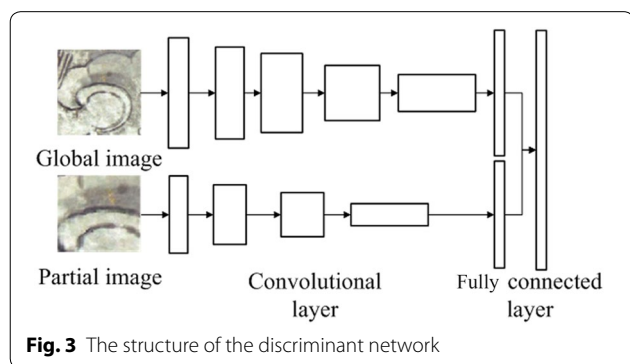
Name	Type	Kernel	Dilation ( $\eta$ )	Stride	Outputs
Conv1	conv	$5 \times 5$	1	$1 \times 1$	64
Conv2	conv	$3 \times 3$	1	$2 \times 2$	128
Conv3	conv	$3 \times 3$	1	$1 \times 1$	128
Conv4	conv	$2 \times 2$	1	$1 \times 1$	256
Conv5	conv	$1 \times 1$	1	$1 \times 1$	128
Conv6	conv	$1 \times 1$	1	$1 \times 1$	256
Conv7	conv	$3 \times 3$	1	$2 \times 2$	128
Conv8	conv	$3 \times 3$	1	$1 \times 1$	256
Dilated conv1	dilated conv	$3 \times 3$	2	$1 \times 1$	256
Dilated conv2	dilated conv	$3 \times 3$	4	$1 \times 1$	256
Dilated conv3	dilated conv	$3 \times 3$	8	$1 \times 1$	256
Conv9	conv	$3 \times 3$	1	$1 \times 1$	256
Conv10	conv	$4 \times 4$	1	$1/2 \times 1/2$	128
Deconv1	deconv	$3 \times 3$	1	$1 \times 1$	128
Conv11	conv	$4 \times 4$	1	$1/2 \times 1/2$	64
Deconv2	deconv	$3 \times 3$	1	$1 \times 1$	32
Output	output	$3 \times 3$	1	$1 \times 1$	3

the convolution kernels to possess a large receptive field without requiring an increase in the number of parameters and the associated computing power. It enlarges the receptive field area without causing image information loss. Therefore, this study replaces the traditional pooling operation with cavity convolution. In this process, dilation represents the magnitude of the cavity convolution; the larger the dilation is, the larger the receptive field is.

**2. Discriminant network design** The discriminant network contains both global and local discriminant networks that are responsible for judging whether the output image is authentic or was synthesized by the generative network. The discriminant network is based on a CNN. Specifically, the local image information is extracted by the convolutional layers, filtered and selected by the pooling layers, and then integrated into eigenvectors by a fully connected layer to achieve image classification. Two eigenvectors with the same dimensions are separately input into the two types of networks. The input eigenvectors are connected; thus, they output a single eigenvector. The authenticity of the output image is then judged by a sigmoid function. The structure of the discriminant network is shown in Fig. 3.

In the discriminant network design, a few layers are generally sufficient to judge the extracted features regarding the authenticity of the input image, and having fewer layers reduces the computational complexity of the network updating process. However, mural images have rich texture information structures whose correlation within each region of the image is high. Experiments show that the typical number of network layers is insufficient to accurately judge mural images. Therefore, a deep convolutional neural network structure is used to extract the mural image information.

The global discriminant network takes two images as input: the real mural image and the network-composited image. It consists of 9 convolutional layers and a fully connected layer and outputs a  $1 \times 1024$  feature vector.



The convolutional layer uses  $5 \times 5$ ,  $3 \times 3$  and  $2 \times 2$  convolution kernels with  $2 \times 2$  and  $1 \times 1$  step sizes to extract global mural image information. The details of the global discriminant network structure are listed in Table 2. The input image to the local discriminant network is a  $32 \times 32$ -pixel region of the generated network output and the region of the authentic image corresponding to the generated region. The local discriminant network consists of 8 convolutional layers and 1 fully connected layer, and it outputs a  $1 \times 1024$  feature vector. The convolutional layer uses  $3 \times 3$  and  $2 \times 2$  convolution kernels with  $2 \times 2$  and  $1 \times 1$  step sizes to extract local mural image information. The details of the local discriminant network structure are listed in Table 3.

Finally, the output of the global discriminant network and that of the local discriminant network are connected to obtain a  $1 \times 2048$  eigenvector. This eigenvector is then processed by the fully connected layer, which outputs a continuous value. The sigmoid function is used to map the obtained value into a  $[0,1]$  range; this new value corresponds to the probability that the image is authentic

**Table 2 Detailed information of the global discriminant network**

Type	Kernel	Stride	Outputs
conv	$5 \times 5$	$2 \times 2$	64
conv	$5 \times 5$	$2 \times 2$	128
conv	$5 \times 5$	$2 \times 2$	256
conv	$5 \times 5$	$2 \times 2$	512
conv	$3 \times 3$	$2 \times 2$	256
conv	$3 \times 3$	$2 \times 2$	256
conv	$2 \times 2$	$2 \times 2$	256
conv	$2 \times 2$	$1 \times 1$	512
conv	$2 \times 2$	$1 \times 1$	512
FC	–	–	1024

**Table 3 Detailed information about the local discriminant network**

Type	Kernel	Stride	Outputs
conv	$3 \times 3$	$2 \times 2$	64
conv	$3 \times 3$	$2 \times 2$	128
conv	$3 \times 3$	$2 \times 2$	256
conv	$3 \times 3$	$2 \times 2$	512
conv	$2 \times 2$	$2 \times 2$	256
conv	$2 \times 2$	$1 \times 1$	256
conv	$2 \times 2$	$1 \times 1$	512
conv	$2 \times 2$	$1 \times 1$	512
FC	–	–	1024

**Table 4 The combination structure of the fully connected layer**

Type	Outputs
concat	2048
FC	1

rather than compositely restored. The combined structure of the fully connected layer is shown in Table 4.

### Loss function

To better generate near-authentic repair effects, MSE and adversarial loss functions are proposed to optimize the network model.

To accelerate model training time, the initial repaired image output from the generative network is not input directly into the discriminant network for judgment and optimization. Instead, the network model is trained in two stages. In the first stage, the loss value of the MSE is reduced to the expected value by the generative network training model, which enables the generative network to output a high-quality repaired image. In the second stage, the repaired image output by the generative network is used as a false sample for the discriminant network. The generative network and the discriminant network are combined, and the MSE and adversarial loss functions are optimized to refine the generative and discriminant network models.

The MSE loss function is the most commonly used regression loss function. The sum of the squares of the distances between each sample target variable and the predicted value is calculated, and the MSE is the ratio between the sum of all the squared losses of the samples and the number of samples [23]:

$$MSE = \frac{1}{N} \sum_{(x,y) \in D} (y - prediction(x))^2, \quad (1)$$

where  $N$  is the number of samples,  $(x, y)$  represents the sample, for which  $x$  is the feature set of the trained sample and  $y$  is the actual value of the trained sample, and  $prediction(x)$  is the predicted  $x$  value of the sample.

The idea behind the GAN is that the generative network attempts to generate images sufficiently similar to authentic images so that its output images will be judged as authentic by the discriminant network; that is, the purpose of the generative network is to generate images whose predicted probability is close to 1 according to the discriminant network. Similarly, the purpose of the discriminator is to enable its output of the generative network input to be close to 0 while enabling its output for

the actual data to be close to 1. The equation for adversarial network loss is as follows:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_{data(x)}} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (2)$$

where  $x$  represents the actual data,  $z$  represents noisy data,  $P_{data(x)}$  is the probability distribution of the actual data,  $p_z(z)$  is the probability distribution of the synthesized data,  $G(z)$  is the data synthesized by the generative network,  $D(x)$  is the probability of the authenticity of the data judged by the discriminant network,  $D(G(z))$  is the probability of the authenticity of the data generated by the generative network and judged by the discriminant network,  $\log D(x)$  is the judgment of the actual data by the discriminant network, and  $\log(1 - D(G(z)))$  is the judgment of the synthesized data by the discriminant network.

### Training and testing procedures

In the training process, a complete GAN framework, that is, a generator and a discriminator, is used. The purpose is to optimize the generator according to the completion and generate mural images that meet the requirements. In the test process, only the generator in the GAN was used to obtain the final mural image completion effect and the value of the loss function.

The input for generating the network is a preprocessed mural image of  $128 \times 128$  pixels. A rectangular area with a length and width of 24–36 pixels is randomly generated as a mask at any position in the image, and the pixel value of the mask area is set to 0. Finally, the mural image with the mask is obtained.

The training process of the network model in this study consists of training the generative network model and the combined training of the generative network and the discriminant network. The algorithm pseudocode is as follows.

Input: Mural image.

Output: The network model and the restored mural images.

Step 1. A mask of the original mural image is selected; the area to be repaired is simulated and then added to the original mural image;

Step 2. The restored image is input into the convolutional layers for mural image feature extraction;

Step 3. The mural image features are input into the deconvolutional layer to restore the size of the new image to that of the original;

Step 4. The masked region is removed, and the pixels of other image regions are restored to the corresponding pixels from the original image;  
 Step 5. The MSE loss function is calculated;  
 Step 6. The network model is saved, and the repaired mural image is output;  
 Step 7. Steps 1–6 are repeated, and the MSE loss function is optimized with a gradient-based method until the expected value is reached;  
 Step 8. The repaired global and local images are input as the false samples of the global and local discriminant networks, respectively; while the authentic mural image and the corresponding local image are input as the authentic samples of the discriminant network;  
 Step 9. The global and local discriminant networks use CNNs to extract features. Finally, a fully connected layer is connected for the eigenvector outputs;  
 Step 10. The two vectors output by the fully connected layer are combined into one eigenvector, which is converted into a probabilistic value with a sigmoid function;  
 Step 11. The generated and discriminant network models are saved to judge the authenticity of the input mural image;  
 Step 12. Steps 7–11 are repeated, and the parameters of the generative and discriminant network models are updated and optimized until the desired effect is achieved.

A flowchart of the model training procedure is shown in Fig. 4.

In the model training stage, the generative and discriminant network models learn and update the parameters through competition, which finally enables the generative network to complete the task of mural image restoration satisfactorily. Next, testing is initiated, and the effect of mural image restoration on the test dataset is observed.

Therefore, the trained generative network (whose parameters do not need further updating) is used as the test network. The algorithm pseudocode is as follows:

Input: Mural image;

Output: The restored mural image.

Step 1. A mask from the original mural image is selected, and the area to be repaired is simulated and then added to the original mural image;

Step 2. The network model is loaded, and the image to be repaired is input into the convolutional layer for mural image feature extraction;

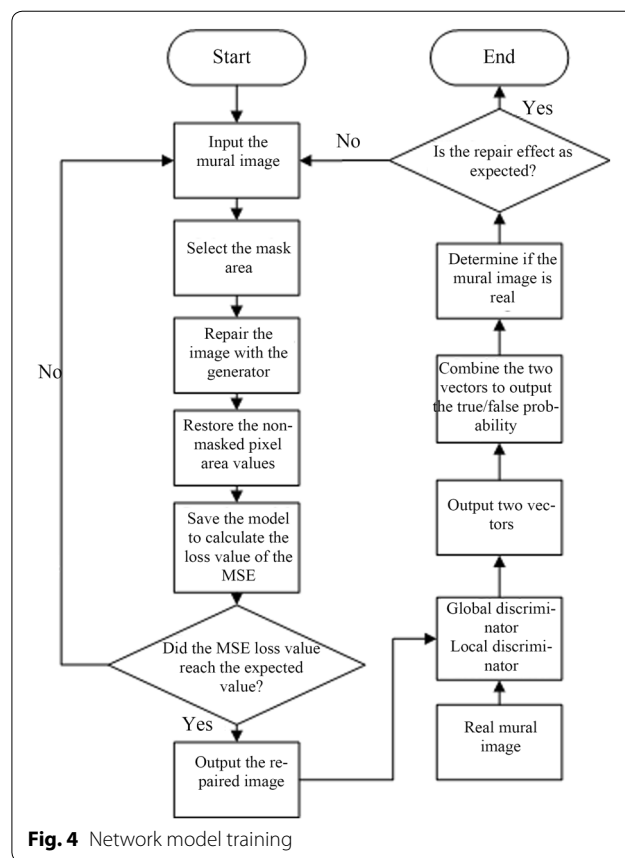


Fig. 4 Network model training

Step 3. The mural image features are input into the deconvolutional layer to restore the size of the new image to that of the original;

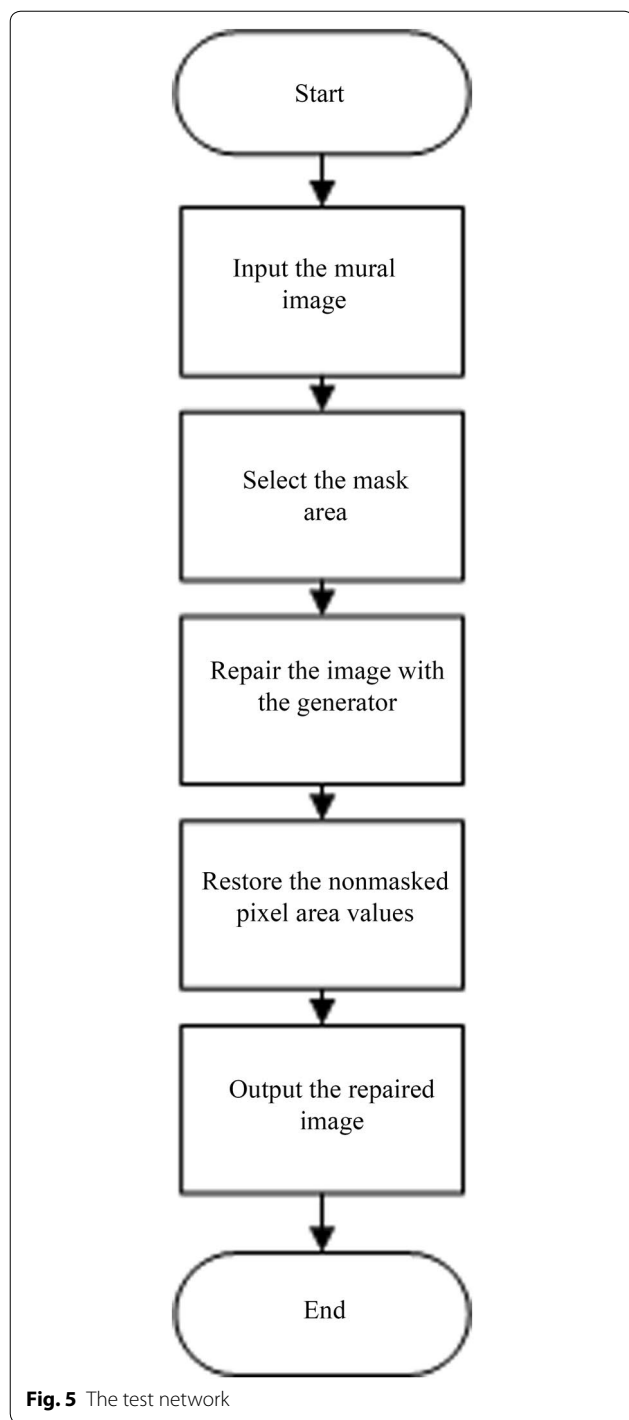
Step 4. The masked region is removed, and the pixels in the other image regions are restored to the corresponding pixels of the original image;  
 Step 5. The repaired mural image is output.

A flowchart of the test network is shown in Fig. 5.

## Results and discussion

### Experimental environment

To verify the effectiveness of the proposed consistency-enhanced GAN, tests on mural image restoration were conducted. The hardware environment in this experiment mainly consists of an Intel Core i5-9400F CPU @ 2.90 GHz with 16 GB memory and an Nvidia GeForce RTX 2070 graphics card. The software environment includes the JetBrains PyCharm compiler, running on a Windows 10 system. The software was written in Python 3.7, and TensorFlow was used as the framework for complete mural image restoration.



#### Data source

Due to a lack of a standard mural image datasets, as well as the small number of existing mural images, inconsistency in the damage degree of mural images, low image resolution and complex mural themes, the selection and processing of mural images is an important consideration

in the experiment. In this study, 800 mural images with good photographic quality from the temples of Wutaishan, Shanxi Province, were used. After image augmentation and expansion, 12,000 mural images were obtained. Among these, 10,000 were used as the training set, and the remaining 2000 images were used as the test set. These datasets are utilized for network model training and effect testing, respectively.

The experimental results show that data augmentation not only improves the quality of mural image but is also applicable for expanding the images and increasing the number of images, all of which improve the robustness of the final network model for mural image restoration. The image augmentation procedures primarily involve random flipping, while the color augmentation procedure mainly involves changing image brightness, saturation and contrast. Professionals engaged in the protection of ancient murals were invited to screen the mural images expanded by the data augmentation algorithms and select the most realistic and qualified mural images. This process helped to determine the appropriate ranges for the data augmentation algorithm parameters and ensure the viability of the final data set. Through augmentation, the number of images was increased. After several experiments, the parameters of the mural data augmentation were adjusted to satisfy the requirements of the image restoration experiment. Some example of the data augmentation effects are shown in Fig. 6.

#### Comparison tests

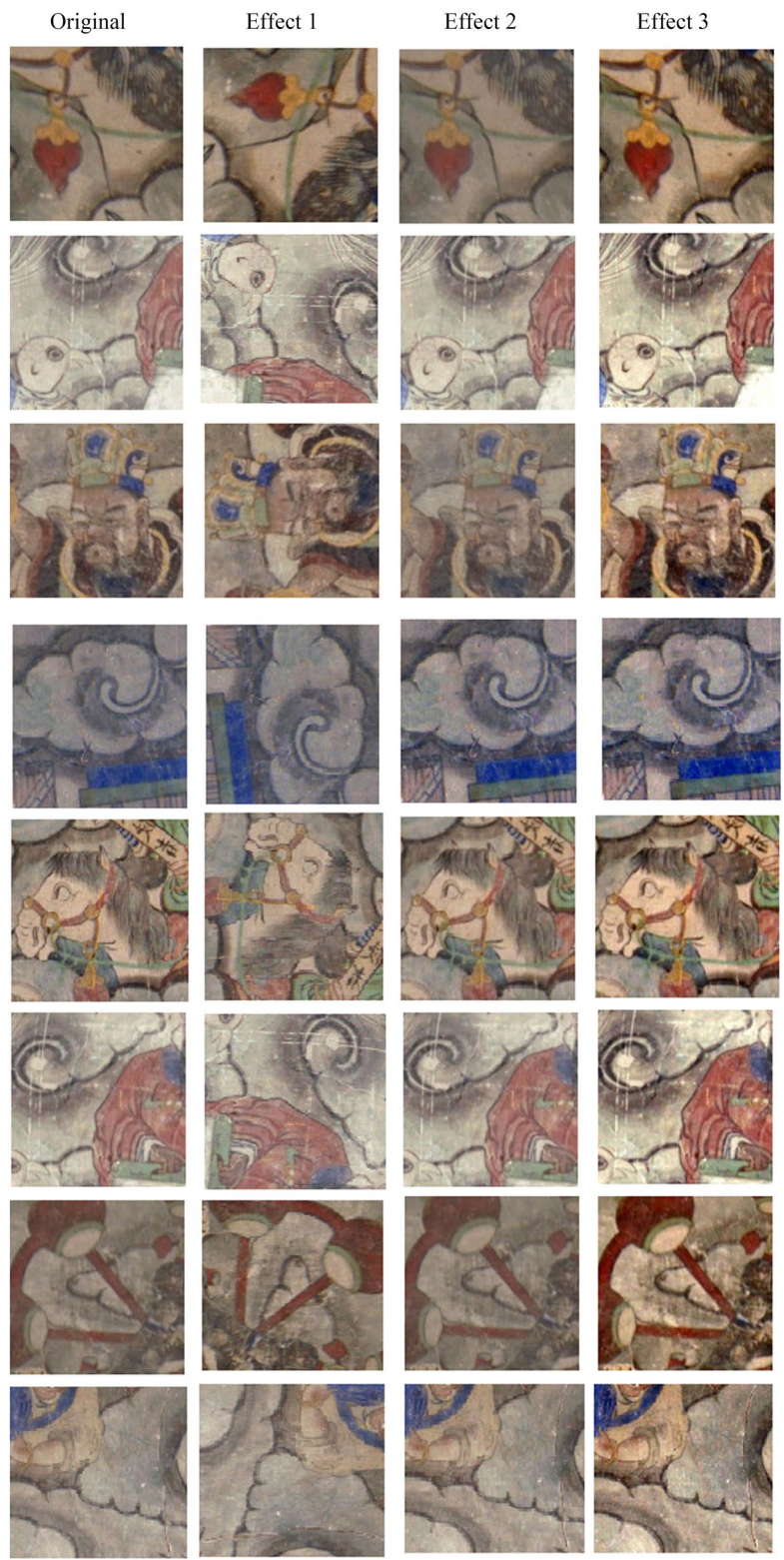
In this paper, murals from the Wutaishan temple Shanxi are taken as experimental objects, and the developed system was applied to repair mural images with authentic damage and artificially damaged mural images. The restoration effects of the method proposed in this study are compared with those of the methods in [6–8] on the same set of mural images.

#### Restoration of the authentic damaged murals

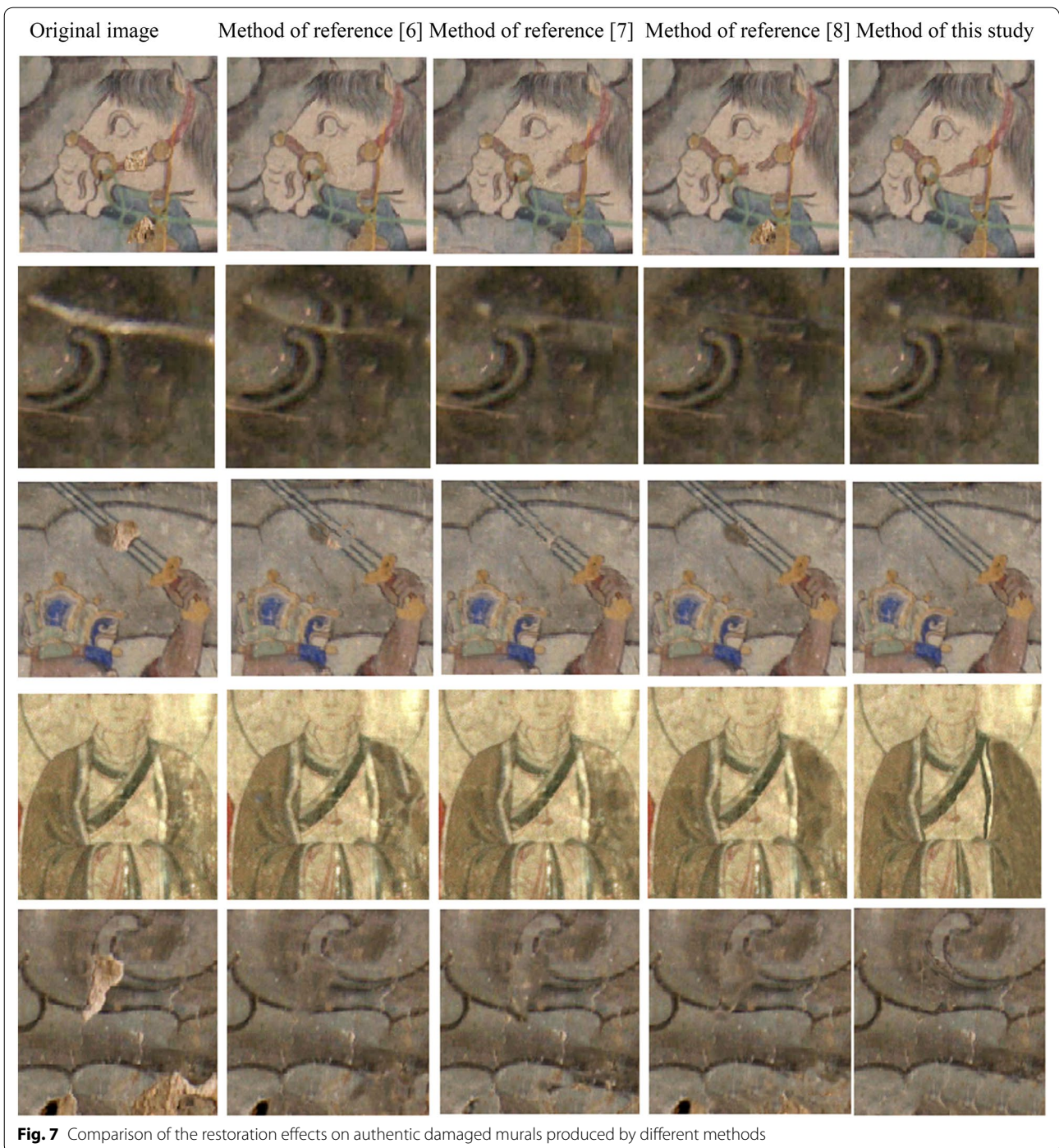
Ten authentic mural images with different damage types, area sizes and area shapes were selected as the experimental objects; the corresponding restoration effects of the four tested algorithms are shown in Fig. 7.

As shown in Fig. 7, the methods reported in the literature all achieve good effects when used for the restoration of long areas whose texture structure does not need repair. However, due to the limitations of the traditional algorithms, varying degrees of texture fragmentation and blurring occur when restoring regions with complex texture structures. Especially on larger damaged regions, the algorithm in this study performs better than do the other methods. The proposed algorithm is not only able





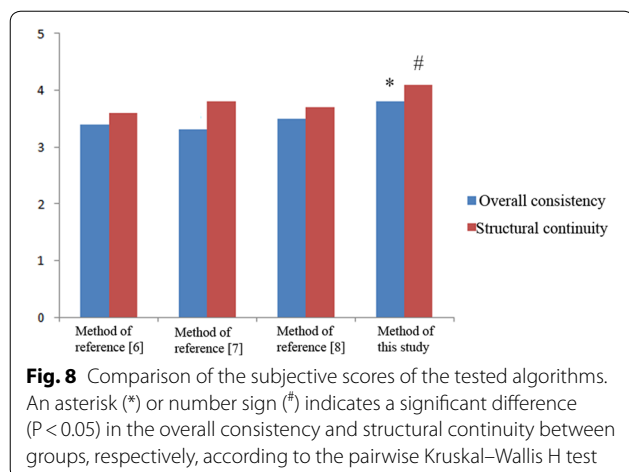
**Fig. 6** Examples of the data augmentation effects



to restore the color to large damaged regions but also achieves a satisfactory effect when restoring regions whose texture structure is missing, resulting in strong visual consistency.

To ensure a more convincing subjective evaluation, two experts on ancient mural repair were invited to score the experimental objects repaired by the four methods

in terms of overall consistency and structural continuity using a blinding method [8]. The scoring system contains 10 levels, where the highest possible score is 10 points and the lowest is 1 point. After scoring, averages are obtained and compared. Compared to the algorithms in [6–8], the algorithm proposed in this study achieves significantly high scores in terms of overall consistency and



structural continuity (all  $P < 0.05$ ; Fig. 8). These results indicate that the algorithm proposed in this study outperforms those in the literature in terms of subjective evaluations.

#### Restoration of the artificially damaged mural images

To restore mural images, different images are selected, and a region on each images is randomly selected and then damaged using different mask sizes. The restoration effects based on the different tested methods are shown in Fig. 9.

In Fig. 9, all the algorithms achieve good effects at simple color filling. However, when addressing texture information restoration, textural disruption and incompleteness and distortion of the repaired area appear for the method in [7]. Although the algorithm in [6] yields a noticeable improvement for repairing texture disruptions, there is still some evidence of repair. The algorithm in [8] performs well at repairing large damaged areas with a single color but performs poorly on texture information restoration. Furthermore, due to the limitations in matching block searching and texture diffusion during large-area restoration, when the texture information to be processed is complex, the texture information resulting from the algorithms in [6–8] is relatively blurred and fails to reflect the texture structure of the original mural image. In contrast, the algorithm proposed in this study uses two discriminators, which enhances the consistency between the global and local expressions of the mural image. Therefore, it can restore even mural images with large damaged area and complex image information.

Furthermore, three groups of masks with sizes of  $18 \times 18$ ,  $24 \times 24$ , and  $36 \times 36$  are selected, and the differences in the effects of different mask sizes on mural image repair are compared under the four algorithms. Ten images are selected in each group, and the average

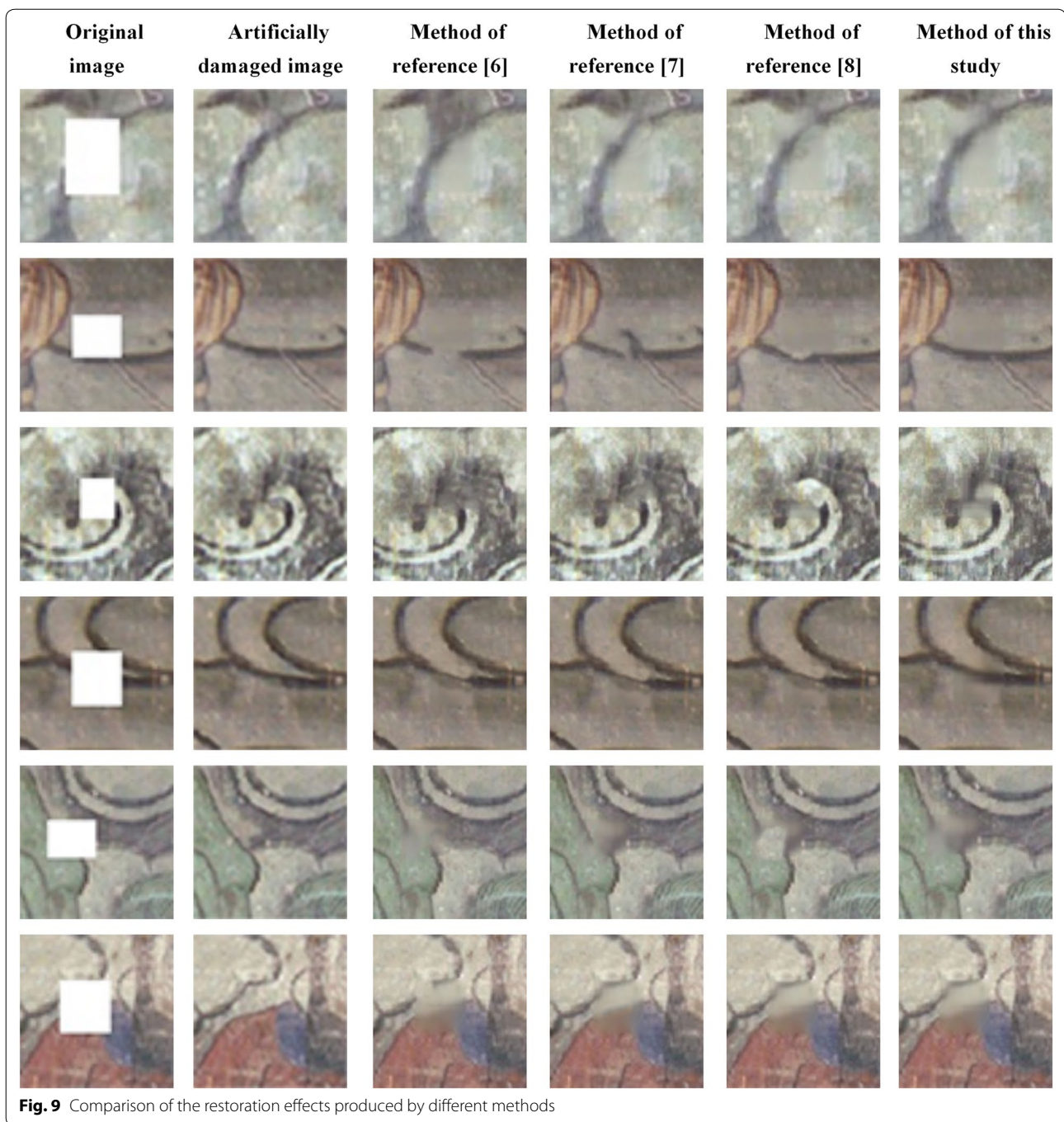
of the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) is calculated and compared. PSNR, which is typically used to evaluate the quality of an image after repair compared with the original image, is based on the error between corresponding pixels. Generally, the higher the PSNR is, the smaller the image distortion and the better the corresponding image repair. The PSNR comparison result is shown in Fig. 10. SSIM is generally used to evaluate the similarity of image structure. It measures the similarity of images in terms of brightness, contrast, and structure. The value range of SSIM is [0, 1]: the larger the value is, the higher the image similarity.

The comparison between PSNR and SSIM indicates that when the damaged area of the mural image is small, the difference between the PSNR and SSIM of the algorithm in this paper and the literature [6–8] is not large (Figs. 10 and 11). When the area is  $36 \times 36$ , a significant difference is observed between the algorithm proposed in this study and those in the literature ( $P < 0.05$ ; Fig. 10), and the average PSNR is 6–8 dB higher than that of other algorithms. In addition, because the two discriminative networks are merged to promote the optimization of the generated network model, the consistency between the overall recovery effect and the local effect is greatly improved, and the SSIM based on the algorithm proposed in this study is improved by an average of 0.08–0.12 compared with that of the methods in literatures, showing significant differences (Fig. 11).

#### Conclusions

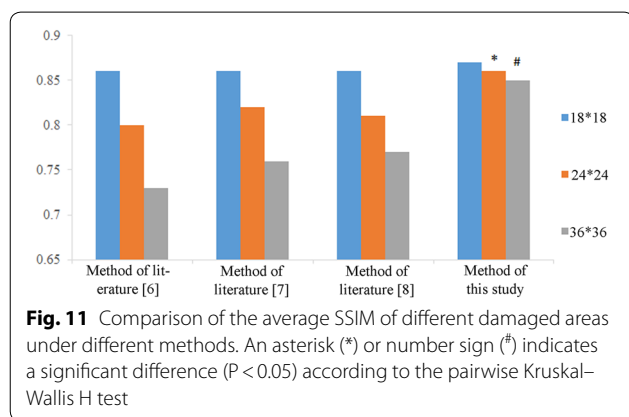
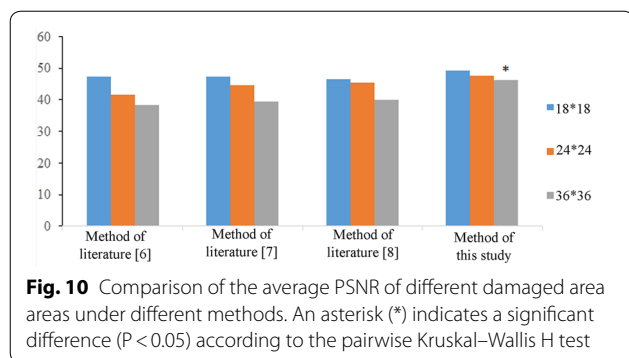
Targeting the problem of restoring images of damaged ancient murals, this study proposes a consistency-enhanced GAN to achieve mural restoration. First, the generative network, using an FCN as the basic framework, restores the masked portion of a mural image. Then, local and global discriminant networks are combined to optimize the network model, strengthening the consistency between the global and local expressions of the mural image output by the generative network. Compared with the existing algorithms, the algorithm proposed in this study noticeably improves the subjective visual effects as well as the PSNR and SSIM index values of the repaired mural image. The results show that the proposed algorithm can better repair mural images with complex image information and strong texture structures.

However, this study still has some limitations. First, when the algorithm proposed in this study is used for mural repair, the image quality around the masked area of the mural image to be repaired must be high. Second, some problems occur, such as blurring of the repaired area and a lack of texture information in regions with complex texture and excessive missing texture. Third,



this study tested only Chinese ancient murals. In the future, high-quality mural image datasets with more abundant subject matter, including murals from other countries, should be acquired, and appropriate data augmentation techniques should be adopted to further expand the datasets. Fourth, due to a lack of consistent standards for mural image restoration evaluation, as well as possible large differences in restoration work

performed by different conservation professionals, the design of this study did not include comparisons between restoration works by the proposed algorithm and those by conservation professionals. In the future, such comparisons could be made to further improve the effectiveness of the algorithm proposed in this study for mural image restoration. In addition, transfer learning should be introduced, and more advanced network models



should be adopted for mural image dataset training. The network might benefit from layer increases to obtain more image information, thereby enabling large-area mural image restoration.

#### Abbreviations

GAN: Generative adversarial network; FCN: Fully convolutional network; BN: Batch normalization; SSIM: Structural similarity index.

#### Acknowledgements

None.

#### Authors' contributions

All the authors contributed to the current work. JFC devised the study plan and led the writing of the article. ZBZ and ADZ conducted the experiments and collected the data. HYC and QZ performed the analyses, and CJF supervised the entire process and provided constructive advice. All authors read and approved the final manuscript.

#### Funding

This study was supported by the Natural Science Foundation of Shanxi (201701D21059), the Project of Key Research Base of Humanities and Social Sciences in Shanxi Colleges and Universities (20190130), the Art Science Planning Subject of Shanxi Province (2017F06), the Xinzhou Platform and Specialized Talents (20180601), and the 13th Five-year Plan of Education Science in Shanxi Province (GH-17059).

#### Availability of data and materials

All data for analysis in this study are included within the article.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup> School of Computer Science & Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China. <sup>2</sup> Department of Computer Science & Technology, Xinzhou Teachers University, No. 10 Heping West Street, Xinzhou 034000, China.

Received: 19 September 2019 Accepted: 22 January 2020

Published online: 29 January 2020

#### References

- Zhang N, Zhang Q, Feng W, Wang XW, Sun SL, Chai BL, et al. The deterioration identification system of ancient murals and its application in the Dunhuang Mogao Grottoes. *Dunhuang Res.* 2017;162:135–40 (in Chinese with an English abstract).
- Izzo FC, Falchi L, Zendri E, Giscontini G. A study on materials and painting techniques of 1930s Italian mural paintings: two cases by Mario Sironi and Edmondo Bacci in Venice. In: Pons MS, Shank W, López LF, editors. *Conservation issues in modern and contemporary murals*. England: Cambridge Scholar Publishing; 2015. p. 35–51.
- Sakr AA, Ali MF, Ghaly MF. Discoloration of ancient Egyptian mural paintings by streptomyces strains and methods of its removal. *Int J Conserv Sci.* 2012;3:249–58.
- Abdel-Halim ME, Sakr AA, Ali MF, Ghaly MF, Sohlenkamp C. Characterization of streptomyces isolates causing colour changes of mural paintings in ancient Egyptian tombs. *Microbiol Res.* 2013;168:428–37.
- Li J, Zhang H, Fan Z, He X, He SM, Sun MY, et al. Investigation of the renewed diseases on murals at Mogao Grottoes. *Herit Sci.* 2013;1:31. <https://doi.org/10.1186/2050-7445-1-31>.
- Jiao LJ, Wang WJ, Li BJ, Zhao QS. 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).
- Ren XK, Deng LK. Murals inpainting of the wavelet texture description algorithm based on scale space. *Comput Eng Sci.* 2014;36:2192–5 (in Chinese with an English abstract).
- Cao JF, Li YF, Zhang Q, Cui HY. Restoration of an ancient temple mural by a local search algorithm of an adaptive sample block. *Herit Sci.* 2019;7:39. <https://doi.org/10.1186/s40494-019-0281-y>.
- Wu M, Wang HQ, Li WY. Research on multi-scale detection and image inpainting of Tang dynasty tomb murals. *Comput Eng Sci.* 2016;52:169–74 (in Chinese with an English abstract).
- Li CY, Wang HQ, Wu M, Pan SC. Automatic recognition and virtual restoration of mud spot disease of Tang dynasty tomb murals image. *Comput Eng Sci.* 2016;52:233–6 (in Chinese with an English abstract).
- Bertalmio M, Sapiro G, Caselles V, Ballester C. Image inpainting. In: *Proceedings of conference on computer graphics and interactive techniques*. Washington, DC: Addison-Wesley Press; 2000. p. 417–24.
- Criminisi A, Pérez P, Toyama K. Object removal by exemplar-based inpainting. In: 2003 IEEE computer society conference on computer vision and pattern recognition (CVPR 2003). IEEE; 2003. p. 16–22.
- Pathak D, Krahenbuhl P, Donahue J, Darrell T, Efros AA. Context encoders: feature learning by inpainting. In: *IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE; 2016. <https://doi.org/10.1109/cvpr.2016.278>.
- Yang C, Lu X, Lin Z, Shechtman E, Wang O, Li H. High-resolution image inpainting using multi-scale neural patch synthesis. In: *IEEE conference on computer vision and pattern recognition (CVPR)*. IEEE; 2017. p. 1063–6919.
- Liu GL, Reda FA, Shih KJ, Wang TC, Tao A, Catanzaro B. Image inpainting for irregular holes using partial convolutions. In: *Proceedings of European conference on computer vision*. Munich: Springer Press; 2018. p. 89–105.
- Yu JH, Lin Z, Yang JM, Shen XH, Lu X, Huang TS. Generative image inpainting with contextual attention. In: *Proceedings of the IEEE computer society conference on computer vision and pattern recognition*. Los Alamitos: IEEE Computer Society Press; 2018. p. 5505–14.

17. Yan Z, Li X, Li M, Zuo W, Shan S. Shift-net: image inpainting via deep feature rearrangement. In: Proceedings of the European conference on computer vision. Munich: Springer Press; 2018. p. 3–19.
18. Zhang H, Hu Z, Luo C. Semantic image inpainting with progressive generative networks. In: Proceedings of the 26th ACM international conference on multimedia. New York: ACM; 2018. p. 1939–47.
19. Chen YZ, Hu HF. An improved method for semantic image inpainting with GANs: progressive inpainting. *Neural Process Lett.* 2018;49:1355–67.
20. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. In: Proceedings of the international conference on neural information processing systems (NIPS 2014). p. 2672–80.
21. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE T Pattern Anal.* 2014;39:640–51.
22. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd international conference on international conference on machine learning. Lille: JMLR; 2015. p. 448–56.

### Publisher's Note

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

**Submit your manuscript to a SpringerOpen<sup>®</sup> 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)

---