## VIRTUAL RESTORATION OF MISSING PAINT LOSS OF MURAL BASED ON GENERATIVE ADVERSARIAL NETWORK

Q. Wang <sup>1,2</sup>, M. Hou <sup>1,2</sup>, S. Lyu <sup>1,2,\*</sup>

<sup>1</sup> School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, No, 15 Yongyuan Road, Daxing District, Beijing, China – 2108521519024@stu.bucea.edu.cn, houmiaole@bucea.edu.cn, lvshuqiang@bucea.edu.cn <sup>2</sup> Beijing Key Laboratory for Architectural Heritage Fine Reconstruction & Health Monitoring, No.15 Yongyuan Road, Daxing District, Beijing, China

KEY WORDS: Mural, Generative adversarial network, Virtual restoration, dilated convolution, receptive fields,

#### **ABSTRACT:**

Mural painting is one of the important cultural heritage reflecting the historical migration of the nation. In order to inherit these precious historical and cultural heritage, how to non - destructively and digitally protect and restore the existing murals has become an urgent task. The use of computer - assisted restoration of murals can not only save manpower and material resources, but also avoid secondary damage to the murals.

However, most of the existing computer-assisted mural restoration algorithms use similar blocks with priority calculations and matching blocks in adjacent areas to guide mural restoration. There are some problems such as incoherent overall semantic structure, unnatural detail texture and inability to effectively repair large area missing remain to be solved. Aiming at the problems existing in the restoration of large area diseases such as paint loss and color fading in murals, we constructed a fine image restoration network model which based on generative adversarial network. A multi-scale dense matching repair network based on a generative adversarial network. The dense combination of dilated convolutions is used to improve the repair effect of detailed textures, Then, mean absolute Error, (Visual Geometry Group, VGG) feature matching, auto-guided regression, and geometric alignment are used as the loss function to guide the training of the generative network. Second, the discriminator with local and global branches is used to train the discriminant network, so that the repaired image is in the local and global content. Experiments were performed on the three mural data sets one by one. The results show that the network model can effectively restore the lines and faces in the murals. The images restored are not only coherent in semantic details, but also natural in color, which is conducive to the appreciation and display of murals. Thus, as one of the important directions of cultural heritage digital protection, the use of generative adversarial network in the digital restoration of ancient murals have been proved to be effective. It not only provides a reference for the true restoration of the murals but also means a lot to the preservation of murals.

### 1. INTRODUCTION

Due to the long-term natural weathering, man-made destruction and other factors, the murals suffered from various diseases such as paint loss and fading, which seriously affected the appearance and preservation of the murals. It is urgent to digitally protect and restore the murals to inherit these precious cultural heritages. Early approaches to image inpaintingare mostly based on patches of low-level features. a typical method, iteratively searches optimal corresponding author patches to fill in the holes. It can produce plausible re-sults when painting image background or repetitive tex-tures. However, it cannot generate pleasing results for caseswhere completing regions include complex scenes, faces, and objects, which is due to PatchMatch cannot synthesize new image contents, and missing patches cannot be found in remaining regions for challenging cases. we constructed a fine image restoration network model which based on adversarial network, and this model consists of a generator and a discriminator with two local branches.

Image restoration refers to filling the defective image in a local area so that the restored image is visually and semantically

coherent. Most of the early image repair algorithms used existing information to repair and complete the defect area. According to the core idea of algorithm, they can be divided into diffusion repair algorithms based on partial difference equations (PDE) (Tfc, A., & Js, B. . 2001). and texture synthesis based similar patch group and filling algorithm (Criminisi, A., P Pérez, & Toyama, K. . 2003). and iterative patch search and filling algorithm based on underlying features (Patch Match) (Barnes, C. . 2009). The diffusion repair algorithm based on PDE uses the known information around the missing area to extend from the periphery of the missing area to the center according to the characteristics of isophotoes, so that the central pixels of the missing area can obtain coherence and realize image information restoration. based similar patch group and filling algorithm based on texture synthesis calculates the priority of the missing area and the known area to determine the best matching block, and then fills it to the area to be repaired. Iterative patch search and fill algorithm based on underlying features, then iteratively search for the optimal fill patch in the global image, and then fill in the missing area to complete the repair.

<sup>\*</sup> Corresponding author: lvshuqiang@bucea.edu.cn

With the rapid development of deep learning, image restoration methods based on generative adversarial networks have also achieved remarkable success. Pathak (Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. 2016). proposed a context-encoder-based repair model, which predicts and repairs missing regions through autoencoders and counterlosses. Cai (Cai, N., Su, Z., Lin, Z., Wang, H., Yang, Z., & Ling, W. K. 2017). proposed a method to repair images using deep convolutional neural networks. The deep neural networks are usually difficult to train because they contain a large number of hidden layers. Therefore, most deep neural network models use BP structure to train the model, but the image repairing is essentially an unsupervised learning process.

The traditional algorithm can only repair a small part of the image to be restored. In order to complete and repair a larger area, Lazuka (Satoshi, Iizuka. , Edgar, Simo-Serra. , and Hiroshi, Ishikawa. , 2017:). used dilated convolution in the image repair network to obtain a larger receptive field. They used local and global consistent confrontation training to generate similar repair results. Wang (Wang, Y. , Tao, X. , Qi, X. , Shen, X. , & Jia, J. . 2018). proposed a generative multicolumn convolutional neural network (GMCNN) which can obtain better repairing results by using different sizes of convolution kernels and different receptive fields.

The goal of image restoration is to ensure that the generated image has a global semantic structure and fine texture structure. However, these existing methods do not consider the consistency between the output and the target. In order to solve the limitations of the above methods, we propose a mural image restoration network model based on a generative network. The model mainly consists of a generator and a discriminator with two local branches. In the generative network part of the model, we added a multi-scale dense fusion network in the middle of the convolutional layer. The features extracted by the convolutional layer are assigned to four convolutional layers with different expansion rates, and then the multi-scale features are extracted using dilated convolution with different expansion rates. Finally, in order to obtain dense multi-scale features, various sparse features are combined through accumulation Multi-scale features.

In order to train the generator more efficiently, in addition to the commonly used feature matching loss, a self-guided regression constraint loss was designed to calculate the feature difference map extracted by the convolutional layer. It was employed to estimate and correct the semantic layer information. A geometric alignment constraint loss was also designed to compensate and predict the distance between highlevel features and real features. In the discriminant network part of the model, a discriminator with global and local branches was used to ensure the consistency of local and global content. The repair effect was optimized through adversarial training.

## 2. MAIN BODY

#### 2.1 Traditional Method

In recent years, a variety of image restoration algorithms have been proposed. The traditional diffusion-based method (Ballester, C. , Bertalmio, M. , Caselles, V. , Sapiro, G. , & Verdera, J. . 2001). spreads the repaired area from the adjacent area to the area to be repaired. It fills the area with those areas in similar textures and colors. Patch-based methods, such as (Kwatra, V. , Essa, I. , Bobick, A. , & Kwatra, N. . 2005). search for related patches from non-destructive areas in an iterative manner to fill the defective areas. The image restoration technique based on texture synthesis is a restoration technique that synthesizes the restoration information of the defective area according to the information of the known small sample block area in the image. Sharma (Sharma, N. , & Mehta, N. .). proposed a representative restoration technology based on sample synthesis algorithm.

Taking this algorithm as an example, the principle is as follows:

Criminisi mural image restoration methodis shown in the Figure 1, where  $\Phi$  represents the known area,  $\Omega$  represents the area to be repaired,  $\partial\Omega$  represents the boundary between the known area and the damaged area,  $\Psi_P$  is the selected block to be repaired on the boundary, and  $\Psi_q$  is a matching block selected in the known area  $\Phi$  with q point as the center. The priority value of each pixel on  $\partial\Omega$  will be calculated firstly and then determining the repair order in the area  $\Omega$ . Then, the best matching block  $\Psi_q$  will be found to fill into the area  $\Omega$ .



Figure 1. The traditional mural image restoration method

# 2.2 Virtual restoration based on generative adversarial network

We propose a mural image restoration model based on the generative adversarial network consisting of a generator and a discriminator. We propose a dense multi-scale fusion generator, which has strong representation ability and can extract useful features. Compared with previous state-of-the-art methods, our generative image embedding framework achieves compelling visual results on challenging data sets. In order to ensure the dense convolution kernel while expanding the receiving domain, we propose our dense multi-scale fusion block. Specifically, the first convolution on the left reduces the channel of the input feature to 64 to reduce the parameters, and then sends these processed features to 4 branches to extract multi-scale features, through cumulative addition, We can get dense multi-scale features. In short, this basic block particularly enhances general

extended convolution and has fewer parameters than large kernels. Different from the previously used generative confrontation repair network (Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. 2017). we propose to use RaGAN (Jolicoeur-Martineau, A. 2018). To pursue more realistic generated images. The discriminator also considers the consistency of the global image and the local image.

The proposed inpainting model is trained in an end-to-end way. (Hui, Z. , Li, J. , X Wang, & X Gao. 2020). Given an input image with hole I<sub>in</sub>, its corresponding binary mask M (value 0 for known pixels and 1 denotes unknown ones), the output I<sub>out</sub> predicted by the network, and the ground-truth image I<sub>gt</sub>. We take the input image and mask as inputs, i.e., [I<sub>in</sub>,,M].

#### 2.2.1 Network Structure

The mural restoration model based on the generative confrontation network is composed of two parts, a generator and a discriminator with two local branches. After building the network, the repaired area is obtained and is input to the discriminant network as an input image. An optimized repair network structure is obtained through continuous confrontation training.

To generative model part, we used the expansion convolution of different convolution kernels to form a multi-scale dense repair network to obtain a larger receptive area (Hui, Z., J Li, Gao, X., & Wang, X. 2019). The first convolution in the network structure was used to reduced the channel of the input feature to 64 to reduce the parameters. These features were assigned to dilated convolutions with different convolution kernel sizes to extract multi-scale features. Through accumulation, various sparse multi-scale features were combined to obtain dense multi-scale features.



Figure 2. Generative Model

In the discriminative model part, in addition to the feature matching loss, we designed a self-guided regression constraint loss, which is used to calculate the feature difference map extracted by the convolutional layer. It can also be used to estimate and modify the semantic layer information. A geometric alignment constraint loss is also designed to compensate and predict the distance between high-level features and real features. Finally, a discriminator with global and local branches is used to ensure the consistency of local and global content and improve the quality of image generation.



Figure 3. Discriminative Model

#### 2.2.2 Data Set

For the lack of public mural data set, we collected the mural image data of Qutan Temple in Haidong City, Qinghai Province, China. They were processed to create a mural restoration data set which includes 31603 scene images of the mural restoration dataset, 30300 scene images of the mural linear feature restoration dataset, and 1709 scene images of the mural human face dataset. Some of them are shown in Figure 4 to 6.



Figure 4. mural restoration dataset



Figure 5. mural linear feature restoration dataset



Figure 6. mural human face dataset

#### 2.2.3 Loss Function

To better train this efficient generator, except for frequentlyused VGG feature matching loss, we designed a novel selfguided regression loss for concentrating on uncertain areas and enhancing the semantic details. Besides, we devised a geometrical alignment constraint item to compensate for the pixel-based distance between prediction features and groundtruth ones. We also employ a discriminator with local and global branches to ensure local-global contents consistency. To further improve the quality of generated images, discriminator feature matching on the local branch is introduced, which dynamically minimizes the similarity of intermediate features between synthetic and ground-truth patches.

Taking the linear feature repair model as an example, the loss function is shown in Figure 7.



Figure 7. Linear feature repair model loss function

## 2.3 Experiments

Experiments were carried out one by one on the three fresco restoration data sets.

#### 2.3.1 Experimental Settings

The created mural dataset is used to train and test the network model, which was built under the PyTorch framework and trained by NVIDIA GeoForce RTX 2080.

We added a  $64 \times 64$  binary mask (0 means known pixel, 1 means unknown pixel) on a random position of the original mural image during the training. Firstly, this mask was used to simulate the missing disease in the mural paint layer to form a masked image as the input image of the repair model. Then, the generated image which obtained through the generation network sampling and prediction was input to the discrimination network. And the judgment network will feed the result back to the generation network after identifying whether the input image comes from a real image or a generated image through local and global judgement. Further, after continuous confrontation training, continuous learning, and parameter adjustment between the generation network and the discrimination network, the optimal state is finally reached "Nash equilibrium" to get the final repaired output image.



Figure 8. Visual results on dataset

## 2.3.2 Qualitative Comparisons

It can be seen from Figure 8 that after repairing by the network model based on the generative confrontation network, the defect area in the sample data has been repaired well. The repaired area has a better integration with other areas of the painting body. The repaired boundary is also relatively natural and smooth.

In addition, mean square error (MSE) and peak signal-to-noise ratio (PSNR) are used to quantitively evaluate the images before and after restoration. Among them, the larger the PSNR value, the closer the images before and after the restoration. The smaller the error, the better the corresponding restoration effect.

Sample	MSE	PSNR
Sample 1	0.8087	49.0530
Sample 2	0.8949	48.6129
Sample 3	1.1141	48.1974
Sample 4	1.0880	48.2340

 Table 1. Quantitative results (regular mask) onthree testing datasets.

#### 3. CONCLUSIONS

We proposed a dense multi-scale fusion network with autoguided regression loss and geometric alignment constraints to perform fine restoration of images, which improves the quality of image restoration. In order to extract better features, a dense multi-scale fusion network was used. With the help of multiple counter losses, the repair of semantic structure becomes easier. In addition, geometric alignment constraints are introduced between the generated image and the real image, which improves the rationality of the repair results.

Experiments were performed on the three mural restoration data sets one by one. The results show that the network model can effectively restore the lines and faces in the murals. The images restored are not only coherent in semantic details, but also natural in color, which is conducive to the appreciation and display of murals. Therefore, as one of the important directions of cultural heritage digital protection, the use of generative adversarial network in the digital restoration of ancient murals have been proved to be effective. It not only provides a reference for the true restoration of the murals but also means a lot to the preservation of murals.

#### ACKNOWLEDGEMENTS

The authors would like to acknowledge of Youth Beijing Scholar for its support of this work, and thank the professor, Shuqiang Lyu, Miaole Hou ,who has offered many precious suggestions in academic study.

#### REFERENCES

Tfc, A., & Js, B. (2001). Nontexture inpainting by curvaturedriven diffusions-sciencedirect. Journal of Visual Communication and Image Representation, 12(4), 436-449.

Criminisi, A., P Pérez, & Toyama, K. . (2003). Object Removal by Exemplar-Based Inpainting. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), 16-22 June 2003, Madison, WI, USA. IEEE.

Barnes, C. . (2009). Patchmatch:a randomized correspon-dence algorithm for structural image editing. Proc.ACM SIGGRAPH, 28.

Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: feature learning by inpainting. IEEE.

Cai, N., Su, Z., Lin, Z., Wang, H., Yang, Z., & Ling, W. K. (2017). Blind inpainting using the fully convolutional neural network. Visual Computer International Journal of Computer Graphics, 33(2), 249-261.

Satoshi, Iizuka. , Edgar, Simo-Serra. , and Hiroshi, Ishikawa. , 2017: Globally and locally consistent image completion. ACM Transactions on Graphics (TOG), 36(4):107:1–107:14.

Wang, Y., Tao, X., Qi, X., Shen, X., & Jia, J. (2018). Image Inpainting via Generative Multi-column Convolutional Neural Networks. Ballester, C. , Bertalmio, M. , Caselles, V. , Sapiro, G. , & Verdera, J. . (2001). Filling-in by joint interpolation of vector fields and gray levels. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society.

Kwatra, V., Essa, I., Bobick, A., & Kwatra, N. (2005). Texture optimization for example-based synthesis. ACM Transactions on Graphics, 24(3), p. 795-802.

Sharma, N., & Mehta, N. . Region Filling and Object Removal by Exempeler Based Image Inpainting.

Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. 2017. Improved training of wasserstein GANs.

Jolicoeur-Martineau, A. (2018). The relativistic discriminator: a key element missing from standard gan.

Hui, Z. , Li, J. , X Wang, & X Gao. (2020). Image fine-grained inpainting.

Hui, Z., J Li, Gao, X., & Wang, X. . (2019). Progressive perception-oriented network for single image super-resolution.