

Image and Graph Restoration Dependent on Generative Adversarial Network Algorithm

Yuanhao CAO

Abstract: As a research hotspot in the field of deep learning, image inpainting is of great significance in people's real life. There are various problems in the existing image inpainting algorithms, resulting in the visual inability to meet people's requirements. In view of the defects of the existing image inpainting algorithms, such as low accuracy, poor visual consistency and unstable training, in this paper the missing content is generated by adjusting the available data. For a data set, first analyze the samples in the data set into sample points in the probability distribution, quickly generate a large number of forged images by using the generation countermeasure network, search the code of the closest damaged image, and then infer the missing content through the generation model. Combining the semantic loss function and perceptual loss function, the problem that the gradient is easy to disappear is solved. Experiments show that the algorithm improves the accuracy of image restoration, can generate more realistic repaired images, is suitable for the repair of various types of images, and realizes the realism of photos.

Key words: Deep Learning; Generative Adversarial Network; Generative Model; Image Restoration; Loss Function

1 INTRODUCTION

With the development of artificial intelligence (AI), digital image restoration has become a crucial task in computer vision [1-2]. Digital image restoration is an image processing technology where the undamaged information in the image is used to restore the missing content in the image or remove the specific information in the image under the premise that the quality of the image and its natural effects are not damaged [3-4]. In other words, the core challenge in the restoration work is to synthesize pixels that are visually realistic and semantically reasonable for the missing areas so that they can be consistent with the existing pixels [5]. Due to the extensive presence of images, digital image restoration technology has been applied in the protection of works of art, restoration of lost or damaged information in old photos, hiding incorrect information in pictures and videos, removing texts in pictures and videos, performing image-based rendering, and many other applications in computer photography [6-7]. Hence, the restoration of digital images has become a hot spot in image processing technology [8].

In this paper, the missing content is generated by adjusting the available data based on the method of generative adversarial networks. For a data set, the samples in the data set are first parsed into sample points in the probability distribution, and the generative adversarial network is used to generate a large number of fake images quickly and search for the code of the closest damaged image. Subsequently, this code is used to generate a model to infer the content that is missing. On this basis, the semantic loss function and the perceptual loss function are combined in this paper, and the unsaturated region is expanded by improving the activation function *Sigmoid* function, which has resolved the issue that the gradient is prone to vanishing. The experiments have demonstrated that the method proposed in this paper can predict the information on a large area of missing areas in the image successfully, which has achieved the sense of reality in the photo and produced clearer and more consistent results than the previous methods.

2 GENERATIVE ADVERSARIAL NETWORKS

Generative adversarial networks (GAN) are a highly popular research direction in the artificial intelligence (AI) field at present. The generator G and the discriminator D are composed of a deep neural network, respectively, and the counter loss function is optimized through continuous counter training as the following:

$$\min_G \max_D \left\{ f(D, G) = E_{x \sim p_{data}(x)} [\log D(x) + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \right\} \quad (1)$$

The calculation process and structure of the generative adversarial network are shown in Fig. 1 as the following.

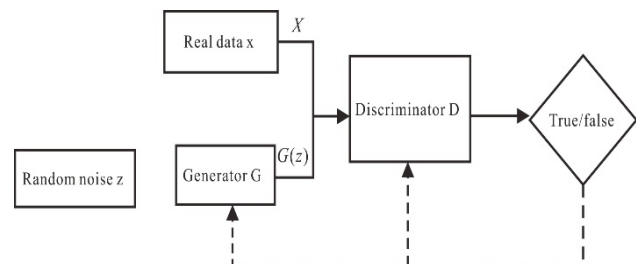


Figure 1 Calculation process and structure of GAN

Through continuous iteration, the image that best matches the undamaged part from the generated samples is identified in this paper to accomplish the image restoration task perfectly.

3 IMAGE RESTORATION

For damaged images, in order to restore them well, it is necessary to fill in the missing part with a relatively large area in the image. In a high-quality restoration result, it should not only look realistic in the damaged area but is also semantically consistent with the surrounding images. Digital image restoration can be divided into two types according to the size of the damaged area of the image for images with small damage scale, such as scratches, etc. For this kind of repair problem, the following two methods can

be adopted: one is the digital image repair method based on partial differential equation. Its main idea is to use the diffusion equation in physics to spread the information around the area to be repaired to the repair area to achieve the purpose of repair. The other is the restoration method based on variational principle. By establishing the prior model and data model of the image, the restoration problem is transformed into a variational problem of functional extremum. For the image with large damaged scale, there are two methods: one is the repair technology based on image decomposition, the idea is to decompose the image into structural part and texture part, the structural part is realized by small-scale repair algorithm, and the texture part is filled by texture synthesis method, and then the two are combined and superimposed to obtain the repair result. The other is a texture synthesis repair algorithm based on sample blocks. It selects a sample block with appropriate size from the area to be repaired, and looks for the closest texture matching block around the area to be repaired to replace the texture block.

3.1 Semantic Restoration Based on Constrained Images

The generator G transforms the random noise sequence into a high-dimensional image through random noise z by using the inverse convolutional neural network, as shown in Fig. 2. Through the adversarial training of the generator and the discriminator, an image that imitates the sample is generated. For an image, it is coded first, with y indicating the damaged image, and M indicating the binary code of the defect area. For a damaged image, images y and M are expressed in Fig. 3 as the following. Hence, the goal of this paper is to identify the code that is closest to the damaged image so that the damaged image can be restored and is semantically consistent with the surrounding images.

$$\hat{z} = \arg \min_z \{ \ell_c(z|y, M) \} \quad (2)$$

In the above equation, ℓ_c stands for the semantic loss function, which is used to constrain the relationship between the input damaged image y and the binary mask M corresponding to the generated image. As the defect area of the image takes up a large proportion of the raw image, the remaining undamaged part is fully used. However, for the damaged region with a relatively large area in the center of the image, it is obviously not the optimal choice to treat each pixel equally.

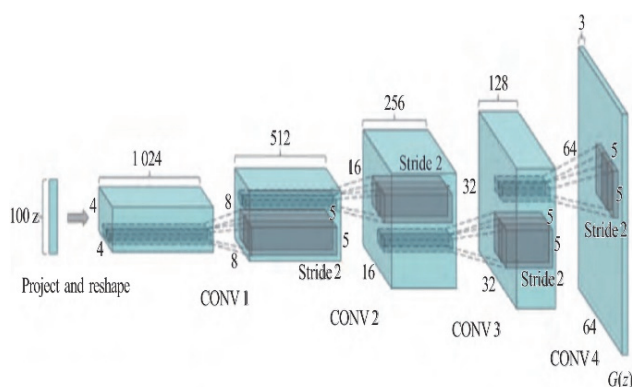


Figure 2 Implementation of the inverse convolutional neural network

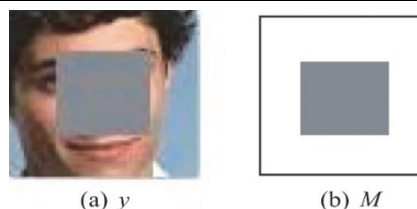


Figure 3 Damaged image y and binary mask image M

In order to fill in the missing content in the image, it is necessary to make full use of the pixels in the undamaged area and take the distance relationship between the undamaged pixels and the damaged area into consideration. It is apparent that the pixels that are closer to the damaged area have a higher correlation with the pixels to be filled. Hence, if only the remaining pixels are used and each pixel is treated equally to fill the picture, it will not be conducive to achieving a sound restoration result in this paper. To identify a better encoding, more attention should be paid to pixels that are close to the missing area. In order to achieve this goal, a method of weighting pixels is used in this paper. That is, allow the pixels close to the missing area to play more roles in the image restoration process, and let the pixels far away from the missing area play fewer roles in the image restoration process, which is denoted by the weight W as the following.

$$W_i = \begin{cases} \sum_{j \in N(i)} \frac{(1 - M_j)}{|N(i)|}, & M_i \neq 0 \\ 0, & M_i = 0 \end{cases} \quad (3)$$

In the above equation: i stands for the pixel; W_i stands for the weight of the pixel i ; $N(i)$ stands for the set of neighboring pixels of the pixel i ; $|N(i)|$ stands for a base of $N(i)$. In the experiment, in order to address the issue of over-fitting [15] due to excessive training parameters, L_1 regularization [16] is introduced herein, and the text loss function is defined as the L_1 norm between the restored image and the undamaged part, as shown in the following:

$$\ell_c(z|y, M) = \|W \Theta(G(z) - y)\|_1 \quad (4)$$

In the above equation: Θ stands for the multiplication operation between the elements.

3.2 Perceptual Loss Function

The perceptual loss function is used to punish the unreal images generated. In the generative adversarial network, the discriminator D is trained to distinguish between real samples and generated ones. On this basis, the perceptual loss function is defined so that the discriminator D can have a better performance. The perceptual loss function is expressed as the following:

$$\begin{aligned} \ell_p(z) &= \lambda \log(1 - D(G(z))) \\ \hat{z} &= \arg \min_z \{ \ell_c(z|y, M) + \ell_p(z) \} \end{aligned} \quad (5)$$

In the above equation: λ stands for a parameter that

balances the two loss functions; ℓ_p stands for the perceptual loss function; the random noise z stands for reversed convolution and is iteratively updated to make the restored image look more realistic and vivid. If the perceptual loss function is not added, the mapping from y to z may generate a picture that is incomprehensible. In the subsequent section, we will show some unstable examples to illustrate this point. The experimental results of the semantic loss function and the perceptual loss function will be combined with the findings of the experiments without using the perceptual loss function for comparison.

As shown in Fig. 4, a large number of samples are generated through the adversarial learning of the generative adversarial network, and z is continuously iterated and upgraded so that the network can generate the image closest to the original sample.

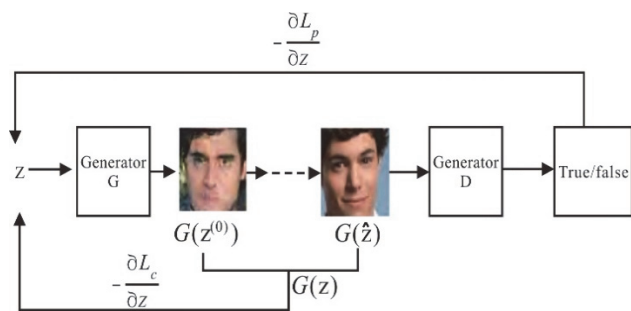


Figure 4 Image restoration process

3.3 Improved Activation Function

The activation function is used to add nonlinear factors in the neural network to enhance the expressive capability of the network. In this paper, the activation function used for the loss function is the *Sigmoid* function (S-type function), and its function expression is shown in Eq. (6) as the following.

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

The original GAN is obtained based on information theory. Firstly, let us review the concept of entropy:

$$H = -k \sum_{i=1}^n p_i \log p_i \tag{7}$$

In the above equation: k stands for a normal number. In the generative adversarial network, p_i is required to be a valid value all the time, rather than being close to 0 or 1. Based on the *Sigmoid* activation function, the input value can meet p_i . However, the *Sigmoid* function has a narrow unsaturated region, and the maximum value of the derivative of the *Sigmoid* function is 0.25. When the parameter product operation is carried out in the multi-layer neural network, the final result will show a decrease exponentially, which can lead to the vanishing of the gradient.

Hence, an improved *Sigmoid* activation function is put forward in this paper to expand the unsaturated region, as shown in Eq. (8) below.

$$ReSigmoid(logits) = \frac{1}{1 + e^{-\alpha \cdot logits}} \tag{8}$$

In the above equation: α stands for a hyper-parameter. The positive part of the derivative of the improved *Sigmoid* function is maintained near 1 through the adjustment of the parameters. Hence, the gradient vanishing problem is improved, and $\alpha = 0.1$ is set in the experiment.

3.4 Image Restoration

The purpose of the image restoration network module is to ensure that the original color and texture of the image can be restored to the greatest extent in terms of the color and spatial position of each reconstructed pixel. The total loss function L_{inp} of the image restoration network module is defined in Eq. (9).

$$L_{total}^{inp} = 2L_{valid} + 12L_{hole} + 0.04L_{per} + 100(L_{style}^1 + L_{style}^2) + 100L_{adv} + 0.3L_{var} \tag{9}$$

The weight of each loss item is determined after the results of 50 independent experiments are analyzed.

In the above equation, the restoration loss in the unmasked area is defined in Eq. (10), and the Manhattan distance between the restored image and the unmasked area of the real image is taken as the restoration loss. In the equation, I_{dam} stands for the damaged image, M stands for the irregular binary mask (the corresponding area to be restored in the mask is 0, and the others are 1), I_{inp} stands for the restoration result image, and I_{real} stands for the real undamaged image. The restoration loss function of the masked area is shown in Eq. (11) as the following:

$$L_{valid} = \|M \times (I_{inp} - I_{dam})\|_1 \tag{10}$$

$$L_{hole} = \|(1 - M) \times (I_{inp} - I_{dam})\|_1 \tag{11}$$

The definition of perceptual loss is shown in Eq. (12) as the following. In this paper, different convolutional feature layers of multiple pre-trained networks are used to obtain the feature perception loss between the restored result image and the real image, and the perception of the area to be restored loss is enhanced accordingly. In Eq. (12), I_{com} stands for the real image of the non-missing area plus the predicted image of the missing area, as shown in Eq. (13) below. In the equation, m stands for the number of pre-trained networks used, n stands for the number of convolutional feature layers used, Ψ_j^i stands for the j -th convolutional features of the i -th pre-trained network, and ω_i stands for the weight of the perception loss of the i -th pre-trained network. After 50 independent experiment comparisons, the feature diagrams output by the pool1, pool2, and pool3 layers of VGG16 and the conv1, pool2 and pool3 layers of Dense Net are used as the perception layer of the generation network in the experiment of this paper. The perceived loss is calculated on this basis. The parameters of the pre-trained network are not involved in

training but are only used to obtain the loss value. The weighted sum of the perceptual losses obtained by the two pre-trained networks is taken as the final perceptual loss. The weight settings adopted in the experiment of this paper are shown in Eq. (14) as the following.

$$L_{per}^* = \sum_{i=0}^{m-1} \omega_i \left(\sum_{j=0}^{n-1} \left\| \Psi_j^i(I_{inp}) - \Psi_j^i(I_{real}) \right\|_1 + \sum_{j=0}^{n-1} \left\| \Psi_j^i(I_{com}) - \Psi_j^i(I_{real}) \right\|_1 \right) \quad (12)$$

$$I_{com} = M \times I_{real} + (1 - M) \times I_{inp} \quad (13)$$

$$L_{per} = L_{per}^{VGG-16} + 30L_{per}^{DenseNet} \quad (14)$$

4 IMPLEMENTATION DETAILS OF THE EXPERIMENT

4.1 Specific Content of the Experiment

The experiment is implemented based on the deep learning framework Tensorflow, and the Python language is used as the primary implementation language. There are a total of 1000 pictures in the test set, 500 males and 500 females. These pictures are cropped to a size of 64×64 , and they contain different expressions. In the network, the size of the input image is $64 \times 64 \times 3$, followed by a convolutional layer and a pooling layer to reduce the image to half of its original size. In addition, the number of channels is twice that of the previous layer, which passes through 3 convolutional layers and 3 pooling layers, and the output layer includes two types of softmax classifiers. In order to train the DCGAN model, the Adam optimizer [24] is used in this paper for optimization. Through the adversarial training of generator G and discriminator D , their capacities are continuously improved so that the generated pictures are closer to the real samples.

In the image restoration stage, the original sample is coded using a size of 32×32 , which is taken as the damaged image. In the experiment, $\lambda = 0.003$ is selected.

The optimal \hat{z} in the latent space is identified by using the method of back-propagation, and the Adam optimizer is used in the restoration process. In addition, z is limited to $(-1, 1)$. In each iteration, it is observed that a stable result is generated gradually.

In this experiment, the following two traditional algorithms are compared: the total variation model and the curvature driven diffusion model. In addition, based on the model LC that adds the semantic loss function to the generative adversarial network, the model LC + LP that combines the semantic loss function with the perceptual loss function, and the model LC + LP (*Sigmoid*) of the improved activation function *Sigmoid* function, the restoration results are compared. In order to better reflect the effect of the restoration, the original images are compared as well, as shown in Fig. 5 below.

The restoration time of the method proposed in this paper is compared with the restoration time of the other methods, and the results are shown in Tab. 1. For a single picture, the restoration time of the method proposed in this paper and the LC and LC + LP methods is less than that of

the TV and CDD methods. As the GPU processor is applied, the LC + LP method takes longer to restore pictures than the LC method. The reason is that two loss functions are used, and it is necessary to perform training on more parameters. Compared with the LC + LP method, the method proposed in this paper takes longer to restore pictures, but the restoration effect is superior.

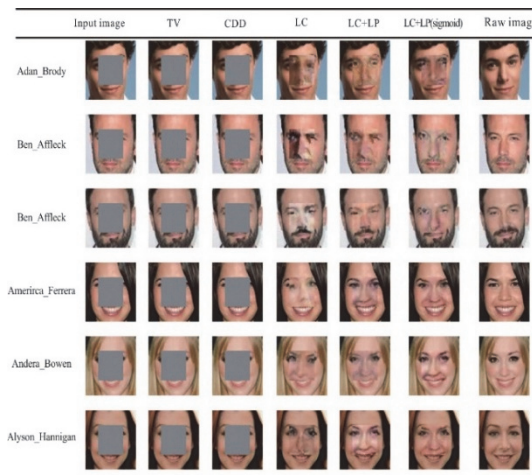


Figure 5 Comparison of image restoration results

4.2 Quality Evaluation

In this paper, the restoration effects of various models are compared by using the peak signal to noise ratio (*PSNR*) and structural similarity index (*SSIM*). For two $m \times n$ monochrome images I and K , if one is the noise approximation of the other, then the mean square error is defined as the following:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (15)$$

Table 1 Comparison of restoration time between the method proposed in this paper and the other comparative methods

Method	TV	CDD	LC	LC+LP	LC+LP (<i>Sigmoid</i>)
Processor	CPU	CPU	GPU	GPU	GPU
Duration to restore/s	113.26	98.17	0.87	0.95	1.15

The peak signal-to-noise ratio (*PSNR*) is defined as the following:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (16)$$

Structural similarity *SSIM* is an index to measure the similarity of two images.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (17)$$

In the above equation: μ_x stands for the mean value of x ; μ_y stands for the mean value of y ; σ stands for the variance of x ; σ_{xy} stands for the covariance of x and y ; $c_1 =$

$(k_1L)^2$ and $c_2 = (k_2L)^2$ are two constants, which are used to maintain stability; L stands for the dynamic range of the pixel value; $k_1 = 0.01$, $k_2 = 0.03$.

The *PSNR* and *SSIM* are used to compare the image restoration results of different models; the larger their scores are, the better the image restoration effect is. The quality evaluation scores are shown in Tab. 2 as the following, in which the input image, the total variation model TV, and the curvature driven diffusion model CDD are compared, respectively. In addition, based on the model LC where the semantic loss function is added to the generative adversarial network, the model LC + LP that combines the semantic loss function and the perceptual loss function, and the values of *PSNR* and *SSIM* of the model LC + LP (*Sigmoid*) for the improved activation function *Sigmoid* function on this basis, it can be observed that the total variation method and curvature driven diffusion have a poor restoration effect on images with a relatively large damaged area. However, based on the model that uses the semantic loss function, the restored images are relatively blurry. The effect of model restoration that combines the semantic loss function and the perceptual loss function is superior to that of using the semantic loss function alone. From the experimental data, it can be seen that the loss function model that applies the improved activation function has an optimal effect.

Table 2 Quality evaluation score scale

	Input image	TV	CDD	LC	LC+LP	LC+LP (<i>Sigmoid</i>)
<i>PSNR</i>	18.19	18.21	18.18	16.39	21.86	22.01
<i>SSIM</i>	0.763	0.771	0.761	0.876	0.934	0.942

5 CONCLUSIONS

Based on the method of generative adversarial networks, the missing content can be generated by adjusting the available data in this paper. In addition, the semantic loss function and the perceptual loss function are used to achieve a better restoration effect than that by using the semantic loss function alone. The unsaturated area is also enlarged by improving the activation function *Sigmoid* function, which has resolved the problem that the gradient is prone to vanishing. Moreover, the method proposed in this paper is compared with two traditional image restoration methods, that is, the full variation model and the curvature driven diffusion model. The results have demonstrated that the method proposed in this paper can predict and restore the information in the image with a large area of loss successfully and obtain clearer and more coherent results. The experimental results suggest that the proposed method has an excellent performance in the restoration of challenging images in practice. However, the method proposed in this paper is only designed for the restoration of rectangular damaged images, and further research is required for the restoration of actual irregular damaged images.

6 REFERENCES

- [1] Chen, W., You, J., Chen, B., Pan, B., Li, L., Pomeroy, M., & Liang, Z. (2018). A sparse representation and dictionary learning based algorithm for image restoration in the presence of rician noise-science direct. *Neurocomputing*, 286, 130-140. <https://doi.org/10.1016/j.neucom.2018.01.066>
- [2] He, L., Cui, G., Feng, H., Xu, Z., Li, Q. & Chen, Y. (2015). The optimal code searching method with an improved criterion of coded exposure for remote sensing image restoration. *Optics Communications*, 338, 540-550. <https://doi.org/10.1016/j.optcom.2014.11.021>
- [3] Hao, Y., Xueping, L., & Leiting, C. (2016). Solving adaptive image restoration problems via a modified projection algorithm. *Mathematical Problems in Engineering*, 2016(pt.4), 1-11. <https://doi.org/10.1155/2016/6132356>
- [4] Liu, X. (2019). Total generalized variation and wavelet frame-based adaptive image restoration algorithm. *The Visual Computer*, 35(12), 1883-1894. <https://doi.org/10.1007/s00371-018-1581-z>
- [5] Lampert, C. & Breuel, T. (2015). Objective quality measurement for geometric document image restoration (extended abstract). *Applied Optics*, 54(20), 6155-6161.
- [6] Memik, S. O., Member, S., Katsaggelos, A. K., & Sarrafzadeh, M. (2016). Analysis and fpga implementation of image restoration under resource constraints. *IEEE Transactions on Computers*, 52(3), 390-399. <https://doi.org/10.1109/TC.2003.1183952>
- [7] Ren, D., Zhang, H., Zhang, D., & Zuo, W. (2015). Fast total-variation based image restoration based on derivative alternated direction optimization methods. *Neurocomputing*, 170(C), 201-212. <https://doi.org/10.1016/j.neucom.2014.08.101>
- [8] Sevcik, J., Smidl, V., & Sroubek, F. (2018). An adaptive correlated image prior for image restoration problems. *IEEE Signal Processing Letters*, PP(99), 1-1. <https://doi.org/10.1109/LSP.2018.2836964>
- [9] Wang, W., Wu, C., & Tai, X. C. (2020). A globally convergent algorithm for a constrained non-lipschitz image restoration model. *Journal of Scientific Computing*, 83(1), 1-29. <https://doi.org/10.1007/s10915-020-01190-4>
- [10] Xia, Y., Leung, H., & Kamel, M. S. (2016). A discrete-time learning algorithm for image restoration using a novel l2-norm noise constrained estimation. *Neurocomputing*, 198(JUL.19), 155-170. <https://doi.org/10.1016/j.neucom.2015.06.111>
- [11] Patwary, N. & Preza, C. (2015). Image restoration for three-dimensional fluorescence microscopy using an orthonormal basis for efficient representation of depth-variant point-spread functions. *Biomedical Optics Express*, 6(10), 3826-3841. <https://doi.org/10.1364/BOE.6.003826>
- [12] Yu, Y., Liu, M., Feng, H., Xu, Z., & Li, Q. (2020). Split-attention multiframe alignment network for image restoration. *IEEE Access*, PP(99), 1-1. <https://doi.org/10.1109/ACCESS.2020.2967028>
- [13] You, J., Jiao, Y., Lu, X., & Zeng, T. (2019). A nonconvex model with minimax concave penalty for image restoration. *Journal of Scientific Computing*, 78(2), 1063-1086. <https://doi.org/10.1007/s10915-018-0801-z>
- [14] Yang, X. & Wang, L. (2017). Fast half-quadratic algorithm for image restoration and reconstruction. *Applied Mathematical Modelling*, 50(oct.), 92-104. <https://doi.org/10.1016/j.apm.2017.05.031>
- [15] Zha, Z., Yuan, X., Wen, B., Zhang, J., & Zhu, C. (2020). Image restoration using joint patch-group based sparse representation. *IEEE Transactions on Image Processing*, PP(99), 1-1. <https://doi.org/10.1109/TIP.2020.3005515>

Contact information:

Yuanhao CAO
North China University of Water Resources and Electric Power,
Henan, Zhengzhou, 450046, China
E-mail: ppxhappy@126.com