Super-Resolution of Defocus Thread Image Based on Cycle Generative Adversarial Networks

Pengfei Jiang, Wangqing Xu, Jinping Li

Super-Resolution of Defocus Thread Image Based on Cycle Generative Adversarial Networks

Pengfei Jiang[123][0000-0001-9334-8636], Wangqing Xu[123] and Jinping Li[123*]

[1] School of Information Science and Engineering, University of Jinan, Jinan 250022, China
[2] Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan 250022, China
[3] Shandong College and University Key Laboratory of Information Processing and Cognitive Computing in 13th Five-Year, Jinan 250022, China
ise_lijp@ujn.edu.cn

Abstract. The dual camera calibration measurement method can realize low-cost and high-precision bolt dimension measurement by using two microscope cameras. But the height difference between the thread crest and root exceeds the depth of field, and the thread image becomes defocus, which seriously affects the measurement accuracy. For this reason, a super-resolution method for defocus thread image based on cyclic generative adversarial networks is proposed. We collected focus thread images and defocus thread images as training data. Two encoders are used in the generation network to extract image defocus features and content features. And a sub-pixel convolution layer is added to the decoder to achieve image super-resolution. A loss function based on adversarial loss and cycle-consistent loss is constructed to realize unsupervised training of the network, thereby achieve super-resolution of defocus thread images. The experimental results show that, in the simulated defocus images, the method has superiority in image detail preservation, sharpness improvement and peak signal to noise ratio. In the bolt dimension measurement task, it can effectively reconstruct the clear thread image and thus provide the measurement accuracy to 0.01mm.

Keywords: Bolt, Dimension Measurement, Defocus Image, Super-Resolution, Cyclic Generative Adversarial Networks.

1 Introduction

Bolts are an important part of industrial products and represent the country's basic industrial level. Bolt dimension parameters determine the degree of stability of industrial products, among which the major diameter, minor diameter and pitch are the most important. At present, the dimension measurement methods of bolt are mainly based on three measurement methods: manual, laser and machine vision. Among them, manual measurement method is the most common, but this method has low accuracy, poor consistency, time-consuming and labor-intensive, and easily leads to bolt damage. The laser measurement method has high precision, but the cost is very

high, so it is difficult to apply in a large range. The bolt dimension measurement method based on machine vision collects the bolt image using a camera, then establishes the conversion model between the size in pixel and the size in millimeter, and finally locates the thread edge to calculate the bolt size in pixel, and combines the conversion model to obtain the bolt size in millimeter. Compared with other measurement methods, it has significant advantages of high precision, high efficiency, low cost, and non-contact. Modern industry has higher and higher requirements for bolt size measurement accuracy, and the accuracy requirements of some enterprises have reached 0.01mm. But limited by the spatial resolution of the camera and the field of view, if the diameter of the bolt exceeds 10mm, it can't achieve high measurement accuracy. The width of the lens field of view of the microscope camera is about 3 mm. Using the microscope camera and the dual-camera calibration dimension measurement method [1,2] to realize the bolt dimension measurement can ensure that the measurement accuracy can reach more than 0.01mm. Due to the small depth of field of the microscope camera, the height difference between the thread crest and root bottom exceeds the range of the depth of field, and there is a certain degree of defocus blur on the edge of the thread in the collected bolt image. This will widen the bolt edge transition area, cause the edge location algorithm to be misaligned, and severely reduce the bolt dimension measurement accuracy. Therefore, removing defocus blur and improving thread image quality are of great significance for bolt dimension measurement. The schematic diagram of the bolt dimension measurement and the thread image are shown in Fig. 1.

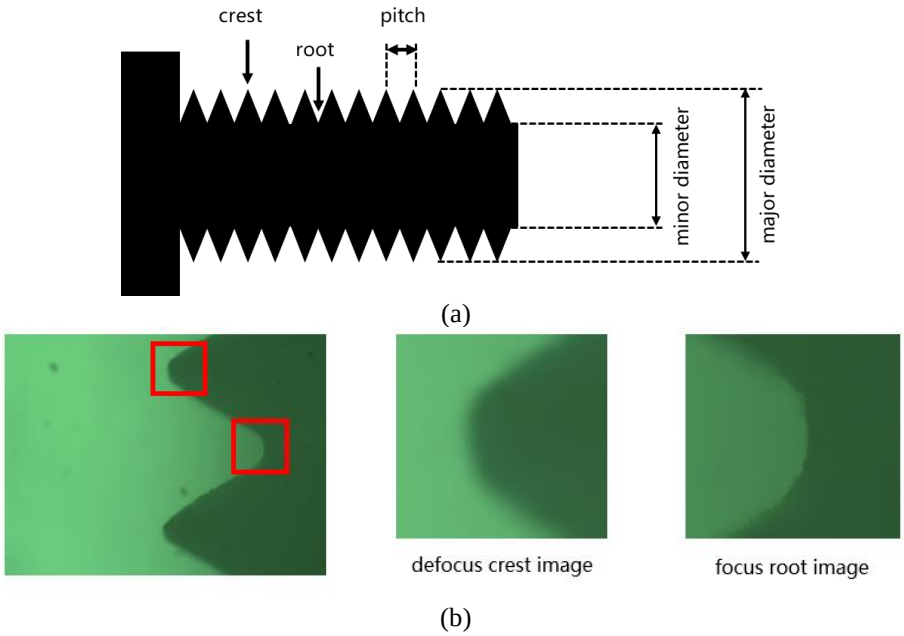


Fig. 1 The schematic diagram of the bolt dimension measurement and the thread image.
(a) Bolt structure and bolt dimension measurement items.
(b) Thread image taken by microscope camera.

The current methods to improve image quality mainly include traditional deblurring methods based on defocus models, image deblurring methods based on deep learning, and super-resolution (SR) methods based on deep learning. The traditional deblurring method based on the defocus model represents the defocus blur process as the result of the convolution of the original clear image and the point spread function under random noise. The point spread function can be approximately described as the disc defocus model and the gaussian defocus model [3,4]. These methods first estimate the point spread function, and then uses inverse filtering [5], Wiener filtering [6], Lucy-Richardson [7] and other algorithms to estimate a clear image. These methods can improve the image quality to a certain extent, but they are very computationally intensive and prone to ringing. Not only does it fail to recover effectively, but it also caused damage to a certain extent, and it is difficult to apply it to dimension measurement tasks. The image deblurring method and super-resolution method based on deep learning use convolutional neural network (CNN) to directly learn the best mapping of the image from the blurry domain to the clear domain, which can effectively restore the high-frequency details of the image. The difference between the super-resolution method and the image deblurring method is that the super-resolution method can improve the defocus image quality and increase the image resolution, which can theoretically increase the single pixel accuracy of the dimensional measurement system and improve the dimensional measurement accuracy, and is more valuable for research. In the bolt dimension measurement task, only the edge features of the image are extracted, so only the super-resolution of the thread edge image is needed, and for the other parts, the bicubic linear interpolation is performed to enlarge the image resolution and the edge image is stitched. In the bolt dimension measurement task, only extract the image edge features. Therefore, it is only necessary to achieve thread edge image super-resolution, and for the other parts, using bilinear interpolation to enlarge the image resolution and stitching with edge images.

The biggest problem with the super-resolution of the defocus thread image is that the degree of defocus is unknown, and because the image acquisition equipment needs to be adjusted during the process of acquiring the real thread defocus and focus image, the acquired bolt defocus image and focus image can't be guaranteed It is completely spatially aligned, and paired dataset can't be obtained. However, most of the existing super-resolution methods use paired dataset for supervised training. This paper starts from the real defocus thread image, proposes a super-resolution method of defocus thread image based on cyclic generative adversarial networks (CycleGAN), constructs a CycleGAN to realize the mutual conversion from blurry domain to clear domain for unpaired defocus and focus threaded images. In the generation network, two encoders are used to extract the image defocus feature and content feature respectively to improve the network's adaptability to different defocus degrees, and a sub-pixel convolutional layer is added to the decoder to achieve image super-resolution. The defocus thread images and focus thread images collected by the industrial camera are used as the dataset for model training, and the super-resolution result is used to measure the bolt dimension to verify the effectiveness of the method in this paper.

2 Related Work

The super-resolution of the defocus thread image in the task of bolt dimension measurement belongs to the real-world single image super-resolution. This section will introduce two parts: single image super-resolution and real-world super-resolution.

2.1 Single Image Super-Resolution

CNN has achieved great success in low-level vision tasks with their powerful representation capabilities. The most advanced super-resolution methods at present use CNN to learn to reconstruct high-resolution images from low-resolution images. SRCNN [8] first applied the CNN to super-resolution, using only three convolutional layers to learn the mapping from low-resolution images to high-resolution images, and achieved a leapfrog performance improvement in peak signal to noise ratio (PSNR) and structural similarity (SSIM). ESPCN [9] proposes a sub-pixel convolutional layer, which realizes end-to-end learning from low-resolution images to high-resolution images. In order to further improve the effect of super-resolution, researchers extract more image information by increasing the depth or width of the network, such as a deeper network with residual learning [10], Laplacian pyramid structure [11], recursive learning [12], deep back projection network [13] and residual dense network [14]. In addition, the EDSR [15] removes unnecessary batch normalization layers (BN) layers in the residual block, enlarges the model size, and the performance has been significantly improved. The innovation of SRGAN and its improved method [16,17] lies in the application of a generative adversarial network (GAN) to achieve super-resolution, which combines pixel loss and adversarial loss, and uses the generation network and the discriminant network to confront each other to improve the clarity of the image, and the reconstructed image is in the texture details and visual perception are more real.

2.2 Real-World Super-Resolution

There are two major problems in real-world super-resolution. One is that the degradation of the real blurred image doesn't follow the bicubic degradation model, and the other is that the real blurred image lacks the corresponding high-definition image. Although single image super-resolution has powerful performance in dealing with bicubic degradation, it is difficult to apply to actual degraded images. The paired training data is obtained by sampling high-definition images, but the degradation of real blurred images doesn't follow the bicubic degradation model, and the results obtained by inputting real images are not reliable. Literature [18] proposes a dimensionality extension strategy, which takes different degrees of blur kernels as input to improve the applicability to real images. Literature [19] trains a network that converts high-definition images into low-resolution blurred images firstly, and then uses the synthesized low-resolution blurred images and high-definition images to form a pair of training data for network training. ZSSR [20] trains a small model for each image separately, but its training time is too long to be applied to actual scenes. Literature [21] constructed a dataset by means of optical zoom, but it still can't guarantee complete alignment. After the CycleGAN [22] is proposed, there are effective methods to

solve the learning problem of unpaired data. However, there are few researches on real-world super-resolution using CycleGAN at present. This paper starts from the defocus thread image in the bolt dimension measurement system, and uses the CycleGAN to learn the mapping from low-resolution defocus thread image to high-resolution focus thread image to enhance the sharpness of the thread edge and improve the measurement accuracy of bolt dimension measurement system.

3 Dual Camera Calibration Measurement

The principle of 2-D measurement based on machine vision is to establish a conversion model from size in pixel to size in millimeter by calibrating the internal and external parameters of the camera, that is, to calculate the single pixel accuracy of the image at the working distance of the lens, and then calculate the size in pixel of the object to be measured by extracting the edge features of the image to achieve two-dimensional size measurement. Single-pixel accuracy A_p is generally calculated by standard part, and its calculation formula is：

$$A_p = \frac{L_r}{L_p}, \tag{1}$$

where L_r is the size in millimeter of the standard part, and L_p is the size in pixel of the standard part. The single pixel accuracy is affected by the camera resolution and the field of view of the lens. Generally, the larger the camera resolution, the smaller the field of view of the lens, and the higher the single pixel accuracy. The field of view of a single camera is restricted by the camera's resolution, and it is difficult to cover the entire bolt under the premise of ensuring the measurement accuracy. Therefore, it is necessary to use the dual camera calibration measurement method to achieve bolt dimension measurement. This method uses two cameras to collect images at both ends of the size to be measured, and then accurately calibrates the positions of the two cameras to obtain the camera spacing L_1, then locates the edge of the image, calculates the relative distances L_2 and L_3 between the edge and the camera center line. Measurement result L is equal to $L_1 + L_2 + L_3$. The schematic diagram of the dual camera calibration measurement method is shown in Fig. 2.

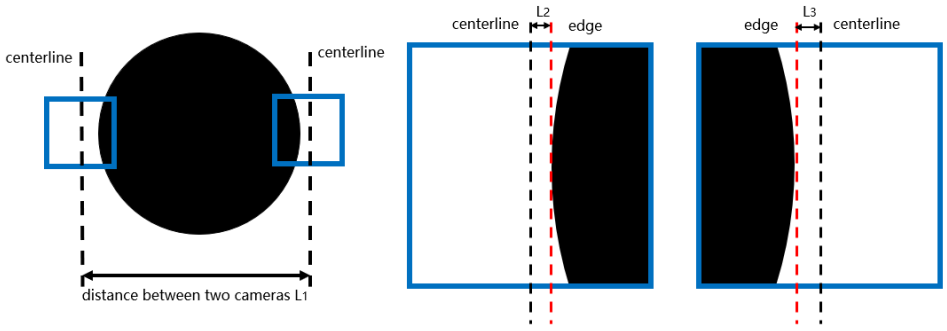


Fig. 2. Schematic diagram of dual camera calibration measurement method.

4 Method

Since the measurement system parameters need to be adjusted in the process of acquiring real thread defocus and focus images, the defocus images and the focus images can't achieve complete spatial alignment. Therefore, we choose to use a CycleGAN to learn the mapping from low-resolution defocus images to high-resolution focus images. The overall structure of the network is shown in Fig. 3. The CycleGAN contains two generative adversarial networks, one is applied to the forward process of learning the image from the blurry domain to the clear domain, and the other is applied to the reverse process of learning the image from the clear domain to the blurry domain. The super-resolution network takes the low-resolution defocus image x as input, and obtains the high-resolution output hr through the generation network G_H; then inputs hr into second generation network G_L to obtain the low-resolution composite image lr, finally uses the discriminant network D_H and D_L to determine whether hr and lr are true.

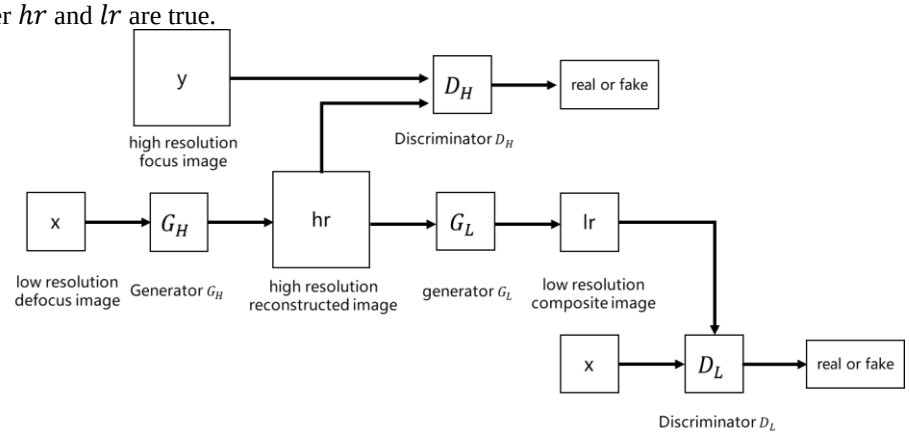


Fig. 3. Overview of the super-resolution method for defocus thread image.

4.1 Network Structure

Generate Network. The role of the generative network is to generate a pseudo sample similar to the real data distribution according to the input data distribution. The generation network is composed of an image defocus information encoder E_S, an image content encoder E_C, and a decoder D_{L2H} (replaced by D_{H2L} in the reverse process), and its network structure is shown in Fig. 4.

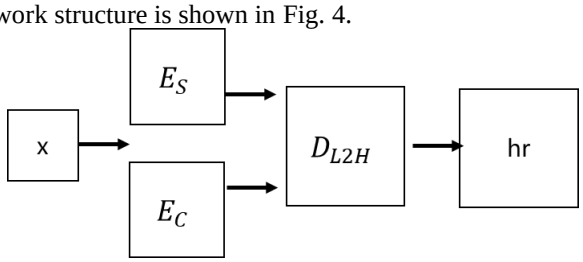


Fig. 4. Generate network.

The image defocus information encoder E_S uses VGG16 as the backbone network, and combines 1 convolutional layer, 1 activation layer and 1 mean pooling layer. Its network structure is shown in Fig. 5. The structure finally outputs an 8-dimensional feature vector S, whose role is to extract the defocus information in the image.

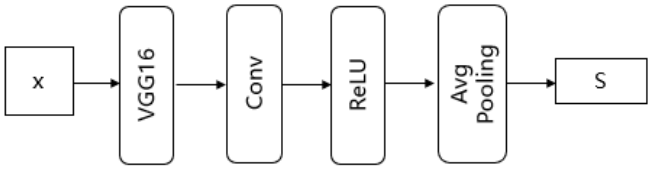
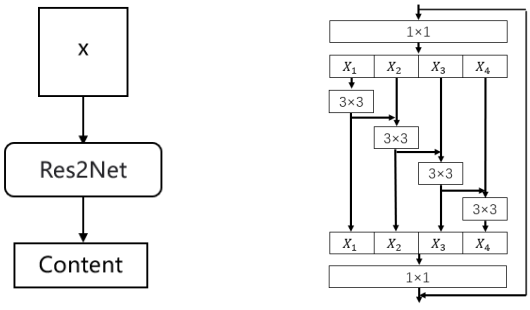


Fig. 5. image defocus information encoder.

The role of the image content encoder E_C is to extract the content information in the image. In order to better extract the multi-scale information in the image, the image content encoder adopts the Res2Net [23] structure, and its network structure is shown in Fig. 6.



(a) (b)

Fig. 6. image content encoder and Res2Net.
(a) image content encoder. (b)Res2Net.

The decoder network structure is shown in Fig. 7. In the forward loop of reconstructing a high-resolution clear image from a low-resolution defocus image, the image defocus feature and image content feature extracted from the encoder are input to the decoder for feature fusion and use sub-pixel convolutional to achieve 2 times magnification of image resolution, and finally output high-resolution results. In the reverse process of synthesizing a high-resolution clear image into a low-resolution defocus image, the sub-pixel convolution layer is changed to a maximum pooling layer to reduce the image resolution by half.

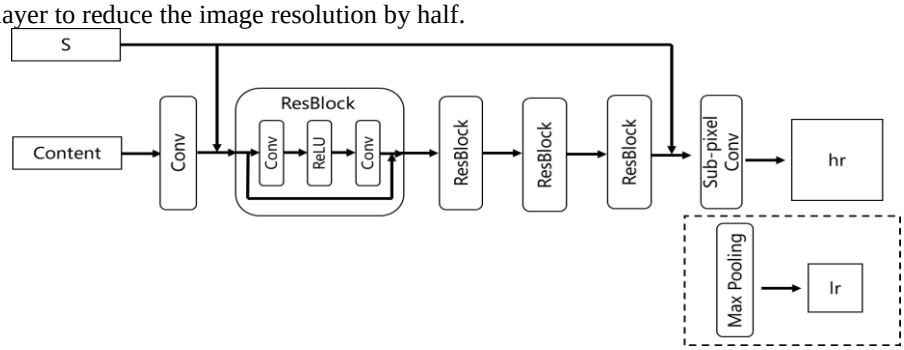


Fig. 7. Decoder network.

Discriminant Network. The role of the discriminant network is to distinguish between real data and generated data so that the generation network can generate more realistic target domain images. The network structure is shown in Fig. 8, where k represents the size of the convolution kernel in the convolution layer, n represents the number of output feature maps, s represents the step size of the convolution kernel. The discriminant network consists of 8 convolutional layers, and each convolutional layer uses a 3×3 size convolution kernel. In the calculation process, the number of feature maps is gradually increased from 64 to 512, each time the feature map is doubled, a convolution kernel with a step size of 2 is used to reduce the size of the output feature map. Finally, the binary adaptive mean convergence layer and the sigmoid activation function are used to judge the authenticity of the input image.

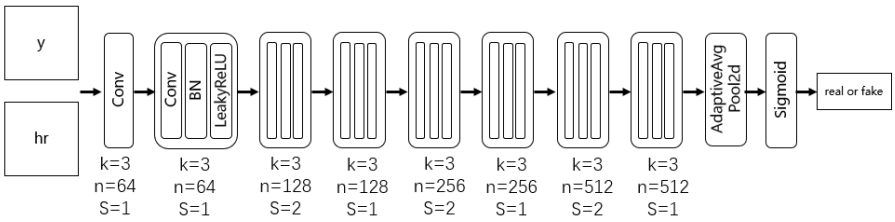


Fig. 8. Discriminant network.

4.2 Loss Function

The loss function is very important to the network performance, and the essence of the network model training process lies in the continuous optimization of the loss function. For paired dataset, the loss function usually uses pixel domain loss, that is, using mean square error (MSE) loss or perceptual loss [24] in the pixel domain, by comparing the two types of images pixel by pixel to ensure a higher PSNR. Since the defocus image and the focus image can't be completely aligned in space, the loss function in the pixel domain can't be used to optimize the network model. Therefore, the cyclic consistency loss function is introduced in this method to realize the optimization of the super-resolution network. The overall network loss function l_t is

$$l_t = l_H(G_H, D_H, x, y) + l_L(G_L, D_L, y, x) + \lambda l_C(G_H, G_L, x, y), \quad (2)$$

where l_H is the loss of the generation network G_H from low-resolution defocus image to high-resolution focus image and its discriminant network D_H; l_L is the loss of the generation network G_L from high-resolution focus image to low-resolution defocus image and its discriminant network D_H; l_C is the cyclic consistency loss; λ is the weight of l_C. The loss function l_H is defined as

$$l_H(G_H, D_H, x, y) = E_{y-P_{data}(y)}[logD_H(y)] + E_{x-P_{data}(x)}[\log(1 - D_H(G_H(x)))], \quad (3)$$

where E is the mathematical expectation function, ~ represents the obey relationship, $P_{data}(y)$ is the distribution of the focus high-resolution image y. The purpose of the loss function is to make the generated high-resolution clear image and the focus image as similar as possible. Similarly, the loss function l_L can also be constructed.

$$l_L(G_L, D_L, y, x) = E_{x-P_{data}(x)}[logD_L(x)] + E_{y-P_{data}(y)}[\log(1 - D_L(G_L(y)))]. \quad (4)$$

The cyclic consistency loss function l_C is defined as

$$l_C(G_H, G_L, x, y) = E_{x-P_{data}(x)}\left[\left\|G_L\left(G_H(x)\right) - x\right\|_1\right] + E_{y-P_{data}(y)}\left[\left\|G_H\left(G_L(y)\right) - y\right\|_1\right]. \quad (5)$$

5 Experiment and Analysis

Since the method in this paper is aimed at unpaired defocus thread image and can't directly quantitatively analyze the super-resolution results, this experiment will first construct a paired simulated defocus image datasets to analyze the performance of the method, and then the method in this paper is applied to the bolt dimension measurement to test its influence on the measurement accuracy.

In this experiment, the hardware configuration is Inter I5-9600KF CPU with RTX-3060Ti GPU, and the software uses OpenCV image processing library and PyTorch deep learning framework.

5.1 Super-Resolution Experiment of Simulated Defocus Image

Dataset. In this part of the experiment, the DIV2K dataset is used as the experimental data. First, we crop the high-resolution clear images in the DIV2K dataset to obtain a clear image with a size of 512×512, and then apply different degrees of defocus blur with a blur kernel size of 3, 4, 5 to the clear image, then apply a defocus blur with blur kernels size of 3, 4, 5 to the clear image to obtain 3 defocus images, and the shape of the blur kernel is as shown in Fig. 9. Finally, we reduce the resolution of defocus images to 256×256 to obtain low-resolution defocus images and form 3 pairs of paired data. The number of images in the training set is 3000, and the number of images in the test set is 600.

Fig. 9. The shape of the blur kernel.

Experimental Results and Analysis. In order to compare the performance of the methods in this paper, the methods of EDSR, SRGAN, and ESRGAN are introduced for comparison. Because these methods use pixel domain loss functions in the training process, in order to ensure fairness, in this part of the experiment, the method in this paper also adds the pixel domain mean square error loss and the perceptual loss to optimize the network. The square error loss l_{mse} and the perceptual loss l_p are defined as

$$l_{mse} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\|I(i,j) - K(i,j)\|^2, \quad (6)$$

$$l_p = \frac{1}{uv}\sum_{i=0}^{u-1}\sum_{j=0}^{v-1}\|V_I(i,j) - V_K(i,j)\|^2, \quad (7)$$

where I is the label image, $I(i,j)$ is the pixel value in the i-th row and j-th column of the label image, K is the super-resolution result, $K(i,j)$ is the pixel value in the i-th row and j-th column of the label image, V_I is the output of I input VGG16, $V_I(i,j)$ is

the value in the i-th row and j-th column of V_I, V_K is the output of K input VGG16, $V_K(i,j)$ is the value in the i-th row and j-th column of V_K.

The result of super-resolution of the simulated defocus images is shown in Fig. 10. The first column is the low-resolution defocus images, the second column is the high-resolution clear images, the third column, the fourth column, the fifth column, the sixth column, and the seventh column are the results of Bicubic, EDSR, SRGAN, ESRGAN and the method in this paper.

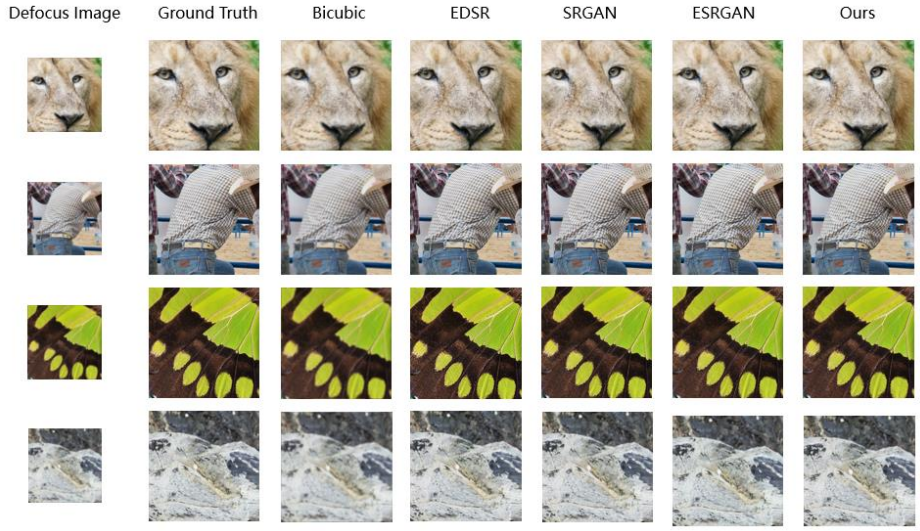


Fig. 10. Super-resolution results of simulated defocus images.

In order to be able to compare several methods more fairly, we use PSNR and SSIM to quantitatively evaluate the image quality. The calculation formula of PSNR is

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\|I(i,j) - K(i,j)\|^2, \tag{8}$$

$$PSNR = 10 * log_{10}\left(\frac{MAX_I^2}{MSE}\right), \tag{9}$$

where I is the label image, K is the super-resolution result, MSE is the mean square error between the label image and the super-resolution result, MAX is the maximum value of the pixel color.

The calculation formula of SSIM is

$$l(I,K) = \frac{2\mu_I\mu_K + C_1}{\mu_I^2 + \mu_K^2 + C_1}, \tag{10}$$

$$c(I,K) = \frac{2\sigma_I\sigma_K + C_2}{\sigma_I^2 + \sigma_K^2 + C_2}, \tag{11}$$

$$s(I,K) = \frac{\sigma_{IK} + C_3}{\sigma_I\sigma_K + C_3}, \tag{12}$$

$$SSIM = l(I,K) * c(I,K) * s(I,K), \tag{13}$$

where $l(I,K)$, $c(I,K)$, $s(I,K)$ represents the similarity of the two images in terms of brightness, contrast, and structure, μ_I, μ_K represent the mean values of I and K, σ_I, $\sigma_K$represent the variances of I and K, σ_{IK} represents the covariance of I and K, C_1, C_2, and C_3 are constants, usually $C_1 = (k_1 * l)^2$, $C_2 = (k_2 * l)^2$, $C_3 = C_2/2$,where $k_1 = 0.01$, $k_2 = 0.01$, $l = 255$. The calculation formula of mean μ_I, variance σ_I^2 and covariance σ_{IK} is

$$\mu_I = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} I(i,j), \tag{14}$$

$$\sigma_I^2 = \frac{1}{H \times W - 1}\sum_{i=1}^{H}\sum_{j=1}^{W} (I(i,j) - \mu_I)^2, \tag{15}$$

$$\sigma_{IK} = \frac{1}{H \times W - 1}\sum_{i=1}^{H}\sum_{j=1}^{W} (I(i,j) - \mu_I)(K(i,j) - \mu_K), \tag{16}$$

where I and K are the target image, H and W are the height and width of the image, μ_I and μ_K are the average values of I and K.

The following Table 1 gives super-resolution evaluation on simulated defocus images of all methods.

Table 1. super-resolution evaluation on simulated defocus images.

Methods	PSNR	SSIM
Bicubic	25.96	0.8193
EDSR	29.12	0.8563
SRGAN	31.79	0.8762
ESRGAN	32.19	0.8968
Ours	32.74	0.9182

It can be seen from the visual comparison, the image quality produced by the bicubic interpolation method is poor, high-frequency information is seriously lost, and even artifacts appear. Compared with the bicubic interpolation image, the super-resolution method based on deep learning can improve the quality of reconstruction and make the image clearer. It can be seen from Table 1 that the super-resolution method based on deep learning all achieve high performance in terms of PSNR and SSIM. Among them, the methods with the generative adversarial structure are able to better recover high-frequency information in defocus images. Compared with other methods, the method in this paper obtains higher values, which means that the method in this paper can generate better reconstructed images.

5.2 Super-Resolution Experiment of Defocus Thread Image

Experiment Platform. In order to verify the effectiveness of the method in this paper, a bolt dimension measurement platform is built to conduct experiments, and a laser rangefinder is used to calibrate the experimental platform. The experimental platform is shown in Fig. 11. The camera is MER2-1800-32U3M, and its resolution is

4912× 3684. The lens is MML3-ST40D, and the light source is a forward point light source matching the lens.

Fig. 11. Experimental platform.

Experimental Results and Analysis. In this part of the experiment, this method is applied to the bolt dimension measurement task. First, the defocused thread images and the focused thread images are collected through the experimental platform, and then the edge images of the thread crest and root are intercepted, their size is 640*640. Finally the defocused thread images are down-sampled by 2 times to obtain low-resolution defocus images for super-resolution model training. The number of training sets is 600. The visualization results of super-resolution of the defocus thread images are shown in Fig. 12.

The bolt dimension measurement results are shown in Table 2. Taking the standard bolts of known size as the measurement object, the method in this paper is used to measure the dimensions of 5 different positions. The measurement items are the major diameter, minor diameter and thread spacing of the bolt. The results are shown in Table 2. In order to test the stability of the method, 10 repeated experiments were carried out for the major diameter, minor diameter and pitch of the No. 1 position, and the results are shown in Fig. 13.

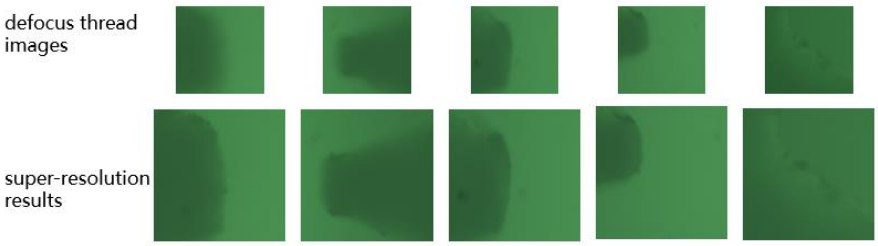


Fig. 12. Super-resolution results of defocus thread images.

Table 2. Results of bolt dimension measurement.

Position	Item	Real value	Measurement results (no SR)	Error (no SR)	Measured value (SR)	Error (SR)
No. 1	major diameter	10.0052	9.9928	-0.0124	9.9990	-0.0052
	minor diameter	8.8003	8.7924	-0.0079	8.7962	-0.0041
	pitch	1.0034	1.0002	-0.0032	1.0050	0.0016
No. 2	major diameter	10.0012	10.0245	0.0233	10.0104	0.0092
	minor diameter	8.7954	8.8007	0.0053	8.8008	0.0054
	pitch	0.9941	0.9972	0.0031	0.9960	0.0019
No. 3	major diameter	9.9978	10.0016	0.0038	10.0013	0.0035
	minor diameter	8.7967	8.8359	0.0392	8.8006	0.0039
	pitch	1.0060	1.0067	0.0007	1.0067	0.0007
No. 4	major diameter	10.0102	9.9903	-0.0199	10.0079	-0.0023
	minor diameter	8.8505	8.8534	0.0029	8.8530	0.0025
	pitch	1.0032	0.9937	-0.0095	1.0023	-0.0009
No. 5	major diameter	10.0081	10.0491	0.0338	10.0183	0.0102
	minor diameter	8.8223	8.8286	0.0063	8.8274	0.0051
	pitch	9.9921	9.9983	0.0062	9.9994	0.0073

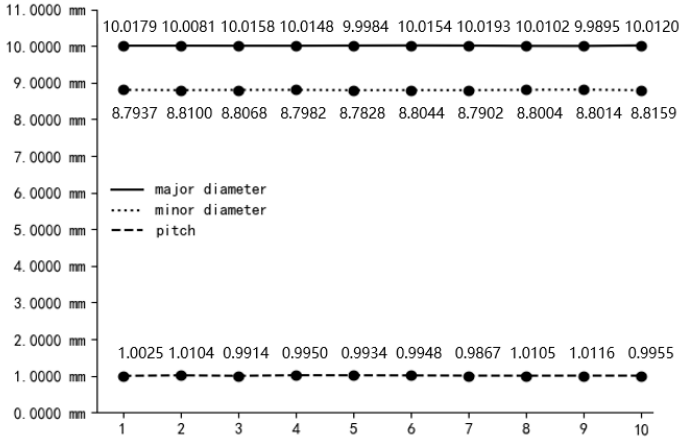


Fig.13. Repeated measurement results of major diameter, minor diameter and pitch at position No. 1.

It can be seen from Fig. 12 that the method in this paper effectively restores the edge details of the thread image and reconstructs a clear thread image. It can be seen from Table 2 and Fig.13 that the method in this paper reduces the edge positioning error by about 15 pixels, greatly reduces the measurement error, and can effectively improve the measurement accuracy of the bolt dimension. It can basically ensure that the measurement accuracy is within 0.01mm and can maintain good stability in re-

peated experiments, which is of great significance to the task of bolt dimension measurement.

6 Conclusion

In this paper we have proposed a super-resolution method for defocus thread image based on CycleGAN in bolt dimension measurement. This method utilizes CycleGAN to solve the problem of the lack of paired data in defocus thread images. In the generative network, two encoders are used to extract the defocus feature and content feature of the image respectively to increase the adaptability to different degrees of defocus blur, and the sub-pixel convolutional layer is used to achieve image resolution enlargement. The results show that this method can effectively reconstruct a clear thread image in the bolt dimension measurement task, improve the edge positioning accuracy, and effectively improve the bolt dimension measurement accuracy. However, the calculation of this method is slightly complicated. In the next step, we will simplify the network to improve the efficiency of bolt dimension measurement.

Funding: This research was supported by the Department of Science & Technology of Shandong Province (2017CXGC0810)

References

1. Zhang P J: Research on high-precision measurement technology of complex workpiece shape and size based on machine vision University of Jinan (2019).
2. Lan J F. Design and implementation of a high-precision detection system for piston shape and position based on machine vision. University of Jinan (2021).
3. Yang Y, Geng Z, Wang R, Li J. Defocusing and Deblurring of Traditional Cameras with Spatial Variation. Electro-Optics and Control 22(09), 91-95 (2015).
4. Yang Y. Research on Image Restoration Algorithm. Sichuan University (2004).
5. R C Gonzalez, P. Wentz. Digital Image Processing. Science Press (1981).
6. Xiaoping Hu, Guoliang Chen, Zhengyu Mao, et al. Research on Wiener Filter Restoration of Defocused Images. Journal of Instrumentation, 28(003), 479-482 (2007).
7. Yan H, Yan W, Li W W. Image Restoration Based on Lucy-Richardson Algorithm. Computer Engineering, 36(15), 204-205 (2010).
8. Dong C, Loy C C, He K, Tang X. Learning a deep convolutional network for image super-resolution. European Conference on Computer Vision 2014, LNCS, vol. 8692, pp.184-199. Springer, Berlin (2014).
9. Shi W, Caballero J, F Huszár, et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. IEEE Conference on Computer Vision and Pattern Recognition 2016, LNCS, pp. 1874-1883. IEEE, Las Vegas (2016).
10. Kim J, Lee J K, Lee K M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. IEEE Conference on Computer Vision and Pattern Recognition 2016, LNCS, pp. 1646-1654. IEEE, Las Vegas (2016).

11. Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2017, LNCS, pp. 5835-5843. IEEE, Hawaii (2017).
12. Kim J, Lee J K, Lee K M. Deeply-Recursive Convolutional Network for Image Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2016, LNCS, pp. 1637-1645. IEEE, Las Vegas (2016).
13. Haris M, Shakhnarovich G, Ukita N. Deep Back-Projection Networks For Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2018, LNCS, pp. 1664-1673. IEEE, Salt Lake City (2018).
14. Zhang Y, Tian Y, Kong Y, et al. Residual Dense Network for Image Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2018, LNCS, pp. 2472-2481. IEEE, Salt Lake City (2018).
15. Lim B, Son S, Kim H, et al. Enhanced Deep Residual Networks for Single Image Super-Resolution. IEEE Conference on Computer Vision and Pattern Recognition 2017, LNCS, pp. 1132-1140. IEEE, Hawaii (2017).
16. Ledig C, Theis L, F Huszar, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. IEEE Conference on Computer Vision and Pattern Recognition 2017, LNCS, pp. 105-114. IEEE, Hawaii (2017).
17. Wang X, Yu K, Wu S, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. European Conference on Computer Vision 2018, LNCS, vol. 11133, pp.63-79. Springer, Cham (2018).
18. Zhang K, Zuo W, Zhang L. Learning a Single Convolutional Super-Resolution Network for Multiple Degradations. IEEE Conference on Computer Vision and Pattern Recognition 2018, LNCS, pp. 3262-3271. IEEE, Salt Lake City (2018).
19. Bulat A, Jing Y, Tzimiropoulos G. To learn image super-resolution, use a GAN to learn how to do image degradation first. European Conference on Computer Vision 2018, LNCS, vol. 11210, pp. 187-202. Springer, Cham (2018).
20. Shocher A, Cohen N, Irani M. Zero-Shot Super-Resolution Using Deep Internal Learning. IEEE Conference on Computer Vision and Pattern Recognition 2018, LNCS, pp. 3118-3126. IEEE, Salt Lake City (2018).
21. Zhang X, Chen Q, Ren N, et al. Zoom to Learn, Learn to Zoom. IEEE Conference on Computer Vision and Pattern Recognition 2019, LNCS, pp. 3757-3765. IEEE, Long Beach (2019).
22. Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. IEEE International Conference on Computer Vision 2017, LNCS, pp. 2242-2251. IEEE, Venice (2017).
23. S H Gao, M M Cheng, K Zhao, X Y Zhang, M H Yang, P Torr. Res2Net: A New Multi-Scale Backbone Architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 2, pp. 652-662 (2021).
24. Johnson J, Alahi A, Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. European Conference on Computer Vision 2016, LNCS, vol. 9906, pp. 694-711. Springer, Cham (2016).