Membrane Technology
ISSN (online): 1873-4049

Defect Image Generation Algorithm of Commutator Based on Generative Adversarial Networks

Xianjun Tang[1,2], Jiatao Liao[1,2], Yufeng Shu[1,2,*], Zhongming Xie[3]

[1]*Intelligent Manufacturing College, Dongguan Polytechnic, Dongguan, 523808, China*
[2]*Guangdong Textile Industry Intelligent Detection Engineering Technology Research Center (DGPT), Dongguan, 523808, China*
[3] *Shini Plastics Technologies (Dongguan), Inc., Dongguan, 523808, China*
**Corresponding Author.*

Abstract:
At present, there have been many studies on the application of deep learning algorithms in defect detection work. Based on the object detection algorithms introduced earlier, a defect image recognition system is constructed to detect various types of defects in the converter, which can save manpower and material resources and improve the safety of this work. However, deep learning algorithms require a certain amount of valid samples to be effective. However, in actual operation, it is difficult to collect defect images of the commutator, which means that there are situations where the requirements of deep learning detection algorithms cannot be met. This article aims to explore a feasible data augmentation scheme, which is to generate effective samples through generative adversarial networks. This article will investigate two classic image generation methods based on Generative Adversarial Networks (GANs). And apply it to the field of generating defect images for commutators, propose a method for generating defect images in a directional manner, and compare it with the above two methods. The experimental results show that our method has a higher FID value compared to the above two methods, and has a more similar data distribution compared to real images.

Keywords: GAN, Defect image, DCGAN

IMAGE GENERATION METHOD BASED ON GENERATIVE ADVERSARIAL NETWORK METHOD

Theoretical Description of the Method of Generative Adversarial Network

The Generative Adversarial Network [1] was first proposed by Goodfellow in 2014. He proposed a new framework for identifying generative models through adversarial processes. There are two parts: 1) the model used to capture the data distribution is called the generation model G; 2) the model used to estimate the probability that the sample comes from the training data rather than G is called the discriminator model D, the training process of the generator G is to maximize the error probability of the discriminator D, while the training process of the discriminator D is to minimize it. This framework can be thought of as a minimal duo game. In the space of any function G and D, there is a unique solution. G restores the training data distribution, where D is equal to 1/2. In the case where G and D are defined by multi-layer perceptrons, the entire system can be trained with back propagation. However, there are some problems in the GAN model, including the instability of training. Therefore, various improved versions of GAN are usually used in practical application, instead of using the original GAN directly. In various improved versions of GAN, Deep Convolution generative adversarial network (DCGAN for short) [2] is introduced into the generative adversarial network by adjusting the network structure, thus obtaining good results in image generation. Wasserstein GAN [3] improves the loss function in the GAN model from the theoretical point of view, thus effectively improving the instability in the model training process. The following two variants of GAN are described in more detail.

(1) *DCGAN*: The deep convolution generative adversarial network (DCGAN) is one of the more successful and commonly used network designs of GAN (as shown in Figure 1). Compared with the original generative adversarial network, it mainly uses the convolution layer instead of the maximum pooling layer for down-sampling, eliminates the full connection layer and uses batch normalization. From the perspective of view of optimization the model structure, it uses convolution steps and transposed convolution [4] for down-sampling and up-sampling to make the network training smoother. The original GAN can learn better image features after training, but it is very unstable during training. At the same time, generating a model produces meaningless images in many cases. However, DCGAN can solve the unstable problem of GAN training to a certain extent.

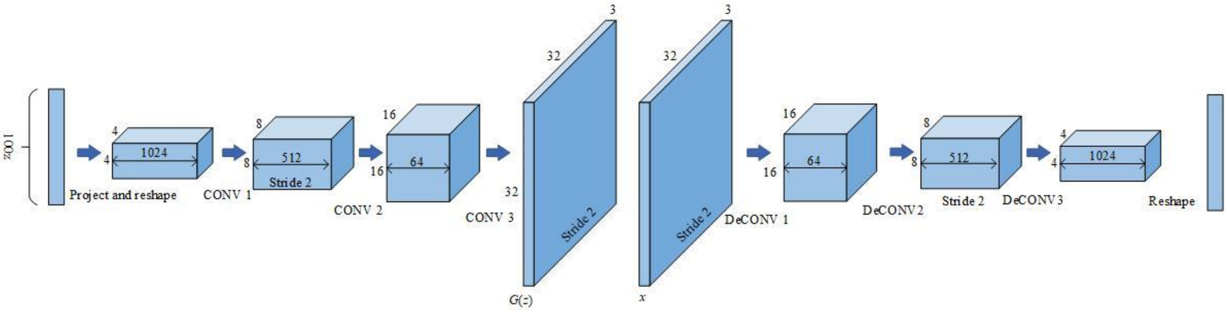

Figure 1. DCGAN network model

As can be seen from the above figure 1, two new convolution modes, namely deconvolution and step-length convolution, are adopted instead of pooling operation in DCGAN. Deconvolution operation is used to realize image upsampling in the generation model, and in the discrimination model, the step-size convolution is used instead of the pooling operation to realize the down-sampling of the image. For the training of the DCGAN network, the generator and discriminator have different weights and bias sets. The optimization procedure associated with the discriminator minimizes its loss function $L_D(\theta,\phi)$ by changing its parameter set. Similarly, the optimization process associated with the generator is to minimize its loss function $L_G(\theta,\phi)$ by changing its parameter set. The goal of the generator G is to learn the distribution of the real data and then generate samples that are as similar as possible to the real samples so that the discriminant network D cannot distinguish them. The input to the generator G is random Gaussian noise $Z=(Z^1,Z^2,\dots,Z^m)$ and the output is a sequence of synthesized samples$G(Z)=(G(z)^1,G(z)^2,\dots,G(z)^m)$. The goal of the discriminator D is to distinguish whether the input data is a real sample x or an output sequence $G(Z)$ of G. If the input data is a real sample x, then D output 1, if the input data is the output sequence $G(Z)$ of G, then D output 0, the loss function of DCGAN is shown in Equation (1), and the optimization objective is shown in Equation (2).

$$L(D,G)=E_{x\sim P_{data}(x)}[log\,D\,(x)]+E_{z\sim P_z(z)}\left[log\left(1-D\big(G(z)\big)\right)\right] \tag{1}$$

$$\underset{G}{min}\,\underset{D}{max}L(D,G) \tag{2}$$

One of the training features of DCGAN is alternate training of two networks. When one network is trained, the parameters of the other network are fixed. For the training of the G network, the goal is to be as close as possible to 1, so its optimization function is shown in Equation (3).

$$min_G\,E_{z\sim P_z(z)}\left[log\left(1-D\big(G(z)\big)\right)\right] \tag{3}$$

When the discriminator D is trained, the loss function of the network consists of two parts, namely the real sample and the generated sample of G. For the real sample x, the output of D shall be as close to 1 as possible, and for the output sequence $D(x)$ of G, the output of D shall be as close as possible to 0, so the optimization function is shown in Equation (4).

$$\underset{D}{max}E_{x\sim P_{data}(x)}[log\,D\,(x)]+E_{z\sim P_z(z)}\left[log\left(1-D\big(G(z)\big)\right)\right] \tag{4}$$

The goal of DCGAN training is to achieve Nash equilibrium, which describes the game states in non-cooperative games [5]. No part can improve its score simply by changing its strategy. Nash equilibrium is the equilibrium state of the game between generator and discriminator. In this state, the distribution Q of the generated samples matches the distribution P of the real data, and the discriminator cannot distinguish Q from P. Therefore, whether the input of D is real data or generated data, the output of D is 0.5 in the Nash equilibrium state [6-7]. As mentioned above, batch normalization is adopted in DCGAN, which mainly defines the input batch as $B=\{z_1,z_1,\dots,z_m\}$ and then sets two learning parametersγ,β. Finally, output the result$BN_{\gamma\beta}(z_i)$.

While making the above improvements, DCGAN also uses the tanh activation function on the output layer of the generator and Leaky ReLU [8] activation function in the discriminator. The above operations make the training process more stable and further improve the results of the model.

(2) Wasserstein GAN: The above DCGAN network model is a relatively popular and successful variant of GAN in recent years. Corresponding to it, there is another variant of generative adversarial network, which is different from DCGAN in

terms of structure innovation of generative adversarial network. Its purpose is to fundamentally solve the problem that generative adversarial network is difficult to train. In order to make the training process more stable, Wasserstein GAN (WGAN) was proposed. In WGAN there is a new loss function derived from the Wasserstein distance [9], which is used to measure the distance between two probability distributions. In generative adversarial network, that essence of a generator is to generate images that are similar to the real image or want to be distributed, and the similarity between the generate image and the original image can be expressed by the distance of the distribution, that is, the closer the distance is, the more similar they are. So it seems more appropriate to measure generator by distance. The Wasserstein distance is shown in Equation (5):

$$W\left(P_r, P_g\right) = \inf_{\gamma \in \prod\left(P_r, P_g\right)} E_{(x,y)\sim\gamma}[\|x - y\|] \tag{5}$$

Where $\prod\left(P_r, P_g\right)$ represents the collection of all federated distributions $\gamma(x, y)$ and $\gamma(x, y)$represents the distance to transform the distribution P_r into the P_g distribution. $\|x - y\|$ represents the distance between samples, then $E_{(x,y)\sim\gamma}[\|x - y\|]$ is the expected value of distance for all joint distributions. The loss function of WGAN can be expressed by Kantorovich-Rubinstein duality, as shown in Equation (6).

$$\min_{G} \max_{D \in D} \underset{x\sim P_r}{E}[D(x)] - \underset{x\sim P_g}{E}\left[D\left(G(\tilde{x})\right)\right] \tag{6}$$

Where D is the set of 1-Lipschitz functions from which the discriminator D calculates the distance between the joint distribution samples.

Although the improved WGAN solves the problem of training instability and improves the quality of the generated image, there are still some problems compared with the original GAN [10]. The weight clipping in WGAN limits the performance of the network, that is, it is very difficult for the network to simulate those complex functions, and only some relatively simple functions can be produced. In addition, the weight clipping needs to be set by you. Improper setting will cause gradient explosion and gradient disappearance. In order to avoid the above problems, the gradient penalty is added, which can be expressed as [11]:

$$\lambda \underset{x\sim P_{\tilde{x}}}{E}[(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2] \tag{7}$$

The gradient penalty limits the gradient weights to [-c, c], where c is the threshold, to prevent gradient disappearance and explosion problems.

Generation of Defect Image of Commutator Cylindrical Expansion Plane Based on DCGAN and WGAN Method

Data set

Some defects will inevitably appear in the process of commutator from raw material to processing, which will affect the quality of commutator. Our purpose is to carry out quality inspection on commutator during production process or before leaving factory to retain qualified commutator products and reject unqualified commutator products. In this paper, the defect detection is carried out on the cylindrical area of the commutator. Figure 2 is the actual diagram of the commutator:

Figure 2. Physical drawing of commutator

As can be seen from Figure 2, due to the orientation of the inside of the commutator, the outer circle area is a encircle image, and it is difficult for us to observe its defects with naked eyes. In order to better introduce the depth learning algorithm, the professional equipment which can be used for 360-degree shooting is used to take the cylindrical developed image of the commutator for research. Figure 3 shows the outer cylindrical development of the commutator.

Figure 3. Development of outer cylinder of commutator

In deep learning, besides network structure, objective function and training strategy are very important to the effect of the model; the quality of data set has great impact on the final result. Better quality data sets can speed up the training of the network and improve the training effect of the network. Therefore, we first prepare the commutator data set before conducting the experiment. Firstly, the image is normalized and uniformly set to a fixed size because the neural network requires a fixed size to input the image. In this paper, four kinds of defects are defined according to the imaging characteristics of the defective products, which are (1) cylindrical slag defect, (2) cylindrical burr defect, (3) oil stain defect and (4) cylindrical imprinting defect. The following Figure 4 shows the four defects:

Figure 4. Image defect display of commutator cylindrical expansion (a) cylindrical slag inclusion defect (b) cylindrical imprinting defect (c) cylindrical burr defect (d) greasy dirt defect

After we completed the normalization of the image, the size and image quality of the image data set of the cylindrical expansion diagram outside the commutator met the requirements of model training. In order to diversify the data set, we use the traditional data enhancement method to expand the image of commutator outer circle for the first time, and increase the number of commutator outer circle expansion image.

Horizontal Flip: We flip the image horizontally. The flipped image is the mirror image of the original image. The mirror image after horizontal flip has no effect on the quality of the image. As an example, we compare the images before and after turning (see figure 5).

Figure 5. Horizontal flip (a) Original image (b) Image after horizontal flip

Vertical Flip: We flipped the image vertically, which means we rotated the original image 180 degrees. The mirror image after the vertical flip has no effect on the quality of the image. The following figure 6 shows a vertical flip of the image of the commutator's outer circle expansion.

Figure 6. Vertical flip (a) Original image (b) Image after vertical flip

Zoom: Zoom in or out of the image. When enlarging, the size of the enlarged commutator outer circle expansion image will be larger than that of the original image. We trim the image according to the size of the original image, and the reduction refers to filling the edge part. Because no suitable filling scheme is found in the experiment, we only enlarge the image of the outer circle of the commutator. The enlarged image in (a) of Figure 7 is shown in Figure 7(b).

(a) (b)

Figure 7. Enlarged image (a) Original image (b) Image enlarged

Lab environment setup

This section mainly analyzes the image generation results of DCGAN method and WGAN-gp method. In this paper, many image tools are used in the study of commutator cylindrical defect image, as shown in Table 1:

Table 1. Tools used in the algorithm

Items	Content
Operating system	Linux Ubuntu 16.04、Windows 10
Computing Architecture	CORE i7-9750h + NVIDIA GTX 1080Ti
In-depth learning framework	Tensorflow, Keras
Image process	Python 3.7、OpenCV 3.0

Experimental results and analysis

We trained using DCGAN and WGAN, respectively, and output the results of the model generation based on the above two commutator cylindrical expansion images, as shown in Figures 8 and 9. Figure 8 is a developed image of the commutator cylindrical generated by DCGAN; and Figure 9 is an expanded image of the commutator cylindrical generated by WGAN.

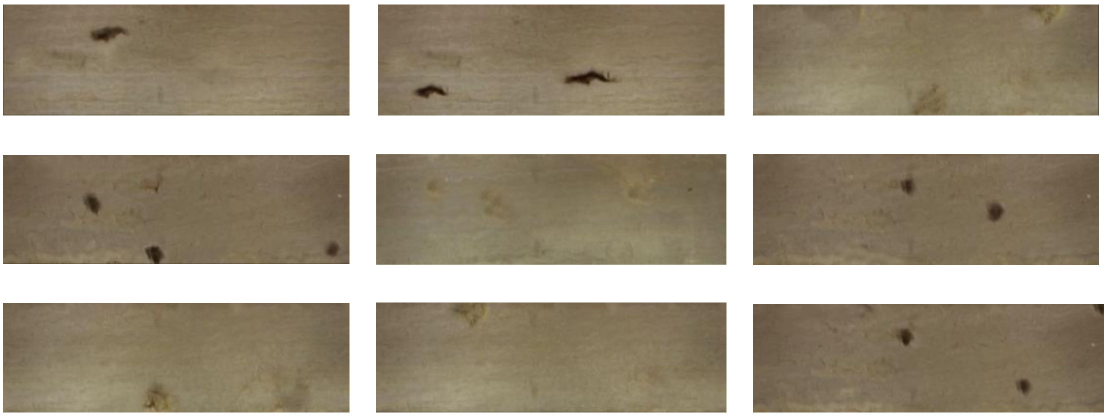

Figure 8. DCGAN-generated developed image of commutator cylindrical

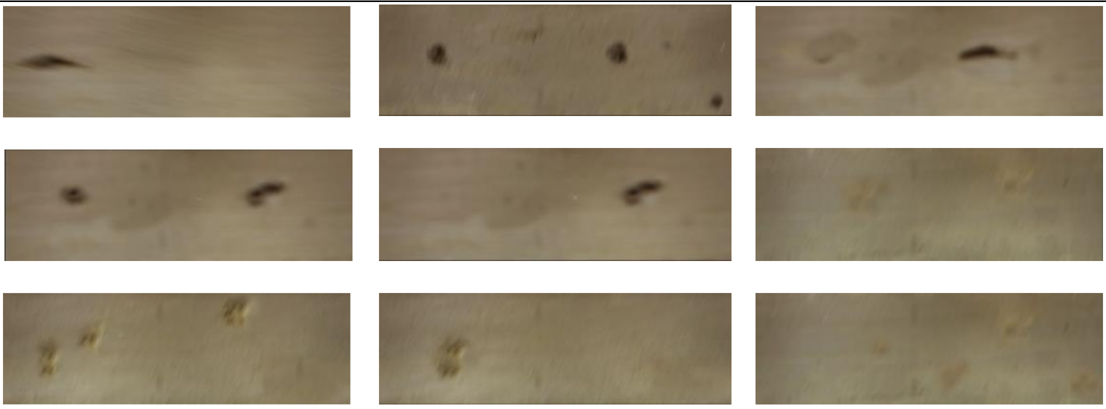

Figure 9. Developed image of commutator cylindrical generated by WGAN

As can be seen from Figures 5 and 6, both the DCGAN and WGAN methods can extend the commutator outer circle expansion image data set; the outer circle expansion image with different defects can be directionally generated. The DCGAN method has a better contour than the WGAN method, for example, the cylindrical imprint defect of the commutator developed image generated by different methods in Figure 8 and Figure 9, the cylindrical imprint defect generated by the DCGAN method has better detail display compared with the cylindrical imprint defect generated by the WGAN method, and compared with the original image, the DCGAN method is more visually similar, and the generated image is more delicate. However, the DCGAN method also has a problem that the commutator cylindrical developed image generated by the DCGAN cannot generate the specified defect image and can only be randomly generated. This obviously cannot be used as a data set extension for specifying defects.

ORIENTATION DEFECT IMAGE GENERATION METHOD BASED ON CONDITIONAL CONSTRAINT

In section 2, we evaluated the images generated using the outer circle expansion image generation method based on WGAN commutator and the outer circle expansion image generation method based on DCGAN commutator. It has been found that both methods can produce a better image of the commutator's outer circle. The image quality of the image generation method based on DCGAN commutator is better, but the generated image cannot be generated according to the specified defect, and the generated image is random. In order to generate the image of the outer circle of the commutator with a specified defect, we construct a model for generating the image of the outer circle of the commutator which can directionally generate the defect by improving the generative adversarial networks and adding some constraints.

Principle of Generative Adversarial Network Based on Conditional Constraints

Because the traditional generative adversarial network has the ability of random generation, when a plurality of kinds of images is trained at the same time, the types of the final generated images are also random, and the specified images cannot be generated. Because of its freedom, it limits its application in directional data expansion.

Therefore, in order to be able to generate the image of the outer circle of the commutator with specified defects, we consider the class label in the supervised training. First, we add a condition constraint in the GAN network. We change the GAN without supervision learning into the supervised learning process with condition constraint. As shown in Figure 10, we add a condition y to both the generator G and the discriminator D to constrain them. In the generator, the noise z is not only input only before the network starts. But the constraint y is also input into the network as additional information. Similarly, in the discriminators, we input the image and the corresponding condition y in series into the classifier network. The effect of the condition y is to constrain the data generation process to produce an image of the commutator's outer circle expansion for the specified defect.

A conditional generative adversarial network can be understood as an extension of the original countermeasure generative adversarial network, both the generator and the discriminator adding an additional condition y, which is class information, so as to realize the specified output of the image meeting the condition y constraint. In the generator, the original input noise z and the conditional information y are combined to form a joint hidden layer representation. Like the original generative adversarial network, the objective function of the conditional generative adversarial network is a minimal maximum game with conditional probability. After the condition y is added, the loss function of the generative adversarial Network is shown in Equation 8:

$$\min_{G}\max_{D} L(D,G) = E_{x\sim P_{data}(x)}[\log D(x|y)] + E_{z\sim P_z(z)}\left[\log\left(1 - D\left(G(z|y)\right)\right)\right] \quad (8)$$

Where, y is the conditional information, $G(z|y)$ means that conditional constraint y is added to the generator, and $D(x|y)$ means that corresponding conditional constraint y is added to the discriminant.

Condition-Constraint-Based Network Architecture of Commutator Outer Circular Defect Expansion Image Generation

Based on the principle of constraint generative adversarial network, we define the image generative adversarial network of commutator cylindrical defect expansion based on condition constraint. We construct the network in this paper by combining the design of convolution neural network in DCGAN network. The DCGAN network mentioned above abandons a large number of full-connection and pooling layers in the traditional neural network and adopts a full-convolution neural network for sampling, and the calculation speed is accelerated. As shown in Figure 11, the commutator cylindrical defect expansion image generative adversarial network based on conditional constraints is composed of two parts, namely a generator G and a discriminator D.

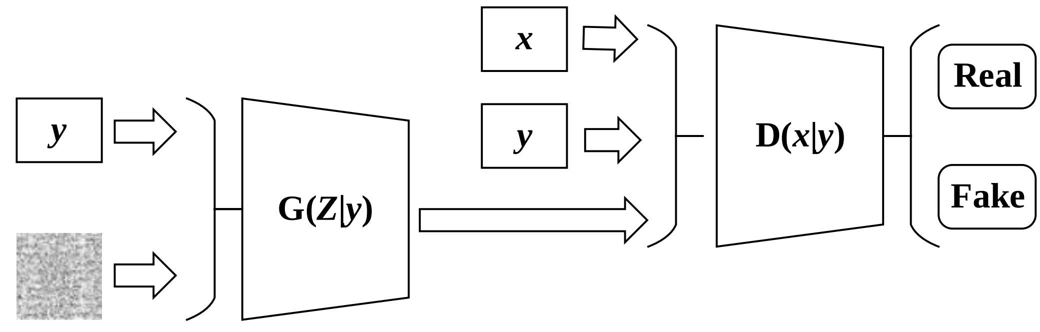

Figure 10. Condition generative adversarial networks architecture

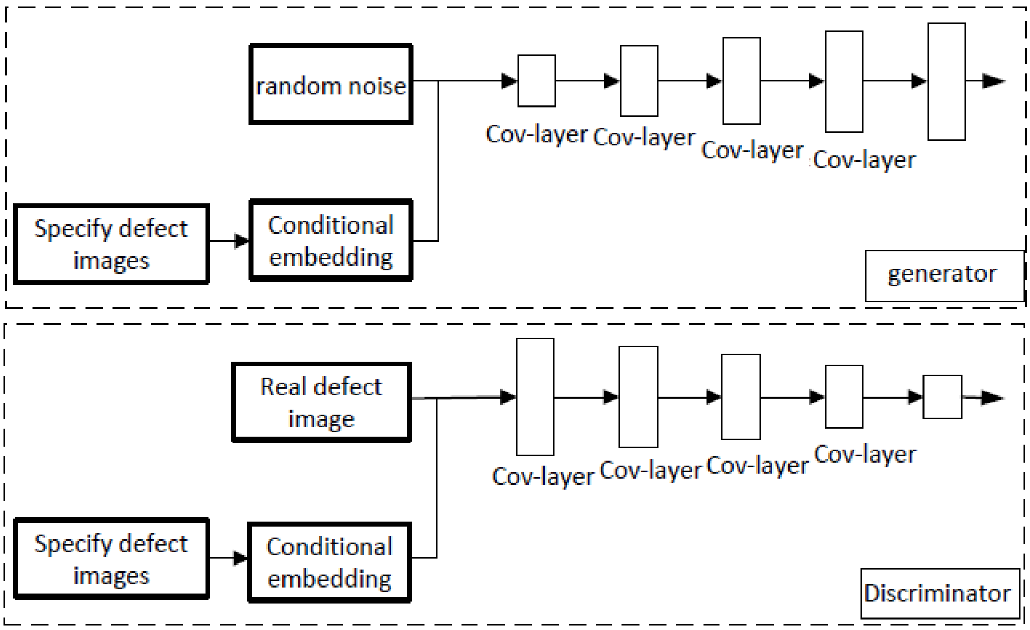

Figure 11. Image generative adversarial network of commutator outer circular defect based on conditional constraints

For the conditional constraint y added in the generator and the discriminator for the purpose of generating a commutator cylindrical expansion image of a specified defect, class label information is added in the generation process of the image, i.e., in the generator model to guide the generation process of constraint data, thereby generating the commutator cylindrical expansion image of the specified defect, and among the discriminators, the discriminator model discriminates whether the commutator cylindrical expansion image generated by the generator is a real image or a generated image and also judges whether the image matches the condition y, i. e. whether or not it belongs to a defective image in the category.

The commutator cylindrical defect expansion image generative adversarial network based on condition constraint mainly comprises a generator and a discriminator. In this paper, the generator and the discriminator in the generative adversarial network are constructed by combining the concept of a convolution neural network, and the original generative adversarial network model is no longer used as the infrastructure of the commutator cylindrical defect expansion image generative adversarial network, In the following, we will describe the generator network and the discriminator network in detail.

Generator network structure

The network architecture of the generated model mainly includes input layer, convolution layer and full connection layer, but the convolution layer in the generated model is fractional step convolution. The activation function is nonlinear activation function ReLU, and the batch normalization layer is added after each convolutional layer. The batch normalization operation can effectively prevent the disappearance of gradient, make the new data distribution more close to the real data distribution, and improve the stability of training. The generator structure is shown in Figure 12:

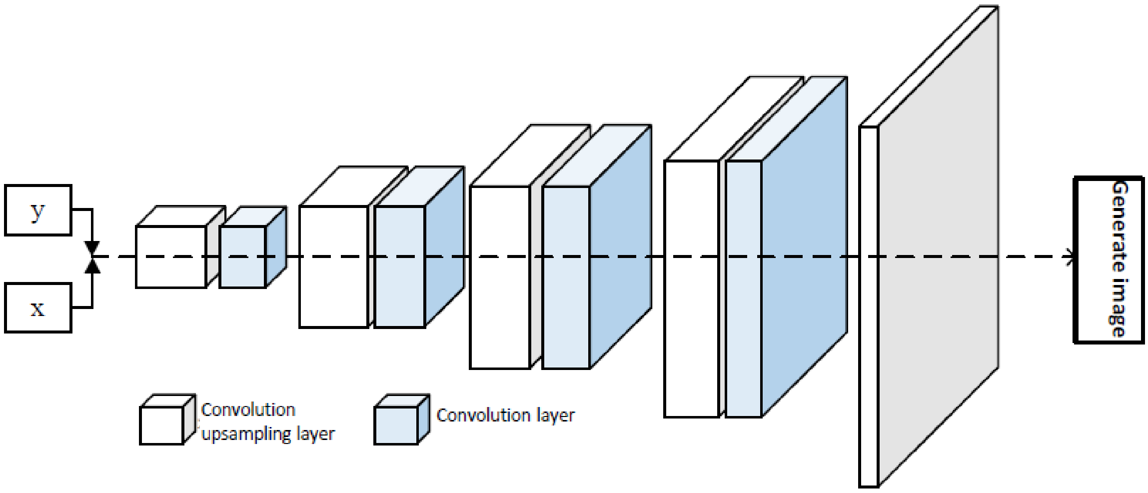

Figure 12. Generator model structure

The details of the convolution layer in the generated model are shown in the table 2 below.

Table 2 Generator model details

Operation	Convolution kernel	Step length	Output channel
Convolution upsampling layer	5×5	1	512
Convolution	3×3	1	256
Convolution upsampling layer	5×5	1	256
Convolution	3×3	1	128
Convolution up-sampling layer	5×5	1	128
Convolution	3×3	1	64
Convolution up-sampling layer	5×5	1	32
Convolution	3×3	1	16
Convolution up-sampling layer	5×5	1	3

Discriminator network structure

The input of the discrimination model is the condition information y and the real image or the condition information y and the generated image. In the network architecture of discriminant model, it is mainly composed of convolution layer. And leaky relu is used in each convolution layer as a non-linear activation function as shown in Figure13,

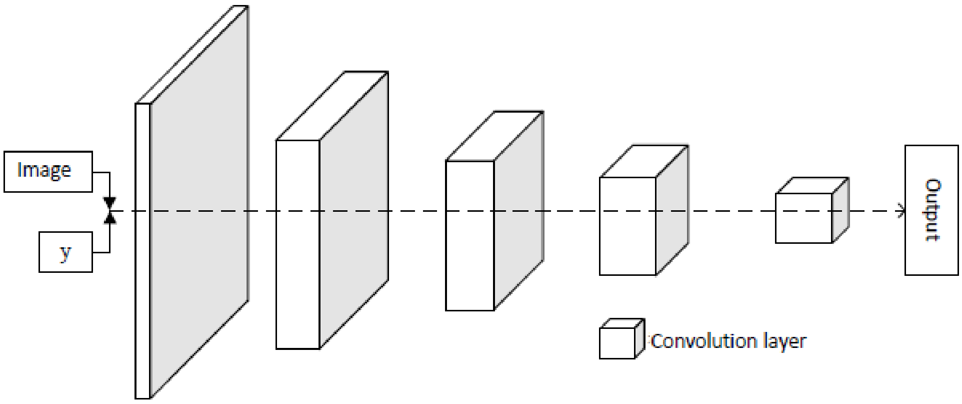

Figure 13. Model structure of raw discriminator

Specific details of convolution layer in the discriminant model are shown in the following table 3.

Table 3. Generator model details

Operation	Convolution kernel	Step Size	Output channel
Convolution	3×3	1	32
Convolution	5×5	1	64
Convolution	3×3	1	128
Convolution	5×5	1	128
Convolution	3×3	1	256
Convolution	3×3	2	512

Experiment and Analysis

Experiment environment setup

In this paper, many image tools are used in the study of commutator cylindrical defect image, as shown in Table 4:

Table 4. Tools used in the algorithm

PROJECTS	Content
Operating system	Linux Ubuntu 16.04、Windows 10
Computing Architecture	CORE i7-9750h + NVIDIA GTX 1080Ti
In-depth learning framework	Tensorflow, Keras
Image proces	Python 3.7、OpenCV 3.0

Evaluation criteria

At present, the image generated by the countermeasure generative adversarial network is mainly evaluated through two aspects, namely a qualitative standard and a quantitative standard, wherein the qualitative standard is mainly based on visual observation, and the quantitative standard mainly adopts a FID score which reflects the difference between the distribution of real data and the distribution of generated image data. The smaller the difference is, the more realistic the generated image is; conversely, the larger the FID score is, the resulting results are unsatisfactory. The calculation formula is as follows:

$$d^2\left((m_r, C_r), (m_g, C_g)\right) = \|m_r - m_g\|_2^2 + Tr\left(C_r + C_g - 2(C_r C_g)^{\frac{1}{2}}\right) \quad (10)$$

Where, d represents FID distance is $m_r, C_r$the feature mean and covariance matrix of the real image, and $m_g, C_g$is the feature mean and covariance matrix of the generated image. FID has more flexible data requirements and can reflect the difference between the real image and the generated image more effectively.

Results and analysis

In the experiment, we compare WGAN-based and DCGAN-based image generation methods, which are compared with the method proposed in this paper. First, Figure 14~Figure 17 show the comparison between the result images generated by different methods and the real image. Because the two comparison methods cannot generate the defect image directionally, Based on the method presented in this paper, we select the corresponding defect images in DCGAN and WGAN respectively.

From Figure 14 to Figure 17, it can be seen that the above three methods can correctly simulate the expansion image of the cylindrical defect of the commutator and generate a more vivid simulation image. However, compared with the other two methods, the generated image of WGAN-based commutator cylindrical defect expansion image generation method is more blurred, which does not vividly express the shape and appearance of defects, so it can be said that the fidelity is not high, and WGAN only learns part of the characteristics. However, as can see from Figure 15 that the generated image generated by the DCGAN-based commutator cylindrical defect expansion image generation method better simulates the shape of the defect. Compared with WGAN, the image is more realistic than the WGAN. However, compared with the original image, the image is distorted which is not in accordance with the common sense. Although the accuracy is high and the image is not blurred, but there is also a certain shape mismatch. Since the DCGAN method cannot directionally generate the defect image, the defect category of the generated image cannot be controlled. However, the defect image generated by this method is similar to the real image. The visual effect is less different from that of the real commutator, and the performance of the defect image is improved compared with that of the DCGAN model. Compared with the DCGAN model, the visual effect is improved.

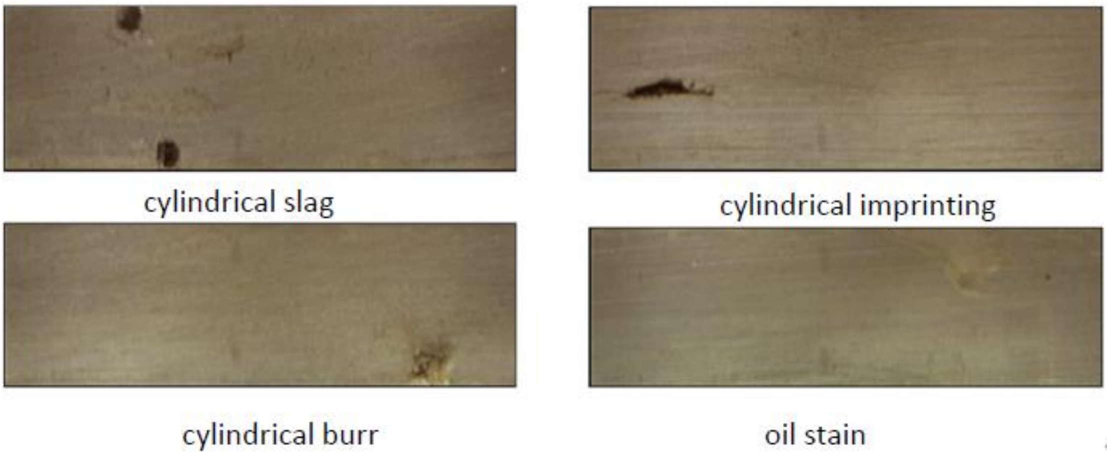

Figure 14. Real image

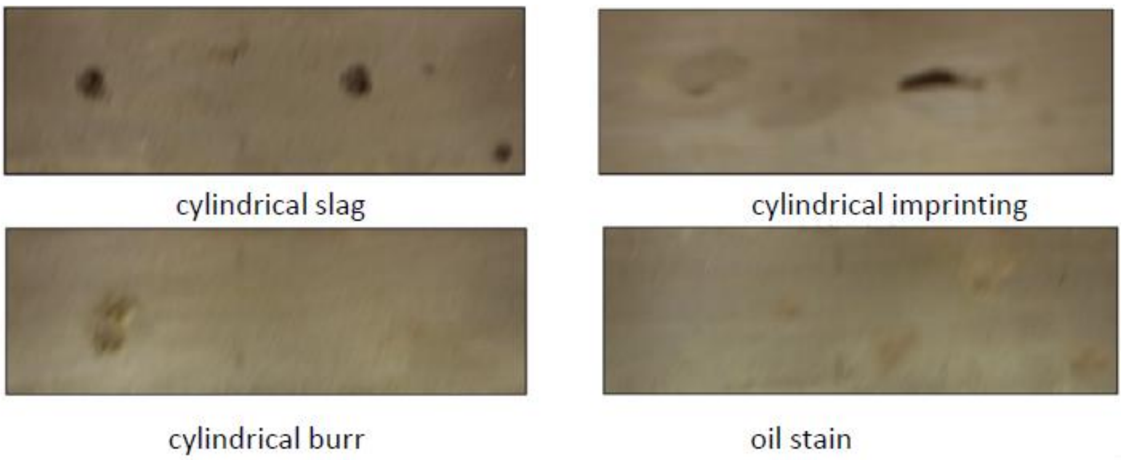

Figure 15. Example of defect expansion image generation method for commutator cylindrical based on WGAN

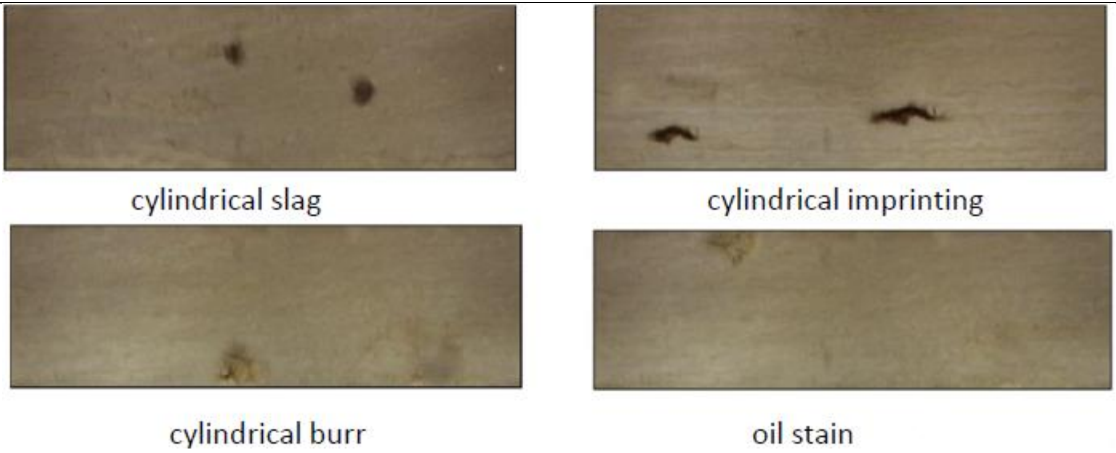

Figure 16. Example of defect expansion image generation method for commutator cylindrical based on DCGAN

In order to verify the above-mentioned judgment based on visual quality more accurately, FID is used in this paper to quantitatively compare the defect images generated by the above three methods, so as to determine the best defect image generation model. We calculated the FID for the above three methods and recorded in Table 5:

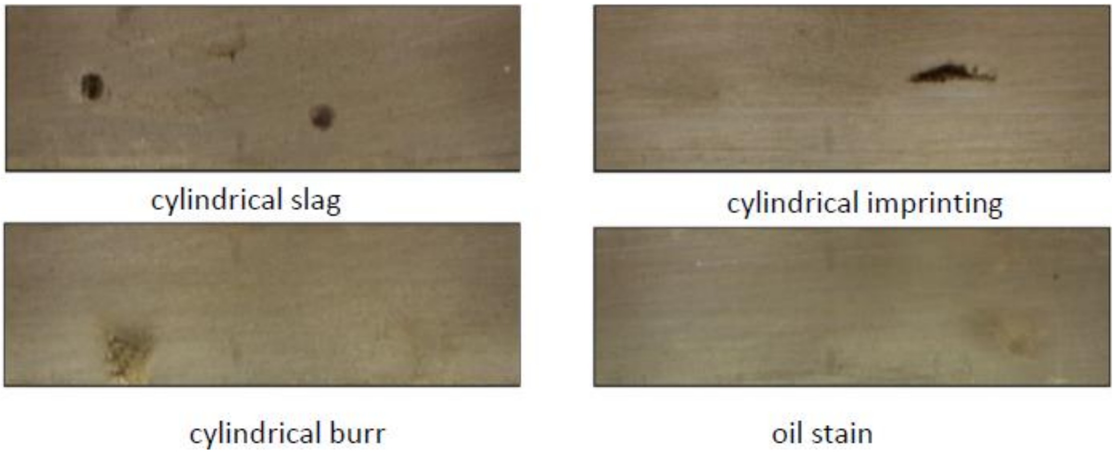

Figure 17. Example of the generation method in this paper

Table 5. FID values for three different defect generation methods

Defect generation model	Cylindrical slag defect	Cylindrical imprinting defect	Cylindrical burr defect	Oil stain
DCGAN	157	169	89	140
WGAN	241	219	135	178
The proposed method	85	67	74	104

As can be seen from Table 5, that the larger the FID value, the greater the difference in data distribution between the generated image and the real data; the proposed method in this paper has the lowest FID value on the generated image of four defects, that is to say, the proposed method has better defect image generation effect than the above two comparison methods, and can generate the generated images closer to the real image data distribution. While the WGAN has the highest FID score, which is consistent with our visual observation. WGAN rubber is quite different from the original image, and WGAN performance needs to be improved.

SUMMARY

In this paper, that principle of image generation method based on generative adversarial network is introduced, and then two classical GAN variant algorithms are introduced. Then we apply the above two methods to the internal control defect expansion image generation method of commutator, and find that it cannot generate data directionally. Therefore, this paper proposes a new method which can generate defect image directionally, and compares with the above two methods.

The experimental results show that the method in this paper has higher FID value than the above two methods, and more similar data distribution than the real image.

ACKNOWLEDGMENTS

Research and Innovation Fund Project of Dongguan Polytechnic: KYCX202401.

Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation. ("Climbing Program" Special Funds: pdjh2024a659).

2022 Innovation Team Project of Colleges and Universities of Guangdong Province (Project No.: 2022KCXTD063, Project Name: Intelligent equipment innovation team based on machine vision).

REFERENCES

[1] *E. Panfilov, A. Tiulpin, S. Klein, M. T. Nieminen, S. Saarakkala.* Improving Robustness of Deep Learning Based Knee MRI Segmentation: Mixup and Adversarial Domain Adaptation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Seoul, Korea, November 2, 2019, IEEE, 2019: 450–459

[2] *G. Fenu, E. Medvet, D. Panfilo, F. Pellegrino.* Mosaic Images Segmentation using U-net, in: 9th International Conference on Pattern Recognition Applications and Methods, Valletta, Malta, 22-24 February 2020, SciTePress Digital Library, 2020: 485–492

[3] *I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al.* Generative Adversarial Networks. Communications of the Acm. 2020, 63 (11): 139–144

[4] *X. Yan, J. Yang, K. Sohn, H. Lee.* Attribute2Image: Conditional Image Generation from Visual Attributes, in: Computer Vision – ECCV 2016, Amsterdam, The Netherlands, October 11–14, 2016, Springer International Publishing, 2016: 776–791

[5] *L. Zhu, Y. S. Chen, P. Ghamisi, J. A. Benediktsson.* Generative Adversarial Networks for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing. 2018, 56 (9): 5046–5063

[6] *S. Reed, A. Van den Oord, N. Kalchbrenner, V. Bapst, M. Botvinick, N. De Freitas.* Generating Interpretable Images with Controllable Structure, in: 5th International Conference on Learning Representations (ICLR 2017), Toulon, France, April 24-26, 2017, ICLR, 2017

[7] *X. Wang, A. Gupta.* Generative Image Modeling Using Style and Structure Adversarial Networks, in: European conference on computer vision (ECCV 2016), Amsterdam, The Netherlands, October 11–14, 2016, Springer International Publishing, 2016: 318–335

[8] *X. Huang, Y. Li, O. Poursaeed, J. Hopcroft, S. Belongie.* Stacked Generative Adversarial Networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, Hawaii, July 22-25 2017, IEEE, 2017: 5077–5086

[9] *E. Denton, S. Chintala, A. Szlam, R. Fergus.* Deep generative image models using a laplacian pyramid of adversarial networks. arXiv preprint arXiv. 2015: 1506.05751

[10] *K. Garg, A. K. Singh, D. Herremans, B. Lall.* PerceptionGAN: Real-world Image Construction from Provided Text through Perceptual Understanding, in: 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 26-29 August 2020, Proceedings of the IEEE, 2020: 1–7

[11] *C. Xu, L. Xu, P. Ohorodnyk, M. Roth, B. Chen, S. Li.* Contrast agent-free synthesis and segmentation of ischemic heart disease images using progressive sequential causal GANs. Medical Image Analysis. 2020, 62: 101668