Fuzzy Systems and Data Mining IX
A.J. Tallón-Ballesteros and R. Beltrán-Barba (Eds.)

doi:10.3233/FAIA231075

RM-GAN: Region Attention Mechanism and Multi-Scale Features for Respirator Defect Generation

Qiang HE[a,1], Xuewu LI[a], Yang LIU[a], Xin YOU[a], Bing LI[a], Yufeng WANG[a] and Yangjie XIAO[b]
[a] *China Southern Power Grid Co., Ltd. Extra High Voltage Transmission Company Guiyang Bureau*
[b] *North China Electric Power University*

Abstract. As a key component to ensure the efficient and safe operation of the transformer, the substation respirator's performance and status have a direct impact on the overall safety of the power system. Aiming at the problem that the defect characteristics of respirators are diverse and unbalanced, this paper proposes a respirator defect generation method based on a regional attention mechanism and multi-scale features, called RM-GAN. First, based on the structural characteristics of the respirator, the model adopts a feature-preserving image preprocessing method and introduces a regional attention mechanism to improve the precise positioning and modeling of respirator components. Then, by combining the multi-scale features of the respirator in the discriminator, defect characteristics can be captured on respirators of different scales, thereby enhancing the accuracy and robustness of the generated results. Finally, experiments were conducted on a custom-built dataset of respirator defects to validate the effectiveness of RM-GAN. The results indicate that RM-GAN is capable of generating high-quality images of respirator defects.

Keywords. substation respirator; GAN; regional attention; multi-scale features; defect generation

1. Introduction

Transformers are important equipment in substations. Once a transformer fails, it will have a very serious impact on the normal operation of the power system. The respirator is an important part of the transformer. It can effectively filter the moisture in the air to reduce moisture and oxidation of the transformer oil. If the respirator becomes discolored and fails due to saturation of adsorbed moisture, it will easily cause the transformer oil to become damp, causing the insulation strength of the transformer oil to decrease, and there is a risk of internal failure of the transformer. Therefore, timely detection and replacement of failed respirator silica gel is crucial to ensure the normal operation of the transformer [1-3].

Currently, the detection of respirator discoloration mainly relies on manual inspection, which not only consumes a lot of manpower and time but also the effect is

[1] Corresponding Author, Qiang HE, China Southern Power Grid Co., Ltd. Extra High Voltage Transmission Company Guiyang Bureau, Guiyang, China; Email: 15831255315@163.com.

also affected by subjective factors, resulting in inaccurate detection. With the rapid development of deep learning and computer vision technology, it is more effective to use computer vision methods to detect respirator discoloration [4-7]. However, due to the complex style of respirator equipment in substations, resulting in an imbalance of positive and negative samples, the effectiveness of current respirator discoloration detection methods is mostly limited by the quality of the data set. Current research on respirator defect recognition and detection mainly focuses on traditional image recognition and deep learning methods and lacks discussion of generative models. This is primarily because respirator defects are diverse and, in most cases, real defect samples are scarce. Therefore, generating diverse and credible respirator defect images with limited defect samples to improve the effect of defect detection is a very important but challenging task.

In recent years, generative adversarial networks [8] (GAN) have received extensive attention and research as a powerful data generation tool [9-10]. Although GAN has achieved certain success in various application scenarios, such as image restoration, data enhancement, etc., there is still a lack of research on defect generation of substation respirators, and many problems and challenges still exist. First, because the structure and working environment of respirators are very complex, and the sizes and forms of respirators are also different; second, because the samples of respirators in different environments are different and the types are diverse. How to accurately generate defect images with various characteristics is still an open problem.

To address these problems, this paper proposes a novel respirator defect generation method based on regional attention mechanism and multi-scale features, called RM-GAN. By introducing a regional attention mechanism and multi-scale features, our model is able to capture the regional characteristics of respirator defects more effectively. At the same time, it can handle respirators of different sizes, which effectively enhances the model's feature extraction capabilities for respirators, thereby improving the quality of generated respirator images. It has achieved good generation results on the experimental data set and can help improve detection performance.

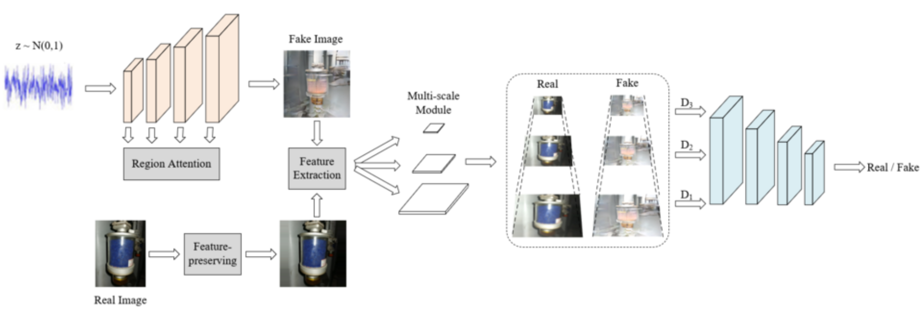


Figure 1. The structure of RM-GAN

2. RM-GAN

The imbalance in the quantity of various respirator types adversely impacts the efficacy of defect detection mechanisms in respirators. This paper proposes a respirator defect generation model RM-GAN based on regional attention mechanism and multi-scale features to solve this problem. Its structure is shown in Figure 1. RM-GAN mainly includes three core improvements: 1) A feature-preserving preprocessing method is designed to ensure that the input image still has the basic structural characteristics of the original respirator, improving the quality and usability of the data. 2) The model integrates a regional attention mechanism to more accurately extract and focus on the features of the respirator area, thereby greatly improving the model's feature recognition capabilities. 3) A multi-scale shared convolution structure is adopted in the discriminator part to more effectively supervise the training of the generator and thereby improve the overall quality of the generated respirator images.

2.1. Feature-preserving Module

Due to the particularity of the respirator structure, the target image usually presents a rectangular shape with significantly different aspect ratios, as shown in Figure 2(a). Deep learning models generally process square inputs. Conventional resize preprocessing methods will cause distortion to the shape characteristics of the respirator image, as shown in Figure 2(b). This method will affect the accuracy of subsequent generated models. In order to solve this problem, this paper designs a feature-preserving preprocessing method, as shown in Figure 2(c). This method avoids directly stretching the original image but uses an edge expansion strategy to fill in the background information. This method not only meets the size requirements of the model input but also preserves the shape features of key parts of the image. Furthermore, this preprocessing method also plays a role in guiding the model's attention by retaining key areas of the image. It provides higher-quality input data for the subsequent regional attention module, thereby further strengthening the model's ability to identify and analyze specific regions.

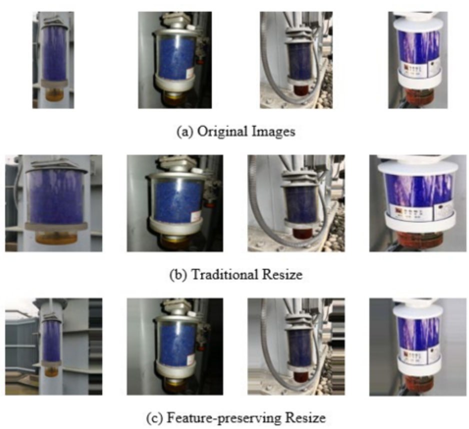


Figure 2. Examples of input images with different preprocessing methods

2.2. Region Attention Mechanism

Breathers usually only occupy a sub-region of the image and have significant regional color features, and traditional attention mechanisms perform poorly in this task. We propose an efficient regional attention module targeting the overall structural characteristics of respirator images. This module performs feature aggregation in the X and Y directions and encodes it into a region-sensitive attention map, thereby more effectively enhancing the feature representation capabilities of the region of interest, aiming to capture regional information related to the respirator.

The regional attention module is shown in Figure 3, and the inputs are defined as $I \in R^{C\times H\times W}$. The input is encoded through a 1×1 convolutional layer and an average pooling layer in the X and Y directions. The encoded result is concatenated and then passed through a 1×1 convolutional layer followed by an activation layer, resulting in the generation of two independent tensors $t^h \in R^{C\times H\times 1}$ and $t^w \in R^{C\times 1\times W}$. After the matrix product, the regional attention feature map is obtained. The calculation process is as follows:

$$f_r = \frac{1}{N}\sum_{0\le n\le N} Sigmoid(t_c(h)\bullet t_c(w)) \in R^{1\times H\times W} \tag{1}$$

where c denotes the features of the c-th channel. To preserve global information, a 1×1 convolution layer with a single output channel and a sigmoid activation function is employed. The calculation formula is as follows:

$$f_g = Sigmoid(Conv_{c=1}(I)) \in R^{1\times H\times W} \tag{2}$$

where I denotes the input features. Finally, the global and regional attention feature maps are accumulated and multiplied with the input to obtain the final output of the regional attention module. The output formula is as follows:

$$O = (f_r \oplus f_g) \otimes I \tag{3}$$

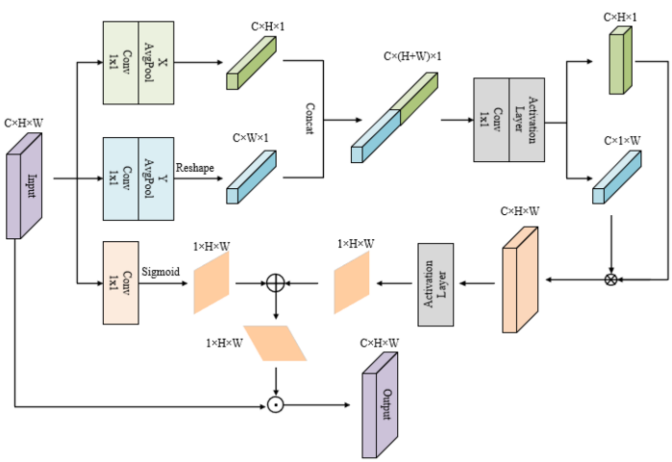


Figure 3. The structure of region attention

2.3. Multi-scale Discriminator

Due to the different locations where the inspection images were taken, the proportion of the respirator in the image is significantly different. In order to accurately identify respirators of different sizes, the discriminator needs to introduce multi-scale feature inputs. Therefore, we propose a multi-scale discriminator to process images of different scales, represented by D_1, D_2, and D_3 respectively, and adopt a shared convolution architecture. Specifically, the input image is passed through a shared convolution layer to extract the features of the sample and obtain the corresponding feature map. Then, the feature maps of the real samples and the generated samples are downsampled with different sampling factors to obtain images of three different scales. Then D_1, D_2, and D_3 process feature maps of different scales respectively, and learn different discriminator parameters from them. The learning formula of the multi-scale module is as follows:

$$L = \min_{G} \max_{D_{1,2,3}} \sum_{i=1,2,3} E_{x\sim p(x)}[\log D(x)] + E_{x\sim p(z)}[\log(1-D(G(z)))] \tag{4}$$

3. Experiment

3.1. Dataset

In accordance with the substation equipment labeling specifications, this paper uses the VOC2007 dataset construction method to construct a professional substation inspection respirator defect detection image dataset. The dataset has a total of 1909 images, including 4 types of respirators, of which discoloration of the respirators is the main defect. The details are shown in Table 1:

Table 1. Respirator Image Dataset

Types	Nums	Types	Nums
normal with shell	878	normal without shell	105
discoloration with shell	824	discoloration without shell	102

3.2. Image Generation Results and Analysis

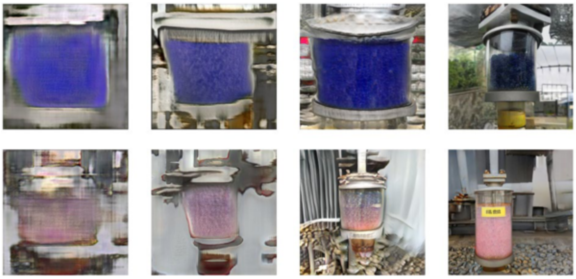
In the generation experiment, to reflect the superiority of this method, comparative experiments were conducted on the same data set between this method and several popular generative models at this stage. The quantitative experimental results are shown in Figure 4. In addition, this paper selects the commonly used evaluation indicators FID (Frechet Inception Distance) and IS (Inception Score) of generative adversarial networks as a basis for fairly measuring the quality of generated images. The definitions of FID and IS are as follows:

$$IS(G) = \exp\left(E_{x\sim Pg} KL(p(y\,|\,x)\,\|\,p(y))\right) \tag{5}$$

$$FID(x,g) = \|\mu_x - \mu_g\|_2^2 + Tr(\Sigma_x + \Sigma_g - 2(\Sigma_x\Sigma_g)^{0.5}) \tag{6}$$

where KL represents the Kullback-Leibler Divergence, Tr denotes the trace, μ and Σ represent the mean and covariance matrix of the image feature vectors, respectively.

Specifically, we utilized the generative model to produce 5,000 synthetic samples and computed the distance between them and the real dataset, yielding the FID and IS values. Ideally, we aspire for the generative model to exhibit a high IS and a low FID. The qualitative experimental results are shown in Table 2.



(a) DCGAN (b) ACGAN (c) StyleGAN-ADA (d) RM-GAN

Figure 4. The generation results of different models

Table 2. Experimental comparison of different models

Models	FID	IS
DCGAN[11]	174.83	1.34
ACGAN[12]	148.59	1.58
StyleGAN-ADA[13]	124.03	1.95
RM-GAN(ours)	**112.48**	**2.23**

This method is compared with the classic generative models DCGAN, ACGAN, and StyleGAN-ADA. As can be seen from Figure 4, the respirator image generated by our proposed method is more realistic and the target respirator is generated more accurately. We believe that this is because the regional attention mechanism and multi-scale module of RM-GAN play a role in the color area and different proportions of the respirator respectively, thus improving the quality of the generated respirator. Our generation method RM-GAN not only has the best visual effect but also achieves the best value in evaluation indicators. RM-GAN has the lowest FID, indicating that the respirator defect samples generated by our method are closer to the data distribution of real samples and have better generation quality. RM-GAN has the highest IS, indicating that the generated samples have better diversity. RM-GAN is significantly better than other generation methods in terms of evaluation indicators, which corresponds to the effect in Figure 4.

In order to further verify the effectiveness of each module on the RM-GAN network, ablation experiments were performed as shown in Table 3. As can be seen from Table 3, each module plays a role in improving performance. Among them, the regional attention and multi-scale discriminator have more obvious improvements in performance. We believe this is because they have improved the network architecture based on the characteristics of respirator defect samples, so they can achieve greater improvements. The feature preservation preprocessing is an adjustment to the original image. Although it has been improved, it is not as obvious as other modules. In addition, experiments show that using three improvements simultaneously can achieve the best results.

Table 3. Ablation experiment

Feature-preserving	Region Attention	Multi-scale D	FID	IS
√			126.93	1.91
	√		117.65	2.01
		√	116.72	2.12
√	√	√	**112.48**	**2.23**

3.3. Detection Results and Analysis

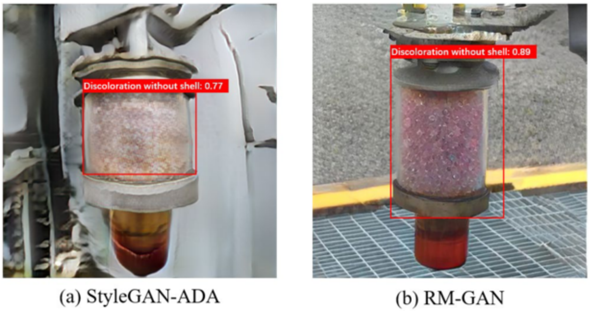


Figure 5. Detection display of images generated by different models.

We perform detection and display on the generated respirator defect images to reflect the advantages of this method. The detection effect is shown in Figure 5. Among them, Figure 5(a) is the detection result of the respirator generated by StyleGAN-ADA, and Figure 5(b) is the detection result of the respirator generated using this method RM-GAN. As can be seen from the detection result diagram, the detection effect on the image generated by RM-GAN is significantly better than the detection effect on the image generated by StyleGAN-ADA. Since the image quality generated by the latter is poor, it can easily lead to false detection; while the image generated by our method is still far from the real image, but it can accurately detect the regional part of the respirator. This shows that the method in this paper can improve the performance of detection and the quality of the generated results is better than other methods.

4. Conclusion

This paper proposes a respirator defect generation method based on regional attention mechanism and multi-scale features, called RM-GAN. By introducing a feature-preserving image preprocessing method and a regional attention mechanism, the precise positioning and modeling of respirator components are improved. At the same time, combining the multi-scale features of respirators in the discriminator can capture defect characteristics on respirators of different scales, thereby enhancing the accuracy and robustness of the generated results. Through experiments, it has been verified that RM-GAN can generate high-quality respirator defect images. This method is of great significance for ensuring the safe and stable operation of transformers.

References

[1] Liu S. Treatment of discoloration of silica gel in transformer respirators [J]. Popular Electricity, 2014, 29(10): 30-31.
[2] Xie W. Analysis on the treatment of silicone discoloration of transformer respirators [J]. China High-tech Enterprises, 2016(25):152-153.
[3] Li C, Zeng Q, Zeng L, et al. Operation risk analysis and operation and maintenance strategy research and application of main transformer breather [J]. Electrical Switch, 2022, 60(05): 84-87.
[4] Xin M, Xiao X, Liu B, et al. Automatic detection of transformer respirators based on image processing [J]. Practice and Understanding of Mathematics, 2018, 48(02): 127-133.

[5] Li R, Xu D, Zhai D, et al. Intelligent identification method of transformer respirator defects based on HSV feature transformation and target detection [J]. High Voltage Technology, 2020, 46(09): 3027-3035.
[6] Shi R. Power transformer respirator discoloration detection method based on adaptive threshold segmentation [J]. Jiangxi Electric Power, 2022, 46(08): 36-41+44.
[7] Xiang X, Pan L, Wu H. Respirator defect detection method based on feature matching and deep learning [J]. Laboratory Research and Exploration, 2022, 41(06): 57-61+97.
[8] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[J]. Advances in neural information processing systems, 2014, 27.
[9] Huang W, Xu A, Ming Z, et al. Method for predicting substation worker behavior based on generative adversarial network [J]. China Southern Power Grid Technology, 2019, 13(02): 45-50.
[10] Yang P, Li J, Hao R. Small sample image data enhancement technology based on generative adversarial network [J]. Network Security and Data Governance, 2023, 42(06):79-84+102.
[11] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
[12] Odena A, Olah C, Shlens J. Conditional image synthesis with auxiliary classifier gans[C]//International conference on machine learning. PMLR, 2017: 2642-2651.
[13] Karras T, Aittala M, Hellsten J, et al. Training generative adversarial networks with limited data[C]. Advances in neural information processing systems, 2020, 33: 12104-12114.