Advances in Artificial Intelligence, Big Data and Algorithms
G. Grigoras and P. Lorenz (Eds.)
© 2023 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA230848

Improved Faster RCNN-Based Metal Particle Detection on the Inner Surface of GIS Equipment

Chao LV[1], Chunyu SHI, Baoquan XU
Changchun University of Science and Technology, China

Abstract. To focus on the problem of low detection rate of metal particles on the inner surface of GIS equipment, an improved Faster RCNN GIS equipment inner surface metal particle target detection network is proposed.The residual network is introduced as the feature extraction network based on the feature of small metal particles on the inner surface, and the feature pyramid network and receptive field block are combined with the residual network,three feature extraction networks, VGG16, ResNet50,ResNet50+RFB-FPN, are used to experimentally compare 2000 metal particle maps.The results show that the accuracy and recall of ResNet50+RFB-FPN feature extraction network are 97.3% and 83.4%, which are improved compared with both VGG16 network and ResNet50 network. It is concluded that the improved Faster RCNN-based metal particle identification on the inner surface of GIS equipment can meet the requirements of detection accuracy and speed.

Keywords. Faster RCNN GIS RFB-FPN

1. Introduction

High voltage technology is developing the voltage level is increasing, and the requirements for gas insulated metal-enclosed switchgear (GIS) are getting higher and higher. In China, with the constant development of the electricity system and the rapid development of electric energy construction, GIS equipment has a broad application space. And with the expansion of urban scale in recent years, the area has become an important influencing factor, and the demand for GIS equipment is growing, but inside the GIS, the generation of metal particles is inevitable. The existence of metal particles will greatly reduce the insulation performance of GIS.

In the GIS inner surface metal particles cleaning, in the equipment assembly and outage maintenance are mainly manual cleaning, the efficiency and accuracy of manual detection of particles are not high. The method of using computer to identify metal particles becomes effective and more accurate in handling metal particles on the inner surface of GIS equipment. Faster RCNN has been extensively applied in different fields since its introduction [3], and the innovative region proposal network is able to speed up detection network running Time more effectively than its predecessor Fast RCNN [2]. In this study, we propose an improved Faster RCNN model for detecting metal particles

[1] Corresponding Author: Chao LV, Changchun University of Science and Technology, China;
Email: lcalcjou@sina.com

on the inner surface of GIS, introducing FPN network and RFB module into the residual network, using ROI Align for regional feature aggregation, and fusing multi-layer feature maps for particle identification. [1]

2. Fundamental Theory

2.1. Faster RCNN framework

The Faster RCNN framework is composed of 4 parts: feature extraction network, RPN, target classification, and edge regression.

2.2. RPN

The main role of RPN is to generate a region suggestion frame on the features after the convolution layer, and to fine-tune the suggestion frame while determining whether it is the foreground or background. The suggestion frames used in this study are composed of 5 areas (32^2 , 64^2 , 128^2 , 256^2 , 512^2) and 3 aspect ratios (1:1, 1:2, 2:1), and each centroid corresponds to 15 suggestion frames on the original map, which is more favorable for extracting small-sized particles. The generated suggestion frames will be entered into two branches of classification network and regression network. The classification network determines whether the proposed box is foreground or background, represents the overlap IOU between the suggestion box and the mark box . The formula for calculating the overlap degree is given in equation (1)

$$IoU = \frac{Detection\ \text{result} \cap GroundTruth}{Detection\ \text{result} \cup GroundTruth} \tag{1}$$

After the suggestion frames generated by the aforementioned network and ignoring the suggestion frames crossing the image boundaries, the IoU is set to 0.7 due to the large overlap between the suggestion frames, and the loss function of the whole RPN is shown in Eq. (2) according to the score of the suggestion frames in the classification network, the classification Eq. (3) calculates the loss function. and the boundary frame regression loss function is calculated by Eqs. (4)(5).

$$L\left(\{p_k\},\quad \{tk\}\right) = \frac{1}{N_{cls}} \sum L_{cls}\left(p_k, p_k^*\right) + \lambda \frac{1}{N_{reg}} \sum p_k^* L_{reg}\left(t_k, t_k^*\right) \tag{2}$$

$$L_{cls} = -\left[p_k^* \log p_k + \left(1 - p_k^*\right)\log\left(1 - p_k\right)\right] \tag{3}$$

$$L_{reg}\left(t_k, t_k^*\right) = Smooth_{L1}\left(t_k - t_k^*\right) \tag{4}$$

$$smooth_{L1}\left(t_k - t_k^*\right) = \begin{cases} 0.5\left(t_k - t_k^*\right)^2 & \text{if } |x| < 1 \\ \left|t_k - t_k^*\right| - 0.5 & \text{otherwise} \end{cases} \tag{5}$$

Where: k is the *k-th suggestion* frame; p_k represents the probability that the suggestion frame is a positive sample; positive samples is 1 and negative samples p^* is 0; t_k represents the 4 coordinate parameters; t_k^* represents the coordinate parameters the marker frame corresponding to t_k ; L_{cls} representative classification loss; L_{reg} representative regression loss; N_{cls} is the number of suggestion frames participating in training; N_{reg} is the number of feature images; λ is the balance parameter.

3. Faster RCNN with Fused Multi-Scale Features

GIS Metal particles are small and dense compared to traditional target recognition objects. To enchance the efficiency and accuracy of recognition, deepening the convolution layer features can be extracted efficiently, but the Faster RCNN feature extraction network, VGG16, will have gradient disappearance or explosion when deepening the depth, so the residual network (ResNet50) is chosen to extract particle features. [4] [10]

The Faster RCNN network predicts the object is the image which has been convolved several times, and the features which has been convolved several times retains little detail information, and the low-level features contains more information in the convolution process is not used for subsequent prediction, and some information of the target is lost, so small target recognition is not satisfactory. [9] Since the metal particles are smaller targets, adding the low-level features information is more beneficial for the recognition of small targets, and this study adds a feature pyramid network (FPN) and receptive field block (RFB) to the residual network.[5]The network architecture is illustrated in Figure 1-2. The sum feature maps are sampled and added to the previous layer's feature maps, and the summed feature maps are fused with the RFB module for secondary feature fusion to improve semantics, so that the feature maps used for detection contain multiple stages and multiple levels of features, which is advantageous for target detection.[6][8]

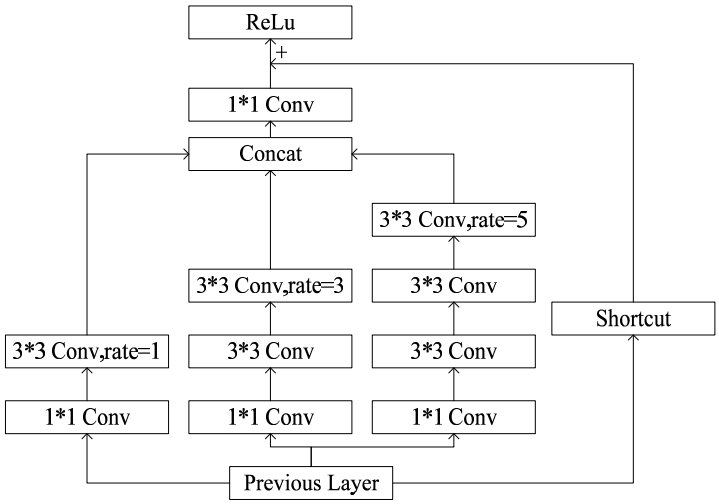

Figure 1. RFB module structure diagram

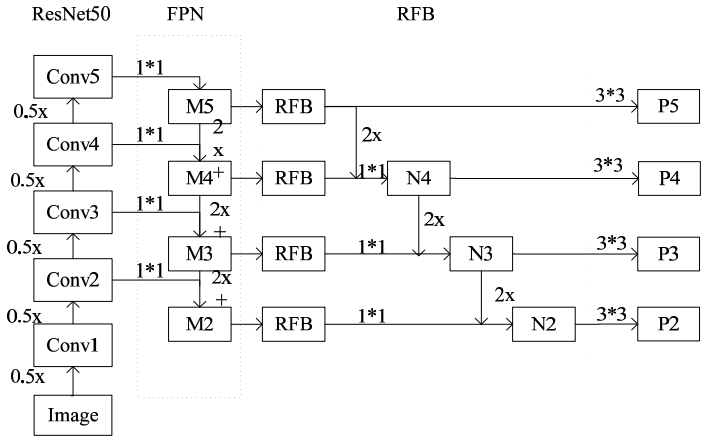

Figure 2. RFB-FPN feature extraction network fusion structure diagram

After the fused feature map enters the RPN to generate the suggestion box, the suggestion box needs to be mapped back to the original feature map. The method to determine which feature layer k the suggestion box comes from is shown in Eq. $K=K_0+\log_2\sqrt{wh}/224$ where: k_0 is the base value, set to 4; w and h are the suggested field's length and width; 224 is the standard pre-training size; and k represents the feature layer used, for which the calculation results are rounded.

4. Experiment and Results

4.1. Test Environment and Data

First paragraph A total of 2000 images of particles were collected, and the two metal particle categories were linear and granular [7]. The data set was labeled with labelimg tool and the labeling file was saved in xml format. 80% of the images for training and 20% for testing. In this study, VGG16, ResNet50 and ResNet50+RFB-FPN were used as the feature extraction networks based on Faster RCNN network to train the metal particles images on the inner surface of GIS To validate the validity of the improved model. Test platform configuration is windows10, CUDA10. 2, python3.7, processor is Intel i5-10200X, graphics card is NVIDIA GTX 3050, and Pytorch is used as the development environment. Training parameters: each batch contains 4 images, initial learning rate is 0.005, a momentum value of 0.9, Gama is 0.33， train the model until convergence.

4.2. Target Detection Results of the Improved Network

The validity of the model was evaluated with precision, recal and F1. Precision is the proportion of all particles that have a predicted value and a true value; recall is proportion of particles with both predicted and true values in the sample of particles with true values; F1 is the sum of the precision and recall of the model, and F1 is 1. The bigger the value, the better the model. The formula is as follows:

$$P = \frac{TP}{TP + FP} \times 100\%$$

$$R = \frac{TP}{TP + FN} \times 100\%$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \times 100\%$$

Where TP is samples with identical prediction values and true values of 1; FP denotes the samples with a predicted value of 1 and a true value of 0. FN represents samples that are predicted to be 0 and true to 1.

The training loss diagram of the two metal particles is shown in Fig 3.

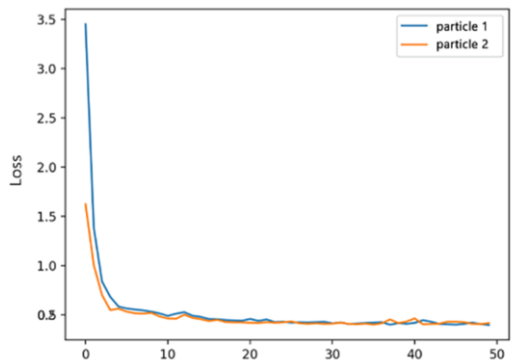

Figure 3 Training loss diagram of the two metal particles

Three different feature extraction models of VGG16, ResNet50, and ResNet50+RFB-FPN were replaced in the Faster RCNN network, and evaluation metrics are presented in Table 1. The comparison of loss values and accuracy of three feature extraction networks is in Fig 4.

Table 1 Model evaluation parameters are extracted from different features

Backbone	Accuracy/%	R rate/%	F_1 /%	Recognition speed/s
VGG 16	95.4	80.4	87.8	0.0193
ResNet 50	95.5	81.2	88.5	0.0200
ResNet50+RFB-FPN	97.3	83.4	90.4	0.0224

The 80 images (20 images for particle 1 and particle 2 and 40 images for particle 1 and particle 2 mixed) were tested individually and input into the established particle detection model, and the recognition accuracy (consistency with manual discrimination) of each particle in the images was used as the evaluation criterion, and the recognition results were output for each image. The results of the recognition data are shown in Table 2.

Table 2 Identification results of particulate testing

Index	Particle 1	Particle 2	Mixed Particles
Totally particles/ piece	160	135	364
Accuracy number/ piece	152	132	353
Accuracy rate/%	95.0	97.8	97.0

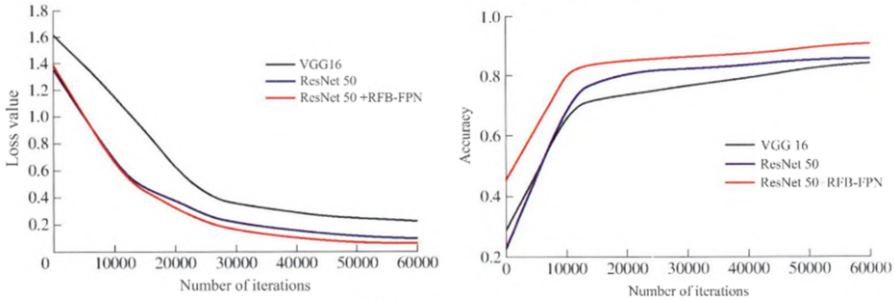

Figure 4 Comparison of loss values and accuracy of three feature extraction networks

From the above experimental results, the accuracy and recall rate of ResNet50+RFB-FPN as backbone improved by 1.9 and 3 percentage points, respectively, compared with the VGG16 network, and the accuracy and recall rate improved by 1.8 and 2.2 percentage points, respectively, compared with the ResNet50 feature extraction network. It proves that ResNet50+RFB-FPN as feature extraction network is more suited for metal particle detection than VGG16 and ResNet50. In terms of detection speed, the improved ResNet50+RFB-FPN feature extraction network improves the recognition effect while sacrificing some of the recognition speed, but the recognition speed meets the requirement of real-time particle recognition. Particle 1 is the yellow box and particle 2 is the red box,three feature extraction networks are shown in Figures 5 - 7.

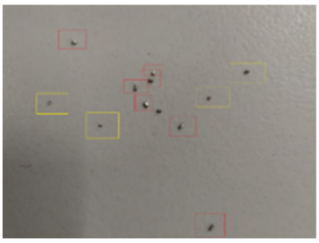

Figure 5 Example of VGG16 recognition effect

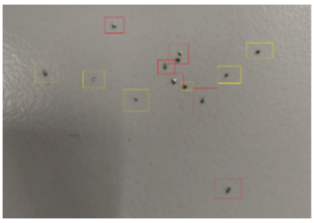

Figure 6 Example of ResNet50 recognition effect

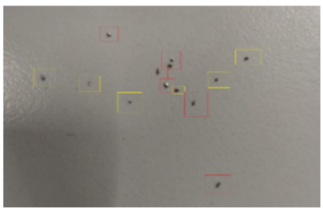

Figure 7 Example of ResNet50+RFB-FPN recognition effect

It can be seen that when VGG16 is used as the backbone, the features of small metal particles are not extracted effectively, and there is a situation of missing detection, and the positioning of the prediction frame is poor; when ResNet50 is used as the backbone, the positioning accuracy of the prediction frame is improved compared with that of VGG16, but the extraction is not complete when there are more metal particles on the same map, and there is mainly a situation of missing detection for small metal particles. When ResNet50+RFB-FPN is used as the backbone, the extraction of metal particles of all sizes is improved, and the prediction frame is more accurate and less likely to be missed. The average precision of the modified neural network is 97.3%. Thus, the feature extraction network of ResNet50 + RFB-FPN is more than 95% for the characteristic of metallic particles., and a lower accuracy of 92.6% for metal particles with slender and poor reflective properties.

5. Conclusion

Aiming at the problems of poor detection ability and low detection rate of metal particles on inner surface, an improved detection algorithm of metal particles on inner surface of GIS equipment based on Faster RCNN was proposed. By replacing the feature recognition network to deepen the network and improve the recognition effect of particles, and adding the feature pyramid networks (FPN) and receptive field block (RFB) to fuse the multi-scale features, and using ROI Align for regional feature aggregation, the recognition accuracy of small target metal particles is improved, and then the overall recognition effect of the network on metal particles is improved. After experimental validation, the improved Faster RCNN can achieve 97.3% recognition accuracy and 83.4% recall rate for particles. However, the accuracy of the model can be improved for the particles (particles 1)with poor reflective properties, and we will continue to use more different methods to improve the recognition of such particles in the future.

References

[1] REN, SHAOQING, HE, KAIMING, GIRSHICK, ROSS, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149. DOI:10.1109/TPAMI.2016.2577031.

[2] Cai Z, Vasconcelos N. Cascade R-CNN: Delving into High Quality Object Detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 6154-6162

[3] Danyang Cao,Shaobo Yang. A Method based on Faster RCNN Network for Object Detection[J]. Recent Advances in Computer Science and Communications,2022,15(9).

[4] HE K，ZHANG X，REN S，et al. Deep Residual Learning for Image Recognition［C］ // 2016 IEEE Conference on Computer Vision and Pattern Recogni- tion. IEEE Computer Society，2016

[5] Tsung-Yi Lin,Piotr Dollár,Ross B. Girshick,Kaiming He,Bharath Hariharan,Serge J. Belongie. Feature Pyramid Networks for Object Detection.[J]. CoRR,2016,abs/1612.03144.

[6] Redmon J, Farhadi A.2017.YOLO9000: Better, Faster, Stronger. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA,6517-6525.

[7] LIN, TSUNG-YI, GOYAL, PRIYA, GIRSHICK, ROSS, et al. Focal Loss for Dense Object Detection[J]. 2020,42(2):318-327. DOI:10.1109/TPAMI.2018.2858826.

[8] GIRSHICK R，DONAHUE J，DARRELL T，et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Confe - rence on Computer Vision and Pattern Recognition. Columbus，USA：IEEE，2014

[9] Ren S Q, He K M, Girshick R, et al. Object detection networks on convolutional feature maps[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(7): 1476-1481
[10] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, America, 2016: 770-778.