ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

A Study on Defect Detection of YOLOV8 Insulators Based on Improvements

Sun Xia[1,*], Qi Shuo[1]

[1]College of Intelligent Equipment, Shandong University of Science and Technology, Tai'an, 271000, China
wrenrainy@163.com
[*]Corresponding author

Abstract: *The aim of this paper is to enhance the performance of the target detection model in the insulator defect detection task. The YOLO v8 detection model is first improved by designing the Backbone-DCNv3 module, which is significantly enhanced in feature extraction and target localization capabilities. Meanwhile, the introduction of MPDIoU loss function further improves the detection accuracy of the model. Tests on the same dataset show that compared with the original model and other models in the YOLO series this paper's model exhibits superior performance in key evaluation metrics such as precision rate, recall rate and average precision. They for precision was 91.45%, recall was 89.47%, and mAP was 91.89%, respectively. Compared with the original model, the improvement was 2.13%, 2.02%, and 2.62%, respectively. Therefore, the research in this paper not only provides new ideas and methods in the field of target detection, but also provides strong technical support for the practical engineering application of insulator defect detection.*

Keywords: *Insulator defect detection; YOLO v8; Deep learning*

1. Introduction

Condition monitoring of insulators is a crucial aspect in the safe operation of power systems. As a key component of transmission lines, insulators need daily inspection to guarantee the reliability of their structure and performance. The traditional inspection of insulators needs to rely on manpower which is time-consuming and laborious, with the rapid development of deep learning technology, the automatic insulator defect detection method based on image analysis has become a hot spot of research, which provides strong technical support to solve the intelligent inspection.

In a related study, Tang Lu et al. (2024) used an improved YOLOX target detection network to accurately locate and identify insulator states in foggy conditions with an accuracy of 0.963 and an average accuracy of 0.917 [1]. Wenbin et al. (2024) proposed a lightweight network (Multi-Defect Detection Network (MDDNet)) for multi-defect detection of insulators on transmission lines, which is an algorithm that focuses on the joint detection of multiple types of insulator defects in the form of insulator arc burns and insulator umbrella skirt breakage. The mean average precision of the model (mAP0.5) reaches 92.1%[2]. Changle Zhang et al. (2024) proposed an insulator fault detection algorithm based on YOLOv5s network. Compared with the original YOLOv5s model, the mAP (mean Average Precision) is improved by 3.6%, the speed by 5 FPS, and the size was only 13.8 MB, which is helpful for the development of intelligent inspection work[3].

Therefore, this paper proposes an insulator defect detection method based on the improved YOLO V8. The C2f module in the model is improved, the third generation deformable convolutional network (DCNv3) was introduced into the detection model, and the ackbone-DCNv3 module has been designed, while the MPDIoU loss function was adopted in order to improve the detection performance of the model.

2. Experiments and Methods

2.1. Image Acquisition

In this study, a total of 896 insulator images were collected using the Baidu AI Studio public dataset of high-voltage transmission line insulator dataset, and some of the images shown in Figure 1. The

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

labelimg professional image annotation tool was used to carry out the annotation work. Labelimg was an open-source image annotation software, supporting rectangular box annotation, polygonal annotation and other annotation methods, and the interface was simple and easy to operate. The labelimg tool was used to successfully construct a high-quality, accurately labeled high-voltage transmission line insulator image dataset. This dataset provides strong data support for subsequent image processing and analysis.

Figure 1: Image of partial insulator defects

2.2. YOLO V8 model

The YOLOv8n model structure consists of three key parts: Backbone, Neck and Head [4].

In the Backbone part, the CSP idea is adopted to extract features through a 3x3 convolution operation (Conv module), and residual concatenation is introduced to avoid gradient vanishing and alleviate the training difficulty.The C2 module converts the feature maps into fully connected layer inputs, while the SPPF module is used to receive pooling windows at different scales for pooled splicing of the feature maps.

The Neck part, on the other hand, mainly implements bottom-up feature extraction and top-down feature fusion. Through the upsampling module, the feature map size is enlarged by a factor of 2 using the neighborhood interpolation algorithm, followed by concatenating the Backbone-extracted feature map with the upsampled feature map through the Concat module in order to fuse the low-level and high-level semantic information. In addition, the Neck part of the Feature Pyramid (FRN) is downscaled and channel number adjusted by convolution operation.

Finally, the detector at the Head end consists of a series of convolutional and fully-connected layers, which are responsible for localizing and identifying the target on the feature map, producing predictions of the location and class of the target bounding box [5].

2.3. Improvements to the YOLO V8 model

2.3.1. Backbone- DCNv3

Traditional convolutional networks (Convolution, Conv) have limitations in adapting to local feature variations, which usually need to rely on more convolutional kernels to extract more complex feature maps, an approach that is less efficient [6]. The receptive field covered by the output transform of this convolution is always rectangular, and due to the accumulation of cascading convolution effects, the receptive field will gradually expand, resulting in irrelevant background information mixed in the final output transform, especially in complex real-world environments, where the use of convolution alone for extracting insulator features is not ideal. For this reason, Deformable Convolutional Networks v3 (DCNv3) was proposed to replace the C2f module in the YOLOv8 backbone network to enhance the extraction of features for insulator defects. As shown in Fig.2 for Backbone- DCNv3, DCNv3 makes three key design improvements. First, a multi-group mechanism is introduced, which allows the network to have greater flexibility and adaptability in dealing with different types of features; second, convolutional neurons can now share weights, and this sharing mechanism improves the parametric efficiency of the model; finally, by normalizing the modulation scalar along the sampling points, DCNv3 allows for more fine-grained control of the information flow, which further enhances the model's expressive and adaptive capabilities [7].

The formula for DCNv3 is expressed as Equation 1:

$$\boldsymbol{y}(p_0) = \sum_{g=1}^{G} \sum_{k=1}^{K} w_g m_{gk} x_g \left(p_0 + p_k + \Delta p_{gk}\right) \tag{1}$$

where G denotes the total number of aggregated groups. For group g, denotes the position-independent projection weight of the group, where denotes the dimension of the group. denotes the modulation scalar of the kth sampling point in the gth group, normalized along the dimension K with a softmax function. denotes the sliced input feature map. is the offset in group g corresponding to the grid

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

sampling position.

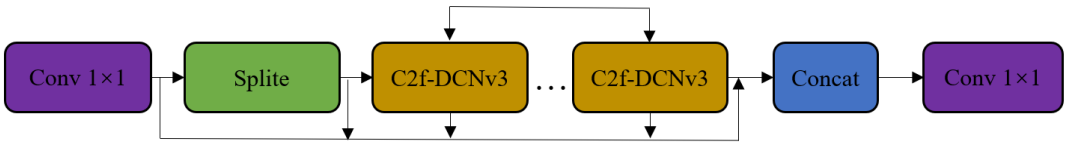

Figure 2: Structure of Backbone- DCNv3

2.3.2. MPDIoU loss function

Bounding box regression is a key aspect in target detection and instance segmentation, but most existing bounding box regression loss functions are difficult to optimize when the predicted box has the same aspect ratio as the real box. To solve this problem, Siliang et al. proposed the MPDIoU bounding box regression loss function(shown in Figure 3) [8].MPDIoU utilizes the minimum point distance to directly predict the distances between the upper-left and lower-right corners of the bounding box and the real box, thus improving the efficiency and accuracy of bounding box regression. Compared to the traditional IoU loss function, MPDIoU not only inherits the advantages in overlapping regions, centroid localization, width and height differences, but also simplifies the process of comparing the similarity between the predicted and real boxes. This computational method simplifies the process, speeds up the convergence of the model, and improves the accuracy of the regression. Therefore, in this paper, the MPDIoU loss function is introduced into the YOLOv8 model to speed up the convergence of the model and improve the detection performance of the model to enhance the application value of the model [9].

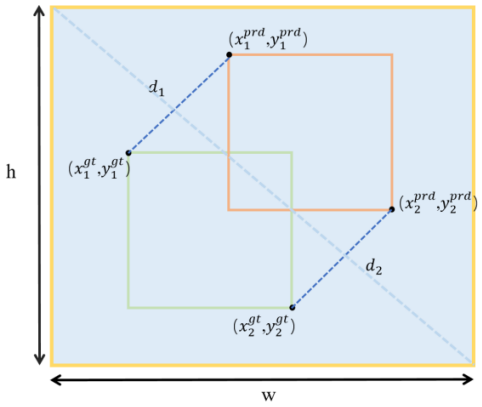

Figure 3: Schematic diagram of MPDIoU loss function

The MPDIoU is calculated as follows:

$$d_1^2 = (x_1^B - x_1^A)^2 + (y_1^B - y_1^A)^2 \tag{2}$$

$$d_2^2 = (x_2^B - x_2^A)^2 + (y_2^B - y_2^A)^2 \tag{3}$$

$$MPDIoU = IoU - \frac{d_1^2}{w^2 + h^2} - \frac{d_2^2}{w^2 + h^2} \tag{4}$$

3. Results and discussion

3.1. Test environment

The environment configuration for this model training trial were as follows: Python 3.8.16 was used as the programming language, PyTorch 2.0.0 as the deep learning framework, and software environments such as Torch 2.0.0 and CUDA 11.8.0 were equipped. The operating system was Windows 11, the central processor was Intel Core I7-13900HX CPU, the graphics gas pedal was NVIDIA GeForce RTX3060, and the RAM was 16 GB.The experimental environment provided a stable and efficient platform for the training of the model, which ensured the smooth running of the experiments and the accuracy of the

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

results [10].

3.2. Evaluative indicators

Several evaluative metrics are used in this study to comprehensively assess the performance of the trained models. These metrics include precision, recall, and average precision, which provide comprehensive information for evaluating different aspects of the model.

$$Precision = \frac{T_P}{T_P+F_P} \times 100\% \tag{5}$$

$$Recall = \frac{T_P}{T_P+F_N} \times 100\% \tag{6}$$

$$AP = \int_0^1 P(R)\, dR \tag{7}$$

$$mAP = \frac{1}{K}\sum_{i=1}^{K} AP(i) \tag{8}$$

Where TP (True Positive) was the number of samples that were correctly detected as positive categories, FP (False Positive) was the number of samples that were incorrectly detected as positive categories. FN (False Negative) was the number of positive samples that were incorrectly determined as negative categories. Mean Precision was usually calculated using the area under the Precision-Recall curve, usually using different interpolation methods for PR curves.N was the total number of categories and AP_i was the mean precision of the ith category.

3.3. Evaluative indicators

3.3.1. MPDIoU loss function

In order to verify the effect of adopting Backbone- DCNv3 module on the performance of the detection model, a comparative analysis of adopting Backbone- DCNv3 module and the original model was carried out under the same test conditions, and the results are shown in Table 1. As shown in Table 1, the improved model has significantly improved various detection performance indicators, with 91.24% for P, 89.53% for R and 91.67% for mAP using the Backbone- DCNv3 module. Compared with the original model, they were improved by 1.92%, 1.08%, and 1.1%, respectively. This indicates the effectiveness of using this module.

Table 1: Comparative analysis using Backbone- DCNv3 module

Norm	Original model	Adoption of Backbone-DCNv3 module
P /%	89.32	91.24
R/ %	87.45	88.53
mAP/ %	89.27	90.37

3.3.2. Analysis of MPDIoU loss function results

The MPDIoU loss function was introduced into the YOLOv8 model, for testing the impact of referencing the MPDIoU loss function on the model performance, this subsection conducts a comparison test between the original model and the model with Backbone- DCNv3 module with the introduction of CIoU and MPDIoU, respectively, and the results are shown in Table 2. As observed in Table 2, both the original model and the model with Backbone- DCNv3 module have improved detection performance metrics after the introduction of MPDIoU loss function.

Table 2: Comparative analysis of MPDIoU loss function

Model	CIoU	MPDIoU	P /%	R/ %	mAP/ %
Original model	√	×	89.32	87.45	89.27
Original model	×	√	90.26	88.32	91.25
Adoption of Backbone-DCNv3 module	√	×	91.24	88.53	90.37
Adoption of Backbone-DCNv3 module	×	√	91.45	89.47	91.89

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

After the introduction of the MPDIoU loss function, the P of the original model improves from 89.32% to 90.26%, the R from 87.45% to 88.32%, and also the mAP from 89.27% to 91.25%. The model with Backbone-DCNv3 module has improved its precision rate, recall rate and mean average precision mean after the introduction of MPDIoU loss function. The precision rate was improved from 91.24% to 91.45%, the recall rate was improved from 88.53% to 89.47%, and the mean average precision was improved from 90.37% to 91.89%. This further demonstrates the effectiveness of the MPDIoU loss function in improving model performance, and that this improvement works across different model architectures.

3.3.3. Analysis and discussion of different model comparison tests

In order to deeply explore the performance differences of different target detection models on specific tasks, a series of representative YOLO series models (YOLO v5, YOLO v6, YOLO v7, YOLO v8) are selected for comparative experiments with the models proposed in this paper. By comparing the performance of these models on key evaluation metrics such as precision, recall and average precision, we aim to comprehensively assess the performance advantages and disadvantages of each model. The results of which as shown in Table 3. As can be seen from Table 3, the YOLO series of models show a steady increase in performance with the iteration of versions. The model in this paper has achieved a significant improvement in the accuracy rate, reaching 91.45%, which is higher than the 89.32% of the YOLO v8 model. This result indicates that the model in this paper has higher accuracy in recognizing targets and can better distinguish between targets and backgrounds. Meanwhile, this paper's model also performs well in terms of recall, reaching 89.47%, which is better than the 87.45% of the YOLO v8 model. As an important index to measure the overall performance of the model, the model in this paper also shows significant advantages. Its average accuracy reaches 91.89%, which is significantly improved compared to the 89.27% of the YOLO v8 model.

Table 3: Different model test results

Model	P/%	R/%	mAP/%
YOLO V5	87.56	85.49	87.98
YOLO V6	87.58	86.47	86.57
YOLO V7	88.58	89.74	88.42
YOLO v8	89.32	87.45	89.27
Model in this paper	91.45	89.47	91.89

3.3.4. Visual analytics discussion

In order to visualize the impact of the improved strategy in this paper on the performance of the YOLO v8 model, the image detection results of the model before and after the improvement were compared and analyzed on a test set. Figure 4(a) demonstrates the target detection test plot of the original YOLO v8 model, while Figure 4(b) presents the target detection results of the improved YOLO v8 model in this paper. Comparing these two figures, the significant improvement of the improved YOLO v8 model in insulator defect location identification can be clearly observed. The improved model not only can locate the position of defects more accurately, but also significantly improves the recognition accuracy. It was especially worth mentioning that the improved model was able to effectively detect even those insulator defects that were small and difficult to recognize.

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

Figure 4: Comparison of models before and after improvement

4. Conclusions

In order to be able to intelligently detect insulator defects, this study improves the YOLO V8 model by adopting the Backbone- DCNv3 module for feature extraction and introduces the MPDIoU loss function to enhance the effectiveness of the model in terms of feature extraction capability and target localization accuracy. The experimental results show that the improved detection model has a P of 91.45%, an R of 89.47%, and a mAP of 91.89%. The model in this paper outperforms YOLO v8 and other earlier versions of the YOLO model in key evaluation metrics such as precision, recall and average precision. This proves the effectiveness and superiority of this paper's model in the task of target detection, and provides new ideas and methods for research and application in related fields.

Acknowledgements

We would like to acknowledge the financial support from Tai'an Science and Technol-ogy Innovation Development Project (Policy guidance) under Grant 2022GX074.

References

[1] Tang Lu, Wang Shuqing, Wang Niantao, et al. Defect detection of insulators in foggy weather based on improved YOLOX network [J]. High Voltage Electrical Apparatus, 2024, 60(03):223-228.
[2] Wen Bin, Hu Yiming, Peng Shun, et al. A lightweight network for multi-defect detection of transmission line insulators [J/OL]. Radio Engineering, 2024, 55(05): 1-9.
[3] Zhang Changle, Jin Jun. Simulation study of insulator fault detection based on YOLOv5s[J]. Electronic Design Engineering, 2024, 32(05):178-182.
[4] Pan Wei, Wei Chao, Qian Chunyu, et al. An improved model of YOLOv8s for small target detection in UAV view [J/OL]. Computer Engineering and Applications, 2024,1-10.
[5] Luo Lei, Xie Zhukui. Traffic sign detection algorithm based on improved YOLOv8[J]. Electromechanical Engineering Technology, 2024, 53(03):205-210.
[6] Li Z, Zhu Y, Sui S, et al. Real-time detection and counting of wheat ears based on improved YOLOv7[J]. Computers and Electronics in Agriculture, 2024, 218108670.
[7] Xiaoqiang D, Hongchao C, Zenghong M, et al. DSW-YOLO: A detection method for ground-planted strawberry fruits under different occlusion levels[J]. Computers and Electronics in Agriculture, 2023, 214.
[8] Yang Z, Shen Y, Shen Y . Football referee gesture recognition algorithm based on YOLOv8s[J]. Frontiers in Computational Neuroscience, 2024, 18: 1341234.

ISSN 2616-5775 Vol. 7, Issue 5: 37-43, DOI: 10.25236/AJCIS.2024.070505

[9] He J, Zhang S, Yang C, et al. Pest recognition in microstates state: an improvement of YOLOv7 based on Spatial and Channel Reconstruction Convolution for feature redundancy and vision transformer with Bi-Level Routing Attention. [J]. Frontiers in plant science, 2024, 15: 1327237.
[10] Wang Guofa, Pang Yihui, Ren Huaiwei, et al. Research and practice of intelligent mine system engineering and key technology [J]. Coal Journal, 2024, 49(01):181-202.