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

doi:10.3233/FAIA231074

Discolored Transformer Breather Recognition for Substation Based on Improved YOLOv8

Qiang HE[a,1], Xin YOU[a], Zhiyuan QIU[a], Wenxin FENG[a], Daoyu LI[a], Xianzhong CHEN[a] and Zhaoye ZHENG[b]
[a] *China Southern Power Grid Co., Ltd. Extra High Voltage Transmission Company Guiyang Bureau*
[b] *North China Electric Power University*

Abstract. Transformer breather is an important equipment to ensure the normal operation of transformer in substations. In order to identify the discoloration defects of breathers accurately and reduce the operation and maintenance costs of substations, this paper proposes a recognition method of discolored transformer breathers based on improved YOLOv8. Firstly, to solve the problem of sample imbalance in original image data, this paper augments the data by cropping and generating, and trained a more robust model with "real + virtual" data. Secondly, considering the various scales of breather, the BiFPN structure is incorporated to achieve more efficient multi-level feature fusion. Finally, since breathers exist in complex substation scenes, this paper introduces CBAM attention mechanism to enhance the feature extraction ability of the model. The experimental results show that the proposed method can improve the mAP of discolored transformer breather detection by 2.1% compared to the baseline and inference in real-time, while hardly increasing the number of parameters and computation cost, which verifies the effectiveness of the proposed method.

Keywords. substation transformer breather, YOLOv8, data augmentation, BiFPN, attention mechanism

1. Introduction

Power transformer is an important equipment in substation, and its normal operation provides the fundamental guarantee for the safe production of substation. The breather is an important part of the transformer used to isolate the moisture in the air, which can avoid the insulation strength of transformer oil decline due to moisture [1]. Therefore, recognizing the discoloration of breathers is one of the key steps of the transformer defect detection. The traditional substation generally adopts the manual inspection mode to find the defect of transformer breathers, which consumes a lot of human resources and has low efficiency. With the popularity of inspection robots in recent years, the operation and maintenance of substations have gradually shifted to an unattended mode [2]. It has gradually become the mainstream method to capture transformer images through

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

inspection robots equipped with cameras and detect the status of different components by Computer Vision (CV) technology [3].

However, the characteristics of the substation scene and the limitations brought by inspection robots pose new challenges to the design of CV models for substation inspection. Compared with models in the general field, a model used for substation inspection should have following characteristics: 1) Subject to deployment conditions, it needs to have as small a number of parameters and computation cost as possible; 2) The key features should be effectively learned from the limited substation images which often accompanied with the problem of sample imbalance; 3) It can achieve higher detection accuracy as well as real-time detection under the complex scene of substation. Although some achievements have been made in the research of discolored breather recognition, these methods based on detectors with a large number of parameters and a high computation cost, and cannot achieve the optimal speed-precision tradeoff.

To solve the above problems, we propose a discoloration recognition method for transformer breathers based on improved YOLOv8. The main contributions can be summarized in three folds:

- To solve the problem of sample imbalance in original data, we adopt the method of "cropping + generation" to enhance the training data, and uses "real + virtual" data to train the model, so that the model has stronger generalization ability.
- In terms of the breathers photographed with various scales, the structure of Bi-directional Feature Pyramid Network (BiFPN) is integrated into the detector's neck network for more efficient feature fusion.
- In view of the characteristics of breathers existing in complex substation scenes, we introduce Convolutional Block Attention Module (CBAM) to improve the model's ability to extract key features of breathers in complex backgrounds.

2. Related Work

Existing object detection methods based on deep learning can be divided into two types: two-stage methods and one-stage methods. The representative works of the two-stage methods are the Region with CNN features (R-CNN) series [4-5], which firstly obtain the region proposals with the region proposal network then classify these proposals. One-stage methods are represented by the You Only Look Once (YOLO) series [6-7], which treat the object detection task as a whole regression task, using an end-to-end network to directly predict the location and class of the object in the image.

On the other hand, in the scene of power system, the detection of breathers and their defects based on above object detection methods has been concerned by researchers for a long time. Reference [3] proposed an improved RetinaNet and achieve a higher the detection accuracy for components in transformer including breathers. Reference [8] combined a feature matching algorithm with SSD network to realize rapid and stable recognition of the defects in transformer breathers. Reference [9] used a modified YOLOv4 network for the defect detection of breathers and other equipment in substation. Although these methods achieved relatively good results in breather detection, they all rely on earlier object detectors with a large number of parameters and a large amount of calculation, neglecting the deployment condition of inspection robots in which the computing power and storage space are limited, while cannot achieve the optimal balance between speed and precision as well.

3. Proposed Method

According to the deployment condition of substation inspection robots, we choose YOLOv8n as the baseline model, and improve the training strategy as well as the model structure on the basis of the characteristics of transformer breather images. The improved network structure is shown in Figure 1.

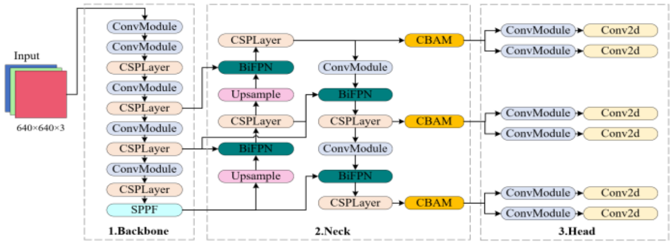

Figure 1. Structure of the improved YOLOv8n model.

3.1. Base Architecture

Since we take YOLOv8n as base architecture, we describe its operating principle, characteristic and the reason for choosing it briefly. YOLOv8 is developed to achieve a better trade-off between speed and precision based on YOLOv5. YOLOv8 includes five models of different sizes, all of them consist of backbone, neck and head, which used to extract image features, fuse multi-scale features and detect object of different scales, respectively. Compared with YOLOv5, it mainly improves in the following aspects: 1) The model works in an anchor-free way; 2) More elaborate image feature extraction modules with more residual connections are used; 3) The head adopts the decoupled design. 4) More advanced loss function and sample assigning strategy are used in training. With these improvements, YOLOv8 can achieve advanced detection accuracy and inference speed performance with lightweight models.

Compare to the existing discolored breather detection methods which based on earlier object detector, YOLOv8 is a more powerful baseline with a better balance between speed and precision. Meanwhile, considering the accessibility of edge deployment in practical applications which is not fully taken into account in previous work, we choose the lightest YOLOv8n model among the five sizes as the baseline, the original structure of it is shown in Figure 2.

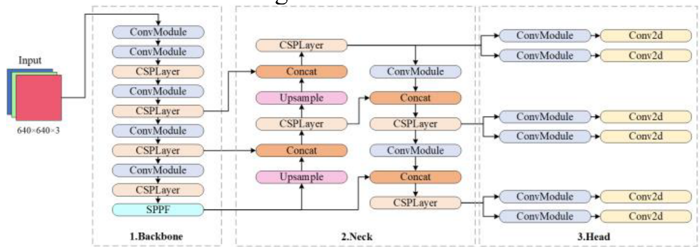

Figure 2. Structure of the original YOLOv8n model.

3.2. Data Augmentation with Cropping and Generating

According to the observation of the original images of transformer breathers, we find that the number of breathers without iron shell is relatively small and the background is relatively single as well. This problem of data imbalance may not be conducive to the model's effective recognition of the breathers without iron shell and their discoloration defects. In this paper, some images of breathers without iron shell are cropped at random scale. Then, a variety of different scale images of breathers without iron shell are generated by using FastGAN [10] network with the original and cropped images in a few-shot way. Finally, the cropped images and generated images are used to enhance the original training data, these "real + virtual" images are used to train a more robust discolored breather detection model, as shown in Figure 3.

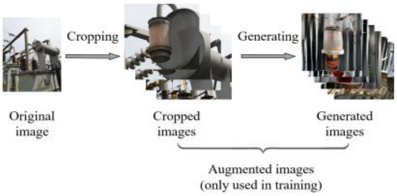

Figure 3. Process of data augmentation.

3.3. Introducing BiFPN Structure

Since the breathers have many different sizes in the image, sufficient multi-scale feature fusion plays an important role in improving the detection accuracy. We introduce BiFPN [11] structure to replace the original Path Aggregation Network (PANet), using its more efficient multi-scale feature fusion design to obtain fused features with stronger representation ability and improve the performance of breather discoloration detection. The comparation between PANet and BiFPN is shown in Figure 4.

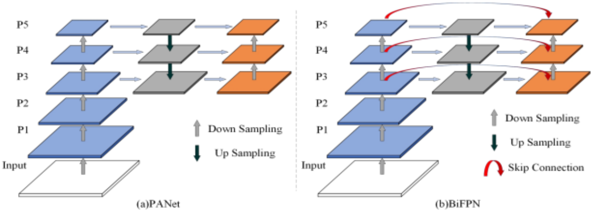

Figure 4. Comparation between PANet and BiFPN.

3.4. Introducing CBAM Attention Mechanism

Transformer breathers are often existing in the complex substation background, so it is of great significance for the model to accurately extract the relevant features of breathers from the complex substation scenes to improve the detection accuracy. We introduce the CBAM [12] attention mechanism at the output end of neck network to enhance the model's ability for extracting effective features from breathers, the structure of CBAM is shown in Figure 5. By calculating the attention scores from both spatial and channel dimensions, and then reweighting the convolutional feature map, the CBAM attention mechanism enables the model to focus more effectively on the key features related to breathers and their discoloration defects.

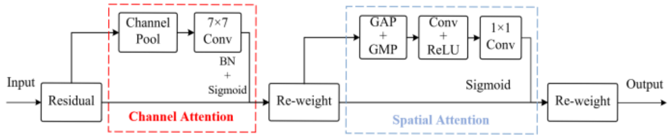

Figure 5. Comparation between PANet and BiFPN.

4. Experimental Evaluation

4.1. Experiment Settings

All of our experiments are implemented with an NVIDIA GeForce GTX 1080 Ti GPU and PyTorch 1.10.1 framework. We train and evaluate models with the discolored breather detection dataset containing 2537 original substation transformer breather images, in which the training set and validation set is divided in a 9:1 ratio. We set the batch size to 8 and the number of iterations to 50 while keep the default settings of the remaining hyperparameters for fair comparison. In order to effectively evaluate the comprehensive performance of different models in the substation scenario, we use Parameters (Params), FLOating Point operations (FLOPs), and Frames Per Second (FPS) to evaluate the size, computational complexity, as well as inference speed of models, respectively. And we use Average Precision (AP) for different categories and mean Average Precision (mAP) to evaluate the detection accuracy of models.

4.2. Comparison Experiment

We select four object detection models widely used in engineering and the baseline model YOLOv8n for comparison with our method. The experimental results are shown in Table 1. Firstly, in terms of deployment, SSD, TOOD and two RCNN series models have a larger number of parameters and a higher computation cost while our method only has 3.1M Params and 8.2G FLOPs, which means our method is more feasible for deployment in the inspection robot. Secondly, in terms of real-time detection, the YOLO models can significantly exceeds the threshold of real-time detection (i.e., higher than 30 FPS), reaching over 100 FPS, while other detectors have much lower inference speed which slightly over or under 30 FPS. In this case, we make a slight sacrifice of speed for higher precision. Finally, in terms of detection accuracy, the proposed method can achieve 97.2% mAP, which is 2.1% higher than the baseline model and significantly higher than other reference models. Therefore, our method has better performance on the task of discolored breather detection while maintaining a high inference speed for real-time detection and a lightweight model for edge deployment.

Table 1. Performance comparison of different models

Methods	Params(M)	FLOPs(G)	FPS(frame/s)	mAP(%)
Faster RCNN [6]	41.1	206.7	14.8	95.2
Sparse RCNN [14]	106.0	149.9	14.3	95.7
SSD [13]	24.3	344.9	51.3	92.6
TOOD [15]	31.8	180.7	14.7	96.1
YOLOv3 [8]	12.2	19.0	78.7	91.8
YOLOv5	3.1	9.0	117.4	94.3
YOLOv8	3.0	8.2	109.1	95.1
Our method	3.0	8.2	102.6	97.2

4.3. Ablation Experiment

In order to verify the effectiveness of each improvement scheme in the proposed method and its contribution to the ultimate model, we implement the ablation experiment, the results are shown in Table 2.

Table 2. Results of ablation experiment

Data Augment	BiFPN	CBAM	Params(M)	FLOPs(G)	FPS(frame/s)	mAP(%)
-	-	-	3.0	8.2	109.1	95.1
√	-	-	3.0	8.2	109.7	95.7
-	√	-	3.0	8.2	108.3	96.0
-	-	√	3.0	8.2	103.4	95.8
√	√	√	3.0	8.2	102.6	97.2

Firstly, with the data augmentation of cropping and generating, we obtain to a new training set with a more balanced sample distribution, the instance number and APs for different categories are shown in Table 3. The APs for all kinds of breather and their discoloration problem have been improved after data augmentation, and the metric of mAP is improved by 0.6% without additional model parameters and computation cost. The result verifies that the model is more robust to different types of breathers.

Secondly, by integrating BiFPN structure into the neck network of the model, it can effectively improve the perception ability of the model to the targets of different scales, and increase the mAP of the model by 0.9%, while having little impact on other metrics of the model.

Table 3. Result of different training sets

Methods	Before augmentation		After augmentation	
	Number	AP(%)	Number	AP(%)
Normal breather with shell	875	95.6	875	95.7
Normal breather without shell	193	93.4	867	94.9
Discolor breather with shell	824	97.1	824	97.3
Discolor breather without shell	391	94.2	830	95.1

Thirdly, by introducing CBAM attention mechanism at the output end of the neck network, the ability of extracting key features related to breathers is effectively enhanced, and the mAP of the model is increased by 0.7% compared with the baseline as well, while the number of parameters and computational complexity of the model are almost unchanged. Although the inference speed is slightly decreased in exchange for higher mAP, the cost is acceptable because the inference speed has already significantly exceeded 30 FPS which is enough for real-time detection.

Finally, when all three improvement schemes are used together, the mAP of the proposed method improved by 2.1% compared to the one of the baseline. The mAP curves of training process are shown in Figure 6. At the beginning of training, the mAP of the original YOLOv8n model is higher, which is caused by the modified structure and introduced new parameters in the improved model. At the later stage of training, the mAP of the improved model gradually exceeds that of the baseline, and the original model begins to overfit while the improved model does not. The above results verify that the improved model is more accurate and robust than the baseline model, while maintaining a relatively high inference speed for detecting in real-time and hardly increasing the parameters and computation cost.

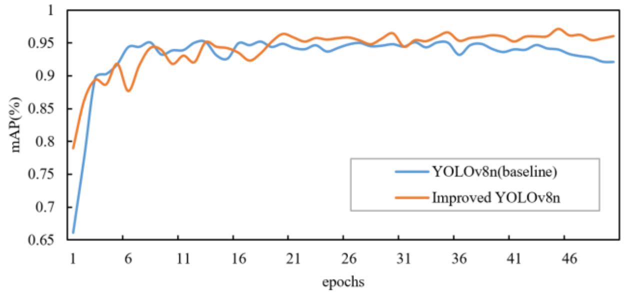

Figure 6. mAP curves of original and improved models.

4.4. Qualitative Experiment

The above experiments have verified the effectiveness of the proposed method and the contribution of each improvement scheme on the discolored breather detection task. In order to observe the effects brought about by above improvements more intuitively, we visualize the detection results, as shown in Figure 7. In all images for comparison, the improved YOLOv8 model can pay more attention to the effective features of breathers to effectively avoid the false detection and improve the confidence score of the detected object. This result further verifies the effectiveness of our proposed method.

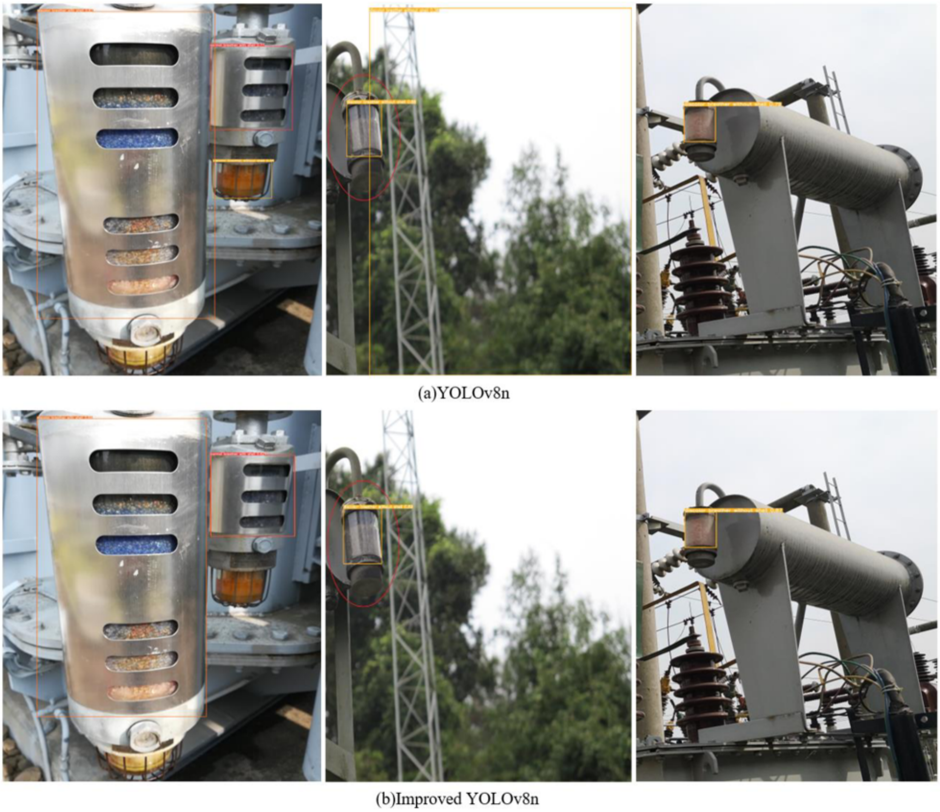

Figure 7. detection results of original and improved models.

5. Conclusion

This paper proposes a discolored breather recognition method based on improved YOLOv8 by analyzing the characteristics of application scenarios, deployment conditions and original data. Firstly, the problem of sample imbalance is solved by a novel data augmentation strategy of cropping and generating. Secondly, a more powerful multi-scale feature fusion is enabled by incorporating the structure of BiFPN. Finally, CBAM attention mechanism is used to improve the ability of extracting effective features. The experimental results show that our method can realize the discolored breather detection with a lightweight model and a low computation cost which is suitable for deployment on inspection robots while achieving a better speed-precision tradeoff.

References

[1] Chen L, Zhu Y, Wang K, et al. Study of intelligent vision detection device for transformer respirator. Modern Manufacturing Technology and Equipment, 2023,59(05): 48-50.
[2] Chu G, Zhu W, Zhang G, et al. Design of intelligent unmanned inspection system for transformer substation. Proceedings of the 16th National Conference on Signal and Intelligent Information Processing and Application. 2022: 455-459.
[3] Wan J, Wu G, Guan M, et al. Intelligent recognition method for transformer small components based on RetinaNet. Power System Protection and Control, 2021,49(12): 166-173.
[4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
[5] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 2015, 28.
[6] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
[7] Redmon J, Farhadi A. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.
[8] Xiang X, Pan L, Wu H. Respirator defect detection method based on feature matching and deep learning. Research and Exploration in laboratory, 2022,41(06): 57-61+97.
[9] Chen T, Zhou M, Han Q, et al. Defect Detection for Substation Based on Improved YOLOv4. Computer Systems & Applications, 2022,31(06):245-251.
[10] Liu B, Zhu Y, Song K, et al. Towards faster and stabilized gan training for high-fidelity few-shot image synthesis. International Conference on Learning Representations. 2020.
[11] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10781-10790.
[12] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module. Proceedings of the European conference on computer vision. 2018: 3-19.
[13] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector. Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.
[14] Sun P, Zhang R, Jiang Y, et al. Sparse r-cnn: End-to-end object detection with learnable proposals. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 14454-14463.
[15] Feng C, Zhong Y, Gao Y, et al. Tood: Task-aligned one-stage object detection. 2021 IEEE/CVF International Conference on Computer Vision. IEEE Computer Society, 2021: 3490-3499.