



Enhanced Defect Detection in Insulator Iron Caps Using Improved YOLOv8n

Qiming Zhang¹, Ying Liu^{1*}, Song Tang³, Kui Kang²

¹ Sichuan Global Insulators Co., Ltd, 610100 Chengdu, China

² School of Mechanical Engineering, Xihua University, 610039 Chengdu, China

* Correspondence: Ying Liu (1220170066@mail.xhu.edu.cn)

Received: 06-24-2024

Revised: 08-01-2024

Accepted: 08-29-2024

Citation: Q. M. Zhang, Y. Liu, S. Tang, and K. Kang, “Enhanced defect detection in insulator iron caps using improved YOLOv8n,” *Inf. Dyn. Appl.*, vol. 3, no. 3, pp. 162–170, 2024. <https://doi.org/10.56578/ida030302>.



© 2024 by the author(s). Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

Abstract: To address the challenges in detecting surface defects on insulator iron caps, particularly due to the complex backgrounds that hinder accurate identification, an improved defect detection algorithm based on YOLOv8n, whose full name is You Only Look Once version 8 nano, was proposed. The C2f convolutional layers in both the backbone and neck networks were replaced by the C2f-Spatial and Channel Reconstruction Convolution (SCConv) convolutional network, which strengthens the model’s capacity to extract detailed surface defect features. Additionally, a Convolutional Block Attention Module (CBAM) was incorporated after the Spatial Pyramid Pooling - Fast (SPPF) layer, enhancing the extraction of deep feature information. Furthermore, the original feature fusion method in YOLOv8n was replaced with a Bidirectional Feature Pyramid Network (BiFPN), significantly improving the detection accuracy. Extensive experiments conducted on a self-constructed dataset demonstrated the effectiveness of this approach, with improvements of 2.7% and 2.9% in mAP@0.5 and mAP@0.95, respectively. The results confirm that the proposed algorithm exhibits strong robustness and superior performance in detecting insulator iron cap defects under varied conditions.

Keywords: Insulator iron cap; YOLOv8n; Defect detection; Spatial and channel convolution; Attention mechanism; Bidirectional Feature Pyramid Network

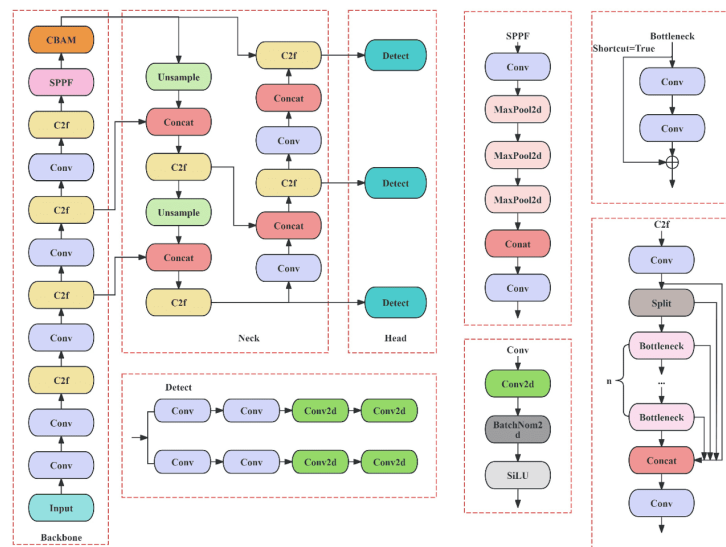
1 Introduction

In recent years, with the continuous development of the power system, the scale of China’s power grid has been gradually expanding, covering more than 95% of the country’s territory. China’s power grid system has the world’s largest voltage span and the longest transmission lines by using a variety of power generation and transmission methods. Insulators in the power grid system are one of the important external insulation equipment in transmission and distribution lines. Their main role includes supporting and fixing the current-carrying conductor, the current-carrying conductor and the ground to form good insulation between. Common types of insulators include porcelain insulators, glass insulators and composite insulators. Porcelain and glass insulators are composed of insulating parts and fittings, and the fittings include the steel caps and feet of the insulators. In actual operation, the insulator is easy to be affected by natural conditions and environmental changes because it is exposed to the complex environment for a long time. Insulators corrode from time to time, resulting in the decline of their electrical and mechanical properties [1]. In wet conditions, leakage current flows through the insulator’s steel cap surface, which constitutes electrolytic corrosion of the chemical primary battery. In this reaction, the steel cap loses electrons to the negative reaction, thus losing the metal. The mechanical bearing capacity of corroded insulator steel caps also reduces, which in serious cases causes broken insulator strings and other consequences, resulting in great economic losses.

Insulators are usually continuously exposed to the natural environment. The cross-influence of a variety of harsh weather and climate conditions leads to corrosion, breakage, and deformation of the insulator iron caps and other failures, which is a serious threat to the reliable operation of electrical equipment. In order to increase the service life of the insulator iron caps, the surface is usually plated with a layer of zinc to enhance their corrosion resistance. However, due to the lack of zinc in the production and transport process, damage and other defects inevitably occur, which reduces the service life of the iron cap, thereby increasing the likelihood of the broken string. Therefore, in

order to reduce the probability of this situation, it is necessary to carry out defect detection on the insulator iron cap before it leaves the factory, aiming to efficiently and accurately detect the defective target state of the iron cap.

2 The YOLOv8n Algorithm



3 YOLOv8n Algorithm Improvement

The complexity of texture features and the amount of interference information in the defective images of insulator surfaces lead to a large amount of redundant information in the extracted feature maps in the spatial and channel dimensions, thus weakening the adaptability and generalisation power of Convolutional Neural Networks (CNNs). In addition, the C2f structure in the original YOLOv8n network is more complex, resulting in a larger model computation and slower detection speed, which does not facilitate the deployment of embedded devices in industrial automation application scenarios. SCConv [4] is a lightweight convolution module proposed in 2023 to compress redundant features in CNNs. The structure of SCConv is shown in Figure 2. The SCConv module consists of a Spatial Reconstruction Unit (SRU) and a Channel Reconstruction Unit (CRU). When dealing with the intermediate feature X in the bottleneck residual block, the spatial details are firstly optimised by the SRU unit, and then the CRU unit further refines the channel features to form the feature Y . After the SCConv module, the channel features are further refined to form the feature Y . Then the 1×1 convolution of the SCConv module reduces the number of channels to reduce the

computation amount. Finally, the convolution becomes the corresponding number of channels to calculate the loss. This method effectively reduces the redundancy of space and channels in the CNN, promotes the learning of critical features, and shares the original two-branch structure parameterised to form a smaller and lighter head structure [5].

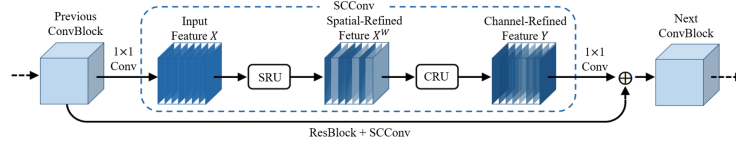


Figure 2. SCConv module

In this study, SCConv was adopted to replace the bottleneck structure in the original C2f module, which can effectively reduce the feature redundancy in the spatial and channel dimensions as well as the complexity and computational cost of the model, thereby realising the lightweight of the model. The improved C2f module is named C2f-SCConv module, and its structure is shown in Figure 3. Using the C2f-SCConv module to replace the original C2f structure in the neck network can improve the detection performance of the model while reducing resource consumption.

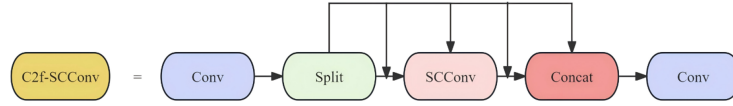


Figure 3. C2f-SCConv structure

3.2 CBAM

The deepening of the network layers helps the model to extract the features of defects, but the model also loses the detail information during the downsampling operation, which is not conducive to the target defect detection of insulator iron caps [6]. Therefore, in order to improve the model's ability to capture key information in various types of defects and reduce the interference of irrelevant information, this study introduces the CBAM [7] after the SPPF layer, which integrates the channel attention mechanism with the spatial attention mechanism. As shown in Figure 4, the feature map is first input into the channel attention module, and the corresponding attention map is output. Therefore, the channel attention mechanism can be used to highlight the channels related to the defective features of the insulator iron cap and suppress the irrelevant channels, thereby enhancing the ability of the model to extract the defective features of the insulator iron cap in terms of channels [8]. Then the input feature map is multiplied with the attention map, and the output goes through the spatial attention module. The result obtained after the spatial attention multiplication is multiplied with the original result to obtain the output result. The spatial attention mechanism highlights image regions of interest and reduces the influence of interfering regions by weighting the features at each spatial location of the surface defects of iron caps [9]. This enables the model to improve the attention of the YOLOv8n-SCConv model to the features of the insulator iron cap defects from both channel and spatial aspects, which better distinguishes the background from the target and thus improves the accuracy of the detection results.

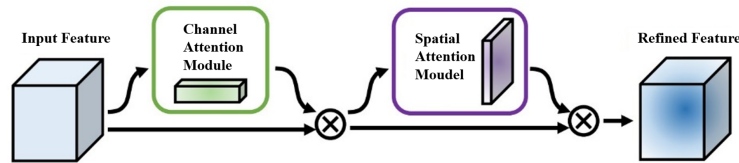


Figure 4. CBAM module

3.3 BiFPN

Path Aggregation Network (PANet) [10] is used in YOLOv8n. PANet [11] has both top-down and bottom-up paths, which makes it fuse top and bottom feature information. Although this method can improve the detection accuracy, it is easy to miss tiny target objects [12]. To solve this problem, this study replaces PANet with BiFPN [13]. The structure of PANet and BiFPN is shown in subgraphs (a) and (b) of Figure 5. PANet introduces a reverse

path on the basis of Feature Pyramid Network (FPN) [14] to convey the missing position information, as shown in subgraph (a) of Figure 5. BiFPN is a weighted BiFPN, as shown in subgraph (b) of Figure 5. It simplifies the network structure based on PANet by removing a single input node to ensure that no important information is lost. Meanwhile, additional features are integrated by adding extra edges when the input and output nodes are located in the same layer. Specifically, BiFPN designs a top-down path from p7 to p3 that passes semantic information from higher-level features to the bottom layer, and a bottom-up path from p3 to p7 that passes new information about the location of the bottom layer to higher levels. In addition, BiFPN adds a connection from p4 to p6 that directly connects the input and output nodes of the same layer across intermediate layers for deeper feature fusion. As the model is trained deeper, this dynamic weight adjustment mechanism allows BiFPN to gradually optimise its feature fusion strategy so that it can more accurately mine and integrate useful information from different layers. The learned features are then used to continuously update the weights to obtain more valuable information [15].

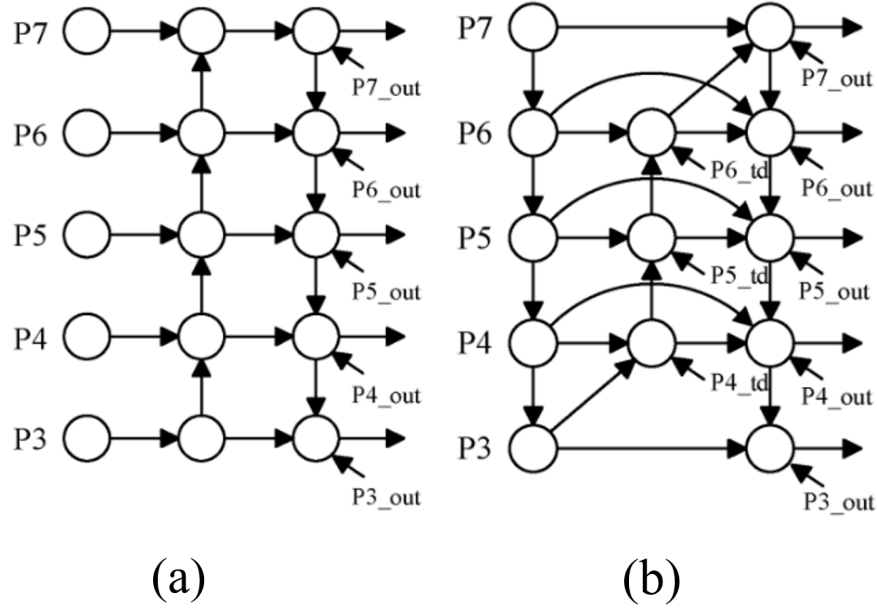


Figure 5. Model structure of PANet and BiFPN: (a) PANet; (b) BiFPN

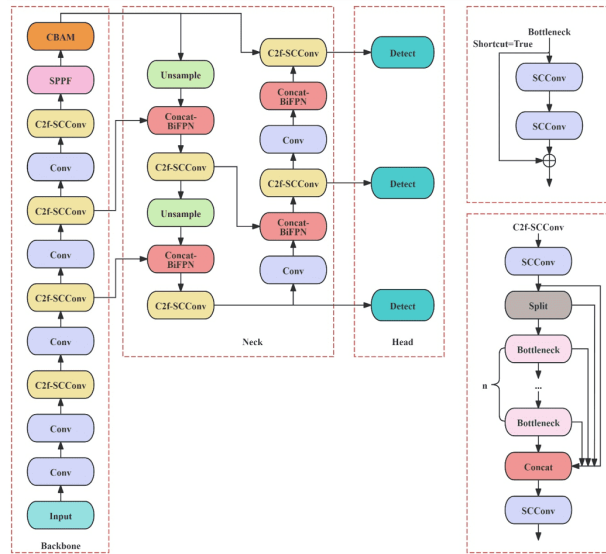


Figure 6. Improved YOLOv8-SCPN network structure

3.4 Improved YOLOv8n-SCPN Network Structure

YOLOv8n was improved using the above method to obtain a model, whose structure is shown in Figure 6. In this study, SCConv was used to replace the bottleneck structure in the original C2f module to form the C2f-SCConv structure, which enables the backbone network to use a larger sensory field and obtain more semantic information, effectively reducing the feature redundancy in the spatial and channel dimensions, the complexity of the model and the cost of computation, thereby realising the model's lightweight. Secondly, the CBAM was introduced between the backbone network and the neck network to improve the network's attention to the defective target during feature extraction, thus achieving the purpose of optimising the network performance. Finally, the PANet in the neck was replaced with BiFPN, and the introduction of BiFPN can exchange a small amount of computational cost for a larger model performance gain.

4 Experimental Validation and Analysis of Results

4.1 Pre-Processing of the Dataset

In this study, the dataset uses the industrial greyscale camera A3A20MG8, lens M2016-12MP-2 of Huarui Technology to take images of the insulator iron cap defects for experiments. However, due to the irregular surfaces of the collected insulator iron caps, the images usually have background interference or their features are not obvious enough. Therefore, the image quality can be improved by the pre-processing method.

Image enhancement mainly aims to improve the recognition of image feature information in the background and interference elements. Filtering the casting image can suppress noise, but the image regions have a certain blurring and defects in the image colour and geometric information in the image of the contrast decrease. In order to reduce the impact of filtering and highlight the texture feature information to improve the subsequent algorithms for regional processing and the stability and accuracy of feature recognition, there is a need for image enhancement processing, aiming to improve defective features and image quality through pre-processing methods. In order to reduce the effect of filtering and highlight the texture feature information to improve the stability and accuracy of the subsequent algorithm and feature recognition, it is necessary to enhance the image to improve the recognition of defect feature information. In this study, the gamma transform was used to enhance the image of insulator iron cap defects.

Because the gamma transform has the advantages of improving the visual effect of the image [16], correcting the image illumination, and having simple operations, it is widely used in the enhancement of low-quality images. According to the different values of the input parameter, it has a homogeneous effect on the global image value, and has a good effect on the image brightness and overexposed area suppression. The principle of its action is as follows: different pixel values of the image are homogeneous according to the value of the gamma value, but the difference of the transformation changes the contrast of the image feature information. The gamma transform formula is as follows:

$$s = cr^{\gamma} \quad (1)$$

where, c is the magnification, generally taken as 1; γ is the key parameter of the gamma transform, which determines the direction of the image brightness; r is the input pixel value; and s is the output pixel value. In order to observe the effect of different γ on the function output, the function output curve at different γ values was plotted, as shown in Figure 7, where the horizontal coordinate is the input and the vertical coordinate is the output.

As can be seen from the figure, when γ is taken less than 1, the pixel value of the image increases; when γ is taken more than 1, the pixel value of the image decreases; when the value of γ is taken to be 0.8, 1.3, and 1.5, the effect of the processing of the insulator cap is shown in Figure 8.

As can be seen from the figure, when the value of γ is greater than 1, the overall brightness of the image is enhanced, but the part of the defective features with a lower grey value is enhanced relatively large, resulting in a decrease in the contrast between the defective and normal areas. When the value of γ is less than 1, the overall brightness of the image is reduced, and the contrast of the defects in the normal surface is enhanced, but a larger value also blurs the edges and the background boundary information.

After the above image enhancement, the dataset was divided and labelled for the collected defect images. The 4,000 defect images were divided into training set, validation set and test set according to the ratio of 8:1:1, among which the training set is 3200, the validation set is 400, and the test set is 400. Figure 9 shows the partial sample examples of the defects in the insulator's iron cap, and the dataset was manually labelled by Labeling software.

4.2 Experimental Environment and Parameter Settings

The Python version used in this experiment is 3.11, the training framework is Pytorch 2.3.1, and the graphics card is NVIDIA GeForce 4070 with 12GB of video memory. CUDA 12.1 was selected to train the model on the Win11 system. The training number was set to 150 epochs, batch size to 16, learning rate to 0.001, and momentum to 0.937.

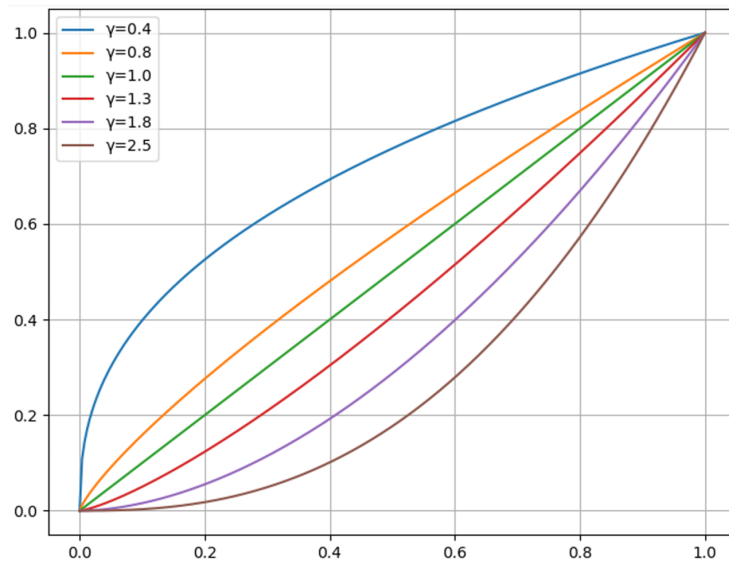


Figure 7. Curves of different gamma values

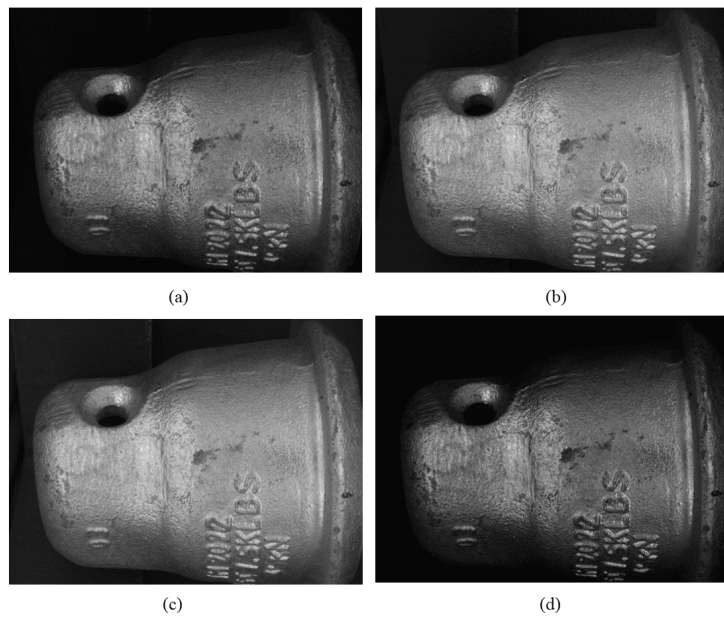


Figure 8. Comparison of different gamma values with the original figure: (a) Original diagram of the iron cap; (b) $\gamma = 1.3$; (c) $\gamma = 1.5$; (d) $\gamma = 0.8$

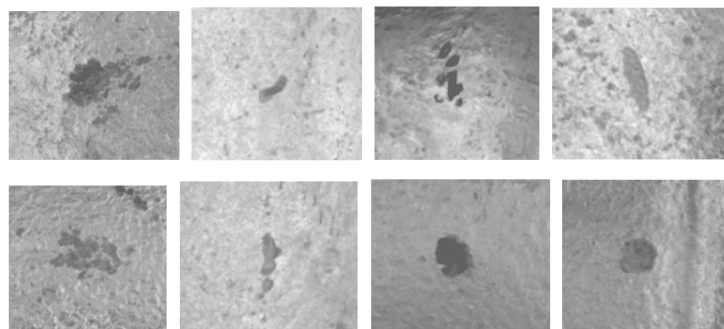


Figure 9. Sample defects

4.3 Evaluation Indicators

In order to validate the effectiveness of the proposed model, three metrics, mean accuracy (mAP), precision (P) and recall (R), were used for comprehensive evaluation [17].

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N \int_0^1 P_i(R_i) dR_i \quad (4)$$

where, TP denotes the number of positive samples predicted as positive samples, FP denotes the number of negative samples predicted as positive samples, and FN denotes the number of positive samples predicted as negative samples in the number of samples. p_i is the value of accuracy, R_i is the value of recall, and n is the number of classifications of the data samples [18]. mAP is the average of the area of each Precision-Recall (PR) curve, and takes the value ranging from 0 to 1.

4.4 Comparative Experiments with Different Algorithms

In order to further verify the feasibility of the YOLOv8n algorithm proposed in this study, it was compared with the Faster Region-based CNN (Faster R-CNN) algorithm [19], the Single Shot MultiBox Detector (SSD) algorithm [20], and other algorithms of the YOLO series, YOLOv5s and YOLOv8n, in the dataset. mAP@0.5 was used as a measure of the performance of the model, and the results of the comparison experiments are shown in Table 1.

Table 1. Comparison of detection results using different models

Model	mAP@0.5
Faster R-CNN	87.6
SSD	85.1
YOLOv5s	88.2
YOLOv8n-SCPN	90.6

4.5 Ablation Experiments

In order to verify the effectiveness of the improved algorithm proposed in this study, different experimental groups were set up and ablation experiments were conducted. As shown in Table 2, Experiment 1 was trained using YOLOv8n, whose data served as the control group. In Experiments 2 to 8, different modules were progressively replaced or added to the network.

Table 2. Ablation experiments

	BiFPN	C2f-SCConv	CBAM	P	R	mAP@0.5	mAP@0.95
1	×	×	×	0.841	0.728	0.879	0.513
2	✓	×	×	0.851	0.751	0.883	0.528
3	×	✓	×	0.853	0.742	0.894	0.516
4	×	×	✓	0.859	0.753	0.892	0.532
5	✓	✓	×	0.865	0.760	0.892	0.521
6	✓	×	✓	0.855	0.763	0.901	0.524
7	×	✓	✓	0.871	0.759	0.895	0.514
8	✓	✓	✓	0.882	0.776	0.906	0.542

From Experiments 2-4, it can be seen that compared with YOLOv8n, the improved algorithm has different degrees of improvement in mAP@0.5 and mAP@0.95 when using only the BiFPN module, the C2f-SCConv module or the CBAM. This indicates that the problem of defective targets, which BiFPN can easily identify incorrectly, is solved

after using the BiFPN module. The receptive field of the network was enlarged after using the C2f-SCConv module, which effectively improves the feature extraction capability of the network. The addition of the CBAM makes the network focus more on the defective region. From Experiments 5-7, it can be seen that combining any two of the three modules added to the network improves the detection accuracy compared to YOLOv8n. This further illustrates the effectiveness of these module improvements. From Experiment 8, it can be seen that after the simultaneous addition of the three improved modules, its mAP@0.5 and mAP@0.95 reached the highest among all experimental groups. Compared with YOLOv8n, the improvement is 2.7% and 2.9%, respectively. It illustrates the effectiveness of the three modules added to the network at the same time. In conclusion, the YOLOv8n-SCBP algorithm outperforms the original algorithm as well as the algorithms of Experiments 2-7 in the ablation experiments in terms of performance metrics, proving the superiority of the proposed algorithm in the task of stacked workpiece detection.

5 Conclusion

Aiming at the defects on the surface of insulator iron caps, the difficulty to distinguish the complex defect background leads to certain leakage and the wrong detection phenomenon of existing defect detection algorithms. Therefore, this study proposes a new YOLOv8n-SCPN algorithm on the basis of YOLOv8n. Firstly, the C2f convolutional network was replaced by the C2f-SCConv convolutional network in the backbone and neck network, which enhanced the ability of the model to extract the features on the surface of the insulator iron cap. Secondly, the CBAM was introduced after the SPPF layer, which enables the feature extraction network to extract more effective information about the surface defects of insulator caps. Finally, a BiFPN was added to the neck to replace the original feature fusion method in YOLOv8n in order to improve the model's ability to detect defects in insulator caps. Compared with the YOLOv8n algorithm, the YOLOv8n-SCPN algorithm proposed in this study has significantly improved the accuracy, recall and mAP on the home-made dataset. However, the YOLOv8n-SCPN algorithm still has low precision in detecting complex surface defects and still needs to be optimised in terms of the number of parameters. Therefore, the algorithm needs to be further improved in detecting complex surface defects so that it can achieve more lightweight and efficient detection.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] S. Sanyal, T. Kim, S. Jeon, Y. J. Lee, J. Yi, I. H. Choi, J. A. Son, and J. B. Koo, "Influence of corrosion on electrical and mechanical properties of porcelain suspension insulators: An overview," *Trans. Electr. Electron. Mater.*, vol. 21, pp. 543–549, 2020. <https://doi.org/10.1007/s42341-020-00239-3>
- [2] M. Hussain, "YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection," *Machines*, vol. 11, no. 7, p. 677, 2023. <https://doi.org/10.3390/machines11070677>
- [3] O. Khare, S. Gandhi, A. Rahalkar, and S. Mane, "YOLOv8-based visual detection of road hazards: Potholes, sewer covers, and manHoles," in *2023 IEEE Pune Section International Conference (PuneCon), Pune, India*, 2023, pp. 1–6. <https://doi.org/10.1109/PuneCon58714.2023.10449999>
- [4] J. Li, Y. Wen, and L. He, "SCConv: Spatial and channel reconstruction convolution for feature redundancy," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada*, 2023, pp. 6153–6162. <https://doi.org/10.1109/CVPR52729.2023.00596>
- [5] B. J. Souza, S. F. Stefenon, G. Singh, and R. Z. Freire, "Hybrid-YOLO for classification of insulators defects in transmission lines based on UAV," *Int. J. Electr. Power Energy Syst.*, vol. 148, p. 108982, 2023. <https://doi.org/10.1016/j.ijepes.2023.108982>
- [6] B. Dey, E. Dehaerne, and S. Halder, "Towards improving challenging stochastic defect detection in SEM images based on improved YOLOv5," in *Photomask Technology 2022, Monterey, California, United States*, 2022, pp. 28–37. <https://doi.org/10.1117/12.2645402>
- [7] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 3–19. <https://doi.org/10.48550/arXiv.1807.06521>
- [8] Y. Li, Q. Fan, H. Huang, Z. Han, and Q. Gu, "A modified YOLOv8 detection network for UAV aerial image recognition," *Drones*, vol. 7, no. 5, p. 304, 2023. <https://doi.org/10.3390/drones7050304>

- [9] A. A. Sami, S. Sakib, K. Deb, and I. H. Sarker, "Improved YOLOv5-based real-time road pavement damage detection in road infrastructure management," *Algorithms*, vol. 16, no. 9, p. 452, 2023. <https://doi.org/10.3390/a16090452>
- [10] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 8759–8768. <https://doi.org/10.48550/arXiv.1803.01534>
- [11] K. Wang, J. H. Liew, Y. Zou, D. Zhou, and J. Feng, "PANet: Few-shot image semantic segmentation with prototype alignment," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South)*, 2019, pp. 9196–9205. <https://doi.org/10.1109/iccv.2019.00929>
- [12] T. Li, Y. Zhang, Q. Li, and T. Zhang, "AB-DLM: An improved deep learning model based on attention mechanism and BiFPN for driver distraction behavior detection," *IEEE Access.*, vol. 10, pp. 83 138–83 151, 2022. <https://doi.org/10.1109/ACCESS.2022.3197146>
- [13] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10 781–10 790. <https://doi.org/10.48550/arXiv.1911.09070>
- [14] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117–2125. <https://doi.org/10.48550/arXiv.1612.03144>
- [15] E. Alibek and K. C. Kim, "Metal surface defect detection and classification using EfficientNetV2 and YOLOv5," *J. Korea Inst. Electron. Commun. Sci.*, vol. 17, no. 4, pp. 577–586, 2022. <https://doi.org/10.13067/JKIECS.2022.17.4.577>
- [16] W. Wang, X. Yuan, Z. Chen, X. Wu, and Z. Gao, "Weak-light image enhancement method based on adaptive local gamma transform and color compensation," *J. Sens.*, vol. 2021, no. 1, p. 5563698, 2021. <https://doi.org/10.1155/2021/5563698>
- [17] J. E. Kwon, J. H. Park, J. H. Kim, Y. H. Lee, and S. I. Cho, "Context and scale-aware YOLO for welding defect detection," *NDT E Int.*, vol. 139, p. 102919, 2023. <https://doi.org/10.1016/j.ndteint.2023.102919>
- [18] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang, "UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios," *Sensors*, vol. 23, no. 16, p. 7190, 2023. <https://doi.org/10.3390/s23167190>
- [19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [20] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision–ECCV 2016*. Springer, Cham, 2016, pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2