Overview of the Research on Improved YOLO Model in Insulator Defect Detection

Xuanrui Hu*

School of Computer and Information Engineering, Heilongjiang University of Science and Technology, Heilongjiang, China

*Corresponding author: huxuanrui@usth.edu.cn

Abstract. People's everyday lives depend on electricity, and insulators are essential to the reliable and safe functioning of high-voltage transmission lines. Therefore, defect detection of insulators has always been a concern for many research teams and researchers. The You Only Look Once(YOLO) series algorithm has always been the most widely used algorithm in the research of insulator defect detection. This paper seeks to examine the future directions for enhancing the YOLO model, evaluate the research problems in the field of insulator defect detection, and methodically study the pertinent theories for doing so. First, the paper outlines the main research progress in the field of insulator defect detection. Then focus on discussing the basic structure, dataset, improvement research, and experimental results of the three generations of YOLOv3, YOLOv5, and YOLOv8 models. Based on the above review, the paper proposes directions for enhancing the dataset of insulator defects and provides insights on improving the model for insulator defect detection. In addition, this article also points out the difficulties in collecting datasets in current research, providing research suggestions for future researchers.

Keywords: Insulator Defect Detection; YOLO Model Improvement; Dataset Enhancement.

1. Introduction

An essential part of high-voltage power transmission lines are insulators. Their main function is to maintain the insulation of the transmission line, prevent current back flow or grounding [1]. Insulators are mounted on transmission lines and exposed to various harsh environments and climates for extended periods of time. Therefore, insulators are prone to various problems. Among these issues, insulator self explosion is one of the most serious problems, the self explosion of insulators poses a significant safety threat to high-voltage transmission lines, easily causes faults in high-voltage transmission lines, and even leads to serious economic losses. Therefore, in order to ensure the safety and high-quality operation of the power grid, the detection and maintenance of insulators are crucial. Power line inspectors need to regularly inspect insulators for self explosion defects or other potential issues, defective insulators should be repaired or replaced in a timely manner to avoid causing greater losses. Nevertheless, these methods are neither precise, safe, nor efficient due to China's power system's massive and complicated structure. Therefore, power grid companies are using deep learning technology to automatically identify insulator defects [2].

Many deep learning-based object detection methods for insulators are already available. To identify faulty insulator portions, the Faster R-CNN technique was modified by Liao et al. to incorporate the deep residual network Resnet101. Compared with other object detection algorithms, this improves the detection accuracy of insulator defects, but the high computational cost leads to slow detection speed [3]. Zhu et al. modified the anchor box using the new K-means++ method and using an enhanced YOLOv3 technique to identify insulator flaws. Anchor boxes that are better suited to the size and shape of insulator faults can be obtained by applying the K-means++ method. K-means++algorithm also enables the model to more accurately locate the target and improve the accuracy of the model [4]. Cao and Jiang et al. proposed an efficient contour based segmentation network called Cap Net, and then used a design memory mechanism and polar coordinate loss (PAL) to promote training convergence and segmentation quality respectively. Finally, a lightweight algorithm called Mirror Fill was proposed to erase randomly selected capacitors. They effectively solved the problem of insufficient defect samples by using normal samples and their bounding box annotations, while utilizing the Mirror Fill algorithm to generate multiple defect samples from a single normal insulator, increasing sample diversity. The Cap Net model proposed by Cao and Jiang et al. can perform unsupervised anomaly detection, achieving an F1 score of 37% for defect samples. This improves the flexibility and robustness of the insulator detection model, significantly enhancing its detection performance [5]. It can be seen that there are numerous object detection algorithms related to insulators based on deep learning.

This article introduces the mainstream algorithm structure of Yolo series for target detection of insulator blasting defects, and lists the improvements of various generations of algorithms in existing research. Then it analyzed the experimental results of various improved algorithms. Finally, it summarized the difficulties of existing research and prospects for research directions in dataset enhancement and model improvement.

2. YOLO series algorithm for target detection of insulator blasting defects

2.1. Introduction to typical YOLO algorithms in the field of insulator defect detection

2.1.1 Introduction to YOLO v3 Algorithm

The third iteration of the YOLO series algorithm is called YOLOv3. YOLOv3 is a fully convolutional network that detects targets by using multi-scale characteristics. By directly predicting bounding box and category probabilities in images, YOLOv3 turns object recognition difficulties into regression problems. It then returns the bounding box position and category without producing candidate regions. The Darknet-53 network, which consists of 53 convolutional layers without pooling or fully connected layers, serves as the foundational backbone network for YOLOv3. The Darknet-53 network uses a fully convolutional network design, which modifies the tensor size by varying the convolution kernel's stride during forward propagation. In addition, Darknet53 also introduces residual structure, which deepens the network by drawing on the residual structure of Res Net, greatly reducing the difficulty of training deep networks. Fig. 1 shows the network structure of YOLOv3.

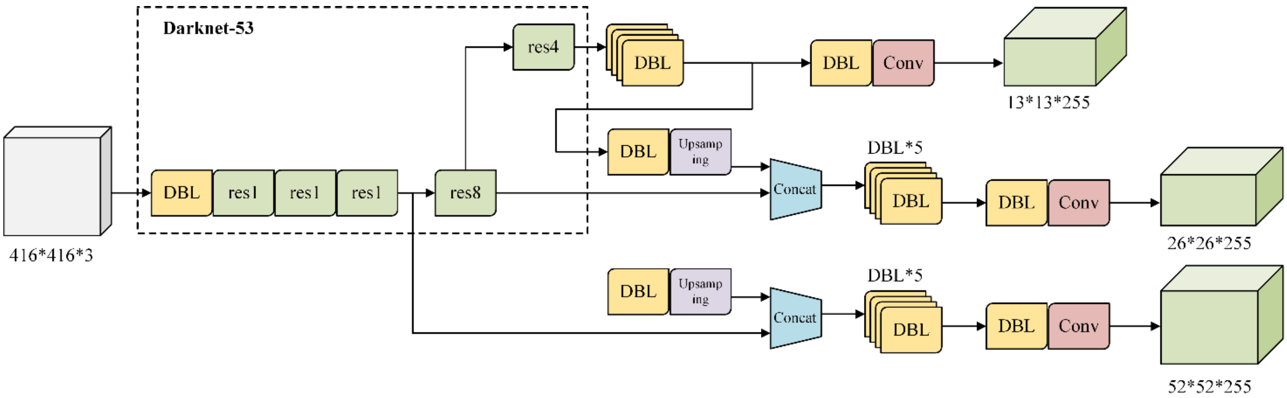

Fig. 1 YOLOv3 network structure

2.1.2 Introduction to YOLO v5 Algorithm

YOLOv5 is a single object detection algorithm mainly composed of a backbone network (Backbone), a middle network (Neck), and a head network (Head). On the backbone network, YOLOv5 has Focus module and Conv convolution module, and adds SPP structure on CSPDarknet53, while adopting PANet structure in the middle of the network. Use YOLOv3 header in the head network, including three output detection headers. At present, YOLOv5 has four structures: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x.The depth of these four models and the width of their feature maps increase sequentially, so it resulting in higher accuracy. But this also means that the computational cost of the model will increase. For insulator testing carried on drones, YOLOv5s structure is often chosen due to its light weight requirements. Fig. 2 shows the network structure of YOLOv5s.

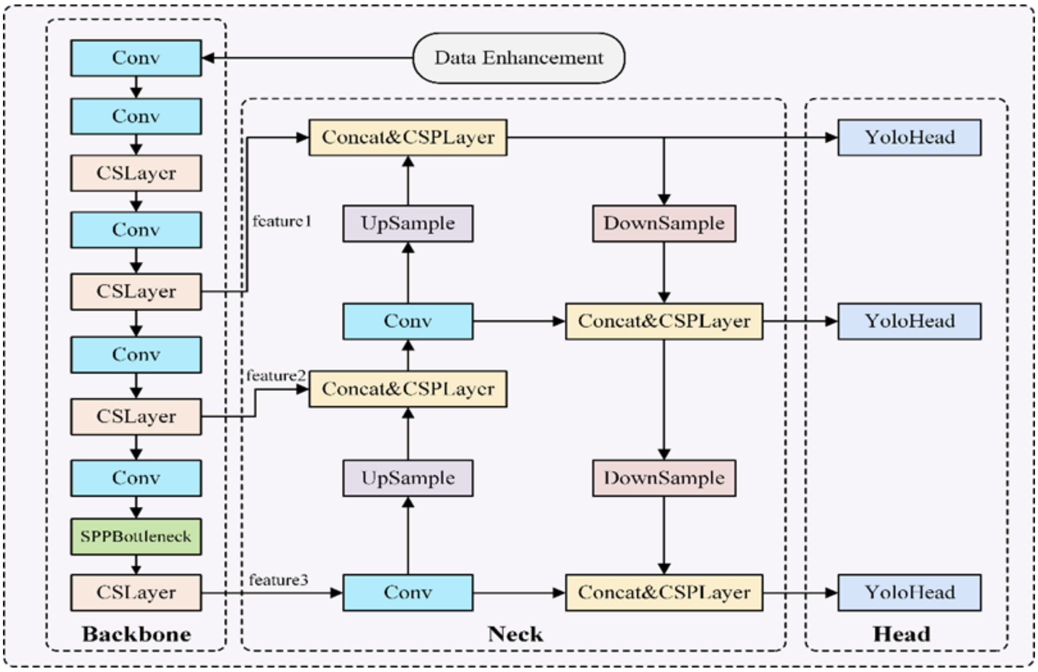

Fig. 2 YOLOv5s network structure

2.1.3 Introduction to YOLOv8 Algorithm

An input, a backbone network, a middle network, and a head network make up the object detection model YOLOv8. Images can be segmented using YOLOv8 into numerous tiny networks, each of which can predict various bounding boxes and associated object probabilities. Convolutional neural networks are used by YOLOv8 to examine and process target photos before producing and delivering object detection results. The YOLOv8 network structure is displayed in Fig. 3.

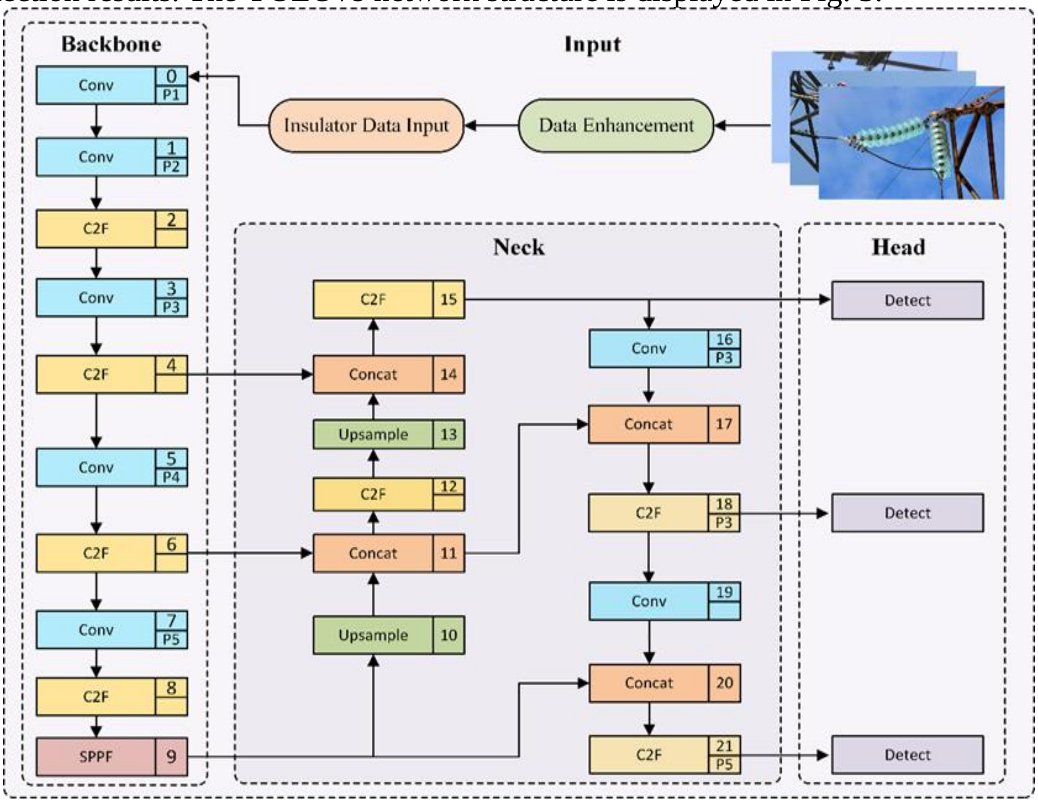

Fig. 3 YOLOv8 network structure

2.2. Improvement of YOLO series algorithm for target detection of insulator blasting defects

2.2.1 Research and Improvement of YOLOv3

In YOLOv3, Yao and Qin enhanced the IOU loss function utilized in the conventional anchor-based detection approach and suggested an insulator detection technique that substitutes the GloU (Generalized Intersection over Union) loss function for the original loss function. The GIoU series of calculating formulas are as follows [6].

$$\text{IoU:} \, IoU = \frac{I}{A^{P} + A^{g} - I} \tag{1}$$

$$\text{GIoU:} \, GIoU = IoU - \frac{|A_c - U|}{|A_c|} \tag{2}$$

$$\text{GioU Loss: Loss} = 1 - GLoU \tag{3}$$

Zhu et al. used the K-means++ algorithm to adjust the anchor box and improve the underlying network Darknet-53 feature extraction network of the model (referred to as Darknet-41-YOLOv3 in this article). Firstly, by reducing the number of convolutional layers and the number of residual units used, the computational complexity and network depth can be reduced. Then, by increasing the scale detection, shallow feature information is added to the prediction results to enhance the sensitivity of detecting small targets [3].

2.2.2 Research and Improvement of YOLOv5

In the research and improvement of YOLOv5, Chen and Du proposed a detection method based on the improved YOLOv5s model (referred to as SAM-YOLOv5s in this paper) [7]. This model introduces a self attention mechanism in the convolutional layers associated with the CSP module and feature output. In addition, depthwise separable convolution has been introduced.

Zhang and Lv proposed an improved insulator defect detection method called BS-YOLOv5s, which is based on SimAM 3-D attention mechanism and YOLOv5s based network [8].BS-YOLOv5s evaluates the weight of each neuron by introducing SimAM 3-D attention mechanism. Concurrently, the model suggested the Bi Slim neck form and revised the fusion network. The newly created network includes insulator defect characteristics of various sizes while preserving the deeper aspects of tiny targets. The SimAM 3-D attention mechanism's energy function is as follows.

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{\mathrm{x}}_{\mathrm{i}})^2 \tag{4}$$

2.2.3 Research and Improvement of YOLOv8

YOLOv8 is the latest model widely used in the field of insulator blasting defect detection. In the research of YOLOv8 model, Hu and Tang proposed a light weight detection algorithm based on YOLOv8n[9]. This detection algorithm introduces an improved network based on ASF-YOLO, which can automatically adapt to learning different receptive fields at different positions in the feature map. At the same time, the traditional convolution module in the Backbone section will be replaced with the GhostConv module. The algorithm also adds a RepViTBlock module, allowing the model to more effectively capture spatial information in images.

Gao and Han et al. also proposed a new detection algorithm (referred to as DCN-IRMB-YOLOv8 in this paper).They added an object detection head to the original YOLOv8 structure, allowing the network model to focus more on small target objects[1]. Secondly, use Deformable NetWork Convolution v2 (DCNv2) to replace the regular convolutions in the C2f module. In addition, an Inverse Residual Mobile Block (IRMB) module has been introduced in the backbone network. Finally, use the WIOU function as the loss function. The following is the calculation formula for deformable convolution.

$$y(p_0)=\sum_{k=1}^{k} w_k \cdot x(p_0+p_k+\Delta p_k)\cdot\Delta m_k \tag{5}$$

3. Experimental results analysis of YOLO series improved algorithms for target detection of insulator blasting defects

3.1. evaluating indicator

The evaluation indicators for all improved models in this article are as follows. From (6) to (15), each indicator of particular definitions is displayed.

$$\mathrm{R}=\frac{TP}{TP+FN} \tag{6}$$

With R standing for recall, TP for properly projected positive samples, and FN for mistakenly forecasted negative positive samples.

$$\mathrm{P}=\frac{TP}{TP+FP} \tag{7}$$

Where P stands for accuracy and FP is number of samples that were negative inadvertently projected as positive.

$$\mathrm{F1}=2\times\frac{\mathrm{P}\times\mathrm{R}}{\mathrm{P}+\mathrm{R}} \tag{8}$$

The precision and recall harmonic means are used to compute the F1 Score.

$$\mathrm{FPS}=\frac{1}{t} \tag{9}$$

Where t is the time needed to analyze each image and the model's detection speed is evaluated using the frame rate per second, or FPS.

$$\mathrm{AP}=\int_0^1 P(R)dR \tag{10}$$

Where AP stands for average accuracy, a performance metric for the model.

$$\mathrm{mAP}=\frac{\sum AP}{N} \tag{11}$$

Where N is the whole group of the chosen target number, and mAP is the mean average accuracy chosen when there are numerous detection targets.

$$mAP0.5=\frac{1}{N}\sum_{i=1}^{N} AP0.5_i \tag{12}$$

Where MAP0.5 stands for mAP at the 50% IOU cutoff.

$$Avgloss=\frac{1}{N}\sum_{i=1}^{N} L(y_i,\hat{y}_i) \tag{13}$$

Where Avgloss is the average loss across all data points and N is the total number of data points. This i data point is the i-th one. Where y_i is the true value and $\hat{y}_i$ is the model's predicted value.

$$Accuracy(\%)=(\frac{TP+TN}{TP+TN+FP+FN})\times 100\% \tag{14}$$

The number of samples that the model accurately predicted divided by the total number of samples is the accuracy. The percentage of positive category samples that the model accurately predicted is known as True Positives (TP). The number of negative category samples that the model accurately predicted is known as True Negatives (TN). The number of negative category samples that the model mistakenly anticipated to be positive is known as False Positives (FP). The quantity of positive category data that the model mistakenly predicted to be negative is known as False Negatives (FN).

$$\mathrm{GF}LOPs = \frac{\text{Total number of floating point operations}}{\text{Time in seconds}} \times 10^{-9} \tag{15}$$

where GFLOPs (Giga-FLOPS) represent the number of billion floating-point operations that may be executed per second. The "total number of floating-point operations" for a given computing work is the total amount of operations carried out in floating-point.

3.2. Dataset Overview

Table 1 presents an overview of the dataset for the models reviewed in the article.

Table 1. Details of the dataset

Model	Dataset name	Number of images	perform data augmentation(Yes/No)	Dataset format	Labeling software
GLOU-YOLOv3	DJI M200	500	Yes	VOC format	LabelImg software
Darknet-41-YOLOv3 model	Drone aerial images	1500	Yes	VOC format	LabelMe
SAM-YOLOv5s	China Electric Power Insulator Dataset	848	No	VOC2007	not provided
BS-YOLOv5s	WI	1260	Yes	not provided	not provided
Light-YOLOv8	Public dataset	3630	Yes	not provided	not provided
DCN-IRMB-YOLOv8	CPLID	600	Yes	not provided	Label Image

3.3. Comparison and analysis of experimental results

3.3.1 Research and Improvement of YOLOv3

Table 2. Performance of GIOU-YOLOv3 model

Model	mAP (%)	Avgloss	Speed (f/s)
YOLOv3	90.0	0.83	58.8
GLOU-YOLOv3	92.5	0.71	58.8

By comparing the experimental results of YOLOv3 and GLOU-YOLOv3 algorithms under the same experimental conditions in table 2, it can be seen that the improved GLOU function of the model has increased its accuracy by 2.5%. Therefore, GLOU is an effective improved model solution. However, by observing the Speed metric, it can be found that after introducing GLOU, the detection speed of the model remains unchanged and there is no significant improvement.

Table 3. Performance of Darknet-41-YOLOv3 model

Model	Test number	Misdiagnosis count	Leakage count	mAP(%)	Accuracy/%	Average detection time/ms
YOLOv3	956	57	138	76.05	94.87	36.133
Darknet-41-YOLOv3	956	41	105	83.21	95.99	30.975

By comparing the experimental results of YOLOv3 and Darknet-41-YOLOv3 algorithms under the same experimental conditions in table 3, it can be seen that the model successfully reduces the computational complexity and network depth by removing redundant convolutional layers, it greatly improves the detection speed of the model and reduces the loss of feature information. At the same time, The K-means algorithm has been effectively enhanced by Darknet-41-YOLOv3, which decreases the number of iterations and increases computing performance.

By comparing these two sets of improvements, the paper roughly determines the direction of improvement for YOLOv3. GIOU-YOLOv3 has improved the loss function of the model. Similarly, the corresponding features in YOLOv3 can also be improved, such as the LeakyReLU activation function, which is a reliable direction for improving YOLOv3 models. Darknet-41-YOLOv3 has improved the number of convolutional layers in Darknet-53, and it is concluded that appropriately improving the convolutional layers in Darknet-53 can greatly enhance the performance of YOLOv3 models. Darknet-41-YOLOv3 has also successfully improved the K-means algorithm, so attempting to improve the algorithms in the model, such as improving the logistic regression algorithm for predicting bounding box confidence, is also a reliable direction for improvement.

3.3.2 Research and Improvement of YOLOv5

Table 4. Performance of SAM- YOLOv5s model

Model	Precision (%)	Recall (%)	mAP(%)	FPS	F1(%)
YOLOv5S	92.39	92.10	87.90	56.23	87.90
SAM- YOLOv5s	94.79	94.37	89.10	63.9	89.10

The data in table 4 shows that after the improvement, the recall rate of the YOLOv5s model has increased to 94.37%, the FPS has increased to 63.9%, and the accuracy of the model has increased to 94.79%. All indicators are better than YOLOv5s. Therefore, it can be concluded that integrating self-attention mechanism and decomposing standard convolution into depthwise separable convolution in CSPLayer effectively reduces the complexity of the model and improves its efficiency.

Table 5. Performance of BS-YOLOv5S model

Model	F1	mAP0.5	FPS	weight
YOLOv5S	0.879	0.867	57.47	13.7
BS-YOLOv5S	0.891	0.901	66.23	13.1

Compared with the experimental results in table 5, BS-YOLOv5s showed significant improvements in mAP0.5, frames per second (FPS), weight, and F1, Among them, mAP0.5 improved by 3.4% compared to YOLOv5s, which proves that the newly added SimAM attention mechanism and the redesigned Bi-Silm-neck have successfully improved the performance of the model The reduction of the weight parameter clearly indicates that the lightweight design of the model significantly reduces the number of parameters.

By comparing the improvements of two models, the paper determines the general direction of improvement for YOLOv5. Firstly, both improvements introduce attention mechanisms. This proves the importance of attention mechanisms in improving YOLOv5s models. Therefore, the performance of the model can be improved by modifying more appropriate and efficient attention mechanisms.

Secondly, both models have made different improvements to the network layer, both of which have a positive effect of reducing parameters and improving efficiency. Therefore, it can be concluded that reducing redundant network layers and modifying the structure of network layers is another improvement direction for YOLOv5s.

3.3.3 Research and Improvement of YOLOv8

Table 6. Performance of Light-YOLOv8 model

Model	GFLOPs	mAP(%)	Parameter count/10^6	Model size/MB
YOLOv8	28.4	88.6	11.1	22.5
Light-YOLOv8	27.1	93.5	6.1	12.9

When comparing the experimental results in Table 6, it is evident that Light-YOLOv8's parameter count and model size are substantially lower than YOLOv8's. This suggests that swapping out the convolutional block can successfully lower the number of model parameters while preserving high accuracy, which in turn reduces the model's size. Simultaneously, mAP rose by 4.9%, demonstrating how the addition of RepViTBlock modules and better networks greatly improves the model's performance.

Table 7. Performance of DCN-IRMB-YOLOv8 model

Model	Precision (%)	mAP0.5(%)	Recall	Parameter count/10^6	FPS
YOLOv8	91.0	95.1	91.3	3.01	81.4
DCN-IRMB-YOLOv8	92.8	98.0	94.8	2.82	64.3

Table 7 demonstrates that the enhanced YOLOv8 algorithm's detection accuracy has increased by 2.9% while utilizing fewer parameters. This shows that the number of parameters has been successfully decreased by substituting Deformable Convolution v2 (DCNv2) for the regular convolution in the C2f module. The accuracy and regression rates were improved to 92.8 and 94.8, respectively, proving that introducing the reverse residual attention module and referencing WIOU as the loss function can significantly improve the detection performance of the model.

By combining the improvements of two YOLOv8 models, the general direction of improvement for the YOLOv8 model can be derived. Both improvements have modified the convolutional layers of the original YOLOv8 model, reducing computational complexity and network depth, decreasing parameters and model size, and enhancing the light weighting of the model. Secondly, DCN-IRMB-YOLOv8 introduces attention mechanism in the backbone network, which indicates that YOLOv8, like YOLOv5, adding attention mechanism is also one of the successful improvement directions. Finally, the model also introduces the WIOU loss function, which is similar to the improvement idea of YOLOv3's Glou loss function For YOLOv8, the loss function is also an improved method that can enhance the accuracy of the model.

4. Challenges in research on insulator defect detection and prospects for future research directions

The most difficult aspect in the research of insulator defect detection should be the collection of datasets. The working environment of insulators is mostly in the wild, with spatial locations at high altitudes. Currently, the most widely used data collection methods are drone aerial photography. However, this method is difficult to collect large-scale data in bulk. Therefore, how to collect large-scale insulator datasets from different perspectives is a pressing issue that requires this study's solution.

In the identification of defects in insulators, traditional manual climbing inspections are no longer suitable for the current scale of the power grid, and drone based aerial image inspections have become mainstream, with an efficiency of up to 40 times that of manual inspections and being safe and reliable [10]. From the perspective of the dataset, although most improved models have enhanced the dataset of insulators when preparing training data, they rarely consider environmental factors such as shaking and weather in real-life shooting. In the improved model selected in this article, BS-YOLOv5s adjusts the resolution of the image to 640x640. The imgaug library in AI Studio simulates the use of rain, snow, fog, and low light scenes to flip, mirror, noise, motion blur, and expose the original image. In the Light-YOLOv8 dataset, the imgaug image enhancement library in Python was used to simulate three severe weather conditions of snow, rain, and fog for image enhancement. The authors of these two models have taken good note of insulator target detection under different weather and shooting conditions. When conducting research on insulator blasting defect detection, the paper suggests considering weather factors such as rainy and snowy days, as well as practical issues related to detecting equipment shaking.

Reducing the model's size and increasing its detection accuracy are the primary objectives of model improvement. There are now a number of popular improvement paths. One option is to alter the model's convolutional layers, which can decrease parameter sizes to prevent redundancy or alter the convolution technique to increase efficiency and lower parameter sizes. To improve the model's detection accuracy, the second strategy is to incorporate or enhance attention processes in the backbone network. To increase the model's accuracy and resilience, the third step is to optimize the loss function.

5. Conclusion

The use of YOLO series improved algorithms for defect detection in insulator images has always been a concern for many research teams and researchers. This article first reviews the current status of research in the field of insulator defect detection, and then elaborates and organizes the basic structures of the three most widely used YOLO models. This article also summarizes the improvements and extensions of three models in insulator defect detection. Thus, through comparative analysis, the approximate improvement directions of each of these three YOLO models were demonstrated, provided researchers with excellent ideas for optimizing and improving this version of the model. Finally, by summarizing and generalizing, the challenges faced by current research on insulator defect detection were summarized, and a general improvement approach for insulator datasets and all YOLO detection models was proposed. In summary, this article provides a systematic review and outlook for the research of YOLO series models in the identification of insulator blasting defects, hoping to have a positive impact on research and practice in related fields.

References

[1] Gao Y M, Han S J, Zhou B, etc Improved YOLOv8 insulator and self explosion defect target detection algorithm. RadioEngineering, 2024, 1-10.

[2] Zhu Y K, Wang J L, et al. Analysis of flashover fault causes and preventive measures for composite insulators. Inner Mongolia Electric Power Technology, 2021, 39(6): 82-86.

[3] Liao G P, Yang G J, Tong W T, et al. Study on power line insulator defect detection via improved faster region-based convolutional neural network. Proceedings of the 2019 IEEE 7th International Conference on Computer Science and Network Technology. Dalian, China: IEEE, 2019: 262-266.

[4] Zhu Y C, Zheng Y, Qin J L. Research on insulator target detection based on improved YOLOv3. Electric porcelain lightning arrester 2022, (03): 166-171.DOI: 10.16188/j.isa.1003-8337.2022.03.024.

[5] Jiang D, Cao Y, Yang Q, CapNet: Learning Insulator Self-Blast From Bounding Box, in IEEE Transactions on Instrumentation and Measurement, 2024, vol. 73, pp. 1-10

[6] Yao L N, Qin Y Y, Insulator Detection Based on GIOU-YOLOv3, 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 5066-5071, doi:10.1109/CAC51589.2020.9326959.

[7] Chen Y X, Du Z A, Li H X, Zhang K J and Wen P. Insulator Defect Detection Based on Improved YOLOv5 Model, 2024 IEEE 4th International Conference on Software Engineering and Artificial Intelligence (SEAI), Xiamen, China, 2024, pp. 123-127, doi: 10.1109/SEAI62072.2024.10674078.

[8] Zhang Z B, Lv G H, Zhao G X, Zhai Y, Cheng J Y, BS-YOLOv5s: Insulator Defect Detection with Attention Mechanism and Multi-Scale Fusion, 2023 IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 2023, pp. 2365-2369, doi: 10.1109/ICIP49359.2023.10222163.

[9] Hu W H, Tang C L. Lightweight YOLOv8 for insulator defect detection in complex weather conditions Journal of Chongqing Technology and Business University (Natural Science Edition) 2024, 1-10. http://kns.cnki.net/kcms/detail/50.1155.N.20240926.1059.004.html.

[10] Li D P, Ren X M, Yan N N. Real time detection of insulator string drop based on drone aerial photography. Journal of Shanghai Jiao Tong University, 2022, 56(08):994-1003. DOI: 10.16183/j.cnki.jsjtu.2021.416.