IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 12 August 2024, accepted 28 August 2024, date of publication 2 September 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3452782

RESEARCH ARTICLE

Improved YOLOv8n for Foreign-Object Detection in Power Transmission Lines

HANJUN WANG, SHIYU LUO, AND QUN WANG
Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences, Shenyang 110000, China
Corresponding author: Shiyu Luo (luoshiyu22@mails.ucas.ac.cn)

ABSTRACT Effective and accurate detection of foreign objects in transmission lines plays a crucial role in achieving intelligent power inspection. However, in the real world, detecting objects that are too far or too small can lead to inaccurate object detection tasks. Therefore, this article proposes an improved model based on YOLOv8n to improve detection performance. We introduce attention mechanism into the YOLOv8n network and add a small object detection module to improve detection accuracy. Considering the requirements of detection tasks for detection speed and accuracy, after comparing the three attention mechanisms of CBAM, ECA, and GAM, we chose the backbone network formed by the fusion of YOLOv8n and ECA attention mechanism, and added a small object detection module in the head section. The results show that compared to the unimproved YOLOv8n model, this method can effectively improve detection accuracy and still perform excellently in detection speed and robustness.

INDEX TERMS Transmission line, foreign object detection, YOLOv8n, ECA, small object detection.

I. INTRODUCTION

Object detection is an important task in the field of computer vision [1], whose main purpose is to correctly recognize specific objects in an image or video. Nowadays, object detection is widely used in various fields, including the electrical field. Electricity is ubiquitous in daily life, and transmission lines are the key to delivering electricity to every household. The transmission line network covers a wide range, usually spanning different terrains and landforms such as cities, rural areas, and mountainous areas, and usually uses high-voltage electricity for power transmission. The power system requires extremely strong stability and reliability. The presence of foreign objects on transmission lines, such as kites, balloons, bird nests, etc., may lead to faults such as short circuits and arc discharges, and even cause fires or power grid accidents. These safety hazards may affect the reliability and stability of the power system [3]. In summary, conducting foreign object detection on transmission lines is the key to ensuring the safe operation of the power grid, reducing power losses, improving the efficiency of power grid operation, and reducing maintenance costs.

The associate editor coordinating the review of this manuscript and approving it for publication was Kai Song.

Manual inspection of transmission lines requires a significant investment of human resources, and line inspectors can only inspect transmission lines one by one by walking or using vehicles, which cannot achieve all-weather and all-round coverage. This leads to lower inspection efficiency and may result in missed or delayed detection of issues. In addition, manual inspections require line patrol personnel to enter the high-voltage power grid work area. Despite safety measures, there is still a potential risk of accidents occurring. The combination of deep learning technology and robot intelligent inspection of transmission lines can effectively solve the above problems [5], [6], [7].

Kites, balloons, and other foreign objects usually appear at high places. Due to changes in distance, the detected object will become small targets. The key to foreign object detection in power transmission lines is how to effectively identify these small targets. Although object detection has developed rapidly in recent years, small object detection remains a challenging problem.

YOLOv8 adopts anchor free frame object detection, which has a faster detection speed. YOLOv8 has better detection accuracy compared to previous versions. Among the multiple sizes of YOLOv8, n is the smallest size with the least number of parameters, the lowest complexity, and the fastest processing speed. Considering the real-time detection of power inspection work, edge computing is often used for actual project deployment, which requires high running speed, so we choose YOLOv8n as our baseline network. In order to solve the problem of difficult small target recognition, we attempted to use attention mechanisms to strengthen the connection between images in space and channels, and added a small target detection module to capture the features of large and small targets. This work has made good progress and can effectively improve the accuracy of model detection.At the same time, while ensuring high accuracy, the size of the model remains at a relatively small level and has a high operating speed, making it possible to run at the terminal devices.

II. RELATED WORK

With the rapid development of deep learning technology and the improvement of computing power, more and more scholars are applying deep learning technology to various industries. Many scholars have applied deep learning to the field of foreign object detection in transmission lines.

Zhang et al. improved the RCNN model using RPN (Region Proposal Network) technology to identify foreign objects in transmission lines and achieved better results than Faster RCNN [8]. Hao et al. also used RPN technology to improve the Faster RCNN model, achieving an mAP of 95.24% [9]. Chen et al. used MASK RCNN as the backbone network to detect foreign objects [10]. Liang et al. used Faster RCNN network to detect foreign objects and transmission component defects in transmission lines, with an mAP of 91.11% [11]. Zhu et al. used multi-scale feature pyramid connections in CNN to fuse multi-level information, achieving an mAP of 88.1% [12].

Later, many scholars began to introduce the YOLO model into this field. Li et al. improved the YOLOv3 model by using Mobilenetv2 instead of Darknet-53 as the backbone network to identify foreign objects in transmission lines [13]. Li et al. used Atrous Spatial Pyramid Pooling and Convolutional Block Attention Module to improve the YOLOX model, resulting in an mAP increase of 4.24% compared to the baseline YOLOX model [13]. Wu et al. used YOLOv4 as the baseline model, replaced CSPDarkNet53 with MobileNetV2, replaced standard convolutions in SPP and PANet modules with depthwise separable convolution (DSC), and embedded CBAM into SPP and PANet modules, achieving an mAP of 96.71% [15]. Wang et al. used YOLOv8m as the baseline model, integrated the GAM attention module into the backbone network, replaced the SPPF module with the SPPCSPC module, and introduced the focus eiou loss function. The mAP reached 95.5% [2].

III. METHODOLOGY

In this chapter, we will discuss the architecture of YOLOv8n, modifications to the YOLOv8n model, hyperparameter settings, image preprocessing methods, dataset preparation, and model evaluation methods.

A. YOLOv8n

YOLOv8 is the latest version of the object detection model in the YOLO series proposed by Ultralytics and others. The letters n/s/m/l/x denote different sizes of the YOLOv8 model, with the size determining the model's complexity and number of parameters. Different sizes can be used to meet the needs of different scenarios. Considering the need for foreign object detection speed on transmission lines, we have chosen the n size to reduce detection time.

The backbone of YOLOv8 utilizes the CSPDarknet network structure [17], an improved version based on Darknet. YOLOv8 replaces the C3 module in the YOLOv5 backbone network with the C2f module and introduces residual connections to reduce overfitting compared to YOLOv5.

The neck of YOLOv8 adopts the PANet structure, which consists of a series of feature fusion and upsampling operations. The feature fusion layer is used to fuse feature maps from different layers to obtain richer information. The upsampling layer is used to upsample low-resolution feature maps to the same size as high-resolution feature maps. Through upsampling, low-resolution feature maps can be restored to the original size of the image for better prediction of target positions.

The head of YOLOv8 uses anchor-free object detection, which can directly predict the center of the target through the classifier and regressor. Considering the need for rapid detection of foreign objects on power transmission lines, the 'n' size of YOLOv8, known as YOLOv8n, was chosen to reduce detection time. The structure of the YOLOv8n model is illustrated in Figure 1.

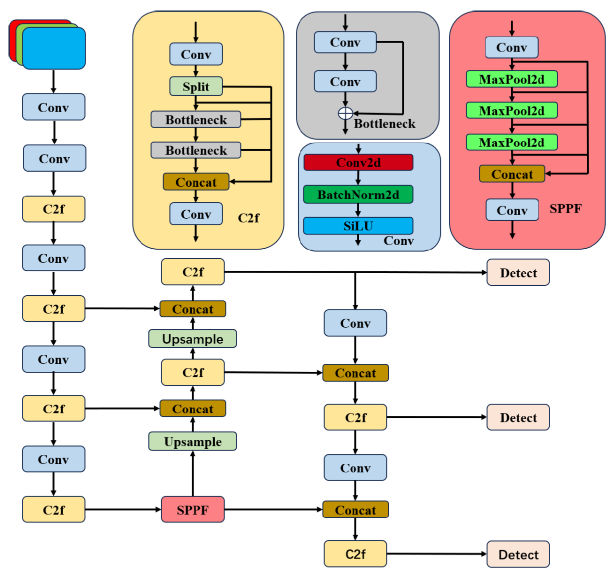


FIGURE 1. YOLOv8n model.

B. LOSS FUNCTION

The loss function of YOLOv8 consists of various components, including classification loss and regression loss. The classification loss function adopts VFL Loss [25], also known as logarithmic loss, which is a commonly used loss function in classification problems. It measures the difference between the probability distribution predicted by the model and the actual labels, and is used to measure the accuracy of the model's predictions. The VFL Loss function formula for YOLOV8 is as follows:

$$VFL(p,q) = \begin{cases} -q(q(log(p)+(1-q)log(1-p)), & q>0 \\ -\alpha p^{\gamma} log(1-p), & q=0 \end{cases} \tag{1}$$

where q represents the intersection and IoU between the predicted box and the true box, and p represents the probability. The regression loss function uses CIoU Loss [26]. CIoU is a regression loss function used to measure the degree of deviation between the predicted target box and the true box. It takes into account factors such as the center point offset of the target box, the difference in aspect ratio and the size of the box, making it more accurate than traditional IoU loss. The formula for CIoU Loss is as follows:

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho(p,p^{gt})}{c^2} + \alpha\upsilon. \tag{2}$$

$$\upsilon = \frac{4}{\pi^2}(arctan\frac{w^{gt}}{h^{gt}} - arctan\frac{w}{h})^2. \tag{3}$$

$$\alpha = \frac{\upsilon}{(1-IoU)+\upsilon}. \tag{4}$$

where ρ^2(p,p^{gt}) represents the square of the predicted box center point and the true box center point, and c^2 represents the square of the diagonal length of the minimum bounding rectangle between the two rectangles. $\alpha\upsilon$ is a parameter that measures the aspect ratio.w and h represent the width and height of the ground truth. w^{gt} and h^{gt} represent the width and height of the predicted box.

C. MODIFICATIONS

Although YOLOv8n has demonstrated its powerful performance in many fields, there is still room for optimization when applied to foreign object detection in transmission lines. We enhanced the accuracy of the model for small object detection by introducing an attention mechanism and adding a small object detection layer. Modifying the model will lead to an increase in the number of parameters, which will result in an increase in detection time, but this is acceptable.

D. ECA

ECA (Efficient Channel Attention) is a channel attention mechanism proposed by Wang et al. [19]. This refers to a lightweight channel attention mechanism that employs a 1D convolution for non-dimension-reducing local cross-channel interaction strategies, as depicted in Figure 2. Their experiments have shown that ECA can bring significant gains with a small number of parameters.

The ECA module, which only emphasizes inter-channel relationships and not spatial ones, has proven to be more significant in detecting foreign objects on power transmission lines. Testing three attention modules: ECA [19],

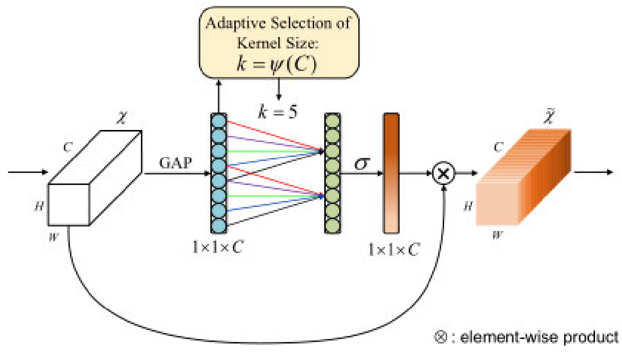


FIGURE 2. Diagram of ECA module.

GAM [20], and CBAM [21]. ECA showed superior performance, despite GAM and CBAM employing both spatial and channel attention mechanisms. Wang et al. introduced the GAM module in the YOLOv8m model [2], applying it before the SPPF layer. In contrast, our experiments involved incorporating an ECA module after each C2f module, leading to enhanced accuracy. However, using multiple ECA modules did not synergize well with the small object detection module. After thorough comparison, we chose to integrate a single ECA module before the SPPF module, combined with a small object detection module. Due to ECA's lightweight design, this approach maintained processing speed even with the additional detection module.

E. SMALL OBJECT DETECTION MODULE

Due to distance, foreign objects on power lines become small targets in images. Chen et al. [22] and others define small objects as those whose bounding box area is between 0.08% and 0.58% of the image area. Other common definitions include a bounding box width and height to image width and height ratio less than a certain value for example such as 0.01 or the square root of the bounding box area to image area ratio being less than a certain value such as 0.03. The MS COCO dataset defines small objects as those with a resolution smaller than 32 × 32 pixels. After multiple convolutions in YOLOv8, small object features become less distinct, significantly impacting detection accuracy. Upsampling can enlarge feature maps for better recognition. Wang et al. summarized three common upsampling methods: Interpolation, Deconvolution, unPooling [23], and also proposed the lightweight upsampling operator CARAFE [24] in 2019. In YOLOv8n, the original upsampling layer used interpolation, which we continued to use without modification.

We have added an additional small object detection module in the neck section, which includes an upsampling layer, a fully connected layer, and a C2F module. The upsampling layer uses the nearest neighbor upsampling method and concatenates it with the feature map obtained from the previous backbone to obtain a new feature map for detection. The advantage of this is that the new feature map obtained is magnified to twice the original size. We then perform three more C2f operations to pass the feature map down. In the end, we will also obtain another feature map that has been magnified twice.

The new feature map allows us to obtain a larger receptive field, enlarge the features of the target, and improve the accuracy of small object detection. Through experiments, it has been proven that this improvement is very effective.

The revised structure of the model, incorporating these improvements, is depicted in Figure 3.

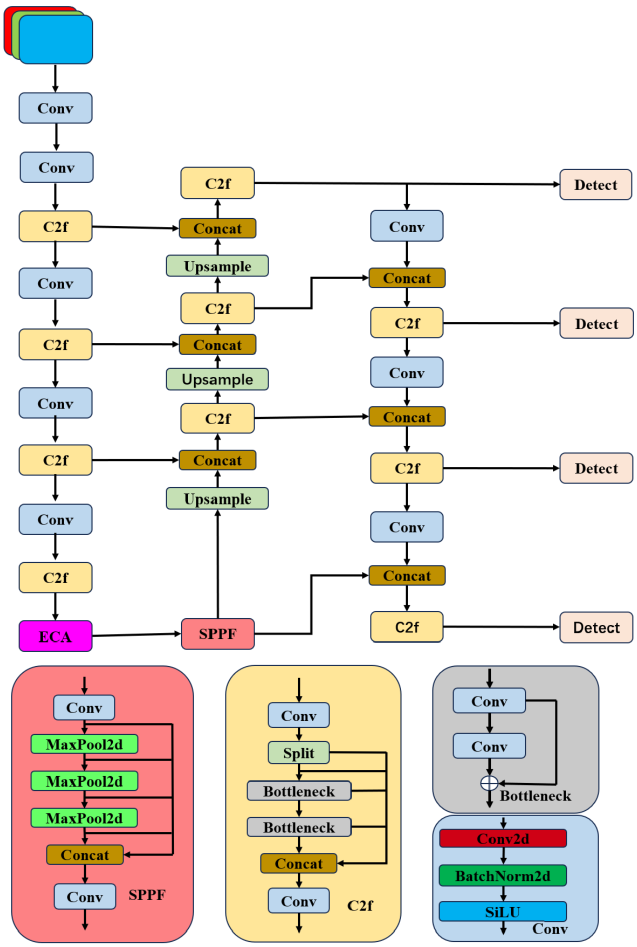


FIGURE 3. improved model with ECA and small object detection layer based on YOLOv8n.

F. HYPERPARAMETER SETTINGS

The hyperparameters are crucial for the training effectiveness of the model. Under the existing conditions, we have adjusted the hyperparameters to the best training setting. To ensure comparability of the results, we used the same hyperparameters when conducting control experiments using other attention mechanisms. In order to simulate the actual situation and address hardware limitations during the inspection process, all of our training, validation, and inference processes are run on NVIDIA GeForce RTX 2080Ti. The specific hyperparameter settings for the model are detailed in Table 1.

TABLE 1. Training hyperparameters.

Hyperparameter	Value
Epochs	100
Batch Size	16
Learning Rate	0.01
Momentum	0.937
Weight Decay	0.0005

G. IMAGE PREPROCESSING METHODS

In order to enhance the robustness of the model, we preprocessed and enhanced the images before training. YOLOv8 itself also includes the process of image enhancement. YOLOv8 adopts the Mosaic data augmentation method, which was first proposed in the YOLOv4 model [17]. The Mosaic data augmentation method randomly crops four images and then concatenates them into one image for training. The order of the first image is in the top left, the second image is in the bottom left, the third image is in the bottom right, and the fourth image is in the top right is shown in Figure 4. This enriches the background of the image and greatly enhances the robustness of the model. During training, YOLOv8n will fix the size of the image to 640 * 640. In addition, we also use the following methods to enhance the data:

Image rotation: Rotate the image at different angles to simulate different shooting angles during the inspection process.

Image scaling: Enlarge and reduce the size of the image to enhance the recognition of small targets.

Bluring: Bluring the image, considering that rain, snow, and haze may occur during the actual detection process, blurring the image can simulate this situation.

FIGURE 4. Mosaic data augmentation method.

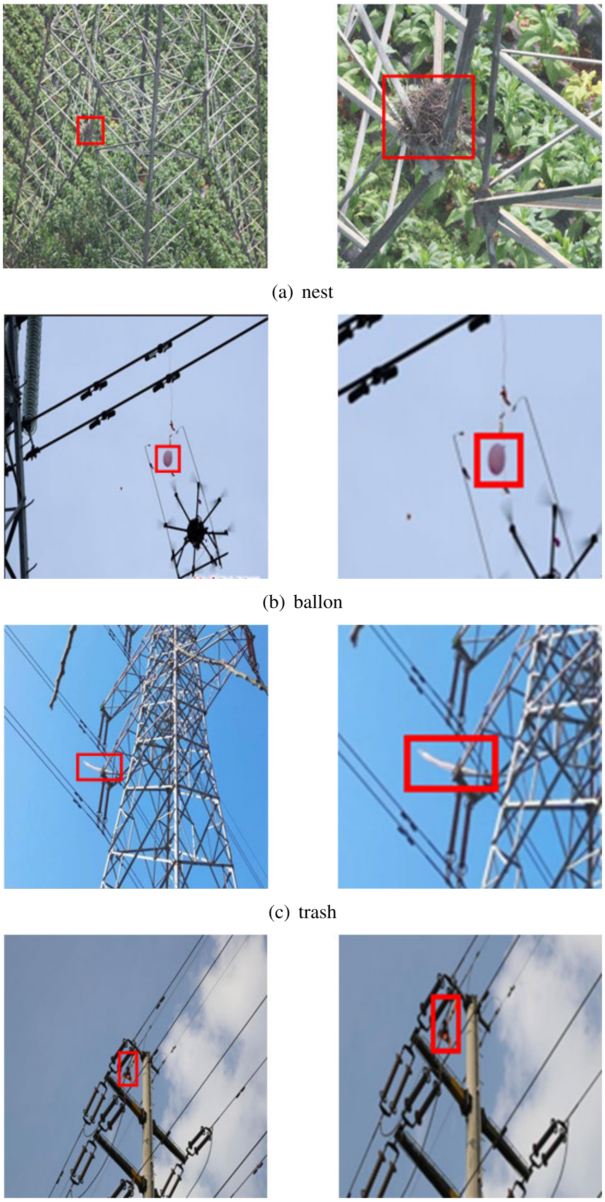
H. DATASET

Our dataset comes from our own field collection and collaboration with laboratory units. The annotated dataset is divided into four categories: Bird's Nest, Balloon, Kite, and Garbage. The dataset consists of three parts: training set, validation set, and testing set. The training set contains

3116 images. The validation set contains 610 images. The test set contains 791 images. Small targets are prone to exist among the four types, which will also be the focus of our detection. In Figure 5, we present some small target data used in our training set.

I. THE EVALUATION METHOD OF THE MODEL

We choose mAP as our metric for evaluating model accuracy, and processing time as our metric for evaluating model speed.



(a) nest
(b) ballon
(c) trash
(d) kite

FIGURE 5. small object and its enlarged image.

1) mAP

mAP (mean average precision) is a commonly used metric in object detection tasks, used to measure the accuracy of model localization and classification in different categories. Firstly, calculate the P-R (Precision Recall) curves for each category at different IoU (Intersection over Union) thresholds. Then calculate the AP (Average Precision) for each category, that is, calculate the area under the P-R curve. The average AP value for all categories is mAP. The formulas for calculating the P-R curve and AP value are showed below.

$$IoU = \frac{AreaofOverlap}{AreaofUnion}. \tag{5}$$

this is a standard for measuring the accuracy of detecting corresponding objects in a specific dataset.

$$Precision = \frac{TP}{TP + FP}. \tag{6}$$

TP represents a positive example of correct partitioning and FP represents a positive example of incorrect partitioning.

$$Recall = \frac{TP}{TP + FN}. \tag{7}$$

FN represents the counterexample of incorrect partitioning.

2) PROCESSING TIME

Processing time is an indicator used to evaluate the model's ability to detect and classify input images. Includes image preprocessing time, inference time, loss function calculation time, and post-processing time. In order to improve the efficiency of detection, we should try to reduce processing time as much as possible.

IV. RESULT

A. TRAINING COMPARISON

To verify the effectiveness of the transmission line foreign object detection model, we conducted multiple training and testing on the transmission line foreign object dataset. We trained using CBAM, ECA, GAM, and YOLOv8n models separately, and tested models with only one layer of attention module and models with multiple layers of attention modules added. Finally, compared with the baseline YOLOv8n model, the final results are shown in Table 2. Due to the crucial importance of transmission line safety, any safety accidents can have a significant impact. So in the task of detecting foreign objects in transmission lines, we would rather mistake some images without foreign objects for images with foreign objects, rather than let any image with foreign objects not be recognized. So, besides focusing on the mAP50 value, we also pay more attention to the Recall value.

The two improved models using CBAM and GAM have better accuracy than the unmodified YOLOv8n model. The mAP and Recall of the YOLOv8n + GAM model reached the highest, reaching 92.861% and 87.652%, respectively. However, the improved model using the ECA module did not demonstrate good accuracy, even worse than the baseline model. Through this comparison, we can find that the introduction of attention mechanism can enhance the accuracy of model detection, which also allows us to conduct subsequent experiments.

After testing the test set, we found that the detection accuracy was improved after using the attention mechanism,

TABLE 2. Model accuracy metrics.

Model	mAP50(%)	Precision(%)	Recall(%)	Speed(ms)	GFLOPs
YOLOv8n	92.489	94.864	84.082	28.9	8.2
YOLOv8n+CBAM	92.697	91.987	83.715	31.9	8.2
YOLOv8n+CBAM(multiple uses)	92.652	95.437	80.789	33.1	8.2
YOLOv8n+GAM	92.861	90.704	87.652	32.6	9.5
YOLOv8n+GAM(multiple uses)	93.308	91.463	88.436	36.3	13.5
YOLOv8n+ECA	92.087	92.443	82.864	29.4	8.2
YOLOv8n+ECA(multiple uses)	93.398	93.331	87.139	30.8	8.2

TABLE 3. The accuracy metrics of using attention mechanisms and SODM (small object detection modul).

Model	mAP50(%)	Precision(%)	Recall(%)	Speed(ms)	GFLOPs
YOLOv8n	92.489	94.864	84.082	28.9	8.2
YOLOv8n+SODM	93.553	95.085	88.401	30.7	12.4
YOLOv8n+CBAM+SODM	91.803	88.703	84.81	32.5	12.4
YOLOv8n+GAM+SODM	93.009	92.795	88.009	33.7	13.7
YOLOv8n+ECA+SODM	93.987	92.706	87.798	31.1	12.4

TABLE 4. The accuracy metrics of using different times of ECA module.

Model	mAP50(%)	Precision(%)	Recall(%)
YOLOv8n+ECA	92.087	92.443	82.864
YOLOv8n+ECA(multiple uses)	93.398	93.331	87.139
YOLOv8n+ECA+SODM	93.987	92.706	87.798
YOLOv8n+ECA(multiple uses)+SODM	92.002	90.1	86.153

but the detection of small targets is still not satisfactory. We have added a small object detection module for this. Considering that the small object detection module achieves different effects on different attention mechanisms, we added small object detection modules to all three attention mechanisms for experimentation. The specific results are shown in Table 3

We found that the small object detection module had a good response to the ECA attention mechanism. The model with the addition of the ECA attention mechanism showed a 1.9% improvement in mAP50 after adding the small object detection module, reaching the highest accuracy of 93.987%, far higher than the accuracy of the other two attention mechanisms and the small object detection module. Recall also reached 87.798%, ranking second.

In terms of running speed, due to the fact that the ECA attention module only has channel attention mechanism and the overall parameter quantity is small, it runs faster. The speed of YOLOv8n + ECA + small object detection module is slightly faster than the other two improved models, with a single image processing speed of 31.1ms, but still slower than the baseline model YOLOv8n.

After obtaining the above results, we consider whether using the ECA attention mechanism multiple times in the model will further improve accuracy. We conducted experiments on this. We have added ECA attention mechanism after each C2f module in the backbone network, and also added a small object detection module. With the same Hyperparameters and dataset, we obtained the results in Table 4.

Without adding a small object detection module, using the ECA attention module multiple times increased mAP50 by approximately 1.3% compared to using it only once, and the Recall value also reached 87.139%. But after adding a small object detection module, mAP50 decreased by about 1.9% and the Recall value decreased by about 1.6%.

Based on the above experiment, we ultimately chose a model consisting of YOLOv8n + ECA (one-time use) + small object detection module as our final model to complete the task of foreign object detection in transmission lines.

Our model's GFLOPs are also much lower than the model proposed by Wang et al., whose model is improved based on YOLOv8m, with a GFLOPs of 98.9 [2], which is much higher than our 12.4. This enables our model to be deployed at the edge to better complete tasks

Figure 6 and Figure 7 are some data from our final model training process, Figure 6 is the P-R Curve and Figure 7 including cls loss and dfl loss.

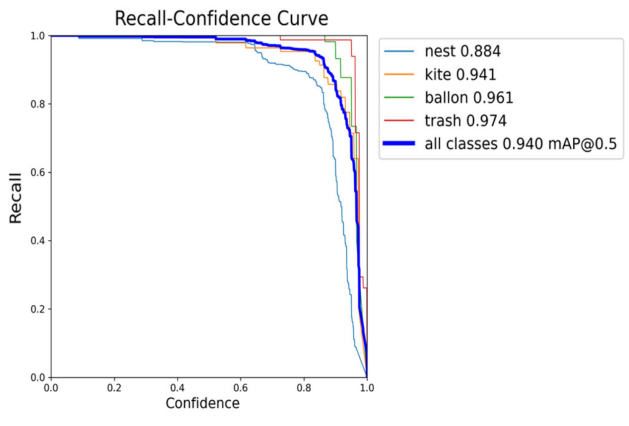


FIGURE 6. Final model's training P-R Curve.

B. COMPARISON WITH OTHER ADVANCED MODELS

As shown in the above table 5, our model has a higher mAP, which means our model has a higher detection accuracy. Moreover, while achieving higher detection accuracy,

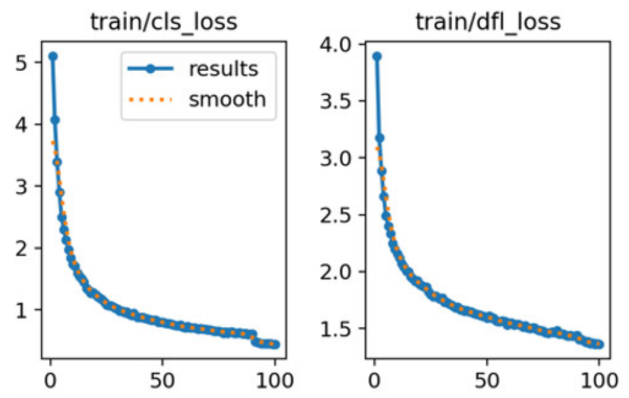


FIGURE 7. Final model's training cls loss and dfl loss.

TABLE 5. Comparison with other advanced models.

Model	mAP50(%)	Speed(ms)
Faster RCNN	87.2	61.7
YOLOv5n	92.714	26.7
YOLOv6n	90.907	30.6
YOLOv8n	92.489	28.9
ours	93.987	31.1

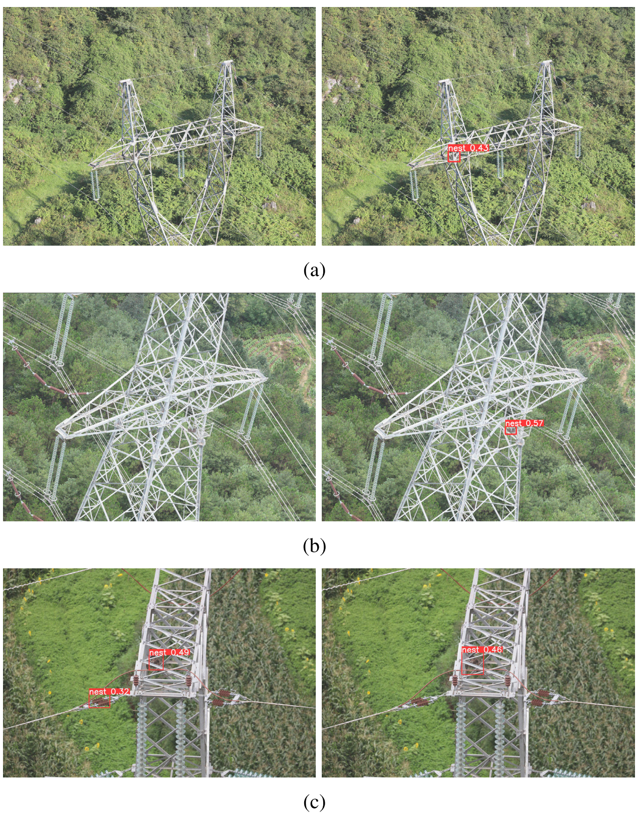


FIGURE 8. Comparison between baseline model and final model.

our model also has a faster running speed. Compared to the two-stage model, our model is much more faster. Compared to the baseline model and other one-stage models, it only increased by 2-5ms.

C. SMALL FOREIGN OBJECT DETECTION RESULT

Our model demonstrates high accuracy and fast running speed. Although the addition of attention mechanism and small object detection module resulted in longer processing time, it is still within an acceptable range. Figure 8 shows some of the results of our final model after inference. Pictures on the left are detected by the baseline model YOLOv8n. Two of them are not detected out and one of them is incorrectly detected.

Through comparison, it can be found that after our improvement, the final model is more accurate in recognizing small targets, and can recognize bird nests that the baseline model has not recognized. These bird nests are usually displayed very small in the picture due to distance. Even with the human eyes, it is difficult to recognize it immediately. In addition, through comparison, it can be found that our model also reduces the possibility of false alarms. The baseline model sometimes mistakenly identifies insulators as bird nests, while our model does not. This will also reduce the probability of false alarms during daily inspections.

V. CONCLUSION

In this article, we propose an improved YOLOv8n model to complete the task of detecting foreign objects in transmission lines. This model effectively improves the accuracy of detection and maintains a fast detection speed.

We have made two key improvements to the baseline model this time: firstly, we have introduced ECA attention mechanism into the model and added attention mechanism to each feature map of different sizes to improve the dependency relationship between channels. Secondly, a small object detection layer has been added to the detection head, enhancing the model's ability to recognize small targets and reducing the impact of shooting distance on the detection task. The improved YOLOv8n model significantly enhances its performance on this task, with improved mAP compared to the baseline YOLOv8n model. Reached. The detection speed still maintains a high level, and the inference time for each image is. Compared to the baseline YOLOv8 model, there is no significant decrease.

This work has broad application scenarios, which can greatly improve the efficiency of foreign object detection in transmission lines, and can be applied to other aspects of power inspection through transfer learning.

REFERENCES

[1] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *Proc. IEEE*, vol. 111, no. 3, pp. 257–276, Mar. 2023, doi: 10.1109/JPROC.2023.3238524.

[2] Z. Wang, G. Yuan, H. Zhou, Y. Ma, and Y. Ma, "Foreign-object detection in high-voltage transmission line based on improved YOLOv8m," *Appl. Sci.*, vol. 13, no. 23, p. 12775, Nov. 2023.

[3] D. Zhang, Z. Zhang, N. Zhao, and Z. Wang, "A lightweight modified YOLOv5 network using a Swin transformer for transmission-line foreign object detection," *Electronics*, vol. 12, no. 18, p. 3904, Sep. 2023.

[4] J. Katrasnik, F. Pernus, and B. Likar, "A survey of mobile robots for distribution power line inspection," *IEEE Trans. Power Del.*, vol. 25, no. 1, pp. 485–493, Jan. 2010, doi: 10.1109/TPWRD.2009.2035427.

[5] L. Yang, J. Fan, Y. Liu, E. Li, J. Peng, and Z. Liang, "A review on state-of-the-art power line inspection techniques," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 12, pp. 9350–9365, Dec. 2020, doi: 10.1109/TIM.2020.3031194.
[6] O. M. Butt, M. Zulqarnain, and T. M. Butt, "Recent advancement in smart grid technology: Future prospects in the electrical power network," *Ain Shams Eng. J.*, vol. 12, no. 1, pp. 687–695, Mar. 2021.
[7] L. Hernández-Callejo, "A comprehensive review of operation and control, maintenance and lifespan management, grid planning and design, and metering in smart grids," *Energies*, vol. 12, no. 9, p. 1630, Apr. 2019.
[8] W. Zhang, X. Liu, J. Yuan, L. Xu, H. Sun, J. Zhou, and X. Liu, "RCNN-based foreign object detection for securing power transmission lines (RCNN4SPTL)," *Proc. Comput. Sci.*, vol. 147, pp. 331–337, Jan. 2019.
[9] J. Hao, S. Lin, L. Wang, Z. Niu, and Y. Deng, "Research on detection methods of foreign objects and component defects in transmission lines based on deep learning," in *Proc. IEEE 5th Conf. Energy Internet Energy Syst. Integr. (EI2)*, Taiyuan, China, Oct. 2021, pp. 3070–3075, doi: 10.1109/EI252483.2021.9713226.
[10] W. Chen, Y. Li, and C. Li, "A visual detection method for foreign objects in power lines based on mask R-CNN," *Int. J. Ambient Comput. Intell.*, vol. 11, no. 1, pp. 34–47, Jan. 2020.
[11] H. Liang, C. Zuo, and W. Wei, "Detection and evaluation method of transmission line defects based on deep learning," *IEEE Access*, vol. 8, pp. 38448–38458, 2020, doi: 10.1109/ACCESS.2020.2974798.
[12] J. Zhu, Y. Guo, F. Yue, H. Yuan, A. Yang, X. Wang, and M. Rong, "A deep learning method to detect foreign objects for inspecting power transmission lines," *IEEE Access*, vol. 8, pp. 94065–94075, 2020, doi: 10.1109/ACCESS.2020.2995608.
[13] H. Li, L. Liu, J. Du, F. Jiang, F. Guo, Q. Hu, and L. Fan, "An improved YOLOv3 for foreign objects detection of transmission lines," *IEEE Access*, vol. 10, pp. 45620–45628, 2022, doi: 10.1109/ACCESS.2022.3170696.
[14] M. Wu, L. Guo, R. Chen, W. Du, J. Wang, M. Liu, X. Kong, and J. Tang, "Improved YOLOX foreign object detection algorithm for transmission lines," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–10, Oct. 2022.
[15] Z. Qiu, X. Zhu, C. Liao, W. Qu, and Y. Yu, "A lightweight YOLOv4-EDAM model for accurate and real-time detection of foreign objects suspended on power lines," *IEEE Trans. Power Del.*, vol. 38, no. 2, pp. 1329–1340, Apr. 2023, doi: 10.1109/TPWRD.2022.3213598.
[16] G. Jocher, A. Chaurasia, and J. Qiu. (2023). *Ultralytics YOLOv8*. [Online]. Available: https://github.com/ultralytics/ultralytics.2023
[17] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
[18] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 8759–8768, doi: 10.1109/CVPR.2018.00913.
[19] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 11531–11539, doi: 10.1109/CVPR42600.2020.01155.
[20] Y. Liu, Z. Shao, and N. Hoffmann, "Global attention mechanism: Retain information to enhance channel-spatial interactions," 2021, *arXiv:2112.05561*.
[21] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19. [Online]. Available: https://arxiv.org/abs/1807.06521
[22] C. Chen, M. Y. Liu, O. Tuzel, and J. Xiao, "R-CNN for small object detection," in *Proc. 13th Asian Conf. Comput. Vis. (ACCV)*, Taipei, Taiwan. Springer, 2017, pp. 214–230.
[23] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep learning for image super-resolution: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365–3387, Oct. 2021, doi: 10.1109/TPAMI.2020.2982166.
[24] J. Wang, K. Chen, R. Xu, Z. Liu, C. C. Loy, and D. Lin, "CARAFE: Content-aware ReAssembly of FEatures," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3007–3016, doi: 10.1109/ICCV.2019.00310.
[25] H. Zhang, Y. Wang, F. Dayoub, and N. Sünderhauf, "VarifocalNet: An IoU-aware dense object detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 8510–8519.
[26] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," in *Proc. AAAI Conf. Artif. Intell.*, Apr. 2020, vol. 34, no. 7, pp. 12993–13000.

HANJUN WANG received the bachelor's degree from Dalian University of Technology, in 1996. He is currently working as a Researcher with Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences. His research interests include software engineering and big data technology.

SHIYU LUO received the bachelor's degree from Nanjing University Jinling College, in 2022. He is currently pursuing the master's degree with the Department of Computer Science and Technology, Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences. His research interests include deep learning and computer vision.

QUN WANG received the bachelor's degree from Dalian University of Technology, in 2009. He is currently working as an Associate Researcher with Shenyang Institute of Computing Technology, University of Chinese Academy of Sciences. His research interests include intelligent language technology and intelligent technology for the energy industry.