

Power System to Prevent External Damage Detection Method

Jiacheng Tang, Binglin Li*

School of Mechatronic Engineering, Southwest Petroleum University, Chengdu 610500, China

* Corresponding author

Abstract: An improved YOLOv5s multi-scale target detection algorithm is proposed to address the increasingly threatening problem of foreign object intrusion in power transmission lines, which poses a risk to the safety and stable operation of the power grid. To enhance the model's ability to extract features of different sizes, a new C3 feature extraction module was designed. It uses deformable convolutional networks to replace the original fixed convolutional layers, and then introduces the ECA attention mechanism to take into account global context information, reducing the decrease in model detection accuracy caused by feature loss. An improved form of NWD fusion Wise IoU is proposed as the network's predicted box regression loss function, improving the model's generalization ability and its ability to detect features of small targets, and achieving faster network convergence speed. Experimental results show that the improved algorithm has increased the accuracy P, recall rate R, and average precision mAP by 5.21%, 4.71%, and 6.24%, respectively, compared to the original YOLOv5s model in detecting foreign object intrusion in power transmission lines, while also enhancing the detection performance of small targets.

Keywords: Transmission line; YOLOv5s; multi-scale target; C3 feature extraction module; loss function.

1. Introduction

The safety of power transmission lines is crucial to ensuring the stable operation of the power grid [1]. The manpower and resources required for manual inspections and drone inspections are extremely high. Manual inspections cannot meet the real-time inspection requirements of high-voltage transmission lines [2]. Therefore, using video surveillance equipment for 24-hour inspections, with the advantages of high efficiency and low cost, has become the preferred method for protecting high-voltage transmission lines [3]. Algorithms such as Faster R-CNN, Mask R-CNN, and YOLO based on deep learning convolutional neural networks have been used in the safety detection of transmission lines, with an accuracy rate of detecting hidden dangers reaching over 80%. While Faster R-CNN maintains good accuracy, it has slow processing speeds and long processing times, resulting in low smoothness in real-time detection. Mask R-CNN requires high precision in dataset annotation using instance segmentation methods, incurring high annotation costs and limiting its application in complex environmental conditions [4-9]. In terms of real-time detection smoothness, YOLOv3 and YOLOv4 perform well, but their accuracy is lower, leading to cases of false alarms and missed detections. YOLOv5 improves detection accuracy while inheriting good real-time performance, but it still struggles with detecting targets that are partially obscured or small [10-11]. Zhang et al [12] proposed using the ResNet network as the backbone structure of YOLOv5, which improved the model's detection performance for small objects. Li et al [13] introduced a residual network structure into SSD for feature fusion, enhancing the network's ability to extract information features and improve the detection capability of small objects. Zhang et al [14] balanced the feature network by combining channel attention mechanism with spatial attention mechanism, enhancing the detection capability of the YOLOv5 model in complex conditions. Huang et al [15] used USRNet to reconstruct the dataset for the YOLOv5x

model, then introduced small object detection head and EIou loss function, enhancing the model's performance in detecting small objects and its ability to handle complex background environments.

The size of the network model has a big impact on its performance, so this paper proposes an improved YOLOv5s algorithm to meet the requirements of real-time monitoring and small network detection. A new C3 feature extraction structure is designed, which replaces the fixed convolutional layers in C3 with a deformable convolution network DCN-V2 to enhance the ability of the C3 module to extract features of different scales. An ECA attention mechanism is introduced to process features, obtaining global information that considers context and enhances the network's non-linear capabilities [16-17]. Additionally, a new loss function form combining the NWD structure with WIou Loss is proposed to make the predicted boxes more accurate and enhance the model's ability to detect small target objects [18-19].

2. Improvement of C3 Module Structure

2.1. Deformable convolution layer

In YOLOv5, the C3 module is mainly used for extracting features from the input feature map, which is crucial in determining how many target features are extracted. However, the convolution used in the C3 module is a traditional square grid convolution, with a fixed size of the convolution kernel, and the sampling position when extracting the input feature map is also fixed, as shown in Figure 1(a).

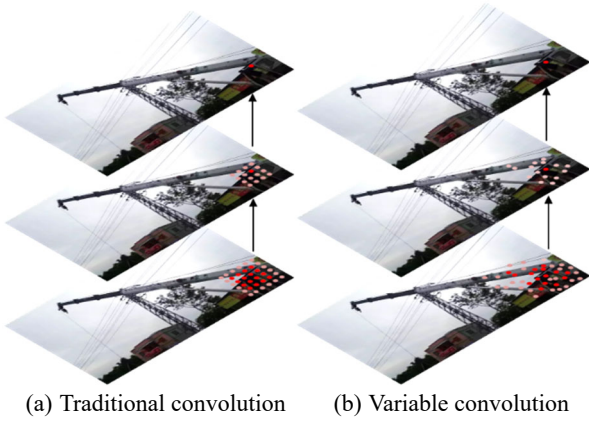


Figure 1. Convolution kernel shape

In the actual input feature map, different positions may correspond to objects of different sizes or deformations, which requires the convolution layer to adaptively adjust the range and method of the receptive field. Therefore, the method of introducing deformable convolution kernels allows the convolution kernel to adjust its shape according to the actual situation, better extracting the input features. Deformable convolution adds a variable offset matrix to each sampling point of the traditional convolution kernel, breaking away from the traditional fixed square sampling frame, allowing the convolution kernel to randomly sample near the original position, as shown in Figure 1(b), which has good adaptability to irregular features, enhancing the network's robustness and generalization ability. The degree of influence of pixels within different network nodes on the receptive field varies. Using gradient response to calculate the intensity of each image, important features are obtained, as shown in Figure 2.

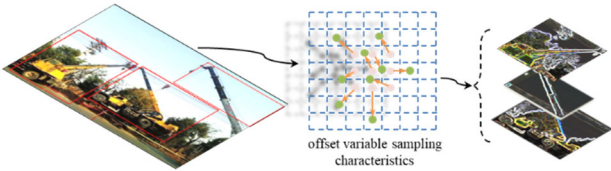


Figure 2. Deformable convolution feature extraction

2.2. Introduce attention mechanism

After feature extraction, a regular convolution operation only extracts features at the positions corresponding to the convolution kernel, with the receptive field limited to a certain area, considering only local contextual information. However, a local receptive field can lead to the loss of some features of the detected target. Therefore, in the output feature position of the C3 module, the ECA attention mechanism module is introduced to reduce the impact of irrelevant information, enabling the network to better adapt to the image content, as shown in Figure 3.

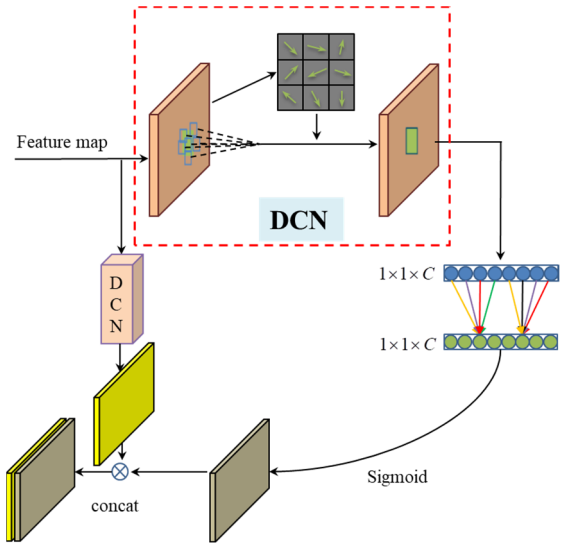


Figure 3. Introduction of ECA into C3 feature extraction structure

3. NWD Combined with WIoU Loss

In the normal IoU calculation, there is no positional information. When predicting boxes in different directions, the intersection ratio of the predicted box and the true box is the same. In the IoU calculation, it may result in the same confidence level, but it cannot determine which one is more accurate. To reduce the uncertainty caused by the loss function, YOLOv5s uses the CIoU loss function. The detection capability for some small targets is still relatively weak. For example, in overhead power line scenarios, there may be small targets such as high-altitude garbage, drones, or kites. IoU is very sensitive to positional deviations of tiny objects and often cannot accurately detect the target position. To address this issue, a method specifically for detecting small targets, NWD combined with WIoU detector, is proposed in this paper. NWD is embedded in the loss functions of label classification, non-maximum suppression, and anchor-based detectors, replacing some IoU mechanisms. By adjusting the different proportions of their combination, it balances the detection capabilities of large and small targets, achieving the most optimal model detection performance. A parameter called "rat" is set to control the proportion. The smaller the value, the greater the proportion occupied by NWD, making it more suitable for small target detection. It has strong adaptability to target sizes in different environments. The specific calculation method is as follows.

$$IoU = NWD * (1 - rat) + WIoU * rat \quad (1)$$

Therefore, this paper sets the ratio rat to 0.5 to achieve the best results. The detector schematic is shown in Figure 4.

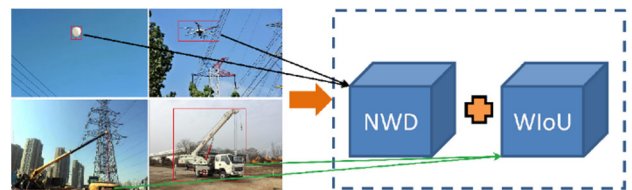


Figure 4. NWD combined with WIoU detector

4. Experimental Results and Analysis

In this article, experiments based on different attention mechanisms, comparative experiments of improved loss functions, ablation experiments, and experiments comparing with other algorithms were all conducted in the same environment settings. To verify the reliability of the proposed improved algorithm in this article, images were selected from the ones captured by cameras installed on high-voltage transmission lines and nearby power poles, depicting threats to transmission safety and other common foreign object targets on transmission lines, creating a dataset for external intrusion prevention, totaling 2500 images, with some samples shown in Figure 5. To prevent data imbalance caused by small samples, data augmentation operations were performed (first dividing the dataset, then separately augmenting the training set to avoid the impact of augmented data on test results in the test set), involving operations such as image cropping, synthesis, etc, to expand the image samples and ensure a roughly equal number of images for each category. The dataset was divided into training, testing, and validation sets in the proportion of, with all image data annotated using LabelImg, totaling 4733 images across 7 categories.



Figure 5. Transmission line foreign body invasion situation

In the experiment, to evaluate the performance of the algorithm for preventing external damage to transmission line routes, accuracy rate P , recall rate R , average precision mAP for each class, and frames per second FPS for real-time detection of images are used in the calculation formula.

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

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

$$AP = \int_0^1 PdR \quad (4)$$

TP is the number of true positive samples that the model correctly predicts as breaches, FP is the number of false positive samples that the model incorrectly predicts as breaches, FN is the number of breaches missed by the model, and the detection accuracy AP for each class represents the accuracy of breach recognition when IoU is greater than 0.5. mAP is the average detection accuracy for all classes, where N represents the number of classes in the dataset, i.e., $N=7$. The higher the mAP value, the better the model's detection performance.

$$mAP = \frac{\sum_i^N AP_i}{N} \quad (5)$$

4.1. Results and analysis

4.1.1. Analysis of ablation experimental results

To demonstrate the validity of the experiment and verify the effectiveness of the proposed three improvement strategies, we conducted ablation experiments separately. By controlling variables, we ensured the rigor of the experiment, and the results are shown in Table 1.

Table 1. Analysis of ablation experiment results

DCN	ECA	NWD+WIOU	P/%	R/%	mAP/%	FPS
			91.23	82.80	85.19	151.51
✓			93.82	85.63	88.46	147.05
	✓		92.59	83.89	86.25	144.92
		✓	93.34	84.55	87.72	138.88
✓	✓	✓	96.44	87.09	91.34	131.57

Based on the results of the ablation experiments in Table 2, it can be seen that using the original YOLOv5s model for testing experiments achieved a high FPS of 151.51, meeting the real-time monitoring requirements of power transmission lines. By adding deformable convolutional networks to extract features and introducing offset variables leading to an increase in network parameters, although this resulted in a decrease in FPS by 4.46, the model showed improvements in precision P , recall R , and mean average precision mAP by 2.59%, 2.83%, and 3.27%, respectively. Incorporating the ECA attention mechanism to consider global context information made the network structure more complex, resulting in a FPS drop of 6.59, but the model showed enhancements in P , R , mAP by 1.36%, 1.09%, 1.06%, respectively. Adopting an improved form of loss function by combining NWD with WIOU enhanced the detection capability of small targets, while also increasing network computational load, leading to a FPS decrease of 12.63. However, the model showed improvements in P , R , mAP by 2.11%, 1.75%, 2.53%, respectively. In summary, the three improvements in the algorithm each contributed to enhancing the model's performance, demonstrating the effectiveness of the enhancement strategies. Combining the three improvements is beneficial for enhancing the overall performance of the model. Compared to the original YOLOv5s model, the improved model in this study showed an overall increase of 5.21% in P , 4.71% in R , and 6.24% in mAP , with an achievable FPS of 131.57, fully meeting the real-time monitoring requirements of power transmission lines.

4.1.2. Comparative experiment of loss function

In the model training configuration, the learning rate is set to 0.01, weight decay rate is 0.0005, using the stochastic gradient descent optimizer, $SGD=0.9$, training for 200 epochs with a batch size of 32. Training was completed on the same YOLOv5s model with different loss functions, and the comparison loss curve is shown in Figure 6.

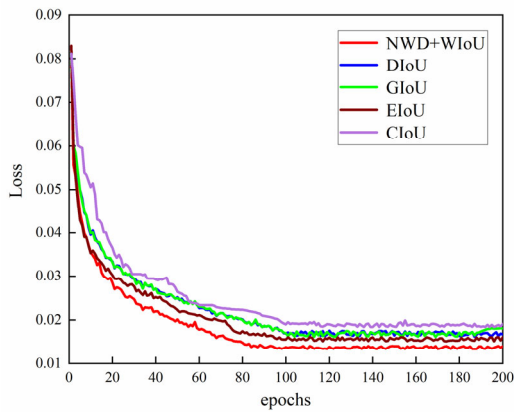


Figure 6. Comparison of different loss function curves

By observing the downward trend in Figure 7, it can be seen that the model training starts to converge around the 20th epoch. Different loss functions have varying convergence speeds, with the model already converged after 80 rounds of training, using the loss value at the 100th round as the evaluation criterion. When the NWD+WIoU mixed loss function replaces the bounding box loss function, the model's loss is smaller, and the network converges faster. From Table 3, it can be seen that when the model is trained using the mixed loss function, the loss value is only 0.01381, which is 0.00422 lower than the original YOLOv5s model using CIoU. Furthermore, by comparing with several other common loss functions, it has been proven that the improved mixed loss function enhances the detection model's performance.

Table 2. Comparison of different loss functions

Loss function	Value	mAP/%
NWD+WIoU	0.01381	87.69
DIoU loss	0.01702	86.01
GIoU loss	0.01698	85.96
EIoU loss	0.01557	86.17
CIoU loss	0.01803	85.35

4.1.3. Algorithm experimental comparison

To further validate the advantages of the improved algorithm, several commonly used object detection algorithms were trained and tested on the same dataset. Comparative experiments were conducted with the algorithm in this paper, and the experimental results are shown in Table 3.

Table 3. Comparison of different algorithm models

Model	Param/M	mAP/%	FPS
YOLOv4-tiny	9.8	76.15	122.32
SSD	102.6	70.36	23.81
Faster R-CNN	108.8	87.34	18.86
YOLOv5s	13.7	85.19	151.51
YOLOv8s	21.85	89.86	163.36
Ours	14.5	91.43	131.57

From Table 3, it can be seen that the mAP of the improved algorithm reaches 91.43%, which is 15.28% higher than YOLOv4, 21.07% higher than SSD, 4.09% higher than Faster R-CNN, 6.24% higher than the original YOLOv5s model, and 1.81% higher than the YOLOv8s model, thereby confirming the effectiveness of this improved algorithm. The YOLOv4-tiny model has the smallest number of model parameters but performs poorly in detection. Faster R-CNN has high detection accuracy but a large number of parameters, leading

to poor real-time performance. The improved YOLOv5s model is only slightly larger than the original YOLOv5s by 0.8M, but shows a significant improvement in accuracy. Compared to the latest YOLOv8s model, it maintains higher detection accuracy with lower parameter count. In terms of real-time detection performance, the FPS is only 19.96 lower than the original YOLOv5s and 31.79 lower than YOLOv8s, but it outperforms YOLOv4-tiny, SSD, and Faster R-CNN by 9.52, 107.76, and 112.71, respectively.



Figure 7. Model detection effect diagram

5. Conclusion

This text mainly addresses the shortcomings of the current online monitoring and protection system for power transmission lines, such as poor real-time performance and low precision in detecting small targets. It proposes an improved YOLOv5s algorithm to enhance the speed and accuracy of real-time monitoring of power transmission lines:

1) A new C3 feature extraction structure is proposed, replacing the fixed convolution method with a deformable convolution network and introducing an attention mechanism. This enhancement allows the improved C3 module to increase its random sampling capability, obtain more contextual information, and improve multi-scale feature extraction capability.

2) A detector NWD+WIoU that combines a mixed loss function is proposed, which accelerates the training loss reduction of the detection network, leading to higher accuracy. It can adapt to scenarios where both large and small targets are present simultaneously, enhancing the overall detection capability of the model.

3) Improvements are made to the original YOLOv5s model by adding a small number of model parameters, making the network structure more complex. However, the detection speed can still reach 131.57 frames per second, ensuring the real-time detection requirements. It effectively enhances the detection of small targets and addresses the challenge of detecting multi-scale targets invading power transmission lines.

References

- [1] ZHANG Yun, WANG Jiye, SONG Rui, et al. Research on efficient detection technology of transmission line abnormal target based on edge intelligence[J]. Electric Network Technology, 2022, 46(5): 1652-1661.
- [2] YANG Jianfeng, QIN Zhong, PANG Xiaolong, et al. Foreign body intrusion monitoring and recognition method based on Dense-YOLOv3 deep learning network[J]. Power System Protection and Control, 2021, 49(04): 37-44.

- [3] GU Xiaodong, TANG Danhong, HUANG Xiaohua. Deep learn-ing-based defect detection and recognition of a power grid inspection image[J]. *Power System Protection and Control*, 2021, 49(05): 91-97.
- [4] HUANG Yonggang, LEI Wenqian, CHEN Wei, et al. Self-powered wireless online monitoring system for electric transmission lines[J]. *Journal of Mechanical Engineering*, 2022, 58(20): 83-91.
- [5] ZHUANG Lihu, HE Tong, ZENG Yihui, et al. Fast image recognition of transmission tower based on big data[J]. *Protection and Control of Modern Power Systems*, 2018, 3(2): 149-158. DOI: 10.1186/s41601-018-0088-y.
- [6] CHEN Jiachen, YU Yaochen, CHEN Zhong, et al. An improved method for defect identification of transmission lines based on YOLOV3[J]. *Southern Power System Technology*, 2021, 15(1): 114-120.
- [7] TIAN Pengfei, YU You, DONG Ming, et al. A CNN-SVM-based fault identification method for high-voltage transmission lines[J]. *Power System Protection and Control*, 2022, 50(13): 119-125.
- [8] ZHANG Zixiao, LU Xiaoqiang, CAO Guojin, et al. ViT-YOLO: transformer-based YOLO for object detection[C] // International Conference on Computer Vision Workshops. October 11-17, 2021, Montreal, BC, Canada: 2799-2808.
- [9] WEI Xianzhe, LU Wu, ZHAO Wenbin, et al. Target detection method for external damage of a transmission line based on an improved Mask R-CNN algorithm[J]. *Power System Protection and Control*, 2021, 49(23): 155-162.
- [10] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2018-4-8]. <https://doi.org/10.48550/arXiv.1804.02767>.
- [11] REN Shaoqing, HE Kaiming, ROSS G, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39: 1137-1149.
- [12] ZHANG Baoxing, MO Yifu, PAN Qishen, et al. Instantaneous intelligence detection for the collapse of poles in distribution network based on improved YOLO-ResNet hybrid neural network[J]. *Southern Power System Technology*, 2022, 16(8): 133-141.
- [13] LI Ruisheng, ZAHNG Yanlong, ZHAI Denghui, et al. Pin Defect Detection of Transmission Line Based on Improved SSD[J]. *High Voltage Technology*, 2021, 47(11): 3795-3802.
- [14] ZHANG Huanlong, QI Qiye, ZHANG Jie, et al. Bird nest detection method for transmission lines based on improved YOLOv5[J]. *Power System Protection and Control*, 2022, 51(02): 151-159.
- [15] HUANG Yuehua, LIU Hengchong, CHEN Qing, et al. Transmission line insulator fault detection method based on USRNet and improved YOLOv5x[J]. *High Voltage Technology*, 2022, 48(09): 3437-3446.
- [16] ZHU Xizhou, HU Han, DAI Jifeng, et al. Deformable convnets V2: more deformable, better results[C] // Conference on Computer Vision and Pattern Recognition. June 15-20, 2019, Long Beach, CA, USA: 9300-9308.
- [17] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C] // Conference on Computer Vision and Pattern Recognition. June 13-19, 2020, Seattle, WA, USA: 11531-11539.
- [18] XU Chang, WANG Jinwang, YANG Wen, et al. Detecting tiny objects in aerial images: a normalized wasserstein distance and a new benchmark[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 190: 79-93.
- [19] TONG Zanjia, CHEN Yuhang, XU Zewei, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. [2023-1-24] <https://doi.org/10.48550/arXiv.2301.10051>.