# EBS-YOLO: Foreign object detection algorithm for transmission lines based on improved Yolov10

*S.X. Liu [1], S.H. Qin [2], D.Y. Jiang [3]*
*[1] School of Electrical Engineering, Shanghai Dianji University, 201306, Shanghai, China, Shuihua Road 300;*
*[2] The Key Laboratory of Cognitive Computing and Intelligent Information Processing of Fujian Education Institutions, Wuyi University, 354300, Fujian, China, Wuyi Avenue 16*

***Abstract***

When a foreign body touches a transmission line it may have serious consequences. If it is not handled in time it may lead to accidents such as short circuits and blackouts, affecting the normal operation of the power system and the stability of social life. In order to detect foreign objects on transmission lines, this paper proposes an EBS-YOLO method based on Yolov10. Firstly, in the structure of the backbone network, we adopt C2f-Efficient Multi-Scale-Conv plus (C2f-EMSCP) as the convolutional layer for feature extraction, replacing part of the C2f standard convolutional layer, and obtaining a richer feature representation by combining different scales of feature mapping. Secondly, a Bidirectional Feature Pyramid Network (BiFPN) is used, which enables the model to fuse features of different scales better. Then, the SEAMHead module with multi-head attention is utilized to augment the original features, enhance head detection, and reduce the effect of object occlusion. Finally, in order to solve the consistency problem between the predicted and real bounding boxes, the SIoU loss is used to replace the original CIoU loss in Yolov10. The experimental results show that EBS-YOLO achieves an average detection accuracy of 90.4 %, which is 4.1 % better than Yolov10, and Recall and Precision are improved by 5.3 % and 1.8 %, respectively. Compared with other methods, our EBS-YOLO has higher accuracy in detecting foreign objects on transmission lines.

Keywords: deep learning, Yolov10, transmission line foreign objects, multiscale, occlusion attention.

Citation: Liu SX, Qin SH, Jiang DY. EBS-YOLO: Foreign object detection algorithm for transmission lines based on improved Yolov10. Computer Optics 2026; 50(1): 1636. DOI: 10.18287/COJ1636.

## *Introduction*

As an important part of the power system, transmission lines undertake the critical task of power transmission. However, transmission lines are often adversely affected by foreign objects, which include, but are not limited to, a variety of airborne floating objects, nests built by birds, and so on. In order to ensure the safety and reliability of power transmission, it is necessary to effectively deal with the influence of these foreign objects on transmission lines.

With the rapid development of power grid construction in China, the scale of transmission lines has been expanding, and their layout has become more complex and dense [1, 2] . In this context, the traditional inspection work relying on manual methods has become incompetent. On the one hand, manual inspection has greater operational risks; on the other hand, due to the huge workload, the efficiency of manual inspection is difficult to meet the needs of modern power grid construction. In addition, manual inspection also faces problems such as operational difficulties under adverse weather conditions [3] .

In recent years, the application of UAV technology has brought new possibilities for the inspection of transmission lines. Drones can cover a larger area in a shorter period of time and are able to reach places that are difficult for human beings to reach. However, despite the many advantages of UAV technology, current onboard algorithms are still deficient in real-time detection capability, and the adaptability of UAVs in complex environments is still relatively limited due to the limited diversity of training data [4] .

Therefore, it is necessary to develop a more intelligent detection method, combined with the advantages of UAV inspection, to achieve the automation and intelligence of transmission line foreign object detection. This intelligent detection method not only needs to have high-precision recognition ability, but also needs to be able to achieve real-time monitoring, and maintain stable performance in complex and changing environments, so as to better meet the needs of modern power grid construction for safe and reliable power transmission.

To address these issues, this paper proposes EBS-YOLO to detect foreign objects on transmission lines by improving Yolov10. The contributions of this paper are as follows:

1. In the structure of the backbone, Efficient Multi-Scale-Conv plus (EMSCP) is used as the feature extraction convolution replacing part of the standard convolutional layer of C2f, which combines different scales of feature mapping to obtain a richer feature representation.

2. The use of a Bidirectional Feature Pyramid Network (BiFPN), which fuses features at different scales, reduces the possibility of information loss.
3. Introducing the SEAMHead network with multi-head attention enhances the original features, strengthens head detection, and reduces the effect of object occlusion.
4. The SIoU loss is used instead of the original CIoU loss in Yolov10, this improvement makes it more flexible when dealing with targets of different shapes and sizes, and effectively reduces the computational complexity.

## *1. Related work*

Traditional detection algorithms are classified into two-stage methods and one-stage methods. Two-stage methods need to generate candidate regions first before they can classify and localize the target. For example, R-CNN [5], Faster R-CNN [6], and Mask-RCNN [7] belong to this category. However, these methods have obvious drawbacks, although they have high detection accuracy the computation is too complicated and the real-time computation ability is poor, so it is difficult to use them on UAVs. In contrast, one-stage methods are a simplified target detection framework that can predict the target directly from the input image, e.g., Yolo and SSD belong to this category [8, 9]. Compared with two-stage methods, one-stage methods greatly simplify the computational complexity although the accuracy is reduced, and therefore can achieve real-time detection which is more suitable for use on UAVs.

For example, Li et al. [10] proposed an improved Yolov3 model that can run on embedded devices, using Mobilenetv2 as the backbone network, and using depth-separable convolution instead of the standard 3×3 convolution in the detection head, these improvements reduce the parameters of the network. Li et al. [11] proposed the Yolov7 model based on deformable convolution and SimAM attention mechanism, which improves the recall of model detection. Model, which improves the recall and detection accuracy of the model detection. Liu et al. [12] proposed YOLO-CSM with two attention mechanisms and extra detection layers, which makes the model more effective in capturing global information and key features, and improves the recognition ability of small objects. Yu et al. [13] used the Yolov7 algorithm as the basis and combined it with a genetic algorithm (GA) to optimize hyperparameters and space-to-depth (SPD) convolution to obtain higher detection speed and detection quality. Zheng et al. [14] proposed the GEB-YOLO model using the GhostConv network to lighten Yolov8 to reduce the number of parameters in the model. The EC2f module is proposed to use the ECA attention mechanism to compute the attention weights on the channel dimensions, this approach reduces the computational complexity and the improved model has higher efficiency and accuracy. Shao et al. [15] used Yolov8 as a framework and designed a full-dimensional dynamic convolution (ODConv) as the backbone network. The feature pyramid module is improved to fuse multi-scale features using a multi-scale attention mechanism, which enhances the feature extraction capability as well as the fusion performance during the feature fusion process.

Yolov10 used in this paper is one of the most recent algorithms in the field of target detection [16], which eliminates deficiencies in post-processing and model architecture compared to previous versions. By removing the non-maximum suppression (NMS) and optimizing the components of each model, Yolov10 reduces the computational cost while maintaining the detection accuracy. The backbone network is improved by stacking the downsampling and convolutional layers four times. An extra layer of PSA (Partial self-attention) is added to incorporate the global representation learning capability into the YOLO model while keeping the computational cost low, which in turn enhances the model performance. The fourth convolutional layer is also replaced with C2fCIB, which uses the CIB module to replace the standard convolution with deep convolution plus point-by-point convolution. The third and fourth downsampling layers are replaced with SCDown, which decomposes the traditional downsampling operation into two independent steps: firstly, the number of channels is adjusted, and then the spatial dimension is downsampled. The main advantage of this approach is that it can manage computational resources more efficiently while maintaining model performance. And this idea is applied to the Neck part by replacing it in the structure of PANet. It is done to minimize the computational cost while maximizing the information retention. Through the study, it was found that the classification head has a larger number of parameters and computations than the regression head, and the regression head has a greater impact on the loss than the classification head. Therefore, the computational share of the classification head is reduced to reduce the computational effort with as little impact as possible. The structure of the Yolov10 algorithm is shown in Fig. 1.

## *2. Proposed method*

### 2.1. EBS-YOLO algorithm

The improved EBS-YOLO structure is shown in Fig. 2. Firstly, we introduce the C2f-EMSCP module to enhance the multi-scale feature extraction capability. Subsequently, to better address the feature fusion problem, the feature pyramid is enhanced using the BiFPN structure. Next, to solve the occlusion problem a more efficient detection head module is needed, SEAMHead can solve this problem well due to its occlusion attention mechanism. Finally, the SIoU loss function is used to solve the problem of the large difference in the shape and size of the predicted and real frames, which exists in the traditional IoU. The experimental results show that the average detection accuracy of EBS-YOLO reaches 90.4 %, which is 4.1 % better than Yolov10, and Recall and Precision are improved by 5.3 % and 1.8 % respectively. Compared with other methods, our EBS-YOLO has higher accuracy in detecting foreign objects on transmission lines.

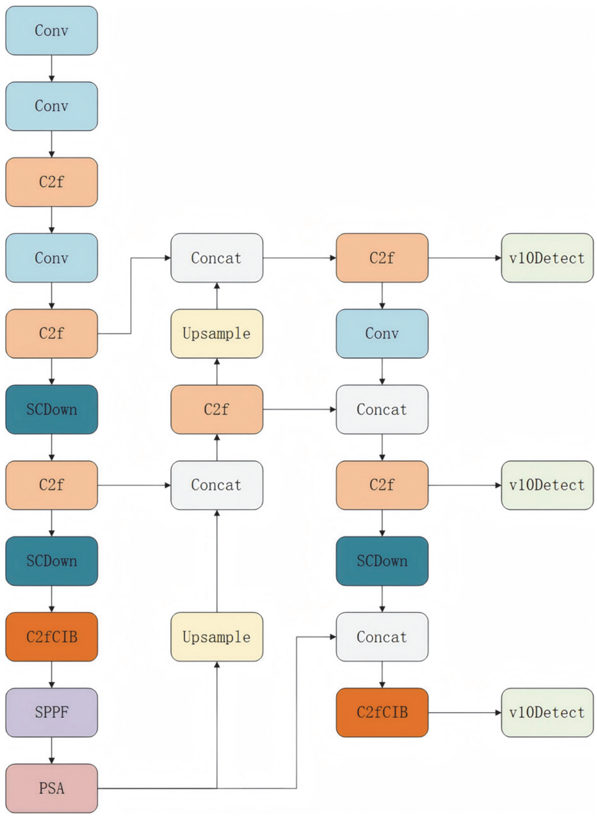


Fig. 1. Network structure of Yolov10 algorithm

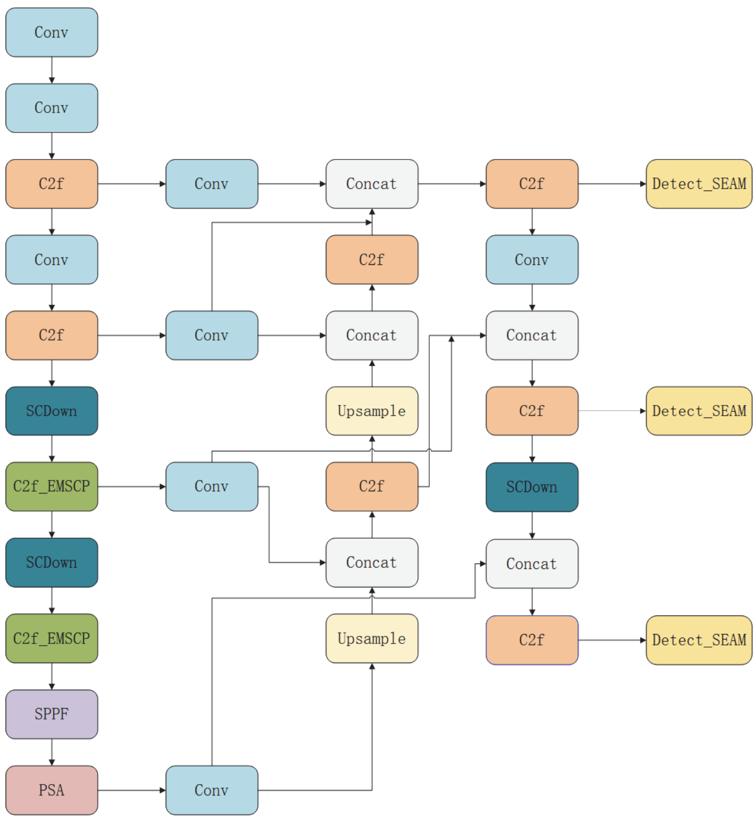


Fig. 2. Network structure diagram of EBS-YOLO

## 2.2. C2f-EMSCP module

The EMSCP structure combines the idea of multi-scale information processing with the concept of group convolution. Compared with the traditional convolution, the accuracy of target detection is improved while reducing the number of model parameters and computational complexity, resulting in a more efficient performance. Diverse features are captured by applying multi-scale convolutional operations to the feature map of the input image. Detailed information and texture features in the image are simultaneously captured along with large-scale structural and macroscopic features in the image. These features at different scales are then combined, combining the advantages of both high and low scale features. This approach enhances the model's ability to recognize and describe image details and is well suited to processing image data that contain features at multiple scales.

The structure of the EMSCP is shown in Fig. 3, where the obtained feature mappings are first reordered and divided equally into four groups. Each group passes its corresponding feature mapping to the corresponding convolutional layer, which performs 1×1, 3×3, 5×5, and 7×7 convolutional operations correspondingly. These newly obtained feature mappings are then stacked into a new feature mapping and finally reordered into a standard format. Next, feature fusion is performed by a 1x1 convolutional layer, which will perform convolution operations on the stacked feature maps to achieve information fusion at different scales.

The EMSCP structure is applied to the Bottleneck by using the EMSCP structure in the cv2 output channel of the Bottleneck, replacing the original 3×3 convolution, and the new Bottleneck structure is named Bottleneck-EMSCP. A comparison of the improved structure with the original structure is shown in Fig. 4. In the original Bottleneck structure, when ADD = TRUE, i.e., when residual concatenation is to be performed, the input is summed with the result of cv2 to get the final output. In the improved structure, the final output is obtained directly by EMSCP, which simplifies the computational complexity of the model.

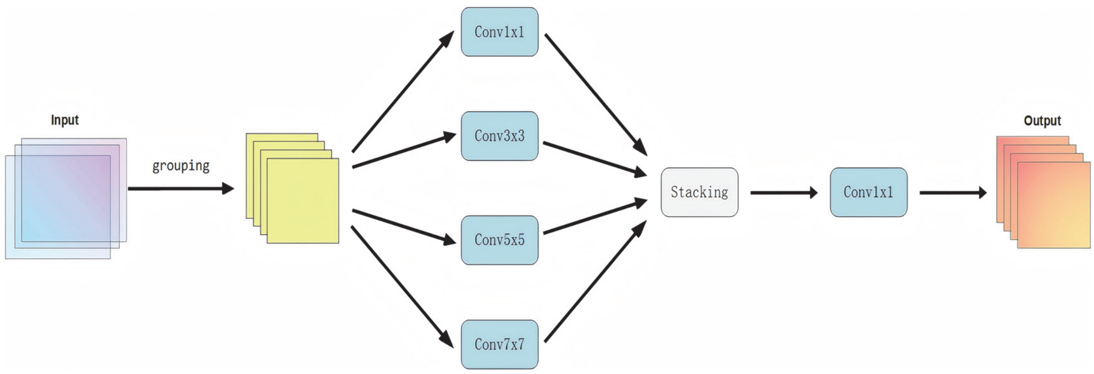


*Fig. 3. EMSCP structure*

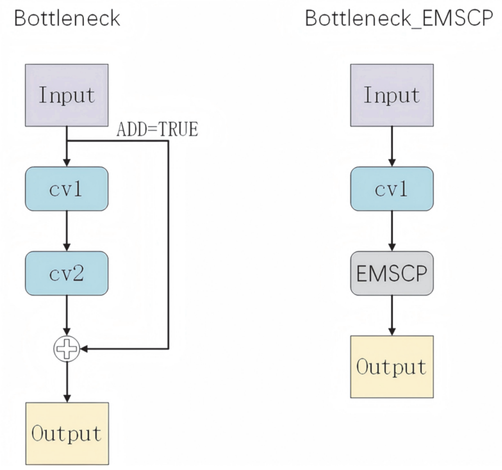


*Fig. 4. Comparison of Bottleneck and Bottleneck-EMSCP structures*

## 2.3. BiFPN

BiFPN was proposed by Tan et al. [17] in 2020 for multi-scale feature fusion, optimizing feature propagation, tuning, and selection. For better fusion of multi-scale features, we use BiFPN structure instead of PANet structure in Yolov10. Compared with it, it can achieve a more effective fusion of features while simplifying the structure of the network and reducing the cost of computation. Since BiFPN adopts a bidirectional feature fusion mechanism, the constructed feature pyramid uses information from both bottom-up and top-down directions, which achieves better multi-scale feature fusion. The difference between the structure of BiFPN and that of PANet is shown in Fig. 5.

If a single node has only a single input, there is no feature fusion and it contributes less to the overall feature network. Therefore removing unidirectional input nodes simplifies the feature network. Connecting the input nodes of a single layer to the output nodes of the same layer preserves the features of the single layer to the maximum extent. The P2 feature map with higher resolution and more details in the image is added, as small targets are easily overlooked with only fewer pixels in the image, whereas adding the P2 feature map improves the accuracy of detecting small targets in complex backgrounds.

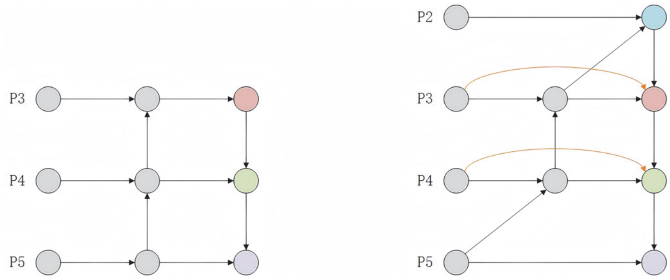


*Fig. 5. Structure comparison of PANet (left) and BiFPN (right)*

## 2.4. SEAMHead module

When the target is occluded by other objects, it can lead to problems such as category confusion, detection failure, inaccurate prediction box position, etc., and even if it can be recognized, it has a large impact on the confidence level. Therefore, the SEAMHead module with an occlusion attention mechanism added to the head region is introduced to realize multi-scale, with an attention mechanism. Being able to handle different scales and improve the quality of the features through the attention mechanism improves the accuracy and robustness of the target detection. The SEAMHead module processes the feature map through the cv2 module and the cv3 module respectively. cv2 and cv3 both contain a 3×3 convolutional layer, the SEAM attention mechanism, and a 1×1 convolutional layer. Among them, cv2 is responsible for generating the location information of the bounding box, cv3 is responsible for generating the confidence level of the corresponding bounding box, and finally, the processed information is stitched together to form the final result.

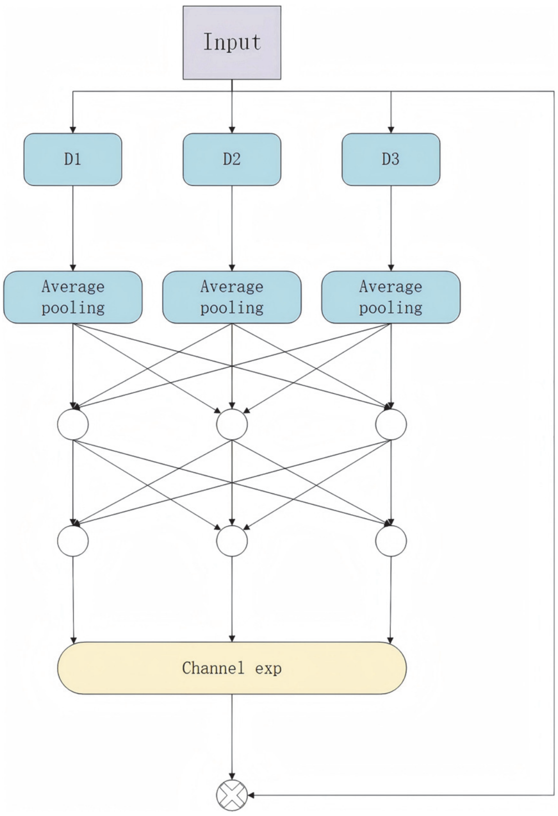


*Fig. 6. SEAM structure*

The SEAM attention mechanism, proposed by Yu et al. [18], highlights image regions to facilitate multi-scale detection, thus minimizing background interference. The input tensor x, features are extracted by depth separable convolution, the obtained outputs from different channels are pooled separately, and then the spatial dimensions are compressed into a single value, which is then unfolded into a 2D tensor. The unfolded 2D tensor is compressed and assigned and amplified weight values. Finally, the 2D tensor with weight values is multiplied element by element with the input tensor to achieve the effect of highlighting the important features and suppressing the unimportant features. The structure of the SEAM attention mechanism is shown in Fig. 6.

### 2.5. SIoU loss function

The SIoU loss function further considers the vector angle between the real frame and the predicted frame compared to the CIoU and redefines the related loss function, which consists of four parts: Angle cost, Distance cost, Shape cost, and IoU cost [19]. The definition of the SIoU loss function is shown in Eq. (1).

$$Loss_{SIoU} = 1 - \mathrm{Iou} + \frac{\Delta+\Omega}{2}. \quad (1)$$

To overcome the problem that IoU loss cannot consider the centroid offset, SIoU loss introduces Distance cost $\Delta$, which calculates the distance between the centroid of the predicted frame and the real frame to judge the proximity between the two frames. The specific formula is shown below.

$$\Delta = 2 - e^{-\lambda\rho x} - e^{-\lambda\rho y}, \quad (2)$$

where $\rho x$ denotes the regression error of the predicted bounding box centroid in the x-axis direction, $\rho y$ denotes the regression error of the predicted bounding box centroid in the y-axis direction, and $\lambda$ is a coefficient calculated by Angle cost. The basic idea is that the distance loss can guide the model to gradually reduce the distance between the predicted bounding box and the real bounding box through continuous iteration, thus improving the accuracy of target detection.

To evaluate the difference between the shape of the predicted box and the shape of the real box, the SIoU loss introduces the Shape cost $\Omega$. The specific formula is shown below.

$$\Omega = (1 - e^{-W_W})^{\theta} + (1 - e^{-W_H})^{\theta}, \quad (3)$$

Where $W_W$ and $W_H$ represent the relative errors of the width and height of the predicted box and the real box, respectively. The basic idea is that when the shape of the predicted bounding box (i.e., width and height) differs more from the shape of the real bounding box, the loss is higher; on the contrary, if the two shapes are similar, the loss is smaller. In this way, during training, the model is motivated to learn a more accurate target shape.

## 3. Experiments

### 3.1. Datasets

Considering that there is no publicly available dataset of foreign objects on transmission lines, a total of 1,701 images were obtained through online acquisition, and the dataset has a total of four categories including bird's nests, balloons, kites, and trash. The dataset has 1269 labels for the bird's nest category, 160 labels for the balloon category, 165 labels for the kite category, and 107 labels for the trash category. Since the number of balloons, kites, and trash is too small this may lead to underfitting, overfitting, or convergence failure, and a large amount of data must be passed to better train the model. We therefore use data augmentation techniques to increase the number and diversity of these three categories. Specifically, we used geometric changes such as cropping, flipping, and rotating as well as reducing the brightness of the images by decreasing the hue and saturation, and finally added different weather backgrounds to some of the images to simulate the effect of extreme weather as well as under light. Finally, a dataset of 6873 images was obtained. Fig. 7 below shows some examples of data enhancement. The classes and the exact number of classes in the enhanced dataset are shown in Tab. 1. The number of images of balloons and kites in the dataset is almost double the number of images of bird's nests, as these objects are the most common cause of foreign object accidents on power lines. Specifically, when the strings of a balloon or kite become entangled in multiple power lines, they often lead to phase-to-phase short circuits, which can cause grid outages. Given the significant safety risks associated with these objects, we prioritise data augmentation for balloons and kites to ensure that the model can effectively detect and classify these objects. To better evaluate the performance of the trained model, 80 % of the dataset is selected as the training set, and 20 % of the dataset is selected as the validation set.

*Tab. 1. Classes in the dataset and the specific numbers*

| Classes | Number |
|---|---|
| Nest | 1269 |
| Balloon | 2080 |
| Kite | 2145 |
| Trash | 1379 |

### 3.2. Parameter settings

This experiment is really run on a platform of AMD Ryzen 5 5600 CPU, NVIDIA GTX-2080TI GPU, and 16 GB RAM. The system environment is Windows 10 The deep learning framework used is pytorch, running on Python 3.9, CUDA 10.2, and cuDNN 8.2.2 libraries to improve the running efficiency. The training parameters are shown in Tab. 2.

In the experiments, a series of evaluation metrics are used to evaluate the training effect of the model. The conventional metrics mean Average Precision (mAP), Recall(R), and Precision(P) are included to reflect the performance of the model.

In addition, Params and GFLOPs are included to reflect the computational speed of the model, and the number of floating-point operations per second are given in the following formulas:

$$R = \frac{T_P}{T_P + F_N}, \tag{4}$$

$$P = \frac{T_P}{T_P + F_P}, \tag{5}$$

$$AP = \int_0^1 P(R)\, dR, \tag{6}$$

$$map = \frac{1}{N} \sum_{i=1}^{N} AP_i, \tag{7}$$

*Fig. 7. Some examples of data enhancement were used to simulate detection scenarios in several environments*

*Tab. 2. Training parameters*

| Batch Size | Epochs | Workers | Ir0 |
|---|---|---|---|
| | 100 | 4 | 0.01 |

where True Positive $T_P$ means the model correctly predicts and labels the number of targets that are present, False Positive $F_P$ means the model incorrectly predicts and labels the wrong targets, and False Negative $F_N$ means the model fails to predict the number of targets that are present. Average Precision (AP) is used to calculate the prediction precision for a single classification. The average precision mean (mAP) is used to calculate the full precision for all classifications.

### *3.3. Ablation experiments*

To evaluate the impact of each improved module in our method, a series of ablation experiments were carried out:(1) The C2f-EMSCP module was used to replace the C2f and C2fCIB modules in the original backbone to achieve multi-scale feature extraction and a reduction in the number of parameters. (2) The Neck part of Yolov10s is reconstructed using BiFPN structure, which enhances the feature fusion ability and improves the detection accuracy. (3) SEAMHead with occlusion-aware attention is utilized to replace v10Detect of Yolov10s so that it can handle occlusion scenes effectively. (4) Using SIoU to replace the loss function CIoU in the original model further improves the detection accuracy of the model. Ablation experiments results are shown in Tab. 3.

*Tab. 3. Ablation experiments in the module of the proposed methodology*

| Number | C2f-EMSCP | BiFPN | SEAMHead | SIoU | Params | GFLOPs | mAP50 |
|---|---|---|---|---|---|---|---|
| | | | | | 7.22 | 21.4 | 86.3 |
| | √ | | | | 7.82 | 21.7 | 86.7 |
| | | √ | | | 6.27 | 20.4 | 86.8 |
| | | | √ | | 8.04 | 23.8 | 87.7 |
| | √ | √ | √ | | 6.31 | 21.6 | 89.7 |
| | √ | √ | √ | √ | 6.31 | 21.6 | 90.4 |

The YLOLV10s network was chosen as the baseline model for this experiment. The second, third, and fourth experiments added C2f-EMSCP, BiFPN, and SEAMHead modules to the baseline model for the experiment, respectively. The fifth experiment introduced all the modules. The sixth experiment introduced the SIoU attention mechanism based on the fifth experiment. The ablation results show that the mAP50 value of the baseline model reaches 86.3 %, and the mAP50 values of the second, third, and fourth experiments are 86.7 %, 86.8 %, and 86.8 % respectively than the baseline model. The fifth experiment showed that the value of mAP50 for the improved model with the integration of multiple modules was 89.7 %, which is an improvement of 3.4 % in mAP value compared to the baseline model. The sixth experiment shows that after the final improvement, the EBS-YOLO algorithm achieves a mAP value of 90.4 %, which is an improvement of 4.1 % over the baseline model.

### 3.4. Comparison with other methods

To verify the performance of our proposed EBS-YOLO, we compared EBS-YOLO with other methods. The experimental results are shown in Table 4, and our EBS-YOLO outperforms the Yolov3, Yolov5, Yolov8, and Yolov10 models with higher detection accuracy and detection effect.

As shown in Tab. 4, Yolov3, as an early target detection algorithm, is not suitable for real-time target detection due to its high model complexity and large parameters, although its precision and recall are among the highest. Yolov5 and Yolov10 algorithms have low model complexity and a small number of parameters, but their precision, recall, and mAP values are relatively low. Yolov8 has a balanced performance among these algorithms, which takes into account both precision and efficiency.

*Tab. 4. Comparison with other methods*

| Model | Precision | Recall | mAP50 | mAP50:95 | Params | GFLOPs |
|---|---|---|---|---|---|---|
| Yolov3 | 90.5 | 88.9 | 89.0 | 64.4 | 103.7 | 282.6 |
| Yolov5 | 89.4 | 83.2 | 86.2 | 56.6 | 9.1 | 24.0 |
| Yolov8 | 87.8 | 87.8 | 88.6 | 61.2 | 11.2 | 28.4 |
| Yolov10 | 88.1 | 82.6 | 86.3 | 57.0 | 7.3 | 21.6 |
| EBS-YOLO | 89.9 | 87.9 | 90.4 | 62.4 | 6.3 | 21.6 |

Comparing with all the models mentioned above, it can be seen that our EBS-YOLO is at the forefront in almost all the metrics compared with other methods. Among them, the computational efficiency is the highest among the above models, which is because the multi-scale feature extraction convolution greatly improves the feature extraction capability of the backbone network and increases the computational efficiency. Meanwhile, the BiFPN structure not only improves the accuracy of the model but also greatly reduces the computational volume of the model further improving the computational efficiency. In addition, the use of occlusion-aware attention allows the model to better detect foreign objects located on transmission line towers, which greatly improves the computational accuracy. Finally, the use of a smoothed loss function to calculate the loss makes the computation more concise and reduces unnecessary computational steps to improve computational efficiency.

### 3.5. Visualization of detection results

We also visualize and analyze the results of our proposed EBS-YOLO with those of other methods. From the first column of Fig. 8 we can see that in the complex background, all other methods incorrectly detect the non-existing bird nests only our EBS-YOLO detects only the only real existing bird nests. Also in the second column, only our EBS-YOLO detected the kite located on top of the transmission line. And Yolov3, Yolov5, Yolov8, and Yolov10 all have missed detection problems. In the third column Yolov3, Yolov5, and Yolov8 all incorrectly identified the UAVs as trash, and Yolov3 and Yolov10 also had double detection problems. The experiments show that only our EBS-YOLO can accurately detect foreign objects and solve the above problems of repeated detection, wrong detection, and missed detection.

## Conclusion

In this paper, a transmission line foreign object detection algorithm EBS-YOLO based on Yolov10 is proposed. The algorithm The algorithm captures the feature information of foreign objects at different scales and is more sensitive to the features of small targets. The trunk and head of Yolov10 are enhanced to solve the problem of difficult recognition of occluded foreign objects. First, the C2f-EMSCP module is introduced into the backbone to enhance the model's ability to obtain richer features; second, the BiFPN structure is used to improve the model's ability to fuse features at different scales; then, the SEAM attention mechanism is combined with the detection head to reduce the effect of object occlusion; finally, the SIoU loss function is used instead of the CIoU loss function of Yolov10 which makes it more stable for targets at different scales and reduces the computational effort. Based on these improvements, comparison and ablation experiments with existing algorithms are conducted on the dataset. The experimental results show that EBS-YOLO effectively solves the problem that transmission line foreign objects occupy few pixels in UAV aerial images and are difficult to identify in an occluded background. Future research work is dedicated to improving the recognition accuracy of the model while keeping it lightweight, and further applying the proposed model to real detection scenarios for real-time and effectiveness. This not only helps to improve the practical applicability of the model but also provides reliable technical support for real-time detection.

## References

[1] Zhibin Qiu, Xuan Zhu, Caibo Liao, Wenqian Qu, and Yanzhen Yu. A lightweight yolov4-edam model for accurate and real-time detection of foreign objects suspended on power lines. *IEEE Transactions on Power Delivery*, 38(2):1329–1340, 2022. DOI: 10.1109/TPWRD.2022.3213598.

[2] Paolo Sospiro, Lohith Amarnath, Vincenzo Di Nardo, Giacomo Talluri, and Foad H Gandoman. Smart Grid in China, EU, and the US: State of Implementation. *Energies*, 14(18):5637, 2021. DOI: 10.3390/en14185637.

[3] Minghu Wu, Leming Guo, Rui Chen, Wanyin Du, Juan Wang, Min Liu, Xiangbin Kong, and Jing Tang. Improved YOLOX foreign object detection algorithm for transmission lines. *Wireless Communications and Mobile Computing*, 2022(1):5835693, 2022. DOI: 10.2139/ssrn.4800586.

[4] Jinguo Zhu, Yue Guo, Fanding Yue, Huan Yuan, Aijun Yang, Xiaohua Wang, and Mingzhe Rong. A deep learning method to detect foreign objects for inspecting power transmission lines. *IEEE Access*, 8:94065–94075, 2020. DOI: 10.1109/ACCESS.2020.2995608.

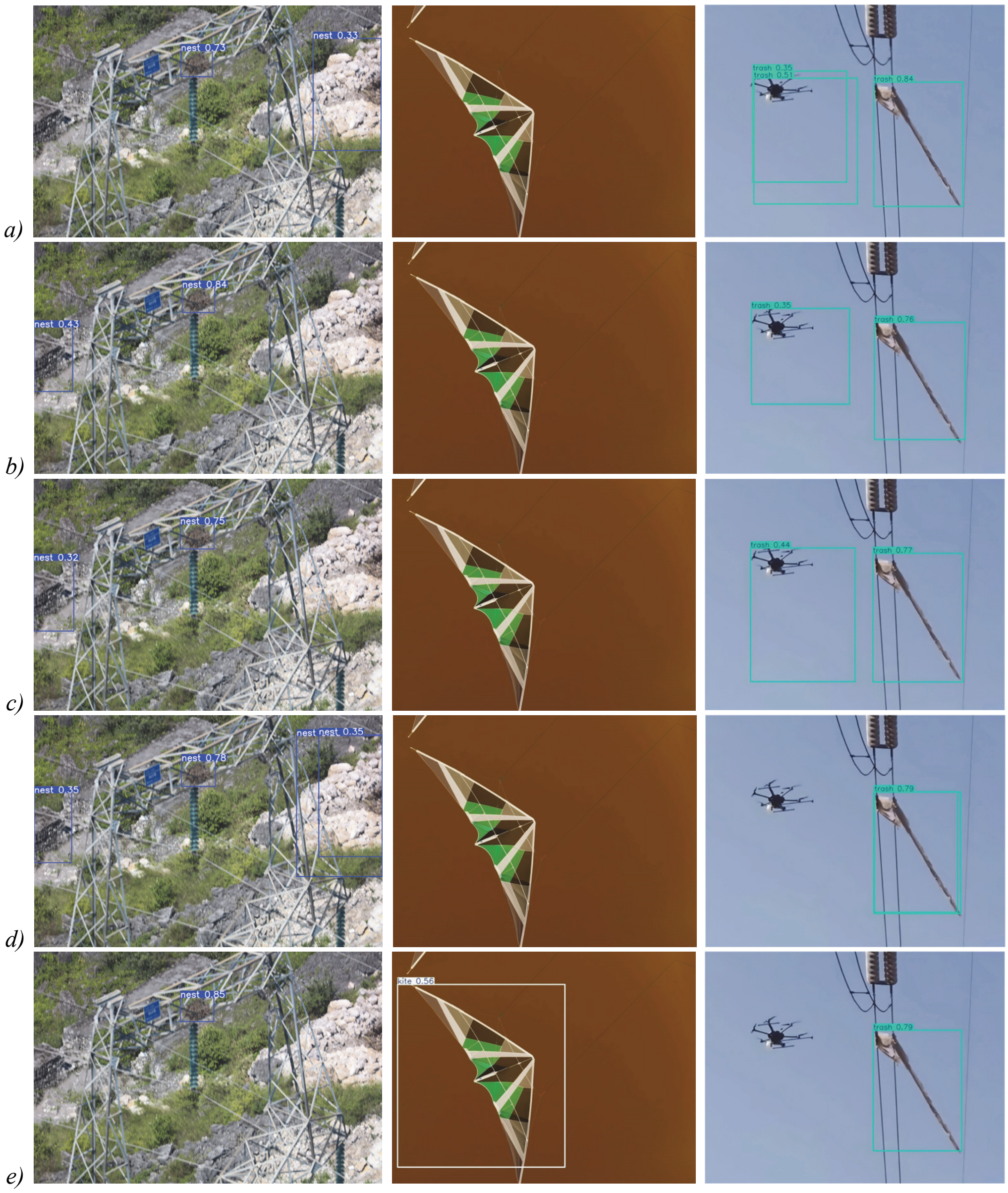


*Fig. 8. Visual analysis of the results with other methods. Compared with Yolov3, Yolov5, Yolov8, and Yolov10, our EBS-YOLO solves the problems of duplicate detection, false detection, and missed detection. a) Yolov3, b) Yolov5, c) Yolov8, d) Yolov10, e) EBS-YOLO*

[5] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2014. DOI: 10.18127/j00338486-202109-11.

[6] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1137–1149, 2016. DOI: 10.1109/TPAMI.2016.2577031.

[7] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2961–2969, 2017. DOI: 10.7717/peerj-cs.1865.

[8] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. SSD: Single shot multibox detector. In *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 21–37, 2016. DOI: 10.57152/malcom.v4i3.1332.

[9] X Han, J Chang, and KJPCS Wang. You only look once: Unified, real-time object detection. *Procedia Computer Science*, 183(1):61–72, 2021. DOI: 10.1109/CVPR.2016.91.

[10] Hui Li, Lizong Liu, Jun Du, Fan Jiang, Fei Guo, Qilong Hu, and Lin Fan. An improved YOLOv3 for foreign objects detection of transmission lines. *IEEE Access*, 10:45620–45628, 2022. DOI: 10.1109/ACCESS.2022.3170696.
[11] Shao Jia Li, Yan Xia Liu, Miao Li, and Lu Ding. DF-YOLO: Highly accurate transmission line foreign object detection algorithm. *IEEE Access*, 11:108398–108406, 2023. DOI: 10.1109/ACCESS.2023.3321385.
[12] Chunyang Liu, Lin Ma, Xin Sui, Nan Guo, Fang Yang, Xiaokang Yang, Yan Huang, and Xiao Wang. YOLO-CSM-Based Component Defect and Foreign Object Detection in Overhead Transmission Lines. *Electronics*, 13(1):123, 2023. DOI: 10.3390/electronics13010123.
[13] Chenhui Yu, Yakui Liu, Wanru Zhang, Xue Zhang, Yuhan Zhang, and Xing Jiang. Foreign objects identification of transmission line based on improved YOLOv7. *IEEE Access*, 11:51997–52008, 2023. DOI: 10.1109/ACCESS.2023.3277954.
[14] Jiangpeng Zheng, Hao Liu, Qiuting He, and Jinfu Hu. GEB-YOLO: a novel algorithm for enhanced and efficient detection of foreign objects in power transmission lines. *Scientific Reports*, 14(1):15769, 2024. DOI: 10.1038/s41598-024-64991-9.
[15] Yeqin Shao, Ruowei Zhang, Chang Lv, Zexing Luo, and Meiqin Che. TL-YOLO: Foreign-Object Detection on Power Transmission Line Based on Improved Yolov8. *Electronics*, 13(8):1543, 2024. DOI: 10.3390/electronics13081543.
[16] Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, and Guiguang Ding. Yolov10: Real-time end-to-end object detection. *arXiv preprint arXiv:2405.14458*, 2024. DOI: 10.1007/s11554-022-01233-z.
[17] Mingxing Tan, Ruoming Pang, and Quoc V Le. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10781–10790, 2020. DOI: 10.1109/CVPR42600.2020.01079.
[18] Ziping Yu, Hongbo Huang, Weijun Chen, Yongxin Su, Yahui Liu, and Xiuying Wang. Yolo-facev2: A scale and occlusion aware face detector. *Pattern Recognition*, 155:110714, 2024. DOI: 10.1109/cvpr42600.2020.01079 .
[19] Zhora Gevorgyan. SIoU loss: More powerful learning for bounding box regression. *arXiv preprint arXiv:2205.12740*, 2022. DOI: 10.1016/j.neunet.2023.11.041.

## *Authors' information*

**Shuxin Liu** (b. 1977) is currently an Associate Professor at Shanghai Dianji University. He received his PhD from East China Normal University in September 2012, after pursuing his doctoral degree there from 2005 to 2012. Prior to that, he studied for his master's degree at Donghua University of Technology from 2003 to 2005. From 2006 to 2020, Dr. Liu worked at Minnan Normal University, starting as a teaching assistant and later serving as a lecturer and associate professor. Since December 2020, he has taken up the position of Associate Professor at Shanghai University of Electric Power, where he also serves as a graduate student supervisor. For over ten years, Dr. Liu's research has focused on key technologies in the fields of computer vision and artificial intelligence. He has undertaken and participated in more than ten science and technology projects sponsored by national, provincial and municipal funds. E-mail: *liusx@sdju.edu.cn*

**Shuhan Qin** (b. 2000) is currently pursuing a Master's degree in Electrical Engineering at Shanghai Dianji University. His research interests are computer vision and object detection of foreign objects in transmission lines.
E-mail: *knyr9u@163.com*

**Dengyin Jiang** is a teacher at Shanghai Dianji University. His main research direction is automation.
E-mail: *jiangdy@sdju.edu.cn*

*Received October 29, 2024. The final version – August 30, 2025.*