2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

DAICW: Defect Detection Algorithm for High-voltage Transmission Lines in Complex Weather

Haofeng Li[1,2,+], Zhiqing Guo[1,2,+], Liejun Wang[1,2], Yongming Li[1,2(✉)]

[1] College of Computer Science and Technology, Xinjiang University, Urumqi, China
[2] Xinjiang Multimodal Intelligent Processing and Information Security Engineering Technology Research Center, Urumqi, China

Abstract. High-voltage transmission lines are continuously exposed to outdoor environments, where harsh natural conditions can lead to structural damage. Addressing the challenge of defect detection under complex weather conditions, we propose the Detection Algorithm in Complex Weather (DAICW). Firstly, we introduce the Detail-Enhanced Convolution (DEConv), designed to extract richer features without increasing the parameters. Subsequently, a Focus-Detect is incorporated to emphasize defect features within images while suppressing background interference. Finally, using the LInner-IoU loss function can effectively accelerate convergence and improve the model's ability to detect small objects. Comparison experiments with other mainstream models show that the proposed model achieves 78.2% precision and 67.1% recall, demonstrating strong adaptability in complex weather and multi-class defect detection tasks.

Keywords: High-voltage Transmission Line, Complex Weather, Defect Detection, Feature Enhancement.

1 Introduction

High-voltage transmission lines are critical components of the power system, serving as essential conduits for delivering electrical energy from generation stations to end-users [1]. Their secure and stable operation is vital for national economic development and the normal functioning of society [2]. However, these lines often traverse regions with complex terrains and varying environmental conditions, making them susceptible to natural disasters, equipment aging, and external damages. Consequently, defects such as cable strand disconnect, insulator damage, and cable foreign object may occur, as illustrated in Fig. 1. If not promptly identified and addressed, these defects can lead to line faults or even widespread power outages, resulting in significant economic losses and social disruptions. Therefore, efficient and accurate detection of high-voltage transmission line defects is imperative to ensure the stable operation of the power system.

In recent years, the rapid advancement of deep learning techniques has led to groundbreaking progress in image recognition and object detection across various domains. Deep learning models, particularly convolutional neural networks (CNNs), excel at ex-

(a) (b) (c) (d) (e) (f)

Fig. 1. (a)insulator damage, (b)grading ring tilt, (c)grading ring abscision, (d)cable strand disconnect, (e)cable loose strand, (f)cable foreign object.

tracting multi-level features from images, achieving high detection accuracy even in complex scenarios [3]. The integration of unmanned aerial vehicles (UAVs) with deep learning algorithms for high-voltage transmission line defect detection not only effectively addresses the issues of high labor costs and safety risks associated with traditional manual inspections but also significantly improves inspection efficiency [4].

High-voltage transmission line defect detection faces numerous challenges, including the presence of small-scale defects and complex weather conditions, which often result in suboptimal model performance. Additionally, existing deep learning models have high FLOPs and demand substantial hardware resources, making direct deployment on resource-constrained edge devices difficult [5]. These issues hinder the widespread application of deep learning techniques in practical inspection tasks. Therefore, developing a lightweight, high-precision deep learning model capable of adapting to complex weather conditions has become a focal point in the field of high-voltage transmission line defect detection. To address these challenges, we have designed the Detection Algorithm in Complex Weather (DAICW). The main contributions of this paper are as follows:

(1) The designed Detail-Enhanced Convolution (DEConv) can extract richer features while maintaining the original parameters.

(2) The designed Focus-Detect head effectively facilitates multi-scale defect detection, highlighting defect features in the image and reducing interference from background regions.

(3) Through a comparative analysis of different loss functions, the advantages of the LInner-IoU loss function in bounding box regression were validated, providing data support for further enhancing the of performance of model.

(4) A dataset for high-voltage transmission line defect detection in complex scenarios was constructed, with defect categories including: insulator damage, grading ring tilt, grading ring abscision, cable strand disconnect, cable loose strand, cable foreign object.

2 Related Work

In recent years, the YOLO series of algorithms have been widely applied in the field of high-voltage transmission line defect detection due to their high precision and rapid detection capabilities [6,7,8]. Images of transmission lines captured under varying weather conditions exhibit significant differences in brightness, color, and texture features, leading to reduced contrast between defects and the background, thereby increasing the difficulty of detection. To address the challenges posed by complex background

interference and small target defects, researchers have proposed various improved algorithms.

2.1 Defect Detection of High-voltage Transmission Lines

To address the challenge of complex background interference, numerous researchers have integrated attention mechanisms into deep learning networks to enhance the model's focus on defect features. Wang et al. [9] proposed the multiscale channel information (MCI)-global-local attention (GLA), a plug-in designed for YOLO series models. GLA captures both global context information and local spatial details, thereby enhancing the network's learning capabilities. Ding et al. [10] proposed GC-YOLO, which incorporates coordinate attention at the end of the YOLOv5 backbone network. This improvement strengthens the network's capability to localize and identify targets within complex scenes, effectively reducing the degradation of detection accuracy induced by complex backgrounds. Lin et al. [11] proposed YOLO-DA, a detection algorithm tailored for remote sensing object detection. By integrating the Convolutional Block Attention Module (CBAM) at the terminal stage of the YOLO-DA detector, the model effectively addresses challenges posed by complex scenes in optical remote sensing images. In addition, Ji et al. [12] proposed FusionNet based on an algorithm for cable foreign object detection in bad weather, where coordinate attention in the network significantly improves the detection precision. While these methods mitigate the negative effects of complex image backgrounds to some extent, they have significant limitations. Effectively extracting and learning features from the feature map is the key to solving the multi-scale problem. The correlation and fusion of contextual information between target and background as well as global and local information are crucial for weakening background features and enhancing defective region features.

2.2 Small Object Detection

In small object detection tasks, targets typically occupy minimal pixel areas and often appear in dense distributions. Specifically, in high-voltage transmission line defect detection, certain defects such as wire strand breakage and loosened strands manifest as small-scale anomalies. To address these challenges, various methodologies have been proposed to enhance the model's capability in identifying small targets. These approaches encompass super-resolution techniques, loss function optimization, feature fusion, and multi-scale feature learning.

To overcome these issues, Jain et al. [13] proposed a generalized model for identifying or classifying different small-size component types on the transmission lines. Mao et al. [14] employed dilated convolutions to expand the receptive field, effectively enhancing the detection of small objects, achieving an Average Precision (AP) of 92% on a transmission line fasteners dataset. Wang et al. [15] introduced an additional detection head specifically designed for small objects into the head network, significantly improving detection accuracy. Peng et al. [16] developed a Multi-scale Path Aggregation Network (M-PANet) that seamlessly integrates multi-scale information and incorporates four detection heads, effectively reducing the recall of small objects. Similarly, Zhang et al. [17] designed a specialized detection head for small objects to enhance the sensitivity of object recognition.

3 Methods

Integrating DEConv into the backbone network facilitates the extraction of richer features without increasing the number of parameters, thereby avoiding additional computational cost and memory burden during the inference phase， as illustrated in Fig. 2. Additionally, the incorporation of the Focus-Detect effectively promotes multi-scale defect detection by emphasizing defect features within images.

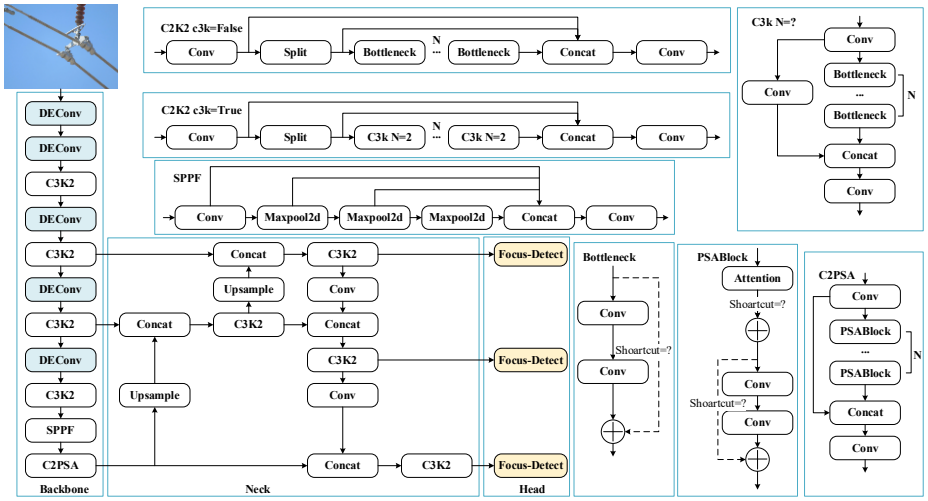


Fig. 2. Overall structure of DAICW.

3.1 Detail-Enhanced Convolution

In the field of image denoising, traditional methods typically use vanilla convolution layers for feature extraction and learning. The convolution kernels in vanilla convolutions (VC) search the vast solution space without any constrains (even initialize randomly), which limits their expressiveness and modeling capacity. From the perspective of image denoising, the presence of fog, snowflakes, and sandstorms causes changes in brightness and color, which are low-frequency variations. Additionally, natural scenes in adverse weather conditions may lose high-frequency details. Low-frequency information is crucial for correcting lighting and color distribution, while high-frequency information is essential for recovering missing fine textures. However, vanilla convolution tends to prioritize low-frequency information while neglecting high-frequency components. Previously, some researchers incorporated edge prior knowledge in denoising models to help recover clearer contours. Inspired by this, we designed DEConv, as shown in Fig. 3, by combining differential convolution with vanilla convolution in parallel to form DEConv.

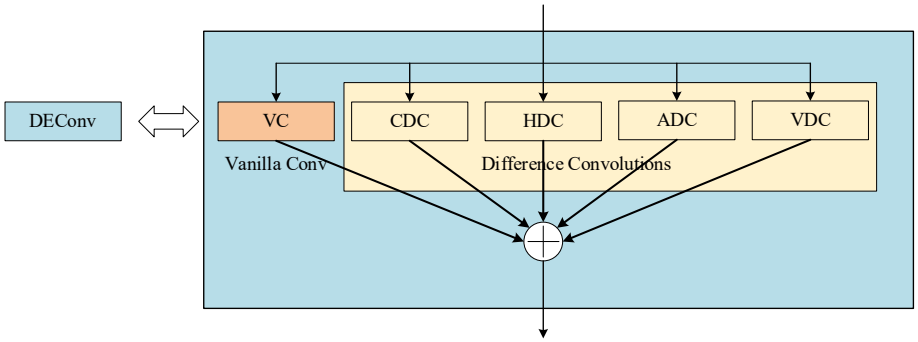


Fig. 3. Detail-Enhanced Convolution. It consists of five convolution layers deployed in parallel: vanilla convolution (VC), central difference convolution (CDC), angular difference convolution (ADC), horizontal difference convolution (HDC), and vertical difference convolution (VDC).

To the best of our knowledge, this is the first time that we introduce Difference Convolution (DC) to address issues such as fog, snow, and sandstorms in defect detection scenarios. In our implementation, we use four DCs and one vanilla convolution, deployed in parallel for feature extraction. In DCs, a pixel-pair differential computation strategy is designed to encode prior information into the CNN. Taking HDC as an example, the horizontal gradient is calculated by computing the difference between the selected pixel pairs, as shown in Fig. 4. After training, the learned convolution kernel weights are rearranged and directly convolved with the input features, with the constraint that the sum of the horizontal weights equals zero. VDC follows a similar derivation, replacing the horizontal gradient with the corresponding vertical gradient. Both HDC and VDC explicitly encode gradient priors into the convolution layers, enhancing representation and generalization capabilities by learning useful gradient information.

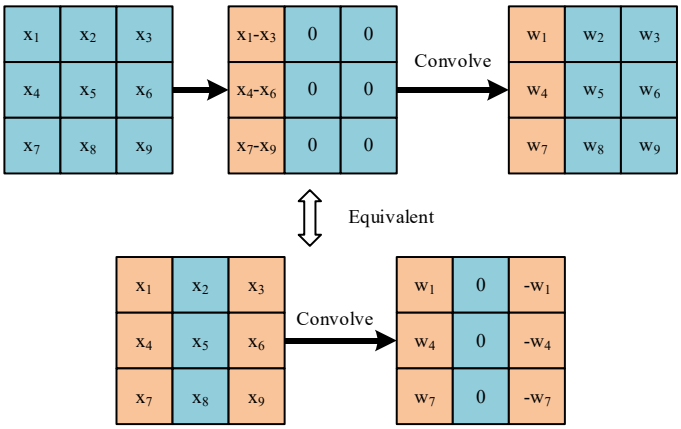


Fig. 4. The derivation of horizontal difference convolution (HDC).

In our design, vanilla convolution is used to obtain intensity-level information, while DC is employed to enhance gradient-level information. By combining the learned features, we obtain the output of DEConv. However, deploying five parallel convolution layers for feature extraction leads to an increase in both parameters and inference time. We attempt to leverage the additivity property of convolution layers to simplify the parallel convolutions into a single standard convolution, and the reparameterization technique perfectly satisfies this requirement. If several 2D kernels of the same size produce outputs for the same input with the same stride and padding, we can add their outputs together to obtain the final result. In this case, we can sum the kernels at the corresponding positions to obtain an equivalent kernel that produces the same final output. Surprisingly, the DEConv we designed fully conforms to this situation. Given the input feature F_{in}, the reparameterization technique is used to obtain F_{out}. The formula for the reparameterization technique is as follows:

$$F_{out} = DEConv(F_{in}) = \sum_{i=1}^{5} F_{in} * K_i = F_{in} * \left(\sum_{i=1}^{5} K_i \right) = F_{in} * K_{cvt} \quad (1)$$

where $DEConv(\cdot)$ denotes the DEConv operation, $K_{i=1:5}$ represents the kernel of VC, CDC, ADC, HDC, and VDC, respectively, * denotes the convolution operation, and K_{cvt} denotes the converted convolution kernel.

3.2 Focus-Detect

This paper presents an in-depth study of the detection head. To address issues such as occlusion and significant variations in defect scale in high-voltage transmission line inspection, a Focus-Detect is designed, as shown in Fig. 5. The design aims to minimize parameter and computational requirements while reducing accuracy loss as much as possible. The Focus-Detect effectively facilitates multi-scale defect detection, highlighting defect features in the image.

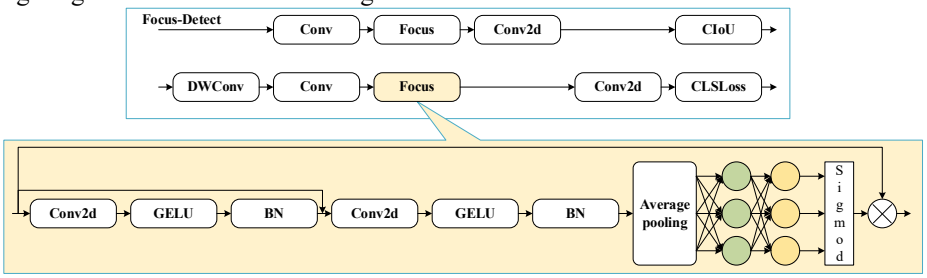


Fig. 5. Focus-Detect

Specifically, the design of the object detection head network structure optimizes the feature extraction and model training processes through a series of convolutions combined with activation functions, along with batch normalization. The first part of the category detection head consists of depthwise separable convolutions with residual connections. Depthwise separable convolution operates in a per-depth manner, meaning it performs convolution separately for each channel, allowing it to learn the importance of different channels and reduce the number of parameters. The most critical component of the Focus-Detect is the Focus module, which is simultaneously applied in both

the IoU loss and category loss functions. This strengthens the information representation capability of the defect region and enhances the model's detection performance in complex backgrounds. Within the Focus module, the basic unit Conv2d-GELU-BN is used twice with residual connections to enhance high-level features. The network then utilizes global average pooling for feature compression. Finally, two fully connected layers are applied.

3.3 Loss Function

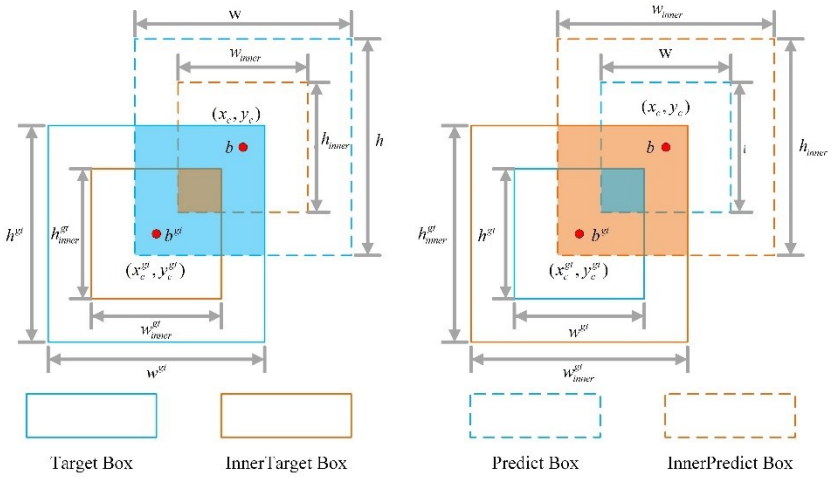


Fig. 6. IoU^{Inner}

Loss functions such as LGIoU [18], LEIoU [19], and LCIoU [20] can effectively accelerate convergence and improve detection performance. However, they do not consider the rationality of the IoU loss itself, which largely determines the quality of the detection results. To address this limitation, this chapter proposes the use of the LInner-EIoU [21] loss function for bounding box regression, which offers higher efficiency and accuracy in bounding box regression. In IoU^{Inner}, a scaling factor, denoted as the ratio, is introduced to control the proportion size of the auxiliary bounding boxes. As shown in Fig. 6, by using auxiliary bounding boxes of different scales on different datasets, the limitations of weak generalization in existing methods can be overcome. The ground truth box and predicted box are denoted as B and B_{gt}, respectively, with the center points of the ground truth box and inner ground truth box represented as (x_c^{gt}, y_c^{gt}), and the center points of the predicted box and inner predicted box as (x_c, y_c). The width and height of the ground truth box are denoted as w_{gt} and h_{gt}, while the width and height of the predicted box are denoted as w and h. The variable "ratio" corresponds to the scaling factor, with values typically in the range of [0.5, 1.5]. The calculation formula for IoU^{Inner} is as follows:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} \times ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} \times ratio}{2} \tag{2}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} \times ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} \times ratio}{2} \tag{3}$$

$$b_l = x_c - \frac{w \times ratio}{2}, b_r = x_c - \frac{w \times ratio}{2} \tag{4}$$

$$b_t = y_c - \frac{h \times ratio}{2}, b_b = x_c + \frac{h \times ratio}{2} \tag{5}$$

$$inter = \left(\min\left(b_r^{gt}, b_r\right) - \max\left(b_l^{gt}, b_l\right)\right) \times \left(\min\left(b_b^{gt}, b_b\right) - \max\left(b_t^{gt}, b_t\right)\right) \tag{6}$$

$$union = (w^{gt} \times h^{gt}) \times (ratio)^2 + (w \times h) \times (ratio)^2 - inter \tag{7}$$

$$IoU^{Inner} = \frac{inter}{union} \tag{8}$$

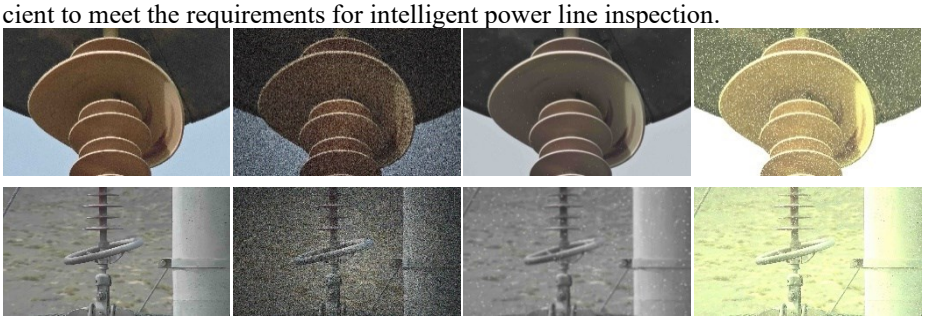
Compared to IoU loss, when the ratio is less than 1 and the auxiliary bounding box size is smaller than the actual bounding box, the effective range of regression is smaller than that of IoU loss. However, the gradient magnitude obtained from this regression is greater than that of the IoU loss, which can accelerate the convergence of high IoU samples. Conversely, when the ratio is greater than 1, the larger auxiliary bounding box size increases the effective range of regression, enhancing the regression performance for low IoU samples. The application of the Inner-IoU loss to existing IoU-based bounding box regression loss functions results in the LInner-EIoU loss, which is expressed as follows:

$$LInner\text{-}EIoU = LEIoU + IoU - IoU^{Inner} \tag{9}$$

4 Experimental Results

4.1 Datasets

The publicly available Chinese Power Line Insulator Dataset (CPLID) [22] was released in 2018, containing 600 images of normal insulators and 248 images of defective insulators. The normal images are real photographs captured by drones from various angles, while the defective images are synthetic, created by pasting cropped images of defective insulators onto background images. The dataset has a limited number of images and a single defect category, focusing only on insulators, which makes it insufficient to meet the requirements for intelligent power line inspection.



(a) original (b) foggy (c) snowy (d) sandstorm

Fig. 7. Images of transmission lines in complex weather.

To address these limitations, this study constructs a high-voltage transmission line defect detection dataset with diverse defect categories, based on images provided by a national power grid project. A large number of images related to high-voltage transmission lines were captured in real-world scenarios, featuring defects such as insulator damage, grading ring tilt, grading ring detachment, cable strand disconnect, cable loose strands, and cable foreign objects, as shown in Fig. 1, totaling 4,800 images. To simulate complex weather conditions in the real world, data augmentation was applied to the images in the self-constructed dataset. Specifically, parameters were designed to simulate the effects of fog, snow, and sandstorm conditions, as illustrated in Fig. 7. The augmented dataset contains a total of 19,200 images. Ultimately, we created a reliable and diverse high-voltage transmission line defect detection dataset, which not only includes multiple defect types but also offers a rich variety of weather scenarios, providing a solid data foundation for the training of subsequent deep learning models.

4.2 Implementation Details

The experiments were conducted on an Ubuntu 20.04 operating system, with the specific experimental platform and configuration detailed in Table 1. During the network training phase, the batch size was set to 32, and the SGD optimizer was used to update and iterate the network parameters. The initial learning rate was set to 0.01, with a momentum of 0.9, and the number of epochs was set to 300. No pre-trained weights were used during training.

Table 1. Experimental platform and setup.

Configuration environment	Version
Operating system	Ubuntu 20.04
Deep learning framework	PyTorch 2.2.2
Computational framework	CUDA 11.8
Programming language	Python 3.10.14
CPU	Intel(R) Xeon(R) Gold 5320 CPU @ 2.20GHz
GPU	Tesla T4

4.3 Evaluation Metrics

To objectively evaluate the effectiveness of the algorithm, this paper uses Precision (P), Recall (R), Average Precision (AP), and mean Average Precision (mAP) as evaluation metrics. The formulas for calculating P and R are as follows:

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$R = \frac{TP}{TP + FN} \tag{11}$$

Here, TP represents True Positive, FP represents False Positive, and FN represents False Negative.

The Average Precision (AP) combines both Precision and Recall, and it measures the model's performance by calculating the area under the Precision-Recall curve. The formula for AP is as follows:

$$AP = \int_0^1 P(r)dr \tag{12}$$

The mean Average Precision (mAP) is the mean of the AP values across all classes, and it is used for the overall performance evaluation in multi-class object detection tasks. The formula for mAP is as follows:

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP(i) \tag{13}$$

4.4 Ablation Experiments

To validate the impact of different ratio values on the model's detection performance, experimental results on the test set are shown in Table 2. As can be seen from the table, when the ratio is set to 0.7, the LInner-EIoU loss function performs the best in terms of Precision, Recall, and Average Precision. As the ratio value increases, all metrics show a decline.

Table 2. The performance comparison is conducted for ratio values within the range (0.5, 1.5).

Loss function	ratio	P(%)	R(%)	mAP50(%)
LInner-EIoU	0.5	76.1	65.0	67.6
LInner-EIoU	0.7	78.2	67.8	69.2
LInner-EIoU	0.9	75.4	65.0	68.1
LInner-EIoU	1.1	77.2	66.1	68.3
LInner-EIoU	1.3	74.8	65.7	67.4
LInner-EIoU	1.5	72.2	65.3	66.3

To validate the impact of the proposed DEConv, Focus-Detect, and LInner-EIoU on the model's performance, YOLOv11 was used as the baseline. Ablation experiments were conducted by adding these modules to the network, and the results are presented in Table 3. As shown in the table, the performance of the YOLOv11 model gradually improved with the addition of different modules. After adding the DEConv module to the YOLOv11 model, the precision increased by 1.5%, recall slightly improved, and mAP50 showed a notable improvement, indicating that DEConv enhances feature extraction, particularly in balancing precision and recall. When the Focus-Detect was added, precision further increased by 2.8%, and both recall and mAP50 showed significant improvements, demonstrating that the Focus-Detect effectively enhanced the model's detection capability, especially in complex backgrounds. After incorporating

Table 3. Ablation experiments.

YOLOv11	DEConv	Focus-Detect	LInner-Iou	P(%)	R(%)	mAP50(%)
✓				72.1	65.0	66.8
✓	✓			73.6	65.3	67.8
✓	✓	✓		76.4	67.1	68.7
✓	✓	✓	✓	78.2	67.8	69.2

the LInner-EIoU module, all three metrics—precision, recall, and mAP50—showed significant improvement, with precision increasing by 1.8%. This indicates that the inclusion of LInner-EIoU significantly improved the model's performance in bounding

box regression, particularly in handling different types of defects, thereby enhancing the model's robustness in complex scenarios.

4.5 Comparative Experiments

This experiment compares the performance of five backbone networks in the high-voltage transmission line defect detection task. The results are shown in Table 4. Fasternet performs the best in terms of precision and mAP50, but it has a high number of parameters and computational cost. Starnet, while having the lowest parameter count and computational cost, shows significantly lower precision, recall, and mAP50 compared to other methods. The DAICW model, with a relatively low number of parameters and computational cost, demonstrates excellent performance in precision, recall, and mAP50, particularly surpassing Fasternet in recall. This indicates a good balance between resource efficiency and detection performance, making it an efficient and practical backbone network choice.

Table 4. Performance comparison of different backbone networks.

backbone	Params/M	FLOPs/G	P(%)	R(%)	mAP50(%)
EfficientViT [23]	3.7	8.0	77.9	66.3	68.5
Fasternet [24]	3.9	9.2	79.1	67.0	69.4
HGNetV2 [25]	2.1	5.9	76.0	63.5	65.3
Starnet [26]	1.9	5.0	68.6	59.1	61.5
Ours	2.5	5.8	78.2	67.8	69.2

Based on the comparative experimental results of different loss functions shown in Table 5, it can be observed that the loss function used in this study significantly improves precision, recall, and mAP. Specifically, the precision of the proposed loss function is 78.2%, recall is 67.8%, and mAP is 69.2%. Compared to other common loss functions, there are notable improvements across all metrics. For instance, compared to the EIoU loss function, precision increased by 2.6%, and recall improved by 1.9%. These results demonstrate that the loss function used in this study performs particularly well in enhancing both precision and recall, especially in the model's ability to identify defect targets and in overall evaluation metrics, significantly outperforming the traditional EIoU loss function.

Table 5. Comparative experiments with different loss functions.

Loss function	P(%)	R(%)	mAP50(%)	F1(%)
IoU	73.2	60.3	61.5	66.1
GIoU [18]	75.8	65.4	68.2	70.2
CIoU [20]	74.3	65.3	67.8	69.5
EIoU [19]	75.6	65.9	68.1	70.4
SIoU [27]	74.6	64.3	65.8	69.0
Linner-IoU	78.2	67.8	69.2	72.6

According to the performance comparison of different models shown in Table 6, the DAICW model demonstrates excellent performance with a parameter count of 2.5M. Compared to YOLOv5n, its precision increased by 1.5%, reaching 78.2%, while Recall improved by 5%, reaching 67.8%. This indicates that the DAICW model significantly enhances detection performance while maintaining a relatively low number of

parameters, especially in terms of recall, demonstrating stronger detection capability. Compared to other models, the DAICW model outperforms most existing models in both precision and recall. For instance, compared to YOLOv7-tiny, both precision and recall show significant improvement, indicating that the model designed in this study is better at capturing target information, especially when handling smaller targets or complex scenes. While the DAICW model's precision is slightly lower than YOLOv8n, it has smaller computational cost and parameter count, indicating a good balance between computational efficiency and performance. Overall, the DAICW model excels in multiple aspects compared to other models, particularly in precision and recall, providing strong support for detection tasks in complex scenarios such as high-voltage transmission lines.

Table 6. Comparison of the performance of different models.

Model	Params/M	FLOPs/G	P(%)	R(%)	mAP50(%)
Faster R-CNN [28]	27	4.3	72.1	62.3	63.5
Mask R-CNN [29]	31	6.6	76.3	65.2	67.2
RtinaNet [30]	33	117	72.1	61.4	63.5
YOLOv5n	2.5	7.1	76.7	63.8	66.9
YOLOv7-tiny [8]	6.2	4.5	69.6	60.5	63.4
YOLOv8n [7]	3.0	8.1	78.6	68.1	69.3
YOLOv10n [31]	2.2	6.5	79.2	66.1	68.8
DETR [32]	41.5	122	73.1	63.5	65.7
DAICW	2.5	5.8	78.2	67.8	69.2

5 Conclusion

For detection tasks in complex weather scenarios, we have designed the DAICW. Experimental results demonstrate that this method effectively enhances the robustness and generalization ability of high-voltage transmission line defect detection models under complex weather conditions. By introducing Detail-Enhanced Convolution, Focus-Detect, and LInner-IoU loss function, the model's detection performance has been progressively improved, with significant enhancements in precision, recall, and mAP. In comparative experiments with other mainstream models, the proposed model has shown superior performance across multiple metrics, including detection precision and recall.

6 Acknowledgments

This work was supported in part by the Scientific and Technological Innovation 2030 Major Project under Grant 2022ZD0115800, in part by the Tianshan Talent Training Program under Grant 2022TSYCLJ0036, in part by the National Natural Science Foundation of China under Grant 62472368.

References

1. Yu P, Wang Z, Zhang H, et al. Safe reinforcement learning for power system control: A review[J]. arXiv preprint arXiv:2407.00681, 2024.
2. Ma W, Xiao J, Zhu G, et al. Transmission tower and Power line detection based on improved Solov2[J]. IEEE Transactions on Instrumentation and Measurement, 2024.
3. Singh A K, Ibraheem S K, Muazzam M, et al. An overview of electricity demand forecasting techniques[J]. Network and complex systems, 2013, 3(3): 38-48.
4. Li X, Li Z, Wang H, et al. Unmanned aerial vehicle for transmission line inspection: status, standardization, and perspectives[J]. Frontiers in Energy Research, 2021, 9: 713634.
5. Li J, Xu Y, Nie K, et al. PEDNet: A lightweight detection network of power equipment in infrared image based on YOLOv4-tiny[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-12.
6. Khanam R, Hussain M. Yolov11: An overview of the key architectural enhancements[J]. arXiv preprint arXiv:2410.17725, 2024.
7. Varghese R, Sambath M. Yolov8: A novel object detection algorithm with enhanced performance and robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024: 1-6.
8. Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 7464-7475.
9. Wang Y, Song X, Feng L, et al. Mci-gla plug-in suitable for yolo series models for transmission line insulator defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2024.
10. Ding L, Rao Z Q, Ding B, et al. Research on defect detection method of railway transmission line insulators based on GC-YOLO[J]. IEEE Access, 2023, 11: 102635-102642.
11. Lin J, Zhao Y, Wang S, et al. YOLO-DA: An efficient YOLO-based detector for remote sensing object detection[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1-5.
12. Ji C, Jia X, Huang X, et al. FusionNet: Detection of Foreign Objects in Transmission Lines During Inclement Weather[J]. IEEE Transactions on Instrumentation and Measurement, 2024.
13. Jain N, Bedi J, Anand A, et al. A transfer learning architecture to detect faulty insulators in powerlines[J]. IEEE Transactions on Power Delivery, 2024, 39(2): 1002-1011.
14. Mao J, Liu Q, Wang Y, et al. MSRN: Multilevel Spatial Refinement Network for Transmission Line Fastener Defect Detection[J]. IEEE Transactions on Industrial Informatics, 2024.
15. Wang H, Luo S, Wang Q. Improved YOLOv8n for foreign-object detection in power transmission lines[J]. IEEE Access, 2024.
16. Peng L, Wang K, Zhou H, et al. Detection of bolt defects on transmission lines based on multi-scale YOLOv7[J]. IEEE Access, 2024.
17. Zhang X, Cui B, Wang Z, et al. Loader Bucket Working Angle Identification Method Based on YOLOv5s and EMA Attention Mechanism[J]. IEEE Access, 2024.
18. Rezatofighi H, Tsoi N, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 658-666.
19. Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression. arXiv 2021[J]. arXiv preprint arXiv:2101.08158, 2022.
20. ZhengZ H, WangP R D W. Enhancinggeometric factorsinmodellearningandinferenceforobjectdetection andinstancesegmentation[J]. IEEETransactionsonCybernetics, 2022, 52(8): 8574-8586.

21. Zhang H, Xu C, Zhang S. Inner-iou: more effective intersection over union loss with auxiliary bounding box[J]. arXiv preprint arXiv:2311.02877, 2023.
22. Tao X, Zhang D, Wang Z, et al. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks[J]. IEEE transactions on systems, man, and cybernetics: systems, 2018, 50(4): 1486-1498.
23. Liu X, Peng H, Zheng N, et al. Efficientvit: Memory efficient vision transformer with cascaded group attention[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 14420-14430.
24. Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 12021-12031.
25. Zhao Y, Lv W, Xu S, et al. Detrs beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2024: 16965-16974.
26. Ma X, Dai X, Bai Y, et al. Rewrite the stars[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 5694-5703.
27. Gevorgyan Z. SIoU loss: More powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.
28. Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.
29. He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
30. Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.
31. Wang A, Chen H, Liu L, et al. Yolov10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
32. Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//European conference on computer vision. Cham: Springer International Publishing, 2020: 213-229.