Integrating Local Channel Attention and Focused Feature Modulation for Wind Turbine Blade Defect Detection

Zheng Cao*, Rundong He, Shaofei Zhang, Zhaoyang Qi, Sa Li, Tong Liu and Yue Li
State Grid Jilin New Energy Group Co., Ltd, Changchun 130021, China

***Abstract*—In the wind power industry, the health state of wind turbine paddles is directly related to the power generation efficiency and the safe operation of the equipment. In order to solve the problems of low efficiency and insufficient accuracy of traditional detection methods, this paper proposes a wind turbine blade defect detection algorithm that integrates local channel attention and focus feature modulation. The algorithm first introduces the Mixed Local Channel Attention (MLCA) mechanism into the C2f module of the backbone network in YOLOv8 to enhance the extraction capability of the backbone network for key features. Then the Focal Feature Modulation (FFM) module is used to replace the original SPPF module in YOLOv8 to further aggregate global contextual features at different levels of granularity; finally, in the Neck part, the progressive feature pyramid AFPN structure is used to enhance the multi-scale feature fusion capability of the model, which in turn improves the accuracy of small object detection. The experimental results show that the proposed algorithm has an accuracy of 82.5%, a mAP50 of 78.6%, and GFLOPS of 8.5. In the detection of wind turbine blade defects, which possesses higher detection performance and real-time performance compared with the traditional methods, and is able to effectively identify common defects such as cracks, corrosion, and abrasion, and exhibits strong robustness and application value.**

***Keywords*—Fan blades; YOLO; attention mechanism; defect detection; inner-IoU**

I. Introduction

Wind energy, as a renewable energy source, is highly valued by countries around the world due to its cleanliness, environmental friendliness, and low carbon emissions. The utilization of wind energy is of great significance in supporting the dual-carbon development, thus holding an important position in the construction of China's clean energy system [1]. The wind turbine blade, as a crucial component in wind power generation, plays a key role in converting wind energy into electrical energy. However, due to the complex and diverse outdoor environment, wind turbine blades often suffer from erosion, cracks, delamination, and other damage defects, which affect the power generation efficiency of the units and even threaten the safety and service life of the wind turbines [2].

Traditional wind turbine blade defect detection primarily relies on manual inspection, which suffers from low efficiency, high costs, and strong subjectivity. Therefore, the real-time, efficient, and accurate detection of blade defects has become an urgent need in wind turbine blade defect identification. Existing wind turbine blade defect detection algorithms can be broadly categorized into those using physical sensing technology for detection and those employing computer vision technology for image recognition. Methods using physical sensing technology mainly focus on static detection with grating Energies [3] and electromagnetic ultrasonic testing [4].

In recent years, with the development of artificial intelligence technology and advancements in drone technology, drone aerial photography and object detection techniques have been widely applied to wind turbine blade defect detection [5, 6]. Kang Shuang et al. achieved defect detection on high-altitude wind turbine blades by analyzing temperature thresholds to extract effective features [7]. Yu et al. [8] trained a DCNN network to extract semantic features of defect areas for wind turbine blade recognition. Additionally, some experts and scholars have explored the optimization of deep learning-based object detection models. In the improvement of two-stage object detection algorithms, Zhang Chao [9] improved the backbone network of Mask RCNN by combining it with a multi-scale feature fusion FPN structure, enhancing the detection effect of wind turbine blade defects. Qu Zhongkan et al. [10] utilized DCNv2 and GIoU to improve Faster R-CNN, significantly enhancing the detection accuracy of delamination defects.

In the improvement of one-stage algorithms, the focus has mainly been on the YOLO series. Wu Yuping et al. [11] adopted a skip-connection feature fusion network structure in YOLOv3-Tiny and introduced the Inception module, significantly improving the detection accuracy of blade defects. Gao Wenjun [1-15] replaced the original feature extraction network of YOLOv4 with a GhostNet feature extraction network, successfully achieving lightweight blade detection models while maintaining good detection accuracy. Su Jia et al. [13] designed a multi-scale feature fusion mechanism in the YOLOv5 model, thereby improving the detection effect on small objects. Zheng Qishan [14] proposed a wind turbine blade defect detection method based on YOLOv5s as the baseline model, by replacing the fast spatial pyramid pooling (SPPF) structure in the backbone network and introducing the convolutional attention mechanism (CBAM) module, further enhancing the model's detection accuracy. Wang Zhengshuai [15] addressed the issue of low detection accuracy in complex environments by adding a Scoring module attention mechanism to the YOLOv5 baseline model, scoring each channel's features to filter out low-scoring features and fuse high-scoring features, thereby improving feature fusion capabilities and detection accuracy. Li Bing et al. [16] proposed an HSCA-YOLOv7 wind turbine blade defect detection algorithm, effectively solving the problem of

inconsistent defect scales in wind turbine blade images. Fu Jinyi et al. [17] proposed an improved small object detection algorithm, CA-YOLOv8, by embedding an aggregation capability module (CAM) and improving the C2f module, enhancing the ability to capture multi-scale detailed features of detection objects.

The methods for detecting defects in wind turbine blades in the past have gradually evolved from traditional manual methods to physical sensing technologies and computer vision technologies. In the latter, whether it is the early exploration based on temperature thresholds and DCNN networks, or the improvement of two-stage and one-stage target detection algorithms, the detection capabilities have been enhanced to a certain extent, but there are still shortcomings. For example, the complex background has a significant impact on the model's detection of defect areas. The defect areas in aerial images of wind turbine blades are mostly small targets, and different-scale targets have higher requirements for the model's feature processing ability. Therefore, the model in this paper aims to further optimize the defect detection algorithm of wind turbine blades for these problems and improve its detection accuracy and efficiency for small targets and multi-scale targets in complex environments.

The complex background of aerial wind turbine blade images significantly affects the model's detection of defect areas; the defect areas in aerial wind turbine blade images are mostly small objects, requiring high feature processing capabilities for different scale objects. To address these issues, this paper proposes a fusion of local channel attention and focal feature modulation object detection algorithm. This paper mainly improves the YOLOv8 model, with innovations focused on enhancing the overall feature extraction capabilities of the model and improving the detection capabilities for small-scale objects. The main innovations include the following four aspects:

1) Improving feature capture for small-scale objects: By integrating the MLCA module (Mixed Local Channel Attention) with the C2f module in YOLOv8's Backbone, the model's ability to capture details of small-scale objects is enhanced.

2) Enhancing overall detection accuracy: Using the FFM module (Focal Feature Modulation) to replace the original SPPF (Spatial Pyramid Pooling - Fast) module, further improving the model's overall detection accuracy.

3) Optimizing multi-scale feature fusion: Introducing AFPN (Asymptotic Feature Pyramid Network) to optimize multi-scale feature fusion in object detection.

4) Improving model accuracy and generalization: Using Inner-IoU to optimize the original model's loss function, further enhancing model accuracy and generalization capabilities

The subsequent part of this paper will conduct a detailed analysis of the algorithm structure, including the improvement principles of the C2f-MLCA and FFM modules in the Backbone part, the AFPN structure in the Neck part, and the role of the Inner-IoU loss function. Then, through the experimental results and analysis, the configuration such as the dataset will be introduced, and various experiments and evaluations will be carried out using multiple indicators to demonstrate the performance advantages of the algorithm. Finally, a conclusion will be drawn to summarize the performance improvement and application significance of this algorithm in the defect detection of wind turbine blades.

II. Improved YOLOv8 Algorithm Structure

A. Improved Network Model

This paper proposes a fusion of local channel attention and focal feature modulation object detection algorithm based on YOLOv8. In the Backbone of YOLOv8, the MLCA module is used to construct the C2f-MLCA module, improving the model's ability to capture details of small-scale objects; secondly, the FFM module is used to replace the SPPF module in the Backbone to aggregate contextual information at different granularity levels; finally, the AFPN structure is used in the Neck part to enhance the model's multi-scale feature fusion capabilities. The overall framework of the improved network is shown in Fig. 1.

B. Improvements in the Backbone

a) C2f-MLCA: There are numerous small-scale objects in wind turbine blade defects [18]. To enhance the detection effect on these small-scale objects, the module used needs to possess better global and local detail feature extraction capabilities. In YOLOv8, multiple Conv modules and C2f modules constitute the backbone network, which is used to extract deep features from images. However, the Bottleneck module in the C2f module causes the network to superimpose a large amount of information at high-frequency positions, while neglecting information at low-frequency positions, thereby reducing the model's feature extraction capability for small objects, as shown in Fig. 2.

MLCA is a lightweight attention module that can simultaneously consider channel information and spatial information, and combine local and global information to enhance the expressive effect of the network [19]. Therefore, the MLCA module is introduced into the Bottleneck of the C2f module, as shown in Fig. 3.

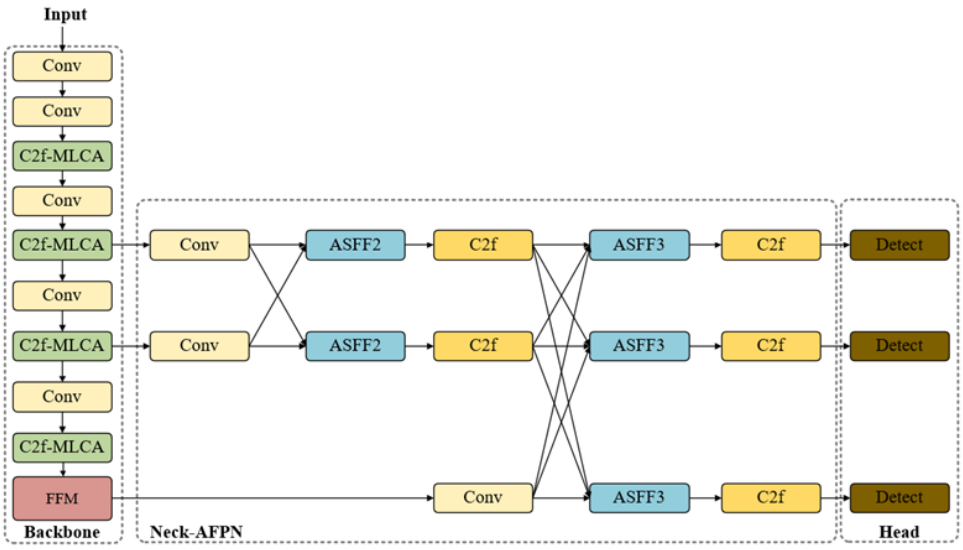

Fig. 1. Overall framework of the model.

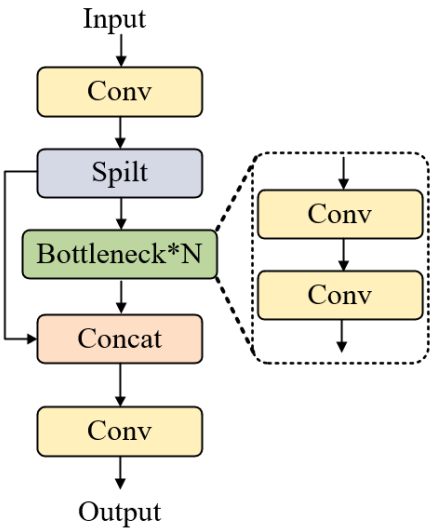

Fig. 2. C2f module.

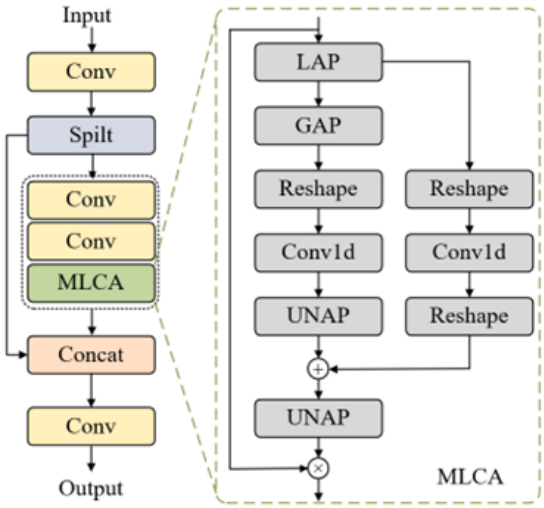

Fig. 3. MLCA module.

The MLCA module primarily utilizes local average pooling and global average pooling to capture global and local feature information of the entire input feature map, and on this basis, performs channel attention and spatial attention to improve feature extraction accuracy. A feature map of size $C \times H \times W$ is first processed by local average pooling (LAP) and global average pooling (GAP) to obtain statistical information of the entire feature map. Subsequently, the features after local and global pooling are transformed by a one-dimensional convolution (Conv1d) to compress the feature channels while maintaining the spatial dimensions, and then rearranged to adapt to subsequent operations. For the features after local pooling, after one-dimensional convolution and rearrangement, they are combined with the original input features through multiplication operations for feature selection to enhance attention to useful features; for the features after global pooling, after one-dimensional convolution and rearrangement, they are combined with the local pooling features through addition operations to fuse global context information. Finally, the feature map processed by local and global attention is restored to the original spatial dimension through anti-average pooling (UNAP) operations to achieve the purpose of mixed attention.

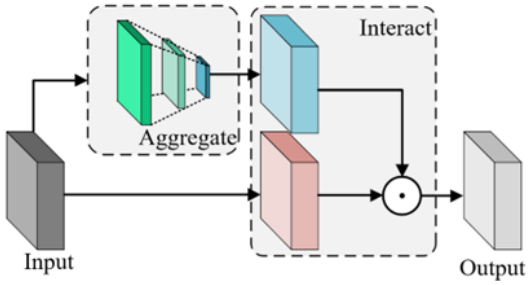

Fig. 4. FFM module.

b) FFM module: During the detection process of wind turbine blade defect objects, the model is easily affected by the complex background interference associated with the samples, making it difficult to fully extract the deep features of the samples. YOLOv8 uses the SPPF module in the Backbone part

to aggregate multiple scale features, but it lacks sufficient modeling of long-range visual context information. In feature information processing, it only concatenates feature map information of different scales to achieve feature aggregation, ignoring the interaction and aggregation of multi-granularity information under complex background interference. To address this issue, this paper introduces a more efficient focal feature aggregation module, FFM [20]. By extracting context information from local to global at different granularity levels and performing gated aggregation, it improves the interference of complex background information on the feature extraction process of the model's backbone network. The structure of the FFM module is shown in Fig. 4.

First, the input feature $X \in \mathbb{R}^{H\times W\times C}$ is fed into the aggregation module to aggregate contextual features, obtaining aggregated features. This process mainly includes two steps: The first step involves sending the input feature into a linear layer to extract the initial level feature Z^0. Then, the initial level feature is processed through a Depthwise Convolution (DWConv) and a GeLU function to obtain the next level feature until the L -th level feature. In this way, a total of (L+1) level features are obtained. These features are considered as contextual information at different granularity levels, as shown in Eq. (1) and Eq. (2):

$$Z^0 = linearlayer(X) \tag{1}$$

$$Z^l = GeLU(DWConv(Z^{l-1})) \tag{2}$$

Among them, $l \in \{1,\cdots,L\}$ represents different levels.

The second step is gated aggregation. By gated aggregation, contextual feature information at different granularity levels is condensed into a single feature vector as the aggregated feature Z^{Out}, with the calculation formula as follows:

$$Z^{Out} = \sum_{l=1}^{L+1} G^l \odot Z^l \tag{3}$$

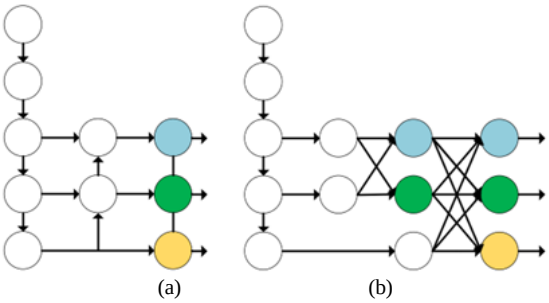

Fig. 5. The schematic of the network before and after the improvement of the Neck part. (a) YOLOv8; (b) YOLOV8+AFPN.

where G^l is the gate weight obtained by the l-th level feature through a linear layer.

Then, the obtained aggregated feature Z^{Out} is processed through a linear layer and element-wise multiplied with the feature query value to obtain the final output focal modulation feature Y, with the calculation formula as follows:

$$Y = q(X) \odot linerlayer(Z^{Out}) \tag{4}$$

where $q(\cdot)$ is the query projection function for obtaining the feature query value; $\odot$ represents element-wise multiplication.

The FFM module achieves this not only by expanding the receptive field and effectively mitigating the interference from the complex background in wind turbine blade defect samples but also by more efficiently obtaining contextual information from granularity levels, thereby enhancing the feature extraction capability of the Backbone.

C. Improvements in the Neck Part

In the Neck part, the AFPN structure [21] is adopted instead of the original FPN+PAN structure in YOLOv8. This is because, during the feature extraction process, the downsampling process is dominated by large objects, causing the feature representation of small objects in the object area to be confused or blurred, thereby affecting the detection effect of small objects. The AFPN structure used in this paper can optimize the multi-scale feature fusion process. By gradually fusing low-level and high-level features, it avoids the loss and degradation of information, as shown in Fig. 5.

The AFPN structure can adaptively fuse features of different levels in the Neck. Firstly, it fuses two adjacent low-level features, and then progressively incorporates higher-level features. This approach alleviates the conflict of multi-object information that arises from each spatial location during the feature fusion process. Specifically, during the multi-level feature fusion process in the Neck part, an adaptive spatial fusion module (ASFF) is used to assign different spatial weights to features of different levels. Then, the features of different levels are fused together using a weighted sum method, with the specific calculation formula as follows:

$$y_{ij}^{\ell} = \sum_{n=1}^{\ell} \alpha_{ij}^{n} x_{ij}^{n\to\ell} \tag{5}$$

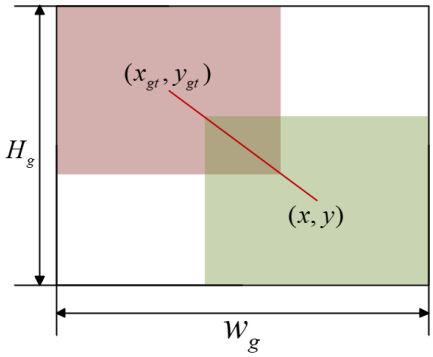

Fig. 6. Parameter definition.

Where y_{ij}^{ℓ} represents the features after adaptive spatial fusion, α_{ij}^{n} represents the spatial weight of the features at the ℓ level in the hierarchy, and $\sum_{n=1}^{\ell} \alpha_{ij}^{n} = 1$; $x_{ij}^{n\to\ell}$ represents the feature vector from the n-th level to ℓ at position (i,j).

D. Improvement of Loss Function

a) CIoU loss: The original YOLOv8 model employs the CIoU Loss as the bounding box regression loss function. Let the ground truth box be denoted as $\vec{B}_{gt} = [x_{gt}, y_{gt}, w_{gt}, h_{gt}]$ and the predicted box as $\vec{B} = [x, y, w, h]$, where x and y represent the center coordinates of the bounding box, and w and h represent the dimensions of the box. The original CIoU Loss is defined by Eq. (6)-Eq. (9) as follows:

$$L_{CIoU} = L_{IoU} + \frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g + H_g)^2} + \alpha v \tag{6}$$

$$L_{IoU} = 1\text{-}IoU = 1 - \frac{W_i H_i}{wh + w_{gt}h_{gt} - W_i H_i} \tag{7}$$

$$\partial = \frac{v}{L_{IoU} + v} \tag{8}$$

$$v = \frac{4}{\pi^2}(\tan^{-1}\frac{w}{h} - \tan^{-1}\frac{w_{gt}}{h_{gt}})^2 \tag{9}$$

In the equation, L_{IoU} is used to measure the overlap between the predicted box and the ground truth box, α is a balancing parameter, and v is used to measure the consistency of the aspect ratio. The remaining parameters are defined as shown in Fig. 6.

b) Inner-IoU: The existing IoU-based bounding box regression still focuses on accelerating convergence by adding new loss terms, while neglecting the limitations of the IoU loss term itself. In practical applications of wind turbine blade defect detection, the original bounding box regression process cannot self-adjust according to different detectors and detection tasks, resulting in low detection efficiency and accuracy for dense small objects or complex environments with multiple scales. To address this issue, Inner-IoU loss [22] is chosen, which calculates the IoU loss through an auxiliary bounding box. For different datasets and detectors, the scale of the auxiliary bounding box used to calculate the loss can be controlled by a scale factor. The definition of Inner-IoU is shown in Eq. (10)-(16).

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} * ratio}{2} \tag{10}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} * ratio}{2} \tag{11}$$

$$b_l = x_c - \frac{w * ratio}{2}, b_r = x_c + \frac{w * ratio}{2} \tag{12}$$

$$b_t = y_c - \frac{h * ratio}{2}, b_b = y_c + \frac{h * ratio}{2} \tag{13}$$

$$inter = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) * (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)) \tag{14}$$

$$union = (w^{gt} * h^{gt}) * (ratio)^2 + (w * h) * (ratio)^2 - inter \tag{15}$$

$$IoU^{inner} = \frac{inter}{union} \tag{16}$$

In the Inner-IoU Loss, the variable ratio corresponds to the scale factor, which typically ranges from [0.5, 1.5]. The traditional IoU calculation method considers the overlap area between the predicted bounding box and the overall bounding box, while Inner-IoU focuses more on the core part of the bounding box to make a more precise judgment of the overlap area.

III. Experimental Results and Analysis

A. Experimental Dataset

The data used in this paper are sourced from images taken by inspection personnel at a wind farm during the inspection of wind turbine blades. The division of the aerial wind turbine blade defect dataset is shown in Table I, totaling 784 images with a resolution of 1200*900. The dataset is randomly divided into training, validation, and test sets in a ratio of 7:2:1.In the defect detection of wind turbine blades, 70% of the data is used for training to enable the model to learn multiple features and enhance performance. The 20% validation set assists in training and optimizing the model. The 10% test set independently evaluates to ensure objectivity and accuracy. The 7:2:1 data division ratio balances multiple key factors and enhances the reliability of the experiment.

B. Experimental Environment and Parameter Configuration

The operating system used in this experiment is Ubuntu 18.04, with GPU being NVIDIA Geforce RTX 3090 and CPU being Intel(R) Xeon(R) Gold 6148. The deep learning framework is Pytorch-2.1.2. During the training of the object detection model, the number of training epochs is set to 200, the momentum size is set to 0.937, the batch size is set to 16, and the initial learning rate is set to 0.0001.

C. Experimental Evaluation Metrics

To verify the effectiveness of the improvements in the proposed algorithm, the model is evaluated using four metrics: Precision (P), Recall (R), Average Precision (AP), and mean Average Precision (mAP). The definitions of these metrics are given in Eq. (17)-(20), where n is the total number of predicted bounding box categories.

$$P = \frac{TP}{TP + FP} \tag{17}$$

$$R = \frac{TP}{TP + FN} \tag{18}$$

$$AP = \int_0^1 P(R)dR \tag{19}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP(i) \tag{20}$$

D. Results Comparison and Analysis

a) Comparative analysis of different object detection model: To verify the advanced nature of the proposed method

for wind turbine blade defect detection, the proposed improved algorithm is compared with other mainstream object detection model algorithms on the wind turbine blade defect dataset. The experimental results are shown in Table II. From Table II, it can be seen that the proposed improved algorithm significantly outperforms the two-stage models Faster R-CNN [23] and Cascade R-CNN[24] in terms of mAP and detection speed. Compared to the single-stage object detection algorithms YOLOv3[25],YOLOv4[26],YOLOv5s,YOLOv5l,YOLOv8s,YOLOv9[27],YOLOv10[28] and YOLOv8n, the proposed algorithm achieves a substantial improvement in mAP while maintaining a high level of accuracy and a fast detection speed, indicating that the proposed method has certain advantages.

To provide a more intuitive comparison of the detection effects before and after the improvement of the proposed algorithm with other mainstream YOLO series algorithms, the detection results of different scenarios are compared and displayed in Fig. 7(a)-(f). The first row represents complex background scenarios, the second row represents scenarios with multiple types of defects, the third row represents scenarios with densely distributed small objects, and the fourth row represents scenarios with multiple types of defects and small objects. It can be seen from the comparison that the proposed algorithm has a lower rate of missed detections and false detections and higher recognition accuracy compared to other algorithms.

TABLE I. TABLE OF DATASET DIVISION FOR AERIAL WIND TURBINE BLADE DEFECTS

Fan blade defect name	Quantity of categories
gelcoat_off	156
cracks	374
surface_erosion	254

TABLE II. COMPARISON EXPERIMENT OF DIFFERENT OBJECT DETECTION MODELS

Model	P/%	R/%	mAP50/%	mAP50-95/%	GFLOPs
Faster R-CNN	75.4	66.0	73.5	43.1	7.1
Cascade R-CNN	72.5	67.8	72.7	44.2	7.1
YOLOv3	73.5	67.9	71.9	43.3	6.9
YOLOv4	76.6	65.9	74.9	45.4	6.7
YOLOv5s	78.4	65.3	75.2	44.4	23.8
YOLOv9s	78.6	66.4	74.6	50.2	26.7
YOLOv10n	72.4	58.9	71.8	40.6	8.2
YOLOv8s	78.2	67.7	75.8	48.6	28.4
YOLOv8n	77.1	68.7	74.6	46.4	8.1
Ours	82.5	71.6	78.6	51.7	8.5

b) Module ablation experiment results: To verify the contribution of each improvement to the proposed algorithm, ablation experiments were conducted. Under the same experimental conditions, YOLOv8n was used as the baseline model, and a series of ablation experiments were performed by sequentially adding each improvement measure. The results are shown in Table III. From the comparison of experimental results in Table III, it can be seen that the original YOLOv8n model performs relatively basic across various metrics, with mAP50 at 88.4%, P at 65.2%, R at 73.2%, and so on. As different improvement methods were introduced, the performance of each model improved. For example, YOLOv8n+MCLA showed progress in some metrics compared to the original model; YOLOv8n+FFM, YOLOv8n+Neck improvement, YOLOv8n+Inner-IoU, and other single or combined improvement methods also enhanced the model's performance to varying degrees. However, the proposed model outperformed in all metrics. It achieved an mAP50 of 91.4%, higher than all other improved models; P was 67.8%, R was 76.5%, showing good performance in accuracy and recall. Additionally, the mAP50-95% metric was 71.6%, higher than other models, indicating that the proposed model maintains high detection accuracy across different confidence levels. Although the GFLOPs slightly increased, it remained within a reasonable range.

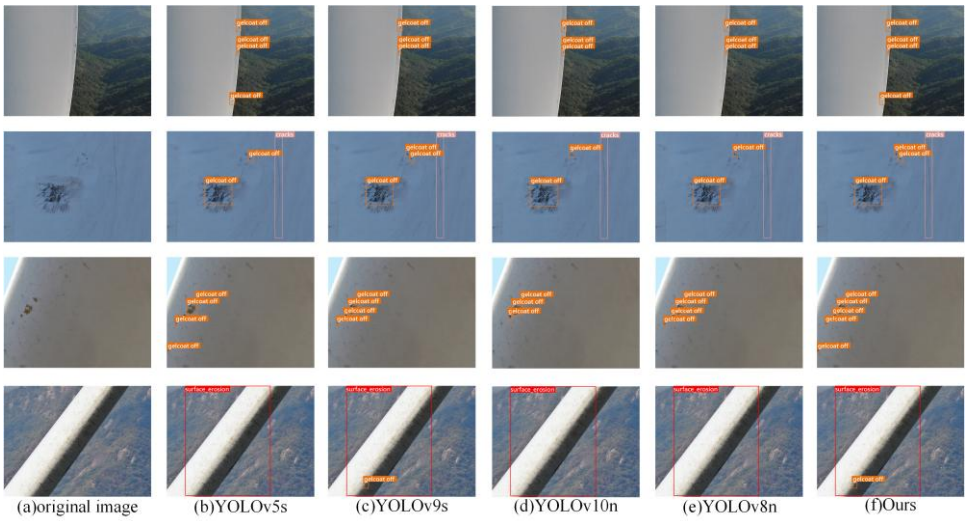

Fig. 7. Comparison of detection effect of different methods before and after improvement.

TABLE III. RESULTS OF ABLATION EXPERIMENT

Model	mAP50/%			P/%	R/%	mAP50/%	mAP50-95/%	GFLOPs
	surface_ erosion	cracks	gelcoat _off					
YOLOv8n	88.4	65.2	73.2	77.1	68.7	74.6	46.4	8.1
YOLOv8n+MCLA	89.6	66.2	73.6	78.6	68.8	74.9	48.7	7.6
YOLOv8n+FFM	88.9	66.3	75.6	77.8	68.6	76.9	46.8	7.8
YOLOv8n+Neck	90.4	65.4	76.3	78.2	69.0	77.4	48.2	8.5
YOLOv8n+Inner-IoU	90.6	67.0	76.4	80.2	70.8	78.0	48.7	8.1
YOLOv8n+ MCLA +FFM	88.7	67.1	75.8	80.6	71.2	77.2	47.5	7.8
YOLOv8n+ MCLA +FFM+ Neck	91.2	67.6	76.0	81.2	70.3	78.3	47.9	8.5
YOLOv8n+ MCLA +FFM +Inner-IoU	90.6	67.2	76.5	82.1	71.4	78.1	48.2	7.8
YOLOv8n+Neck+Inner-IoU	89.6	67.4	75.2	81.7	71.3	77.4	50.6	8.5
Ours	91.4	67.8	76.5	82.5	71.6	78.6	51.7	8.5

TABLE IV. PERFORMANCE OF MODELS WITH DIFFERENT C2F-MCLA NUMBERS

Model	mAP50/%	GFLOPs
YOLOv8n	74.6	8.1
C2f-MCLA*1	74.9	7.6
C2f- MCLA *2	74.7	7.6
C2f- MCLA *3	75.1	7.7

In summary, the proposed model, by integrating multiple improvement methods, demonstrated outstanding performance in defect detection tasks, providing a more effective solution for detecting issues such as surface erosion cracks and gel coating detachment.

c) Impact of C2f-MLCA module on the proposed model: In YOLOv8's Backbone, there are three instances where the C2f module is used. The stacking of different numbers of C2f-MCLA modules has varying impacts on network performance and lightweighting. To determine the optimal number of improvement modules for peak performance, YOLOv8n was used as the baseline algorithm, and the first, the first two, and all C2f modules were improved. The network performance with different numbers of C2f-MCLA modules was compared, with the module improvement sequence following the network structure hierarchy. The experimental results are shown in Table IV.

The mAP50 of YOLOv8n is 74.6%, and GFLOPs is 8.1. The C2f-MCLA series achieves mAP50 values close to or slightly higher than YOLOv8n under different configurations, with C2f-MCLA*3 reaching 75.1%; and the series has GFLOPs around 7.6 or 7.7, indicating relatively lower computational complexity. Overall, the C2f-MCLA series strikes a good balance between object detection accuracy and computational efficiency, potentially offering more advantages in scenarios with limited computational resources or high real-time requirements, providing a reference for selecting object detection algorithms.

d) Evaluation of model generalization and robustness: In the field of object detection, a model's generalization capability is primarily reflected in its ability to adapt well to changes brought by different datasets. The robustness of a model focuses more on its ability to better adapt to the effects brought by changes in sample backgrounds and multiple angles. To evaluate the generalization and robustness of the proposed method, 40 wind turbine defect samples under multiple backgrounds and angles were selected for experimental detection, with results shown in Table V.

TABLE V. EVALUATION RESULT

Model	Number of correct identifications	Accuracy
YOLOv8n	32	80%
YOLOv8n+MCLA	37	92.5%
YOLOv8n+FFM	37	92.5%
YOLOv8n+AFPN	32	80%
YOLOv8n+Inner-IoU	37	92.5%
YOLOv8n+ MCLA +FFM	36	90%
YOLOv8n+ MCLA +FFM+ AFPN	37	92.5%
YOLOv8n+ MCLA +FFM +Inner-IoU	37	92.5%
YOLOv8n+AFPN+Inner-IoU	33	82.5%
Ours	37	92.5%

In the original YOLOv8n model, the number of correctly identified objects is 32, with an accuracy rate of 80%. By separately introducing improvements such as MCLA, FFM, AFPN, and Inner-IoU, the model's performance has been enhanced to varying degrees. Notably, YOLOv8n + MCLA, YOLOv8n + FFM, YOLOv8n + Inner-IoU, YOLOv8n + MCLA + FFM + AFPN,YOLOv8n+MCLA+FFM+Inner-IoU, and the model pro-posed in this study all achieved 37 correctly identified objects and a high accuracy rate of 92.5%. Although the accuracy rate is the same as that of the above-mentioned models, the improvement measures such as MLCA, FFM, AFPN, and Inner-IoU integrated in this model collaborate with each other. The AFPN structure helps maintain a high accuracy rate; the MLCA module enhances the data processing ability, giving it an advantage in feature learning. The current similar accuracy rate might be a phased performance, and the scalability

of its architecture indicates that it has greater potential to surpass other models in the field of wind turbine blade defect detection in the future. This clearly indicates that these improvement methods have a significant effect on enhancing the accuracy of object recognition. However, despite some improvements in YOLOv8n+AFPN and YOLOv8n+AFPN+Inner-IoU, the extent of improvement is relatively small. Overall, the model proposed in this study performs excellently in the task of detecting defects in wind turbine blades, providing valuable references for further optimization of the YOLOv8n model.

IV. Conclusion

The detection of defects in wind turbine blades is of great significance for intelligent inspection in wind farms [29-30]. To address the practical issues of existing algorithms for detecting defects in wind turbine blades, such as insensitivity to small object areas, difficulty in feature extraction, and challenges in dealing with the effects of complex environments, this paper proposes a defect detection algorithm for wind turbine blades that integrates local channel attention and focal feature modulation. Through experiments on a specific dataset of wind turbine blade defects, the proposed method improves the accuracy by 5.4%, the recall by 2.9%, and mAP50 by 4.2% compared to the baseline method, with enhanced generalization and robustness. The proposed method also demonstrates certain advantages over other advanced object detection methods, providing a solution to assist in the intelligent inspection of wind farms. In the future, we will continue to focus on model compression and acceleration in terms of algorithm optimization, improve the attention mechanism to enhance the detection speed and accuracy, and increase the application value of the algorithm.

Acknowledgment

Author Contributions: Conceptualization, Zheng Cao and Yue Li; Data curation, Rundong He, Shaofei Zhang and Sa Li; Formal analysis, Rundong He; Investigation, Shaofei Zhang; Methodology, Zheng Cao, Rundong He and Tong Liu; Project administration, Zhaoyang Qi; Resources, Tong Liu and Yue Li; Software, Zheng Cao; Validation, Tong Liu; Visualization, Sa Li; Writing – original draft, Zheng Cao; Writing – review & editing, Zhaoyang Qi and Yue Li.

Funding: This work was supported by the State Grid Jilin New Energy Group Company Science and Technology Project (SGJLNY00YJJS2400119).

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] J. Cai, G. Hu, H. Shi, “Review on evaluation standards of wind energy resources at home and abroad,” Wind Energy, vol. 12, pp. 56-63, 2021.

[2] C. SAAD, B. H. Mostafa and H. J. R Abderrahmane, “Performance Analysis of Faults Detection in Wind Turbine Generator Based on High-Resolution Frequency Estimation Methods,” International Journal of Advanced Computer Science and Applications (IJACSA), vol. 5, no. 4, 2014.

[3] Z. Cao, Q. Wang. “Deep Learning-Based Image Recognition Technology for Wind Turbine Blade Surface Defects,” International Journal of Advanced Computer Science & Applications, vol. 12, no. 9, 2024.

[4] B. Z. Bo, Z. Yanan and C. Changzheng, “Acoustic emission detection of fatigue cracks in wind turbine blades based on blind deconvolution separation,” Fatigue & Fracture of Engineering Materials & Structures, vol. 40, no. 6, pp. 959-970, 2017.

[5] B. He, H. Jia, F. Zhao, “Application of UAV in fan blade detection,”Electrical technology, pp. 64-65, 2019.

[6] M. Ma, H. Liu, Y. Gao, S. Li, and C. Ma, “Fully autonomous fan blade inspection based on UAV and artificial intelligence,” Electric age,pp. 44-48, 2023.

[7] S. Kang, C. Chen, B. Zhou, “Research on infrared detection of wind power blade defect recognition based on temperature threshold,” Journal of Solar Energy,pp. 337-341, 2019.

[8] Y. Yu, H. Cao, X. Yan, et al., “Defect identification of wind turbine blades based on defect semantic features with transfer feature extractor,” Neurocomputing, vol. 376, pp. 1-9, 2020.

[9] W. Gao, H. Zhang, “Based on Improved YOLOv4 fan blade Defect Detection Method. Agricultural Equipment and Vehicle Engineering,” vol. 61, no. 08, pp. 94-98, 2023.

[10] Z. Qu, X. Li, F. Li, “Design of fan blade defect image multi-point detection system based on Faster R-CNN,” Electronic design engineering, vol. 29, no. 04, pp. 57-61, 2021.

[11] Y. Wu, H. Liu, J. Wu, H. Jia, J. Diao, and D. Zhu, “Application of improved YOLOv3-Tiny network in damage detection of fan blades,” Hebei Industrial Science and Technology, vol. 38, no. 05, pp. 401-408, 2021.

[12] W. Gao, H. Zhang, “Based on the improved YOLOv4 fan blade defect detection method,” Agricultural Equipment and Vehicle Engineering, vol. 61, no. 08, pp. 94-98, 2023.

[13] J. Su, Y. Qin, Z. Jia, J. Wang, “Small target detection algorithm based on ATO-YOLO,” Computer Engineering and Applications, vol. 60, no. 06, pp. 68-77, 2024.

[14] Q. Zheng, S. Zhu, C. Chen, H. Chang, and X. Yan, “Fan blade defect detection method based on improved YOLOv5s,” Information and Computers, vol. 35, no. 16, pp. 14-18, 2023.

[15] Z. Wang, L. Qiu, Y. Li, “YOLOv5s pyrotechnic detection method in complex environment,” Electronic Measurement Technology, vol. 46, no. 24, pp. 149-156, 2023.

[16] B. Li, Y. Bai, K. Zhao, C. Guo, and Y. Zhai, “Wind turbine blade surface defect detection algorithm based on HSA-YOLOV7,” Electric Power. vol. 56, no. 10, pp. 43-52, 2023.

[17] J. Fu, Z. Zhang, W. Sun,and K. Zou, “Improved YOLOv8 aerial image small target detection algorithm,” Computer Engineering and Applications, vol. 60, no. 06, pp. 100-109, 2024.

[18] Y. Xin, G. Wu, Y. Zuo, “Fan blade defect detection method based on EfficientDet,” Electronic Measurement Technology, vol. 45, no. 05, pp. 124-131, 2022.

[19] D. Wan, R. Lu, S. Shen, et al., “Mixed local channel attention for object detection,” Engineering Applications of Artificial Intelligence, vol. 123, 2023.

[20] J. Yang, C. Li, X. Dai, et al., “Focal modulation networks,” Advances in Neural Information Processing Systems, vol. 35, pp. 4203-4217, 2022.

[21] G. Yang, J. Lei, Z. Zhu, et al., “AFPN: Asymptotic feature pyramid network for object detection,” 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp. 2184-2189, 2023.

[22] H. Zhang, C. Xu, and S. Zhang, “Inner-IoU: more effective intersection over union loss with auxiliary bounding box” arXiv preprint arXiv:2311.02877, 2023.

[23] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39(6): 1137-1149.

[24] Cai Z, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2019, 43(5): 1483-1498.

[25] Redmon J. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[26] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[27] Wang C Y, Yeh I H, Mark Liao H Y. Yolov9: Learning what you want to learn using programmable gradient information[C]//European Conference on Computer Vision. Springer, Cham, 2025: 1-21.

[28] Wang A, Chen H, Liu L, et al. Yolov10: Real-time end-to-end object detection[J]. arXiv preprint arXiv:2405.14458, 2024.

[29] Memari M, Shakya P, Shekaramiz M, et al. Review on the advancements in wind turbine blade inspection: Integrating drone and deep learning technologies for enhanced defect detection[J]. IEEE Access, 2024.

[30] Du Y, Zhou S, Jing X, et al. Damage detection techniques for wind turbine blades: A review [J]. Mechanical Systems and Signal Processing, 2020, 141: 106445.