doi:10.6041/j.issn.1000-1298.2025.12.044

Quantitative Detection of Conidiophores and Sporangium of Cucumber Downy Mildew Based on Improved YOLO v8s-OBB

ZHANG Yiding[1] QIAO Chen[2] ZHANG Lingxian[2] HAN Zonghuan[2]

(1. *College of Engineering*, *China Agricultural University*, *Beijing* 100083, *China*

2. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Cucumber downy mildew is a fungal disease that severely threatens cucumber production and quality. Quantitative detection of sporangia and conidiophores is crucial for early disease prevention. However, traditional horizontal bounding box detection methods cannot accurately detect these features due to their diverse morphology and orientations. Therefore, an improved YOLO v8s-OBB detection method was proposed by introducing the convolutional block attention module (CBAM) and the lightweight shared convolution detection head (LSCD) module. The aim was to enhance the detection efficiency and accuracy of sporangia and conidiophores of cucumber downy mildew. By incorporating CBAM, the model's ability to identify key features was enhanced, allowing it to focus more on critical regions in microscopic images and improve the detection of small targets. The LSCD integrated multi-scale features through shared convolution operations, enhancing the model's detection performance for targets of different sizes while reducing computational costs, making it suitable for resource-constrained environments. The rotated bounding box technique accurately captured sporangia and conidiophores' inclination and rotation postures. Experimental results showed that, compared with the original YOLO v8s-OBB model, the improved YOLO v8s-OBB model not only reduced the model size but also achieved superior detection performance for sporangia and conidiophores of cucumber downy mildew, with precision, recall, and mAP@0.5 reaching 96.0%, 90.1%, and 96.5%, respectively. The improved YOLO v8s-OBB model outperformed advanced rotated object detection models such as S2ANet, H2RBox, and R2CNN in detection accuracy. The research result can validate the effectiveness of the improved model in practical applications and provide technical support for the early diagnosis of cucumber downy mildew.

Key words: cucumber downy mildew; conidiophore; sporangia; YOLO v8s-OBB; CBAM; LSCD

CLC number: TP3 **Document code**: A **Article ID**: 1000-1298(2025)12-0479-11 **OSID**:

0 Introduction

Cucumber downy mildew is a significant disease that significantly impacts cucumber production worldwide. This disease is caused by the pathogen *Pseudoperonospora cubensis*, which is highly contagious and spreads rapidly, often leading to a substantial decline in cucumber yield and a severe deterioration in quality[1-3]. During the disease development, the pathogen primarily spreads through the release of motile spores. These spores can quickly disperse in suitable moist environments and invade host plants through stomata or intercellular spaces, proliferating and spreading within the host cells[4]. The infection and dissemination of the pathogen rely on its unique reproductive structures—sporangia and conidiophores. Sporangia, formed on conidiophores, are the critical structures for disseminating *Pseudoperonospora cubensis*. The quantity and distribution of these two structures directly influence the pathogen's infectivity

Received date: 2024-09-10 Accepted date: 2024-10-10

Foundation item: National Natural Science Foundation of China (62176261)

First author: ZHANG Yiding, postdoctoral researcher. E-mail: zhangyiding117@163.com

Corresponding author: QIAO Chen, postdoctoral researcher. E-mail: 649112905@qq.com

and dissemination speed. Therefore, rapid and accurate detection of sporangia and conidiophores is essential for effective disease management.

Traditional pathogen detection methods rely on manual observation and counting, which are time-consuming, labor-intensive, and prone to human error[5-7]. These approaches struggle to identify pathogens with small or complex morphological structures. With advances in image processing and artificial intelligence, deep learning has become a research hotspot in plant disease detection, greatly improving detection efficiency for crops such as apples, grapes, and maize[8-10]. Convolutional neural networks (CNNs) have shown excellent performance in this field; for instance, CNN models trained on segmented datasets have achieved over 98% accuracy, emphasizing the role of image segmentation in enhancing model performance[11]. Similarly, the combination of YOLO and CNN techniques has reached a validation accuracy of 96% in leaf disease detection[12]. QIAO et al.[13] developed a Faster-NAM-YOLO model for quantitative detection of cucumber downy mildew spores, demonstrating strong detection capability. However, traditional horizontal bounding boxes (HBBs) are inadequate for complex, irregularly shaped targets such as conidiophores and sporangia, as their fixed orientations often include excessive background noise and reduce precision[14-15].

To address these challenges, the oriented bounding box (OBB) techniquewas introduced to achieve high-precision detection of complex targets with varying orientations. OBBs can rotate according to the target's actual direction, providing a more accurate description of its shape and posture. SONG et al.[16] designed a YOLO v5-OBB network for detecting corn endosperm cracks, achieving robust performance with low memory use. GENG et al.[17] proposed DR-Net, which improved recognition and counting accuracy of occluded group pigs. YAO et al.[14] introduced a polar coordinate-based boundary representation that reached a mean average precision of 76.25% on the DOTA dataset. WANG et al.[15] reframed OBB detection as pixel-level classification to enhance multi-category detection accuracy. HE et al.[18] applied OBB to SAR ship detection by using polar encoding to mitigate angle discontinuities, while CHENG et al.[19] improved YOLO v7 for X-ray contraband detection. HE et al.[20] further developed a sparse-Transformer-based rotational target detector for remote-sensing images, achieving precisions of 72.87% and 90.4%. These advancements demonstrated that OBB-based methods are well-suited for detecting targets with complex morphologies and orientations, offering effective solutions for precise identification of cucumber downy mildew conidiophores and sporangia.

Focusing on the conidiophores and sporangium of cucumber downy mildew in microscopic images, an improved YOLO v8s-OBB rotational target detection model was proposed. Firstly, the model enhanced the recognition capability of critical features by introducing the CBAM attention mechanism, enabling it to focus better on essential areas of microscopic images, thereby improving the detection of small targets. Secondly, the introduction of LSCD, which integrated multi-scale features through shared convolution operations, enhanced the model's detection performance for targets of different sizes while reducing computational costs to meet the needs of resource-constrained environments. Finally, the model achieved precise quantitative detection of the conidiophores and sporangium of cucumber downy mildew.

1 Materials and methods

1.1 Analysis of the infection process of cucumber downy mildew

During the infection process of cucumber leaves by downy mildew, the primary physiological characteristicwas the appearance of dark brown lesions on the leaf surface, a phenomenon caused by *Pseudoperonospora cubensis*[21], as shown in Fig. 1. This pathogen spreads rapidly and caused significant damage, severely affecting cucumber production and quality. The development and spread of conidiophores and sporangium were closely related to the occurrence of cucumber downy mildew. Sporangium, which growed at the tips of conidiophores, was the fundamental structure responsible for the propagation of the downy mildew pathogen. Their primary function was to produce and release spores under favorable environmental conditions. These spores can survive at lower temperatures and were released when conditions

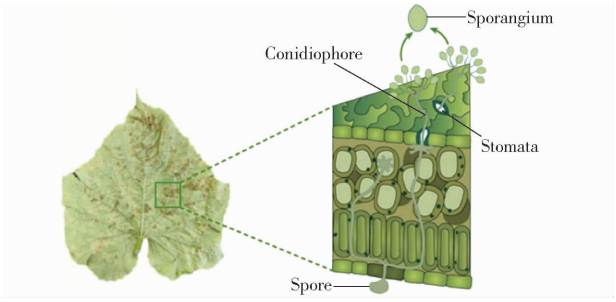


Fig. 1　Diagram of sporangium and conidiophore

were suitable. They were then spread to healthy plants by wind and rain, entering the plant through stomata and initiating a new infection cycle[22-23]. Conidiophores, which growed from the underside of the leaf surface, developed under suitable environmental conditions, such as warmth and humidity, thereby increasing the number of sporangium[24]. The length and position of conidiophores had a decisive impact on the spore dispersal range, determining how far the wind and the dispersion pattern upon landing can carry the spores. Therefore, the precise identification of sporangium and conidiophores was not only crucial for a deeper understanding of the biological characteristics and dissemination mechanisms of the disease but also enabled researchers to more accurately predict the occurrence and spread of the disease, thereby advancing the control measures for cucumber downy mildew.

1.2　Data acquisition and preprocessing

The cucumber downy mildew leaf samples were collected from the Modern Agricultural Science and Technology Innovation Base of the Tianjin Academy of Agricultural Sciences and the Plant Protection Institute of the Tianjin Academy of Agricultural Sciences. To further study the distribution of conidiophores and sporangium of cucumber downy mildew, microscopic images of the pathogen in the experimental samples were obtained. The experiment began with the preparation of three 1.5 mL centrifuge tubes, each containing 1 mL of distilled water. The downy mildew pathogens on the leaf surface were gently brushed into the centrifuge tubes using a soft brush, creating a suspension containing conidiophores and sporangium. To accurately analyze the distribution of sporangium and conidiophores, an Olympus BX51 microscope set to 200x magnification was used to capture microscopic images by extracting suspension samples from the three different centrifuge tubes. Each microscopic image had a resolution of 1 280 pixels × 960 pixels. To ensure the representativeness and reliability of the data, each centrifuge tube sample was used to generate 550 independent microscopic images, resulting in a total of 1 650 images. These images were stored in JPG format to facilitate subsequent image analysis and data processing. The collected microscopic images were shown in Fig. 2. The images were divided into training, validation, and test sets in a ratio of 8∶1∶1.

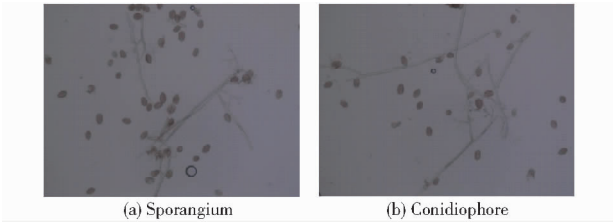


Fig. 2　Microscopic images of sporangium and conidiophore

1.3　Experimental platform and evaluation metrics

The entire experiment was conducted on a Windows 10 (64-bit) operating system, with a computer equipped with 64 GB of memory, an Intel (R) Core(TM) i9 - 13900KF @ 3.0 GHz CPU, and an NVIDIA GeForce RTX 4090 GPU with 24 GB of VRAM. The deep learning framework used was PyTorch 1.13.1, with CUDA version 11.7. During the model training process, the number of iterations (epochs) was set to 150, and the batch size for each training session was set to 32. The Adam optimizer was used for learning rate adjustment, while the other parameters during the training process were kept at their default settings.

Evaluation metricswere crucial for assessing the effectiveness of object detection algorithms. The improved YOLO v8s-OBB model used precision (P), recall (R), and mean Average Precision (mAP) as the evaluation metrics for comparing its performance with other models.

2　Construction of detection model

To further improve the detection accuracy of sporangium and conidiophores in microscopic images by using the YOLO v8s-OBB model, the original YOLO v8s-OBB network structure was modified, as shown in Fig. 3. By integrating the CBAM attention mechanism after the C2f module and introducing the OBB_LSCD in the detection head, the model's

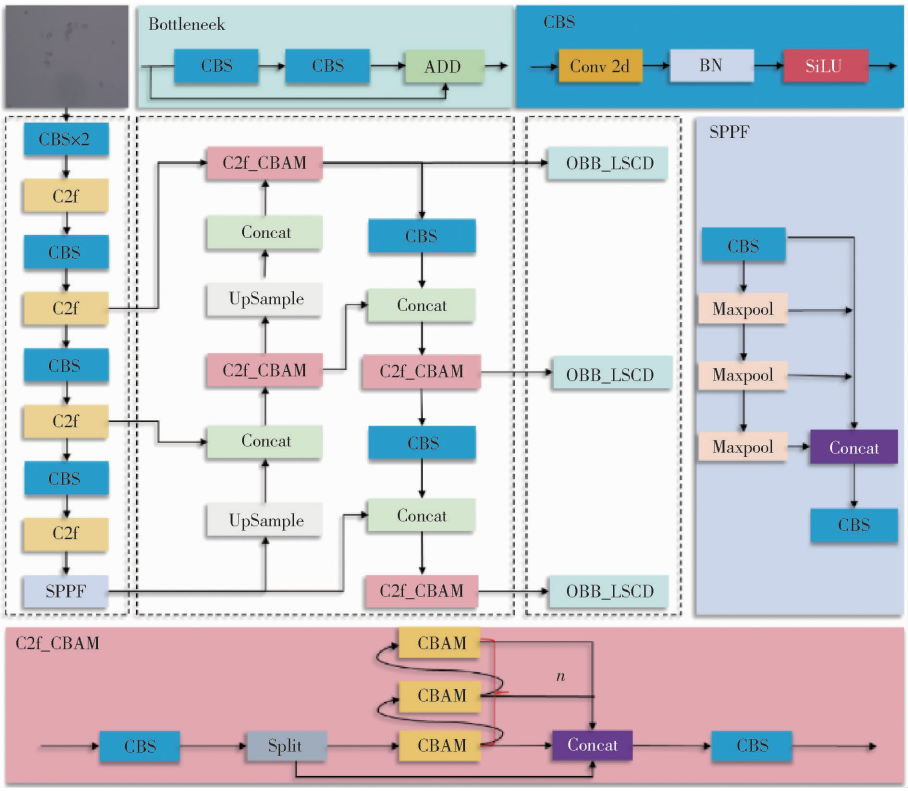


Fig. 3 Improved YOLO v8s-OBB network structure

performance in detecting sporangium and conidiophores in microscopic images was significantly enhanced. The C2f module was responsible for feature extraction and fusion, but it may need to be improved in emphasizing the features of small targets. The CBAM, as an advanced attention mechanism module, can adaptively learn the importance of spatial and channel features, thereby enhancing the model's feature representation capabilities. This allowed the network to focus more effectively on critical regions and features in microscopic images, improving detection accuracy and robustness, especially in complex backgrounds or overlapping sporangium and conidiophores. The OBB detection head enabled the network to output-oriented bounding boxes, which was particularly important when dealing with sporangium and conidiophores with varying orientations. This allowed for more precise encapsulation of the targets and reduced background noise interference. The LSCD, being a lightweight detection head design, significantly reduced the number of parameters and computational costs through shared convolution layers, thereby improving the model's inference speed and making it highly suitable for use in resource-constrained microscopic analysis environments. By incorporating the CBAM attention mechanism and the lightweight shared convolutional detection head into the original YOLO v8s-OBB model, not only did the model achieve precise detection of sporangium and conidiophores in microscopic images with complex backgrounds, but it also enhanced performance and efficiency in real-time detection applications, endowing the model with greater adaptability and robustness.

2.1 CBAM attention mechanism

The CBAM attention mechanism[25] mainly consisted of two sequential sub-modules: the channel attention module (CAM) and the spatial attention module (SAM). Its basic structure was shown in Fig. 4 CBAM sequentially integrated attention weights along the channel and spatial dimensions based on the input feature map and multiplied them with the input features to generate new features, which was beneficial for extracting meaningful information from the feature map. The introduction of the CBAM attention mechanism was illustrated in Fig. 3. CBAM was chosen over other attention mechanisms, such as SE or Non-local modules, due to its lightweight design and superior performance in enhancing the model's ability to focus on critical regions without significantly increasing computational overhead. CBAM effectively balanced detection accuracy and computational efficiency, making it especially suitable for detecting

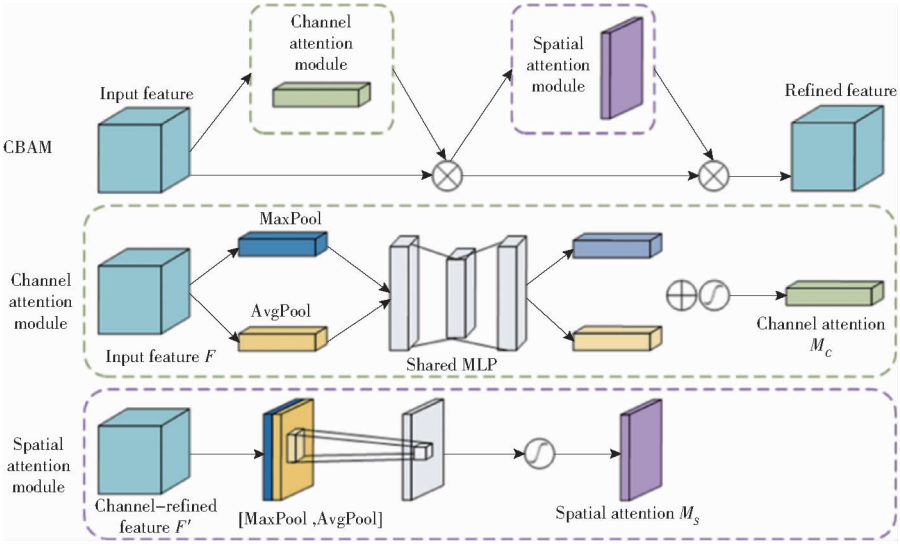


Fig. 4　Network structure diagram of CBAM

small and rotated objects like sporangia and conidiophores in microscopic images. Furthermore, CBAM's modular design ensured easy integration into existing deep learning architectures without introducing excessive complexity.

In this model, CBAM was inserted after the C2f module. The C2f module extracted and fused features from the input data. Still, it may need to fully emphasize small target features, critical for detecting minute objects such as sporangia and conidiophores. By introducing CBAM at this stage, the attention mechanism optimized feature representation and enhanced the model's ability to focus on the most critical regions in microscopic images. Specifically, after the C2f module processed and fused the features, CBAM applied channel and spatial attention to prioritize the areas most relevant for target detection, thus enabling the accurate detection of small and rotated objects in microscopic images. This strategic placement ensured that the model can handle complex backgrounds and overlapping targets effectively while maintaining high precision and recall. Following the CBAM module, the model proceeded to the lightweight shared convolutional detection head (LSCD), further improving multi-scale feature detection.

CBAM calculated the attention map for a feature map in two dimensions: channel attention and spatial attention. For an intermediate layer feature map, denoted as $F \in \mathbf{R}^{C\times H\times W}$, CBAM inferred a one-dimensional channel attention map and a two-dimensional spatial attention map. The calculation process was as follows:

$$F' = M_C(F) \otimes F \tag{1}$$

$$F'' = M_S(F') \otimes F' \tag{2}$$

where symbol $\otimes$ represented element-wise multiplication. The channel attention was multiplied by the input feature map F', and then the spatial attention of F' was calculated. Both of the above were multiplied to obtain the final output F''.

Channel attention mainly focused on which parts of the input microscopic images contained the conidiophores and sporangium of cucumber downy mildew. To effectively calculate the channel features, max pooling, and average pooling were used to compress the feature map along the spatial dimension, resulting in F^C_{Max} and F^C_{Avg}. Multi-layer fully connected network was then used to process these two different descriptive directions to calculate the channel attention, as shown in equations (3) and (4):

$$M_C = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{3}$$

$$M_C(F) = \sigma(W_1(W_0(F^C_{\text{Avg}})) + W_1(W_0(F^C_{\text{Max}}))) \tag{4}$$

where $W_0 \in \mathbf{R}^{C/r\times C}$ and $W_1 \in \mathbf{R}^{C\times C/r}$, the Sigmoid function was used as the activation function for W_0.

The spatial attention mainly focused on the positional information of the input image features. Firstly, max pooling and average pooling were applied along the channel dimension to obtain two different feature descriptors, $F^S_{\text{Max}} \in \mathbf{R}_{1\times H\times W}$ and $F^S_{\text{Avg}} \in \mathbf{R}_{1\times H\times W}$. These two features were then combined through an operation, and a convolution operation was used to

generate the spatial attention feature map $M_S(F) \in \mathbf{R}_{H\times W}$. The calculation process was as follows:

$$M_S(F) = \sigma(f^{7\times7}([AvgPool(f);MaxPool(F)])) \tag{5}$$

$$M_S(F) = \sigma(f^{7\times7}([F^S_{\text{Avg}};F^S_{\text{Max}}])) \tag{6}$$

where $f^{7\times7}$ represented the convolutional layer with a 7×7 convolution kernel.

2.2 Lightweight shared convolutional detection head

The lightweight shared convolutional detection head (LSCD) was an efficient and resource-friendly component of object detection models[26], as shown in Fig. 5. LSCD's design philosophy involved optimizing the detection performance for targets of different sizes through shared convolution operations while simultaneously reducing the demand for computational resources. LSCD can integrate multi-scale features, enabling the model to more accurately identify targets of various sizes.

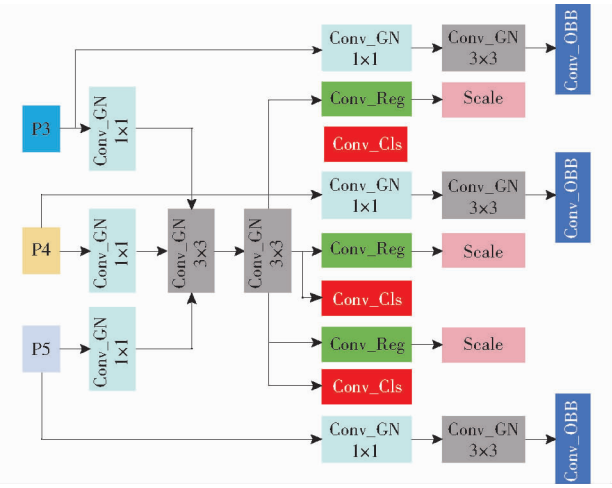


Fig. 5 Network structure diagram of LSCD

The original YOLO v8s-OBB model was improved to accurately detect the rotated and tilted sporangium and conidiophores of cucumber downy mildew in microscopic images. Adopting the LSCD, which integrated multi-scale features and shared convolution operations, optimized the detection performance for targets of different sizes. Meanwhile, by implementing 1×1 and 3×3 convolution operations on the P3, P4, and P5 feature layers of various scales, the model effectively reduced the demand for computational resources while improving target detection accuracy. In addition, each scale layer was equipped with specific regression and classification layers and a Conv_OBB layer explicit for oriented bounding boxes to ensure accurate prediction of the targets' shapes and orientations. By integrating LSCD into the detection head, the performance of the YOLO v8s-OBB model in detecting cucumber downy mildew sporangium was significantly enhanced, achieving high-precision detection of small targets with specific morphologies.

3 Results and analysis

3.1 Selection of base model

To achieve accurate detection of sporangium and conidiophores of cucumber downy mildew in microscopic images, several advanced rotational target detection models were compared, and the resultswere shown in Table 1. According to the results in Table 1, the YOLO v8s-OBB model demonstrated outstanding performance across multiple key metrics. Compared with other models, YOLO v8s-OBB had the highest precision, reaching 94.9%, which was higher than YOLO v7s-OBB's 94.6% and YOLO v5s-OBB's 94.1%, with improvements of 0.3 percentage points and 0.8 percentage points, respectively. In contrast, the precision of S2ANet, H2RBox, and R2CNN was 91.2%, 93.1%, and 90.2%, respectively, making YOLO v8s-OBB higher by 3.7 percentage points, 1.8 percentage points, and 4.7 percentage points compared with that of these models. Regarding recall, YOLO v8s-OBB's score of 85.9% was slightly lower than H2Rbox's 87.2%, but it was 0.8 percentage points and 1.6 percentage points higher than that of S2ANet and R2CNN, respectively. For mean average precision (mAP@0.5), YOLO v8s-OBB reached 95.3%, matching YOLO v7s-OBB but surpassing YOLO v5s-OBB by 0.3 percentage points, and it was significantly higher than S2ANet's 92.3% and R2CNN's 91.1%. Regarding mAP@0.5:0.95, YOLO v8s-OBB achieved 69.1%, slightly higher than YOLO v7s-OBB's 68.7% and YOLO v5s-OBB's 68.3% and better than S2ANet's 65.1% and R2CNN's 64.2%. Regarding computational complexity, YOLO v8s-OBB had 2.96×10^{10} FLOPs, significantly lower than S2ANet's 1.31×10^{11} FLOPs and R2CNN's 1.21×10^{11} FLOPs, demonstrating greater computational efficiency. Additionally, the YOLO v8s-OBB model size was only 11.4 MB, the smallest among all models, making it suitable for environments with limited storage space. These advantages mainly stemed from the

Table 1　Comparative analysis of various advanced models

Model	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%	FLOPs	Model size/MB
S2ANet	91.2	85.1	92.3	65.1	1.31×10^{11}	150.3
H2RBox	93.1	87.2	94.1	67.2	1.11×10^{11}	145.2
R2CNN	90.2	84.3	91.1	64.2	1.21×10^{11}	160.1
YOLO v5s-OBB	94.1	85.7	95.0	68.3	2.83×10^{10}	14.5
YOLO v7s-OBB	94.6	86.1	95.3	68.7	3.11×10^{10}	12.4
YOLO v8s-OBB	94.9	85.9	95.3	69.1	2.96×10^{10}	11.4

efficient architecture design and optimized algorithms of YOLO v8s-OBB, which significantly reduced computational complexity and model size while maintaining high precision. In contrast, S2ANet and R2CNN had higher computational resource consumption, while the YOLO series models achieved a balance of performance and efficiency through network structure and loss function optimization. Considering precision, recall, mean precision, computational complexity, and model size, YOLO v8s-OBB was the best-performing model, making it a suitable base model for further improvements and optimizations. Its balanced accuracy and resource efficiency performance make it an ideal choice for applications in real-time and resource-constrained environments.

3.2　Ablation experiment

Ablation experiment was conducted to further evaluate the impact of differentnetwork structures of YOLO v8-OBB on the detection accuracy of sporangium and conidiophores in microscopic images. As shown in Table 2, the YOLO v8-OBB model exhibited different performance characteristics across various network structures (n, s, m, l, x). YOLO v8s-OBB performed best in terms of precision, achieving a value of 94.9%, which surpassed other structures, notably exceeding the precision of YOLO v8m-OBB at 88.4% and YOLO v8l-OBB at 91.1%. Although YOLO v8x-OBB slightly led in mAP@0.5 with a score of 95.5%, YOLO v8s-OBB follows closely with a score of 95.3%, demonstrating near-optimal performance. Regarding mAP@0.5:0.95, YOLO v8x-OBB reached 70.9%, the highest value, while YOLO v8s-OBB followed with a score of 69.1%, reflecting its strong detection capabilities. Regarding computational complexity, YOLO v8s-OBB showed significant efficiency, with only 2.96×10^{10} FLOPs, much lower than the larger models like YOLO v8x-OBB, which required 2.64×10^{11} FLOPs, indicating its efficiency in resource usage. Additionally, the model size of YOLO v8s-OBB was only 11.40 MB, considerably smaller than YOLO v8x-OBB's 69.50 MB, making YOLO v8s-OBB more suitable for deployment in resource-constrained environments. The influence of different network depths on the results was evident. Larger models, such as YOLO v8x-OBB, although slightly advantageous in some performance metrics, were constrained by their high computational complexity and larger model size. Therefore, considering the balance between performance and resource usage, YOLO v8s-OBB was the best base model for further improvements. It offerred high-performance detection while maintaining low computational complexity and model compactness.

Table 2　Comparative analysis of different network architectures

Model	Precision/%	Recall/%	mAP@0.5/%	mAP@0.5:0.95/%	FLOPs	Model size/MB
YOLO v8m-OBB	88.4	89.3	94.4	70.1	8.12×10^{10}	26.40
YOLO v8s-OBB	94.9	85.9	95.3	69.1	2.96×10^{10}	11.40
YOLO v8l-OBB	91.1	88.4	93.3	69.9	1.69×10^{11}	44.48
YOLO v8x-OBB	90.1	89.4	95.5	70.9	2.64×10^{11}	69.50
YOLO v8n-OBB	89.9	85.6	92.9	66.3	8.40×10^{9}	3.10

In the ablation study, the detection performance of the YOLO v8s-OBB model under different configurations was evaluated to analyze the impact of each component on the overall model performance. The experimental results were shown in Table 3. The base model used only the YOLO v8s-OBB without the LSCD and CBAM modules. The results showed a precision of 94.9%, recall of 85.9%, and mean average precision

Table 3 Results of ablation experiments

Number	YOLO v8s-OBB	LSCD	CBAM	Precision/%	Recall/%	mAP@0.5/%	FLOPs	Model size/MB
1	√	×	×	94.9	85.9	95.3	2.96×10^{10}	11.4
2	√	√	×	93.4	89.7	95.0	2.62×10^{10}	9.5
3	√	×	√	95.6	86.0	96.4	2.96×10^{10}	11.4
4	√	√	√	96.0	90.1	96.7	2.62×10^{10}	9.5

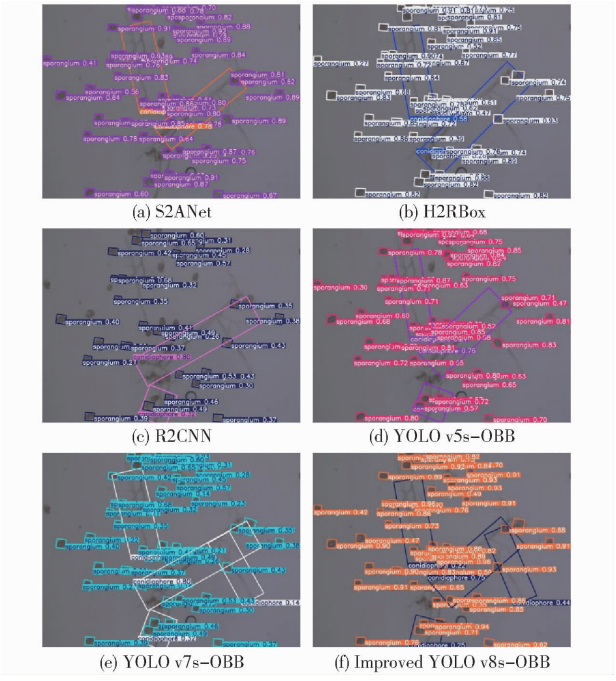
Note: √ represented the adoption of this method, while × represented the non-adoption of this method.

(mAP @ 0.5) of 95.3%. The computational complexity of this configuration was 2.96×10^{10} FLOPs, and the model size was 11.4 MB, indicating that the base model already possessed high detection capability and efficiency. When the LSCD was introduced, the recall rate was increased significantly to 89.7%, although precision was decreased to 93.4%, and mean average precision dropped slightly to 95.0%. However, the computational complexity was decreased to 2.62×10^{10} FLOPs, and the model size was reduced to 9.5 MB. This demonstrated that the LSCD module effectively reduced the model's computational resource requirements while improving recall, making it more suitable for resource-constrained environments. Adding the CBAM to the YOLO v8s-OBB model results in a precision increase to 95.6%, a slight increase in recall to 86.0%, and a significant improvement in mean average precision to 96.4%. The computational complexity remained unchanged at 2.96×10^{10} FLOPs, and the model size stayed at 11.4 MB. This indicated that the CBAM module enhanced the model's ability to recognize key features, thereby improving detection accuracy. When both LSCD and CBAM were applied in the YOLO v8s-OBB model, the results showed the highest precision at 96.0%, an increase in recall to 90.1%, and a further improvement in mean average precision to 96.7%. The computational complexity was 2.62×10^{10} FLOPs, and the model size was 9.5 MB. The combination of these two modules maintained high precision while improving recall and effectively reduced computational complexity and model size, demonstrating that this combination achieved a good balance between enhancing model performance and resource efficiency. In summary, the ablation study showed that the CBAM and LSCD modules each contributed uniquely to improving the model's performance, and their combined use maximized detection performance, especially in resource-constrained environments, showing high application potential.

The experimental results on the test set were shown in Table 4, demonstrating that the improved YOLO v8s-OBB model had significantly enhanced detection performance compared with the original model. The improved model's precision was increased to 95.6%, compared with the original model's 92.5%, reducing false positives and allowing for more accurate identification of conidiophores and sporangium of cucumber downy mildew. The recall was also improved to 89.6%, up from 86.9% in the original model, indicating that the improved model can detect more targets and reduce the number of missed detections. For mean average precision (mAP @ 0.5), the improved model achieved 96.5%, higher than the original model's 94.9%, indicating high consistency and accuracy when detecting different targets. Under the more stringent mean average precision (mAP @ 0.5:0.95), the improved model scored 70.1%, while the original model achieved 69.1%, validating its robustness across different IoU thresholds, particularly in detecting targets within complex backgrounds and diversified scenarios. In summary, the improved YOLO v8s-OBB model outperformed the original model across key performance metrics, including precision, recall, and mean average precision. The enhanced ability to detect targets within complex backgrounds was particularly noteworthy. Furthermore, Fig. 6 presented a performance comparison of different models in detecting conidiophores and sporangia of cucumber downy mildew. In the S2ANet, H2RBox, and R2CNN models, severe cases of missed detections and repeated detections of conidiophores and sporangia can be observed, indicating limitations in precise localization. Compared with these models, the YOLO v5s-OBB and YOLO v7s-OBB models showed some improvement in bounding box accuracy, but they still struggled with identifying smaller or heavily overlapping targets. Some cases of missed detections for individual conidiophores

Table 4　Results of test set　　%

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLO v8s-OBB	92.5	86.9	94.9	69.1
Improved YOLO v8s-OBB	95.6	89.6	96.5	70.1

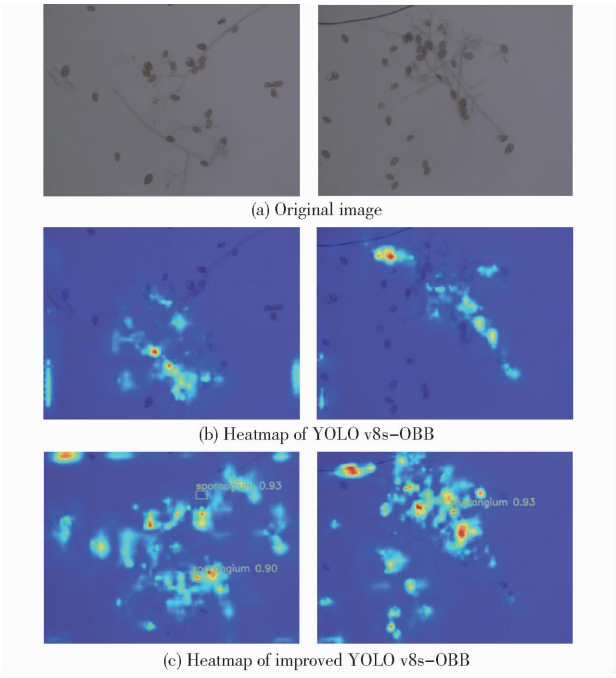


(a) S2ANet　(b) H2RBox　(c) R2CNN　(d) YOLO v5s-OBB　(e) YOLO v7s-OBB　(f) Improved YOLO v8s-OBB

Fig. 6　Comparative analysis of detection results from different advanced models

and sporangia were also presented. In contrast, the improved YOLO v8s-OBB model outperformed all the displayed models, achieving the most accurate bounding box positioning and effectively distinguishing multiple closely located targets. Each target's bounding box was clear and precisely positioned, demonstrating high accuracy with minimal overlap. This performance proved that through refined feature extraction and optimized bounding box processing mechanisms, the practical utility and efficiency of the YOLO v8s-OBB model were significantly enhanced, making it more competitive and higher-performing in multi-target detection tasks within complex backgrounds.

Additionally, Fig. 7 showed the heatmaps generated by the original YOLO v8s-OBB model and the improved YOLO v8s-OBB model for detecting the sporangia and conidiophores of cucumber downy mildew. In the heatmap of the improved YOLO v8s-OBB model, the warmer tones (yellow to red) highlighted the areas that the model considered most likely to contain the target objects. With the introduction of the CBAM attention mechanism and the LSCD, the improved model focused more on these areas of interest, displaying more concentrated hot zones. The colors around each target were more focused and well-defined, demonstrating the model's excellent performance in accurately identifying and analyzing specific structures in microscopic images.



(a) Original image　(b) Heatmap of YOLO v8s-OBB　(c) Heatmap of improved YOLO v8s-OBB

Fig. 7　Comparison of heatmaps before and after YOLO v8s-OBB improvement

3.3　Quantitative detection system for sporangium and conidiophores of cucumber downy mildew

Quantitative detection system was proposed based on the improved YOLO v8s-OBB algorithm for identifying and counting the conidiophores and sporangium of cucumber downy mildew (Fig. 8). The system leveraged the efficient detection capabilities of the deep learning model to perform real-time detection of targets in static images and dynamic videos. The system firstly loaded the best weight file obtained from training the improved YOLO v8s-OBB model and then performed inference on the input images or videos. The model can effectively filter and accurately identify target objects by adjusting parameters such as confidence (conf) and intersection over union (IoU). In addition to drawing detection boxes and annotating labels, the model can also count the

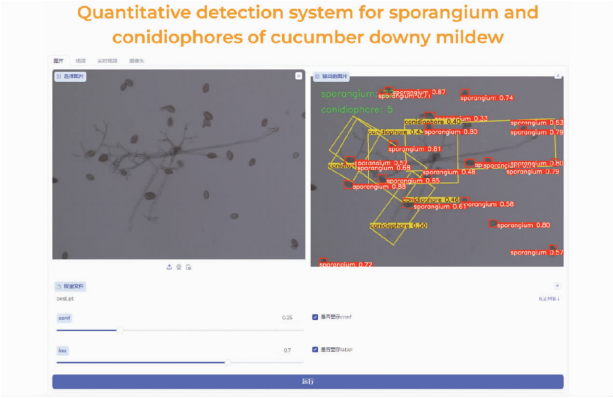


Fig. 8 Quantitative detection system for sporangium and conidiophore of cucumber downy mildew

number of target categories based on the detection results and display the instance count for each category on the image through a counter, thereby providing an effective tool for the quantitative analysis of cucumber downy mildew. The system supported processing static images and videos, and it can obtain real-time data streams through a camera, enabling real-time updates of detection results and making it suitable for online monitoring in actual production environments. The user interaction interface was built on the Gradio framework, allowing users to upload images and video files or conduct real-time detection through a camera and display the detection results. Users can adjust detection parameters as needed to meet the requirements of different environments. This system offerred a convenient and efficient solution for disease detection in the agricultural field, helping growers to monitor crop health in real-time, take timely control measures, and improve agricultural production efficiency.

4 Conclusion

(1) To meet the requirements for a lightweight network structure and real-time detection in the quantitative detection of conidiophores and sporangium of cucumber downy mildew, the introduction of the CBAM and the LSCD significantly enhanced the model's focus on essential regions in microscopic images, thereby improving its target detection capabilities. At the same time, it achieved lower computational complexity and model size, making it suitable for deployment in resource-constrained environments.

(2) To validate the effectiveness of the improved YOLO v8s-OBB model, the detection accuracy of the improved YOLO v8s-OBB model was significantly superior to that of traditional horizontal bounding box detection models such as S2ANet, H2RBox, and R2CNN. On the test set, the improved YOLO v8s-OBB model achieved increases in precision, recall, and mAP @ 0.5 by 3.1 percentage ponits, 2.7 percentage ponits, and 1.6 percentage ponits, respectively compared with the original YOLO v8s-OBB model. This demonstrated the improved YOLO v8s-OBB model ' s adaptability and effectiveness in handling targets with complex backgrounds and varying orientations.

(3) The proposed quantitative detection system for the conidiophores and sporangium of cucumber downy mildew can effectively identify and count these structures, providing a convenient and efficient solution for disease detection in the agricultural sector. This system enabled growers to monitor crop health in real-time and take timely control measures, thereby advancing preventive measures and improving disease management.

References

[1] MA Juncheng, WEN Haojie, LI Xinxing, et al. Downy mildew diagnosis system for greenhouse cucumbers based on image processing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(2): 195 – 202. (in Chinese)
[2] QIN Lifeng, ZHANG Xi, ZHANG Xiaoqian. Early detection of cucumber downy mildew in greenhouse by hyperspectral disease differential feature extraction[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(11): 212 – 220. (in Chinese)
[3] SINGH A, JONES S, GANAPATHYSUBRAMANIAN B, et al. Challenges and opportunities in machine-augmented plant stress phenotyping[J]. Trends in Plant Science, 2021, 26(1): 53 – 69.
[4] PETRIACQ P, STASSEN J H M, TON J. Spore density determines infection strategy by the plant pathogenic fungus plectosphaerella cucumerina[J]. Plant Physiology, 2016, 170(4): 2325 – 2339.
[5] YOON S C, LAWRENCE K C, PARK B. Automatic counting and classification of bacterial colonies using hyperspectral imaging [J]. Food and Bioprocess Technology, 2015, 8(10): 2047 – 2065.
[6] WALLACE E C, DARCANGELO K N, QUESSADA-OCAMPO L M. Population analyses reveal two Host-Adapted clades of pseudoperonospora cubensis, the causal agent of cucurbit downy mildew, on commercial and wild cucurbits [J]. Phytopathology, 2020, 110(9): 1578 – 1587.
[7] ANDERSSON M A, SALO J, KEDVES O, et al. Bioreactivity, guttation and agents influencing surface tension of water emitted by actively growing indoor mould isolates[J]. Microorganisms, 2020, 8(12): 1940.
[8] MA Chaowei, ZHANG Hao, MA Xinming, et al. Method for the lightweight detection of wheat disease using improved YOLOv8 [J]. Transactions of the CSAE, 2024, 40(5): 187 – 195. (in Chinese)

[9] PAWAR S, SHEDGE S, PANIGRAHI N, et al. Leaf disease detection of multiple plants using deep learning[C]//2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM - IT - CON). IEEE, 2022: 241 - 245.
[10] FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Detection system for grape leaf diseases based on transfer learning and updatedCNN[J]. Transactions of the CSAE, 2021, 37(6): 151 - 159. (in Chinese)
[11] SHARMA P, BERWAL Y, GHAI W. Performance analysis of deep learning CNN models for disease detection in plants using image segmentation[J]. Information Processing in Agriculture, 2020,7: 566 - 574.
[12] JEEVANANTHAM R, VIGNESH D, ABDUL R A, et al. Deep learning based plant diseases monitoring and detection system[C]// 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS). IEEE, 2023: 360 - 365.
[13] QIAO Chen, HAN Mengyao, GAO Wei, et al. Quantitative detection of cucumber downy mildew spores at multi-scale based on Faster - NAM - YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 288 - 299. (in Chinese)
[14] YAO Y, CHENG G, WANG G, et al. On improving bounding box representations for oriented object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1 - 11.
[15] WANG J, DING J, GUO H, et al. Mask OBB:a semantic attention-based mask oriented bounding box representation for multi-category object detection in aerial images[J]. Remote Sensing, 2019, 11: 2930.
[16] SONG Huaibo, JIAO Yitao, HUA Zhixin, et al. Endosperm crack detection method for seed dipping maize based on YOLO v5-OBB and CT technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 394 - 401, 439. (in Chinese)
[17] GENG Yanli, LIN Yanbo, FU Yanfang, et al. Object detection algorithm for pigs based on dual dilated layer and rotary box location[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 323 - 330. (in Chinese)
[18] HE Y, GAO F, WANG J, et al. Learning polar encodings for arbitrary-oriented ship detection in sar images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 3846 - 3859.
[19] CHENG Lang, JING Chao. X-ray image rotating object detection based on improved YOLOv7[J]. Journal of Graphics, 2023, 44(2): 324 - 334. (in Chinese)
[20] HE Linyuan, BAI Junqiang, HE Xu, et al. Sparse Transformer based remote sensing rotated object detection[J]. Laser and Optoelectronics Progress, 2022, 59(18): 55 - 63. (in Chinese)
[21] LEBEDA A, COHEN Y. Cucurbit downy mildew (*Pseudoperonospora cubensis*)—biology, ecology, epidemiology, host-pathogen interaction and control[J]. European Journal of Plant Pathology, 2011, 129: 157 - 192.
[22] COHEN Y, RUBIN A, GALPERIN M. Formation and infectivity of oospores of pseudoperonospora cubensis, the causal agent of downy mildew in cucurbits[J]. Plant Disease, 2011, 95(7): 874.
[23] NEUFELD K, ISARD S, OJIAMBO P. Relationship between disease severity and escape of *Pseudoperonospora cubensis* sporangia from a cucumber canopy during downy mildew epidemics[J]. Plant Pathology, 2013, 62: 1366 - 1377.
[24] GRANKE L, HAUSBECK M. Dynamics of pseudoperonospora cubensis sporangia in commercial cucurbit fields in michigan [J]. Plant Disease, 2011, 95(11): 1392 - 1400.
[25] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[J]. arXiv: 1807.06521, 2018.
[26] YUN J, JIANG D, LIU Y, et al. Real-time target detection method based on lightweight convolutional neural network[J]. Frontiers in Bioengineering and Biotechnology, 2022, 10: 861286.

基于改进 YOLO v8s-OBB 的黄瓜霜霉病菌分生孢子梗与孢子囊定量检测

张一丁[1]　乔　琛[2]　张领先[2]　韩宗桓[2]

(1. 中国农业大学工学院, 北京 100083; 2. 中国农业大学信息与电气工程学院, 北京 100083)

摘要：黄瓜霜霉病作为一种严重威胁黄瓜生产和品质的真菌性病害,其分生孢子梗和孢子囊的定量检测对病害防治关口前移具有重要意义,由于分生孢子梗及孢子囊存在形态多样和方向不同特征,传统的水平边界框检测方法无法准确检测。因此,本文提出了一种基于改进 YOLO v8s-OBB 的检测方法,通过引入卷积块注意力机制(Convolutional block attention module, CBAM)和轻量级共享卷积检测头(Lightweight shared convolutional detection head, LSCD)模块,旨在提高黄瓜霜霉病菌分生孢子梗及孢子囊的检测效率和准确性。通过引入 CBAM,增强了模型对关键特征的识别能力,使其更聚焦于显微图像中的重要区域,进而提高了对小目标的检测能力。LSCD 通过共享卷积操作集成多尺度特征,提高了模型对不同尺寸目标的检测性能,同时降低了计算成本,以适应资源受限环境的需求。旋转边界框技术能准确捕捉分生孢子梗及孢子囊的倾斜和旋转姿态。实验结果表明,对比原始 YOLO v8s-OBB 模型,改进 YOLO v8s-OBB 模型不仅模型尺寸进一步减少,而且对黄瓜霜霉病菌分生孢子梗及孢子囊的检测性能更优,其精确率、召回率和 mAP@0.5 分别达到 96.0%、90.1% 和 96.5%。同时,改进 YOLO v8s-OBB 模型检测精度均优于 S2ANet、H2RBox 和 R2CNN 等先进旋转目标检测模型。此研究验证了改进模型在实际应用中的有效性,并为黄瓜霜霉病的早期诊断提供了技术支持。

关键词：黄瓜霜霉病；分生孢子梗；孢子囊；YOLO v8s-OBB；CBAM；LSCD

基金项目：国家自然科学基金项目(62176261)