Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/108/2025.20803

A Lotus Leaf Pest and Disease Detection Model Based on Improved YOLOv8

Haonan Xu[1,a,*], Hansen Wang[1]

[1]*School of Mechanical Engineering, Suzhou University of Science and Technology, Changjiang Road 55, Suzhou, China*
a. 3293368036@qq.com
**corresponding author*

Abstract: Lotus diseases and pests pose a serious threat to the yield and quality of lotus seeds, hence effective detection of these diseases and pests is of significant importance for their control and prevention. To address this, we propose an improved YOLOv8-based lightweight detection model specifically designed for lotus diseases and pests. Initially, we replaced the convolutional modules in the YOLOv8 neck network with GSConv, and the C2f modules with VoV-GSCSP, forming a Slim-neck architecture. This modification maintains high recognition accuracy while significantly reducing the computational complexity of the model. Furthermore, we introduced the C2f_EMA module, which integrates the EMA multi-scale attention mechanism, to replace the C2f modules in the backbone network, thereby enhancing the model's ability to extract features of lotus diseases and pests under complex environmental conditions. Experimental results demonstrate that the proposed improved YOLOv8 model excels in the task of detecting lotus diseases and pests, achieving an average precision (AP) of 89.3%, which is a 1.6 percentage point increase over the baseline model. Additionally, the model has 0.2M fewer parameters and is contained within a size of 5.6MB.

Keywords: Lotus leaf, Pest and disease detection, YOLOv8, Attention mechanism

1. Introduction

Lotus is an aquatic plant with a cultivation history of over 3000 years in China. The seeds are rich in starch, protein, minerals, and bioactive compounds, holding high nutritional and medicinal value with significant market demand. During cultivation, lotus leaves are susceptible to diseases and pests such as rot[1], leaf spot[2], viral diseases[3], and Spodoptera litura[4], which can lead to significant crop loss. Automated detection of these diseases and pests is crucial for enhancing lotus seed quality and yield, thus promoting industrial development.

Currently, detection relies heavily on manual expertise, which is subjective, inefficient, and prone to missed or incorrect detections. Hence, the development of an automated detection method suitable for lotus fields is of great importance.

Traditional automated detection methods rely on manual feature extraction, which is time-consuming and limited in precision and robustness under varying environmental conditions. In recent years, deep learning-based object detection algorithms have shown promise in detecting plant diseases and pests. The YOLO series, known for its rapid detection capabilities, has garnered attention. Hu Gensheng et al. Literature[5] improved the YOLOv5 model, achieving a mAP of 92.89% for

DOI: 10.54254/2755-2721/108/2025.20803

teatea geometrid (a type of pest) detection. Yao Lingyun et al.[6], using YOLOv5su, incorporated self-attention modules and lightweight structures to reach an mAP of 80.4%. Liu Shiyi et al.[7] enhanced the YOLOv7 model for high mAP detection of cucumber leaf diseases. However, these improvements often increase model complexity and parameter count, hindering mobile deployment.

Addressing these issues, this study presents a lightweight lotus disease and pest detection model based on YOLOv8. By optimizing the neck network and integrating attention mechanisms into the C2f modules of the backbone, the model achieves improved detection accuracy and reduced computational complexity.

2. Materials and Methods

2.1. Lotus Leaf Disease and Pest Image Collection

This study focuses on four common types of lotus leaf diseases and pests: rot, leaf spot, viral disease, and Spodoptera litura. Image collection was conducted at multiple lotus planting bases in Xuanwu District and Jiangning District, Nanjing City, Jiangsu Province, during July to August 2023 and 2024. We utilized smartphones such as Samsung Galaxy S21, Google Pixel 5, OnePlus 9, and Xiaomi Mi 11, capturing images at various resolutions (e.g., 3500×2800, 3000×2400, 5000×3500 pixels) under diverse weather conditions. The collection process accounted for different times of day and angles, resulting in 2600 images of diseases and pests, ensuring sample diversity and uniformity.

2.2. Dataset Construction

From the lotus disease and pest images, 1800 were selected for model training and augmented to 3600 through data augmentation, including flipping, translation, and adjustments in brightness and contrast. The augmented images were divided into training and validation sets in an 8:2 ratio. The remaining 800 images formed an independent test set, with 200 images per disease and pest category.

Annotation was performed using LabelImg software. Labels included "Rhizome rot" for rot, "Viral disease" for viral disease, "Leaf spot" for leaf spot, and "Spodoptera litura" for Spodoptera litura. Annotation files were saved in *.txt format, detailing the type of disease or pest and the location of the bounding boxes. Some images contained multiple annotations for various diseases and pests. The specific quantities of the four categories of diseases and pests in the dataset are presented in Table 1.

Table 1: Distribution of Lotus Disease and Pest Images in Dataset

Category	Total Images	Training Set (After Augmentation)	Validation Set (After Augmentation)	Test Set
Rhizome Rot	450	720	180	200
Viral Disease	450	720	180	200
Leaf Spot	450	720	180	200
Spodoptera litura	450	720	180	200
Total	1800	2880	720	800

3. Model and Training

3.1. YOLOv8 Network Structure

YOLOv8, released by Ultralytics in 2023, is a deep learning object detection algorithm that inherits the strengths of its predecessors, YOLOv5 and YOLOv7, demonstrating superior performance. The

YOLOv8 model consists of four main components: the input (Input), backbone (Backbone), neck (Neck), and detection head (Head).

The Input module is responsible for image preprocessing, enhancing processing efficiency through adaptive scaling. The Backbone is composed of Conv, C2f, and SPPF modules, which are utilized for feature extraction. The C2f module includes two Conv layers and several Bottleneck structures to capture rich gradient information; the SPPF module employs three consecutive max pooling operations to gather information about objects at various scales, thereby enhancing detection accuracy. The Neck integrates the PAN (Path Aggregation Network) and FPN (Feature Pyramid Network) structures, merging feature maps from different levels and scales to ensure the model's capability to extract features in multi-scale scenarios. The Head uses a decoupled head and anchor-free strategy, with three decoupled heads operating at different scales to complete detection and classification tasks, outputting the class and location information of the targets

3.2. Improved YOLOv8 Model Network Structure

Compared to other agricultural product recognition targets, lotus leaf disease and pest targets exhibit particular characteristics: 1) Lotus leaves grow in unstructured environments such as muddy fields and swamps, where disease and pest features are similar to surrounding environmental features, with a broad range of scales; 2) There is occlusion between lotus leaves, and the varying lighting conditions of the growth environment increase the difficulty of detection. Additionally, intelligent control of lotus leaf diseases and pests is often conducted in the field, requiring model deployment on mobile devices, which imposes strict requirements on model size and computational scale.

To address these issues, this study made improvements to the YOLOv8 (nano) model: 1) A lightweight design was applied to the Neck section to reduce computational scale while maintaining detection accuracy; 2) A multi-scale attention mechanism was introduced to enhance the model's ability to focus on features and minimize the impact of interfering factors. The improved YOLOv8 model structure is shown in Figure 1.

3.2.1. Slim-neck Module

In natural environments, the detection of lotus leaf diseases and pests demands high precision and speed from the model, affecting its operational capability and deployability. An increase in model parameter count and size leads to increased deployment difficulty and cost. Therefore, the improved YOLOv8 model incorporates a Slim-neck lightweight structure composed of GSConv and VoV-GSCSP modules.

In the Slim-neck, GSConv replaces the Conv modules in the neck network. GSConv combines SC, DSC, and shuffle operations to maximize the advantages of DSC and mitigate its negative impacts. GSConv first performs SC and DSC operations, then concatenates the feature maps, and finally reorganizes the channels through the shuffle operation, reducing computational scale while maintaining detection performance.

The VoV-GSCSP module further reduces computational costs and balances detection accuracy. Its feature extraction process involves two paths: one through standard convolution and the other through the GSBottleneck structure designed with GSConv, with the outputs concatenated afterward.

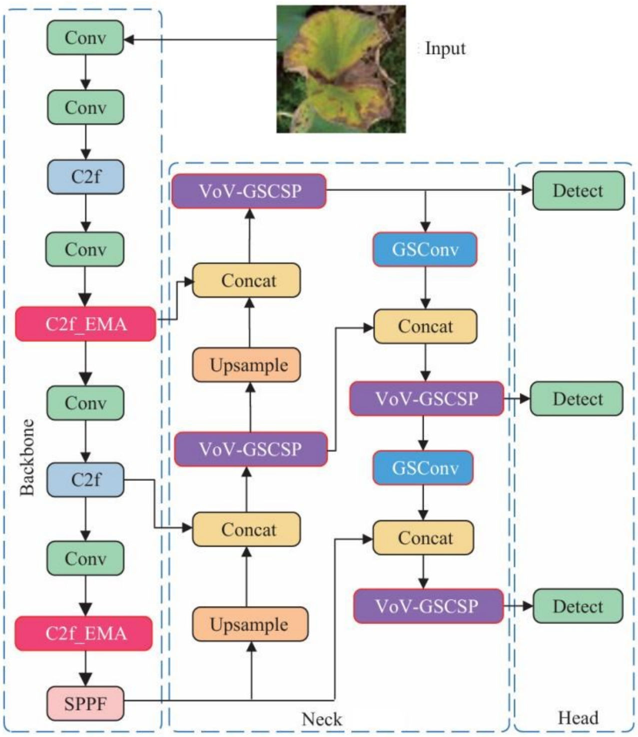

Figure 1: Our improved YOLOv8 model

3.2.2. EMA Multi-scale Attention Mechanism

In the detection of lotus leaf diseases and pests, occlusion and varying lighting conditions can lead to missed detections; the similarity between disease and pest features and the environment increases the risk of false positives. To enhance detection capabilities, this study added the EMA attention mechanism to the YOLOv8 backbone network. EMA is an efficient multi-scale attention mechanism based on cross-scale learning, which reshapes and groups channels to prevent information loss and reduce computational overhead, characterized by high precision and fewer parameters.

The EMA module's workflow includes: dividing the input $X \in \mathbb{R}^{C \times H \times W}$ into G sub-features along the channel dimension, and then extracting attention weight descriptors through three routes. The 1×1 branch uses one-dimensional global average pooling for cross-channel information interaction; the 3×3 branch omits pooling and GroupNorm to achieve multi-scale feature representation. Ultimately, EMA captures pixel-level relationships and acquires global contextual information through the Sigmoid function.

The EMA mechanism is integrated into the Bottleneck of C2f, forming the C2f_EMA module, which optimizes the Bottleneck structure and enhances the capability to capture multi-scale features. The second and fourth C2f modules in the YOLOv8 backbone network are replaced with C2f_EMA modules, strengthening the model's ability to extract and filter features of diseases and pests against complex backgrounds.

3.3. Model Training and Evaluation Metrics

3.3.1. Model Training Platform and Parameter Settings

The hardware configuration for the model training and testing platform in this study includes: 12th generation Intel Core i7-12700K CPU, NVIDIA GeForce RTX 4090 GPU, and 24 GB of video memory. The software configuration comprises the Windows 10 operating system, PyTorch 2.4.1

DOI: 10.54254/2755-2721/108/2025.20803

deep learning framework, Python 3.8 programming language, and the Pycharm integrated development environment. During training, the input image size to the model is set at 640 pixels × 640 pixels, with a batch size of 32, 120 training epochs, an initial learning rate of 0.01, and all other parameters set to default.

3.3.2. Evaluation Metrics

This study employs precision (P), recall (R), and mean average precision (mAP@0.5) to evaluate the model's detection accuracy, with the mAP threshold set at 0.5. The number of parameters is selected to assess the computational scale of the model. The model weight size is chosen as an indicator of deployability. The frame rate (FPS) is used to evaluate the model's real-time detection capability.

4. Results and Analysis

The improved YOLOv8 was compared with other mainstream models, including Faster R-CNN, SSD, YOLOv3 through YOLOv9, as shown in Table 2. The improved YOLOv8 showed superior performance in terms of precision and mAP, with fewer parameters and a smaller model size. Compared to these models, mAP was enhanced by 3.8 to 30.7 percentage points, parameter count was reduced by 1.0 to 61M, and model size was decreased by 8.0 to 238.4MB.

Table 2: Comparison of Detection Performance and Resource Consumption of Mainstream Models

Models	Parameters/M	Model size/MB	P/%	R/%	mAP/%
Faster R-CNN	41.3	108.0	64.3	85.1	85.5
SSD	23.8	92.1	88.5	80.1	79.0
YOLOv3	61.5	235.0	82.7	66.4	75.8
YOLOv4	63.9	244.0	80.9	40.5	58.6
YOLOv5	7.0	13.6	86.9	76.5	84.0
YOLOv7	36.5	71.3	81.1	80.5	83.0
YOLOv8	3.0	6.0	88.6	80.7	87.7
YOLOv9	9.6	19.3	84.2	79.8	84.4
Improved YOLOv8	**2.8**	**5.6**	**92.4**	**81.8**	**89.3**

In actual detection, YOLOv8 and YOLOv5 misclassified normal leaves moved by wind as viral disease leaves, while YOLOv4 and YOLOv3 misclassified viral disease as viral disease. YOLOv5 failed to detect the target in viral disease detection, and YOLOv5, YOLOv4, and YOLOv3 missed detections in leaf spot disease, while YOLOv9, YOLOv7, and SSD missed detections of Spodoptera litura. These phenomena are attributed to the models' insufficient feature extraction and resistance to interference, highlighting the importance of the attention mechanism.

The improved YOLOv8 demonstrated excellent performance in lotus leaf disease and pest detection, maintaining high detection accuracy while being lightweight, offering advantages in field detection scenarios.

5. Conclusion

This study proposes a lightweight lotus leaf disease and pest detection model based on the improved YOLOv8, establishing a dataset that considers various environmental conditions to provide a new approach for intelligent detection and on-site deployment of lotus leaf diseases and pests. The main conclusions are as follows:

DOI: 10.54254/2755-2721/108/2025.20803

1) The improved YOLOv8 model incorporates lightweight convolutional modules, GSConv, and VoV-GSCSP, to construct the Slim-neck architecture. The EMA attention mechanism is introduced, forming the C2f_EMA module, which replaces the C2f modules in the Backbone. This enhances the model's ability to extract and integrate multi-scale features while maintaining high accuracy and reducing computational complexity. The model achieves precision, recall, and mAP of 92.4%, 81.8%, and 89.3%, respectively, enabling accurate identification.

2) Compared to other models such as Faster R-CNN, SSD, YOLOv3 through YOLOv9, the improved YOLOv8 model demonstrates higher mAP, fewer parameters, and a smaller model size. It maintains high precision while being lightweight, showing superior performance in complex background detection scenarios.

The improved YOLOv8 model shows great potential for accurately identifying lotus leaf diseases and pests, offering technical support for automatic prevention and control measures.

References

[1] Deng Xu, Tan Jicai, Huang Youxin. Preliminary Investigation on Lotus Root Diseases and Pests in Xiangtan [J]. Journal of Anhui Agricultural Sciences, 2007, 35(3): 768-769.

[2] Zou Huajiao. Comparison of the Control Effects of Four Pesticides on Lotus Leaf Spot Disease [J]. China Plant Protection News, 2014, 34(4): 64-66.

[3] Liu Xian, Wu Weiwen, He Zhen, et al. Research Progress on the Occurrence and Prevention of Common Diseases and Pests in Lotus [J]. Jiangsu Journal of Agricultural Science, 2017, 45(16): 24-28.

[4] Wei Lin, Liang Zhihuai. Identification and Integrated Control of Lotus Diseases, Pests, and Weeds [M]. Beijing: China Agricultural Science and Technology Press, 2013.

[5] Hu Gensheng, Wu Jitian, Bao Wenxia, et al. Detection of Ectropis oblique in Complex Background Images Based on the Improved YOLOv5 Network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(21): 191-198.

[6] Yao Lingyun, Zhou Junfeng, Li Li. Detection of Peach Leaf Curl Disease Using the Improved YOLOv5su Model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(14): 109-117.

[7] Liu Shiyi, Hu Bin, Zhao Chun. Detection and Identification of Cucumber Leaf Diseases Based on the Improved YOLOv7 [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(15): 163-171.