

REAL-TIME GRAPE DISEASE DETECTION MODEL BASED ON IMPROVED YOLOv8s

/ 基于改进 YOLOv8s 的实时葡萄病害检测模型

Jinglong REN¹⁾, Huili ZHANG^{*1)}, Guangyuan WANG¹⁾, Chenlong DAI¹⁾, Fei TENG¹⁾, Moxian LI¹⁾¹⁾College of Mechanical and Electrical Engineering, Qingdao Agricultural University / China

Tel: +86-13605424499.; E-mail: zhhuili73@163.com

DOI: <https://doi.org/10.35633/inmateh-72-09>**Keywords:** MobileNetV3, YOLOv8s, Grape Leaves, Disease, Image Recognition**ABSTRACT**

This research is dedicated to enhancing the accuracy and processing speed of grape disease recognition. As a result, a real-time grape disease detection model named MSCl-YOLOv8s, based on an improved YOLOv8s framework is proposed. The primary innovation of this model lies in replacing the backbone network of the original YOLOv8s with the more efficient MobileNetV3. This alteration not only strengthens the ability of the model to capture features of various disease manifestations in grape leaf images but also improves its generalization capabilities and stability. Additionally, the model incorporates the SPPFCSPC pyramid pooling structure, which maintains the stability of the receptive field while significantly enhancing processing speed. The integration of the CBAM attention mechanism further accentuates the ability of the model to identify key features, substantially increasing the accuracy of disease detection. Moreover, the model employs Inner-SIoU as the loss function, optimizing the precision of bounding box regression and accelerating model convergence, thereby further enhancing detection efficiency. Rigorous testing has shown that the MSCl-YOLOv8s model achieves an impressive average precision (mAP) of 97.7%, with an inference time of just 37.2 milliseconds and a memory footprint of 39.3 MB. These advancements render the MSCl-YOLOv8s not only highly efficient but also extremely practical for real-time grape disease detection, meeting the actual demands of grape orchard disease identification and demonstrating significant potential for application.

摘要

本研究致力于提升葡萄病害识别的准确性与处理速度，因此提出了一款基于改进的 YOLOv8s 模型的实时葡萄病害检测模型，命名为 MSCl-YOLOv8s。该模型的核心创新在于将原 YOLOv8s 的主干网络替换为更高效的 MobileNetV3，此举不仅增强了模型捕捉葡萄叶片图像中不同尺度病害特征的能力，同时也提高了其泛化性与稳定性。此外，模型整合了 SPPFCSPC 金字塔池化结构，不仅保持了感受野的稳定性，还实现了处理速度的显著提升。引入的 CBAM 注意力机制进一步加强了模型对关键特征的识别，显著提高了病害检测的准确度。模型还采用了 Inner-SIoU 作为损失函数，以优化边界框回归的精度并加快模型收敛，从而进一步提升了检测效率。经过严格测试，MSCl-YOLOv8s 模型实现了 97.7% 的平均精度 (mAP)，推理时间仅需 37.2 毫秒，内存占用为 39.3MB，表现卓越。这些改进使得 MSCl-YOLOv8s 在实时葡萄病害检测方面不仅效率高，且具备强大的实用性，完全符合葡萄果园病害识别的实际需求，展现了巨大的应用前景。

INTRODUCTION

Crop health monitoring plays a crucial role in ensuring the yield and quality of agricultural products and is an indispensable part of the development of precision agriculture (Sishodia et al., 2020; Sharma et al., 2020; Farooq et al., 2019). Grapes, in particular, as a widely cultivated economic crop globally, are not only nutritionally rich but also have significant medicinal value (Georgiev et al., 2014; Ananga et al., 2013; Perestrelo et al., 2014). However, crops like grapes face attacks from various diseases every year, leading to severe yield losses and posing a significant threat to global food security (Nicholas et al., 2012). Therefore, developing a method for timely and accurate identification of crop diseases is particularly important. This not only reduces the yield loss caused by diseases but also helps to improve the efficiency of pesticide use, making it a key component of crop disease management.

¹⁾ Jinglong REN, Ph.D.Stud.Eng; Huili ZHANG, Prof.Ph.D.Eng; Guangyuan WANG, Ph.D.Stud.Eng; Chenlong DAI, Ph.D.Stud.Eng; Fei TENG, Ph.D.Stud.Eng; Moxian LI, Ph.D.Stud.Eng

Although traditional methods of crop disease identification rely on manual observation and laboratory analysis, these methods are inefficient, complex, and fail to meet the practical needs of modern agricultural production. With the rapid development of artificial intelligence technology, machine vision and image processing techniques have been widely applied in the field of crop disease identification (Singh *et al.*, 2024). While traditional machine vision methods classify diseases by analysing features like colour, shape, and texture, these methods are inefficient and struggle to adapt to the rapidly changing needs of agriculture (Thakur *et al.*, 2023; Ibrahim *et al.*, 2023; Salman *et al.*, 2023). This study effectively tackles the rapid detection and identification of diseases in natural environments by integrating a multi-target detection model with manual labelling, data augmentation, and parameter optimization techniques. By integrating multi-target detection models, manual labelling, data augmentation, and parameter optimization techniques, and incorporating a coordinate attention mechanism into the backbone network, the accuracy and speed of detection have been effectively improved. (Jing HAO *et al.*, 2022; Ye XIA *et al.*, 2023). Through self-supervised learning, deep learning methods can automatically learn and extract features, effectively overcoming the limitations of traditional machine learning methods and showing great potential in disease identification.

In the realm of grape leaf disease identification and management, there is an urgent need for a method capable of rapid and precise execution. In response to the limitations of existing approaches, this study has developed an innovative real-time detection model named MSCl-YOLOv8s stability (Yang *et al.*, 2023; Jia *et al.*, 2023). This model combines an improved YOLOv8s framework with the efficient MobileNetV3 backbone network, showcasing superior performance in extracting multi-scale disease features from grape leaf images and enhanced generalization and stability (Wang *et al.*, 2023; Niu *et al.*, 2023). To further increase processing speed, the SPPFCSPC pyramid pooling structure have been employed, optimizing the spatial reception capacity of the model (Zhang *et al.*, 2023). Moreover, the integrated CBAM attention mechanism has improved the precision in identifying key features, significantly boosting the accuracy of disease detection. The utilization of the Inner-SIoU loss function has refined the accuracy of bounding box regression and expedited model convergence. Experimental results prove that the MSCl-YOLOv8s model, with its swift and accurate detection capabilities, meets the practical demands of orchard disease identification and lays a foundation for technological advancements in precision agriculture.

MATERIALS AND METHODS

Dataset

In this study, the focus was on several common diseases in grape leaves, including Black Rot, Leaf Blight, Downy Mildew, Anthracnose, Rust, and Black Measles. Specifically, four primary conditions for in-depth investigation were selected: Black Spot, Leaf Blight, Black Measles, and Healthy leaves. These specific diseases and the healthy state constituted the focal point of our research, aiming to enhance the accuracy and efficiency of our model in detecting grape leaf diseases through comprehensive analysis and recognition of these conditions.

To construct the dataset, a Sony Alpha 6400 camera was utilized to capture images of grape leaves affected by the aforementioned diseases and healthy leaves under various time and weather conditions, and from multiple angles. In total, 3,970 images were collected. To further enrich the dataset and increase the number of images, 500 grape leaf disease images were also incorporated from three public datasets: Plant Village, Plant Pathology, and Ai Studio, bringing the total number of images to 4,470 (Goncharov *et al.*, 2019; Agrios G.N. *et al.*, 1978). These carefully selected and diverse images provided a comprehensive and rich perspective for this research, aiding in enhancing the generalization ability and recognition efficiency of our model. Example images from the dataset are displayed in Figure 1, reflecting the diversity of grape leaves with different disease types.

In this study, the collected raw images of grape leaves were subjected to a series of data augmentation processes to enhance the diversity of the dataset and the generalization ability of the model. These augmentation techniques included image flipping, mirroring, translation, noise addition, and brightness adjustment. These manipulations not only increased the adaptability and robustness of the samples but also helped to improve the ability of the model to recognize grape leaf diseases in natural environments and strengthen the retention of label information.

After the aforementioned data augmentation methods, a total of 17,880 images of grape leaves were obtained. Subsequently, these augmented images were manually annotated using the Labellmg tool, and labels were created for the grape leaf disease dataset in the YOLO format. These labels included Black Rot, Leaf Blight, Black Measles, and Healthy.

To ensure the scientific validity and effectiveness of the training, the dataset of 17,880 images was randomly and evenly distributed into training, validation, and testing sets in an 8:1:1 ratio. The specific number of images for each category is summarized in Table 1. Through this approach, a grape leaf disease recognition dataset was successfully constructed for the experiments in this study.

Table 1

Disease Dataset				
Labels	Training Dataset	Validation Dataset	Test Dataset	Total Number of Images
Block_Rot	3776	472	472	4720
Leaf_Blight	3400	425	425	4250
Black_Measles	3800	475	475	4750
Healthy	3328	416	416	4160
Totals	14304	1788	1788	17880

Yolov8 Model

YOLOv8 represents the latest advancement in the YOLO (You Only Look Once) series of models, demonstrating exceptional performance in tasks such as object detection, image classification, and instance segmentation. Developed by Ultralytics, the company that previously released the highly influential YOLOv5 model, YOLOv8 incorporates numerous improvements and updates in its architectural design and development experience compared to YOLOv5. These modifications draw from the optimization strategies of many outstanding YOLO algorithms, propelling YOLOv8 to a new state-of-the-art (SOTA) level and achieving significant enhancements in object recognition performance.

Despite the commendable achievements in detection accuracy, YOLOv8 still has room for improvement in practical production environments. Currently, the model faces challenges in handling the impact of changes in the shape and size of targets on detection accuracy, reducing confusion due to similar features, and enhancing the speed of model recognition. To better adapt to the demands of real-world applications, YOLOv8 requires further optimization and refinement, strengthening its performance and practicality in various complex scenarios.

Improving the YoloV8s Model

To address the existing issues with YOLOv8 and further improve its accuracy and speed to better meet the needs of practical production, this study has made the following improvements:

1) By replacing the backbone feature extraction network of the original YOLOv8 model with MobileNetV3, a lightweight deep learning model optimized for mobile and embedded devices has been effectively integrated. MobileNetV3 significantly reduces model parameters and computational demands, greatly enhancing operational efficiency on devices with limited computing resources. This optimization makes the model more suitable for deployment on edge devices, such as smartphones or other portable devices, which is crucial for real-time disease detection applications.

2) In optimizing the original YOLOv8 model, its standard SPPF (Spatial Pyramid Pooling in Fast R-CNN) pyramid pooling was replaced with an innovative SPPFCSPC pyramid pooling. This improvement achieves a speed increase while maintaining the receptive field. This strategic modification not only optimizes the efficiency of the model, but also preserves its ability to effectively process spatial information, a key factor in achieving accurate detection.

3) Incorporating the new RFA (Receptive Field Attention) mechanism into the model reduces information loss and enhances feature information, facilitating feature extraction and accelerating recognition speed. This improvement further enhances the recognition performance of the model, making it more efficient and accurate in processing key features, which is vital for tasks requiring high precision and efficiency.

4) After replacing the original Ciou (Complete Intersection over Union) loss function of YOLOv8 with the Inner-SIoU loss function, the improved loss function aids the model in focusing on high-quality bounding boxes. This change not only improves the detection accuracy of the model but also accelerates the convergence process, thereby enhancing target recognition speed. This modification is a strategic improvement that significantly enhances the overall performance of the model, especially in tasks that require precise localization and rapid processing.

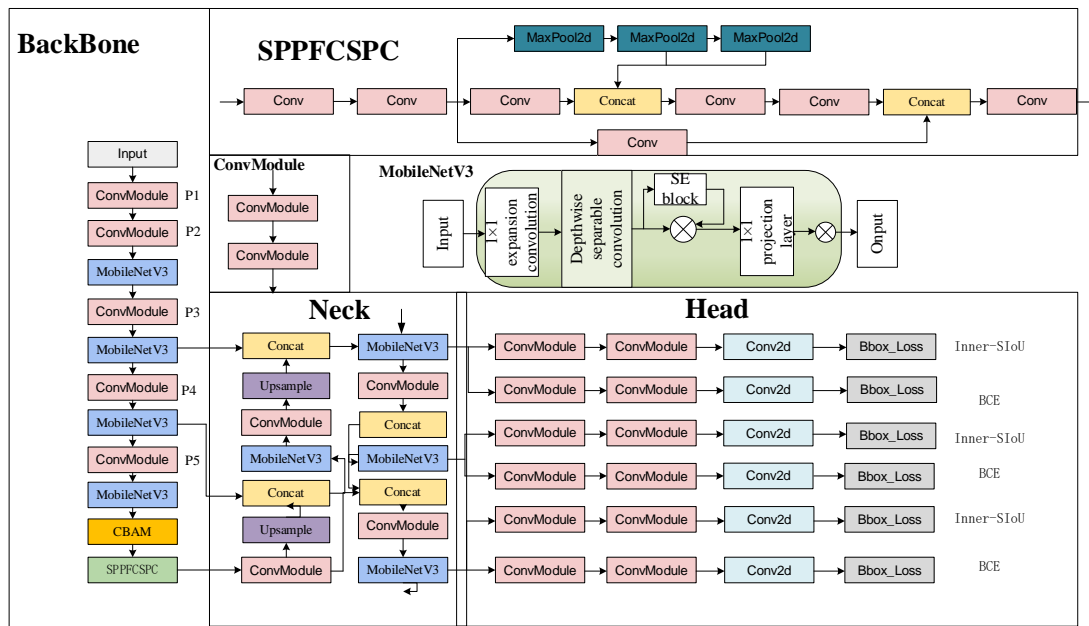


Fig. 1 - MSCI-YOLOv8s Network Structure

MobileNetV3 Feature Extraction Network

To enhance the accuracy and speed of grape leaf disease recognition based on the YOLOv8 model, this study has employed various modules for improvement, aiming to achieve superior recognition performance. MobileNetV3 was chosen as the backbone feature extraction network. MobileNetV3 is a lightweight deep learning model optimized for mobile and embedded devices, renowned for its efficient depthwise separable convolutions and other techniques. It significantly reduces the parameter count and computational needs of the model while achieving a good balance between accuracy and speed, making it highly suitable for real-time disease recognition applications.

MobileNetV3 is an efficient, lightweight deep learning model particularly apt for disease recognition on mobile and embedded devices. It drastically cuts down on the number of parameters and computational demands of the model through depthwise separable convolutions, allowing for rapid operation on resource-constrained devices. The network architecture, automatically designed using Network Architecture Search (NAS) technology, along with the incorporation of the H-swish activation function and Squeeze-and-Excitation (SE) module, ensures that MobileNetV3 remains lightweight while optimizing its feature extraction and representation capabilities. The specific structure employed in this paper is MobileNetV3-Small, as illustrated in Figure 2. This makes MobileNetV3 highly suitable for disease recognition applications, such as real-time monitoring of crop diseases and automated analysis of plant health conditions. Its efficient performance and accurate feature recognition capabilities make it an ideal choice for implementing disease recognition in mobile devices and embedded systems.

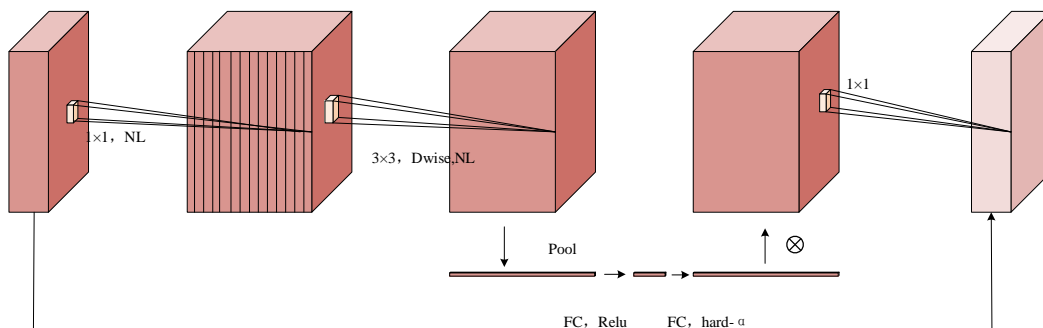


Fig. 2 - MobileNetV3-Small Architecture Diagram

SPPFCSPC Pyramid Pooling

After successfully replacing the backbone feature extraction network of the YOLOv8 model with MobileNetV3, the model achieved significant performance improvements.

To further enhance this effect, this study introduced three additional improvement modules, particularly the SPPFCSPC novel pyramid pooling, which effectively accelerated the processing speed for grape leaf disease recognition. The lightweight nature of MobileNetV3, combined with these innovative improvements, significantly increased processing efficiency while maintaining high-precision recognition.

The Spatial Pyramid Pooling (SPP) module aims to address two key issues: firstly, preventing image distortion caused by cropping and scaling of image regions, and secondly, reducing redundancy in feature extraction by convolutional neural networks. The SPP module accelerates the generation of candidate boxes and saves computational resources through these improvements. With continuous technological evolution, various efficient pyramid pooling structures have emerged. Different versions of the YOLO series exhibit diverse applications of this technology. For example, YOLOv5 and YOLOv8 have adopted the SPPF pyramid pooling structure in their source code, while YOLOv7 has utilized the SPPCSPC pyramid pooling. Each structure has its unique advantages and limitations, reflecting the specific needs and optimization directions of different YOLO model versions in processing images. The application of these pyramid pooling structures has significantly enhanced the adaptability and efficiency of models in various complex environments, achieving more precise and faster performance in object detection.

Specifically, the depthwise separable convolutions and efficient network structure design of MobileNetV3 ensure accurate feature extraction while reducing computational load. Combined with the SPPFCSPC pyramid pooling, it further optimizes the ability of the model to capture features at different scales, enhancing the sensitivity of disease recognition. These improvements work together to significantly enhance the practicality of the model in grape disease recognition, making it more suitable for rapidly and accurately handling complex agricultural disease recognition tasks. In summary, these integrated improvements showcase higher efficiency and precision in the application of grape disease recognition.

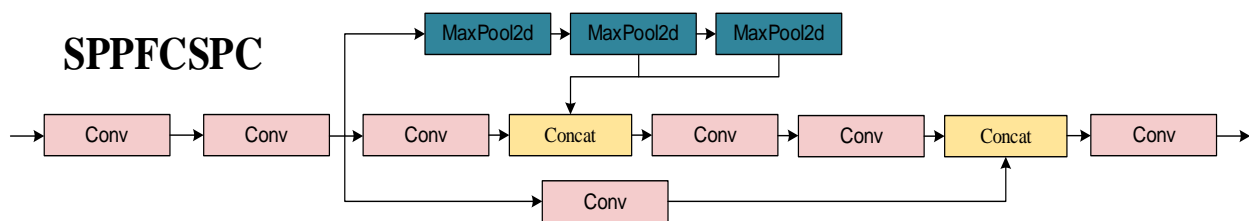


Fig. 3 - SPPFCSPC Structure Diagram

CBAM Attention

This article introduces an attention mechanism known as CBAM (Convolutional Block Attention Module), which plays a pivotal role in various image processing tasks, particularly in the recognition of grape leaf diseases. The main feature of the CBAM attention mechanism is its ability to effectively focus on key parts of an image, especially when dealing with diseased spots on grape leaves. By efficiently selecting features and concentrating on critical areas, it enhances the accuracy of recognition.

The structural design of CBAM is both ingenious and efficient, combining sequential channel and spatial attention methods to fully harness the advantages of both types of attention mechanisms. This design allows CBAM to be flexibly embedded into various convolutional neural network architectures, enhancing the performance of both shallow and deep models in a stable manner. CBAM demonstrates good adaptability to data and robustness, particularly suitable for recognizing grape leaf diseases in variable and complex scenarios.

Overall, the introduction of the CBAM attention mechanism offers an efficient and precise solution in the field of image processing, especially in the domain of grape leaf disease recognition. It enables models to focus more on the key features of images while maintaining outstanding performance. This innovation plays a significant role in enhancing accuracy and practicality in real-world applications.

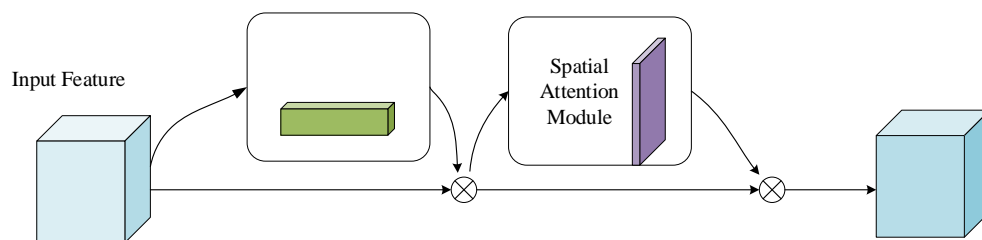


Fig. 4 - CBAM Attention Mechanism Principle Diagram

Inner-SIoU Loss Function

In the original YOLOv8 model, the Ciou (Complete Intersection over Union) loss function was used, which, although accounting for key geometric factors such as overlap area, centre point distance, and aspect ratio, has shortcomings in ensuring precise alignment of bounding boxes (BBoxes), a critical factor affecting the recognition accuracy of the model.

To address this, the Inner-SIoU loss function was adopted in place of the original Ciou loss function in this study. The Inner-SIoU loss function focuses on optimizing the alignment accuracy between predicted and actual bounding boxes, which is crucial for enhancing the accuracy of grape leaf disease recognition. It minimizes the discrepancy in the internal regions of the predicted and actual bounding boxes, rather than just their outer contours, thereby capturing the exact location and shape of disease spots more accurately. Additionally, the Inner-SIoU loss function features rapid convergence and excellent localization effects. By precisely calculating the overlap area, it effectively handles complex cases with less overlap, particularly excelling in images of grape leaves with complex backgrounds or smaller disease spots. The specific formula is as follows:

$$L_{Inner-SIoU} = L_{Siou} + IoU - \frac{(\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) * (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t))}{(w^{gt} * h^{gt}) * (ratio)^2 + (w * h) * (ratio)^2 - inter} \quad (1)$$

In this formula, w^{gt} and h^{gt} denote the width and height of the target box, while w and h represent the width and height of the predicted box. The term *inter* refers to the intersection of the predicted and target boxes. The ratio is a scaling factor that controls the size of the auxiliary bounding boxes used in the calculation.

Overall, the application of the Inner-SIoU loss function has significantly enhanced both the accuracy and speed of the grape disease recognition model developed in this study, better fulfilling the practical requirements for efficient and accurate disease detection. This improvement has substantially increased the practicality and efficacy of the model, marking a significant advancement in the field of disease recognition.

Experimental Environment Configuration and Training Parameter Settings

The experiments for this study were conducted under uniform conditions on a single server. The server specifications included a 15-core Intel(R) Xeon(R) Platinum 8358P CPU @ 2.60GHz, an NVIDIA GeForce RTX 3090 GPU with 24GB of memory, 80GB of RAM, a 1TB solid-state drive, and the Linux Ubuntu 20.04 operating system. The software environment comprised PyTorch 2.0.0, Cuda 11.8, Cudnn 8.6.0, and Python 3.8.

In these experiments, training parameters were meticulously calibrated to ensure optimal learning performance of the model. The image input size was set to 640x640 pixels, standardizing all input data to this resolution. Regarding batch size, a Batch=100 setting was opted for, meaning that each batch processed 100 images. To thoroughly train the model, the number of iterations (Epochs) was set to 500, ensuring that the model had ample time to learn and adjust its weights. The initial learning rate was set at Lr=0.01, providing a balanced starting point that was neither too fast to miss critical learning opportunities nor too slow to impede training efficiency. To monitor and save training progress timely, weight parameters were saved after every 50 iterations.

Evaluation Metrics

In this study, Precision (P), Recall (R), and mean Average Precision (mAP) were meticulously selected as the core metrics for evaluating the accuracy of the model. Precision P primarily measures the similarity between the predicted target area and the actual target area, while Recall R is used to assess the ability of the model to correctly identify target categories, that is, the proportion of accurately identified targets to the total number of targets required to be recognized.

To comprehensively evaluate the performance of the model, inference time and network parameter count were also incorporated into our assessment system to reflect the real-time performance of the model and computational resource efficiency. These metrics were calculated based on the number of true positive samples (TP), false positive samples (FP), and the total number of samples (N). For each grape leaf disease category, its Average Precision (AP_i) was calculated separately, derived from a specific formula designed to accurately reflect the recognition precision of each category. These comprehensive evaluation metrics allow us to thoroughly and accurately assess performance of the model in grape leaf disease recognition, ensuring its effectiveness and reliability in practical applications. The number of true positive samples (TP), false positive samples (FP), and the total number of samples (N) are used to calculate the metrics, with the Average Precision (AP) for each grape disease category derived from a specific formula. The formula is as follows:

$$AP_i = \frac{TP}{TP + FP} \tag{2}$$

$$mAP = \frac{\sum_{i=1}^Q AP_i}{Q} \times 100\% \tag{3}$$

$$P = \frac{TP}{TP + FP} \tag{4}$$

$$R = \frac{TP}{P} \tag{5}$$

RESULTS AND ANALYSIS

Grape Leaf Disease Recognition

Upon the successful completion of the training phase of the MSCl model, the optimally performing weight file, best.pt, was acquired. Subsequently, using images from the test set, an exhaustive evaluation of the MSCl model specific performance in grape leaf disease recognition was conducted. To visually demonstrate the results of the disease recognition experiment, the outcomes are presented in the form of a confusion matrix, as shown in Figure 5.

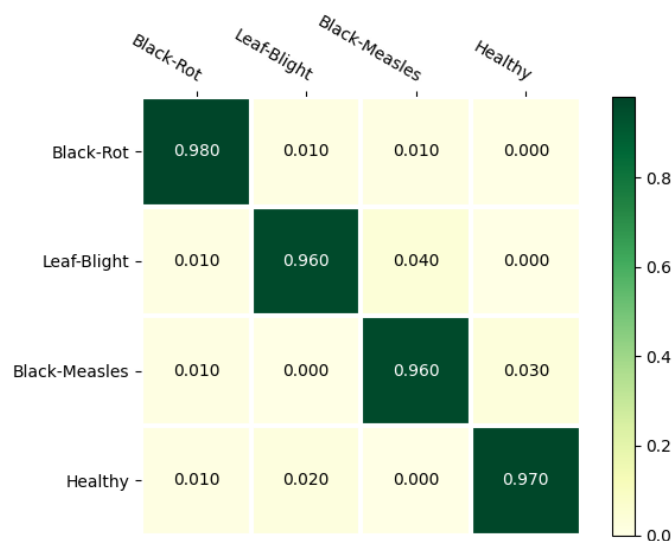


Fig. 5 - Predicted Label

Ablation Experiment

To assess the contribution of the MobileNetV3 backbone feature extraction network and the improvements made by the other three modules to the grape leaf disease recognition model, this study designed a series of ablation experiments. These experiments included evaluations of the original YOLOv8 model and the modified YOLOv8 model with MobileNetV3 as the backbone network. Through such comparisons, the significant impact of integrating MobileNetV3 as the backbone feature extraction network on model performance could be more intuitively understood.

Furthermore, in this study, Grad-CAM (Gradient-weighted Class Activation Mapping) technology for visual analysis of both the pre-improved YOLOv8 model and the post-improved MSRI model was also employed (Selvaraju et al., 2016). This method allowed to visually assess the MSRI model, particularly in analysing how the model focuses on key parts of grape leaves. This visualization technique enabled to intuitively understand and compare the level of attention and precision the model, pre- and post-improvement, gave to leaf disease features.

In the visualization results, as shown in the figure 6, the ability of the model before and after improvement to focus on key parts of the leaf was displayed. This not only revealed the mechanism of the model in disease recognition but also showcased the superiority of the MSRI model in capturing key features. These visual analyses provide a clearer understanding of how model improvements impact the precision of disease detection, offering valuable perspectives and foundations for future research.

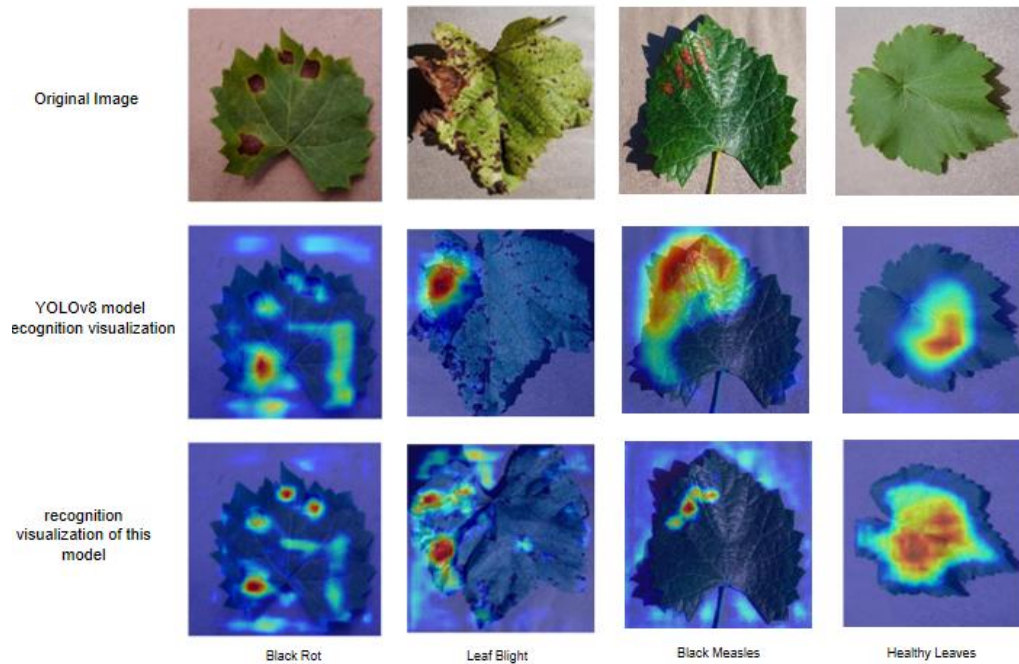


Fig. 6 - Visualization of the Model Focus on Key Leaf Positions Before and After Improvement

As indicated by the data in Table 2 and Figure 7, without any module improvements to the YOLOv8 model, it managed to achieve a grape disease recognition accuracy of 89.6%. While this result is somewhat acceptable, there is a considerable scope for optimization, particularly for real-world application scenarios. To enhance the recognition accuracy, the model was augmented with three modules: SPPFCSPC, CBAM, and Inner-SIoU. Although these improvements resulted in a slight increase in accuracy, the gains were limited, and they also led to higher memory usage and a slight reduction in inference speed. However, when these three modules were integrated into the model simultaneously, the accuracy significantly increased to 97.7%. This marked improvement demonstrates that these three enhancement modules have a distinctly positive impact on bolstering the performance of the grape disease recognition model.

Table 2

Ablation Experiment							
Experiment Number	Backbone Network	Pyramid Pooling	Attention Mechanism	Loss Function	mAP / %	Memory Usage / MB	IT / ms
1	Yolov8s	-	-	CloU	89.6	23.4	58.1
2	MobileNetV3	-	-	CloU	92.4	27.6	50.2
3	MobileNetV3	SPPFCSPC	-	CloU	93.1	29.7	47.2
4	MobileNetV3	SPPFCSPC	CBAM	CloU	94.6	34.7	38.5
5	MobileNetV3	SPPFCSPC	CBAM	Inner-SIoU	97.7	39.3	37.2

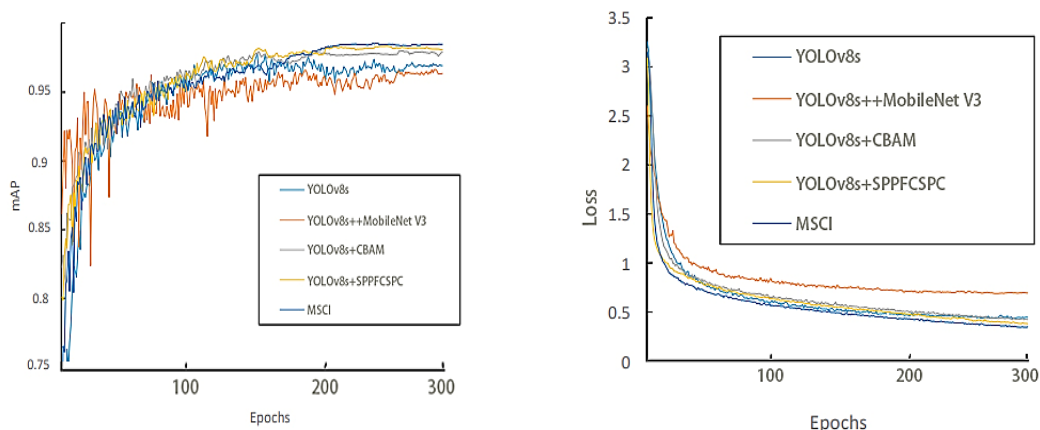


Fig. 7 - mAP Curve and Loss Curve

Comparison of Detection Performance of Different Models

To evaluate the efficiency of the improved algorithm proposed in this article, comparative experiments were conducted with Faster R-CNN, SSD, YOLOv4, YOLOv5, and YOLOv7, as shown in Tabel 9. The experimental results demonstrate that the algorithm in this study excels in accuracy, recall rate, mAP, frame rate, and model size. In particular, it achieved a mAP of 97.7% and a detection speed of 119.81 FPS, meeting the requirements for real-time detection. The model size is 19.40 MB, comparable to YOLOv8 but with superior performance. The algorithm has achieved excellent results in terms of real-time performance, detection accuracy, and efficiency.

Table 3

Comparison results of the models							
Network	P(%)	R(%)	mAP(%)	FPS	Model Size(M)	Network	P(%)
Faster-R-CNN	82.36	88.56	80.78	9	523	Faster-R-CNN	82.36
SSD	78.23	73.89	76.68	18	92	SSD	78.23
YOLOv4	89.63	85.23	88.38	52	22.3	YOLOv4	89.63
YOLOv5	91.35	89.20	91.23	90	15	YOLOv5	91.35
YOLOv7	92.56	90.19	91.96	101	12	YOLOv7	92.56
YOLOv8	92.63	91.20	92.00	151	23	YOLOv8	92.63
MSCI	97.70	93.00	94.03	119.81	19.40	MSCI	97.70

RESULTS

This study successfully developed an efficient and real-time grape disease detection model based on an improved YOLOv8s, named MSCI-YOLOv8s, aimed at significantly enhancing the accuracy and speed of grape disease detection. By replacing the backbone network of the YOLOv8s model with MobileNetV3 and embedding an SPPFCSPC pyramid pooling structure, the ability of the model to capture various scale disease features in grape leaf images was enhanced. This not only improved the generalization ability and stability of the model, but also, by introducing the CBAM attention mechanism, further enhanced the focus on key features, thereby increasing the accuracy of disease detection. Moreover, using Inner-SIoU as the loss function, the precision of the model in bounding box regression was enhanced and the convergence of the model was accelerated, thus improving detection efficiency. With these improvements, the MSCI-YOLOv8 model demonstrates outstanding performance in real-time grape disease detection, showing high application value and practical operational potential. The results of this study not only provide strong technical support for the accurate identification and timely treatment of grape diseases but also offer an effective reference model for similar agricultural image recognition tasks. With further optimization and customization, this model is expected to play a broader role in precision agriculture and intelligent monitoring.

ACKNOWLEDGEMENT

This research received support from the Shandong Key Research and Development Programme in the Public Welfare Field (Project No: 2018GGX104012), titled "Key Technology Research and Industrialization of High-Frequency Isolation Power Battery Detection System."

REFERENCES

- [1] Ananga A, Georgiev V, Ochieng J., (2013). Production of anthocyanins in grape cell cultures: a potential source of raw material for pharmaceutical, food, and cosmetic industries. *The Mediterranean genetic code-grapevine and olive*, 1(0): 247-287, Croatia.
- [2] Agrios G N. (2005). Plant pathology[M]. *Elsevier*, United States.
- [3] Farooq M S, Riaz S, Abid A, (2019). A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming. *Ieee Access*, 7: 156237-156271, Pakistan.
- [4] Georgiev V, Ananga A, Tsoleva V., (2014). Recent advances and uses of grape flavonoids as nutraceuticals. *Nutrients*, 6(1): 391-415, United States.
- [5] Goncharov P, Ososkov G, Nechaevskiy A, (2019). Disease detection on the plant leaves by deep, *Remote Sensing*, 15(17): 4356, China.

- [6] Ibrahim A E, Akchioui N E., (2023). A review on plant diseases detection using artificial intelligence techniques[C]//AIP Conference Proceedings. *AIP Publishing*, 2814(1), Morocco.
- [7] learning[C]//Advances in Neural Computation, Machine Learning, and Cognitive Research II: Selected Papers from the XX International Conference on Neuroinformatics, October 8-12, 2018, Moscow, Russia. *Springer International Publishing*, 151-159, Belarus.
- [8] Jia L, Wang T, Chen Y, (2023). MobileNet-CA-YOLO: An Improved YOLOv7 Based on the MobileNetV3 and Attention Mechanism for Rice Pests and Diseases Detection[J]. *Agriculture*, 13(7): 1285, China.
- [9] Niu K, Wang C, Xu J, (2023). An Improved YOLOv5s-Seg Detection and Segmentation Model for the Accurate Identification of Forest Fires Based on UAV Infrared Image. *Remote Sensing*, 15(19): 4694 China.
- [10] Nicholas K A, Durham W H., (2012). Farm-scale adaptation and vulnerability to environmental stresses: Insights from winegrowing in Northern California. *Global Environmental Change*, 22(2): 483-494, United States.
- [11] Salman Z, Muhammad A, Piran M J, (2023). Crop-saving with AI: latest trends in deep learning techniques for plant pathology. *Frontiers in Plant Science*, 14, Korea.
- [12] Selvaraju R R, Das A, Vedantam R, (2016). Grad-CAM: Why did you say that [J]. *arXiv preprint arXiv:1611.07450*, United States.
- [13] Sharma A, Jain A, Gupta P, (2020). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, 9: 4843-4873, India.
- [14] Singh B K, Delgado-Baquerizo M, Egidi E, (2023). Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology*, 1-17, Morocco.
- [15] Sishodia R.P., Ray R. L., Singh S. K., (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12(19): 3136, United States.
- [16] Thakur A, Venu S, Gurusamy M., (2023). An extensive review on agricultural robots with a focus on their perception systems. *Computers and Electronics in Agriculture*, 212: 108146, India.
- [17] Wang Z, Ye R, Xu Y, (2023). Research on Meteorological Equipment Identification and Monitoring Algorithm Based on Yolov8. *Available at SSRN 4665653*, China.
- [18] Yang Y, Han J., (2023). Real-Time object detector based MobileNetV3 for UAV applications[J]. *Multimedia Tools and Applications*, 82(12): 18709-18725, China.
- [19] Ye XIA, Xiaohui LEI, Andreas HERBST., (2023). Research on Pear Inflorescence Recognition Based on Fusion Attention Mechanism with YOLOv5. *INMATEH - Agricultural Engineering*, 69(01), China.
- [20] Zongwei JIA, Jing HAO, Yiming HOU., (2022). Study on Rapid Detection and Identification of Multi category Apple Leaf Disease. *INMATEH - Agricultural Engineering*, 67(06), China.
- [21] Zhang H, Xu C, Zhang S., (2023). Inner-IoU: More Effective Intersection over Union Loss with Auxiliary Bounding Box[J]. *arXiv preprint arXiv:2311.02877*, China.
- [22] Zhang Y, Wang H, Liu J, (2023). A lightweight winter wheat planting area extraction model based on improved DeepLabv3+ and CBAM. *Remote Sensing*, 15(17): 4156, China.