



## OPEN ACCESS

## EDITED BY

Bimlesh Kumar,  
Indian Institute of Technology Guwahati, India

## REVIEWED BY

Hamidreza Bolhasani,  
Islamic Azad University, Iran  
Hao Wang,  
Laoshan National Laboratory, China

## \*CORRESPONDENCE

Dae-Ki Kang  
✉ dkkang@dongseo.ac.kr

RECEIVED 16 July 2025

ACCEPTED 28 August 2025

PUBLISHED 18 September 2025

## CITATION

Sun H, Fu R and Kang D-K (2025) A novel efficient eggplant disease detection method with multi-scale learning and edge feature enhancement. *Front. Plant Sci.* 16:1666955. doi: 10.3389/fpls.2025.1666955

## COPYRIGHT

© 2025 Sun, Fu and Kang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# A novel efficient eggplant disease detection method with multi-scale learning and edge feature enhancement

Hao Sun<sup>1,2</sup>, Rui Fu<sup>1</sup> and Dae-Ki Kang<sup>2\*</sup>

<sup>1</sup>Shandong Facility Horticulture Bioengineering Research Center, Weifang University of Science and Technology, Weifang, China, <sup>2</sup>Department of Computer Engineering, Dongseo University, Busan, Republic of Korea

In the context of the rapid development of smart agriculture, the detection of crop diseases remains a critical and challenging task. The diversity in eggplant disease scales, disease edge features, and the complexity of planting backgrounds significantly impact disease detection effectiveness. To address these challenges, we propose an eggplant disease detection network with edge feature enhancement based on multi-scale learning. The overall network adopts a “backbone–neck–head” architecture: the backbone extracts features, the neck performs feature fusion, and a three-scale detection head produces the final predictions. First, we designed the Multi-scale Edge Information Enhance (CSP-MSEIE) module to extract features from different disease scales and highlight edge information to obtain richer target features. Second, the Multi-source Interaction Module (MSIM) and Dynamic Interpolation Interaction Module (DIIM) sub-modules were designed further to enhance the model’s capacity for multi-scale feature representation. By leveraging dynamic interpolation and feature fusion strategies, these sub-modules significantly improved the model’s ability to detect targets in complex backgrounds. Then, leveraging these sub-modules, we designed the Multi-scale Context Reconstruction Pyramid Network (MCRPN) to facilitate spatial feature reconstruction and hierarchical context extraction. This framework efficiently combines feature information across multiple levels, strengthening the model’s ability to capture and utilize contextual details. Finally, we validated the effectiveness of the proposed model on two disease datasets. It is noteworthy that on the eggplant disease data, the proposed disease detection model achieved improvements of 4.7% and 7.2% in mAP50 and mAP50–95 metrics, respectively, and the model’s frames per second (FPS) reached 270.5. This detection network provides an effective solution for the efficient detection of crop diseases.

## KEYWORDS

eggplant disease detection, deep learning, edge feature enhancement, multi-scale learning, object detection

## 1 Introduction

Eggplants are widely cultivated and highly valued for their rich content of dietary fiber and essential vitamins. They play a crucial role in improving global dietary patterns and promoting nutritional balance. Advancements in agricultural technology and the rapid growth of international trade have significantly increased both the cultivation area and total production of eggplants in recent years. By 2022, global eggplant production had exceeded 59 million tons, and the cultivation area surpassed 1.89 million hectares, underscoring its importance in modern agriculture (Yan et al., 2024). However, the expansion of cultivation has made eggplants more vulnerable to diseases and pest infestations, as shown in Figure 1, especially under increasingly complex and unpredictable climate conditions. Common problems such as yellow spot disease, fruit rot, and pest infestations severely threaten eggplant yield and quality, resulting in significant economic losses for growers (Liu and Wang, 2021).

Efficient disease management remains a central challenge in agricultural production. Traditional management methods mainly rely on manual inspection and chemical control, both of which have notable limitations. Manual inspection is time-consuming and prone to errors due to subjective judgment and reliance on individual experience, making it unsuitable for large-scale modern agriculture. Moreover, the growing reliance on pesticides to combat frequent disease outbreaks not only raises production costs but also increases pathogen resistance, posing additional threats to the environment and food safety. Therefore, developing accurate and efficient disease detection and management technologies is imperative.

In recent years, automated detection technologies have advanced rapidly in agriculture, offering innovative solutions for addressing eggplant diseases. For example, spectral analysis has been preliminarily applied to eggplant disease identification. However, the

complexity of processing high-dimensional spectral data and the associated information loss during dimensionality reduction have become major development bottlenecks (Wu, 2018). Additionally, texture-based feature extraction algorithms combined with classification models have shown moderate effectiveness in eggplant disease classification (Xie and He, 2016). However, these methods depend on manual feature extraction, suffer from pixel-level information loss, and exhibit high computational complexity, limiting their scalability and practical use. The emergence of machine learning has introduced promising solutions for crop disease recognition. Convolutional Neural Network (CNN)-based models have been successfully applied to eggplant disease recognition, demonstrating notable advantages over traditional approaches. However, early machine learning models mainly focused on classification tasks (Krishnaswamy Rangarajan and Purushothaman, 2020; Maggay, 2025; Theckedath and Sedamkar, 2025), neglecting the critical aspect of disease localization. This limitation prevents these models from fully replacing manual inspection, as accurate localization is essential for targeted treatment and intervention.

The rise of computer deep learning technology has driven object detection models toward greater efficiency and precision. These models can classify diseases and accurately localize them in images, representing a breakthrough in automated crop disease detection. Object detection models are typically categorized into two types: single-stage and two-stage detectors. Representative two-stage models include SSD (Tian et al., 2023), Faster R-CNN (Ren et al., 2017), and RetinaNet (Math and Dharwadkar, 2023). These models first generate region proposals or candidate bounding boxes, followed by classification and fine-grained localization within those regions. In contrast, single-stage models such as the YOLO (You Only Look Once) (Redmon et al., 2016; Redmon and Farhadi, 2017; Bochkovskiy et al., 2020; Redmon and Farhadi, 2018; Jocher

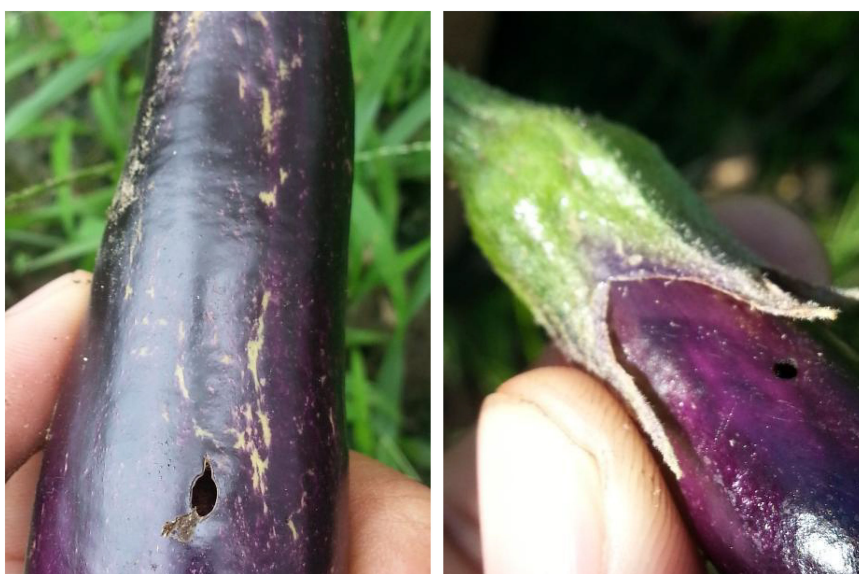


FIGURE 1

Status of eggplant cultivation, highlighting that very small, concealed fruit diseases (e.g., Fruit Borer) are often missed during field inspection.

et al., 2022; Wang et al., 2022; Redmon and Farhadi, 2025; Wang et al., 2024b; Li et al., 2022) series are widely recognized for their real-time performance and high accuracy. Recent research has increasingly focused on enhancing YOLO models for crop disease detection. For example, Liu et al. proposed a YOLOv5 variant with a novel loss function to detect tomato brown rot (Liu et al., 2023). Wang et al. integrated a Transformer into YOLOv8 to enhance tomato disease detection, significantly improving its ability to capture detailed disease features (Wang and Liu, 2024). Jiang et al. combined the Swin Transformer with CNN to optimize YOLOv8's feature extraction, improving detection performance for cabbage diseases under complex conditions (Jiang et al., 2024). Liu et al. introduced a multi-source information fusion approach based on YOLOv8 to enhance detection accuracy across multiple vegetable diseases (Liu and Wang, 2024). Moreover, some researchers have further improved detection performance on target images by employing edge-image enhancement (Wang et al., 2023) and additional image-preprocessing techniques (Wang et al., 2024c).

Although these enhanced YOLO models have shown progress, they primarily focus on leaf diseases, small datasets, and parameter tuning. However, the impact of scale variations in fruit disease regions and edge features under complex backgrounds on detection accuracy remains underexplored. To address this critical gap in current eggplant fruit disease detection methods, we propose an eggplant disease detection network with edge feature enhancement based on multi-scale learning. The key contributions of this paper are outlined as follows:

- We develop the Multi-scale Edge Information Enhancement (CSP-MSEIE) module, which extracts features across multiple disease scales and highlights the edge characteristics of affected regions, enabling richer and more comprehensive target representations.
- We develop the Multi-source Interaction Module (MSIM), Dynamic Interpolation Interaction Module (DIIM), and Multi-scale Context Extraction Module (MCEM), which enhance the model's capacity to capture multi-scale features

and improve target detection accuracy in complex backgrounds by utilizing dynamic interpolation and the fusion of multiple features.

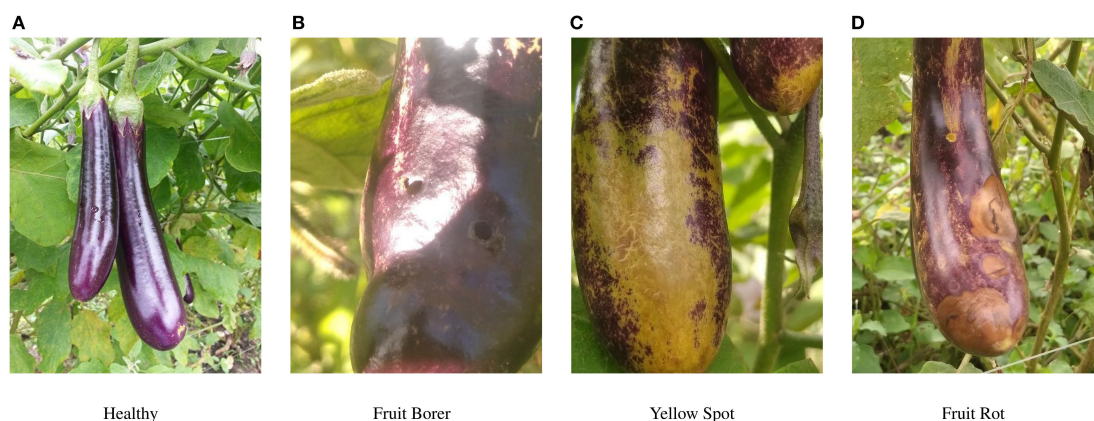
- We construct the Multi-scale context reconstruction pyramid network (MCRPN). This network aims to reconstruct spatial features and extract pyramid context, effectively integrating feature information from different levels and enhancing contextual awareness, thereby improving the model's detection performance.
- We conducted extensive ablation and comparative experiments on the two datasets, and the results show that EggplantDet outperforms other advanced detection algorithms in detection performance, even surpassing the advanced detection model YOLO11.

## 2 Materials and methods

### 2.1 Materials

**Dataset processing:** We validated the effectiveness of the proposed model on two datasets: PlantDoc (Singh et al., 2019) and eggplant disease. PlantDoc is a dataset of 2,569 images across 13 plant species and 30 classes (diseased and healthy) for image classification and object detection. There are 8,851 labels. Among them, the eggplant disease data is an eggplant fruit disease dataset from the Roboflow platform, containing four distinct disease categories (BSCS, 2024). The dataset includes four categories: healthy, fruit borer, yellow spot, and fruit rot. Detailed category distributions are presented in Figure 2. It is divided into training, validation, and test sets, comprising 2507, 744, and 365 images, respectively. The dataset was collected from diverse, natural cultivation environments, making it highly valuable for applied research.

To better enhance the model's generalization ability and detection performance on the eggplant disease dataset, we utilized the online data augmentation method of the Roboflow platform to



**FIGURE 2**  
Display of some four eggplant diseases before enhancement. (A) Healthy (B) Fruit Borer (C) Yellow Spot (D) Fruit Rot.



perform data augmentation on the training dataset of this eggplant disease dataset. The augmentations included: 90° rotation (clockwise and counter-clockwise), saturation adjustment (-30% to +30%), general rotation (-45° to +45°), horizontal and vertical flipping, grayscale (applied to 15% of images), hue adjustment (-15° to +15°), cropping (0-20% zoom), brightness adjustment (-15% to +15%), exposure adjustment (-10% to +10%), Gaussian blur (up to

4.8 px), noise addition (up to 1.99% of pixels), and shear transformation ( $\pm 15^\circ$  horizontally and vertically) in Figure 3. As a result of these 12 augmentation methods, the expanded dataset includes 7521 images for training, 744 for validation, and 365 for testing.

Implementation details: This study was implemented using a Python deep learning framework on the Windows 11 operating

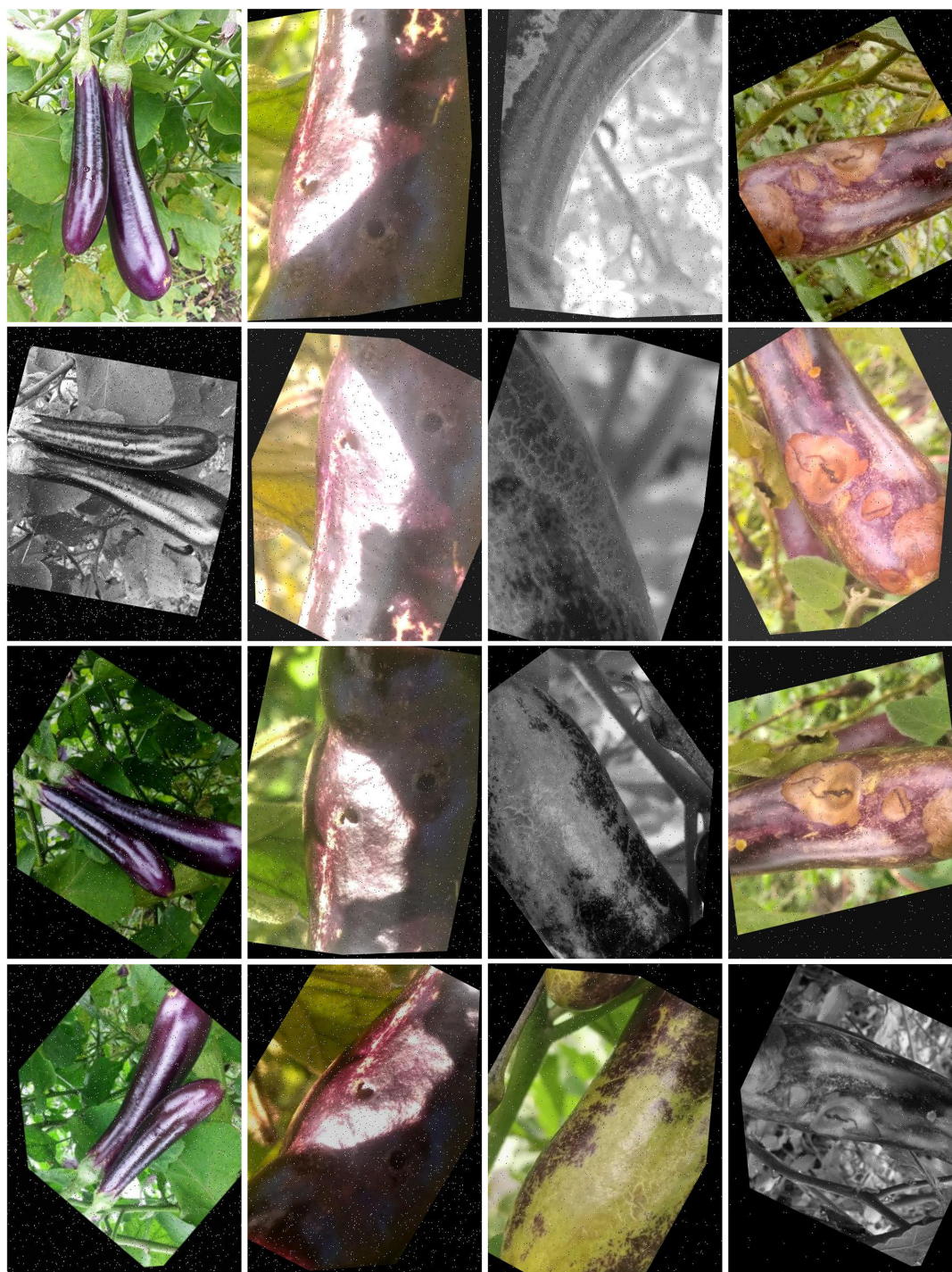


FIGURE 3  
Display of some four eggplant diseases after enhancement.



TABLE 1 Experimental environment.

Name	Details
Programming language	Python 3.9
GPU	NVIDIA GeForce RTX 4090
CUDA	11.8
Pytorch	2.0.1
Platform	Visual Studio Code

system. For the training process, a batch size of 16 was used, with the SGD optimizer, an initial learning rate of 0.01, weight decay set to 0.0005, and the training was carried out over 100 epochs. The detailed experimental settings are provided in [Table 1](#).

## 2.2 Methods

### 2.2.1 Macroscopic architecture of EggplantDet

Considering the scale variations of eggplant disease targets and their susceptibility to complex background interference, we constructed EggplantDet based on the YOLOv8 model. [Figure 4](#) illustrates the overall architecture of the proposed EggplantDet. The detection network comprises three main components: the Backbone, the Multi-scale Context Reconstruction Pyramid Network (MCRPN), and the Head.

- **Backbone:** The input feature map size is  $640 \times 640 \times 3$ , utilizing multiple  $3 \times 3$  convolutions to reduce image dimensions and increase channel numbers. To extract features across various disease scales and emphasize the edge information of the diseases, the CSP-MSEIE module was designed and integrated into the Backbone (as shown in

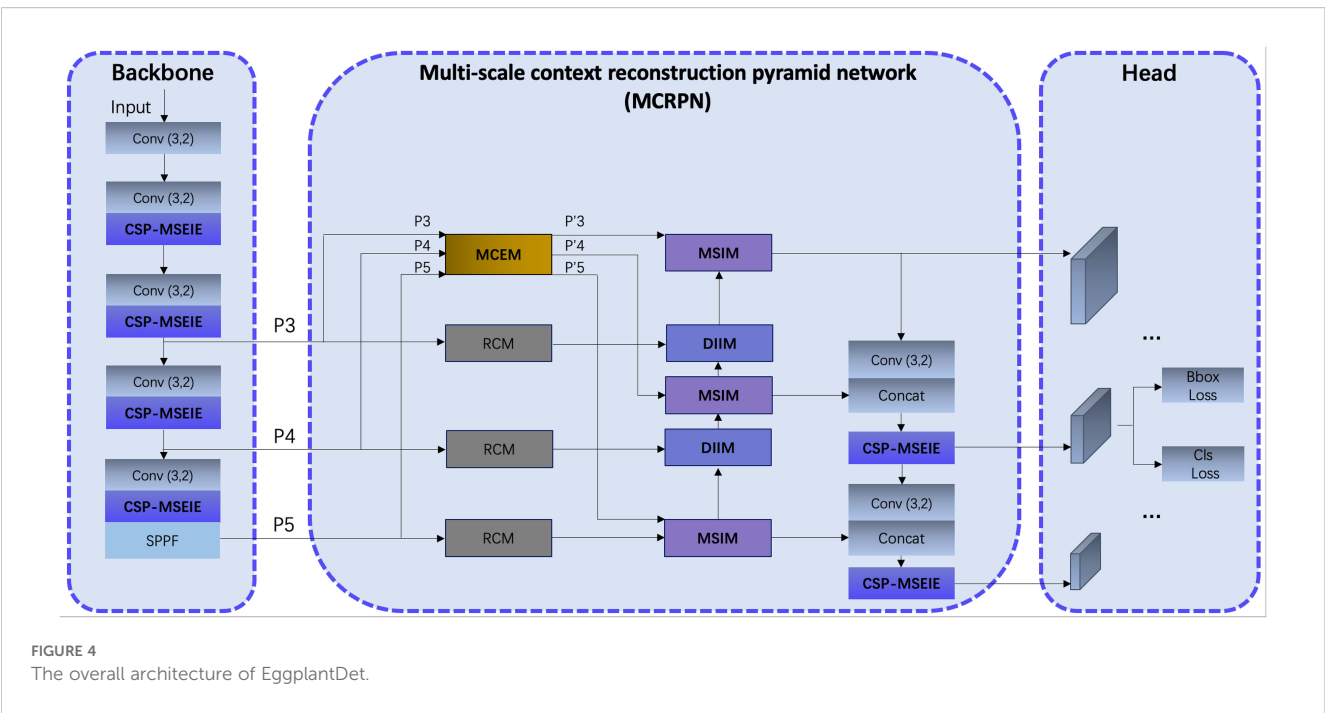
[Figure 3](#)). The convolution, CSP-MSEIE, and SPPF modules work together to generate P3 features of  $80 \times 80 \times 256$ , P4 features of  $40 \times 40 \times 512$ , and P5 features of  $20 \times 20 \times 1024$  for the subsequent MCRPN network.

- **MCRPN:** As depicted in the center of [Figure 3](#), P3, P4, and P5 are first processed through the RCM (Ni et al., 2024) module to reconstruct and extract key contextual features in both horizontal and vertical directions. Subsequently, the MCEM module integrates features from different levels, while the MSIM and DIIM modules fuse multi-scale features. This significantly improves target recognition performance in complex backgrounds.
- **Head:** The detection head integrates features from three scale layers: P3, P4, and P5. This design effectively captures fine-grained information in low-level feature maps, thereby enhancing detection accuracy for multi-scale targets. In terms of loss functions, the model retains traditional box and classification losses to ensure accurate prediction box locations and categories.

Overall, the model applies targeted optimizations to both the Backbone and Neck components. Specifically, the introduction of the CSP-MSEIE module and MCRPN network significantly enhances the extraction and fusion of multi-scale and edge features, enabling EggplantDet to exhibit greater robustness and accuracy in eggplant disease detection tasks.

### 2.2.2 Cross-Stage Partial - Multi-scale Edge Information Enhance (CSP-MSEIE)

To extract multi-scale features and emphasize target edge information, we designed the Multi-scale Edge Information Enhance (MSEIE) module. We integrated it with the Cross Stage



Partial Net (CSP) structure to form the Cross Stage Partial-Multi-scale Edge Information Enhance (CSP-MSEIE) module, enhancing the learning capability of convolutional neural networks in the Backbone. As depicted in Figure 5, the MSEIE module is comprised of three main components: (1) Multi-scale feature extraction: Different parameters of AdaptiveAvgPool (3, 6, 9, 12) are used to achieve multi-scale pooling, extracting local information of different sizes, which helps capture hierarchical features of images. (2) Edge enhancement: The Edge Enhancer module is specifically designed to extract disease edge information, thereby enhancing the network's sensitivity to edge features. As illustrated in Figure 6, the Edge Enhancer module initially applies average pooling to the input feature map to capture low-frequency information. Next, the smoothed feature map is subtracted from the original input feature map to extract the enhanced edge information (high-frequency details). Finally, this high-frequency information is added back to the original feature map to produce the enhanced output. (3) Feature fusion: Features from various scales are aligned to a unified scale through interpolation operations, and after concatenation, they are fused through convolutional layers into a unified feature representation, thus improving the model's perception of multi-scale features. The CSP-MSEIE module integrates multi-scale feature extraction, edge information enhancement, and convolution operations. Incorporating the CSP-MSEIE module into the Backbone notably enhances edge features and the model's capacity to extract features.

### 2.2.3 The principle and details of the MCEM, MSIM and DIIM

To more effectively reconstruct spatial features and capture multi-scale contextual information, we designed three key modules: Multi-scale Context Extraction Module (MCEM), Multi-source Interaction Module (MSIM), and Dynamic Interpolation Interaction Module (DIIM). In the MCEM module, for the P3, P4, and P5 level features extracted by the Backbone, average pooling is first applied to unify feature scales and perform fusion, followed by the use of the RCM module to model axial global context for extracting rectangular key region features. Finally, the features P'3, P'4, and P'5 features are generated through the split operation, thereby effectively integrating information from different levels and enhancing the contextual awareness of the MCRPN network. In the MSIM module, convolution operations are first used to adjust the number of channels, then the sigmoid function and interpolation algorithm further adjust feature dimensions, followed by the multiplication of features from two branches. In the DIIM module, interpolation operations automatically adjust the dimensions of matching features, followed by convolution operations for additive fusion. These three modules greatly enhance the model's capability to capture features across multiple scales and enhance target recognition performance in complex backgrounds by employing dynamic interpolation and the fusion of multiple features. This process can be specifically expressed in Equations 1–3:

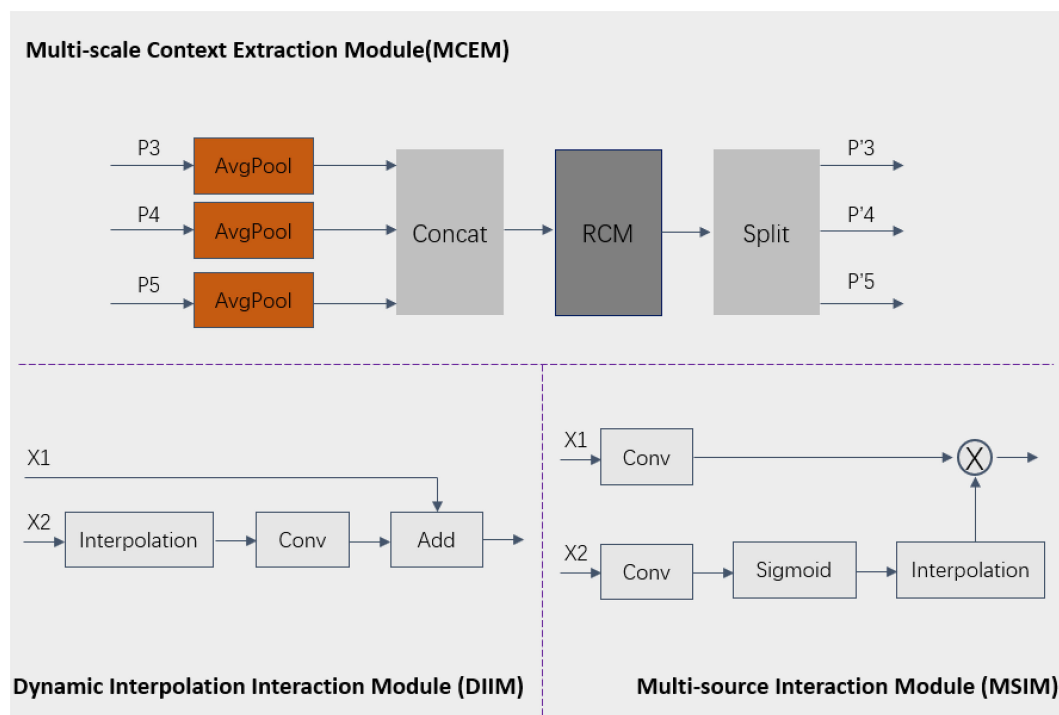
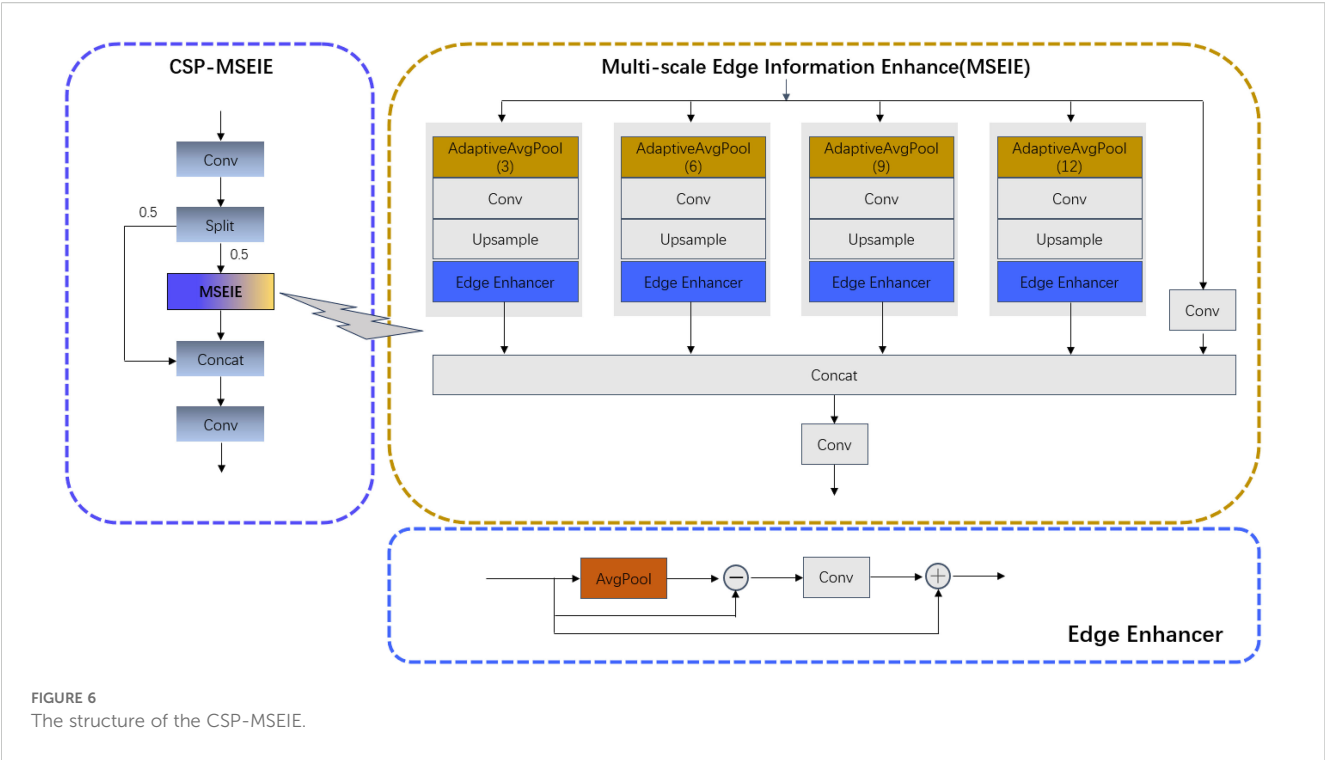


FIGURE 5  
The structure details of the MCEM, MSIM and DIIM.





$$\text{output}_{\text{MCEM}} = \text{Split}(\text{RCM}(\text{C}(\text{AP}(P_3, P_4, P_5))))1 \quad (1)$$

$$\text{output}_{\text{MSIM}} = X_1 + (\text{Interpolation}(\text{Conv}(X_2)))2 \quad (2)$$

$$\text{output}_{\text{DIIM}} = X_1 \times \text{Interpolation}(\text{S}(\text{Conv}(X_2)))3 \quad (3)$$

Where  $\text{AP}(\cdot)$  represents the average pooling operation,  $\text{S}$  represents the  $h$ -sigmoid function,  $+$  denotes addition operation,  $\times$  denotes multiplication operation, and  $\text{C}(\cdot)$  is the Concatenation operation.

Frames per second (FPS). Of these, mAP50 was selected as the primary evaluation metric. The procedure for calculating the mean Average Precision is described in Equations 4–7.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$AP = \int_0^1 P(R) \, dR \quad (6)$$

$$\text{mAP} = \frac{\sum_{i=1}^K AP_i}{K} \quad (7)$$

### 3 Experiments

#### 3.1 Experimental indicators

In this study, we used several indicators to assess the performance of our model: GFLOPs, Parameters, mean Average Precision (mAP50-90), mean Average Precision (mAP50), and

The variable  $K$  signifies the total count of distinct object classifications within the dataset, while each class’s precision is quantified by its specific Average Precision (AP) score. In the

TABLE 2 Comparison with advanced object detection models on the PlantDoc dataset.

Model	Params	GFLOPs	mAP50-95	mAP50	FPS
YOLOv8n (baseline) (Redmon and Farhadi, 2025)	3.0M	8.1	0.285	0.420	190.4
YOLOv9t (Wang et al., 2024b)	2.1M	7.6	0.290	0.423	193.4
YOLOv10n (Wang et al., 2024a)	2.3M	6.6	0.286	0.428	209.5
YOLOv11n (Jocher and Qiu, 2024)	2.6M	6.3	0.295	0.431	227.6
EggplantDet (Ours)	3.1M	7.7	0.304	0.438	241.2

TABLE 3 Detection results with different models on the eggplant disease dataset.

Model	Params	GFLOPs	mAP50-95	mAP50	FPS
Faster-RCNN (Ren et al., 2017)	314M	341.2	0.451	0.721	92.7
SSD (Liu et al., 2016)	53M	112.5	0.418	0.682	44.6
RT-DETR (Zhao et al., 2024)	82M	109.6	0.447	0.728	109.2
YOLOv3 (Redmon and Farhadi, 2018)	12M	19.0	0.423	0.746	80.1
YOLOv5n (Jocher et al., 2022)	2.5M	7.1	0.439	0.763	92.9
YOLOv6 (Li et al., 2022)	4.2M	11.8	0.438	0.755	106.6
YOLOv7 (Wang et al., 2022)	5.6M	13.4	0.426	0.768	118.5
YOLOv8n (Baseline) (Redmon and Farhadi, 2025)	3.0M	8.1	0.469	0.805	186.1
YOLOv8s (Redmon and Farhadi, 2025)	11.1M	28.4	0.478	0.815	108.1
YOLOv9t (Wang et al., 2024b)	2.0M	7.6	0.463	0.810	203.7
YOLOv10n (Wang et al., 2024a)	2.3M	6.5	0.474	0.815	226.1
YOLOv11n (Jocher and Qiu, 2024)	2.6M	6.3	0.476	0.826	250.8
EggplantDet (Ours)	3.1M	7.7	0.503	0.843	270.5

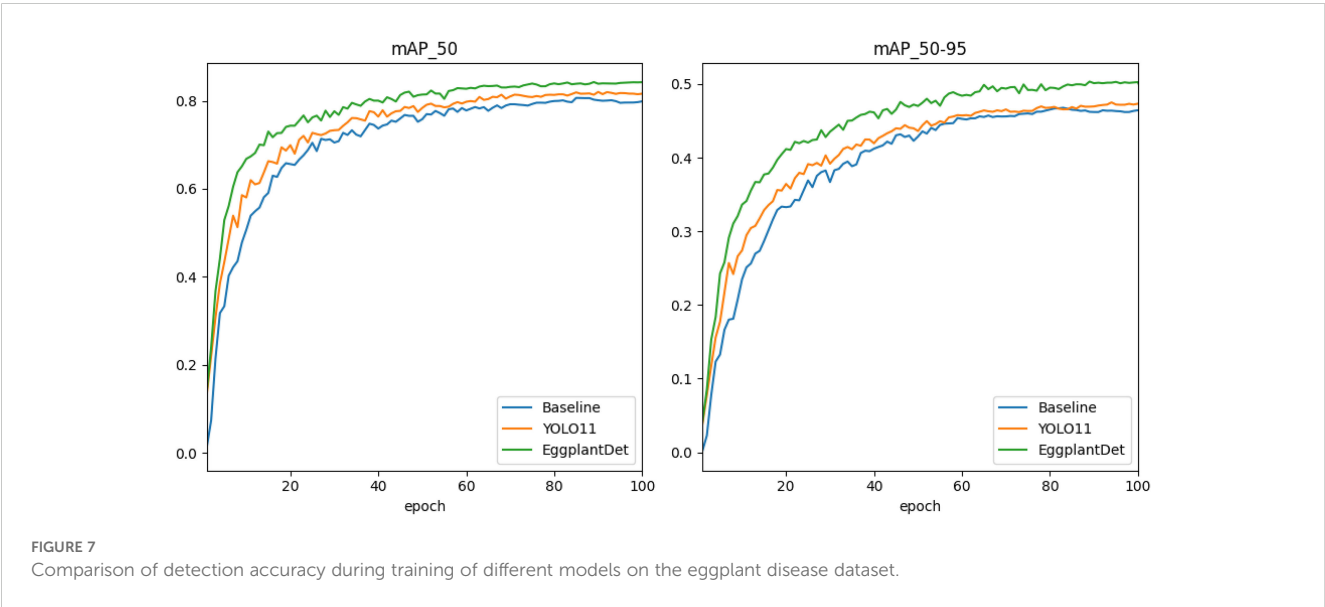


TABLE 4 The results of the ablation study on the eggplant disease dataset.

Baseline	CSP-MSEIE	MCRPN	MCEM	Parameters	GFLOPs	mAP50-95	mAP50	FPS
				3,006,428	8.1	0.469	0.805	186.1
	✓			2,855,900	7.6	0.474	0.823	221.2
		✓		3,339,492	8.3	0.481	0.818	217.7
			✓	3,114,612	8.1	0.467	0.816	209.8
	✓	✓		3,237,964	7.8	0.493	0.837	266.5
	✓		✓	3,034,108	7.9	0.489	0.832	251.4
		✓	✓	3,439,628	8.2	0.487	0.828	247.8
EggplantDet(ours)	✓	✓	✓	3,183,580	7.7	0.503	0.843	270.5



performance evaluation equations, several key indicators are utilized: True Positives (TP) represent accurately identified instances of the target condition, False Positives (FP) indicate cases where the algorithm incorrectly flagged non-existent conditions as present, and False Negatives (FN) encompass actual occurrences of the condition that the system failed to recognize.

3.2 Comparison studies

To verify the generalization ability of the proposed detection model, this paper first compared the disease detection performance of current mainstream advanced object detection models on the public PlantDoc dataset. As shown in Table 2, compared with the baseline model, EggplantDet achieved improvements of 4.3% and



FIGURE 8  
The visualization detection results of different models, (A) Baseline Model; (B) YOLO11; (C) EggplantDet.

6.7% in mAP50 and mAP50–95, respectively. Compared with the advanced YOLO11n, EggplantDet improved mAP50–95 by 3.0% and mAP50 by 1.6%. Additionally, in terms of Frames per second (FPS), EggplantDet also outperformed other advanced mainstream detection models.

Secondly, to further verify the advantages of the proposed eggplant disease detection model, we conducted a comprehensive

experimental comparative evaluation on the augmented eggplant disease dataset. Table 3 presents the experimental results of several advanced detection algorithms, including both two-stage and mainstream single-stage models for comparison. As shown in Table 3, two-stage detectors (e.g., SSD and Faster R-CNN) exhibited significantly lower mAP and FPS compared to the proposed method. Additionally, they required substantially more parameters and

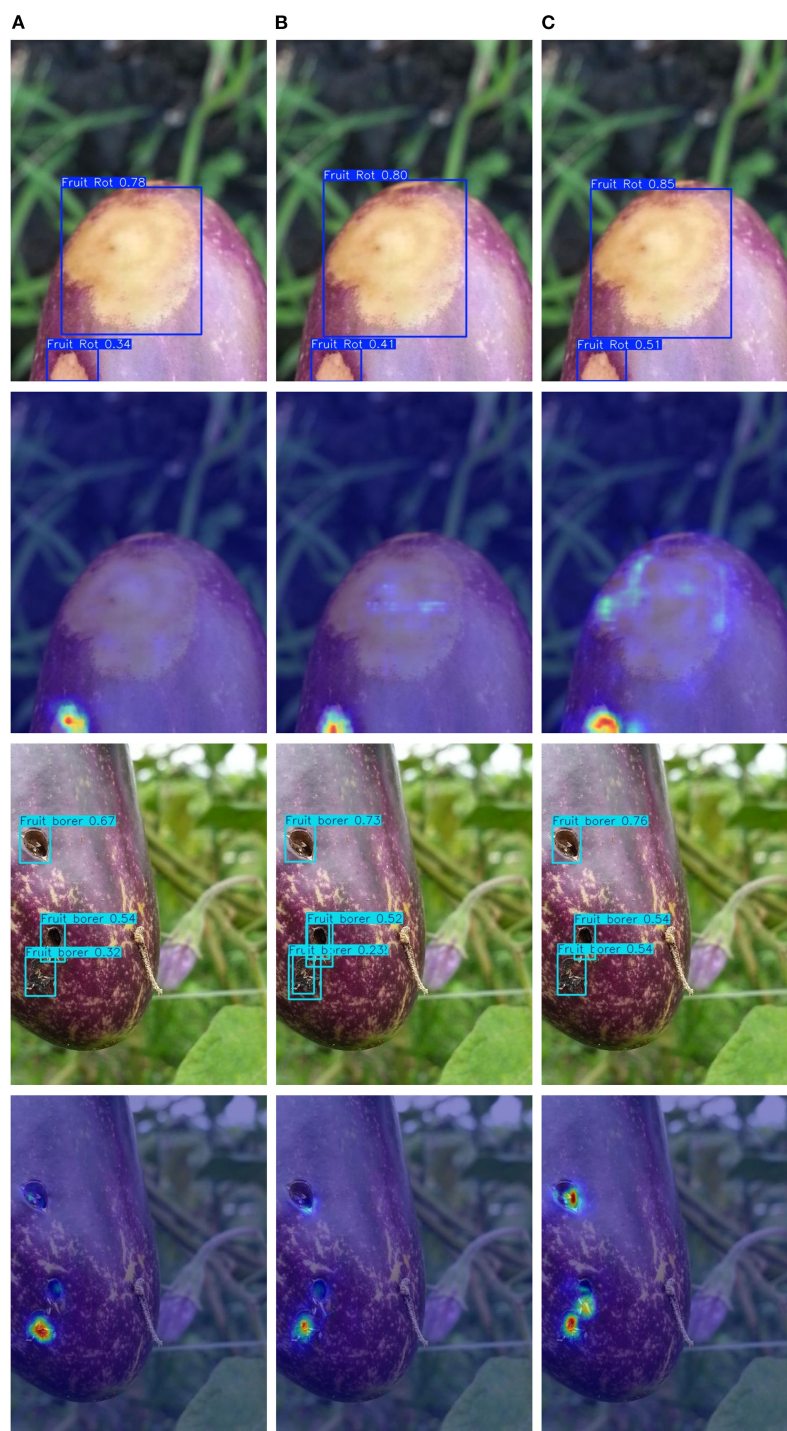


FIGURE 9

The visualization detection results of different models, (A) Baseline Model; (B) YOLO11; (C) EggplantDet.



GFLOPs than the other algorithms. Among single-stage detectors, EggplantDet achieved the best mAP50–95, mAP50, and FPS, while maintaining similar parameter counts and GFLOPs. Compared to the baseline YOLOv8n, EggplantDet achieved 84.3% mAP50 and 50.3% mAP50–95, with respective improvements of 4.7% and 7.2%. Notably, EggplantDet outperformed the state-of-the-art YOLO11n by 0.017 in mAP50, 0.027 in mAP50–90, and 19.7 in FPS. Figures 7A, B visually demonstrate the mAP comparison between EggplantDet and the baseline model on the enhanced eggplant dataset, indicating EggplantDet's excellent performance throughout the process. In conclusion, the improved EggplantDet network demonstrates excellent performance in both detection accuracy and speed, possessing high practical value.

### 3.3 Ablation studies

To evaluate the effectiveness of the proposed modules, YOLOv8 was adopted as the baseline model, and each module was tested individually on the enhanced eggplant disease dataset. The ablation results for the proposed modules are presented in Table 4. Initially, the CSP-MSEIE, MCRPN, and MCEM modules were introduced individually. Each module contributed to improvements in detection performance. Specifically, introducing CSP-MSEIE alone yielded the highest improvement in mAP50, whereas MCRPN contributed the most to mAP50–95. Subsequently, the modules were combined in pairs. All three combinations further enhanced detection performance, demonstrating strong synergy among the modules. Notably, the combination of CSP-MSEIE and MCRPN resulted in mAP50–95, mAP50, and FPS increasing to 49.3%, 83.7%, and 266.5, respectively. Finally, all three modules were integrated simultaneously. As shown in the last row of Table 4, combining the proposed modules improved the model's mAP50 and mAP50–95 by 4.7% and 7.2%, respectively, with FPS reaching 270.5. This further confirms the efficacy of the proposed CSP-MSEIE, MCRPN, and MCEM modules in detecting eggplant diseases.

### 3.4 Visual comparative studies

Figures 8, 9 demonstrate EggplantDet's detection results compared with those of the baseline model and advanced model on the eggplant disease dataset. The figures visually confirm the proposed detection network's advantages. The results show that EggplantDet achieves the highest detection accuracy across all four categories, significantly outperforming both the baseline and advanced model YOLO11n. These results suggest that the proposed model outperforms others in eggplant disease detection. Consequently, it offers a promising solution for crop disease detection tasks.

## 4 Conclusion

Crop pest and disease detection technology provides strong support for the development of smart agriculture. To address

challenges such as disease scale variations, blurred edge features, and background interference in eggplant diseases, this paper proposes an eggplant disease detection network based on multi-scale edge feature enhancement (EggplantDet), which effectively improves the detection accuracy and localization precision of diseased areas while enhancing detection speed. In the feature extraction stage, the CSP-MSEIE module is incorporated to capture hierarchical features, and the EdgeEnhancer module is used to extract edge information, thereby enhancing the network's sensitivity to edges. In the feature processing stage, the MCRPN network captures multi-scale contextual information in horizontal and vertical directions and obtains axial global context to explicitly model rectangular key regions, effectively integrating feature information from different levels. Finally, a range of data augmentation techniques is applied to enhance the eggplant disease dataset, thereby boosting the detection model's ability to generalize. The enhanced detection network outperforms the advanced object detection model YOLO11n, in both detection accuracy and speed. In the future, we will continue to research disease detection networks for more crop varieties and explore lightweight and efficient pest and disease detection technologies to accelerate the transformation of research results into precision crop cultivation applications.

Future work will focus on extending disease detection networks to additional crop species and exploring lightweight, efficient detection technologies to accelerate the deployment of intelligent pest and disease monitoring in precision agriculture.

### Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://public.roboflow.com/object-detection/plantdoc/>; <https://universe.roboflow.com/bohol-island-state-university-vgjl/eggplant-disease-detection>.

### Author contributions

HS: Investigation, Formal Analysis, Methodology, Writing – original draft, Data curation, Conceptualization. RF: Formal Analysis, Data curation, Writing – review & editing, Methodology. DK: Writing – review & editing, Funding acquisition, Formal Analysis, Supervision.

### Funding

The author(s) declare financial support was received for the research and/or publication of this article. This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2022R1A2C2012243).

## Acknowledgments

The authors would like to acknowledge the contributions of the participants in this study.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## References

- Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv*. Available online at: <https://arxiv.org/abs/2004.10934>.
- BSCS (2024). Eggplant disease detection computer vision project. Available online at: <https://universe.roboflow.com/bohoh-island-state-university-vgjl/eggplant-disease-detection>.
- Jiang, P., Qi, A., Zhong, J., et al. (2024). Field cabbage detection and positioning system based on improved YOLOv8n. *Plant Methods* 20, 96. doi: 10.1186/s13007-024-01226-y
- Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., et al. (2022). Ultralytics/yolov5: V6.2 - yolov5 classification models, apple m1, reproducibility, clearml and decim.ai integrations. doi: 10.5281/zenodo.7002879
- Jocher, G., and Qiu, J. (2024). Ultralytics yolo11.
- Krishnaswamy Rangarajan, A., and Purushothaman, R. (2020). Disease classification in eggplant using pre-trained VGG16 and MSVM. *Sci. Rep.* 10, 2322. doi: 10.1038/s41598-020-59108-x
- Li, C., Li, L., Jiang, H., Weng, K., Geng, Y., Li, L., et al. (2022). YOLOv6: A single-stage object detection framework for industrial applications. *arXiv*. Available online at: <https://arxiv.org/abs/2209.02976>.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "SSD: Single shot multibox detector," in *European conference on computer vision (Springer)*. 21–37.
- Liu, J., and Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. *Plant Methods* 17, 22. doi: 10.1186/s13007-021-00722-9
- Liu, J., and Wang, X. (2024). Multisource information fusion method for vegetable disease detection. *BMC Plant Biol.* 24, 738. doi: 10.1186/s12870-024-05346-4
- Liu, J., Wang, X., Zhu, Q., and Miao, W. (2023). Tomato brown rot disease detection using improved YOLOv5 with attention mechanism. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1289464
- Maggay, J. (2025). Detecting affect states using VGG16, ResNet50 and se-ResNet50 networks — sn computer science. *SN Comput. Sci.* doi: 10.1007/s42979-020-01114-9
- Math, R. M., and Dharwadkar, N. V. (2023). Deep learning and computer vision for leaf miner infestation severity detection on muskmelon (cucumis melo) leaves. *Comput. Electrical Eng.* 110, 108843. doi: 10.1016/j.compeleceng.2023.108843
- Ni, Z., Chen, X., Zhai, Y., Tang, Y., and Wang, Y. (2024). Context-guided spatial feature reconstruction for efficient semantic segmentation. *arXiv*. Available online at: <https://arxiv.org/abs/2405.06228>.
- Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. *CVPR*, 779–788. Available online at: [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/html/Redmon\\_You\\_Only\\_Look\\_CVPR\\_2016\\_paper.html](https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html).
- Redmon, J., and Farhadi, A. (2017). YOLO9000: Better, faster, stronger. *CVPR*, 7263–7271. Available online at: [https://openaccess.thecvf.com/content\\_cvpr\\_2017/html/Redmon\\_YOLO9000\\_Better\\_Faster\\_CVPR\\_2017\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2017/html/Redmon_YOLO9000_Better_Faster_CVPR_2017_paper.html).
- Redmon, J., and Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv*. Available online at: <https://arxiv.org/abs/1804.02767>.
- Redmon, J., and Farhadi, A. (2025). YOLOv8 - ultralytics yolo doc. Available online at: <https://docs.ultralytics.com/zh/models/yolov8/> (Accessed April 10, 2025).
- Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi: 10.1109/TPAMI.2016.2577031
- Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S., and Batra, N. (2019). PlantDoc: A dataset for visual plant disease detection. *ArXiv*. doi: 10.1145/3371158.3371196
- Theckedath, D., and Sedamkar, R. (2025). Mobile-based eggplant diseases recognition system using image processing techniques. *Int. J. Advanced Trends Comput. Sci. Eng.* Available online at: [https://www.researchgate.net/publication/339766052\\_Mobile-Based\\_Eggplant\\_Diseases\\_Recognition\\_System\\_using\\_Image\\_Processing\\_Techniques](https://www.researchgate.net/publication/339766052_Mobile-Based_Eggplant_Diseases_Recognition_System_using_Image_Processing_Techniques).
- Tian, L., Zhang, H., Liu, B., Zhang, J., Duan, N., Yuan, A., et al. (2023). Vmf-ssd: A novel v-space based multi-scale feature fusion SSD for apple leaf disease detection. *IEEE/ACM Trans. Comput. Biol. Bioinf.* 20, 2016–2028. doi: 10.1109/TCBB.2022.3229114
- Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv*. Available online at: <https://arxiv.org/abs/2207.02696>.
- Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., et al. (2024a). YOLOv10: Real-time end-to-end object detection. *Adv. Neural Inf. Process. Syst.* 37, 107984–108011.
- Wang, X., and Liu, J. (2024). An efficient deep learning model for tomato disease detection. *Plant Methods* 20, 61. doi: 10.1186/s13007-024-01188-1
- Wang, H., Sun, S., Chang, L., Li, H., Zhang, W., Frery, A. C., et al. (2024c). Inspiration: A reinforcement learning-based human visual perception-driven image enhancement paradigm for underwater scenes. *Eng. Appl. Artif. Intell.* 133, 108411. doi: 10.1016/j.engappai.2024.108411
- Wang, H., Sun, S., and Ren, P. (2023). Underwater color disparities: Cues for enhancing underwater images toward natural color consistencies. *IEEE Trans. Circuits Syst. Video Technol.* 34, 738–753. doi: 10.1109/TCSVT.2023.3289566
- Wang, C.-Y., Yeh, I.-H., and Liao, H.-Y. (2024b). YOLOv9: Learning what you want to learn using programmable gradient information. *arXiv*. Available online at: <https://arxiv.org/abs/2402.13616>.
- Wu, D. (2018). Early detection of botrytis cinerea on eggplant leaves based on visible and near-infrared spectroscopy. *Trans. ASABE*. doi: 10.13031/2013.24504
- Xie, C., and He, Y. (2016). Spectrum and image texture features analysis for early blight disease detection on eggplant leaves. *Sensors* 16, 676. doi: 10.3390/s16050676
- Yan, Z., Liang, Y., Li, Z., Lin, D., Dou, H., Li, N., et al. (2024). Wax patterns, textural properties, and quality attributes of two eggplant (solanum melongena L.) cultivars during storage. *Trans. ASABE HortScience*.
- Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., et al. (2024). "Detrs beat yolos on real-time object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 16965–16974.