Method for lightweight tomato leaf disease recognition based on improved YOLOv11s

Dixin Chen[1], Li Liu[1], Long Zhao[1], Yi Shi[2], Fei Wang[2], Dan Meng[3*], Jianshe Zhao[4*]
(1. *College of Horticulture and Plant Protection, Henan University of Science and Technology, Luoyang 471003, China*;
2. *College of Agricultural Equipment Engineering, Henan University of Science and Technology, Luoyang 471003, China*;
3. *Chinese Society of Agricultural Engineering, Beijing 100125, China*;
4. *Henan Zhongyuan Organic Agriculture Research Institute Co., Ltd., Zhengzhou 450000, China*)

Abstract: Accurate detection of tomato leaf diseases is crucial for early prevention and ensuring agricultural production. This study addresses six tomato leaf diseases: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, and yellow leaf curl virus. A lightweight detection model, YOLO-LGS, is proposed to achieve efficient and automated disease detection. The dataset of tomato leaf diseases was first augmented to enrich the disease features, thereby improving the model's detection performance. The YOLO-LGS model is built on the YOLOv11 architecture, incorporating lightweight group attention net (LWGANet) to reconstruct the backbone network, replacing the convolutional block with parallel spatial attention mechanism with the grouped channel-wise self-attention (GCSA) mechanism, and introducing separated and enhanced attention module (SEAM) into the detection head to balance performance and efficiency. Experimental results show that the YOLO-LGS model achieves an mAP50 of 0.693 and an F1 score of 0.677, outperforming other YOLO models (YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s). Additionally, the model's parameter size is only 6.333 M, and its GFLOPs is 13.4, representing reductions of 32.739% and 37.089%, respectively, compared to YOLOv11s, significantly lowering computational cost while maintaining detection performance. The results demonstrate the effectiveness of LWGANet, GCSA, and SEAM. The development of the YOLO-LGS model provides an efficient, lightweight solution for tomato leaf disease detection in resource-constrained environments.

Keywords: YOLO, LWGANet, object detection, lightweight model
DOI: 10.25165/j.ijabe.20251805.9928

Citation: Chen D X, Liu L, Zhao L, Shi Y, Wang F, Meng D, et al. Method for lightweight tomato leaf disease recognition based on improved YOLOv11s. Int J Agric & Biol Eng, 2025; 18(5): 298–305.

1 Introduction

With the rapid development of global agriculture, accurate identification and control of crop diseases have become key to ensuring food security and sustainable agricultural development[1,2]. As one of the most important economic crops worldwide, tomato disease poses a significant threat to both yield and quality. Traditional disease detection methods mainly rely on manual observation, which is not only inefficient but also subject to subjective biases[3,4]. Furthermore, some laboratory-based detection methods often require specialized equipment and personnel, with long detection cycles that fail to meet the demands for rapid, large-scale, and high-precision field diagnostics[5]. Therefore, developing an efficient, accurate, and automated tomato disease recognition technology is of significant practical importance.

In recent years, with the rapid development of computer vision technology, deep learning-based image recognition algorithms have been widely applied in agricultural disease detection[6,7]. There are mainly two types of deep learning detection networks: the R-CNN-based two-stage detection network and the single-stage network represented by YOLO series algorithms. The R-CNN-based two-stage detection network decomposes the detection task into two key stages: candidate region extraction and region classification. This staged approach effectively improves detection accuracy and efficiency[8-10]. For large-scale crop disease detection, Hua et al.[11] integrated multiple features in an R-CNN model to improve disease detection accuracy and efficiency, enabling fast and accurate identification of crop surface diseases. Deng et al.[12] proposed an orchard disease detection method based on federated learning and improved Faster R-CNN, significantly improving accuracy in small target disease detection and diverse pest identification in complex environments. To enhance the detection accuracy of Faster R-CNN, Alruwaili et al.[13] improved the model by reducing network layers and optimizing the regions of interest pooling layer, developing a real-time tomato leaf disease detection system with an accuracy of 0.974. Kaur et al.[14] proposed an improved Mask R-CNN model for tomato leaf disease detection by adjusting anchor ratios in the region proposal network and optimizing the backbone network structure, utilizing an ensemble deep learning technique to effectively segment disease areas on leaves.

Although two-stage networks have shown high accuracy in

Received date: 2025-05-22 **Accepted date:** 2025-09-08

Biographies: Dixin Chen, PhD, Associate Professor, research interest: horticultural plant cultivation and post harvest physiology, Email: cdxdyp@haust.edu.cn; **Li Liu**, MS candidate, research interest: genetic analysis of fruit quality, Email: 15224715725@163.com; **Long Zhao**, PhD, Associate Professor, research interest: agricultural informatization, smart agriculture, Email: hkdzhaolong@163.com; **Yi Shi**, PhD, Lecturer, research interest: agricultural informatization, smart agriculture, Email: shiyigongteng@163.com; **Fei Wang**, MS candidate, research interest: agricultural informatization, Email: wf01082021@163.com.

***Corresponding author: Dan Meng**, MS, Engineer, research interest: agricultural informatization. Chinese Society of Agricultural Engineering, Beijing 100125, China. Email: mengd202502@163.com; **Jianshe Zhao**, Professor, research interest: Organic cultivation of tomatoes, Henan Zhongyuan Organic Agriculture Research Institute Co., Ltd., Zhengzhou 450000, China. Email: 121936189@163.com.

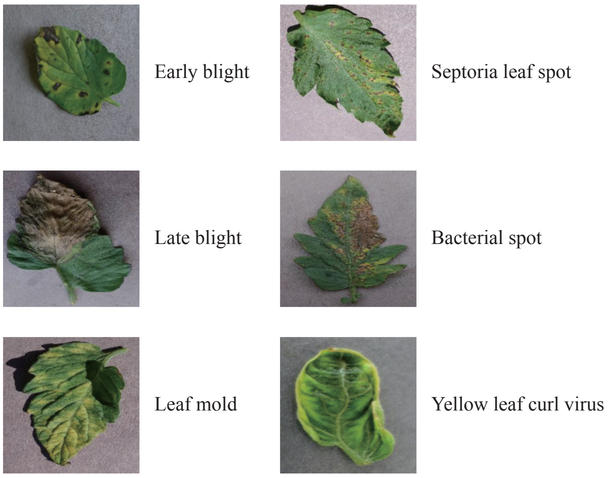
diversified disease detection, they require networks that can handle large amounts of image data within short timeframes, maintaining both real-time performance and accuracy, especially in resource-constrained environments. In comparison, YOLO series object detection networks have received widespread attention in many fields due to their efficiency and fast response capabilities[15-17]. YOLO series models perform end-to-end training, allowing direct prediction on input images and outputting target class and location information, effectively reducing computational load and detection time[18,19]. Researchers have improved YOLO networks by optimizing modules and attention mechanisms, significantly enhancing model detection accuracy and achieving further lightweighting. Wang and Liu[20] developed an efficient deep learning model, TomatoDet, by optimizing the feature extraction module, activation functions, and feature fusion structure of the YOLOv8n model, resulting in an 8.700% improvement in mAP, enabling high-precision real-time tomato disease detection. Kang et al.[21] based their model on YOLOv8, introducing the Ghost module and convolutional block attention module to improve the network architecture, developing a lightweight tomato growth monitoring algorithm (YOLO-TGI) with a minimum checkpoint weight of 3.7 MB. Lin et al.[22] optimized the YOLOv4 model using channel attention mechanism and introduced an osprey search strategy for hyperparameter optimization, achieving a 0.863 mAP. For small or dense disease targets, Xu et al.[23] proposed the Pruned-YOLO v5s+Shuffle (PYSS) model, which can perform real-time disease detection in complex greenhouse environments, improving detection for small target diseases, with a model size of 1.1 MB. Wang et al.[24] introduced lightweight convolutional neural networks and group shuffle convolution into the YOLO framework to efficiently detect small target diseases while reducing model size, achieving a parameter size of 2.96 M.

This study focuses on detecting and recognizing tomato leaf diseases, targeting six disease types: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, and yellow leaf curl virus. These diseases frequently appear during the tomato growth cycle, with complex and diverse symptoms. The disease locations are often small and dense, presenting significant challenges for accurate recognition. To address these issues, the study builds a detection model based on the latest YOLO series algorithm YOLOv11, capable of efficiently recognizing diverse disease symptoms on tomato leaves. The improvements in this study are as follows: 1) the introduction of lightweight group attention net (LWGANet) to reconstruct the YOLOv11 backbone network, reducing the model's parameter size while retaining efficient feature extraction capabilities; 2) the addition of a novel attention mechanism, grouped channel-wise self-attention (GCSA), to replace the original convolutional block with parallel spatial attention (C2PSA) module in YOLOv11, improving the model's detection accuracy for dense and small disease spots; and 3) the introduction of separated and enhancement attention module (SEAM) in the detection head to dynamically adjust the importance of feature channels. Through these improvements, this study aims to provide an efficient, accurate, and widely applicable technical solution for automated tomato disease recognition, offering strong support for early diagnosis and precise prevention of tomato diseases.

2 Materials and methods

2.1 Dataset construction

The dataset used in this study is based on a publicly available tomato leaf disease dataset (https://github.com/ZhouGuoXiong/PDC-VLD), comprising 1782 raw images annotated in PASCAL VOC format. The dataset includes samples of six tomato leaf diseases: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, and yellow leaf curl virus, with image examples shown in Figure 1. While the original dataset contains rich disease features, it lacks the diversity of disease symptoms found in natural environments. Therefore, this study employed data augmentation techniques on the original dataset to enrich its content and improve the model's generalization ability. The augmented images are shown in Figure 2. The data augmentation techniques used include random light intensity adjustment, random flipping, random masking, random rotation, and noise addition, simulating the actual detection environment of tomato leaf diseases and improving the model's detection accuracy and adaptability to complex environments. After data augmentation, the dataset contains 8910 images, which were randomly divided into training, testing, and validation sets with a ratio of 8:1:1. The training set consists of 7136 images, the testing set contains 890 images, and the validation set contains 884 images.


Early blight

Septoria leaf spot

Late blight

Bacterial spot

Leaf mold

Yellow leaf curl virus

Figure 1 Images of tomato leaf diseases

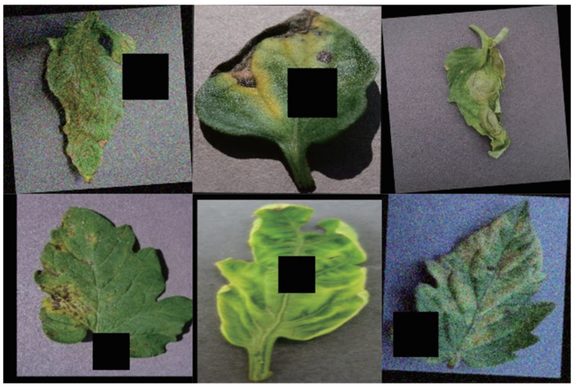


Figure 2 Examples of augmented dataset images

2.2 YOLOv11 model

As the latest iteration of the YOLO series object detection models, YOLOv11 achieves significant performance improvements through a series of architectural innovations[25]. One of its primary innovations is the cross stage partial with kernel size 2 (C3K2) mechanism, an optimized implementation of the CSP structure. Inherited from C2f, this mechanism utilizes smaller convolution kernels and an optimized structural design to achieve lightweight computation while effectively extracting deep-level features, further enhancing computational efficiency and feature extraction flexibility[26,27]. The second major innovation in YOLOv11 is the C2PSA module, which introduces a spatial attention mechanism to help the model focus more effectively on key feature regions within an image. This enhances feature representation in critical areas, improving the model's ability to recognize small or partially occluded objects[28,29]. Additionally, YOLOv11 incorporates depthwise separable convolutions in the detection head, significantly reducing computational complexity and parameter count. By decomposing standard convolution into depthwise and pointwise convolutions, the model reduces computational overhead while maintaining accuracy[30]. This design improves inference speed, making YOLOv11 more suitable for deployment on edge devices and in resource-constrained environments.

2.3 Improvement strategies

2.3.1 Backbone reconstruction using LWGANet

To enhance large-scale tomato leaf disease detection, this study reconstructs the original YOLOv11 backbone using LWGANet, further improving model lightweighting. LWGANet is a novel lightweight backbone network designed to efficiently extract multi-scale object features with lower computational complexity and fewer parameters[31]. The core component of LWGANet is the lightweight group attention (LWGA) module, which divides input feature maps into multiple submodules and applies different levels of attention, including gate point attention, regular local attention, sparse medium-range attention, and sparse global attention. This design optimizes feature representation, enabling precise feature extraction across multiple scales within a lightweight computational framework. LWGANet consists of four stages, each containing multiple LWGA modules that progressively downsample feature maps to accommodate various object scales in detection. Additionally, it incorporates a deep robust feature downsampling module to preserve details of small targets.

In this study, the L1-scale LWGANet was selected, which includes 1, 2, 4, and 2 LWGA modules across the four stages. This configuration captures both local and global feature information while balancing detection performance and computational efficiency. Compared to the L0-scale, it avoids excessive lightweighting that could degrade performance, while compared to the L2-scale, it prevents high computational cost and overfitting. The L1-scale LWGANet structure is illustrated in Figure 3.

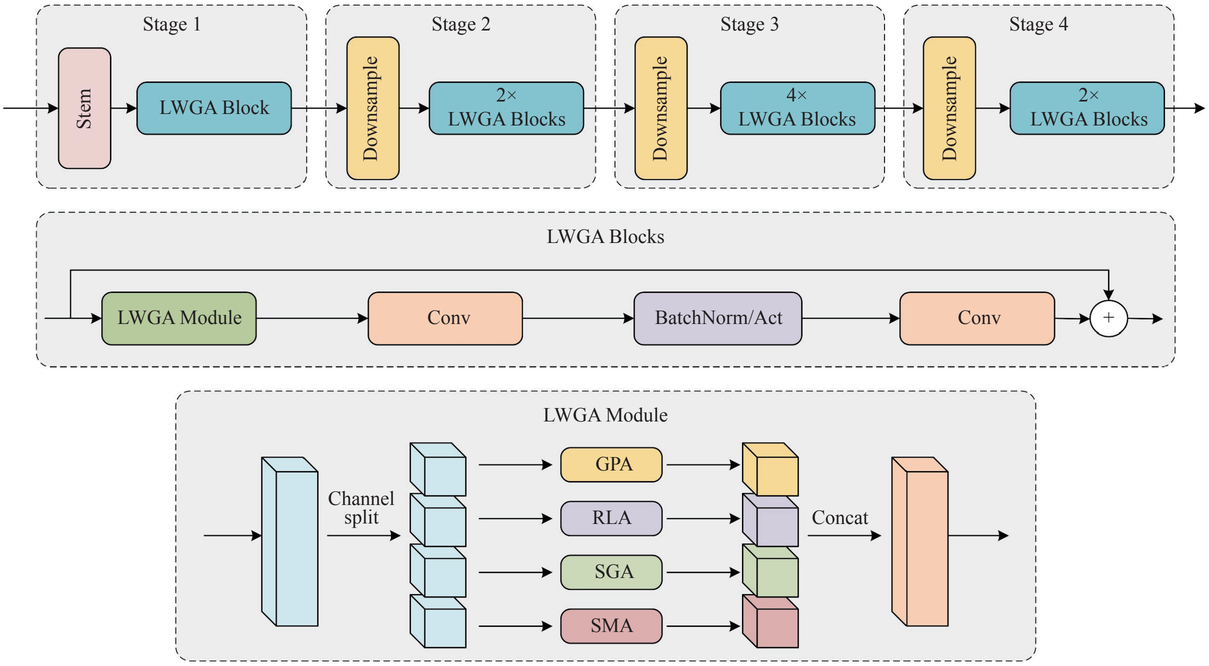


Figure 3 LWGANet architecture at the L1 scale

2.3.2 Replacing C2PSA with GCSA mechanism

This study replaces the original C2PSA mechanism in YOLOv11's backbone network with the GCSA mechanism, whose structure is shown in Figure 4. GCSA is an innovative channel attention mechanism designed to address potential blind spots in multi-scale architectures[32]. GCSA partitions deep feature channels into multiple groups, applying independent channel attention operations within each group. This grouped attention mechanism maintains global channel interaction capabilities while optimizing computational efficiency, allowing the model to adapt better to different input data types and improving generalization. By controlling the number of channels in each group to be smaller than the spatial resolution, GCSA prevents spatial information leakage caused by channel interactions, ensuring more precise target recognition. This mechanism provides an effective solution for complex object detection tasks, particularly for images with intricate backgrounds and noise, significantly enhancing detection accuracy and robustness.

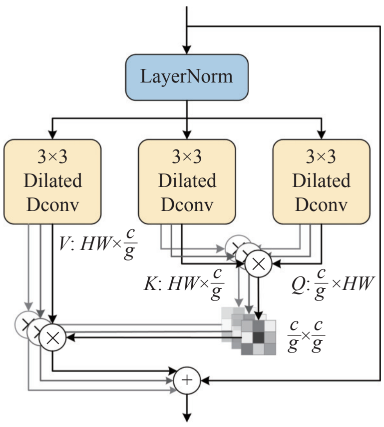


Figure 4 Structure of the GCSA mechanism

2.3.3 Integrating SEAM into the detection head

In the detection head, this study integrates the SEAM into YOLOv11, forming SEAMDetect to enhance the model's target feature perception and improve detection performance. The architecture of SEAMDetect and SEAM is shown in Figure 5. SEAM combines depthwise separable convolution and channel attention mechanisms to enhance feature representation through local feature extraction and channel weighting[33]. It first extracts local features using depthwise separable convolutions, incorporating residual connections to enhance feature stability. Then, global average pooling compresses feature maps into global feature vectors, and two fully connected layers implement the channel attention mechanism[34]. To further improve feature distinction, SEAM employs an exponential weighting normalization strategy, increasing the model's tolerance to feature localization errors[35]. Finally, element-wise multiplication fuses attention weights with the original feature maps, effectively enhancing target features. By embedding SEAM into the YOLOv11 detection head, this study improves the model's ability to focus on target features while maintaining the efficiency of YOLOv11, leading to enhanced detection performance.

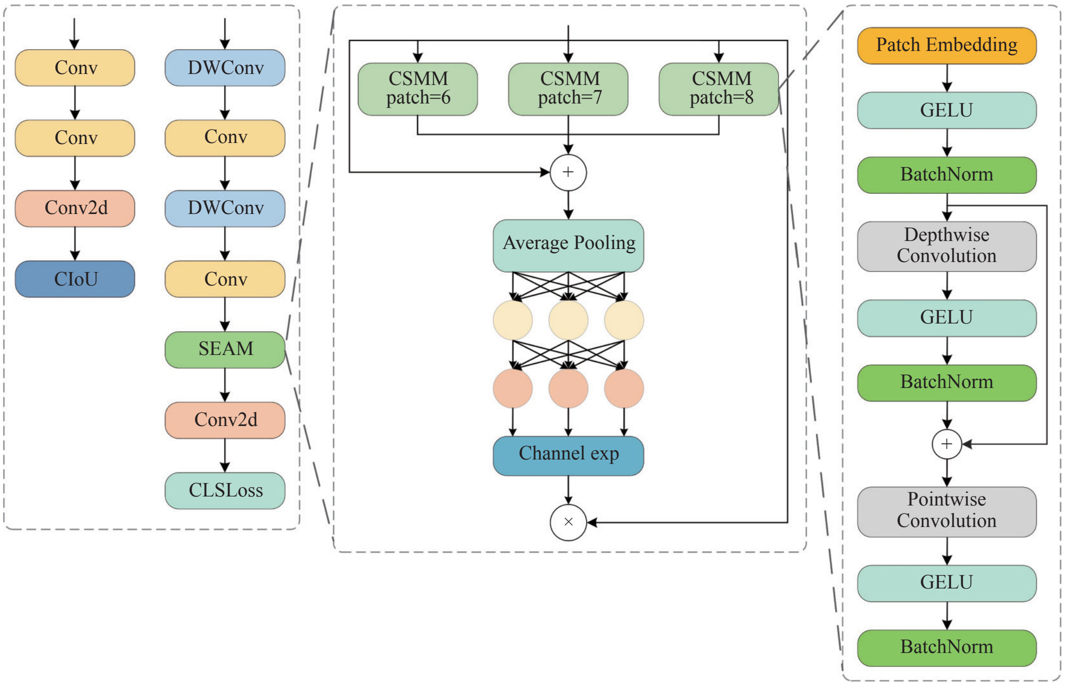


Figure 5 Structural diagrams of SEAM and SEAMDetect modules

2.4 Tomato leaf disease detection model

Since the tomato leaf disease dataset used in this study is large-scale and the original YOLOv11 model has high computational demands, it is unsuitable for deployment in resource-limited environments. To address this, this study develops a lightweight YOLO-LGS model, optimized for multiple types of tomato leaf diseases.

The model is specifically designed based on the disease characteristics and employs the lightweight LWGANet to reconstruct the YOLOv11s backbone network. Additionally, to enhance feature extraction, the GCSA mechanism replaces C2PSA, and SEAM is integrated into the detection head. These improvements enhance the model's ability to recognize diverse disease features, improve detection accuracy, and reduce model complexity. The overall structure of the YOLO-LGS model is illustrated in Figure 6.

2.5 Evaluation metrics

This study evaluates the performance of the tomato leaf disease detection model using the following four key metrics: Precision (*P*): The proportion of correctly predicted positive samples among all predicted positive samples. Recall (*R*): The proportion of correctly predicted positive samples among all actual positive samples. Mean average precision (mAP50): The mean average precision when the intersection over union (IoU) threshold is set to 0.5. F1 Score: The harmonic mean of *P* and *R*, where a higher F1 score indicates a better detection network performance. The formulas for these four metrics are as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \tag{2}$$

$$\text{mAP}_{50} = \frac{1}{m}\sum_{i=1}^{m}\text{AP}_i \tag{3}$$

$$\text{F1} = 2\times\frac{\text{precision}\times\text{recall}}{\text{precision}+\text{recall}} \tag{4}$$

where, TP denotes the number of samples correctly detected as positive categories, FP denotes the number of samples incorrectly detected as positive categories, and FN denotes the number of samples incorrectly detected as negative categories. AP_i denotes the average precision of the category i when the IoU is 0.5.

3 Results

3.1 Experimental setup

The hardware configuration used in this study includes an Intel(R) Xeon(R) Gold 5318Y CPU @ 2.10 GHz processor and an NVIDIA A16 GPU. The software environment consists of Python 3.10, the PyTorch 2.0.1 framework, and CUDA 11.8, ensuring efficient network training.

For training parameters, the study adopts the default YOLO network configuration, with input images at a resolution of 256×256 and a training batch size of 200. The initial and final learning rates are both 0.01, allowing the network to fine-tune parameters in the

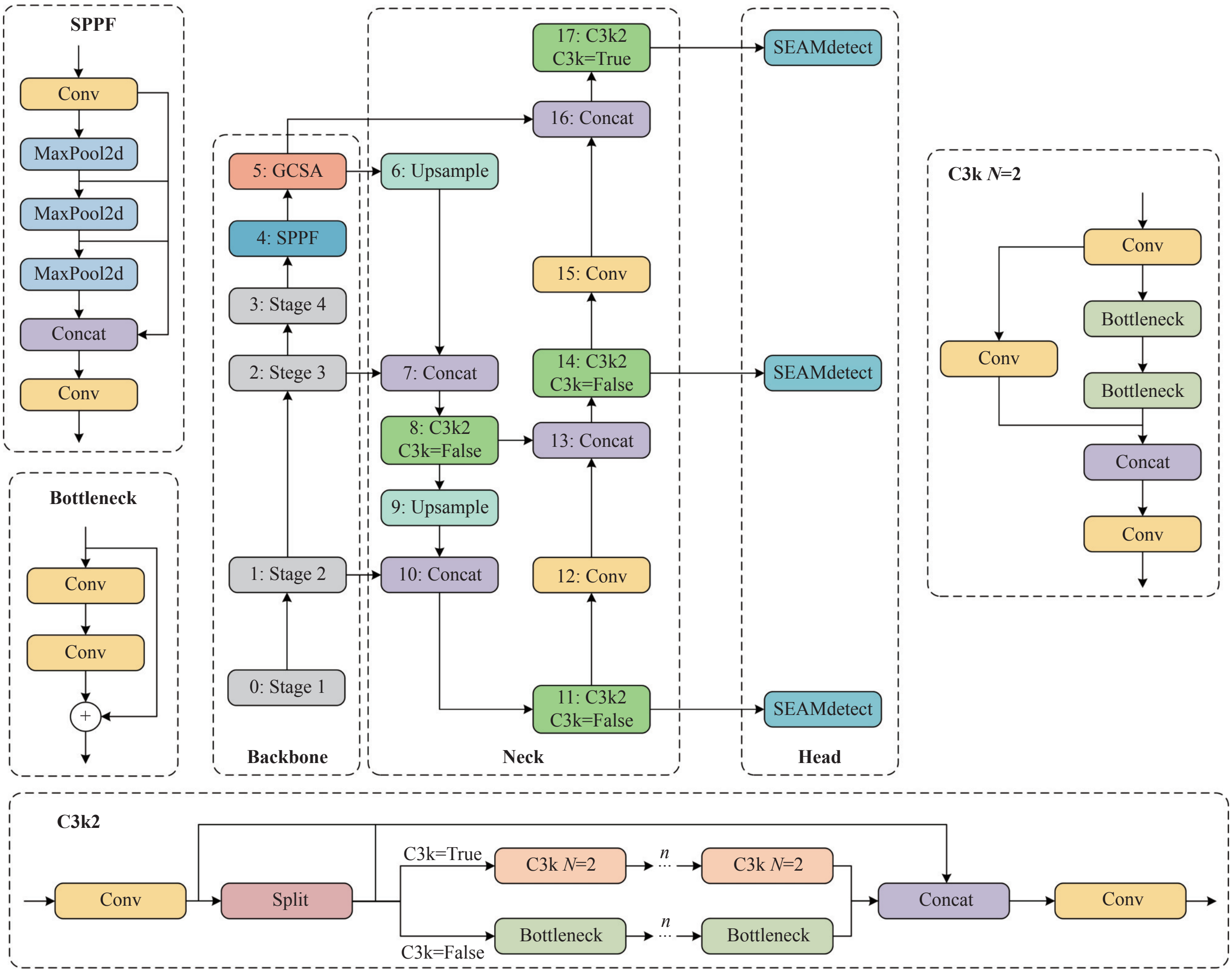


Figure 6 Structure of the YOLO-LGS model

later training stages and avoid overfitting. The momentum is set to 0.937, which accelerates the gradient descent process while reducing oscillations during training. The weight decay is set to 0.0005, effectively mitigating overfitting tendencies while maintaining high training efficiency.

3.2 Performance of the YOLO-LGS model

The improved YOLO-LGS model was used to detect and classify six types of tomato leaf diseases: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, and yellow leaf curl virus. The model's detection performance is presented in Table 1. The results indicate that the overall P is 0.746, R is 0.619, mAP50 is 0.693, and F1 score is 0.677, with parameters and GFLOPs at 6.333 MB and 13.4, respectively. It demonstrates strong performance in tomato leaf disease detection.

Table 1 Detection accuracy of YOLO-LGS for different types of tomato leaf diseases

Disease types	Precision	Recall	mAP50	F1 score
All	0.746	0.619	0.693	0.677
Late blight	0.663	0.703	0.717	0.682
Early blight	0.833	0.786	0.857	0.809
Leaf mold	0.781	0.729	0.781	0.754
Septoria leaf spot	0.738	0.489	0.575	0.588
Yellow leaf curl virus	0.726	0.497	0.620	0.590
Bacterial spot	0.735	0.510	0.605	0.602

Figure 7 presents the confusion matrix of the YOLO-LGS model's detection results. Overall, the model exhibits excellent classification performance, although detection accuracy varies across different disease categories due to differences in disease characteristics. The model performs particularly well in detecting early blight, with a correct identification rate of 0.820. The distinctive concentric ring patterns on infected leaves make it easier for the model to capture features, leading to high detection accuracy. The recognition rates for late blight and leaf mold are 0.750 and 0.770, respectively, indicating strong performance.

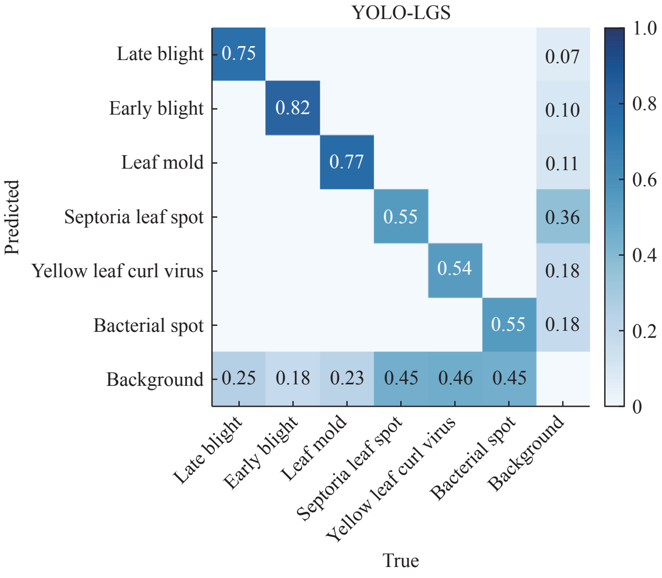


Figure 7 Confusion matrix of YOLO-LGS model detection results

However, the model struggles with diseases characterized by small, dense brown spots, such as bacterial spot and septoria leaf spot, achieving an accuracy of only 0.550. The high false positive and false negative rates are likely due to the similarity of these diseases' visual symptoms, increasing the likelihood of misclassification. The correct detection rate for yellow leaf curl virus is 0.540. This disease causes leaf yellowing and curling, but its characteristics are less distinct, making it more challenging for the model to identify accurately.

3.3 Ablation study

To evaluate the contributions of LWGANet, GCSA, and SEAMDetect, an ablation study was conducted to examine their effects on model performance and verify the effectiveness of the YOLO-LGS model. The results of the ablation experiments in terms of accuracy and computational complexity are summarized in Table 2.

Table 2 Results of the ablation study

Model	LWGANet	GCSA	SEAMDetect	Precision	Recall	mAP50	F1 score	Parameter/M	GFLOPs
1	-	-	-	0.738	0.621	0.688	0.674	9.415	21.3
2	√	-	-	0.714	0.605	0.670	0.655	6.258	13.1
3	-	√	-	0.742	0.644	0.701	0.690	11.320	37.9
4	-	-	√	0.740	0.628	0.690	0.679	11.414	38.5
5	√	√	-	0.735	0.616	0.686	0.670	6.297	13.2
6	√	-	√	0.729	0.607	0.676	0.667	6.293	13.3
7	√	√	√	0.746	0.619	0.693	0.677	6.333	13.4

Replacing YOLOv11s' original backbone with LWGANet did not significantly improve accuracy; however, it reduced the model's parameter size by 33.536% and GFLOPs by 38.498%, highlighting LWGANet's advantage in computational efficiency without substantially affecting precision. Introducing GCSA or SEAMDetect independently resulted in improved mAP50 (0.701, 0.690) and F1 score (0.690, 0.679), respectively, though at the cost of increased parameters and computational load. Combining GCSA with LWGANet led to a 2.389% increase in mAP50 and a 2.290% increase in F1 score, confirming the effectiveness of GCSA in improving model accuracy. When LWGANet and SEAMDetect are used together, the model achieves *P* and F1 scores of 0.729 and 0.667, respectively, with only a slight increase in parameters and GFLOPs. The final YOLO-LGS model successfully balances accuracy and computational complexity. These results indicate that while LWGANet primarily contributes to model lightweighting, GCSA and SEAMDetect significantly enhance detection accuracy, albeit at the expense of additional computational resources.

4 Discussion

4.1 Comparison of detection performance across different models

To validate the effectiveness of the YOLO-LGS model, its performance was compared with other YOLO series models in the tomato leaf disease detection task. Table 3 presents the detection accuracy and model sizes of different models, while Figure 8 illustrates their actual detection performance.

Table 3 Comparative experimental results of different models

Model	Precision	Recall	mAP50	F1 score	Parameter/M	GFLOPs
YOLOv8s	0.726	0.626	0.682	0.672	11.128	28.4
YOLOv9s	0.695	0.578	0.633	0.631	7.169	26.7
YOLOv10s	0.697	0.592	0.655	0.640	8.040	24.5
YOLOv11s	0.738	0.621	0.688	0.674	9.415	21.3
YOLO-LGS	0.746	0.619	0.693	0.677	6.333	13.4

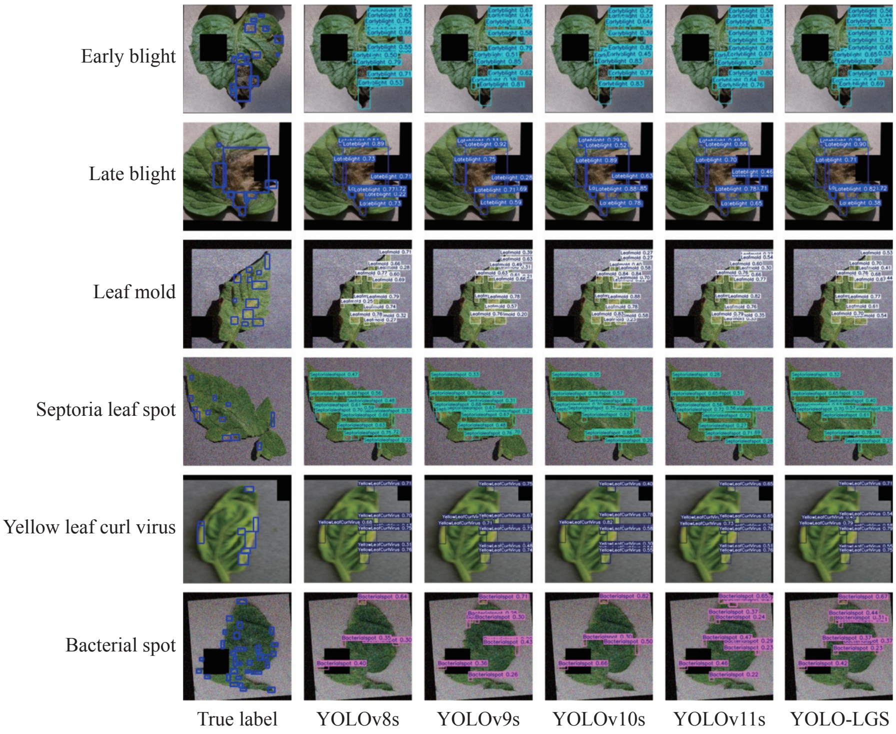


Figure 8 Comparative detection performance of different models

The results indicate that YOLO-LGS achieves a high balance between accuracy and computational efficiency. YOLO-LGS achieves an mAP50 of 0.693, outperforming YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s by 1.613%, 9.479%, 5.802%, and 0.727%, respectively. The F1 score of YOLO-LGS (0.677) is slightly higher than YOLOv8s (0.672) and YOLOv11s (0.674) and is significantly better than YOLOv9s (0.631) and YOLOv10s (0.640). YOLO-LGS exhibits the lowest computational cost, reducing the parameter count by 43.092%, 11.671%, 21.232%, and 32.739% and GFLOPs by 52.817%, 49.813%, 45.306%, and 37.089%, respectively, compared to other YOLO models.

These results confirm that YOLO-LGS is a highly efficient tomato leaf disease detection model, outperforming other models in key metrics while maintaining a lightweight structure. Its optimized architecture makes it suitable for deployment in resource-constrained environments.

4.2 Impact of data augmentation on model performance

Before training the detection network, this study applied data augmentation techniques to the tomato leaf disease dataset, including random light intensity adjustments, flipping, masking, rotation, and noise addition. To evaluate the impact of data augmentation techniques on model performance, this study trained the models both with and without data augmentation. Table 4 presents the training results of different models on the non-augmented dataset, while Figure 9 illustrates the changes in *P* and mAP50 metrics before and after data augmentation. Applying data augmentation significantly improved model accuracy while having minimal impact on model size. After applying data augmentation, YOLOv8s achieved a mAP50 and F1 score improvement of 50.220% and 40.000%, respectively. YOLOv11s saw mAP50 and F1 score improvements of 50.218% and 40.125%, respectively. YOLO-LGS improved by 43.776% in mAP50 and 34.059% in F1 score.

Table 4 Detection performance of different models on the non-augmented dataset

Model	Precision	Recall	mAP50	F1 score	Parameter/M	GFLOPs
YOLOv8s	0.500	0.461	0.454	0.480	11.128	28.4
YOLOv9s	0.539	0.455	0.464	0.493	7.169	26.7
YOLOv10s	0.513	0.442	0.445	0.475	8.040	24.5
YOLOv11s	0.499	0.465	0.458	0.481	9.415	21.3
YOLO-LGS	0.540	0.470	0.484	0.503	6.333	13.4

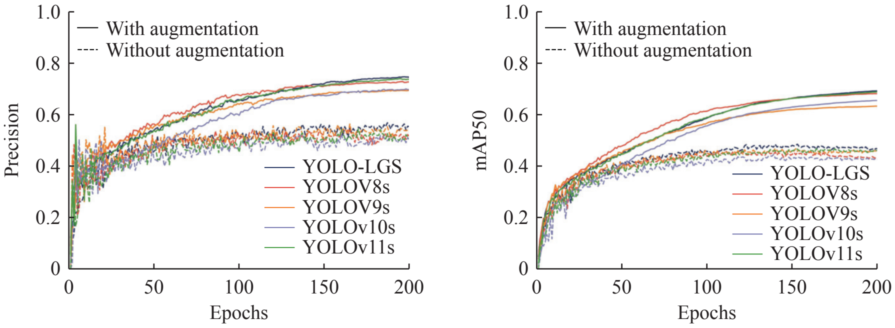


Figure 9 Precision and mAP50 comparison before and after data augmentation

These results indicate that data augmentation significantly enhances model generalization and robustness by exposing the network to diverse variations and noise during training, thereby improving real-world performance[36,37]. Additionally, it helps reduce overfitting to training data, improving the model's performance on test data.

5 Conclusions

To enhance tomato leaf disease detection accuracy, this study proposes a lightweight detection model, YOLO-LGS, optimized for six disease types: bacterial spot, early blight, late blight, leaf mold, septoria leaf spot, and yellow leaf curl virus. Before training, data augmentation techniques were applied to enrich the dataset, effectively increasing sample diversity and improving training accuracy. The study builds upon YOLOv11s, incorporating: LWGANet to reconstruct the backbone network, significantly reducing computational complexity; GCSA to replace C2PSA, improving feature extraction; SEAM in the detection head, enhancing target feature representation. The YOLO-LGS model achieves an mAP50 of 0.693 and an F1 score of 0.677, surpassing YOLOv8s, YOLOv9s, YOLOv10s, and YOLOv11s in accuracy while maintaining the smallest model size (6.333 M, 13.4 GFLOPs).

These findings confirm the effectiveness of LWGANet, GCSA, and SEAM, making YOLO-LGS a high-performance, lightweight solution for tomato leaf disease detection in resource-limited environments. In the future, the model architecture will continue to be optimized, and comprehensive evaluations of its performance across diverse datasets will be conducted. This will enhance its detection capabilities in complex and dynamic environments, thereby improving its applicability in real-world scenarios.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (Grant No.32472723), Natural Science Foundation of Henan Province for Excellent Young Scholars (Grant No. 252300421154), Heluo Youth Talent Project (Grant No. 2024HLTJ11), and Youth Backbone Teacher Project of Henan University of Science and Technology (Grants No. 13450013 and 13450010).

[References]

[1] Zhang S L, Wang J Z, Yang K, Guan M L. YOLO-ACT: An adaptive cross-layer integration method for apple leaf disease detection. Frontiers in Plant Science, 2024; 15: 1451078.

[2] Sajitha P, Andrushia A D, Anand N, Naser M Z. A review on machine learning and deep learning image-based plant disease classification for industrial farming systems. Journal of Industrial Information Integration, 2024; 38: 100572.

[3] Liu W B, Bai C H, Tang W, Xia Y, Kang J. A Lightweight Real-Time recognition algorithm for tomato leaf disease based on improved YOLOv8. Agronomy, 2024; 14(9): 2069.

[4] Tang Z W, He X Y, Zhou G X, Chen A B, Wang Y F, Li L J, et al. A precise image-based tomato leaf disease detection approach using PLPNet. Plant Phenomics, 2023; 5: 0042.

[5] Wang Y R, Wang Y H, Mu J, Raza Mustafa G, Wu Q J, Wang Y, et al. Enhanced multiscale plant disease detection with the PYOLO model innovations. Scientific Reports, 2025; 15(1): 5179.

[6] Too E C, Li Y J, Njuki S, Liu Y C. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 2019; 161: 272–279.

[7] Lin Y S, Huang Z F, Liang Y, Liu Y F, Jiang W P. AG-YOLO: A rapid citrus fruit detection algorithm with global context fusion. Agriculture, 2024; 14(1): 114.

[8] Lin P, Lee W S, Chen Y M, Peres N, Fraisse C. A deep-level region-based visual representation architecture for detecting strawberry flowers in an outdoor field. Precision Agriculture, 2020; 21: 387–402.

[9] Chu P Y, Li Z J, Lammers K, Lu R F, Liu X M. Deep learning-based apple detection using a suppression mask R-CNN. Pattern Recognition Letters, 2021; 147: 206–211.

[10] Behera S K, Rath A K, Sethy P K. Fruits yield estimation using Faster R-CNN with MIoU. Multimedia Tools and Applications, 2021; 80: 19043–19056.

[11] Hua S, Xu M J, Xu Z F, Ye H B, Zhou C Q. Multi-feature decision fusion algorithm for disease detection on crop surface based on machine vision. Neural Computing and Applications, 2022; 34: 9471–9484.

[12] Deng F M, Mao W, Zeng Z Q, Zeng H, Wei B Q. Multiple diseases and pests detection based on federated learning and improved faster R-CNN. IEEE Transactions on Instrumentation and Measurement, 2022; 71: 1–11.

[13] Alruwaili M, Siddiqi M H, Khan A, Azad M, Khan A, Alanazi S. RTF-RCNN: An architecture for real-time tomato plant leaf diseases detection in video streaming using Faster-RCNN. Bioengineering, 2022; 9(10): 565.

[14] Kaur P, Harnal S, Gautam V, Singh M P, Singh S P. An approach for characterization of infected area in tomato leaf disease based on deep learning and object detection technique. Engineering Applications of Artificial Intelligence, 2022; 115: 105210.

[15] Xie Z D, Yang Z, Li C, Zhang Z, Jiang J Z, Guo H Y. YOLO-Ginseng: a detection method for ginseng fruit in natural agricultural environment. Frontiers in Plant Science, 2024; 15: 1422460.

[16] Sun C L, Azman A B, Wang Z Y, Gao X X, Ding K. YOLO-UP: A high-throughput pest detection model for dense cotton crops utilizing UAV-Captured visible light imagery. IEEE Access, 2025; 13: 19937–19945.

[17] Liu Z Y, Zhuo L, Dong C W, Li J F. YOLO-TBD: tea bud detection with triple-branch attention mechanism and self-correction group convolution. Industrial Crops and Products, 2025; 226: 120607.

[18] Wang J X, Liu M, Du Y R, Zhao M H, Jia H L, Guo Z, et al. PG-YOLO: An efficient detection algorithm for pomegranate before fruit thinning. Engineering Applications of Artificial Intelligence, 2024; 134: 108700.

[19] Chen J Q, Wang H B, Zhang H D, Luo T, Wei D P, Long T, et al. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion. Computers and Electronics in Agriculture, 2022; 202: 107412.

[20] Wang X W, Liu J. An efficient deep learning model for tomato disease detection. Plant Methods, 2024; 20: 61.

[21] Kang R, Huang J X, Zhou X H, Ren N, Sun S P. Toward real scenery: a lightweight tomato growth inspection algorithm for leaf disease detection and fruit counting. Plant Phenomics, 2024; 6: 0174.

[22] Lin J W, Hu G, Chen J. Mixed data augmentation and osprey search strategy for enhancing YOLO in tomato disease, pest, and weed detection. Expert Systems with Applications, 2025; 264: 125737.

[23] Xu Y L, Chen Q Y, Kong S L, Xing L, Wang Q, Cong X, et al. Real-time object detection method of melon leaf diseases under complex background in greenhouse. Journal of Real-Time Image Processing, 2022; 19: 985–995.

[24] Wang J L, Qin C C, Hou B B, Yuan Y, Zhang Y K, Feng W F. LCGSC-YOLO: A lightweight apple leaf diseases detection method based on LCNet and GSConv module under YOLO framework. Frontiers in plant science, 2024; 15: 1398277.

[25] Khanam R, Hussain M. Yolov11: An overview of the key architectural enhancements. arXiv: 2410.17725, 2024; In press. doi: 10.48550/arXiv.2410.17725.

[26] Khan A T, Jensen S M. LEAF-Net: A unified framework for leaf extraction and analysis in multi-crop phenotyping using YOLOv11. Agriculture, 2025; 15(2): 196.

[27] Teng H X, Wang Y D, Li W T, Chen T, Liu Q H. Advancing rice disease detection in farmland with an enhanced YOLOv11 algorithm. Sensors, 2025; 25(10): 3056.

[28] Meng H T, Tao M, Huang R D, Xu Y Q, Memon M B. Improved YOLOv11-EGM deep learning model for rock fragment identification. Archives of Civil and Mechanical Engineering, 2025; 25: 223.

[29] Li Y B, Zhang W W, Lyu S T, Yu J, Ge D D, Guo J W, et al. YOLOv11-CAFM model in ground penetrating radar image for pavement distress detection and optimization study. Construction and Building Materials, 2025; 485: 141907.

[30] Weng T H, Niu X P. Enhancing UAV object detection in low-light conditions with ELS-YOLO: A lightweight model based on improved YOLOv11. Sensors, 2025; 25(14): 4463.

[31] Lu W, Chen S B, Ding C H Q, Tang J, Luo B. LWGANet: a lightweight group attention backbone for remote sensing visual tasks. arXiv: 2501.10040, 2025; In press. doi: 10.48550/arXiv.2501.10040.

[32] Li J Y, Zhang Z L, Zuo W M. TBSN: Transformer-based blind-spot network for self-supervised image denoising. arXiv: 2404.07846, 2024; In press. doi: 10.48550/arXiv.2404.07846.

[33] Yu Z P, Huang H B, Chen W J, Su Y X, Liu Y H, Wang X Y. Yolo-facev2: A scale and occlusion aware face detector. Pattern Recognition, 2024; 155: 110714.

[34] Guan S T, Lin Y M, Lin G Y, Su P S, Huang S L, Meng X Y, et al. Real-time detection and counting of wheat spikes based on improved YOLOv10. Agronomy, 2024; 14(9): 1936.

[35] Gui J S, Wu J C, Wu D W, Chen J N, Tong J H. A lightweight tea buds detection model with occlusion handling. Journal of Food Measurement and Characterization, 2024; 18: 7533–7549.

[36] Kaur R, Mittal U, Wadhawan A, Almogren A, Singla, J, Bharany S, et al. YOLO-LeafNet: a robust deep learning framework for multispecies plant disease detection with data augmentation. Scientific Reports, 2025; 15: 28513.

[37] Ayyad S M, Sallam N M, Gamel S A, Ali Z H. Particle swarm optimization with YOLOv8 for improved detection performance of tomato plants. Journal of Big Data, 2025; 12: 152.