Root-YOLOv7: Multi-scale adaptive object detection and grading of root-knot nematode disease

Yao Zhao[1], Honghai Zhao[2], Haitao Xiong[1], Fan Zhang[1], Cheng Lu[3], Juan Li[1*]

(1. *College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao 266109, Shandong, China*;
2. *Shandong Engineering Research Center for Environment-Friendly Agricultural Pest Management, College of Plant Health and Medicine, Qingdao Agricultural University, Qingdao 266109, Shandong, China*;
3. *College of Science and Information Technology, Qingdao Agricultural University, Qingdao 266109, Shandong, China*)

Abstract: Root-knot nematodes can infect over 2000 plants, which causes significant economic losses. Rapid and accurate detection of root-knot nematode disease is the key to screening resistant varieties and evaluating the effect of prevention and control. To address the challenge of detecting root-knot nematode disease caused by dense root knots, multiple fibrous roots, and small root knots, this paper takes cucumber as the research object to propose a detection and grading model Root-YOLOv7. Specifically, the backbone network of YOLOv7 is restructured, which enables the proposed model to better capture object information at different scales through a hierarchical structure and a self-attention module based on shifted windows. Additionally, combined with the Wise-IoU loss function, the proposed model can adaptively adjust the weight of the overlapping part of the object box, which solves the problem that bounding box regression cannot be optimized effectively when detecting low-quality objects. Furthermore, an improved head network structure is proposed to compute the attention weights by capturing the cross-dimension interaction of the root knot feature between the spatial and channel dimensions. To verify the effectiveness of the proposed model, the performance of Root-YOLOv7 is compared with typical object detection models. Experimental results show that the *AP*@0.5 of Root-YOLOv7 reaches 87.40%, which is 72.67%, 5.60%, 5.28%, 9.68%, 5.83%, and 7.45% higher than Faster R-CNN, RT-DETR, YOLOv5, YOLOv6, YOLOv7, and YOLOv8, respectively. The proposed approach is expected to reduce the workload of plant pathologists and provide technical support for the cultivation of plant varieties with disease resistance.

Keywords: object detection, YOLOv7, root-knot nematode disease, deep learning, disease grading
DOI: 10.25165/j.ijabe.20251802.9414

Citation: Zhao Y, Zhao H H, Xiong H T, Zhang F, Lu C, Li J. Root-YOLOv7: Multi-scale adaptive object detection and grading of root-knot nematode disease. Int J Agric & Biol Eng, 2025; 18(2): 259–268.

1 Introduction

As one of the major pathogens of infectious plant diseases, root-knot nematodes have posed a great threat to the production of vegetables, fruit trees, and other crops worldwide, which has led to 50% or even total crop failure[1]. According to the statistics, root-knot nematode diseases cause approximately $160 billion in economic losses worldwide each year, which has become a global issue that plant pathologists cannot ignore[2]. Cultivating resistant varieties is one of the effective measures to prevent and control root-knot nematode disease[3], while the detection, counting, and grading of root knots in plant root systems are prerequisites for evaluating resistant varieties. Currently, the counting of root knots and the grading of disease are performed manually based on experience[4]. This method is not only time-consuming and laborious but also greatly influenced by subjective factors[5]. Consequently, the development of efficient methods for the detection and grading of root-knot nematode disease is an urgent challenge that requires resolution.

Recent advancements in artificial intelligence technology have found extensive application across various domains, including underwater object detection[6], industrial processes[7,8], medical diagnostics[9], food quality inspection[10], and animal behavior recognition[11]. Researchers have also engaged in numerous correlated studies in the field of plant disease detection. For example, Bohnenkamp et al. established a hyperspectral database of wheat leaf diseases caused by various fungal pathogens and employed machine learning techniques for analysis[12]. Gomez-Caro et al.[13] utilized close-range thermal imaging and reflectance spectroscopy to identify *Bacillus anthraci* in mangoes. Principal component analysis and partial least squares regression were used to discriminate spectral responses, and the spectral bands with disease characteristics were obtained. Mohammadpoor et al.[14] utilized the Fuzzy C-means algorithm to highlight the diseased areas on grape leaves and employed support vector machines to ascertain the presence of disease. To detect common guava plant diseases, Almadhor et al.[15] applied the DeltaE model to segment the infected regions, identified the diseases using multiple machine learning classifiers, and proposed an artificial intelligence-driven framework. Ostovar et al.[16] utilized machine learning techniques to realize the automatic detection of Norwegian spruce stump root rot based on

Received date: 2024-08-20 **Accepted date:** 2025-03-06
Biographies: Yao Zhao, MS candidate, research interest: intelligent detection, Email: 2484624212@qq.com; **Honghai Zhao**, PhD, Professor, research interest: plant nematology and crop nematode disease management, Email: hhzhao@qau.edu.cn; **Haitao Xiong**, MS candidate, research interest: intelligent detection, Email: 2115819631@qq.com; **Fan Zhang**, MS candidate, research interest: intelligent detection, Email: 1811242825@qq.com; **Cheng Lu**, Undergraduate, research interest: intelligent detection, Email: 191700672@qq.com.
***Corresponding author: Juan Li**, PhD, Doctoral supervisor, Professor, research interest: machine vision and artificial intelligence, information processing and intelligent monitoring. College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao 266109, Shandong, China. Tel: +86-13698695091, Email: lijuan291@sina.com.

RGB images and classify the stump according to whether it is rotten. Collectively, the aforementioned research studies demonstrate promising results in the detection of plant diseases through traditional machine learning methods. However, these approaches necessitate manual feature extraction and do not have end-to-end learning capabilities.

Aiming at the aforementioned issues, relevant scholars have employed deep learning to realize the detection of plant diseases and pests. For instance, Huang et al. introduced the convolutional block attention module and an adaptive spatial feature fusion structure in YOLOv5s to detect gray planthoppers in crops such as rice and wheat[17]. Li et al. proposed a semantic segmentation model that combined attention mechanism and transfer learning to segment lesions in cucumber leaves and assess the disease severity accurately[18]. Wang et al.[19] enhanced the detection performance of cucumber leaf diseases by integrating an extraction module, which consisted of a Swin Transformer backbone network and gradient-weighted class activation mapping into a generative adversarial network. Wang et al.[20] proposed an intelligent typical method for apple pest identification and enumeration, which solved various types of pest identification challenges. Liu et al.[21] proposed an algorithm for the identification of tomato pests, which mitigated the issue of imbalanced sample sizes in images by enhancements to the attention mechanism and loss function. Qi et al.[22] proposed an improved approach for detecting tomato viruses, which had a good detection effect on the diseased area. Xue et al.[23] utilized deep convolutional neural networks to recognize microscopic images of nematodes and compute ecological indices. Sharma et al.[24] introduced ensemble block sequences in the channel layer to extract deep features and construct a lightweight model for the detection of plant leaf disease.

From the aforementioned studies and related research, it is evident that there exists substantial research focused on the detection of diseases and pests utilizing deep learning techniques. However, investigations specifically addressing root-knot nematode disease are extremely rare. This is because the roots of diseased plants have the characteristics of dense root knots, numerous fibrous roots, and significant differences in root knot sizes, which makes it difficult to count root knots and grade diseases. To the best of our knowledge, only one study has been reported on root knot detection based on deep learning, and this research does not involve root knot counting and grading of disease severity[25]. Considering that cucumber is one of the plants severely affected by root-knot nematode disease, this paper focuses on cucumber as the subject of investigation and proposes a multi-scale adaptive deep learning model for root-knot nematode disease detection and grading. The main contributions of this paper are summarized as follows: (1) A detection model has been developed using datasets of cucumber roots with varying degrees of root-knot nematode disease. (2) The backbone network of YOLOv7 has been reconstructed using a hierarchical construction to process images of diverse scales and complexities, thereby enhancing the generalization ability of the model. (3) An improved network architecture for the head layer has been proposed, which enables the model to effectively focus on key features.

2 Materials and methods

2.1 Image data acquisition

The dataset used in this research is obtained from the Qingdao Dagu River Basin National Agricultural Science and Technology Park, Jimo District, Qingdao City, Shandong Province, China. A total of 100 cucumber plants are cultivated in pots within the greenhouse, with each pot containing an equal volume of nematode-free sandy loam soil. The potted soil is artificially and quantitatively inoculated with the root-knot nematode (*Meloidogyne incognita*). The pertinent experimental conditions are illustrated in Figure 1.

a. Configured root-knot nematode suspension b. Inoculation of nematodes utilizing a pipette

Figure 1 Schematic diagram of root-knot nematode inoculation experiment

The experiment comprises nematode treatment and blank control groups. In the treatment group, the potted soil is inoculated with the southern root-knot nematode at nine different inoculation doses, resulting in nine separate nematode dose treatments. In the blank control group, the soil is inoculated with 100cc of pure water. In the trial, for each nematode dose treatment and blank control, 10 potted plants are set as replicates. The detailed inoculation design is presented in Table 1.

Table 1 Inoculation design of root-knot nematodes

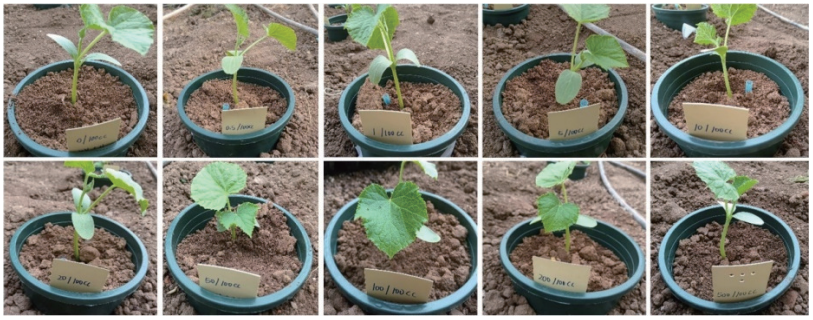
Dose treatment (DT) No.	Root-knot nematode inoculation dose	Number of potted plants
CK	0/100cc	10
DT 1	0.5/100cc	10
DT 2	1/100cc	10
DT 3	5/100cc	10
DT 4	10/100cc	10
DT 5	20/100cc	10
DT 6	50/100cc	10
DT 7	100/100cc	10
DT 8	200/100cc	10
DT 9	500/100cc	10

Following inoculation, cucumber seedlings of uniform size are selected for transplantation, and labels are affixed to each pot for marking and indications, as illustrated in Figure 2.

The extracted roots are washed, and the roots are allowed to naturally unfold against a well-lighted blue background, ensuring that the root surfaces are free from water droplets. The camera (Canon PowerShot SX50 HS) is employed to capture images of the cucumber roots. A total of 519 images with a resolution of 3000 dpi×4000 dpi are obtained through a manual filtering process of the original images. The dataset comprises images of cucumber roots with different degrees of disease severity, as illustrated in Figure 3.

2.2 Data preprocessing

To enhance the diversity of the dataset and mitigate the phenomenon of overfitting during model training, image augmentation methods, including brightness transformation, random rotation, random noise addition, and mirroring, have been employed to increase the total number of images to 2595. The effect of data augmentation is illustrated in Figure 4.



a. Potted plants inoculated with root-knot nematodes of different densities

b. Growth of potted plants

Figure 2 Schematic diagram of potted plants

a. Mildly infected b. Moderately infected c. Severely infected

Figure 3 Images of cucumber roots with different degrees of disease severity

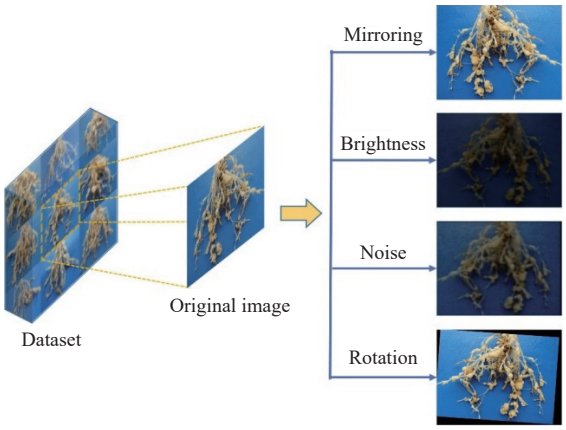


Figure 4 Effect of data augmentation

In this research, the Labelimg image annotation tool is utilized for the manual annotation of root knots, and the corresponding annotation files are saved in XML format. Labelimg is an image annotation software that requires a Python environment to run. In this research, Python version 3.9 is used. It is widely recognized that the quality of the dataset significantly influences the training efficacy of the model[26]. To minimize the errors caused by annotation, the dataset is annotated under the guidance of domain experts. Furthermore, the dataset is randomly divided into training, testing, and validation sets at a ratio of 7:2:1 to prevent sample imbalance, which can adversely affect the detection accuracy of the model.

2.3 Proposed Root-YOLOv7

The object detection model can be divided into two primary types: one-stage detectors and two-stage detectors[27, 28]. Two-stage detectors are usually slower performance compared to their one-stage counterparts. Consequently, this paper adopts the one-stage detector YOLOv7 model as the foundational framework. The YOLOv7 model incorporates composite model scaling, which facilitates the initial attributes of the model to be retained, obtaining the optimal structure[29]. In addition, the mosaic data augmentation in the YOLOv7 model preprocessing strategy is suitable for recognizing small objects[30], which meets the needs of detecting small root knots. Despite the robust capabilities of the YOLOv7 model, its performance in the detection of root knots remains suboptimal. To enhance the detection of root knots, this paper proposes the Root-YOLOv7. The structure diagram of Root-YOLOv7 is illustrated in Figure 5.

In Figure 5, the red dashed box marked with a red star is the innovative component. The backbone network is responsible for extracting features and obtains three different sizes of features of the object information. Subsequently, the head layer integrates

features from the backbone layer to construct bounding boxes and generate feature maps at these three different sizes. Finally, the RepConv layer employs a reparameterized structure to generate prediction results of three different sizes.

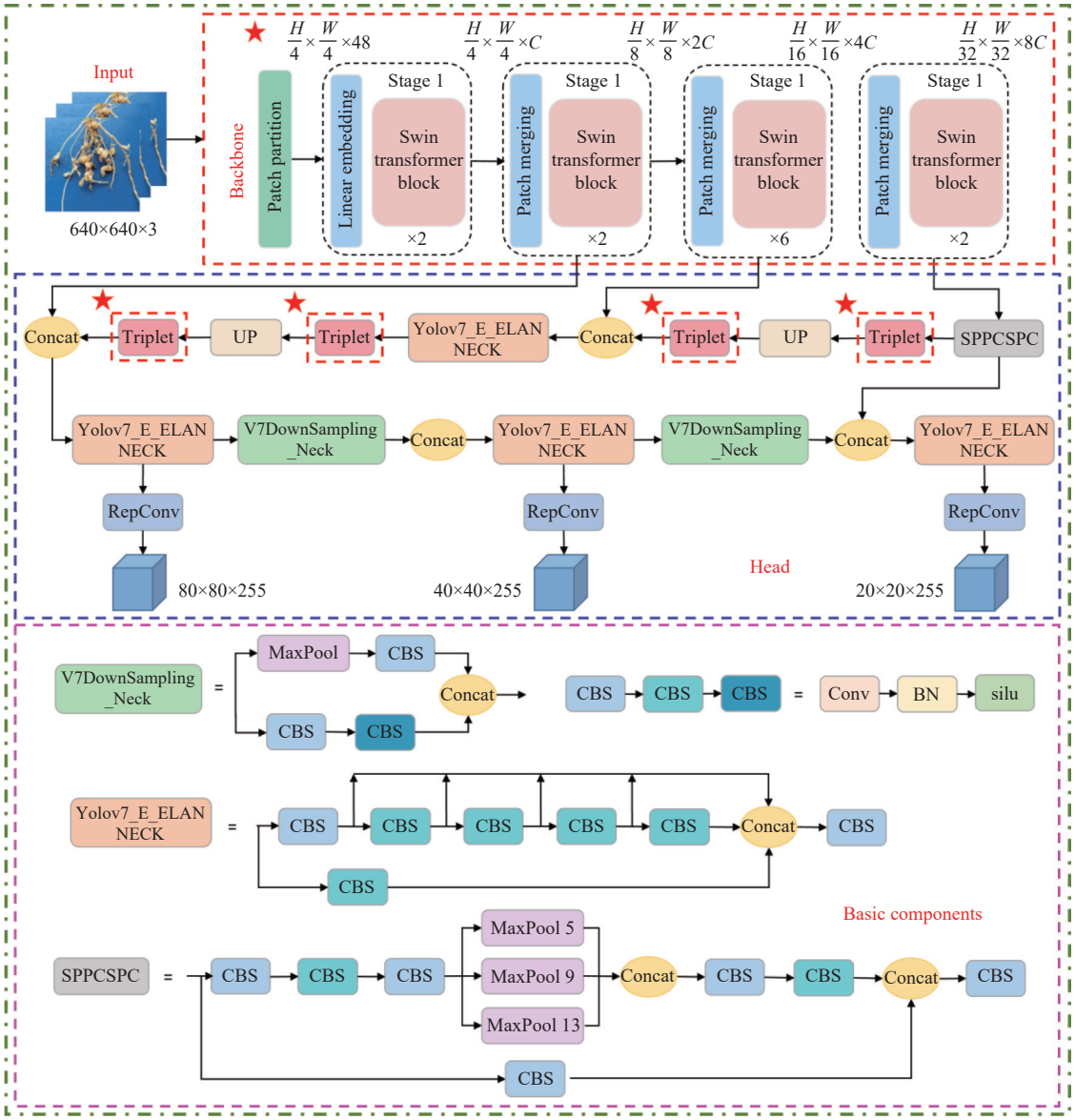


Figure 5 Network structure diagram of Root-YOLOv7

2.3.1 Reconstituted backbone module

In the YOLO series of networks, the backbone network is critical for feature extraction[31]. To extract rich features of root knots, this paper employs the Swin Transformer to reconstruct the backbone network of YOLOv7, which can effectively utilize large-scale parallel computing and provide better performance. The Swin Transformer network is characterized by a hierarchical design, and its shifted window scheme restricts the self-attention computation to non-overlapping local windows while permitting connections across different windows[32].

Swin Transformer contains four stages, each of which reduces the resolution of the input feature map while progressively expanding the receptive field. The image size remains constant in the Swin Transformer Block, and only the feature weights are optimized throughout the entire stage without altering the dimensions, as illustrated in Figure 6. The Patch Merging module effectively reduces dimensions of the feature maps by half while simultaneously doubling the number of channels, thereby facilitating the establishment of multi-scale information.

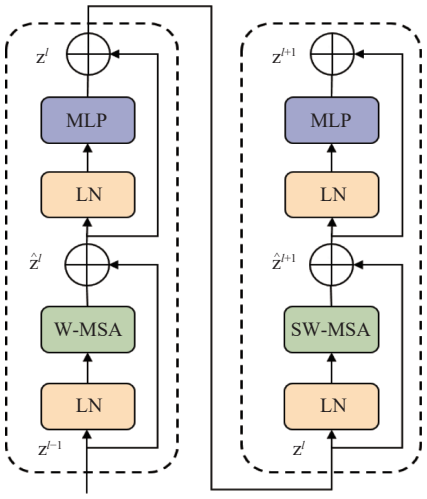


Figure 6 Schematic diagram of the Swin Transformer blocks

2.3.2 Improved loss function

The bounding box loss function plays a crucial role in

determining the outcomes of object detection[33]. WIoU is used in this research, which enables the model to measure the degree of overlap of object boxes effectively and prioritize anchor boxes of ordinary quality. The WIoU loss function provides a gradient gain allocation strategy that evaluates the quality of anchor boxes by calculating their outliers and subsequently assigns different gradient gains based on the evaluation[34,35].

There are low-quality instances in training data, and geometric attributes will amplify the impact of these instances, which reduces the generalization ability of the model. The implementation of an effective loss function is essential for mitigating the adverse impact of such instances. In this context, a distance attention mechanism has been constructed, leading to the formulation of WIoU v1 through the application of a dual-layer attention mechanism:

$$L_{\mathrm{WIoUv1}} = R_{\mathrm{WIoU}} L_{\mathrm{IoU}} \tag{1}$$

where, $R_{\mathrm{WIoU}} \in [1, \mathrm{e})$, $L_{\mathrm{IoU}} \in [0, 1]$.

The gradient gain $L_{\mathrm{IoU}}^{\lambda *}$ is the monotonic focusing coefficient for L_{WIoUv1}. The $L_{\mathrm{IoU}}^{\lambda *}$ decreases as L_{IoU} decreases, affecting the convergence speed of the model in the late stage of training. Therefore, the L_{IoU} is introduced to obtain WIoU v2:

$$L_{\mathrm{WIoUv2}} = \left(\frac{L_{\mathrm{IoU}}^{*}}{\overline{L_{\mathrm{IoU}}}} \right)^{\lambda} \times L_{\mathrm{WIoUv1}} \tag{2}$$

where, $\overline{L_{\mathrm{IoU}}}$ is the exponential moving average, which keeps $\left(\frac{L_{\mathrm{IoU}}^{*}}{\overline{L_{\mathrm{IoU}}}} \right)^{\lambda}$ at a high level.

The β is defined to describe the quality of the anchor box, as follows:

$$\beta = \frac{L_{\mathrm{IoU}}^{*}}{\overline{L_{\mathrm{IoU}}}} \in [0, +\infty) \tag{3}$$

A non-monotonic focusing coefficient is constructed and combined with WIoU v1 to obtain WIoU v3:

$$L_{\mathrm{WIoUv3}} = \gamma L_{\mathrm{WIoUv1}}, \ \gamma = \frac{\beta}{\mu \alpha^{\beta - \mu}} \tag{4}$$

where, μ makes γ=1 when β=μ. The anchor box will enjoy the highest gradient gain when β is a constant value.

This paper undertakes a series of experiments aimed at identifying the optimal variant of the WIoU loss function. Comprehensive details regarding the experiments can be found in Section 4.1.

2.3.3 Enhanced head module

The main function of the attention mechanism is to highlight significant features and suppress general features. To enhance the correlation between the key root knot features captured by the model and the representation of feature information, this paper introduces the Triplet Attention mechanism in the head layer. This mechanism emphasizes the significance of multidimensional interactions without reducing dimensionality and effectively eliminates indirect correspondences between channels and weights.

The Triplet Attention employs three dedicated branches to capture the dependencies among the dimensions of the input tensor of (C, H), (C, W), and (H, W) to achieve cross-dimensional interaction. Thus, the associations among different parts are captured more accurately, and the feature extraction ability of the model to the root knot is improved. Ultimately, a refined attention tensor y of an input tensor is obtained, and the process is delineated as follows:

$$y = \frac{1}{3}\left(\overline{\hat{x}_1 \sigma(\varepsilon_1(\hat{x}_1^*))} + \overline{\hat{x}_2 \sigma(\varepsilon_2(\hat{x}_2^*))} + x\sigma(\varepsilon_3(\hat{x}_3^*))\right) \tag{5}$$

where, σ represents the sigmoid activation function; ε_1, ε_2, and ε_3 represent the standard two-dimensional convolutional layers. Simplify Eq. (5) to obtain y:

$$y = \frac{1}{3}\left(\overline{\hat{x}_1 w_1} + \overline{\hat{x}_2 w_2} + x w_3\right) = \frac{1}{3}\left(\overline{y_1} + \overline{y_2} + y_3\right) \tag{6}$$

where, w_1, w_2, and w_3 are the three cross-dimensional attention weights computed in Triplet Attention. The $\overline{y_1}$ and $\overline{y_2}$ denote the 90° clockwise rotation to preserve the original input shape.

2.4 Network performance evaluation indicators

This paper introduces Precision (P), Recall (R), Average Precision (AP), and $F1$ as performance evaluation indicators for models, and their definitions are as follows:

$$P = \frac{TP}{TP + FP} \tag{7}$$

$$R = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = \frac{2 \times P \times R}{P + R} \tag{9}$$

$$AP = \int_0^1 P(R) \mathrm{d}R \tag{10}$$

where, TP, FP, and FN represent the number of samples correctly predicted as positive, incorrectly predicted as positive, and incorrectly predicted as negative, respectively.

In addition, AP@0.5 is the average accuracy value of 0.5 IoU thresholds. AP@0.5:0.95 is the average of 10 IoU thresholds with IoU∈[0.5: 0.05: 0.95].

2.5 Experimental configuration

The image size is standardized to 640 x 640 pixels for the input during the model training process. The training procedure encompasses 300 iterations with a batch size of 16. The momentum for the learning rate is established at 0.937, and the optimizer utilized is SGD. Detailed hardware and software configurations for the experimentation are presented in Table 2.

Table 2 Hardware and software configuration

Software/Hardware	Configuration
Operating system	Windows10 (64-bit)
CPU	Intel(R) Xeon(R) CPU E5-4627 v4
RAM	64 GB
Programming language	Python
Framework	Pytorch 2.10
Label software	Labelimg

3 Results

3.1 Experimental results and analysis

The training results of Root-YOLOv7 and YOLOv7 are illustrated in Figure 7. It can be seen from Figure 7 that the P, R, AP@0.5, and AP@0.5:0.95 values of Root-YOLOv7 are significantly higher than those of YOLOv7, and the changing trend of values of Root-YOLOv7 is relatively stable. In addition, the curve of the original YOLOv7 model has a large fluctuation in the process of the first 100 iterations, while the curve of the Root-YOLOv7 model is relatively flat as a whole. This indicates that the research in this paper improves the robustness of the model. The partial detection results of YOLOv7 and Root-YOLOv7 on root knots are illustrated in Figure 8.

From Figure 8b and 8d, it can be seen that the Root-YOLOv7 model demonstrates superior capabilities in detecting small object root knots and densely clustered root knots that are overlooked in

the YOLOv7 model. The missed detection in the latter may be attributed to the reduced pixel count in images of small objects, which consequently leads to a diminished extraction of features during the convolution process. In Figure 8c, the presence of two overlapping root knots can obscure the significance of their features. However, the Root-YOLOv7 model is still able to accurately differentiate between them. The detection performance and anti-interference ability of the Root-YOLOv7 model have shown marked improvement. In conclusion, Root-YOLOv7 has better performance in detecting root-knot nematode disease.

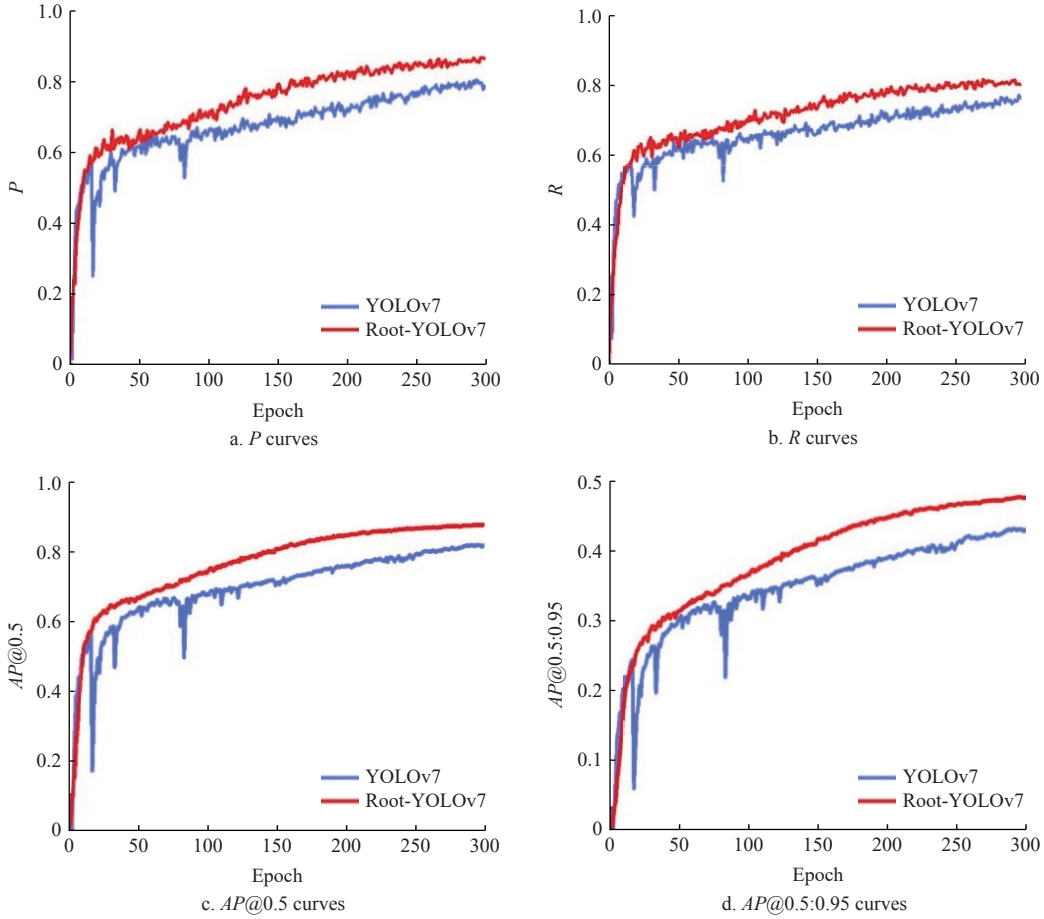


Figure 7 Training results of Root-YOLOv7

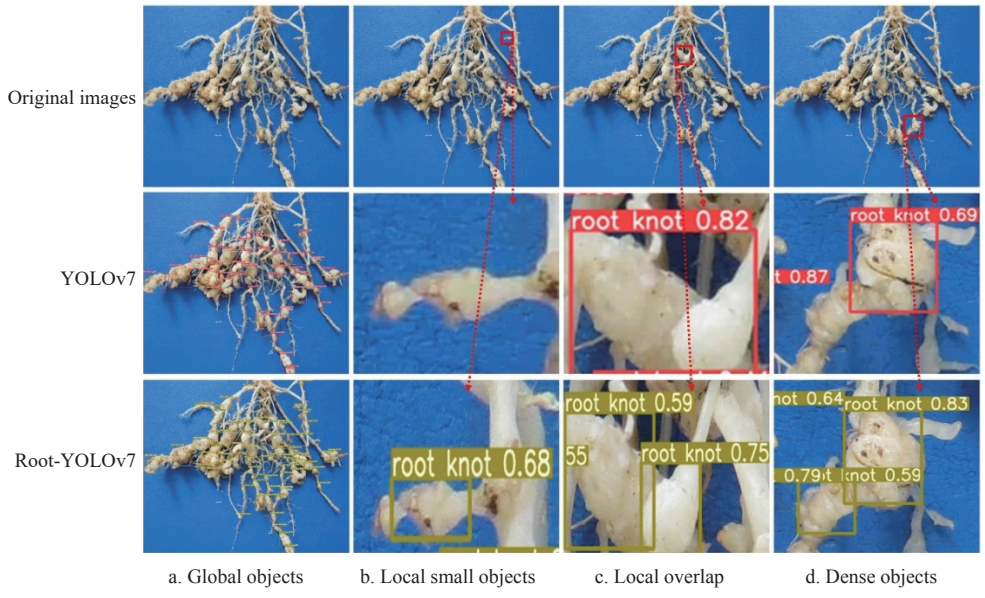


Figure 8 Comparison of detection results of YOLOv7 and Root-YOLOv7

3.2 Ablation experiments

To assess the effectiveness of the enhanced method, eight groups of ablation experiments utilizing YOLOv7 were designed, incorporating various combinations of the improved modules. The experimental results are presented in Table 3, in which the symbol "+" denotes the inclusion of a module, while "−" signifies its exclusion. The data presented in bold indicates the optimal performance achieved during the experiments.

Table 3 Ablation experiments

Models	Swin Transformer	WIoU v2	Triplet Attention	*P*/%	*R*/%	*AP*@0.5/%	*AP*@0.5:0.95/%
YOLOv7	−	−	−	78.78	76.49	81.57	42.94
Model1	+	−	−	85.28	**80.94**	87.16	**48.31**
Model2	−	+	−	80.98	79.44	84.56	45.14
Model3	−	−	+	85.06	78.29	85.14	47.16
Model4	+	+	−	81.41	76.70	83.62	43.99
Model5	+	−	+	77.54	75.37	81.20	42.58
Model6	−	+	+	76.21	71.99	78.26	40.48
Ours	+	+	+	**86.62**	80.31	**87.40**	47.50

From Table 3, it can be seen that there is a significant enhancement in the overall performance of the model when the Swin Transformer, WIoU v2 loss function, and Triplet Attention module are applied independently. The Swin Transformer module has an extremely positive impact on the model, which benefits from its hierarchical design and shifted window scheme. The model captures the cross-dimension interaction of the root knot feature between the spatial and channel dimensions after introducing the Triplet Attention module, which makes the model show satisfactory performance. The variations in *P* and *F*1 values of the eight models are illustrated in Figure 9.

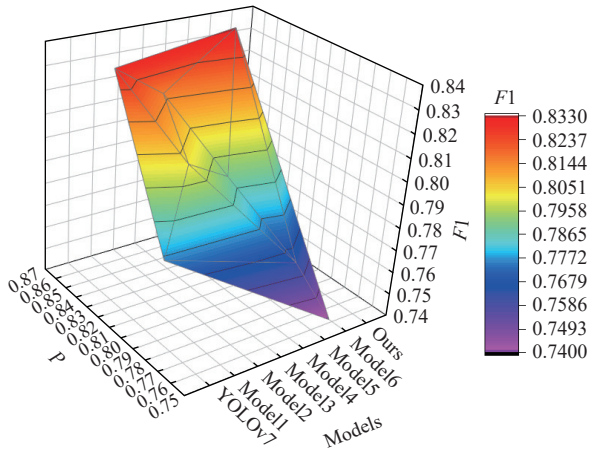


Figure 9 3D heat map of *P* and *F*1 value of eight models

It is evident from Figure 9 that the *P* and *F*1 value of the Root-YOLOv7 model attain their maximum values following the fusion of three improved methods. The ablation experiments indicate that the proposed approach significantly optimizes the detection ability of the YOLOv7 model.

3.3 Comparison with other object detection models

To further validate the detection effectiveness of the Root-YOLOv7 model, this paper conducts a comparative analysis with several prominent object detection models, including Faster R-CNN[36], RT-DETR[37], YOLOv5[38], YOLOv6[39], YOLOv7[40], and YOLOv8[41]. The comparative experiments are all based on the same dataset, and the comparison results of the seven models are presented in Table 4.

Table 4 Comparison of the performances of seven models

Models	*P*/%	*R*/%	*AP*@0.5/%	*AP*@0.5:0.95/%	GFLOPs/G	Parameters/M
Faster R-CNN	27.89	24.63	14.73	4.70	251.40	41.30
RT-DETR	80.30	75.22	81.80	43.07	110.00	31.99
YOLOv5	79.85	74.74	82.12	44.55	15.80	7.01
YOLOv6	73.85	71.24	77.72	45.09	11.80	4.23
YOLOv7	78.78	76.49	81.57	42.94	120.70	37.35
YOLOv8	75.67	72.79	79.95	47.23	**8.10**	**3.01**
Root-YOLOv7	**86.62**	**80.31**	**87.40**	**47.50**	183.80	50.18

The data presented in Table 4 indicates that Root-YOLOv7 exhibits superior accuracy compared to the other six models. Notably, Root-YOLOv7 achieves the highest *P*, *R*, and *AP* values among the YOLO series models. In particular, when compared to the recent YOLOv8, Root-YOLOv7 also demonstrates remarkable performance improvements, with *P*, *R*, and *AP*@0.5 increased by 10.95%, 7.52%, and 7.45%, respectively. Furthermore, in comparison to YOLOv7, Root-YOLOv7 achieves higher accuracy with a slight increase in the number of model parameters. A comprehensive analysis leads to the conclusion that the Root-YOLOv7 model possesses significant advantages in the specific context of root knots, thereby enhancing its applicability in the detection domain.

To more intuitively evaluate the effect of different models in the detection of root knots, the confusion matrices of different models are plotted in this research, as shown in Figure 10.

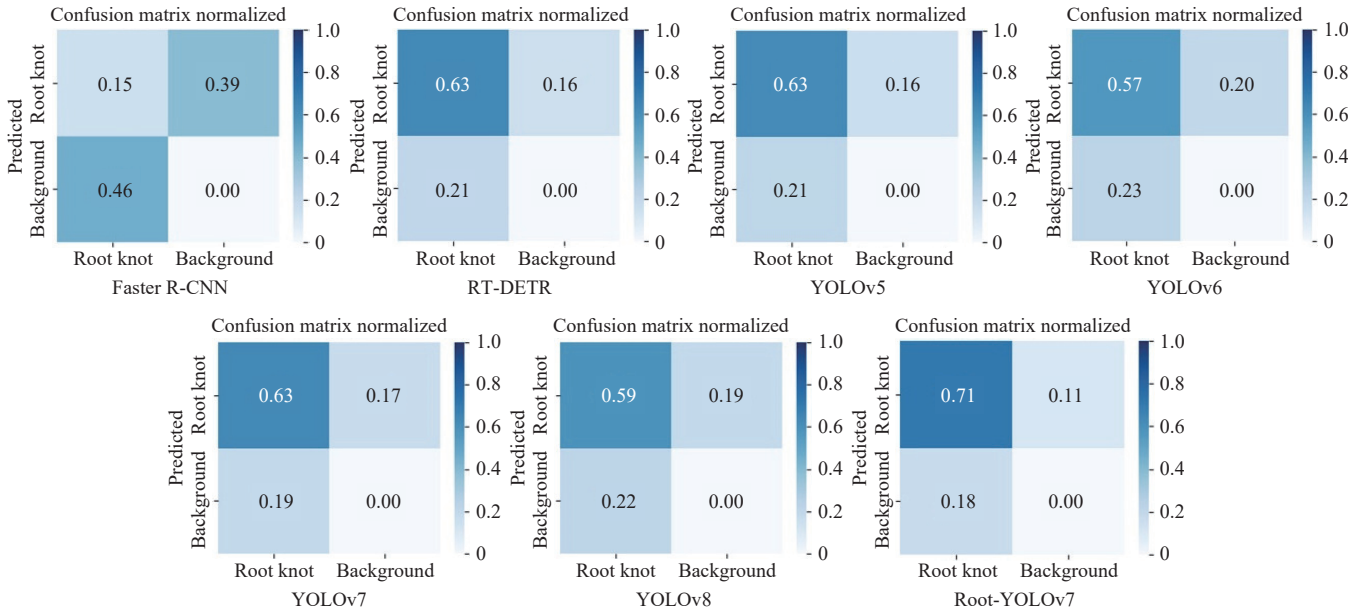


Figure 10 Confusion matrices for different models

From Figure 10, it can be seen that the Root-YOLOv7 model performs the best. Compared with other models, Root-YOLOv7 achieves the highest *TP*, which indicates its excellent ability to correctly detect root knot objects. Meanwhile, its *FP* and *FN* are significantly lower than the other models. These results fully demonstrate that Root-YOLOv7 can not only achieve more comprehensive detection but also effectively reduce the occurrence of false and missed detections, providing a more reliable solution for root knot object detection tasks.

3.4 Grading of root-knot nematode disease

The reliable and accurate assessment of plant disease is essential for enabling farmers to implement early prevention and management, predict yield loss, and evaluate plant resistance. On the basis of achieving the optimal root knot detection effect, it is necessary to quantify the number of root knots and subsequently assess the degree of root-knot nematode disease through the root knot counting method[42]. Figure 11 illustrates the disease grading results derived from randomly selected root images.

Figure 11 Grading results of root-knot nematode disease

Root knot index (*RKI*) is a comprehensive index that incorporates both the incidence rate and the disease severity into account. It can be derived from the root-knot nematode disease grade obtained above, and then the corresponding prevention and control effect (*PCE*) of the plant can be calculated, as illustrated in the following equation:

$$RKI = \frac{\sum(N \times D)}{T \times H} \times 100\% \quad (11)$$

$$PCE = \frac{C - E}{C} \times 100\% \quad (12)$$

where, N represents the number of plants at a certain disease level, D represents the disease level, T represents the total number of plants, H represents the highest level, C represents the *RKI* in the control group, and E represents the *RKI* in the experimental group.

In addition, the *RKI* can also be used to evaluate the resistance of plants to root-knot nematode disease. Therefore, the detection and grading of root-knot nematode disease can not only help plant pathologists achieve efficient control but also provide theoretical guidance for the cultivation of plant varieties with disease resistance.

4 Discussion

4.1 Impact of WIoU loss function on performance

This paper employs the WIoU v2 loss function. Within the WIoU series, different loss functions exert a distinct influence on the performance of the model. The impact of different WIoU loss functions on model performance in this research is illustrated in Figure 12.

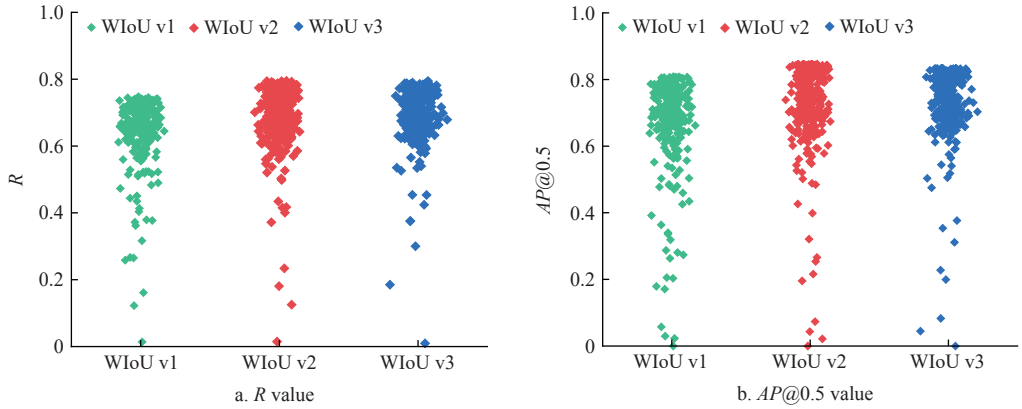


Figure 12 Effects of different WIoU loss functions on model performance

From Figure 12, it is evident that the *R* value and the *AP*@0.5 value of WIoU v2 are significantly higher than those for WIoU v1 and WIoU v3. Furthermore, the scatter points associated with the WIoU v2 loss function exhibit less fluctuation compared to the other two versions, suggesting that WIoU v2 contributes to greater stability in the model. In conclusion, among the series of WIoU loss functions examined, WIoU v2 demonstrates a more favorable impact on all aspects of the dataset in this paper. Consequently, this paper designates WIoU v2 as the selected loss function for the model.

4.2 Result analysis of the ablation experiment

In the ablation experiment, the *AP* value of Root-YOLOv7 is maximized when the IoU threshold is established at 0.5. This outcome can be attributed to the enhanced synergistic effects of the Swin Transformer, WIoU v2, and Triplet Attention modules integrated within Root-YOLOv7. Conversely, when the IoU

threshold is adjusted to the range of 0.5:0.95, the *AP* of Root-YOLOv7 is observed to be 0.81% lower than that of YOLOv7 combined with the Swin Transformer. This discrepancy may be indicative of certain inhibitory interactions among the multiple embedded modules (Swin Transformer + WIoU v2 + Triplet Attention).

Furthermore, the *P* of YOLOv7 combined with the Swin Transformer and WIoU v2 is 0.43% higher than that of YOLOv7 with WIoU v2. However, the *AP* is lower than that of YOLOv7 with WIoU v2. This discrepancy may be attributed to the fact that Swin Transformer enhances the capability to capture object feature information across various scales when it is embedded in the backbone network, thus improving the prediction accuracy of the model. Nevertheless, the *AP* does not exhibit significant change, which may be due to the influence of recall.

4.3 Influence factors of model performance

The *P* and *AP*@0.5 values of Root-YOLOv7 are 7.84% and 5.83% higher than those of YOLOv7, respectively. Despite the significant enhancement in detection performance, a small portion of root knots remain undetected. This limitation may be attributed to several factors: (1) Cucumber roots exhibit numerous fibrous roots, dense root knots, and root knot objects that are relatively small, which may impair the convolutional neural network semantic extraction ability and comprehensive feature extraction ability of YOLOv7 during detection. (2) The dataset lacks sufficient richness. A diverse dataset includes samples from various regions, varieties, or backgrounds[43]. A well-diversified dataset can provide more comprehensive information, thereby enabling the model to better capture the object features and enhance its generalization capabilities. (3) During the dataset acquisition process, it is difficult to fix the distance and angle between the camera and the roots, and the location of the root knots is also inconsistent. Such variations in shooting distance and angle may influence detection performance to some extent[44].

4.4 Limitations and future outlook

This research takes 519 original root-knot nematode disease samples for cucumbers. In fact, root-knot nematode disease is a kind of disease that is prone to occur in the growth of various crops. Therefore, it can be considered to expand the research to other crops to increase the generalization ability of the model in the future. In addition, our research focuses on detection accuracy and does not consider the issue of a lightweight model. In practical applications, it is sometimes necessary to consider a lightweight model to deploy it on mobile devices such as embedded systems. Therefore, subsequent research may focus on improving the generalization ability and lightweight level of the model in order to further enhance its universality and real-time performance.

5 Conclusions

The detection of root-knot nematode disease presents significant challenges due to the extensive presence of fibrous roots, dense root knots, and the occurrence of diminutive root knots within the roots of diseased plants. To address these challenges, this paper proposes a multi-scale adaptive object detection and grading model Root-YOLOv7 for the identification of root-knot nematode disease using cucumber plants as the research object. The *P* of Root-YOLOv7 is 7.84% higher than the original YOLOv7, which indicates the effectiveness of the improved approach in this research. This model facilitates the detection and quantification of root knots within the root systems of plants, thereby providing insights into the extent of damage caused by root-knot nematode disease. The findings of this paper hold both theoretical and practical significance for the advancement of grading devices for root-knot nematode disease and the cultivation of varieties with stress resistance.

Acknowledgements

The research work is financially supported by the project of National Natural Science Foundation of China (Grant No. 32073029); the Shandong Province Modern Agricultural Technology System (Grant No. SDAIT-22-09); the Key Research and Development Program of Shandong Province (Grant No. 2021LZGC026-03); the Science & Technology Specific Projects in Agricultural High-tech Industrial Demonstration Area of the Yellow River Delta (Grant No. 2022SZX18); and the Key R&D Program of Shandong Province (Grant No. 2022CXGC020710).

Author contributions

Yao Zhao: investigation, resources, methodology, software, original draft; Honghai Zhao: conceptualization, supervision, review & editing; Haitao Xiong: methodology, software; Fan Zhang: investigation, software; Cheng Lu: review and editing; Juan Li: conceptualization, methodology, review & editing.

[References]

[1] Lapajne J, Knapič M, Žibrat U. Comparison of selected dimensionality reduction methods for detection of root-knot nematode infestations in potato tubers using hyperspectral imaging. Sensors (Basel), 2022; 22(1): 367.

[2] Haque F, Thimmanagari M, Chiang Y W. Ultrasound assisted cyanotoxin extraction for nematode inhibition in soil. Ultrason Sonochem, 2022; 89: 106120.

[3] Atkinson H J, Lilley C J, Urwin P E. Strategies for transgenic nematode control in developed and developing world crops. Curr Opin Biotechnol, 2012; 23(2): 251–256.

[4] Park E, Kim Y-S, Faqeerzada M A, Kim M S, Baek I, Cho B-K. Hyperspectral reflectance imaging for nondestructive evaluation of root rot in Korean ginseng (*Panax ginseng Meyer*). Front Plant Sci, 2023; 14: 1109060.

[5] Abade A, Ferreira P A, Vidal F D. Plant diseases recognition on images using convolutional neural networks: A systematic review. Comput Electron Agric, 2021; 185: 106125.

[6] Hu X L, Liu Y, Zhao Z X, Liu J T, Yang X T, Sun C H, et al. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network. Comput Electron Agric, 2021; 185: 10613.

[7] Ma G J, Xu S P, Yang T, Du Z B, Zhu L M, Ding H, et al. A transfer learning-based method for personalized state of health estimation of lithium-ion batteries. IEEE Trans Neural Networks Learn Syst, 2024; 35(1): 759–769.

[8] Tian L L, Wang Z D, Liu W B, Cheng Y H, Alsaadi F E, Liu X H. An improved generative adversarial network with modified loss function for crack detection in electromagnetic nondestructive testing. Complex Intell Syst, 2022; 8(1): 467–476.

[9] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. Med Image Anal, 2019; 58: 101552.

[10] Deng L M, Li J, Han Z Z. Online defect detection and automatic grading of carrots using computer vision combined with deep learning methods. LWT-Food Science and Technology, 2021; 149: 111832.

[11] Xiong H T, Xiao Y, Zhao H P, Xuan K, Zhao Y, Li J. AD-YOLOv5: An object detection approach for key parts of sika deer based on deep learning. Comput Electron Agric, 2024; 217: 108610.

[12] Bohnenkamp D, Behmann J, Paulus S, Steiner U, Mahlein A K. A hyperspectral library of foliar diseases of wheat. Phytopathology, 2021; 111(9): 1583–1593.

[13] Gomez-Caro S, Mendoza-Vargas L A, Ramirez-Gil J G, Burbano-David D, Soto-Suarez M, Melgarejo L M. Close-Range Thermography and reflectance spectroscopy support in vitro and in vivo characterization of

colletotrichum spp. isolates from mango fruits. Plant Disease, 2022; 106(9): 2355–2369.
[14] Mohammadpoor M, Nooghabi M G, Ahmedi Z. An intelligent technique for grape fanleaf virus detection. Int J Interact Multimedia Artif Intell, 2020; 6(1): 62–67.
[15] Almadhor A, Rauf H T, Lali M I U, Damasevicius R, Alouffi B, Alharbi A. AI-driven framework for recognition of guava plant diseases through machine learning from DSLR camera sensor based high resolution imagery. Sensors. 2021; 21(11). doi: 10.3390/s21113830.
[16] Ostovar A, Talbot B, Puliti S, Astrup R, Ringdahl O. Detection and classification of Root and Butt-Rot (RBR) in Stumps of Norway Spruce using RGB images and machine learning. Sensors (Basel), 2019; 19(7). doi: 10.3390/s19071579.
[17] Huang W X, Huo Y, Yang S C, Liu M J, Li H, Zhang M. Detection of Laodelphax striatellus (small brown planthopper) based on improved YOLOv5. Comput Electron Agric, 2023; 206: 107657.
[18] Li K, Zhang L, Li B, Li S, Ma J. Attention-optimized DeepLab V3 + for automatic estimation of cucumber disease severity. Plant Methods, 2022; 18(1): 109.
[19] Wang F Y, Rao Y, Luo Q, Jin X, Jiang Z H, Zhang W, et al. Practical cucumber leaf disease recognition using improved Swin Transformer and small sample size. Comput Electron Agric, 2022; 199: 107163.
[20] Wang T W, Zhao L G, Li B H, Liu X W, Xu W K, Li J. Recognition and counting of typical apple pests based on deep learning. Ecol Inf, 2022; 68. doi: 10.1016/j.ecoinf.2022.101556.
[21] Liu J, Wang X W, Miao W Q, Liu G X. Tomato Pest Recognition Algorithm Based on Improved YOLOv4. Front Plant Sci, 2022; 13. doi: 10.3389/fpls.2022.814681.
[22] Qi J T, Liu X N, Liu K, Xu F R, Guo H, Tian X L, et al. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease. Comput Electron Agric, 2022; 194: 106780.
[23] Qing X, Wang Y H, Lu X Q, Li H B, Wang X, Li H M, et al. NemaRec: A deep learning-based web application for nematode image identification and ecological indices calculation. Eur J Soil Biol, 2022; 110. doi: 10.1016/j.ejsobi.2022.103408.
[24] Sharma V, Tripathi A K, Mittal H. DLMC-Net: Deeper lightweight multi-class classification model for plant leaf disease detection. Ecol Inf, 2023; 75. doi: 10.1016/j.ecoinf.2023.102025.
[25] Wang C S, Sun S D, Zhao C J, Mao Z C, Wu H R, Teng G F. A detection model for cucumber root-knot nematodes based on modified YOLOv5-CMS. Agronomy-Basel. 2022; 12(10). doi: 10.3390/agronomy12102555.
[26] Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning. Journal of Big Data, 2019; 6(1). doi: 10.1186/s40537-019-0197-0.
[27] Gai W, Liu Y, Zhang J, Jing G. An improved Tiny YOLOv3 for real-time object detection. Systems Science & Control Engineering, 2021; 9(1): 314–321.
[28] Zeng N, Wu P, Wang Z, Li H, Liu W, Liu X. A Small-Sized Object Detection Oriented Multi-Scale Feature Fusion Approach With Application to Defect Detection. IEEE Trans Instrum Meas, 2022; 71: 1–14.
[29] Wang C-Y, Bochkovskiy A, Liao H Y M . YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022; pp.7464–7475.
[30] Soeb M J A, Jubayer M F, Tarin T A, Al Mamun M R, Ruhad F M, Parven A, et al. Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). Sci Rep. 2023; 13(1). doi: 10.1038/s41598-023-33270-4.
[31] Bochkovskiy A , Wang C-Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection. arXiv - CS - Computer Vision and Pattern Recognition, 2020. doi: 10.48550/arXiv.2004.10934.
[32] Chen Y, Shu H C, Xu W J, Yang Z Y, Hong Z H, Dong M S. Transformer text recognition with deep learning algorithm. Comput Commun. 2021; 178: 153–60. doi: 10.1016/j.comcom.2021.04.031.
[33] Chen X, Yuan M J H, Yang Q, Yao H Y, Wang H Y. Underwater-YCC: Underwater target detection optimization algorithm based on YOLOv7. J Mar Sci Eng. 2023; 11(5). doi: 10.3390/jmse11050995.
[34] Tong Z J, Chen Y H, Xu Z W, Yu R. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism. arXiv - CS - Computer Vision and Pattern Recognition, 2023. doi: 10.48550/arXiv-2301.10051.
[35] Huang H Q, Lan G W, Wei J, Zhong Z, Xu Z R, Li D B, et al. TLI-YOLOv5: A lightweight object detection framework for transmission line inspection by unmanned aerial vehicle. Electronics, 2023; 12(15). doi: 10.3390/electronics12153340.
[36] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell, 2017; 39(6): 1137–1149.
[37] Lv W Y, Xu S L, Zhao Y A, Wang G Z, Wei J M, Cui C, Du Y N, et al. DETRs beat YOLOs on real-time object detection. arXiv - CS - Computer Vision and Pattern Recognition, 2023. doi: 10.48550/arxiv-2304.08069.
[38] Roy A M, Bhaduri J. DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism. Adv Eng Inf, 2023; 56. doi: 10.1016/j.aei.2023.102007.
[39] Saydirasulovich S N, Abdusalomov A, Jamil M K, Nasimov R, Kozhamzharova D, Cho Y I. A YOLOv6-based improved fire detection approach for smart city environments. Sensors, 2023; 23(6): 3161.
[40] Wang S Y, Wu D S, Zheng X Y. TBC-YOLOv7: a refined YOLOv7-based algorithm for tea bud grading detection. Front Plant Sci, 2023; 14. doi: 10.3389/fpls.2023.1223410.
[41] Jiang H Y, Hu F, Fu X Q, Chen C R, Wang C, Tian L X, et al. YOLOv8-Peas: a lightweight drought tolerance method for peas based on seed germination vigor. Front Plant Sci, 2023; 14. doi: 10.3389/fpls.2023.1257947.
[42] Yannong Xiao M W, Yanping Fu, Fantao Zeng. Comparison of grading methods for vegetable root knot nematode disease. Journal of Huazhong Agricultural University, 2000; 19(4): 336–338.
[43] Althnian A, AlSaeed D, Al-Baity H, Samha A, Bin Dris A, Alzakari N, et al. Impact of dataset size on classification performance: an empirical evaluation in the medical domain. App.Sci-Basel, 2021; 11(2). doi: 10.3390/app11020796.
[44] Xu W K, Zhao L G, Li J, Shang S Q, Ding X P, Wang T W. Detection and classification of tea buds based on deep learning. Comput Electron Agric, 2022; 192: 106547.