DETECTION OF APPLE LEAF DISEASES TARGET BASED ON IMPROVED YOLOv7 / 基于改进 YOLOv7 的苹果叶病害目标检测

Lingqing FENG, Yujing LIU*), Hua YANG, Zongwei JIA, Jiaxiong GUAN, Huiru Zhu,Yiming HOU
College of Information Science and Engineering, Shanxi Agricultural University, Taigu/China
Tel: +86-13834837537; E-mail: sxauxky@sxau.edu.cn
DOI: https://doi.org/10.35633/inmateh-72-26

Keywords: *apple leaf, disease detection, YOLOv7, SimAM attention mechanism, SIoU*

ABSTRACT

Apple leaf diseases significantly threaten the yield and quality of apples. In order to detect apple leaf diseases in a timely and accurate manner, this study proposed a detection method for apple leaf diseases based on an improved YOLOv7 model. The method integrated a Similarity-based Attention Mechanism (SimAM) into the traditional YOLOv7 model. Additionally, the regression loss function is modified from Complete Intersection over Union (CIoU) to Structured Intersection over Union (SIoU). Experimental results demonstrates that the improved model exhibits an overall recognition precision of 92%, a recall rate of 99%, and a mean average precision (mAP) of 96.1%. These metrics show a respective improvement of 14.4%, 38.85%, and 18.69% compared to the pre-improved YOLOv7. When compared with seven other target detection models in comparative experiments, the improved YOLOv7 model achieves higher accuracy, lower rates of missed and false detections in disease target detection. The model excels in detecting disease categories in complex environments and identifying small targets at early disease stages. It can provide technical support for effective detection of apple leaf diseases.

摘要

苹果叶部病害严重危害苹果的产量和品质，为了实现及时且准确地对苹果叶病害进行检测，提出了一种基于改进 YOLOv7 模型的苹果叶部病害检测方法，在传统的 YOLOv7 模型基础上融合了无参数注意力机制 SimAM；并将回归损失函数由 CIoU(Complete Intersection over Union）改进为 SIoU(Structured Intersection over Union)。试验结果表明，改进后的模型整体识别精准度、召回率、平均精度均值 mAP（Mean average precision）分别为 92%，99%，96.1%；与改进前 YOLOv7 相比，分别提升 14.4%，38.85%，18.69%，与对比实验中的其他 7 种目标检测模型相比，改进模型检测精度更高、漏检和错检率低、在复杂环境以及病害初期小目标检测中表现优良。可以为苹果叶病害有效检测提供技术支持。

INTRODUCTION

Apples are among the world's most significant economic crops due to their rich nutritional content, high yield, and survival rate, and are widely cultivated in many countries globally *(Zhong & Zhao, 2020).* The prevalent diseases in apple cultivation include Alternaria Blotch, Brown Spot, Grey spot, Rust and mosaic diseases, which pose a severe threat to young fruits, new shoots, petioles, and young leaves *(Chao et al., 2020).* Traditionally, the detection of apple leaf diseases has been performed by experts through visual inspection of the leaves, relying on experience *(Khan et al., 2022).* This method is not only time-consuming and laborious but also prone to misjudgement due to subjective human factors *(Bansal et al., 2021).* However, the spread and development of apple diseases are very rapid; if not identified and addressed promptly, diseases can quickly escalate, greatly impacting the quality and yield of apples and directly affecting their economic benefits *(Bi et al., 2022; Wang & Zhao, 2022).* Therefore, timely and accurate identification of the types of apple leaf diseases is of great significance and practical value for the prevention and control of these diseases.

With the advancement of computer science and technology, computer vision technology has been widely applied in the agriculture sector, and many models utilizing machine learning methods for plant disease identification have been developed for the detection and classification of plant diseases *(Pathan et al., 2020)*. Singh et al. proposed a new segmentation algorithm for diseased parts of apple leaf images, and then extracted colour and texture features from the segmented apple leaves, achieving the best accuracy of 96.4% using a K nearest neighbour classifier *(Singh et al., 2022)*. Sahu et al. proposes a novel hybrid random forest Multiclass SVM (HRF-MCSVM) design for plant foliar disease detection *(Sahu & Pandey, 2023)*.

Although traditional machine learning can reduce identification time, there are obvious limitations in handling large-scale, high-dimensional, nonlinear, and sequential data.

In recent years, with the development of deep learning, many researchers have begun to apply deep learning to plant disease identification. G. Priyadharshini conducted comparative research on the use of CNN, R-CNN, Fast R-CNN, and Faster R-CNN for the detection and classification of tomato leaf diseases *(Priyadharshini et al., 2023)*. He et al. proposed the MFaster R-CNN algorithm based on the Faster R-CNN algorithm, which achieved intelligent diagnosis of complex background and similar disease spot features in corn diseases in field environments, with an overall average accuracy rate of 97.23% *(He et al., 2023)*. Although these two-stage target detection algorithms generally have higher detection accuracy but suffer from long parameter extraction times and slow speeds, they are suitable for large-scale and complex scenarios. Sajitha.et al. presented a leaf disease detection and classification system utilizing YOLO v7, the system achieved high accuracy (96%) in identifying various leaf diseases *(Sajitha et al., 2023)*. Mathew et al. applied the YOLOv5 algorithm to detect bacterial spot disease in pepper plants based on symptoms on the leaves *(Mathew & Mahesh, 2022)*. Li et al. proposed an apple leaf disease detection method based on the BTC-YOLOv5s model for accurate localization and recognition of multiscale and different-shaped apple leaf diseases against complex backgrounds, offering higher detection precision and accuracy *(Li et al., 2023)*. These One-stage target detection methods have relatively faster detection speeds and accuracy.

The above research shows that deep learning has achieved certain effects in crop and plant disease detection methods. However, there is a lack of studies on small targets with less apparent features in the early stages of diseases. Additionally, there are challenges with high rates of missed and false detections when detecting targets against complex backgrounds and identifying multiple disease categories. To address the current challenges faced by apple leaf disease detection methods, this paper presents an efficient and accurate apple leaf disease target detection algorithm based on YOLOv7. The SimAM attention mechanism was introduced in YOLOv7 and SIoU Loss was used to replace the original CIoU loss to improve the detection speed and accuracy of small targets in the early stages of apple leaf disease in various environments.

MATERIALS AND METHODS

Data Acquisition and Preprocessing

Some data used in the experiments in this paper were sourced from AI Studio. In order to collect images of apple diseases from different periods and environments, many early disease images captured using smartphones, and additional images were obtained using web crawlers and Google searches. All images were collected directly from fields under natural light. To ensure the diversity of the dataset, samples of each disease category were selected based on background, severity, and the growth stage of the leaves. The resulting dataset includes images of the four most common apple leaf diseases: Alternaria Blotch, Brown Spot, Grey Spot, and Rust, totalling 1643 pictures as shown in Figure 1.

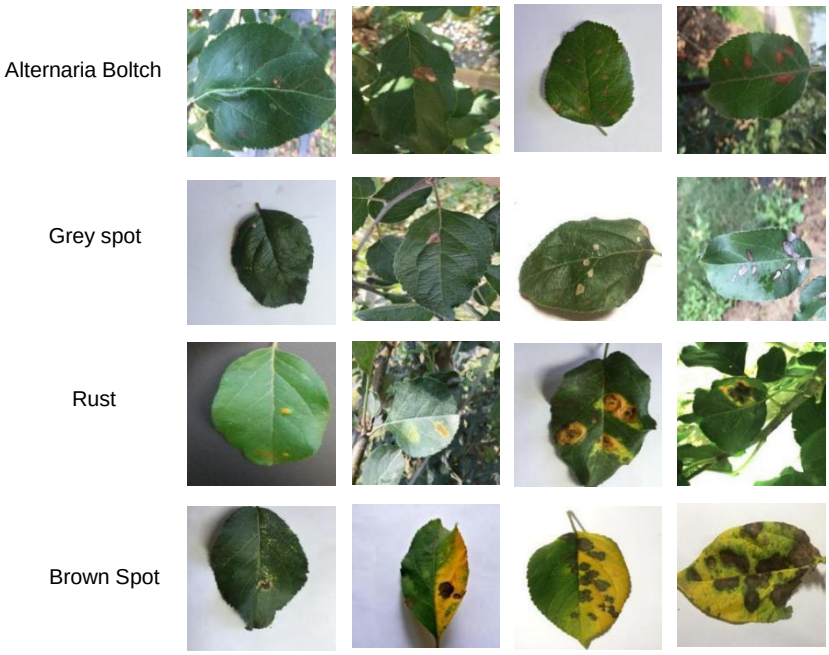


Fig. 1 - Partial data collection

The LabelImg tool was used to mark diseases in the dataset. Each disease in the images is completely enclosed by a rectangular box, with efforts made to minimize the inclusion of unnecessary background within the box. The position of the box is determined by the coordinates of the top-left and bottom-right vertices. After all diseases in the images were annotated, corresponding XML files were generated, containing information such as the size of the images, the names of the labels, and the positions of the labels. The annotation work was carried out by four researchers with relevant expertise.

To ensure the accuracy of model training, improve the robustness of the model, and enhance its ability to recognize apple leaf diseases from different periods and environments, and to avoid model overfitting due to insufficient data, the images were pre-processed to simulate variations in lighting, exposure and angles. The dataset was augmented with operations such as increasing and decreasing brightness, saturation, and contrast, as well as rotation and flipping. A total of 16,430 enhanced images was divided into training, validation, and test sets in an 8:1:1 ratio. with the statistical results of the dataset samples as shown in Table 1.

Table 1

Apple leaf disease dataset

Class name	Original image	Augmented			
		Training set	Validation set	Test set	Total
Alternaria Boltch	400	3200	400	400	4000
Grey spot	370	2960	370	370	3700
Rust	438	3504	438	438	4380
Brown Spot	435	3480	435	435	4350
Total	1643	13144	1643	1643	16430

YOLOv7 Model

YOLOv7*(Wang et al., 2022)* is the seventh generation of the YOLO series, and comparative experiments conducted on public datasets demonstrate that the performance of this version has improved significantly over previous iterations, with detection speed and accuracy surpassing other object detection algorithms. As a single-stage object detection algorithm, YOLOv7 is widely applied in real-time detection due to its superior performance. In this paper, YOLOv7 was utilized as the benchmark model, and its network structure is depicted in Figure 2.

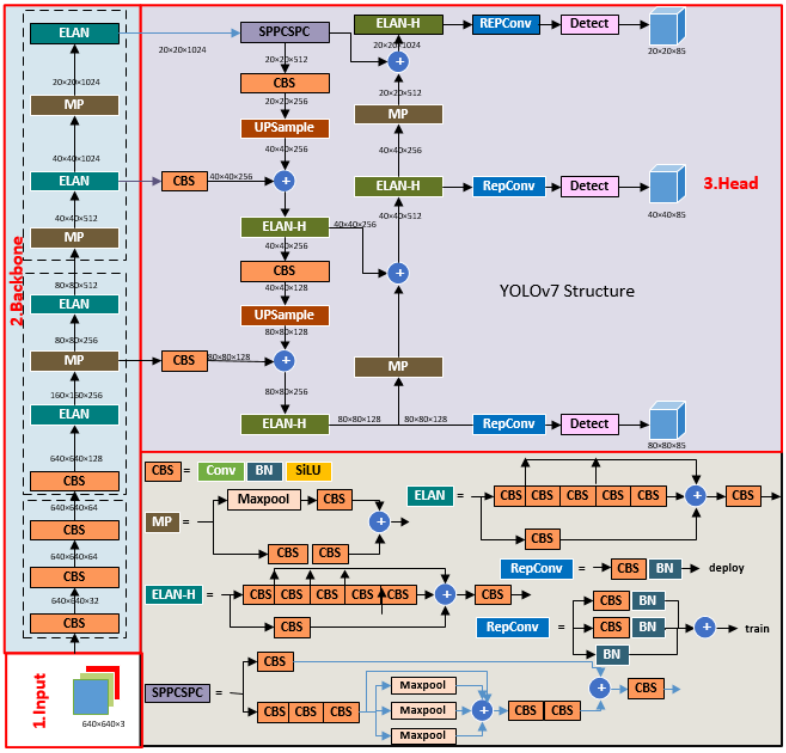


Fig. 2 - Architecture of YOLOv7

The YOLOv7 network structure primarily comprises an Input layer, Backbone layer, and Head layer. The Input layer serves as the initial layer, mainly processing the input images and directing them to the Backbone layer; the Backbone layer, also referred to as the feature extraction layer, is responsible for

extracting features of targets of varying sizes and consists of several CBS modules, MaxPool(MP), and efficient layer aggregation network(ELAN) structures. The Head layer's primary function is to integrate the features provided by the Backbone layer to generate bounding boxes and predict classes, which includes SPPCSPC(Spatial Pyramid Pooling, Cross Stage Partial Channel) layers, several ELAN-H layers, MP layers, and RepConv layers.

Addition of the SimAM Attention Module

Early-stage apple leaf diseases often present as light-coloured and small areas, similar to the colour of the leaves, and can be difficult to distinguish in complex, blurry, or poorly lit environments. As a result, the key information of the targets may become obscured by noise and other interfering factors, leading to lower accuracy. To address this issue, the SimAM attention mechanism was chosen for its lightweight design and strong capability for feature enhancement and representation, coupled with efficient computational performance.

The SimAM attention module was proposed by Yang et al., inspired by the attention mechanisms of the human brain *(Yang et al., 2021)*. It calculates attention weights by exploring the significance of each neuron and assesses the importance of each neuron based on the linear separability between the target neuron and other neurons, as informed by neuroscientific theory. The energy function is defined by Equation (1) as follows:

$$e_t\left(w_t,b_t,y,x_i\right)=\frac{1}{M-1}\sum_{i=1}^{M-1}\left(-1-\left(w_t x_i+b_t\right)\right)^2+\left(1-\left(w_t t+b_t\right)\right)^2+\lambda w_t^2 \quad (1)$$

Where t and x_i represent the target neuron and other neurons of the input feature X, i indexes the spatial dimension, M is the number of neurons in a given channel, y is the label indicating the significance of the neuron, w_t and b_t are the weights and biases, and λ is the regularization coefficient. Simplifying this equation leads to a minimal energy function, expressed as Equation (2):

$$e_t^*=\frac{4\left(\hat{\sigma}^2+\lambda\right)}{(t-\hat{\mu})^2+2\hat{\sigma}^2+2\lambda} \quad (2)$$

Where *t* is the target neuron, *x* represents adjacent neurons, and λ is a hyperparameter, $\hat{\mu}=\frac{1}{M}\sum_{i=1}^{M}x_i$, $\hat{\sigma}^2=\frac{1}{M}\sum_{i=1}^{M}\left(x_i-\hat{\mu}\right)^2$.

Finally, features are enhanced as indicated by Equation (3).

$$\tilde{X}=\text{sigmoid}\left(\frac{1}{E}\right)\odot X \quad (3)$$

This research embeds SimAM within the feature extraction layer of YOLOv7 to identify important neurons while suppressing surrounding ones, enhancing relevant features without increasing the parameter count of the model. This is particularly beneficial for recognizing features of apple leaf disease targets.

Loss Function Improvement: SIoU Loss Function

The quality of the loss function directly impacts the training speed and model detection performance. YOLOv7 incorporates three types of losses: Localization loss, Confidence loss, and Classification loss. Among these, the Localization loss function utilizes Complete Intersection over Union(CIoU). CIoU primarily considers the overlapping area, distance, and aspect ratio between the predicted and ground-truth bounding boxes but does not account for the mismatched orientation between them. This deficiency can lead to slower convergence and low efficiency. To address this, Gevorgyan proposed the SIoU loss function, to further improve the model recognition accuracy *(Gevorgyan, 2022).* SIoU loss was used to replace the CIoU loss in this paper.

The SIoU loss includes four components: angle cost, distance cost, shape cost, and IoU cost. The parameters used in the SIoU loss function are illustrated in Figure 3.

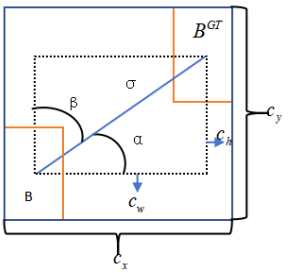


Fig. 3 - Calculation of SIoU loss function

Angle cost:

The calculation of the angle cost is shown in Equation (4):

$$\Lambda = 1 - 2 * \sin^2\left(\arcsin(x) - \frac{\pi}{4}\right) \tag{4}$$

where:

$x = \frac{c_h}{\sigma} = \sin(\alpha), c_h = \max\left(b_{cy}^{gt}, b_{cy}\right) - \min\left(b_{cy}^{gt}, b_{cy}\right)$ is the height difference between the centres of the two boxes and $\sigma = \sqrt{(b_{cx}^{gt} - b_{cx})^2 + (b_{cy}^{gt} - b_{cy})^2}$ is the width difference. If the angle α is greater than 45°, its complementary angle β is used, considering the angle from the y-axis rather than the x-axis.

Distance cost:

The distance cost represents the distance between the centre points of the predicted and ground-truth bounding boxes. Combining the angle cost defined above, SIoU redefines the distance cost as shown in Equation (5):

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho t}) \tag{5}$$

where:

$\rho_x = (\frac{b_{cx}^{gt} - b_{cx}}{c_w})^2$, $\rho_y = (\frac{b_{cy}^{gt} - b_{cy}}{c_H})^2$, $\gamma = 2 - \Lambda$, and the terms involving c_w and c_H represent the widths and heights of the smallest enclosing boxes for the ground-truth and predicted boxes. When α approaches 0, the contribution of the distance cost significantly decreases. Conversely, as α approaches π/4, the contribution of the distance cost increases. With increasing angle, γ is assigned a time-priority distance value.

Shape cost:

The definition of the shape cost Ω is shown in Equation (6):

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^{\theta} \tag{6}$$

where:

$w_w = \frac{|w - w^{gt}|}{\max(w, w, ^{gt})}$, $w_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$, and θ represent the attention of the network to shape.

The SIoU loss function is shown in Equation (7):

$$L_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \tag{7}$$

Model Evaluation Metrics

Model performance is assessed by comparing the detection results of pre- and post-improvement network models on various types of images under the same experimental conditions, focusing on missed and false detections. The study employs precision (P), recall (R), the precision-recall (P-R) curve, and mean average precision (mAP) as evaluation metrics.

The equations for these metrics are as follows:

$$P = \frac{TP}{TP + FP} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

$$mAP = \frac{\sum_{i=1}^{C} AP(i)}{C} \tag{10}$$

Where:

TP (true positive) represents correct positive predictions, FP (false positive) incorrect positive predictions, and FN (false negative) incorrect negative predictions; C is the number of classes detected. AP represents the area under the P-R curve, measuring the detection precision for a specific class. AP(i) indicates the AP value for the detection of the i-th class. The mAP is the arithmetic mean of the APs for all classes, assessing the overall detection performance of the model. The higher the mAP, the better the detection performance. Typically, mAP is used to evaluate the performance of the entire object detection network model.

RESULTS AND ANALYSIS

The experimental setup utilized the Windows 10 operating system with the deep learning framework PyTorch 1.7.1 and the programming language Python 3.8. The CPU model is Intel(R) Core(TM) i7-13700F @2.10GHz. The GPU model is NVIDIA GeForce RTX 3070.GPU acceleration libraries used are CUDA 10.2 and CUDNN 8.3.3. During training, the input image size was set to 640*640 pixels, initial learning rate was set to 0.001, momentum was set to 0.9, the batch size was set to 4, and the epoch was set to 100. The Adam optimizer was used to optimize the network parameters.

Dataset Analysis

The characteristics of apple leaf disease spots change noticeably over different periods. For instance, early-stage spots of Alternaria Blotch and Grey Spot are small and sparse, taking on circular or oval shapes. In later stages, the spots become larger, darker, and more densely distributed. Rust and Brown Spot typically start as light-coloured circular spots on apple leaves. To prevent issues of overfitting and poor generalization due to imbalanced dataset distribution or insufficient samples in certain categories, an initial data analysis was conducted, focusing on the distribution of target quantities, the number of disease categories, and the distribution of small targets.

Analysing by the number of disease categories in the dataset, 1378 images contain only one category of disease, while 265 images contain more than one. By target quantity in the images, 642 images have only one detection target, and 1001 images have more than one target. The distribution of target quantities across different disease categories is as follows: 996 targets for Alternaria Blotch, 1935 for Brown Spot, 885 for Grey Spot, and 843 for Rust. Brown Spot has the most targets, while the other three diseases have a similar number of targets.

In the dataset of apple leaf diseases, a total of 4659 targets were identified among the four disease categories. After annotating the targets, XML files were generated. The ratio of the area of the target to be predicted in the image to the image was calculated, based on the equation (11) provided:

$$P = \frac{area_t}{area_i} * 100\% \tag{11}$$

where $area_t = width * height$, it is the area of the target, $area_i$ is the area of the image, and the P is the ratio of the target area to the image area. $width = x\max - x\min + 1$, $heigh = y\max - y\min + 1$, width and height are expressed in pixel units. xmin and ymin are the x and y coordinate values of the top-left corner of the bounding box, while xmax and ymax are the x and y coordinate values of the bottom-right corner.

All P values were calculated using the above equation, a threshold of P was set, the data of the disease area proportion of all targets were analysed using Python, and then a histogram was output, as shown in Figure 4. As seen in Figure 4, there are 2352 small targets with p-value less than 6%, accounting for 50.5% of all targets. Since early-stage apple leaf spots are often small, the dataset used in this experiment provides a sufficient data foundation for the detection of early-stage small targets. The experimental results also provide a strong basis for the detection and prevention of early-stage small targets.

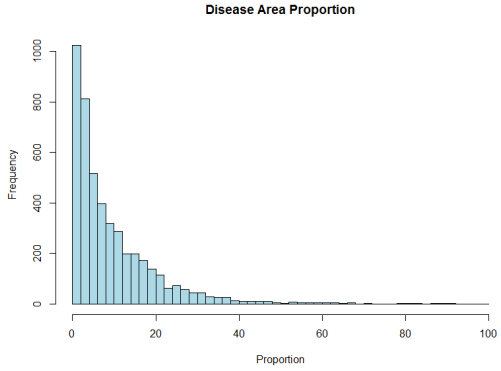


Fig. 4 - Histogram of disease area proportion

Ablation Study and Data Analysis

To verify the effectiveness of the improvements proposed in this study, ablation experiments were conducted on the dataset. Improvements are denoted by "√", while the absence of improvements is indicated by "–". The impact of detecting different models on the performance of apple leaf disease target detection is shown in Table 2.

Table 2

Results of the ablation study for different improvements

Baseline	Order	ECA	CBAM	SE	SimAM	Focal-EIOU	SIOU	Precision (%)	Recall (%)	mAP (%)
YOLOv7	1	—	—	—	—	—	—	77.86	68.15	77.41
	2	√	—	—	—	—	—	80.74	67.98	79.21
	3	—	√	—	—	—	—	80.90	71.87	81.35
	4	—	—	√	—	—	—	84.27	67.46	79.06
	5	—	—	—	√	—	—	87.80	99.00	94.30
	6	—	—	√	—	√	—	80.21	72.32	84.50
	7	—	—	√	—	—	√	78.00	97.00	82.00
	8	—	—	—	√	√	—	77.00	98.00	81.70
	9	—	—	—	√	—	√	92.00	99.00	96.10

Note: "√" represents the use of improvements, while "–" indicates no improvements used.

To validate the effectiveness of incorporating the SimAM attention mechanism into the model and to further analyse the impact of different attention mechanisms on the detection performance of apple tree leaf diseases under varying periods and environmental conditions, this study compared the introduction of ECA, CBAM, SE, and SimAM attention modules into the YOLOv7 network structure. By analysing and comparing the results, the strengths and weaknesses of each attention mechanism can be identified, establishing the effectiveness of SimAM in object detection tasks.

From Table 2, the following conclusions can be drawn:

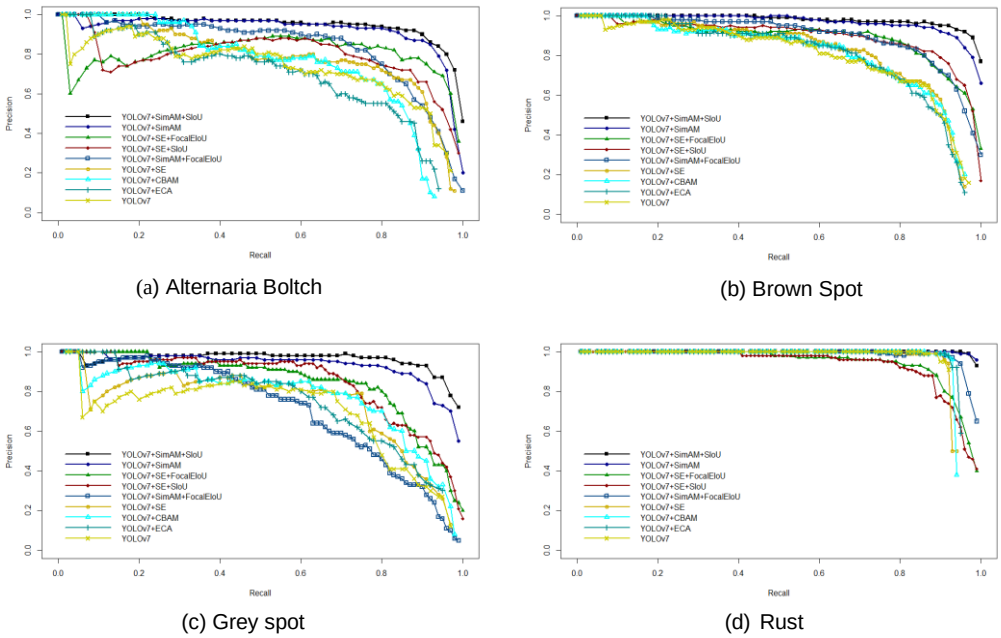
(1) The YOLOv7 model without the introduction of an attention mechanism achieved a mean average precision (mAP) of 77.41%. The introduction of attention mechanisms significantly improved the mAP value, with the SimAM attention mechanism reaching 94.3%. Compared to the ECA, CBAM, and SE attention modules, the mAP values increased by approximately 15.09%, 12.95%, and 15.24%, respectively, further confirming the accuracy of SimAM in target detection under different periods and environments.

(2) The introduction of the SE attention mechanism with the Focal-EIOU loss function achieved an mAP of 84.5%, while using the SIoU loss function resulted in an mAP of 82%, indicating a slight decrease in detection performance. Introducing the SimAM attention mechanism with the Focal-EIOU loss function resulted in an mAP of 81.7%, and using the SIoU loss function increased the mAP to 96.1%, showing a marked improvement in detection performance.

(3) Implementing both SimAM and SIoU improvements increased the mAP value for apple leaf disease detection from 77.41% to 96.1%, an enhancement of 18.69%, indicating a significant improvement in detection performance. Compared to other models in the experiment, the combination of YOLOv7, SimAM, and SIoU yielded the best results in all three evaluative measures: Precision, Recall, and mAP.

The research demonstrates that the improved YOLOv7 algorithm significantly outperforms the baseline YOLOv7 algorithm, with a 14.4% increase in Precision, a 38.85% increase in Recall, and an 18.69% increase in mAP. It also surpasses the other seven comparative models, indicating that the improved YOLOv7 model performs exceptionally well and can achieve rapid and accurate identification of apple leaf disease targets in natural environments. To provide a more intuitive analysis of the improvements of the YOLOv7 model over the original, the experiment draws a PR curves to reflect changes in Precision and Recall for different disease detections using different schemes, as shown in Figure 5.

Observation of Figure 5 reveals that the experiments using eight improvement schemes for the four diseases show many similarities: The four improvement schemes YOLOv7+ECA, YOLOv7+CBAM, YOLOv7+SE, YOLOv7+SimAM+FocalEIoU did not exhibit significant improvements in detection precision compared to the YOLOv7 model. The improvement schemes YOLOv7+SE+SIoU and YOLOv7+SE+FocalEIoU showed clear advantages over the aforementioned four schemes in detecting Alternaria Boltch, Brown Spot, and Grey spot, but not as much in detecting Rust. The schemes YOLOv7+SimAM and YOLOv7+SimAM+SIoU demonstrated markedly superior performance in Precision and Recall over the others, with a very significant improvement in detection performance compared to the YOLOv7 model. According to the information reflected in Figure 5, considering the detection performance of the model, it can be observed that YOLOv7+SimAM+SIoU has the best detection performance for apple leaf disease in this paper.



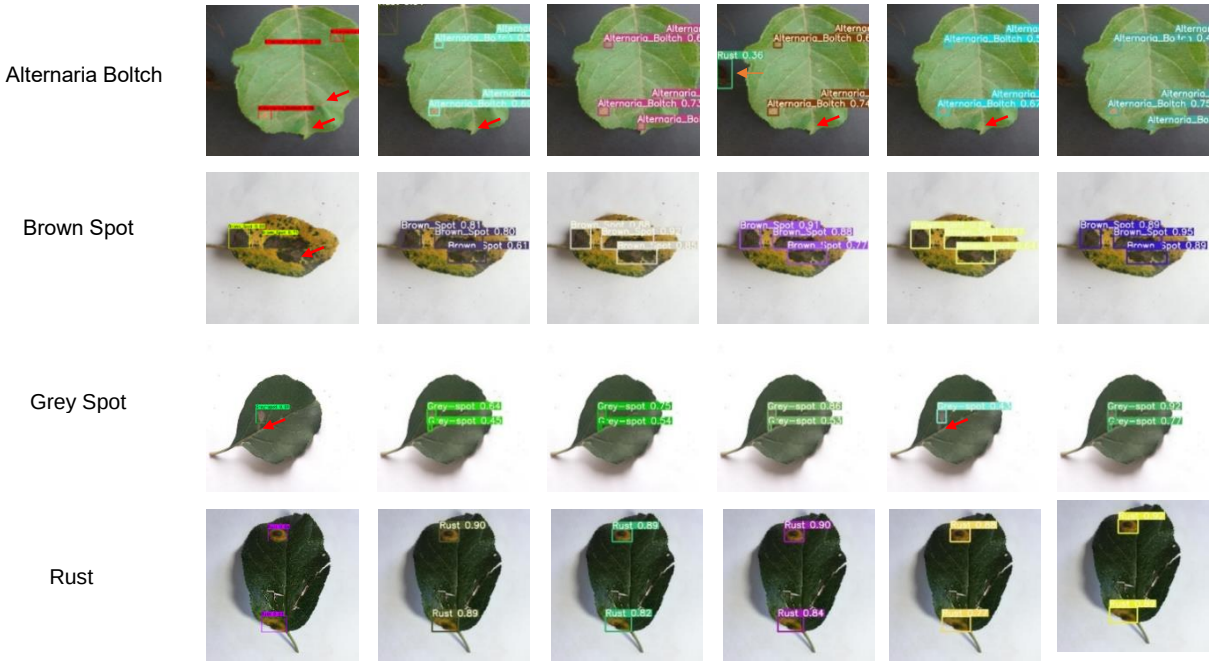
(a) Alternaria Boltch (b) Brown Spot

(c) Grey spot (d) Rust

Fig. 5 - P-R curve

Visualization Analysis of Detection Results

Figure 6 shows the detection results of various models for different classes of apple leaf disease targets against a simple background.



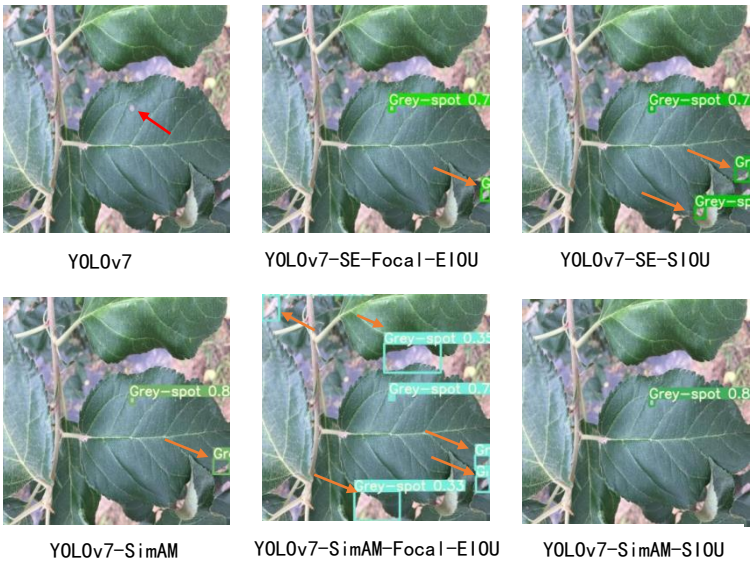
Note: ← Points to a missed target ← Points to a false detection

Fig. 6 - Prediction results of six models against a simple background

The first row shows the detection results for Alternaria Boltch; the original image had 5 targets, the YOLOv7 model detected 3, YOLOv7-SE-Focal-EIOU, YOLOv7-SimAM, and YOLOv7-SimAM-Focal-EIoU detected 4, missing one, and YOLOv7-SimAM had one false positive. YOLOv7-SE-SIOU and YOLOv7-SimAM-SIOU detected all 5 targets. The second row shows the detection results for Brown Spot; the original image had 3 targets, and apart from the YOLOv7 model missing one target, the other models detected all 3. The third row is for Grey spot detection results; YOLOv7 and YOLOv7-SimAM-Focal-EIoU missed one target, while the others detected all. The fourth row shows the detection results for Rust, all models detected all targets.

By comparing the detection results, it is found that the proposed model YOLOv7-SimAM-SIOU can fully detect all targets for the four diseases, and YOLOv7-SimAM-SIOU has the highest corresponding location (bounding box) and confidence values.

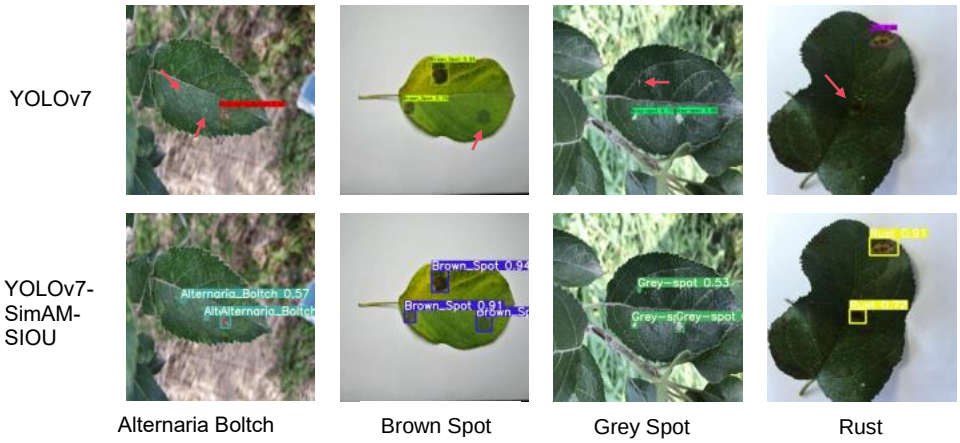
Figure 7 displays the detection results for apple leaf disease targets by various models against a complex background. In the original image, there is a small target of Grey spot, which due to its diminutive size and the complex background where the colour of the gaps between leaves closely resembles that of the Grey spot, poses a challenge for detection. As a result, YOLOv7 failed to detect Grey spot, and models such as YOLOv7-SE-Focal-EIOU, YOLOv7-SE-SIOU, YOLOv7-SimAM and YOLOv7-SimAM-Focal-EIoU mistook similar patterns in the background for the target, leading to false positives. YOLOv7-SimAM–SIOU did not suffer from missed or false detections.



Note: ←Points to a missed target ←Points to a false detection

Fig. 7 - Prediction results of six models against a complex background

Figure 8 shows the detection results of YOLOv7 and YOLOv7-SimAM–SIOU for early-stage apple leaf disease targets. In the early stages of infection on apple leaves, the disease spots appear lighter in colour and smaller in size. For such targets, YOLOv7 had a high rate of missed detections, while the model proposed in this paper, YOLOv7-SimAM-SIOU, detected all early-stage disease targets. This demonstrates that the algorithm proposed in this paper clearly outperforms the YOLOv7 algorithm in detecting small targets in the initial stages of the disease.



Note: ← Points to a missed target.

Fig. 8 - Prediction results of YOLOv7 and YOLOv7-SimAM-SIOU

CONCLUSIONS

To achieve accurate detection of apple leaf disease targets at different stages and in various environments, this study improved the YOLOv7 model by integrating SimAM and SIoU with YOLOv7, achieved a mAP of 96.1%, a precision of 92%, and a recall of 99% for apple leaf disease detection. Comparative experiments show that the performance of this model is clearly superior to the original YOLOv7 network. Comparison tests with seven other target detection models that the improved model is also superior, further verifying the effectiveness of the improved method for recognizing apple leaf diseases at different stages and in different environments. Through visual analysis, the improved YOLOv7 model shows its good performance in detecting apple leaf diseases in various aspects such as reducing false negatives, false positives, multi-category disease detection, and detection of small disease spots in the early stages of infection. Additionally, the model is minimally affected by factors such as light intensity, weather changes, indicating that the model has good robustness. The optimized YOLOv7 model achieves higher accuracy in apple leaf disease detection and can provide a technical reference for the prevention and diagnosis of plant diseases. In subsequent research, the focus will be on further optimizing YOLOv7-SimAM-SIOU and transferring it to more crop disease detection tasks.

ACKNOWLEDGEMENT

This research was financed by the Science and Technology Innovation Project of Colleges and Universities in Shanxi Province, China (No. 2022L085).

This research was financed by the Basic Research Program of Shanxi Province, China (No.202303021222058).

REFERENCES

[1] Bansal, P., Kumar, R., & Kumar, S. (2021). Disease detection in apple leaves using deep convolutional neural network. *Agriculture*, *11*(7), 617. https://doi.org/10.3390/agriculture11070617

[2] Bi, C., Wang, J., Duan, Y., Fu, B., Kang, J.-R., & Shi, Y. (2022). MobileNet Based Apple Leaf Diseases Identification. *Mobile Networks and Applications*, *27*(1), 172–180. https://doi.org/10.1007/s11036-020-01640-1

[3] Chao, X., Sun, G., Zhao, H., Li, M., & He, D. (2020). Identification of apple tree leaf diseases based on deep learning models. *Symmetry*, *12*(7), 1065. https://doi.org/10.3390/sym12071065

[4] Gevorgyan, Z. (2022). *SIoU Loss: More Powerful Learning for Bounding Box Regression* (arXiv:2205.12740). arXiv. http://arxiv.org/abs/2205.12740

[5] He, J., Liu, T., Li, L., Hu, Y., & Zhou, G. (2023). MFaster R-CNN for Maize Leaf Diseases Detection Based on Machine Vision. *Arabian Journal for Science and Engineering*, *48*(2), 1437–1449. https://doi.org/10.1007/s13369-022-06851-0

[6] Khan, A. I., Quadri, S. M. K., Banday, S., & Shah, J. L. (2022). Deep diagnosis: A real-time apple leaf disease detection system based on deep learning. *Computers and Electronics in Agriculture*, *198*, 107093. https://doi.org/10.1016/j.compag.2022.107093

[7] Li, H., Shi, L., Fang, S., & Yin, F. (2023). Real-Time Detection of Apple Leaf Diseases in Natural Scenes Based on YOLOv5. *Agriculture*, *13*(4), 878. https://doi.org/10.3390/agriculture13040878

[8] Mathew, M. P., & Mahesh, T. Y. (2022). Leaf-based disease detection in bell pepper plant using YOLO v5. *Signal, Image and Video Processing*, *16*(3), 841–847. https://doi.org/10.1007/s11760-021-02024-y

[9] Pathan, M., Patel, N., Yagnik, H., & Shah, M. (2020). Artificial cognition for applications in smart agriculture: A comprehensive review. *Artificial Intelligence in Agriculture*, *4*, 81–95. https://doi.org/10.1016/j.aiia.2020.06.001

[10] Priyadharshini, G., & Judie Dolly, D. R. (2023). Comparative Investigations on Tomato Leaf Disease Detection and Classification Using CNN, R-CNN, Fast R-CNN and Faster R-CNN. *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 1540–1545. https://doi.org/10.1109/ICACCS57279.2023.10112860

[11] Sahu, S. K., & Pandey, M. (2023). An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model. *Expert Systems with Applications*, *214*, 118989. https://doi.org/10.1016/j.eswa.2022.118989

[12] Sajitha, P., Andrushia, D. A., & Suni, S. S. (2023). Multi-class Plant Leaf Disease Classification on Real-Time Images Using YOLO V7. *International Conference on Image Processing and Capsule Networks*, 475–489. https://doi.org/10.1007/978-981-99-7093-3_32

[13] Singh, S., Gupta, S., Tanta, A., & Gupta, R. (2022). Extraction of Multiple Diseases in Apple Leaf Using Machine Learning. *International Journal of Image and Graphics*, *22*(03), 2140009. https://doi.org/10.1142/S021946782140009X

[14] Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors* (arXiv:2207.02696). arXiv. http://arxiv.org/abs/2207.02696

[15] Wang, Y., & Zhao, J. (2022). MGA-YOLO: A lightweight one-stage network for apple leaf disease detection. *Frontiers in Plant Science*, *13*, 927424. https://doi.org/10.3389/fpls.2022.927424

[16] Yang, L., Zhang, R.-Y., Li, L., & Xie, X. (2021). SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks. In M. Meila & T. Zhang (Eds.), *Proceedings of the 38th International Conference on Machine Learning* (Vol. 139, pp. 11863–11874). PMLR. https://proceedings.mlr.press/v139/yang21o.html

[17] Zhong, Y., & Zhao, M. (2020). Research on deep learning in apple leaf disease recognition. *Computers and Electronics in Agriculture*, *168*, 105146. https://doi.org/10.1016/j.compag.2019.105146