A Detection Method for Crop Diseases and Pests Based on Improved YOLOv7

Jie Kang, Xiaoying Chen

Abstract—**With the rapid advancement of science and technology, artificial intelligence (AI) and machine learning (ML) have been widely applied in agriculture. To tackle challenges in crop disease and pest identification—such as diverse species, subtle inter-class feature differences, and significant intra-class variations across crop growth stages—an improved YOLOv7 (IP-YOLOv7) object detection algorithm integrated with a hybrid attention mechanism is proposed. By incorporating the hybrid attention module into the backbone network of YOLOv7, the algorithm enhances its capability to learn pathological features and focuses more effectively on small-scale effective feature regions of crop leaves, thereby improving the identification accuracy of YOLOv7 for various crop diseases and pests. Experimental results demonstrate that YOLOv7 achieves a mean average precision (mAP) of 95.17%, while IP-YOLOv7 reaches 97.35% in crop disease and pest detection, indicating high accuracy and robustness of IP-YOLOv7 in this task.**

Index Terms—**crop diseases and pests, IP-YOLOv7, hybrid attention mechanism, object detection algorithm**

I. Introduction

As the foundation of the national economy, agriculture occupies a pivotal strategic position in a country's development [1,2]. The healthy growth and stable high yields of crops are not only critical to the effective supply of agricultural products but also constitute core elements in maintaining social stability and ensuring food security. Furthermore, the accurate identification and timely prevention of crop diseases and pests, as a key link in safeguarding crop yield and quality, directly affect the economic and ecological benefits of agricultural production [3-5]. Errors in disease and pest identification may result in inappropriate prevention measures, leading to reduced crop yields or even total crop failure, thereby posing a serious threat to farmers' incomes and national food security [6,7].

With the rapid development of modern science and technology, emerging cutting-edge technologies such as Artificial Intelligence (AI) and Machine Learning (ML), endowed with powerful data analysis and pattern recognition capabilities, have been extensively and deeply applied in the agricultural sector [8-10]. The introduction of these advanced technologies has driven a revolutionary transformation in the identification of crop diseases and pests. Traditional identification methods have been transcended, and new opportunities and approaches characterized by higher accuracy and efficiency have been provided for the prevention and control of agricultural diseases and pests. Through the learning and analysis of large volumes of pest- and disease-related images and data, AI and ML algorithms can rapidly and accurately identify various types of diseases and pests. This not only offers a scientific basis for the timely implementation of targeted prevention and control measures but also strongly facilitates the modernization of agriculture [11,12].

Crop disease and pest identification is confronted with several challenges. A wide variety of crop diseases and pests exist, and distinct types may exhibit extremely subtle differences in their characteristics, rendering accurate differentiation difficult [13-15].

In recent years, object detection algorithms based on deep learning have demonstrated significant potential in the field of image recognition, including the identification of crop diseases and pests. Among these, the You Only Look Once (YOLO) series of algorithms have garnered considerable attention due to their high efficiency and relatively favorable accuracy [16-18]. Specifically, YOLOv7 has achieved notable improvements in detection speed and accuracy compared to its predecessors [19,20]. However, when addressing the complex task of crop disease and pest identification, the existing YOLOv7 algorithm remains limited in its ability to fully capture effective features in small areas of crop leaves, which may result in reduced identification accuracy for certain specific diseases and pests [21,22].

To address these challenges, attention mechanisms have been incorporated into deep learning models. Such mechanisms enable models to focus on critical feature regions within input data, thereby enhancing their performance [23-25]. The hybrid attention mechanism, which integrates different types of attention mechanisms, has exhibited excellent performance in certain computer vision tasks [26,27]. Through the integration of the hybrid attention mechanism into the YOLOv7 algorithm, it is anticipated that the model's capability to learn pathological features will be strengthened, and its focus on small-scale effective feature regions of crop leaves will be improved. This, in turn, is expected to enhance the identification accuracy of YOLOv7 for various crop diseases and pests [28-31].

In this study, we propose an improved YOLOv7 (IP-YOLOv7) object detection algorithm based on the hybrid attention mechanism. We incorporate the hybrid attention

Manuscript received April 7, 2025; revised July 11, 2025.

This work was supported in part by the Nanjing Agricultural Robotics and Equipment Engineering Research Center.

Jie Kang is an associate professor of the School of Mechanical and Electrical Engineering, Sanjiang University, Nanjing Jiangsu, 210012, China (Corresponding author to provide phone: 15345183633; e-mail: kang_jie@sju.edu.cn).

Xiaoying Chen is an associate professor of the School of Mechanical and Electrical Engineering, Sanjiang University, Nanjing Jiangsu, 210012, China (e-mail: 77788489@qq.com).

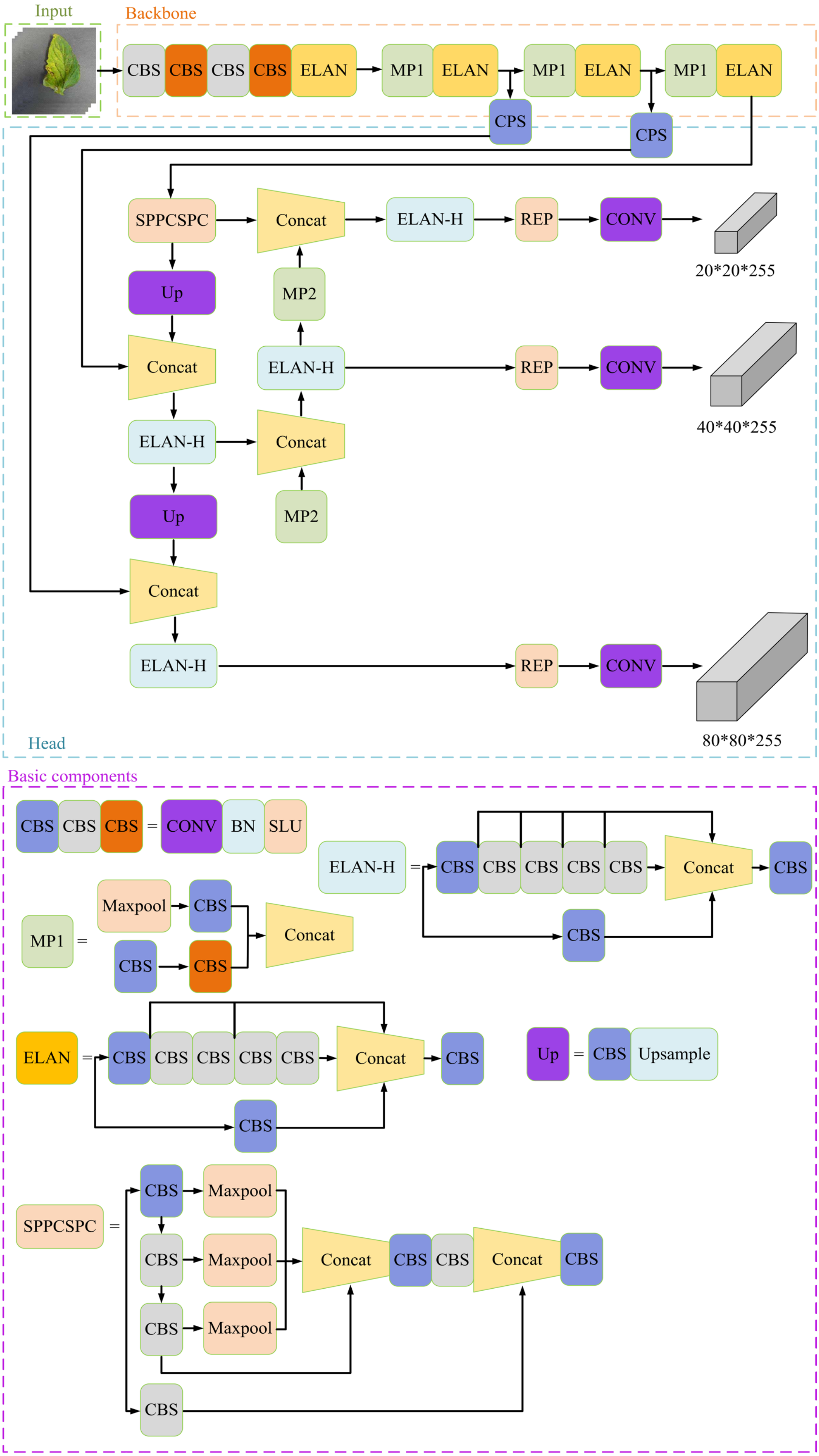

Fig. 1. Network architecture of the YOLOv7.

mechanism module into the backbone network of YOLOv7 to strengthen the model's ability to learn pathological features. Through extensive experiments, we compare the performance of IP-YOLOv7 with the original YOLOv7 in crop disease and pest detection.

II. MODEL ANALYSIS

A. YOLOv7

YOLOv7 achieves an excellent balance between detection speed and accuracy by introducing strategies such as the Extended-Efficient Long Aware Network (E-ELAN), model scaling based on Concatenation-Based models, and convolutional reparameterization. As shown in Fig. 1, the Network architecture of the YOLOv7 consists of Input, Backbone, Head, and Basic components. The Backbone is composed of several CBS modules, ELAN modules, and MP modules. The CBS module is made up of a convolutional layer, a batch normalization layer, and a SiLU activation function. The ELAN module is composed of several convolutional modules, and by controlling the shortest and longest gradient paths, it enables more efficient learning and convergence. The MP module is composed of a max pooling layer and several convolutional modules. It has two branches, upper and lower, which respectively perform downsampling on the feature maps, halving the length, width, and the number of channels of the image, and finally, feature fusion is carried out, thus improving the feature extraction ability of the model. During the training process, the YOLOv7 network adopts multiple loss functions, including coordinate loss, classification loss, and object existence loss. These loss functions jointly guide the optimization of the network parameters, enabling it to achieve good performance in object localization, classification, and detection.

TABLE I
TRAINING HYPERPARAMETER CONFIGURATION

Hyperparameter	Value
Learning rate	0.01
Optimization algorithms	Adam
Bach size	8
Dropout	0.5
Activation functions	ReLU
Epoch	150

To ensure the rigor of the experimental data in this study, the following section will standardize the hyperparameters for the training of this experiment. The specific data are shown in Table 1.

Common evaluation metrics include precision, recall, accuracy, average recognition precision, and F_1 score.

Precision (P_r) represents the proportion of true positive samples among the samples that have positive prediction results. Its calculation formula is shown in Eq. (1):

$$P_r = \frac{T_p}{T_p + F_p} \tag{1}$$

where, T_p represents true positive examples, F_p represents false positive examples, T_n represents true negative examples, and F_n represents false negative examples.

Recall (R_e) represents the ratio of the number of correctly predicted positive samples to the total number of actual positive samples. Its calculation formula is shown in Eq. (2):

$$R_e = \frac{T_p}{T_p + F_n} \tag{2}$$

Accuracy (A_c) represents the proportion of all the classes that are correctly predicted (positive and negative examples) among all the classes. Its calculation formula is shown in Eq. (3):

$$A_c = \frac{T_p + T_n}{T_p + F_n + T_n + F_p} \tag{3}$$

F_1 is the weighted average of precision and recall rate. Its calculation formula is shown in Eq. (4):

$$F_1 = \frac{2}{P_r^{-1} + R_e^{-1}} = 2 \cdot \frac{P_r \cdot R_e}{P_r + R_e} \tag{4}$$

The average precision (AP) is defined as the area enclosed by the Precision-Recall curve and the X-axis, where the Precision is plotted on the vertical axis and the Recall on the horizontal axis. The model performance is best when the AP value is 1. The mean average precision (mAP) is the average of all the AP values for each category. The formulas for calculating AP and mAP are as shown in Eq. (5) and Eq. (6):

$$AP = \int_0^1 P_r(R_e)dR_e \tag{5}$$

$$mAP = \left(\sum_{n=1}^{k} AP_k \right) \Big/ k \tag{6}$$

As can be seen from Fig. 2 and Fig.3, after 100 rounds of iterative training, the accuracy rates of the YOLOv7 model in both the training set and the validation set are almost overlapping in the latter part. This indicates that the fitting effect of this model is extremely good. Regarding the Loss value, after 100 rounds of training on the training set, the Loss value of the YOLOv7 model gradually approaches 0.05 and is almost consistent with the Loss value of the validation set, suggesting that this model does not have overfitting and can generalize well to new data. The final accuracy rate of the training set of this model is 96.04%, and the Loss value is 0.0459. The accuracy rate of the validation set is 95.37%, and the Loss value is 0.0733.

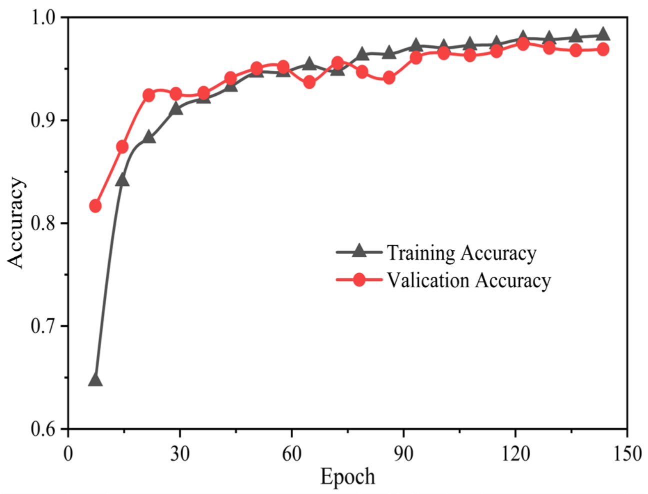

Fig. 2. Accuracy curve of YOLOv7 model.

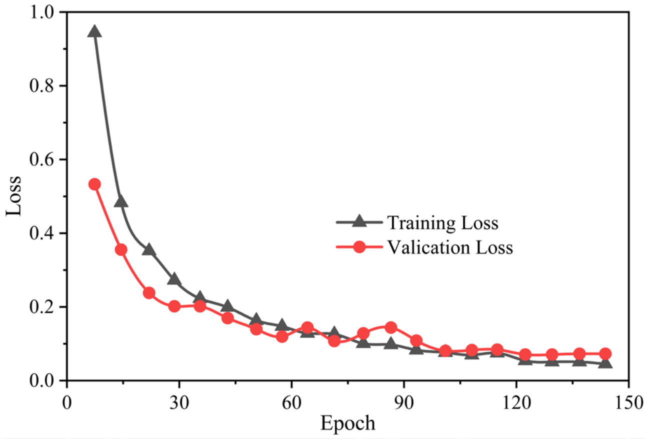

Fig. 3. Loss curve of YOLOv7 model.

B. IP-YOLOv7

Due to the diversity of crop diseases and pests, the diseased areas of some crops are too small, which causes difficulties in feature extraction by the YOLOv7 algorithm. Therefore, the method of YOLOv7 for recognizing disease and pest features still needs further improvement.

In order to enhance the accuracy of crop disease and pest detection, the CBAM module is utilized to assist the model in improving the ability of extracting disease-related features and weakening that of extracting non-disease-related features. Here, we combine the Head module and the CBS module in the Backbone module of the YOLOv7 network framework with the mixed attention mechanism CBAM module, and the improved network framework is shown in Fig. 4.

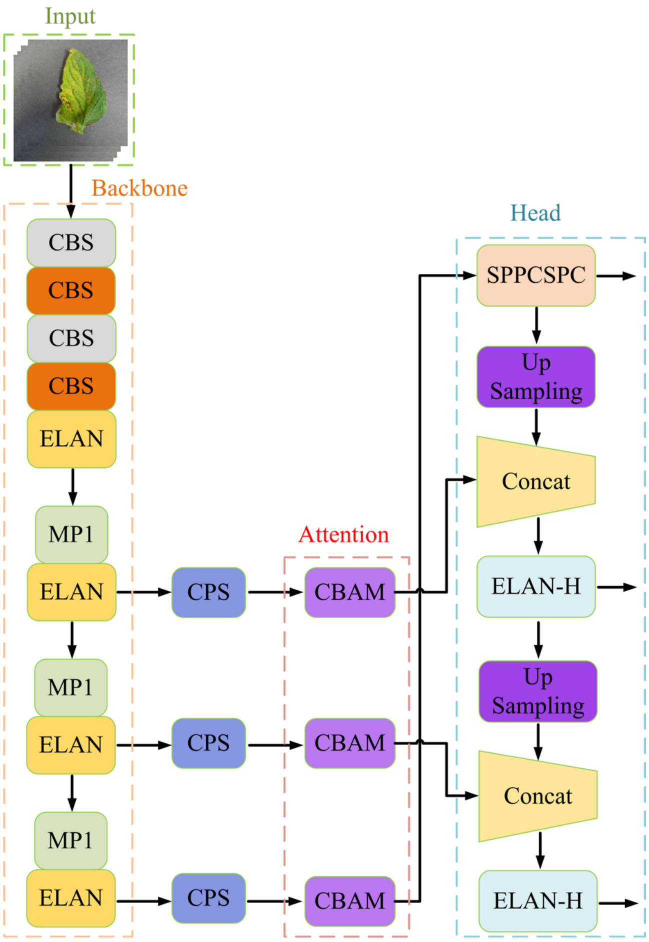

Fig. 4. The improved network framework of YOLOv7.

As can be seen from Fig. 5 and Fig. 6, the YOLOv7 algorithm improved based on the attention mechanism shows an overall upward trend in the recognition accuracy of the training set during these 150 rounds of training. Although there is a certain degree of fluctuation in the accuracy of the validation set, most of it overlaps with the data of the training set and also shows an upward trend, which reflects the strong fitting ability of the improved algorithm model.

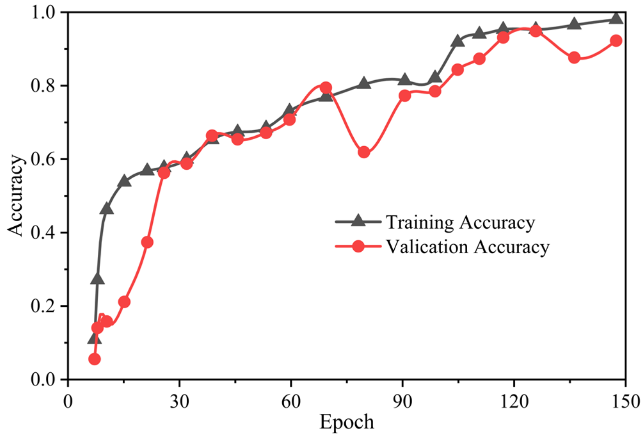

Fig. 5. Accuracy curve of IP-YOLOv7 model.

In terms of the Loss value, after 150 rounds of training on the training set, the Loss value of the YOLOv7 model gradually approaches around 0.01 and is almost the same as the Loss value of the validation set, indicating that this model does not exhibit the phenomenon of overfitting and can be well generalized to new data. The final accuracy of this model on the training set is 98.19%, and the Loss value is 0.0413. The accuracy of the validation set is 96.50%, and the Loss value is 0.1385. The recognition effect of the model is significantly better than that of the YOLOv7 algorithm before the improvement.

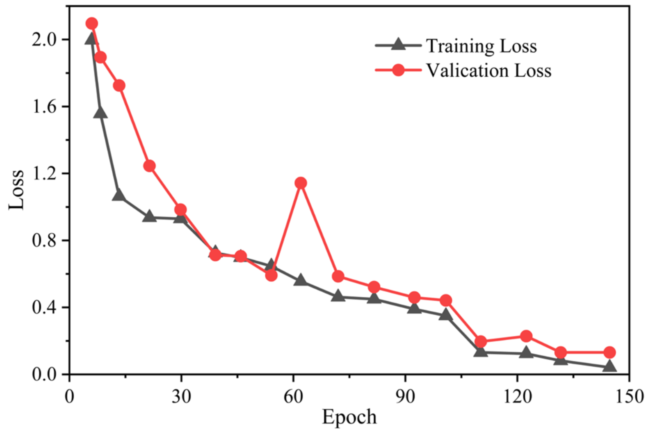

Fig. 6. Loss curve of IP-YOLOv7 model.

III. Results and Discussion

From the data comparison in Fig. 7, it can be seen that the overall performance indicators of IP-YOLOv7 are superior to those of other tested detection algorithms under the same dataset. It leads in terms of average recognition accuracy, F1 score, and mAP. The mAP value of the IP-YOLOv7 algorithm is 3.43% higher than that of the YOLOv7

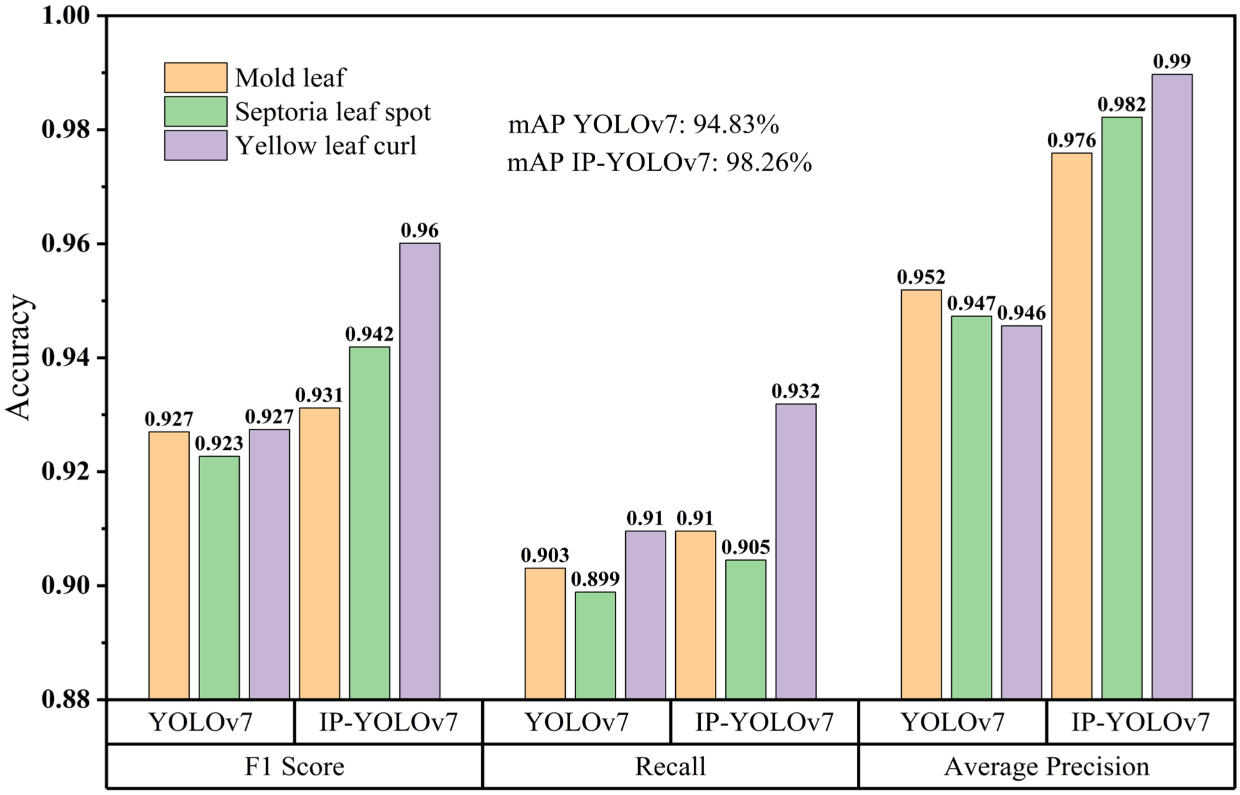

Fig. 7. Statistical data of F1 score, recall rate and average precision of YOLOv7 and IP-YOLOv7.

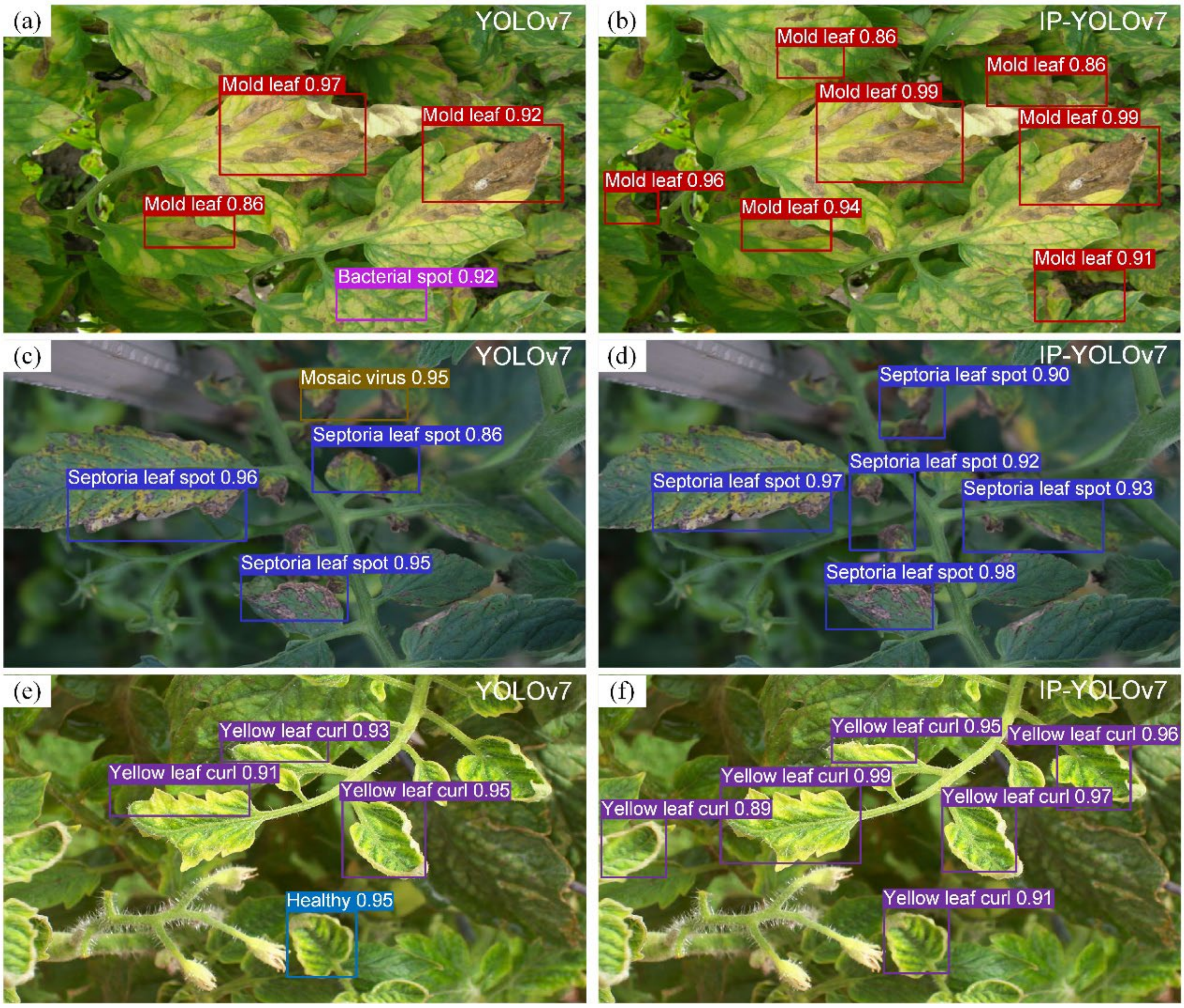

Fig. 8. Test results of YOLOv7 and IP-YOLOv7 networks. (a) Mold leaf detection results based on YOLOv7; (b) Septoria leaf spot detection results based on YOLOv7; (c) Yellow leaf curl detection results based on YOLOv7; (d) Mold leaf detection results based on IP-YOLOv7; (e) Septoria leaf spot detection results based on IP-YOLOv7; (f) Yellow leaf curl detection results based on IP-YOLOv7.

algorithm, and the other indicators are basically superior to those of the YOLOv7 algorithm. It is worth noting that due to the intervention of the CBAM hybrid attention mechanism, the IP-YOLOv7 algorithm pays more attention to pathological features and pays less attention to healthy state features. We will further verify this from the recognition images. The IP-YOLOv7 algorithm with the attention mechanism module is superior in identifying small-scale pest and disease features.

As can be observed from Fig. 8, when the YOLOv7 model without the attention mechanism is used to recognize these three types of images of plant diseases and insect pests, misjudgments and omissions occur. In Fig. 8(a), a part of the images of crop leaf mold disease is incorrectly identified as bacterial spot disease, and in Fig. 8(c), a blurred background is mistakenly recognized as crop mosaic virus disease. The improved IP-YOLOv7 algorithm recognition model demonstrates excellent recognition characteristics in these three types of images of plant diseases and insect pests. Due to the incorporation of the attention mechanism module, the learning of different pathological features is significantly enhanced, and the number of prior boxes is much larger than that in the YOLOv7 model. When comparing Fig. 8(e) with Fig. 8(f), IP-YOLOv7 does not perform corresponding recognition on the healthy leaves. This result is precisely caused by our weakening of the feature learning ability of this model for healthy leaves, which also reflects the success of incorporating the attention mechanism module from a side perspective.

IV. CONCLUSION

This paper addresses the issue of crop disease and pest detection and proposes an improved IP-YOLOv7 object detection algorithm based on the attention mechanism.

When the YOLOv7 algorithm extracts small-scale crop disease features, the number of target feature pixels is relatively small, which leads to the occurrence of missing effective features during extraction, resulting in missed detections or false detections. An algorithm improvement was carried out. The hybrid attention mechanism was applied to the YOLOv7 algorithm. Through the comparison of experimental evaluation indicators with the YOLOv7 object detection algorithm, it was proved that the improved IP-YOLOv7 algorithm has superiority in the task of crop disease and pest detection.

REFERENCES

[1] Y. N. Kuan, K. M. Goh, and L. L. Lim, “Systematic review on machine learning and computer vision in precision agriculture: Applications, trends, and emerging techniques,” Engineering Applications of Artificial Intelligence, Vol. 148, p. 110401, 2025.

[2] W. Chen, L. J. Zheng, and J. P. Xiang, “Algorithm for Crop Disease Detection Based on Channel Attention Mechanism and Lightweight Up-Sampling Operator,” IEEE Access, Vol. 12, pp. 109886-109899, 2024.

[3] D. M. Liu, H. X. Yang, Y. J. Gong, and Q. Chen, “A Recognition Method of Crop Diseases and Insect Pests Based on Transfer Learning and Convolution Neural Network,” Mathematical Problems in Engineering, Vol. 2022, p. 1470541, 2022.

[4] R. Chavan, S. Vagare, A. Jadhav, M. Mulande, and S. Adhav, “An In-Depth Analysis of Tomato Crop Diseases and Classification Using High-Performance Deep Neural Network,” IET Conference Proceedings, Vol. 2023, no. 22, pp. 188-194, 2023.

[5] L. J. Du, J. L. Zhu, M. H. Liu, and L. Wang, “YOLOv7-PSAFP: Crop pest and disease detection based on improved YOLOv7,” IET Image Processing, Vol. 19, no. 1, p. e13304, 2025.

[6] Z. F. Li, L. D. Wang, H. W. Lv, and Z. Y. Wang, "Adaptive Neural Network Control for Switched Stochastic Pure-Feedback Nonlinear Systems with Incomplete Measurements," IAENG International Journal of Applied Mathematics, vol. 55, no. 5, pp. 1014-1027, 2025.

[7] H. R. Ma, Y.-e Hou, and H. Z. Xu, "An Iterated Local Search Algorithm for the Heterogeneous Fixed Multi-compartment Vehicle Routing Problem," IAENG International Journal of Applied Mathematics, vol. 55, no. 5, pp. 1084-1091, 2025.

[8] D. D. Chen, Y. Wei, Z. G. Zhou, Z. W. Li, and P. Y. Chen, "Humanoid Robot Vehicle Driving Navigation Based on Monocular Vision," Engineering Letters, vol. 33, no. 5, pp. 1267-1281, 2025.

[9] D. H. Hou, Y. J. Zhang, and J. Ren, "A Lightweight Object Detection Algorithm for Remote Sensing Images," Engineering Letters, vol. 33, no. 3, pp 704-711, 2025.

[10] Y. C. Lu, X. B. Yu, Z. P. Hu, and X. M. Wang, “Convolutional neural network combined with reinforcement learning-based dual-mode grey wolf optimizer to identify crop diseases and pests,” Swarm and Evolutionary Computation, Vol. 94, p. 101874, 2025.

[11] I. Gallo, A. U. Rehman, R. H. Dehkordi, N. Landro, R. L. Grassa, and M. Boschetti, “Deep Object Detection of Crop Weeds: Performance of YOLOv7 on a Real Case Dataset from UAV Images,” Remote Sensing, Vol. 15, no. 2, p. 539, 2023.

[12] K. W. Chen, L. Z. Wang, H. Y. Wu, C. H. Wu, Y. Liao, Y. P. Chen, H. Wang, J. W. Yan, J. L. Lin, and J. L. He, “Background-Aware Correlation Filter for Object Tracking with Deep CNN Features,” Engineering Letters, Vol. 32, no. 7, pp. 1353-1363, 2024.

[13] S. Sungnul, W. Pornprakun, S. Prasattong, and C. Baitiang, “A mathematical model for the sugarcane trading system in Thailand,” Engineering Letters, Vol. 26, no. 1, pp. 84-91, 2018.

[14] J. Li, M. Q. Yang, Y. P. Liu, Y. Y. Wang, Q. H. Zheng, and D. Q. Wang, “Dynamic hand gesture recognition using multi-direction 3D convolutional neural networks,” Engineering Letters, Vol. 27, no. 3, pp. 490-500, 2019.

[15] Z. Kinda, S. Malo, T. R. Bayala, and I. Wonni, “Cotton Disease Detection on UAV Images: A Deep Learning-Based Approach with YOLOv7,” Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, Vol. 588, pp. 234-249, 2025

[16] G. Yu, H. M. Sun, Z. L. Xiao, and C. Y. Dai, “An improved algorithm for sesame seedling and weed detection based on YOLOV7,” International Journal of Wireless and Mobile Computing, Vol. 26, no. 3, pp. 282-290, 2024.

[17] J. Hu, C. Fan, Z. P. Wang, J. J. Ruan, and S. Y. Wu, “Fruit Detection and Counting in Apple Orchards Based on Improved Yolov7 and Multi-Object Tracking Methods,” Sensors, Vol. 23, no. 13, p. 5903, 2023.

[18] K. H. Zhang, Y. Z. Wu, N. Li, X. Y. Guo, R. Cheng, and L. J. Zhu, "Lightweight Road Defect Detection Algorithm Based on Improved YOLOv10," Engineering Letters, vol. 33, no. 5, pp. 1193-1202, 2025.

[19] L. Y. Sun, and Y. J. Zhang, "Steel Surface Defect Detection Based on YOLOv10," Engineering Letters, vol. 33, no. 5, pp. 1220-1231, 2025.

[20] T. Zhou, X. X. Zhang, and H. L. Chen, "A Dangerous Driving Behavior Detection Method Based on Improved YOLOv8s," Engineering Letters, vol. 33, no. 3, pp. 721-731, 2025.

[21] S. Y. Liu, Y. X. Li, C. H. Wang, Z. Y. Hu, K. Pei, and H. L. Li, "Defect Detection Algorithm for Body-in-white Weld Spots Based on Improved YOLOv5," Engineering Letters, vol. 33, no. 3, pp. 784-792, 2025.

[22] Y. G. Fu, and Y. Zhang, "Lightweight Rice Leaf Disease Detection Method Based on Improved YOLOv8," Engineering Letters, vol. 33, no. 2, pp. 402- 417, 2025.

[23] A. Q. Shan, J. Q. Pan, K. J. Kang, M. H. Pan, G. Wang, M. Wang, Z. L. He, and X. E. Yang, “Effects of straw return with N fertilizer reduction on crop yield, plant diseases and pests and potential heavy metal risk in a Chinese rice paddy: A field study of 2 consecutive wheat-rice cycles,” Environmental Pollution, Vol. 288, p 117741, 2021.

[24] M. Y. Xin, and Y. Wang, “Image Recognition of Crop Diseases and Insect Pests Based on Deep Learning,” Wireless Communications and Mobile Computing, Vol. 2021, p. 5511676, 2021.

[25] Y. G. Fu, and Y. Zhang, “Lightweight Rice Leaf Disease Detection Method Based on Improved YOLOv8,” Engineering Letters, Vol. 33, no. 2, pp. 402-417, 2025.

[26] E. Guevara, A. E. Rojas, and H. Florez,” Technology Platform for the Information Management of Theobroma Cacao Crops based on the

Colombian Technical Standard 5811," Engineering Letters, Vol. 30, np. 1, pp. 298-310, 2022.
[27] T. Liu, Q. Y. Yuan, and Y. G. Wang, "Prediction model of photosynthetic rate based on sopso-lssvm for regulation of greenhouse light environment," Engineering Letters, Vol. 29, no. 1, pp.297-301, 2021.
[28] A. K. Gautam, and R. Kapoor, "A Novel ES-RwCNN Based Finger Vein Recognition System with Effective L12 DTP Descriptor and AWM-WOA Selection," Engineering Letters, Vol. 30, no. 2, pp. 882-891, 2022.
[29] J. L. Lou, X. Y. Chi, H. Guang, J. Qi, Z.Y. Hong, and Y. Chuang, "Agri-NER-Net: Glyph Fusion for Chinese Field Crop Diseases and Pests Named Entity Recognition Network," Automatic Control and Computer Sciences, Vol. 58, no. 6, pp. 679-689, 2024.
[30] J. P. Nyakuri, C. Nkundineza, O. Gatera, and K. Nkurikiyeyezu, "State-of-the-Art Deep Learning Algorithms for Internet of Things-Based Detection of Crop Pests and Diseases: A Comprehensive Review," IEEE Access, Vol. 12, pp. 169824-169849, 2024.
[31] Z. C. Zheng, and C. Zhang, "Electronic noses based on metal oxide semiconductor sensors for detecting crop diseases and insect pests," Computers and Electronics in Agriculture, Vol. 197, p. 106988, 2022.

Jie Kang received the M.S. degree in Control Engineering from Southeast University in 2018. He is currently an associate professor at the School of Mechanical and Electrical Engineering, Nanjing Sanjiang University. His main research interests include machine vision and deep learning.

Xiaoying Chen received the Ph.D. in Instrument Science and Technology from Southeast University in China in 2015. She is currently an associate professor at the School of Mechanical and Electrical Engineering, Sanjiang University in Nanjing. Her main research interests include machine vision and metrological testing technology.