World Journal of Information Technology
Print ISSN: 2959-9903
Online ISSN: 2959-9911
DOI: https://doi.org/10.61784/wjit3112

YOLOV8-KF: A FRAMEWORK FOR PEST AND DISEASE DETECTION AND FRUIT-VEGETABLE RECOGNITION IN COMPLEX SCENARIOS

HanQiu Shen[1*], QinYi Yi[2], JunHao Liu[1], Ling Wu[1], XiangRun Xiao[1], Yuan Huang[1]
[1]*College of Information Science and Engineering, Hunan Institute of Engineering, Xiangtan 411104, Hunan, China.*
[2]*Yueyang Power Supply Branch, State Grid Hunan Electric Power Co., Ltd., Yueyang 414000, Hunan, China.*
[*]*Corresponding Author: HanQiu Shen*

Abstract: Crop pests and diseases constitute major constraints affecting crop yield and quality, and pose a threat to global food security. Therefore, in modern agricultural production, accurate and real-time detection of pests and diseases is critical for the timely implementation of prevention and control measures. Although YOLOv8 performs excellently in multi-class pest and disease detection as well as fruit and vegetable recognition, it yields unsatisfactory detection results under scenarios involving leaf occlusion and complex field monitoring. To address these issues, this study develops a framework for pest and disease detection and fruit and vegetable recognition that combines YOLOv8 with Kalman filter post-processing. The Kalman filter is used to predict target positions and smooth detection deviations under occluded conditions, thereby improving the continuity and stability of target tracking.The results show that the proposed model not only maintains high recognition accuracy for small-sized and multi-class pests, diseases and fruit-vegetable targets, but also reduces missed detection caused by occlusion. This method provides an effective solution for real-time dynamic monitoring of crop pests and diseases, fruit and vegetable recognition, and precise prevention and control in complex scenarios.
Keywords: Crop disease; Object detection; YOLOv8

1 INTRODUCTION

Crop pests and diseases are recognized as key factors that restrict the yield and quality of agricultural products [1], and rapid and accurate identification and monitoring are regarded as the core premise for realizing pest and disease prevention and control as well as ensuring food security [2]. The traditional manual identification method is inefficient and highly subjective, which makes it difficult to meet the practical needs of large-scale field monitoring [3]. YOLOv8 has become a preferred model for crop pest and disease identification due to its optimized network structure and efficient inference performance, and particularly notable advantages are observed in the identification of small-target pests and diseases [4]. In the identification of crop pests and diseases, many targets such as initial disease spots and small pests are characterized by small size, indistinct features, and frequent occlusion by leaves and branches, which easily lead to missed detection and false detection in traditional detection algorithms.

However, YOLOv8 enhances the ability of extracting small-target features by improving the C2f module of the backbone network, and it can effectively capture the subtle features of small-size pests and diseases by combining the multi-scale feature fusion mechanism of the SPPF module [5]. At the same time, the structure of the decoupled detection head is optimized, which improves the accuracy of positioning and classification of small targets. Thus, it can realize the simultaneous identification of multiple types of small and various pests and diseases, which is suitable for the multi-target detection needs under complex field backgrounds [6]. Nevertheless, in practical field video stream monitoring scenarios, problems such as interlaced branches and leaves, target overlap, and temporary occlusion occur frequently, which result in jitter, target identity switching, and even missed detection of YOLOv8 detection boxes, seriously affecting the stability and reliability of dynamic monitoring. To address this pain point, a Kalman filter algorithm is introduced in this study for post-processing optimization on the basis of the detection results of YOLOv8 for fruits and vegetables. Through an iterative process of prediction and update, the Kalman filter can predict the position information of the target under the occluded state, smooth the detection noise, correct the jitter problem of the detection box, maintain the continuity of the target identity, and effectively alleviate the identification deviation caused by occlusion [7].

To verify the practicality of the proposed fusion method, multiple types of pest and disease video sequences under actual field environments are selected in this study, and real scenarios such as leaf occlusion and interlaced branches and leaves are simulated to test the detection and tracking performance of the model [8]. The experimental results show that the fusion method of YOLOv8 and Kalman filter not only retains the high identification accuracy of YOLOv8 for multi-category and small-target pests and diseases, but also significantly improves the stability of target tracking under occluded scenarios, reducing missed detection and identity misjudgment. It is of great significance for the real-time dynamic monitoring of YOLOv8, the promotion of precise pest and disease prevention and control, and the improvement of the development level of smart agriculture [9].

2 RELATE WORK

2.1 Datasets

Healthy cotton leaves exhibit a deep green color with clear and intact veins; leaves damaged by cotton plant bugs, spider mites, and American serpentine leafminers show typical symptoms including holes, chlorotic spots, and larval mines, respectively. The distinct damage patterns caused by different pests are visually presented on the leaves, forming significantly differentiated appearances. These samples thus provide a visual feature set with clear class separability for model training. Representative samples are shown in Figure 1.

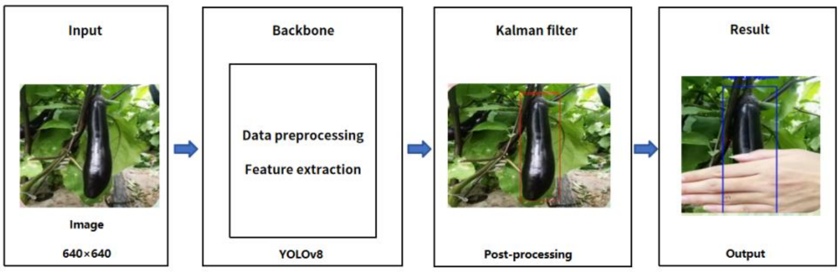

Figure 1 Detection and Recognition Framework

The cotton pest and disease datasets constructed in this study contains a total of 7484 annotated images, covering six typical stress types: noctuid, wilting, leafhopper, spider mite, mirid, and American leafminer. The datasets is divided into a training set and a validation set at an approximate ratio of 6:4, with 4492 images in the training set and 2992 images in the validation set.The sample distribution for each category is as follows: 1208 images for noctuid pests, 997 for wilting disease, 381 for leafhoppers, 1259 for spider mites, 1744 for mirid bugs, and 1895 for American serpentine leafminers. The overall sample distribution balances the representativeness of different pest and disease types, while the category distribution is consistent between the training and validation sets, providing a balanced data foundation for stable model training and reliable evaluation. The detailed distribution of the dataset is presented in the corresponding table.The detailed distribution of the dataset is presented in Table 1.

Table 1 The Precision Metrics of Different Models for Pest and Disease Recognition

Class	Noctuid	Wilting	Leafhopper	Spider mite	mirid	American leafminer	Sum
Train	725	599	229	756	1046	1137	4492
Val	483	398	152	503	698	758	2992
Sum	1208	997	381	1259	1744	1895	7484

2.2 Method

2.2.1 Framework

This figure illustrates the workflow of the proposed YOLOv8-KF framework for pest and disease detection and fruit-vegetable recognition in complex agricultural environments. Initially, a 640×640 field image containing eggplants is fed into the system as input. Subsequently, the image undergoes data preprocessing and feature extraction via the YOLOv8 backbone network, where preliminary target localization is performed. During the post-processing stage, a Kalman filter is applied to refine the detection results. Even when the target is partially occluded by a hand, the Kalman filter predicts and corrects the target's position based on its motion state, effectively mitigating the drift of the detection box caused by interference. Finally, the optimized result with a stable bounding box is outputted. By integrating the rapid detection capability of YOLOv8 with the trajectory prediction advantage of the Kalman filter, this framework significantly enhances the robustness of target recognition in dynamic and occluded scenarios, providing a reliable solution for agricultural monitoring in complex environments.The detailed procedure is illustrated in Figure 1.

2.2.2 Accuracy metrics

The core quantitative metrics for evaluating the performance of object detection models are defined by the following formulas. The overall accuracy (OA) is defined as the proportion of all samples that are correctly predicted by the model. Precision measures the proportion of true positive samples among all predicted positive samples, while recall represents the proportion of true positive samples that are correctly identified. The F1-score, which is the harmonic mean of precision and recall, is employed to balance the two aforementioned metrics. Average precision (AP) is defined as the integral of the area under the precision-recall (P-R) curve, which reflects the detection performance for a single category. Mean average precision (mAP), which is the arithmetic mean of the AP values across all categories, serves as the core comprehensive evaluation metric for multi-class object detection tasks as shown in Equations. (1)–(6):

$$OA=\frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

$$Precision=\frac{TP}{TP+FP} \tag{2}$$

$$Recall=\frac{TP}{TP+FN} \tag{3}$$

$$F1\_score=\frac{2*\ Precision*\ Recall}{Precision+Recall} \tag{4}$$

$$AP=\int_0^1 p(r)dr \tag{5}$$

$$mAP=\frac{\sum_{i=1}^{n} AP}{n} \tag{6}$$

TP refers to correctly classified positive samples, FP refers to negative samples misidentified as positive, TN refers to correctly classified negative samples, FN refers to positive samples misidentified as negative.

3 EXPERIMENT AND CONCLUSION

A recognition model for crop diseases and pests was constructed based on the YOLOv8 object detection framework, and a total of 200 training epochs were implemented. Through systematic analysis of the dynamic variations in loss functions and the evolutionary trends of key evaluation metrics during the training process, the convergence characteristics, fitting quality, and practical detection performance of the model for diseases and pests were comprehensively verified.At the level of loss functions, the bounding box localization loss, object confidence loss, and classification loss of the model all exhibited typical characteristics of rapid decline and subsequent stable convergence. In the initial stage of training, values of these three losses were relatively high, indicating that the model had not yet fully learned the feature information of disease and pest targets. Within the first 50 epochs, loss values decreased rapidly, which reflected the rapid enhancement of the model's capability in feature representation of disease and pest targets. After the 100th epoch, the rate of loss decline slowed significantly and entered a steady state; eventually, all losses remained at a low level throughout the 200 epochs. This evolutionary trend suggested that the optimization processes of the model for bounding box regression, confidence discrimination, and category classification of disease and pest targets proceeded smoothly and orderly. No abnormal phenomena such as gradient oscillation or divergence were observed during training, and the update of network parameters was rational and well controlled.In terms of model performance evaluation, precision, recall, mean average precision at the intersection-over-union threshold of 0.5, and mean average precision over the intersection-over-union range of 0.5 to 0.95 all demonstrated a stable upward trend with the increase in training epochs. In the early stage of training, these evaluation metrics were at relatively low levels, indicating that the detection performance of the model had not yet reached its optimal state. Within the first 100 epochs, the metrics increased rapidly, which reflected the quick optimization of the model's ability to recognize disease and pest targets. After the 150th epoch, the growth rate of the metrics slowed and gradually stabilized, ultimately maintaining at a high level. The continuous improvement of mean average precision at the intersection-over-union threshold of 0.5 and mean average precision over the intersection-over-union range of 0.5 to 0.95 indicated that the model could achieve accurate recognition and localization of disease and pest targets with various scales and growth stages under different intersection-over-union threshold constraints. This capability effectively balanced the classification accuracy and localization precision required for the detection task, thereby providing reliable technical support for the accurate monitoring and intelligent prevention and control of crop diseases and pests. The specific experimental results are illustrated in Figure 2.

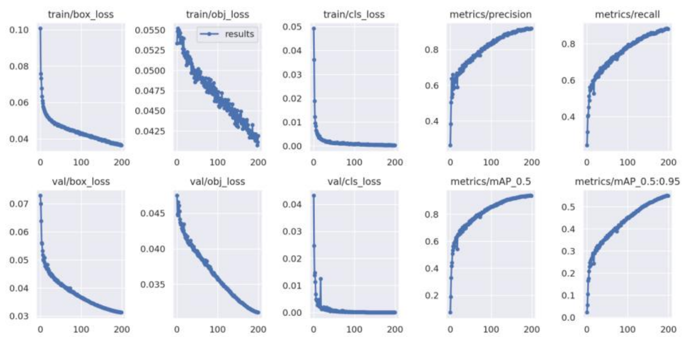

Figure 2 Training Performance Curves

From the perspective of detection localization performance, bounding box regression is effectively performed on pest and disease regions by the model in most samples. Both localized lesions and perforations on leaves and pest-infested parts of the whole plant are successfully framed, with high consistency between the bounding boxes and the actual contours of the targets. Even for samples with small and scattered pest and disease targets, such as subtle feeding traces or small pest colonies on leaves, accurate detection boxes are still output by the model without obvious missed detection or localization deviations, which fully demonstrates the model's perceptual capability for small targets and pests and diseases with complex morphological features. In terms of category recognition, multiple types of pests and diseases are clearly distinguished by annotations with different numbers, indicating that differentiated visual features of various pests and diseases have been effectively learned by the model during feature extraction, thereby enabling accurate classification of targets. Meanwhile, no large-scale false positive boxes or repeated detection on the same target regions are observed in the recognition results, which reveals strong robustness against interfering factors such as background clutter and leaf textures. The corresponding detection results are presented in Figure 3 for detailed visualization.

Figure 3 Detection Visualization Results

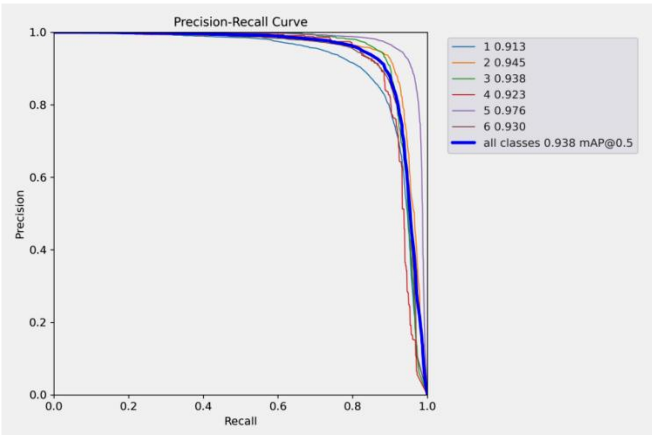

Figure 4 Precision-Recall Curve

The precision-recall curves presented in Figure 4 illustrate the detection performance of the proposed model across six categories of pests and diseases, with an overall mAP@0.5 of 0.938 achieved. For each category, high precision is maintained as recall increases, demonstrating superior detection performance. In terms of per-class performance, the highest average precision of 0.976 is obtained for Class 5, while average precision values of 0.945 and 0.938 are recorded for Class 2 and Class 3, respectively. Even Class 4 achieves an average precision of 0.923, indicating that reliable classification capability for all pest and disease categories is exhibited by the model. The overall curve for all classes shows a highly consistent trend with the individual class curves, maintaining high precision across most recall ranges and experiencing a sharp decline only when recall approaches 1.0. This characteristic suggests that an effective balance between precision and recall is achieved by the model, thereby verifying its excellent capability for pest and disease detection in practical agricultural scenarios.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Zhang Shugui, Chen Shuli, Zhao Zhan. Improved YOLOv8 algorithm for identifying crop leaf diseases and pests. Chinese Journal of Agricultural Mechanization, 2024, 45(07): 255-260. DOI: 10.13733/j.jcam.issn.2095-5553.2024.07.038.
[2] Lee J, Yin H, Jin D, et al. PestDetectSim: an integrated approach for crop pest diagnosis using object detection and similarity-based image retrieval. Plant methods, 2026. DOI: 10.1186/S13007-026-01520-X.
[3] Shi Jie, Xiong Kaixiang, Li Zhi, et al. A maize pest and disease detection system based on a lightweight improved YOLOv8 model and edge computing. Jiangsu Journal of Agricultural Sciences, 2025, 41(02): 313-322.
[4] Yu X, Zhang H, Duan Y, et al. An optimized convolutional neural network based on multi-strategy grey wolf optimizer to identify crop diseases and pests. Displays, 2026, 92: 103341. DOI: 10.1016/J.DISPLA.2026.103341.
[5] Xiong Xianhua, Yang Zheng, Cheng Siqin, et al. A method for ship image recognition and tracking in lock chambers based on improved YOLOv8 and Kalman filter algorithm. China Harbour Construction, 2025, 45(06): 17-24.
[6] Wang Yimeng, Jin Xiaoke. Research on Intelligent Detection of Crop Diseases and Pests and Precision Spraying by Agricultural Machinery Based on Image Processing. China Agricultural Machinery Equipment, 2026(03): 55-57.
[7] Gao Q, Shi C, Ji Y, et al. Crop Pest Identification and Real-Time Monitoring System Design Based on Improved YOLOv8s. Sensors, 2026, 26(2): 404. DOI: 10.3390/S26020404.
[8] Fang Yingying, Pan Yanhong, Pan Xinling, et al. Design of a deep learning-based crop disease and pest identification system. Smart Agriculture Guide, 2026, 6(03): 23-28. DOI: 10.20028/j.zhnydk.2026.03.006.
[9] Huang R, Li M, Zheng H, et al.Chinese crop diseases and pests named entity recognition based on variational information bottleneck and feature enhancement. Scientific Reports, 2025, 15(1): 31573. DOI: 10.1038/S41598-025-04252-5.