RESEARCH ARTICLE

Plant pest detection based on unmanned aerial vehicle remote sensing and YOLOv7 network

Qingsong Chang[1, 2], Xiaojing Yan[3, *], Danhui Zhu[4]

[1]Institute of Scientific and Technology Information, Jiangsu University, Zhenjiang, Jiangsu, China. [2]Network Information Center, [3]School of Agriculture and Forestry Science and Technology, [4]School of Traffic Engineering, Weifang Engineering Vocational College, Weifang, Shandong, China.

Received: February 6, 2026; **accepted:** May 26, 2026.

During crop growth, signs of pest infestation are commonly present but often subtle and hard to detect, which makes unmanned aerial vehicle (UAV) inspection a highly efficient approach. Nevertheless, the flight altitude of the UAV has a substantial influence on the accuracy of detection. This study introduced a height aware You Only Look Once (YOLO) v7 detection method that embedded flight altitude information into the network structure. By dynamically reweighting multi-scale features using altitude conditional vectors and adaptively modulating the channel dimension, the network retained fine-grained texture at low altitudes and enhanced its response to small targets at high altitudes. The results showed that the proposed approach achieved an accuracy of 95.36% with mean Average Precision at an Intersection over Union threshold of 0.5 (mAP@0.5), improving by 6.58%, 13.46%, 3.72%, and 1.59% compared to Faster Region-based Convolutional Neural Network (Faster R-CNN), Single Shot MultiBox Detector (SSD), YOLOv5s, and the original YOLOv7, respectively. In typical scenarios, the proposed method achieved a frame rate of 43.31 Frames Per Second (FPS). This research demonstrated that introducing imaging altitude priors into the detection network enhanced robustness against small-scale pests with minimal computational increase. The proposed method provided a novel approach for early warning of plant pests in large-scale farmland, achieving a harmonious balance between accuracy and real-time responsiveness, and was of significant importance for promoting precision pesticide application and green pest control.

Keywords: unmanned aerial vehicle; remote sensing; altitude perception; small target detection; pest identification; YOLOv7.

[*]**Corresponding author:** Xiaojing Yan, School of Agriculture and Forestry Science and Technology, Weifang Engineering Vocational College, Weifang, Shandong 262500, China. Email: yxj_93@163.com.

Introduction

The early detection and precise identification of plant pests are crucial for securing stable crop yields and advancing the progress of smart agriculture. Yet, in real-world farming conditions, pests typically exhibit characteristics such as scattered infestation areas, swift propagation, and considerable susceptibility to environmental influences. Consequently, traditional approaches that rely on manual inspections or fixed-point monitoring struggle to simultaneously fulfill the demands for promptness and extensive coverage across large-scale farmlands [1, 2]. In addition, during the mid-to-late growth stages of plants, the crop canopy structure is complex, the pest targets are small in scale, and they are highly similar to the background texture, further increasing the difficulty of manual identification and conventional visual methods [3]. With the development of unmanned aerial vehicle-remote sensing (UAV-RS) technology, high-resolution image acquisition based on low-altitude platforms has provided a new technical path for monitoring farmland pests. UAVs have advantages such as high mobility, wide coverage, and low deployment costs, and can acquire large-area crop growth information in a short time, providing continuous and dynamic observation data for pest detection [4, 5]. However, under the UAV perspective, insect pests usually occupy only a few pixels in the image and are affected by multiple factors such as flight altitude changes, uneven lighting, background interference, and viewpoint changes, which makes insect pests exhibit obvious small scale, low contrast, and irregular distribution characteristics in the image. In recent years, deep learning (DL) target detection methods have made great progress in natural scene recognition tasks [6]. Among them, the You Only Look Once (YOLO) series has a good balance between detection accuracy and real-time performance and has been widely used in agricultural RS and crop monitoring [7]. However, the existing YOLO network structure is mainly designed for general targets, and deep features are prone to loss of fine-grained insect pest information during multiple down sampling processes. In addition, texture interference in complex farmland backgrounds will weaken the model's ability to distinguish insect pest areas, thereby affecting detection stability and generalization performance. Therefore, while ensuring real-time detection, how to effectively improve the identification accuracy and detection stability of small-scale pest targets in complex backgrounds has become a key issue that urgently needs to be addressed.

With the rapid development of UAV-RS and DL technology, intelligent detection of plant pests has gradually become an important direction in agricultural informatization research. Researchers have conducted lots of studies on pest image acquisition, feature modeling, and automatic identification methods. Lee *et al*. introduced an intelligent diagnostic system based on crowdsourced data collection and DL. By collecting descriptions of pests and diseases from user groups and field experts, a crop-pest dataset was constructed, and a prediction model was trained using TensorFlow, thereby realizing rapid and reliable diagnosis of pests and diseases [8]. Shafik *et al*. proposed an enhanced model that integrated deep features of convolutional neural network (CNN) and majority voting of long short-term memory (LSTM) network. By extracting the features of the pre-trained network through fc6 layers and inputting them into the adaptive LSTM framework, the lightweight green control of apple trees was optimized [9]. Resti *et al*. introduced a classification scheme based on multinomial naive Bayes and K nearest neighbor algorithm using images to tackle the issue of low identification efficiency of maize diseases. The research employed self-made images of six types of pests and diseases, realizing accurate and rapid identification of maize diseases [10]. Sun *et al*. proposed an improved lightweight YOLOv7 model by introducing EfficientFormerV2 and convolutional block attention module (CBAM) to enhance small target features and accurate and rapid identify cotton pests [11]. Mehedi *et al*. introduced an improved real-time detection model for YOLOv7. It was trained on the PlantVillage dataset and introduced Shapley additive interpretation, thus realizing lightweight automatic monitoring of leaf diseases [12]. Zhao *et al*. proposed a Root-YOLOv7 hierarchical detection model. Multi-scale features were captured by hierarchical trunk and shift window self-attention, thus achieving efficient screening of resistant varieties of root-knot nematode disease [13]. Existing research on plant pest detection has made some progress in image acquisition methods, feature modeling strategies, and DL network design. One-stage detection models based on the YOLO series have shown good potential in balancing detection accuracy and real-time performance. However, existing methods still largely rely on single-scale feature representation or optimization for specific crops and scenarios, lacking adaptability to the scale changes of pest targets under multi-flight altitude conditions.

To address the aforementioned problems, this study aimed to improve the accuracy and stability of small-scale pest detection in UAV-RS scenarios. An improved YOLOv7 pest detection method based on an altitude-aware mechanism was proposed. The UAV flight altitude was encoded as a conditional vector, and an altitude-aware multi-scale fusion (MSF) mechanism was introduced. By dynamically adjusting the feature weights at different scales, adaptive modeling of pest target features at different imaging scales was achieved. Simultaneously, a channel modulation strategy enhanced the response of key features, enabling the model to strengthen small target detection capabilities under high-altitude conditions and retain fine-grained texture information under low-altitude conditions. This research introduced UAV imaging altitude information as a priori into the target detection network, providing a new modeling approach to address the problem of small target scale variations. The proposed method achieved an effective balance among detection accuracy, robustness, and computational efficiency while ensuring real-time performance. The research results provided technical support for large-scale intelligent monitoring of agricultural pests and were of great significance for promoting precision pesticide application and green agriculture development.

Materials and methods

UAV-RS pest data acquisition and dataset construction

Plant pest targets under UAV-RS conditions are usually small-scale, low-contrast, and easily affected by background textures in the images [14]. Therefore, to ensure that the collected data encompassed pest targets ranging from discernible scales down to minute scales, this study employed a multi-altitude acquisition strategy and, based on the pinhole imaging principle, established the relationship between ground resolution and flight altitude as shown below.

$$r = \frac{H \cdot s}{f} \tag{1}$$

where r was the ground spatial resolution. H was the UAV flight altitude. s was the physical size of a single pixel of the camera sensor. f was the camera focal length. When H increased, r also increased, and the size of the pest target within the pixel domain became smaller. The change in UAV flight altitude directly affected the imaging scale and resolution of pest targets in RS images. As the flight altitude increased from low to high, the ground coverage gradually expanded, while the proportion of pest targets in the pixel grid continuously decreased, resulting in a decrease in imaging resolution and more obvious small target characteristics of pest targets. Considering that pests were distributed in patches in the field, the study adopted a grid-based flight path and explicitly introduced forward overlap rate and lateral overlap rate constraints to ensure continuous image coverage and retain sufficient redundant information for subsequent screening [15, 16]. Let the coverage width and length of the k-th image under the ground projection be W_k and L_k, respectively, and the forward and lateral overlap rates be O_f and O_s, respectively. Then the newly added effective coverage area of the image could be approximated as follows.

$$A_k = \left(1 - O_f\right)\left(1 - O_s\right) \cdot W_k \cdot L_k \tag{2}$$

where A_k was the newly added effective coverage area of the k-th image relative to the already covered area. By adjusting O_f and O_s, a balance could be achieved between coverage integrity and redundancy controllability, thereby ensuring that the dataset contained sufficient effective pest samples while avoiding uncontrollable annotation costs due to a large number of duplicate frames. In the dataset generation stage, the study defined a set of samples that could be used to represent the integrity of the task area in terms of spatial coverage, the identifiability in terms of imaging quality, and the consistency in terms of annotation information. The original aerial frames were first screened for quality to remove images with obvious defocus, severe motion blur, overexposure /underexposure, *etc*. Then, bounding box annotation was performed using a unified standard to annotate the pest targets with category and location. A consistent processing rule was applied to areas that were suspected pests but did not meet the identifiability conditions. To enable the quantitative characterization of the target attributes of small pests and to facilitate subsequent network structure optimization, this study defined a mathematical expression for the target relative scale factor as follows.

$$S_i = \frac{w_i \cdot h_i}{W_I \cdot H_I} \tag{3}$$

where w_i and h_i were the width and height of the target bounding box in the pixel domain. W_I and H_I were the width and height of the original image. The smaller S_i was, the closer the target was to a small target shape. Considering the perspective disturbance, illumination change, and scale drift in the UAV operation process, this study introduced data augmentation without destroying the physical consistency of imaging to improve the robustness of the model to the actual scene [17]. The augmented training set could be formally represented as below.

$$D' = \left\{ (T_m(I_i), T_m(B_i) \middle| (I_i, B_i) \in D, m = 1, \ldots, M \right\} \tag{4}$$

where D was the original labeled dataset. I_i was the i-th image. B_i was its corresponding label. T_m(g) was the m-th enhancement transformation. M was the number of enhancement methods. D' was the enhanced training dataset.

Improved structural design of pest detection network based on YOLOv7

In the UAV-RS plant pest detection scenario, the pixel scale of pest targets in the image changed significantly with flight altitude, and their discriminative features were more easily weakened during network down sampling. However, the original YOLOv7 network assumed that samples under different altitude conditions followed a uniform scale distribution in the feature extraction and fusion stages, and the feature modeling mainly relied on the image data itself, without explicitly utilizing the imaging scale information contained in the flight altitude [18, 19]. When faced pest samples with large scale differences, the fixed feature response mechanism was difficult to adapt to the characteristics of low-altitude fine-grained features and high-altitude small targets. Therefore, this study proposed to improve and optimize the YOLOv7 plant pest detection network. To enable the network to perceive the imaging scale change of pest targets, the UAV flight altitude was defined as conditional information. Let the fused feature be represented as $F \in R^{C \times Q \times W}$, where C, Q, and W were the number of channels, height, and width of the feature map. To enable the network to perceive the imaging scale change, the UAV flight altitude was encoded as a height conditional vector h and mathematically expressed below.

$$h = \phi(H) \tag{5}$$

where $\phi(H)$ was the height encoding function, used to map physical height into a conditional representation that the network could learn. In the feature fusion stage, to characterize the differences in importance of multi-scale features under different height conditions, a height-aware MSF mechanism was introduced. Let the feature set from different scale layers be $\{F_l\}_{l=1}^{L}$, where $F_l \in R^{C_l \times Q_l \times W_l}$ was the feature map of the l-th layer. Scale weights were generated through the height conditional vector h. Adaptive reweighting was performed on the features at each scale. The fusion form was defined as follows.

$$F_h = \sum_{l=1}^{L} \omega_l(h) \cdot R_l(F_l) \tag{6}$$

where $R_l(F_l)$ was the scale alignment operation, used to map features at different levels to a unified spatial resolution. $\omega_l(h)$ was the scale weights generated by the altitude condition vector, used to reflect the importance of each scale feature for pest detection at the current flight altitude. After obtaining the altitude-aware fusion features F_h, the feature channels were further adaptively adjusted through a conditional modulation mechanism as follows.

$$F^{'} = \gamma(h) \odot F_h + \beta(h) \tag{7}$$

where F' was the feature representation after height-sensing modulation. $\gamma(h)$ and $\beta(h)$ were the channel scaling coefficient and bias term generated by the height condition vector, respectively. $\odot$ was the channel-by-channel multiplication operation. Altitude-sensing feature modulation enabled the network to enhance its response to small-scale pest features from a high-altitude perspective and retain more fine-grained information under low-altitude conditions, thereby achieving adaptive modeling of pest target scale changes. Flight altitude information was first used to generate a conditional vector through embedding mapping, and then further transformed into feature modulation parameters, which were applied to MSF features to achieve adaptive reweighting of features at different scales. Unlike traditional attention mechanisms that relied solely on feature statistics, this method explicitly introduced UAV-RS imaging prior, giving the feature reweighting process clear scene orientation. During the detection phase, the altitude-sensing modulated feature F' was input into the detection head for target classification and position regression. The output was shown below.

$$\begin{cases} \hat{y} = f_{cls}(F') \\ \hat{b} = f_{reg}(F') \end{cases} \tag{8}$$

where $f_{cls}(F')$ and $f_{reg}(F')$ were the classification branch and regression branch functions, respectively. $\hat{y}$ was the pest target category prediction result. $\hat{b}$ was the corresponding bounding box regression result. Since the features had undergone adaptive modulation under height conditions before entering the detection head, the detection head could obtain a more stable input distribution under different imaging scale conditions, thereby reducing the risk of missed detection and false detection of small pest targets. This study embedded the height-aware MSF and conditional modulation module into the feature fusion stage of YOLOv7 to form a height-aware YOLOv7 (H-YOLOv7) pest detection network for UAV-RS scenarios. Within this framework, the input UAV-RS image first extracted basic features through the YOLOv7 backbone network and encoded the scale conditions corresponding to the flight altitude. Based on this, it was introduced as prior information into the altitude-aware feature modulation module for adaptive adjustment of the features. The joint feature map after altitude-aware modulation was further fed into the YOLOv7 detection head to achieve accurate localization and identification of pest targets.

Experimental setup and evaluation indicators

To confirm the validity and stability of the proposed H-YOLOv7 plant pest detection method in UAV-RS scenarios, a self-built UAV-RS pest RS dataset was used, supplemented by the publicly available pest image dataset IP102 (https://github.com/xpwu95/IP102). The model's detection accuracy, robustness, and generalization ability were systematically evaluated. All experiments were conducted in a unified hardware and software environment. The model training and inference processes were implemented using the PyTorch deep learning framework (https://pytorch.org/). The experimental platform used an Intel Core i7-12700K processor, an NVIDIA RTX 3090 graphics card, and 32 GB of memory. The operating system was Windows 11, and the software environment included Python 3.10 and PyTorch 2.0. The self-built UAV-RS pest dataset was acquired through UAV-RS data collection in actual farmland environments, covering the peak pest season in the mid-to-late stages of crop growth. By setting different flight altitudes, shooting angles, and flight path overlap rates, RS images of pests with significant scale differences were collected. The dataset contains 3,260 high-resolution remote sensing images, covering six common crop pest targets, mainly including aphids, leaf miners, caterpillars, planthoppers, beetles, and whiteflies. After meticulous manual annotation, a total of 18,742 pest target instances were obtained. During data processing, images with severe defocusing, significant motion blur, or no valid pest targets were removed, retaining only samples with clear pest characteristics. The IP102 dataset contained various crop pest types with complex imaging conditions and diverse backgrounds, posing a high detection challenge. A subset of IP102 images with pest categories similar to those in the UAV-RS scenarios was selected, totaling 4,500 images and corresponding to 22,000 labeled pest instances. These images were then uniformly converted to the same annotation format and input resolution as the self-built UAV-RS dataset. To systematically evaluate the performance of the H-YOLOv7 model in UAV remote sensing pest detection, performance comparisons were conducted using Faster Region-based Convolutional Neural Network (Faster R-CNN) (https://github.com/rbgirshick/py-faster-rcnn), Single Shot MultiBox Detector (SSD) (https://github.com/weiliu89/caffe/tree/ssd), YOLOv5s (https://github.com/ultralytics/yolov5), and the original YOLOv7 model (https://github.com/WongKinYiu/yolov7). To comprehensively measure model performance, precision, recall, mean average precision (mAP@0.5), mean average precision across multiple thresholds (mAP@0.5:0.95), and F1 score were determined. Among them, precision represented the proportion of true positives among samples predicted as positive by the model, while recall represented the proportion of true targets correctly detected. mAP@0.5 was used to evaluate the detection performance under the condition of an Intersection over Union (IoU) threshold of 0.5, while mAP@0.5:0.95 was calculated by averaging the detection performance under multiple IoU thresholds in the range of 0.5 to 0.95, thus more comprehensively reflecting the model's performance under different positioning accuracy requirements. The F1 value was the harmonic mean of precision and recall, used to comprehensively evaluate the model's ability to balance detection accuracy and completeness. Furthermore, to intuitively evaluate the detection performance of different methods in complex scenarios, this study selected representative remote sensing image samples of insect pests for visual comparative analysis. Under unified confidence threshold and non-maximum suppression (NMS) parameter settings, the detection results of each comparative method were visualized, and comparative analysis was conducted on aspects such as small target detection capability, dense target recognition effect, and adaptability to complex backgrounds.

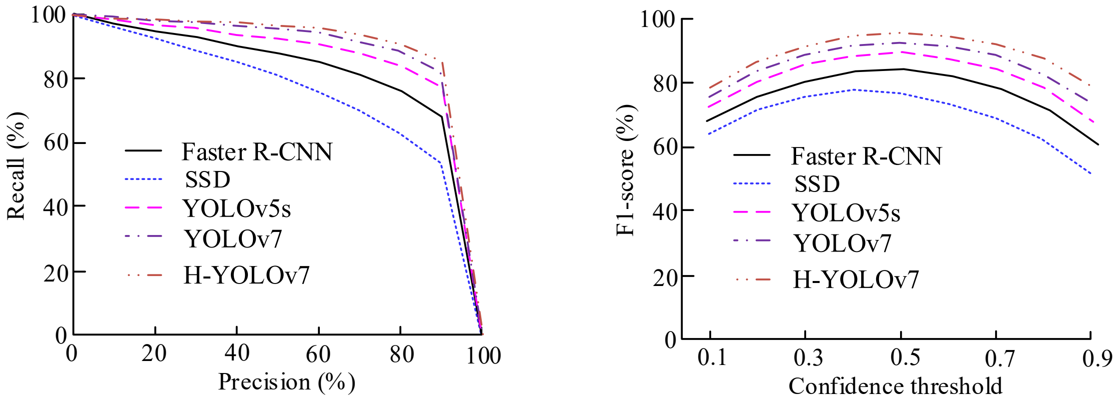
Results and discussion

Comparative analysis of pest detection performance of different methods

The performance of different methods on the pest detection task showed that the accuracy of H-YOLOv7 was 4.95%, 11.63%, 3.13%, and 1.59% higher than that of Faster R-CNN, SSD, YOLOv5s, and YOLOv7, respectively. The results indicated that H-YOLOv7 had a significant advantage in detecting small insect targets and under strict localization conditions. Furthermore, the mAP@0.5 and mAP@0.5:0.95 of H-YOLOv7 were also significantly better than the other four methods. Specifically, H-YOLOv7 's mAP@0.5 was 6.58%, 13.46%, 3.72%, and 1.59% higher than the other four methods, and its mAP@0.5:0.95 was 14.63%, 29.66%, 8.72%, and 4.20% higher than the other four methods, respectively (Table 1). The results demonstrated that H-YOLOv7 had superior target detection capabilities and overall detection level and exhibited better detection stability and accuracy under strict localization conditions. The precision-recall (PR) curves and F1-score curves as a function of confidence thresholds for various approaches showed that the proposed H-YOLOv7 had the largest overall envelope area of its PR curve, exhibiting stronger false detection suppression capability under high

Table 1. Performance comparison of different methods.

Method	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Faster R-CNN	90.86	84.37	88.92	56.74
SSD	85.42	78.63	83.51	50.16
YOLOv5s	92.48	88.19	91.36	59.82
YOLOv7	94.12	90.27	93.28	62.41
H-YOLOv7	95.36	92.18	94.76	65.03



(a) Comparison of PR curves of different methods

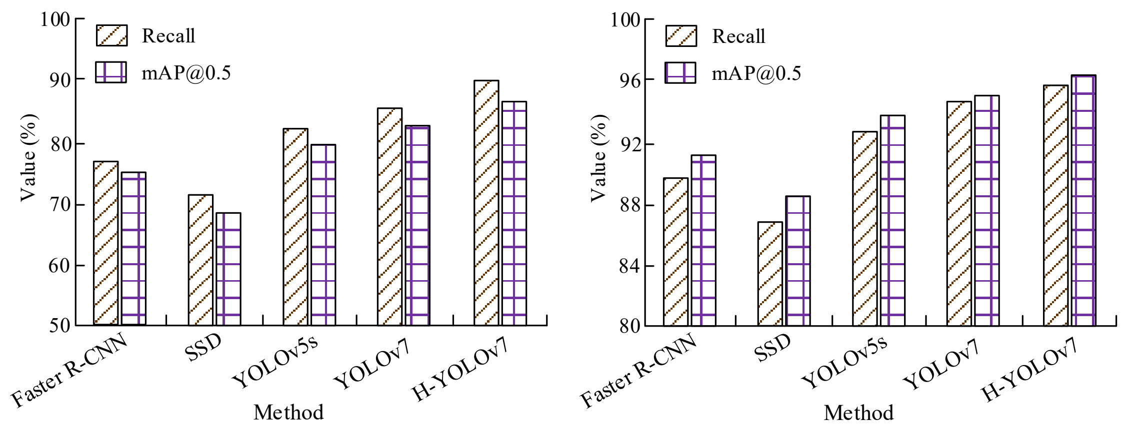
(b) F1-score as a function of confidence threshold curve data

Figure 1. Comparison of performance curves of various approaches.

recall conditions (Figure 1a). The F1-score of various methods as a function of confidence threshold showed that H-YOLOv7 achieved the highest peak F1-score within the same threshold range, which indicated that the highly perceptive feature modulation mechanism could effectively improve the overall balance between precision and recall of the model (Figure 1b). The detection results of different methods on insect pest targets at different scales demonstrated that H-YOLOv7 exhibited significant performance advantages under different scale conditions. The small-scale insect pest targets showed that the recall of H-YOLOv7 was 89.74%. In terms of overall detection performance, the mAP@0.5 of H-YOLOv7 reached 86.47%. The medium-scale insect pest target detection scenario showed that the mAP@0.5 of H-YOLOv7 was 96.38%, which indicated that H-YOLOv7 had good adaptability under different target scales (Figure 2). The study further selected representative remote sensing image samples of pests and conducted a visual comparative analysis of the prediction results of different detection methods. The results showed that there were significant differences in the pest detection effects of different detection methods in the same scene. H-YOLOv7 could detect more small-scale pest targets with a more continuous distribution of detection boxes and more complete target localization, maintaining good detection consistency even in complex backgrounds and densely populated target areas. Overall, H-YOLOv7 demonstrated stronger detection robustness in real-world UAV-RS scenarios.

Verification of detection performance and robustness evaluation in complex scenes

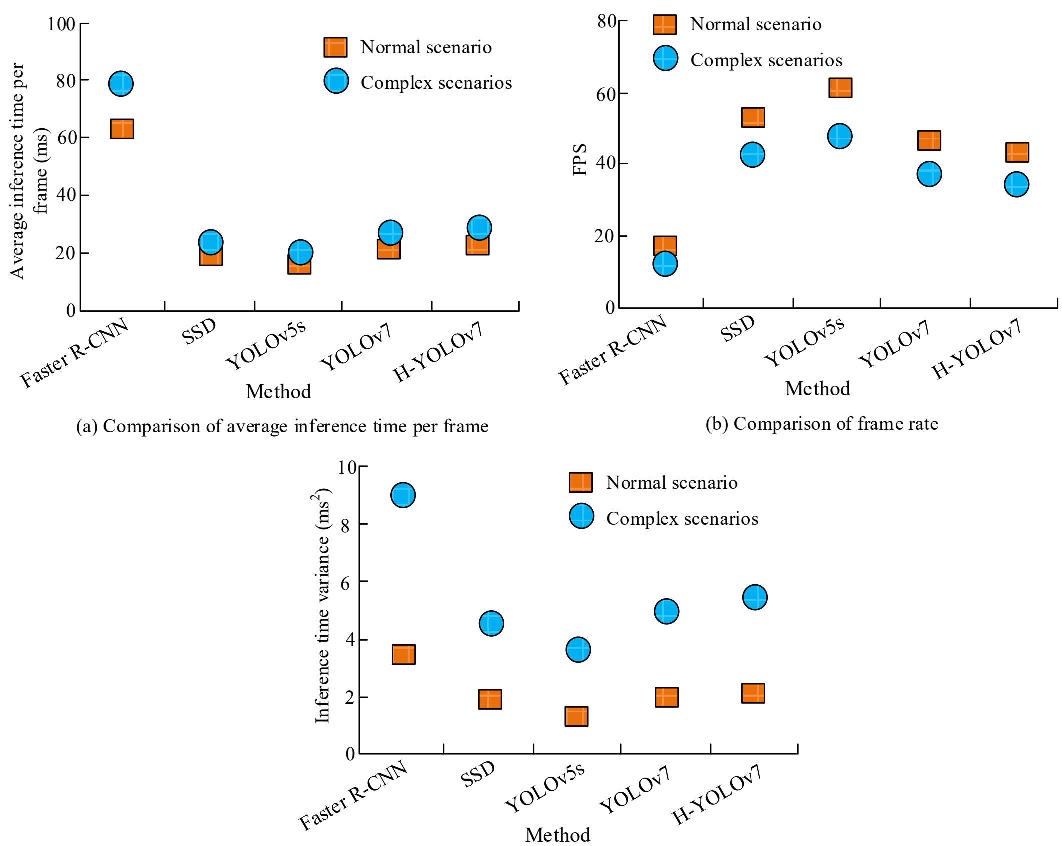
The results of average inference time per frame, corresponding frames per second (FPS), and inference time variance of the H-YOLOv7 method in normal and complex scenarios demonstrated that, in normal scenarios, the single-frame inference time of H-YOLOv7 reached 23.08 ms, indicating that the introduction of the high-sensitivity modulation module brought limited



(a) Performance comparison of small-scale insect pest targets

(b) Comparison of performance of medium-scale pest targets

Figure 2. Validation of pest target detection at different scales.



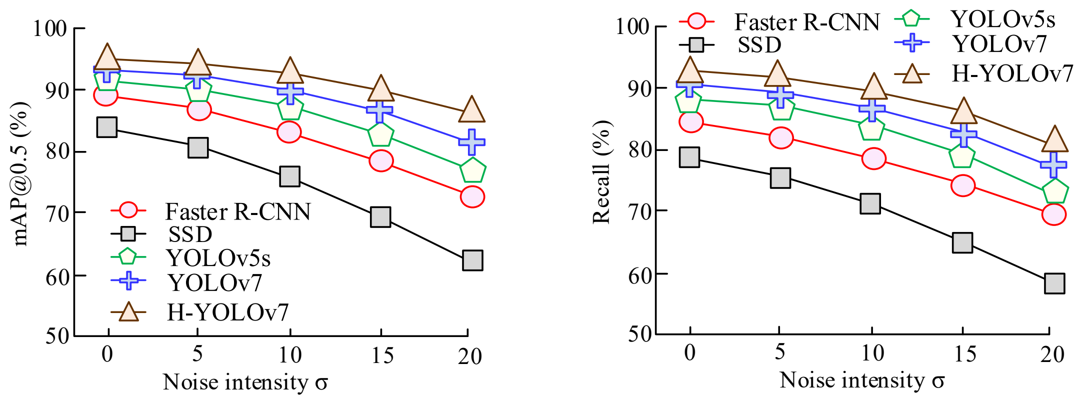
(a) Comparison of average inference time per frame

(b) Comparison of frame rate

(c) Comparison of Inference time variance

Figure 3. Comparison of inference time and real-time performance in different scenarios.

additional computational overhead. Under complex scenario conditions, the inference time of each method increased to varying degrees with H-YOLOv7's inference time increasing by 28.91 ms, and the overall increase remained within an acceptable range (Figure 3a). Further,



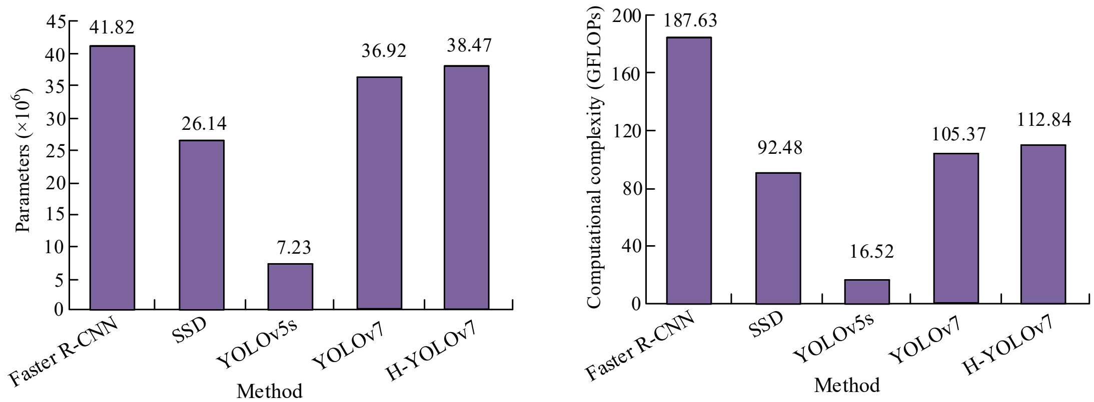
(a) Comparison of mAP@0.5 of different methods

(b) Comparison of recall of different methods

Figure 4. Performance comparison under different noise disturbance intensities.

the real-time performance of different methods from the perspective of frame rate reflected that, after entering complex scenarios, the frame rate of each method decreased, but H-YOLOv7 still maintained 34.58 FPS, which was higher than the 30 FPS threshold usually required for real-time applications (Figure 3b). Although H-YOLOv7 introduced an additional high-sensitivity feature modulation mechanism on the basis of the original YOLOv7, its inference time only increased slightly, and the frame rate remained within the real-time detection range, which confirmed that the proposed method could balance the improvement of detection performance and computational efficiency in complex UAV-RS scenarios and had the feasibility of practical engineering applications (Figure 3c). The performance detection results under different noise perturbation intensities σ showed that, under high noise conditions (σ = 20), H-YOLOv7 exhibited the smallest performance degradation, indicating stronger detection stability under noise interference (Figure 4a). H-YOLOv7 maintained relatively high recall levels under σ = 15 and σ = 20 conditions (Figure 4b). Overall, the performance degradation of H-YOLOv7 under different noise intensities was significantly smaller than that of the comparative methods with a clear advantage under high noise conditions, which indicated that, after introducing a highly perceptive multi-scale feature modulation mechanism, the model had stronger resistance to feature degradation caused by noise perturbations. The model complexity and resource overhead of different methods presented that the H-YOLOv7 model had 38.47 × 10^5 parameters, which was only 4.20% more than YOLOv7, while the computational complexity increased by 7.09%. Compared with Faster R-CNN, H-YOLOv7 reduced the number of parameters by 8.02% and the computational complexity by 39.87%. Compared with SSD, H-YOLOv7 increased the number of parameters and computational complexity by 47.17% and 22.02%, respectively (Figure 5). Overall, H-YOLOv7 achieved a good balance between enhanced detection performance and increased engineering resources, making it feasible for practical deployment on UAV platforms.

The results of this study showed that the proposed H-YOLOv7 model exhibited excellent overall performance in UAV remote sensing pest detection. In target detection scenarios at different scales, the proposed method outperformed Faster R-CNN, SSD, YOLOv5s, and the original YOLOv7 in terms of mAP@0.5 and mAP@0.5:0.95, demonstrating its significant advantages in small target recognition and complex background adaptation. Furthermore, the proposed model maintained high detection stability and real-time performance even under complex environmental and noise interference



(a) Number of model parameters (b) Computational complexity

Figure 5. Comparison of model complexity and resource overhead.

conditions, showcasing good robustness and engineering application potential. The proposed method achieved an effective balance between detection accuracy, stability, and computational efficiency, providing technical support for large-scale intelligent monitoring of agricultural pests.

References

1. Gao R, Liu L, Fan S, Zheng W, Liu R, Zhang Z, *et al*. 2024. Occurrence and potential diffusion of pine wilt disease mediated by insect vectors in China under climate change. Pest Manag Sci. 80(12):6068-6081.
2. Hou S, Rodrigues O, Liu Z, Shan L, He P. 2024. Small holes, big impact: Stomata in plant-pathogen-climate epic trifecta. Mol Plant. 17(1):26-49.
3. Weng H, Yao Y, Ye D, Wu R. 2024. Rapid evaluation of rice neck blast resistance using low altitude remote sensing of UAV combined with YOLOv7. Trans Chin Soc Agric Eng. 40(21):110-118.
4. Hafeez A, Husain MA, Singh SP, Chauhan A, Khan MT, Kumar N, *et al*. 2023. Implementation of drone technology for farm monitoring and pesticide spraying: A review. Inf Process Agric. 10(2):192-203.
5. Bagheri E, Khodkam H. 2025. Smart farming: How drones are transforming the future of food production. Biosyst Eng Sustain Technol. 1(1):49-67.
6. Altaf A, Tripathy RK. 2025. An IoT-enabled deep learning approach implemented on Android device for automated identification of breast cancer using thermal images. Smart Wearable Technol. 1(3):1-7.
7. Li C, Wang M. 2024. Pest and disease management in agricultural production with artificial intelligence: Innovative applications and development trends. Adv Resour Res. 4(3):381-401.
8. Lee HY, Kim DH, Park KR. 2023. Pest diagnosis system based on deep learning using collective intelligence. Int J Electr Eng Educ. 60(2 Suppl):5-19.
9. Shafik W, Tufail A, Liyanage CDS, Apong RAAHM. 2023. Using a novel convolutional neural network for plant pests detection and disease classification. J Sci Food Agric. 103(12):5849-5861.
10. Resti Y, Irsan C, Putri MT, Yani I, Ansyori A, Suprihatin B. 2022. Identification of corn plant diseases and pests based on digital images using multinomial naive bayes and k-nearest neighbor. Sci Technol Indones. 7(1):29-35.
11. Sun J, Jia Y, Wu Z, Zhou X, Shen J, Wu X. 2024. Detecting pests in cotton fields using improved YOLOv7. Trans Chin Soc Agric Eng. 40(10):176-184.
12. Mehedi MHK, Nawer N, Ahmed S, Khan MSI, Hasib KM, Mridha MF, *et al*. 2024. PLD-Det: Plant leaf disease detection in real time using an end-to-end neural network approach based on improved YOLOv7. Neural Comput Appl. 36(34):21885-21898.
13. Zhao Y, Zhao H, Xiong H, Zhang F, Lu C, Li J. 2025. Root-YOLOv7: Multi-scale adaptive object detection and grading of root-knot nematode disease. Int J Agric Biol Eng. 18(2):259-268.
14. Manafifard M, Huang J. 2025. A comprehensive review on wheat yield prediction based on remote sensing. Multimed Tools Appl. 84(19):20843-20916.
15. Wu X, Lei Y, Tong X, Zhang Y, Li H, Qiu C, *et al*. 2022. A non-rigid hierarchical discrete grid structure and its application to UAVs conflict detection and path planning. IEEE Trans Aerosp Electron Syst. 58(6):5393-5411.
16. Ahmed G, Sheltami T, Mahmoud A. 2024. Energy-efficient multi-UAV multi-region coverage path planning approach. Arab J Sci Eng. 49(9):13185-13202.
17. Goceri E. 2023. Medical image data augmentation: Techniques, comparisons and interpretations. Artif Intell Rev. 56(11):12561-12605.
18. Ji Y, Di L. 2024. Textile defect detection based on multi-proportion spatial attention mechanism and channel memory feature fusion network. IET Image Process. 18(2):412-427.
19. Chen L, Wu L, Wu Y. 2025. Maturity detection of *Hemerocallis citrina* Baroni based on LTCB YOLO and lightweight and efficient layer aggregation network. Int J Agric Biol Eng. 18(2):278-287.