DETECTION OF ADULT PEACH FRUIT MOTH BASED ON IMPROVED YOLOv8m / 基于改进 YOLOV8M 的桃小食心虫成虫检测方法

Lijun CHENG*[1], Yihe ZHANG[1], Jianglin YAN[1], Zhengkun ZHAI[1], Zhiguo ZHAO*[2], LinQiang DENG[1]
[1] College of Software, Shanxi Agricultural University, Taigu, Shanxi / China;
[2] College of Software, Shanxi Agricultural University, Taigu, Shanxi / China
Corresponding authors: Lijun Cheng; Tel: +86-13835441585; E-mail: cljzyb@sxau.edu.cn
Zhiguo Zhao, Tel: +86-15034664518; E-mail: nice2me@126.com
DOI: https://doi.org/10.35633/inmateh-75-05

***Keywords:** YOLOv8m; pest detection; down sampling; CPCA attention mechanism; Inner-WIoU*

ABSTRACT

The peach fruit moth was a fruit-eating pest and one of the major pests of fruit trees in China, Korea, Japan, and Australia. Due to long-term problems such as improper control methods, low technical quality, and untimely treatment, the yield and efficiency of fruit products were greatly affected, which constrained the development of the fruit industry. This paper developed a method for detecting adult peach fruit moths based on an improved YOLOv8m to address the challenging problem of manually detecting peach fruit moths. To increase the Receptive Field of the model, v7Down Sampling was introduced in its backbone network. Then, the channel-prioritized Convolutional Attention Mechanism Module (CPCA), which dynamically allocated the spatial attention weights on each channel, reducing the noise and the algorithm's complexity, was incorporated. Finally, the inner-WIoU loss function was introduced to enhance the convergence and generalization of the bounding box. The precision (P) of the improved model increased by 3.4 percentage points compared to YOLOv8m. The recall (R) improved by 2.1 percentage points, and the mAP improved by 1.2 percentage points. The single-category precision (AP) for peach fruit moth detection improved by 2.4 percentage points. Moreover, the weight size, number of model parameters, and computational volume were reduced by 3.6MB, 1.8M, and 1.7G, respectively. This achieved an improvement in the model's effectiveness in detecting adult peach fruit moths without increasing the model's complexity. The results provided strong technical support for the subsequent real-time monitoring of the peach fruit moth.

摘要

桃小食心虫是一种食果害虫，也是中国、韩国、日本、澳大利亚等果树的主要害虫之一。因为长期的防治方法不当、技术素质低、处理不及时等问题，使果品的产量和效益都受到很大影响，制约果业的发展。本文针对人工检测桃小食心虫困难问题，开发了一种基于改进 YOLOv8m 的桃小食心虫成虫检测方法。我们在其主干网络引入 v7Down Sampling，增加模型的感受野。然后引入通道优先卷积注意力机制模块（CPCA），动态分配各个通道上的空间注意力权重，减少了噪声及算法的复杂度。最后引入 Inner-WIoU 损失函数，增强了边界框的收敛和泛化能力。改进后模型的精确度 P 相较 YOLOv8m 提高了 3.4 个百分点。召回率 R 提高了 2.1 个百分点。mAP 提高了 1.2 个百分点。桃小食心虫单类别精度 AP 上提高了 2.4 个百分点。并且权重大小、模型参数量和计算量分别减少了 3.6MB、1.8M、1.7G。实现了在不增加模型复杂度的同时提高模型对桃小食心虫成虫的检测效果，其结果可为后续桃小食心虫的实时监测提供有力的技术支撑。

INTRODUCTION

The peach fruit moth (*Carposina sasakii* Matsumura), is a fruit-eating pest of the genus Peach Fruit Moth in the family Lepidoptera. It is widely distributed and has a serious infestation in China, Japan, Korea, Russia, Australia, and other countries (*Kim et al., 2000*). The peach fruit moth had a wide range of host plants, with the Rosaceae, Rhamnaceae, and Pomegranateaceae being the primary ones. When the damage was severe, it caused worm droppings and rotten fruits, which directly affected the yield and quality of fruits and resulted in significant economic losses (*Fang et al., 2022*). Prediction could prevent the occurrence of the pest in advance based on the dynamic patterns of its emergence. At present, the detection and counting of adult peach fruit moths relied on visual discrimination. From 2021 to 2022, Zhang Xiaowei monitored the population size of adult pear fruit moths and peach fruit moths in pear orchards in Shanxi Province. This study was conducted manually for two consecutive years, using pear orchards in Shanxi Province as monitoring sites. It provided a theoretical basis for the prediction and integrated management of these two types of pear fruit moths (*Zhang et al., 2024*).

This manual method of detecting and counting peach fruit moths was time-consuming and inefficient. Moreover, the number of adult peach fruit moths was high during the peak incidence period, and accurate counting was challenging due to the wide range of infested fruit species (*Zhang et al., 2023*). These difficulties seriously hindered the accurate control of peach fruit moths.

In recent years, the level of agricultural intelligence gradually increased, and deep learning detection algorithms were applied to the field of agricultural pest detection (*Yue et al., 2024*). *Min Dai et al.,* (*2023*), proposed an improved plant pest detection method based on YOLOv5m. Experimental results showed that the improved YOLOv5m achieved 95.7% accuracy, a 93.1% recall rate, a 94.38% F1 score, and a 96.4% mean average precision (mAP). Li Bin proposed an improved YOLOv5 rice pest detection method, and the mAP value of the improved model was 1.49 percentage points higher than that of YOLOv7 and 12.89 percentage points higher than that of Faster R-CNN (*Li et al., 2024*). *Ma Pan et al., (2023),* proposed a cotton aphid image detection algorithm based on the YOLO neural network and integrated it into software. The average time for cotton aphid image detection was 4.1 seconds, and the counting accuracy for both live and dead cotton aphids exceeded 93%. The use of deep learning algorithms could enable the rapid detection of peach fruit moths, significantly improving detection efficiency. Therefore, this study aimed to use a YOLO series of algorithms to address the problems of difficult target recognition and low detection accuracy caused by background clutter and overly dense samples in the peach fruit moth dataset. The base model was optimized and improved to provide a model basis for the detection of peach fruit moths.

MATERIALS AND METHODS

Data set establishment

The adult peach fruit moth dataset used in this study was obtained from the Institute of Fruit Tree Research, Shanxi Agricultural University, Taigu District, Jinzhong City, Shanxi Province, China. In this study, triangular traps were hung on fruit trees 1.2-1.5 m above the ground in pear and apple orchards from May to October 2023 -2024. The traps consisted of a triangular trap frame, white sticky boards, and peach fruit moth traps (*Zhao et al., 2023*). 2-3 days later, the white sticky boards were removed and images of the dataset were captured with a Nikon D7000 camera from different angles. A total of 1781 images in JPG format were taken in this study, as shown in Fig. 1.

The 1781 original data were randomly divided into a training set (1068), a validation set (356), and a test set (357) in the ratio of 6:2:2 to ensure that the training set and validation set were completely independent and not duplicated. To improve the generalization ability and robustness of the model during the training process, this study first uses Open Computer Vision Library (OpenCV) and randomly selects Gaussian fuzzy, adding noise, rotation, brightness change, flipping panning, and other methods for data enhancement of the divided training set. A dataset containing 4982 images was finally obtained, of which 4269 were for the training set, 356 for the validation set, and 357 for the test set. In the process of data collection, it was found that in addition to peach small heartworms, many pear small heartworms and Chinese pear louse also appeared on the white sticky board. Therefore, in this experiment, adult pear small heartworm and Chinese pear louse were added to the data categories, and the data set was labelled with the software LabelImg. The labelling was divided into "Carposina lipogenesis", "Psylla chinensis" and "Oriental Fruit Moth".

Fig. 1 -Collection environment and collection device for peach fruit moth datasets

Improved YOLOv8m target detection algorithm

The improved YOLOv8m model structure in this paper is shown in Fig. 2. Firstly, the v7Down Sampling (v7DS) from YOLOv7 was added to the backbone network of YOLOv8. Then, the Channel-Prior Convolutional Attention Mechanism Module (CPCABlock) was incorporated into the backbone section. Finally, the Inner-WIoU loss function was introduced.

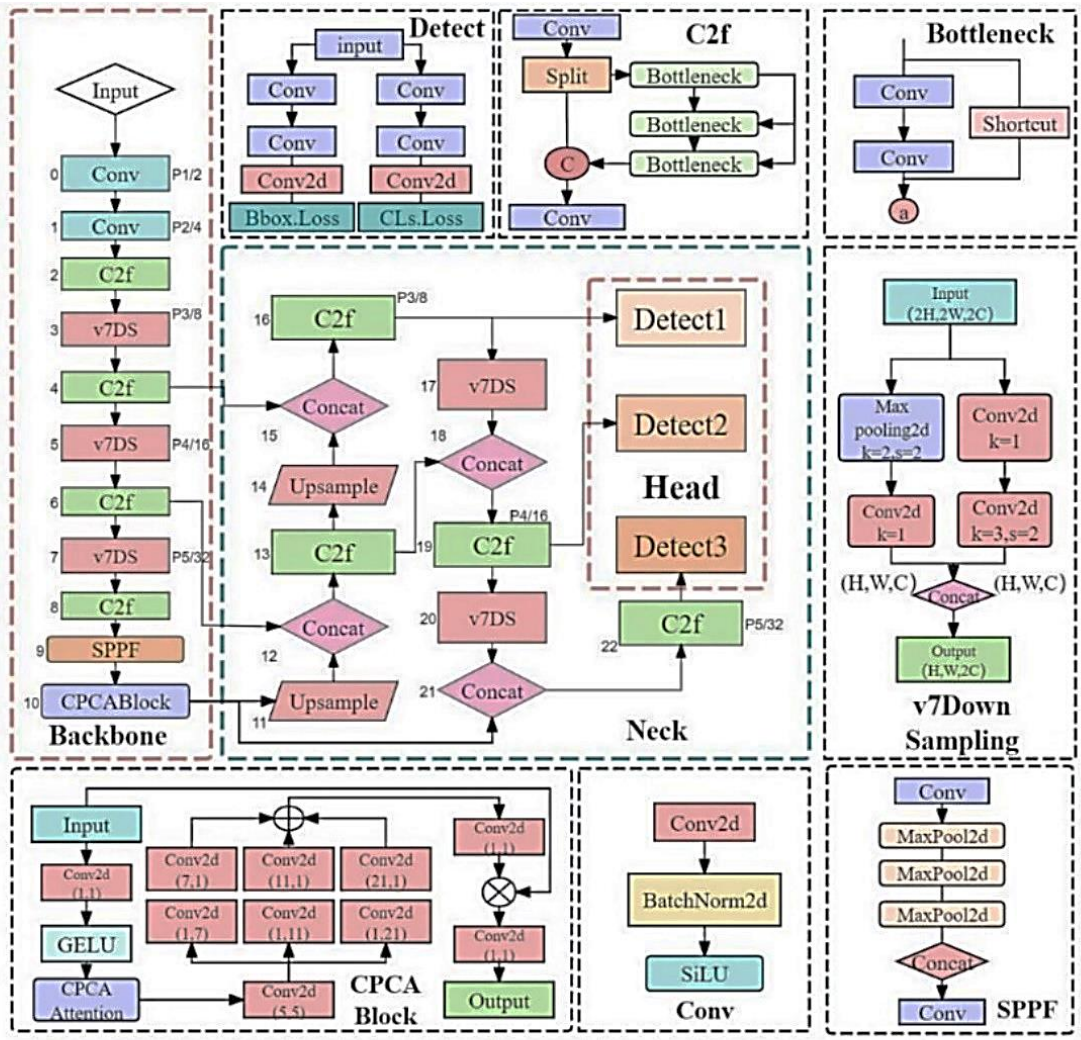


Fig. 2 - Improved YOLOv8m model structure diagram

Conv is convolution, V7DownSamplimg is the YOLOv7 downsampling module, SPPF is the spatial pyramid pooling module, CPCABlock is the channel-first convolutional attention mechanism module, Contact is the feature connection module, Upsample is upsampling module, Detect is detection head, Bbox. Loss and Cls. Loss is bounding box loss and classification loss, Split is the slice operation, Bottleneck is the bottleneck layer, SiLu is the activation function, MaxPooling is the maximum pooling operation, Maxpool2d is the maximum pooling, Conv is the convolution, Contact is the feature connection module and Bbox. Loss and Cls. The Losses are bounding box losses and classification losses, respectively.

Downsampling Module v7DS

Downsampling is a commonly used image processing method in image detection. It can shrink the image without altering its effective content, reduce the image resolution, decrease the amount of data computation, and generate a downsampling map corresponding to the image. This helps extract high-level semantic features from the image and increases the model's Receptive Field, thereby enhancing the model's ability to detect targets (*Lin et al., 2023*).

The downsampling v7DS of YOLOv7 utilizes the Spatial Pyramid Pooling (SPP) structure. It produces a feature map with the number of output channels equal to the number of input channels and reduces the spatial resolution by a factor of two. The structure is shown in Figure 3 below. By reducing the size and resolution of the input image and the computational load of the model, the network can better capture features in the image, process detailed information of the input image, and handle large image data more efficiently (*Wang et al., 2022*).

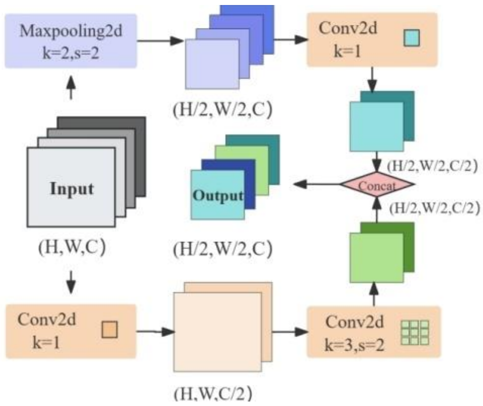


Fig. 3 -v7DS model structure diagram

Channel Prioritized Convolutional Attention Mechanism Module

The peach fruit moth dataset was collected in an open-air environment and was characterized by a cluttered background and fuzzy samples, which required the model to have the ability to judge target objects more accurately. The attention mechanism can suppress the expressive ability of non-essential features on the feature map and enhance the expressive ability of main features, effectively weakening the influence of the cluttered background on the detection results during the detection process.

The Channel-Prior Convolutional Attention Mechanism (CPCA), proposed by Hejun Huang et al. (*Huang et al., 2024*), combines channel attention and spatial attention. It reduces noise and algorithm complexity while achieving dynamic allocation of spatial attention weights on each channel. CPCA is a lightweight yet high-performance attention mechanism that can alleviate computational burden. The model structure is shown in Figure 4 below.

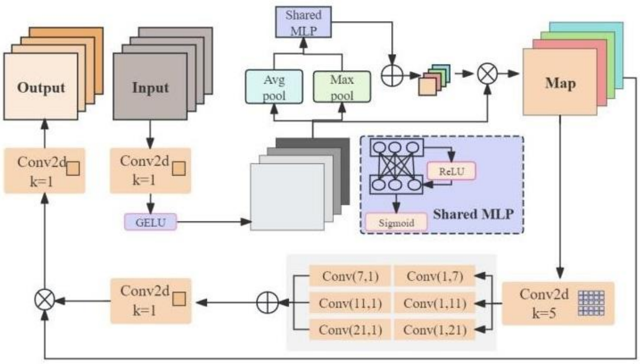


Fig. 4 -CPCABlock module structure diagram

Channel Attention Module of CPCA: The input feature maps are subjected to global average pooling and global maximum pooling respectively to obtain new feature maps, then the weight coefficients are obtained through the σ function, and the weight coefficients are multiplied with the new feature maps to finally obtain the output feature maps, which is calculated as in equation (1).

$$\mathrm{CA}(F)=\sigma\left(MLP\left(AvgPool(\mathrm{Fm})\right)+MLP\left(MaxPool(Fm)\right)\right) \tag{1}$$

where: σ is the Sigmoid function, $AvgPool$ is the global average pooling, and $MaxPool$ is the global maximum pooling (*Wang et al., 2024*).

Spatial Attention Module for CPCA: Utilizing depth-separable convolution to capture spatial relationships between features, a multi-scale structure is used to enhance the ability of the convolution operation to capture spatial relationships, which is calculated as in equation (2).

$$\mathrm{SA}(F)=Conv_{1\times 1}\left(\sum_{i=0}^{3} Branch_i(DwConv(F))\right) \tag{2}$$

where: $DwConv$ denotes the depth convolution. Branch_i, $i\in\{0,1,2,3\}$ denotes the i-th branch. Branch_0 is the identity connection.

Inner-WIoU loss function

The peach fruit moth dataset has problems such as high sample density and different target scales, $Inner$-$WIoU$ implements the similarity calculation by considering the scale difference between the auxiliary border and the actual border, which is suitable for the situation that needs to adjust the loss of focus dynamically and meets the needs of this experiment. As shown in Fig. 5.

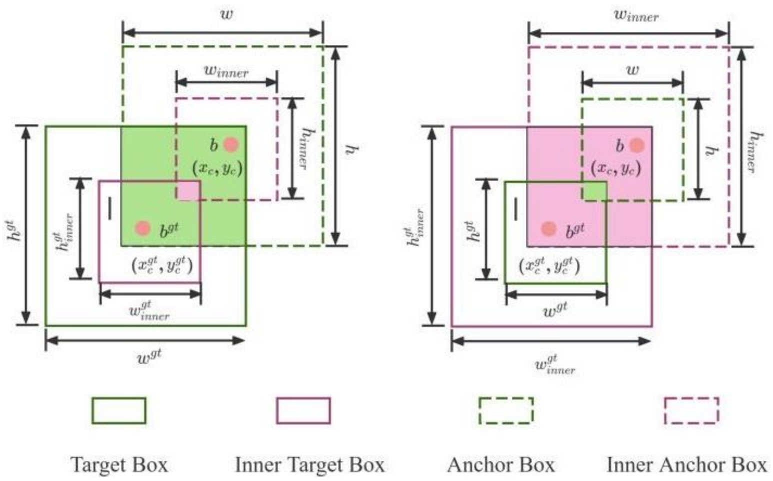


Fig. 5 -Description of Inner-WIoU

Inner-IoU is a combination of loss functions such as *EIoU* and *WIoU* using the idea of Inner. It controls the scale size of the auxiliary bounding box for calculating the loss by introducing the scale factor ratio, which can simultaneously take into account the convergence of the high *IoU* samples and the regression of the low *IoU*, and overcomes the limitations of the existing methods in terms of generalization ability (*Zhang et al., 2023*). *WloU* is a kind of bounding box loss based on the dynamic non-monotonous focusing mechanism, whose loss function is shown in Eq. (3):

$$L_{WIoU}=r\times R_{WIoU}\times L_{WIoU}\ ,R_{WIoU}\in[0,1] \tag{3}$$

where: the distance focusing mechanism is used to amplify the ordinary moderate anchor frame L_{IoL}, and the non-monotonic focusing coefficient r is used to focus the ordinary quality anchor frame, which provides better target frame regression loss to improve the performance of the target detector. *WIoU* removes the aspect ratio penalty term in CIoU, and also balances the effects of high and low-quality anchor frames on the regression of the model, which enhances the model's generalization ability, and improves the model's overall performance (*Tong et al., 2023*).

The *Inner-IoU* is applied to the existing *WIoU*-based marginal regression loss function defined as in equation (4):

$$L_{Inner\text{-}WIoU}=L_{WIoU}+IoU\text{-}IoU^{inner} \tag{4}$$

Experimental Platform

The main parameters of the AutoDL server platform used in this experiment are as follows: 12 vCPU Intel(R) Xeon(R) Silver 4214R CPU with 2.40 GHz, 90 GB of RAM, and an RTX 3080 Ti (12 GB) GPU. The experiments were conducted on the Linux operating system, and the PyTorch deep learning framework was used for model building, training, and evaluation. The model was built, trained, and evaluated using PyTorch version 1.13.1, Python version 3.8.6 (Ubuntu 20.04), and CUDA version 11.7.

Training Parameter Settings

The image input size was set to 640×640, the batch size was set to 16, and multithreading was configured to 8. The Stochastic Gradient Descent (SGD) optimizer was selected, and the training rounds were set to 200 with YOLOv8m's early stopping mechanism applied. The initial learning rate was 0.01, and the final learning rate was 0.001. The momentum parameter was set to 0.937, and the weight decay parameter was set to 0.0005. The random seed was fixed at 0.

Evaluation Metrics

To evaluate the performance of the model's detection results on the peach fruit moth dataset, the following evaluation criteria were chosen: precision (P), recall (R), mean average precision (mAP@0.5), mAP@0.5:0.95, single-category precision for peach fruit moth (*Carposina sasakii* AP/%), number of parameters (Params), model weights (Weight), and computational volume (FLOPs). These metrics were used to assess the model's effectiveness in detecting peach fruit moths and to compare its performance with other models.

RESULTS

Ablation Experiments

Ablation experiments were conducted for different improvement points, and the results are shown in Table 1. YOLOv8m-A introduced v7DS for downsampling in the backbone network, which improved the single-category accuracy of the peach fruit moth by 3.1 percentage points. YOLOv8m-B introduced the CPCA attention mechanism in the backbone network, which significantly reduced model weights and computational load. It reduced the weights by 45.52 MB and the computational volume by 70.3 G, respectively, compared to YOLOv8m. YOLOv8m-C introduced the Inner-WIoU loss function in the backbone network, which improved the all-category precision (P) by 1.6 percentage points over the original model. Both YOLOv8m-D and YOLOv8m-E showed improved performance compared to YOLOv8m. However, the YOLOv8m-VCI model, which incorporated all three improvement points, demonstrated even more significant enhancements. It achieved a 3.4 percentage point improvement in precision (P) compared to YOLOv8m. The recall (R) improved by 2.1 percentage points, and the mAP improved by 1.2 percentage points. The single-category precision (AP) for peach fruit moth detection improved by 2.4 percentage points. Additionally, the weight size, the number of model parameters, and the computational volume were reduced by 3.6 MB, 1.8 M, and 1.7 G, respectively. These improvements enhanced the model's detection effectiveness for adult peach fruit moths without increasing the model's complexity.

Table 1

Ablation test results for different improvement points.

Model	v7DS	CPCA	Inner-WIoU	P [%]	R [%]	mAP@0.5 [%]	AP [%]	Weight [MB]	Params [M]	FLOPs [G]
YOLOv8m	—	—	—	84.8	84.4	85.6	92.2	52	25.8	78.7
YOLOv8m-A	√	—	—	85.6	84.5	86.3	95.3	47.2	23.4	76.2
YOLOv8m-B	—	√	—	85.0	84.1	85.6	94.4	6.5	31.	8.4
YOLOv8m-C	—	—	√	86.4	84.0	86.4	94.1	52	25.8	78.7
YOLOv8m-D	√	√	—	85.7	85.3	86.0	93.6	48.4	24.0	77.0
YOLOv8m-E	√	—	√	88.0	83.2	87.0	93.6	47.2	23.4	76.2
YOLOv8m-VCI	√	√	√	88.2	86.5	86.8	94.6	48.	24.0	77.0

Note： √ indicates that the module is used; — indicates that the module is not used

Detection results of YOLOv8m-VCI

（1）The improved YOLOv8m-VCI achieved 88.2% precision, 86.5% recall, and 86.8% mAP. The single-class precision (AP) for peach fruit moth detection reached 94.6%. The detection results of YOLOv8m-VCI are visualized in Fig. 6. YOLOv8m incorporates an early-stopping mechanism to prevent model overfitting. The experiment was automatically halted when the optimal performance was achieved at 100 training epochs. The training curve of YOLOv8m-VCI stabilized in the later stages of the training process, exhibiting minimal fluctuations.

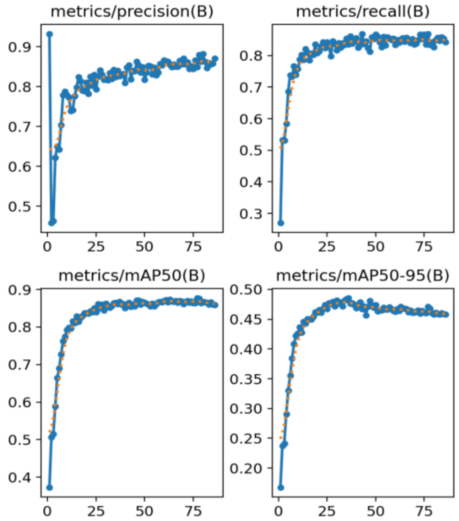


Fig. 6 -YOLOv8m-VCI curves during the training period

(2) In order to examine the improved model more closely, the P-R curves and F1-Confidence Curve for different categories are observed, as shown in Fig.7. The P-R curve, also known as the precision-recall curve, is a comprehensive representation of the model's performance. The more convex the P-R curve is, and the closer it is to the upper-right corner of the coordinate system, the better the model's performance.
The more convex the P-R curve is and the closer it is to the right corner of the coordinate system, the better the model performance. Analysing the P-R curve, it can be seen that YOLOv8m-VCI is 98.5% effective in detecting the adult peach small heartworm. The mAP for the whole category reached 86.7%. The F1-Confidence Curve demonstrates the relationship between the F1 scores and different confidence thresholds. The F1 score is the harmonic mean of Precision and Recall, ranging from 0 to 1, with larger values indicating better model performance. Analysing the F1-Confidence Curve shows that YOLOv8m-VCI is the best for detecting the adult peach small heartworm. The F1 score for the whole category reaches 84% when the confidence threshold is 0.304.

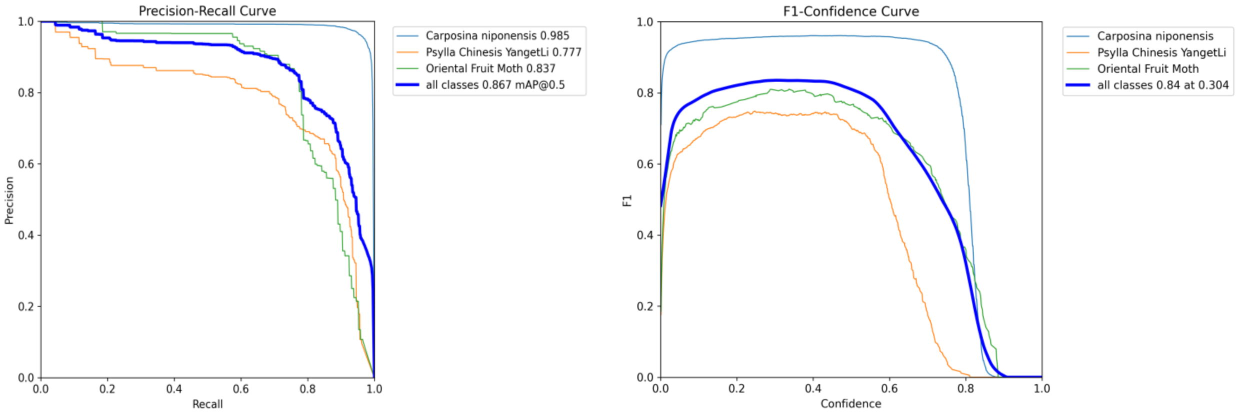


Fig. 7 - P-R curves and F1-Confidence Curve for YOLOvm-VCI during training

Performance Comparison of Different Target Detection Models

(1) In this experiment, seven mainstream target detection models—YOLOv5s, YOLOv5n, YOLOv7, YOLOv7-tiny, YOLOv8m, YOLOv8l, and YOLOv9c—were compared. The improved model, YOLOv8m-VCI, was also introduced for comparison. The results are presented in Table 2.

Table 2

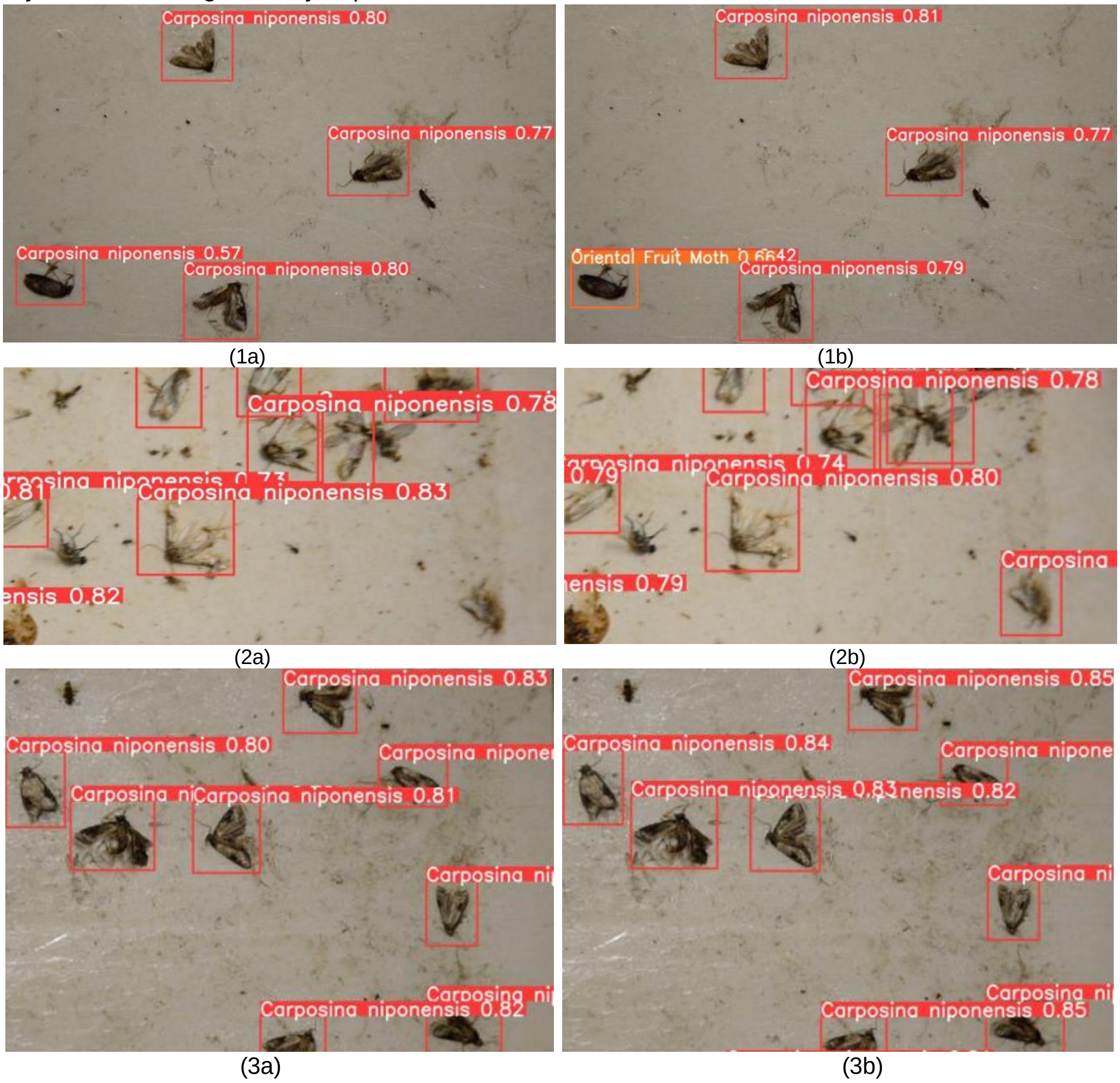
Comparison of detection effects of different models

Model	P [%]	R [%]	mAP@0.5 [%]	mAP@.5 .95 [%]	Weight [MB]	Params [M]	FLOPs [G]
YOLOv5s	83.6	86.5	83.8	43.8	14.4	7.02	15.8
YOLOv5n	83.	85.7	84.8	44.7	3.8	1.76	4.1
YOLOv7	83.3	85.4	84.3	41.9	71.3	37.2	105.1
YOLOv7-tiny	81.7	87.0	85.1	44.5	11.7	6.02	13.2
YOLOv8m	84.8	84.4	85.6	45.7	52.0	25.8	78.7
YOLOv8l	85.6	83.0	85.7	45.9	87.6	43.6	164.8
YOLOv9c	85.5	83.1	85.1	44.9	102.8	51.0	237.7
YOLOv8m-VCI	88.2	86.5	86.8	45.2	48.4	24.0	77.0

As can be seen from Table 2, YOLOv8m-VCI is optimal in terms of precision, which is also 2.7 percentage points better than YOLOv9c. It is only 0.5 percentage points lower than YOLOv7-tiny in recall, which is better than other models. It improves by 3 percentage points over YOLOv5s in mAP and is optimal. mAP50-95 reaches 45.2%. In addition, the model is lower than YOLOv8m in terms of weight, number of parameters, and computational effort. YOLOv8m-VCI meets the requirements for real-time detection of peach fruit moths in real environments.

(2) The detection effects of the improved YOLOv8m-VCI model and the YOLOv8m model are shown in Fig. 8. In (1a), the peach fruit moth was wrongly detected as peach fruit moth, and in (1b), it was detected correctly and with an accuracy P of 66%; in (2a), the YOLOv8m model in the lower-right corner of the figure omitted the detection of the peach fruit moth, and in (2b) the improved YOLOv8m-VCI model detected it correctly. The accuracy P of the improved YOLOv8m-VCI model for detecting peach fruit moth in (3b) is higher

than the accuracy of the YOLOv8m model in (3a). It can be seen that compared with the original model, the improved YOLOv8m-VCI model has improved the misdetection and omission of peach fruit moth, and the accuracy P has been significantly improved.



a. YOLOv8m b. YOLOv8m-VCI

Fig. 8 - Comparison of detection effect of YOLOv8m-VCI model and YOLOv8m model

The Peach Fruit Moth Detection System

The Peach Fruit moth identification application utilized the PyQt framework of Python for front-end interface development. PyQt combined the concise syntax of Python with the powerful functionalities of Qt. It offered a rich library of controls, facilitating users in constructing interfaces, and also achieved cross-platform compatibility, ensuring the consistency and stability of the user interface across different operating systems. The QtDesigner tool of PyQt allowed for a more intuitive and rapid design of the program interface, enhancing the development speed of the program, and separating the program interface from its logic, which made it easier to maintain in the later stages. As shown in Figure 9.

The application primarily implemented the following four functions:

1) Pest image file import function. This section provided four methods for importing image files: importing a single image, an entire file, a video file, and real-time detection via a camera.

2) Detection results display function. This section was responsible for displaying the detection data to the user, including the time taken for detection, the number of targets detected, the type of targets, confidence levels, and the location of targets. Additionally, it offered a target selection feature to meet the user's need to view the detection results of a single target.

3) Detection results and target location information display function. This section generated a list after the detection was completed. The list included the target serial number, file path, category, confidence level, and coordinate location.

4) Operation function. This section provided two operations that users could perform on the system. These were the save and exit functions.

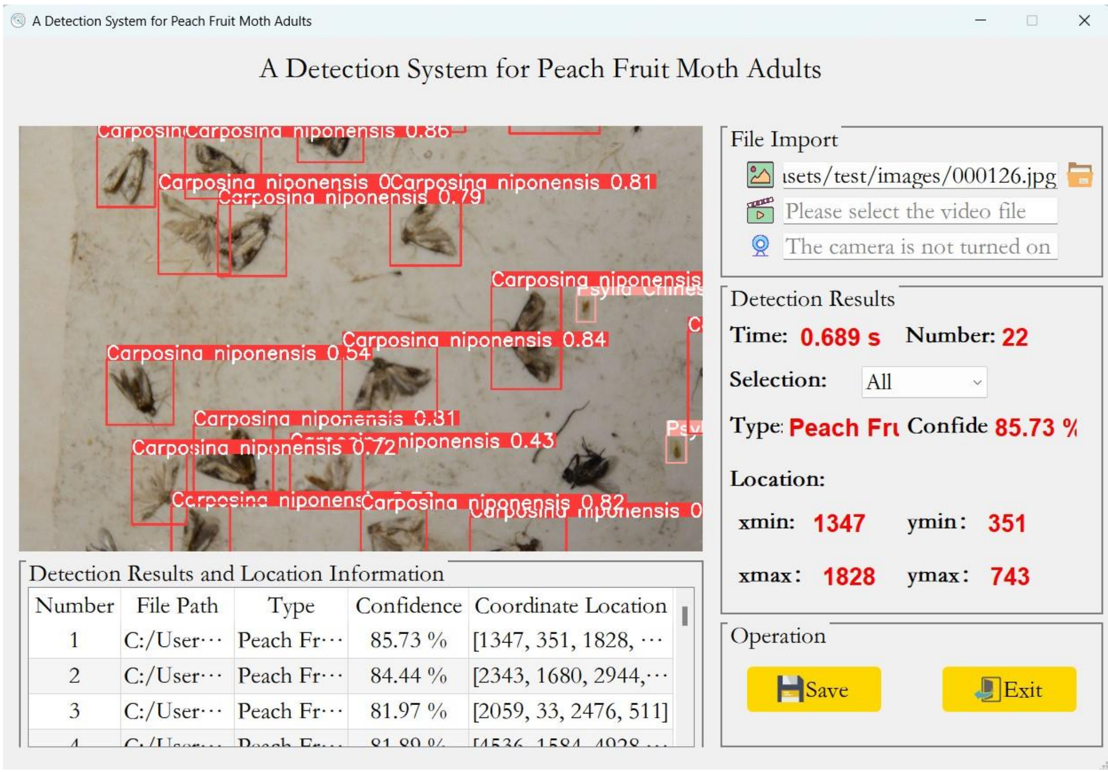


Fig. 9 - The Peach Fruit Moth Detection System

CONCLUSIONS

By introducing v7Down Sampling, CPCA Attention Mechanism, and Inner-WIoU Loss Function in the backbone network of YOLOv8m, the precision P is improved by 3.4 percentage points compared to YOLOv8m. Recall R improved by 2.1 percentage points. mAP improved by 1.2 percentage points. Peach fruit moth single-category precision AP improved by 2.4 percentage points. The size of weights, the number of model parameters, and the computational volume were reduced by 3.6MB, 1.8×106M, and 1.7G, respectively, which realized the improvement of the model's detection effect on adult peach fruit moth without increasing the model's complexity. The detection effect was better than the other seven YOLO series models compared.

The datasets of this study were collected outdoors, taking into account the real application scenarios of the model, and the real background was retained in the image processing, which is conducive to the model's ability to accurately detect small peach fruit moths even when the background is cluttered.

Follow-up research can expand the collection area of the dataset to broaden the scope of application of the model.

In the process of data collection, it was found that there was still a lot of Pear small heartworm and Chinese pear louse on the sticky boards with the addition of small peach fruit moth cores, so the two were added to the study together. However, the proportion of peach fruit moth, pear small heartworm, and Chinese pear louse samples in the actual dataset differed greatly, and there was the problem of sample imbalance. The detection accuracy of the single category of peach fruit moth with more sample data was high, while the detection accuracy of Pear small heartworm and Chinese pear louse with less sample data was low, thus affecting the detection effect of the whole category. In future research, the peach fruit moth and Chinese pear woodlouse can also be taken as research objects to increase the number of samples of the two, which can nearly improve the detection accuracy of the model for these three pests and enhance the detection efficiency.

Follow-up research can deploy the model on mobile terminals, deploy cameras in triangular traps in orchards, and utilize Internet of Things (IoT) technology to connect cameras and mobile terminals to realize real-time monitoring of peach fruit moth and improve the detection efficiency of peach fruit moth.

ACKNOWLEDGEMENT

This research was supported by the Teaching Reform Innovation Project of Higher Education Institutions in Shanxi Province (J20220247) and the Youth Scientific Research Project of the Basic Research Program of Shanxi Province (NO.202303021222039).

REFERENCES

[1] Dai, M., Dorjoy, M.M.H., Miao, H., & Zhang, S. (2023). A new pest detection method based on improved YOLOv5m. *Insects*, 14(1), 54.

[2] Fang, S.S., Qiao, X. F., Su, S., Jian, C.Z., & Chen, M.H. (2022). Advance of Research in Damage and Control of Carposina sasakii Matsumura (桃小食心虫的危害及其防治研究进展). *Shaanxi Journal of Agricultural Sciences*, 68(07):77-82.

[3] Haishi, T., Koizumi, H., Arai, T., Koizumi, M., & Kano, H. (2011). Rapid detection of infestation of apple fruits by the peach fruit moth, Carposina sasakii Matsumura, larvae using a 0.2-T dedicated magnetic resonance imaging apparatus. *Applied magnetic resonance*, 41, 1-18.

[4] Kim, D.S., Lee, J.H., & Yiem, M.S. (2000). Spring emergence pattern of Carposina sasakii (Lepidoptera: Carposinidae) in apple orchards in Korea and its forecasting models based on degree-days. *Environmental Entomology*, 29(6), 1188-1198.

[5] Kotani, S., & Suzuki, Y. (2013). Development of an X-ray penetration imaging system for detecting peaches injured by the peach fruit moth. *Journal of the Japan Society for Precision Engineering*, 79(11), 995-998.

[6] Li, B., & Fan, J. (2024). Classification of Rice Pests Based on YOLO v5 (基于YOLO v5 的水稻害虫分类). *Jiangsu Agricultural Sciences* (02), 175-182.doi:10.15889/j.issn.1002-1302.2024.02.025.

[7] Ma, P.， Yang, Z. H., Wan, H., He, S., Huang, Y., & Xu, S.Y. (2023) A new cotton aphid image recognition algorithm and software based on YOLOv8 (基于 YOLOv8 网络的棉蚜图像识别算法及软件系统设计). *Journal of Intelligent Agricultural Mechanization (in Chinese and English),* 4(03):42-49.

[8] Zhang, X.W., Ding, W.Y., Wang, W.L., Jia Y.J., Zhu, G.F., Zhao, P., Wang Y.P, Zhang, S.D., Li, Z., &Liu, X.X. (2024). The occurrence and population dynamics of adult Grapholita molestaand Carposina sasakii in two pear orchards in Shanxi Province (山西梨园梨小食心虫和桃小食心虫成虫种群发生动态研究). *Chinese Journal of Applied Entomology*, 61(4).

[9] Zhang, L.Y., Liu, A. H., Kong, T. T., Sheng, Z.H., & Zhu, Z.F. (2023). Progress in the study of Carposina sasakii Matsumura (桃小食心虫研究进展). *China Plant Protection* (12),20-26.

[10] Yue, X. S., (2024). Research progress of deep learning in crop disease identification (深度学习在农作物病害识别中的研究进展). *Anhui Agricultural Science Bulletin* (06), 100-103.doi:10.16377/j.cnki.issn1007-7731.2024.06.027.

[11] Zhao, W.H., Liu, X.K., Aerziguli, R.Z., Fu, K.Y., Ding, X.H., Jia, Z.Z., Adili, A.H.M.T., & Guo, W.C. (2023). Evaluation of the trapping effect of different types and densities of traps on the peach fruit moth (不同类型和密度的诱捕器对桃小食心虫诱捕效果评价). *Xinjiang Agricultural Sciences*, 60(12).

[12] Li, M., Xiao, Y.Y., Zong, W.Y., &Song, B. (2024). Detecting chestnuts using improved lightweight YOLOv8 (基于改进 YOLOv8 模型的轻量化板栗果实识别方法). *Transactions of the Chinese Society of Agricultural Engineering*, 40(1).

[13] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, 5(4), 1680-1716.

[14] Lin, L. H, Lin, S. L., Lu, B. J., Lin, Z. X. & Guo T. L. (2024). Remote Sensing Image Target Detection Algorithm Based on Spatial Feature Fusion Down-sampling Convolution (基于空间特征融合下采样卷积的遥感图像目标检测算法). *Infrared Technology*, 1-8.2024，54(2):1-8.

[15] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9), 1904-1916.

[16] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* pp. 7464-7475.

[17] Huang, H., Chen, Z., Zou, Y., Lu, M., Chen, C., Song, Y., ... & Yan, F. (2024). Channel prior convolutional attention for medical image segmentation. *Computers in Biology and Medicine*, 178, 108784.

[18] Wang, Q., Feng, W., Zhao, H., Liu, B., & Liu, S. (2024). VALNet: Vision-Based Autonomous Landing with Airport Runway Instance Segmentation. *Remote Sensing*, 16(12), 2161.

[19] Zhang, H., Xu, C., & Zhang, S. (2023). Inner-IoU: more effective intersection over union loss with auxiliary bounding box. *arxiv preprint arxiv*:2311.02877.

[20] Tong, Z., Chen, Y., Xu, Z., & Yu, R. (2023). Wise-IoU: bounding box regression loss with dynamic focusing mechanism. *arxiv preprint arxiv*:2301.10051.