KIWIFRUIT FLOWER DETECTION USING AN OPTIMIZED YOLOv11N ARCHITECTURE
/
基于改进 YOLOv11n 的猕猴桃花朵检测方法

Yin TANG, Junyu SUN, Yufei ZHANG, Chen WANG, Zhihao ZHANG, Fuxi SHI*
College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100/China
Tel: +86-18991839823; E-mail: shifuxi2012@126.com
DOI: https://doi.org/10.35633/inmateh-77-04

Keywords: *kiwifruit flowers, orchard environments, YOLOv11n, triplet attention module*

ABSTRACT
To accurately detect densely distributed kiwifruit flowers in complex orchard environments, this study proposes an improved detection model, YOLOv11-TYW, based on the YOLOv11n architecture. First, the RepViTBlock is integrated to enhance the model's feature representation capabilities. Second, the ADown module is introduced to improve the downsampling structure, thereby increasing detection accuracy for small flowers and branches while enhancing inference efficiency. Third, a triplet attention module is embedded in the head network to improve detection performance under conditions of occlusion caused by branches and overlapping flowers. Experimental results show that the YOLOv11-TYW model achieves a precision of 88.4%, a recall of 89.1%, and a mean average precision (mAP) of 91.2%, representing improvements of 4.3, 4.4, and 6.7 percentage points, respectively, over the baseline YOLOv11n model. When tested on kiwifruit flower images captured in various orchard environments, YOLOv11-TYW produces more accurate bounding boxes, with fewer false positives and missed detections. Applying the improved model to the actual kiwi orchard environment, demonstrate that YOLOv11-TYW exhibits excellent detection performance in real-world orchard settings and offers technical support for automated kiwifruit flower pollination.

摘要
为实现对果园环境中复杂分布猕猴桃花朵的准确检测，本研究提出一种基于改进 YOLOv11n 的猕猴桃花朵检测模型 YOLOv11-TYW。首先，在 YOLOv11n 的基础上引入 RepViTBlock 增强模型的特征表达能力。其次，引入 ADown 模块改进模型的下采样结构，提升模型对小花朵、枝叶的检测精度与模型推理效率；最后，引入三重注意力模块改进 Head 网络结构，提升模型对于枝叶遮挡、相互遮挡猕猴桃花朵的检测精度。结果说明，YOLOv11-TYW 模型的精确率、召回率和平均精度均值分别为 88.4%、89.1%与 91.2%，相比 YOLOv11n 模型其精确率、召回率和平均精度均值分别提高了 4.3、4.4 和 6.7 个百分点。使用不同环境的猕猴桃花朵照片对改进模型进行检测时，改进的 YOLOv11-TYW 相较于 YOLOv11n 模型的预测边界框更接近花朵目标，并减少了误检与漏检的情况。将改进的模型运用在实际的猕猴桃果园环境中， 结果表明，YOLOv11-TYW 模型在真实猕猴桃果园环境中表现出优良的检测性能，能够实现对密集分布猕猴桃花朵的准确检测，可为猕猴桃花朵的自动授粉提供技术支持。

INTRODUCTION
China is the world's leading producer of kiwifruit, with the largest cultivation area and yield globally (*Zhang et al., 2014*). Ensuring high-quality and high-yield kiwifruit production is crucial for increasing farmers' income and promoting agricultural development. As kiwifruit is a dioecious species that relies on cross-pollination (*Sun et al., 2025*), fruit quality is closely related to the quantity and effectiveness of pollen, making pollination a critical factor for yield improvement. However, in orchard environments, biological pollination is often inefficient (*Nicholson et al., 2020*), and manual pollination is labor-intensive and time-consuming (*Liu et al., 2019*). Consequently, increasing attention has been given to the development of pollination robots to automate this process.

Although pollination robots can significantly improve efficiency, kiwifruit flowers often grow in dense clusters, with overlapping and occlusion common in orchard settings. These conditions pose challenges for flower detection, often leading to false positives and missed detections. Therefore, achieving fast and accurate detection of kiwifruit flowers in natural orchard environments has become a key technical issue in pollination robot research.

With the rise of convolutional neural networks (CNNs), deep learning has demonstrated powerful capabilities in feature extraction and generalization, and has been widely applied to various agricultural vision tasks such as pest and disease detection (*Sun et al., 2017*), crop classification and recognition (*Long et al., 2018*), and weed identification (*Gao et al., 2017*). Deep learning-based object detection methods can be generally classified into two-stage detectors and one-stage detectors (*Zhao et al., 2020*). Two-stage methods (e.g., R-CNN, R-FCN) first generate region proposals and then perform feature extraction and classification, offering high accuracy but lower speed (*Li et al., 2022*). In contrast, one-stage detectors (e.g., OverFeat, YOLO, SSD) integrate object localization and classification within a single network structure, offering faster inference speeds with competitive accuracy. Among them, the YOLO series has become one of the most widely used algorithms for flower detection due to its speed and accuracy (*Shang et al., 2022*).

Recent studies have adapted YOLO-based models for flower detection in complex environments. *Zhang et al., (2025*), proposed an improved YOLOv7-based pear flower detection model, incorporating attention mechanisms and loss function optimization to enhance detection performance for small and distant targets in cluttered backgrounds, achieving a precision of 99.4%, recall of 99.6%, and mAP of 96.4%. *Zhang et al., (2024*), introduced a lightweight safflower detection model based on an optimized YOLOv8n architecture, utilizing a Vanillanet backbone and a large separable kernel attention module to reduce model complexity and improve robustness, with precision and mAP reaching 93.10% and 96.40%, respectively.

For kiwifruit flowers, which are densely distributed and prone to occlusion, *Gong et al., (2023*), developed a YOLOv5s-based detection method tailored to natural orchard environments, identifying issues such as overlapping buds and flowers, low illumination, and morphological similarities during specific growth stages as major causes of false detections. *Liu et al., (2025),* proposed YOLOv8-KFP, an improved YOLOv8n-based model designed for densely distributed kiwifruit flowers, effectively enhancing detection accuracy under conditions of occlusion and phenological variation, thereby supporting automated pollination systems.

In this study, we propose a novel kiwifruit flower detection approach based on an improved YOLOv11n model. The introduction of RepViTBlock module in Backbone backbone enhances the feature extraction capability of the model. ADown downsampling structure is used instead of downsampling structure to improve the inference efficiency of the model. Triplet attention mechanism is introduced in Detection Head to improve the perception ability of the model to distinguish flowers that are obstructed by branches and leaves or flowers that are obstructed by each other. This study provides technical support for automatic pollination of kiwifruit flower.

MATERIALS AND METHODS

IMAGE ACQUISITION AND DATASET CONSTRUCTION

This study focuses on kiwifruit flowers of the ‘Hayward’ variety. The flower images were collected at the Meixian Kiwifruit Experimental Station of Northwest A&F University. Two devices were used for image acquisition: a Canon EOS 90D camera and an Apple A3092 smartphone. All images were captured under automatic settings for focal length, exposure, and shooting mode, with a resolution of 6960 × 4640 pixels. The data collection was conducted from May 5 to May 20, 2025, during two daily time periods: 9:00–12:00 and 14:00–18:00. The shooting distance ranged from 30 to 80 cm, covering various lighting conditions, including front lighting and backlighting.

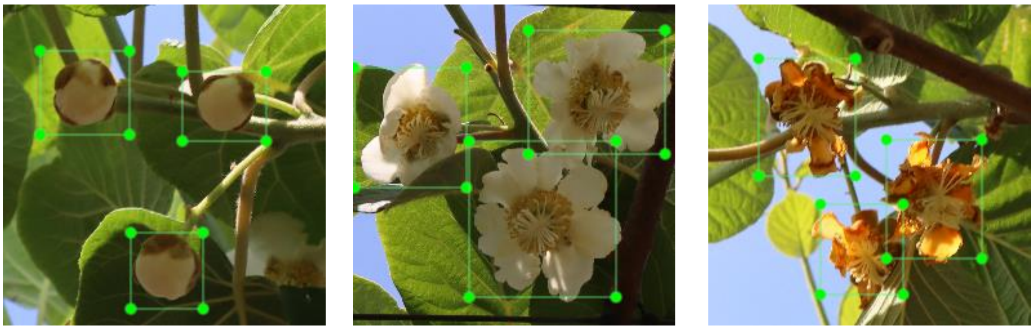
The dataset comprises a total of 2,769 images under different occlusion scenarios: 1,024 images of unobstructed flowers, 991 images of flowers with mutual occlusion, and 754 images of flowers occluded by branches or leaves, as shown in Figure 1.

(a) Not obstructed (b) Covered by branches and leaves (c) Flowers cover each other

Fig. 1 - Three different occlusion situations

All images were standardized and resized to a resolution of 640 × 640 pixels. Manual annotation of the dataset was performed using the LabelImg tool, with labeled targets represented by the minimum enclosing bounding boxes. As shown in Figure 2, the annotated targets—i.e., the kiwifruit flower objects to be detected—were categorized into three classes: bud, flower, and pollinated (naturally pollinated flowers).

The dataset comprising 2,769 images was split into a training set (1,938 images), a validation set (553 images), and a test set (278 images) following a 7:2:1 ratio for model training and evaluation.



（a）Kiwi flower bud （b）Kiwi flower （c）Pollinated kiwifruit flowers

Fig. 2 - Label target category

Detection model based on improved YOLOv11n

YOLOv11n, the latest release in the YOLO series by Ultralytics is a lightweight object detection network built upon the YOLOv8 architecture (*Han et al., 2025*). It primarily consists of three components: the Backbone, the Neck, and the Detection Head.

In this study, to address the challenges of detecting densely distributed kiwifruit flowers in complex orchard backgrounds, we propose an improved model named YOLOv11-TYW, based on the YOLOv11n architecture.

First, a RepViTBlock module is integrated into the Backbone to enhance feature extraction and improve the model’s ability to recognize fine-grained details of kiwifruit flowers. Second, an ADown downsampling structure replaces the original downsampling layers to better accommodate the varying shapes of kiwifruit flowers, thereby improving both detection accuracy and inference efficiency. Third, to address occlusion caused by overlapping flowers, a Triplet Attention Mechanism is incorporated into the Detection Head, strengthening the model’s capacity to perceive and distinguish partially obscured or closely clustered flowers.

The improved network architecture is illustrated in Figure 3.

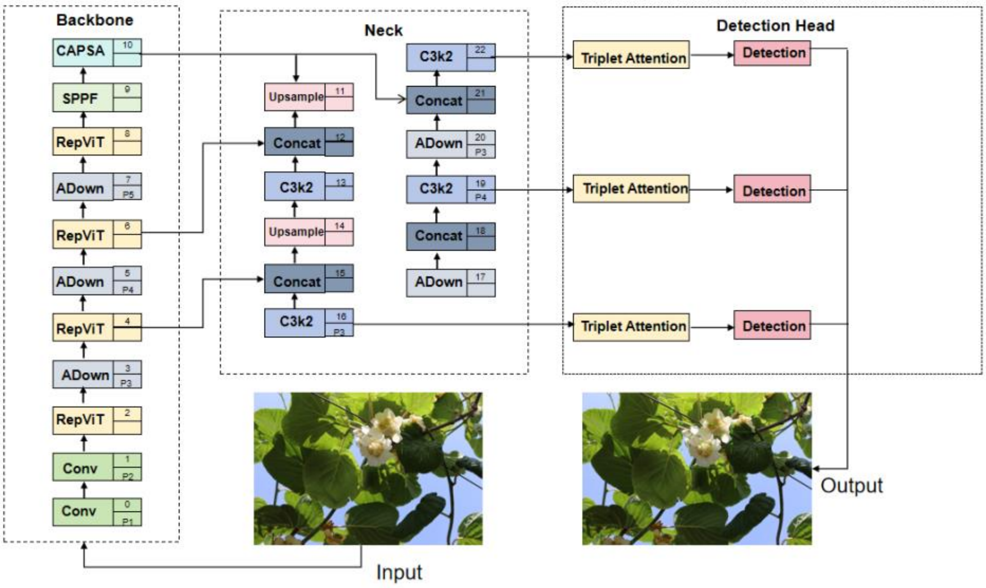


Fig. 3 - YOLOv11-TYW network architecture diagram

RepViTBlock module

The RepViTBlock is a novel module that integrates efficient architectural designs from lightweight Vision Transformers (ViTs) with enhancements to standard lightweight convolutional neural networks (CNNs) (*Wang et al., 2024*). By combining the strengths of both CNNs and ViTs, RepViTBlock demonstrates strong performance across various visual recognition tasks. Its structural design is illustrated in Figure 4.

RepViTBlock optimizes the network by reorganizing architectural components to improve training efficiency and by adjusting convolutional expansion ratios to reduce parameter latency. When incorporated into the YOLOv11n backbone, this module significantly enhances the model's capacity to extract fine-grained features and improves recognition performance, particularly for kiwifruit flowers that are partially occluded by branches or overlapping with other flowers.

Compared to the original YOLOv11n model, the integration of RepViTBlock results in better detection accuracy for kiwifruit flowers under complex orchard conditions, especially in scenarios involving occlusion and dense distribution.

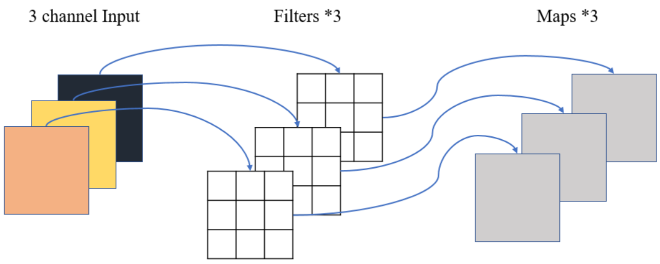


Fig. 4 - RepViTBlock structural diagram

ADown module

In orchard environments, kiwifruit flowers are typically situated in open settings with complex visual backgrounds. Variations in lighting, shadow interference, occlusion by branches and leaves, and sky backgrounds can significantly impact detection accuracy. To address these challenges, this study introduces the ADown module as a replacement for the conventional downsampling structure in YOLOv11-TYW, aiming to enhance both the detection precision for small targets and the model's inference efficiency.

ADown is a downsampling module designed for YOLO-based object detection models (*Bai et al., 2025*), with its architecture illustrated in Figure 5. Compared to traditional pooling or standard convolutional downsampling methods, ADown adopts a more optimized design to reduce information loss. It decreases the spatial resolution while increasing the number of feature channels, employing techniques such as stride convolution and channel reorganization to preserve critical features.

Given that many kiwifruit flower images in the dataset were captured from long distances—resulting in small object sizes—and are often occluded by surrounding foliage, replacing the standard convolutional layers with the ADown module significantly improves the model's ability to detect small flowers and fine structural elements. Overall, the incorporation of ADown enhances YOLOv11-TYW's robustness and accuracy when detecting kiwifruit flowers in complex orchard scenes.

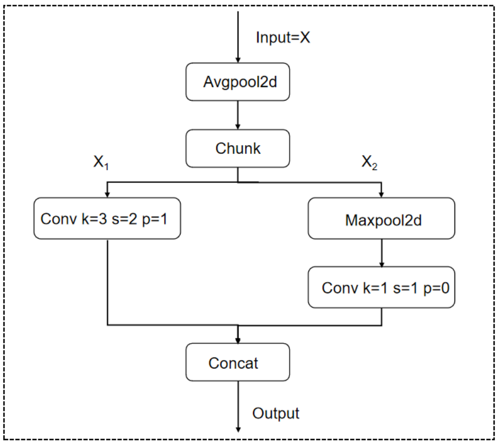


Fig. 5 - ADown structural diagram

Triplet Attention module

To enhance the detection accuracy of kiwifruit flowers that are partially occluded by branches or overlapping with each other, this study incorporates a Triplet Attention Module into the YOLOv11-TYW model. Triplet attention is an attention mechanism designed to improve the feature representation capability of deep neural networks by capturing cross-dimensional interactions within the input data (Misra et al., 2021). Its simplified architecture is illustrated in Figure 6.

The core idea of the triplet attention mechanism is to generate attention weights by modeling dependencies across different dimensions of the input tensor using three parallel branches, each focusing on specific spatial-channel relationships. These branches apply tensor rotation and residual transformations to capture interactions from multiple perspectives:

- Branch ***a*** processes the input tensor directly without rotation and applies residual transformations to extract spatial features.
- Branch ***b*** rotates the input tensor along the Width (W) and Channel (C) dimensions before applying residual transformations to learn width-channel dependencies.
- Branch ***c*** rotates the input tensor along the Height (H) and Channel (C) dimensions and similarly applies residual transformations to capture height-channel interactions.

Each branch outputs an intermediate representation, which is then used to generate attention weights through pooling and convolution operations. These weights are applied to the transformed tensors and finally rearranged to match the original input shape. By capturing rich multi-dimensional dependencies, the triplet attention mechanism improves the model's ability to accurately detect kiwifruit flowers under complex occlusion scenarios.

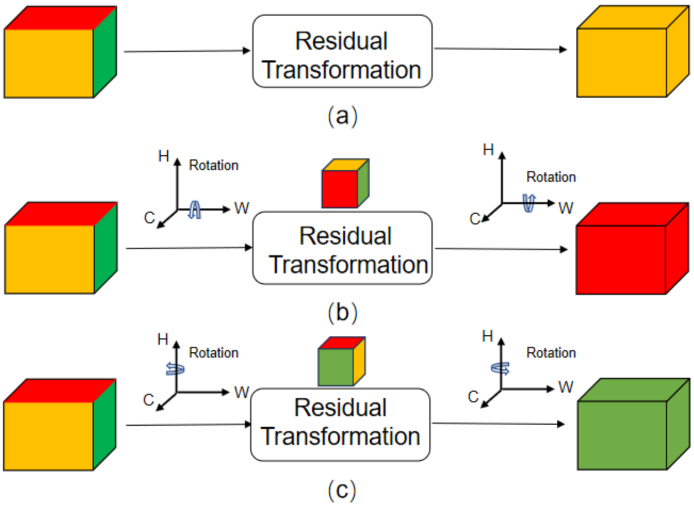


Fig. 6 - Triplet Attention structural diagram.

RESULTS

EXPERIMENTAL DESIGN AND ANSLYSIS

Evaluation indicators

To comprehensively evaluate the detection performance of the proposed YOLOv11-TYW model, the following metrics were used:

- Precision (P), Recall (R), and mean Average Precision at IoU=0.5 ($mAP0.5$) to assess detection accuracy.
- Parameter count (M), Floating Point Operations *(Redmon J. et al., 2017)*, and Frames Per Second *(Jiang Bet al., 2018)* to assess model complexity and inference efficiency.

The metrics are defined as follows:

$$P = \frac{TP}{TP + FP} \times 100\% \tag{1}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \times 100\% \tag{3}$$

$$AP = \int_0^1 P(R)\,dR \times 100\% \tag{4}$$

where:

TP is the number of true positives;

FP is the number of false positives;
FN is the number of false negatives;
N is the number of classes.

MODEL TRAINING

The YOLOv11-TYW model was trained for 300 epochs. As shown in Figure 7, the model achieved a precision of 88.4% and an mAP0.5 of 91.2% after 300 iterations. The loss curve in Figure 8 shows a consistent decrease and convergence without signs of overfitting, indicating that the model was well-trained and stable for further evaluation.

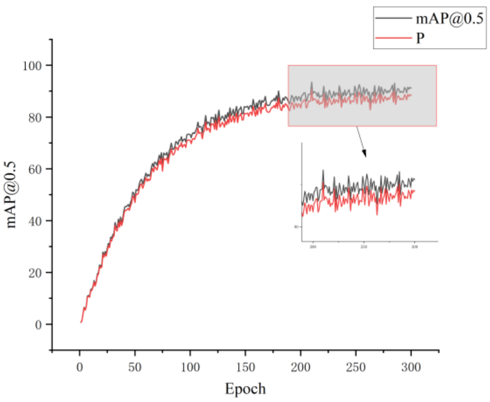


Fig. 7 - Accuracy curve chart

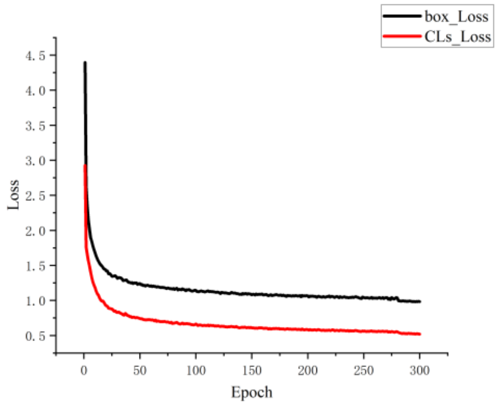


Fig. 8 - Loss function curve graph

Comparison of Detection Results from Different Models

To further assess the effectiveness of YOLOv11-TYW, we compared it with SSD, YOLOv8n, and YOLOv11n under the same dataset and training conditions. The results are shown in Table 1.

Table 1

Contrast test

Model	P/%	R/%	$mAP_{0.5}$/%	GFLOPs	Parameter/M	FPS
SSD	82.4	77.8	83.7	238.4	21.12	78.91
YOLOv8n	83.9	84.7	84.5	23.7	11.04	121.32
YOLOv11n	84.7	85.3	85.4	6.5	2.61	102.45
YOLOv11-TYW	88.4	89.1	91.2	8.9	4.5	108.21

As seen in Table 2, YOLOv11-TYW outperforms the other models in all three-accuracy metrics. Compared to SSD, YOLOv8n, and YOLOv11n, it achieves:

- +7.2%, +5.3%, and +4.3% improvement in Precision
- +14.5%, +5.2%, and +4.4% in Recall
- +8.9%, +7.6%, and +6.7% in mAP0.5

In terms of computational efficiency, YOLOv11-TYW maintains a balance between speed and accuracy with only 8.9 GFLOPs, 4.5M parameters, and 108.21 FPS, making it suitable for real-time detection in orchard scenarios.

Ablation Study

To evaluate the individual contribution of each improvement module, were conducted ablation experiments using YOLOv11n as the baseline. Modules tested include the RepViTBlock, ADown, and Triplet Attention mechanisms. All models were trained under identical conditions, and performance was evaluated using P, mAP0.5, GFLOPs, parameters, and FPS. The results are presented in Table 2.

The results show that:

- RepViTBlock improved precision and mAP0.5 by +2.0% and +2.4%, enhancing feature extraction.
- ADown led to moderate accuracy gains while significantly reducing GFLOPs and parameters (by 36.9% and 31.1%).
- Triplet Attention improved precision and mAP0.5 by +2.7% and +2.5%, strengthening occlusion awareness.

Combining all three modules yielded the best performance, validating the effectiveness of each component and their synergy.

Table 2

Ablation experiment

RepViT Block	ADown	Triplet Attention	P/%	$mAP_{0.5}$/%	GFLOPs	Parameter/M	FPS
—	—	—	84.7	85.4	6.5	2.61	102.27
√	—	—	86.7	87.1.	7.5	4.1	103.47
—	√	—	85.1	86.2.	4.1	1.8	105.84
—	—	√	87.0	87.5	6.7	2.9	103.52
√	√	—	87.2	87.6	9.6	5.2	104.51
√	—	√	89.7	90.5	9.1	4.9	105.68
—	√	√	88.1	88.5	8.5	3.79	106.95
√	√	√	88.4	91.2	8.9	4.5	108.21

Note: "-" indicates not using this module;
√ "indicates the adoption of this module;
P is the precision rate; R is the recall rate;
MAP0.5 is the average precision mean;
GFLOPs are floating-point operands.

Comparison of Detection Performance Across Models Under Different Conditions

To assess real-world robustness, images captured under diverse conditions—sunny and cloudy weather, mutual occlusion, and branch occlusion—were randomly selected for visual comparison between YOLOv11n and YOLOv11-TYW. Results are shown in Figures 9 and 10.

Fig. 9 - Detection effect of different models on densely distributed flowers

Note: The blue detection box corresponds to the kiwifruit flower of YOLOv11n, and the purple detection box corresponds to the kiwifruit flower bud of YOLOv11n; The red detection box corresponds to the kiwifruit flowers of YOLOv11-TYW, the yellow detection box corresponds to the kiwifruit flower buds of YOLOv11-TYW, and the orange detection box corresponds to the pollinated kiwifruit flowers; The flowers or flower buds that were missed in the green circle. Same below.

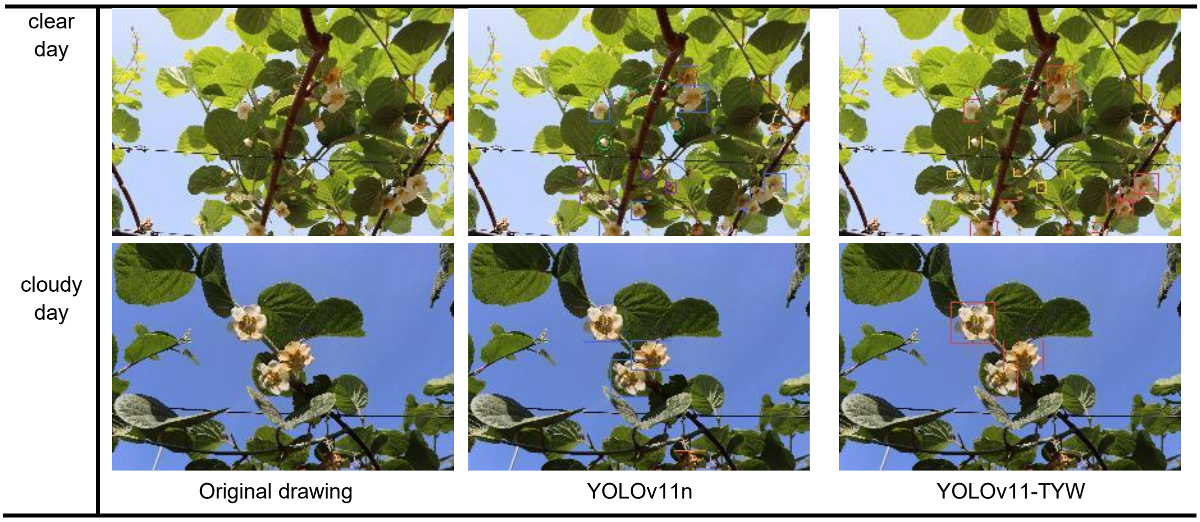


Fig. 10 -The detection effect of different models on sunny flowers

- Unoccluded scenes: Both models performed well, but YOLOv11-TYW predicted bounding boxes closer to the flower edges.
- Mutual occlusion: YOLOv11-TYW correctly detected two overlapping flowers that YOLOv11n misclassified.
- Branch occlusion: YOLOv11-TYW successfully detected flowers occluded by branches, while YOLOv11n missed one flower.
- High-light sunny conditions: YOLOv11-TYW avoided a false negative and reduced misclassification of flower buds as open flowers.
- Cloudy conditions: Both models performed equally well, with no false detections.

These results confirm that YOLOv11-TYW is more robust in handling complex orchard conditions with better localization, fewer false positives, and improved recall under occlusion and varying lighting.

CONCLUSIONS

To address the challenges of detecting densely distributed kiwifruit flowers in orchard environments, especially under conditions of occlusion and overlap, this study proposes an improved detection model, YOLOv11-TYW, based on the YOLOv11n architecture.

The following conclusions can be drawn:

1. Model Architecture Improvements: The introduction of the RepViTBlock in the backbone network enhances feature extraction and improves the recognition accuracy for fine-grained details of kiwifruit flowers. The replacement of the standard downsampling module with the ADown structure optimizes the downsampling process and improves inference efficiency. Additionally, the integration of a Triplet Attention mechanism in the detection head further boosts detection performance. Compared with the original YOLOv11n, the YOLOv11-TYW model achieves improvements of 4.3%, 4.4%, and 6.7% in precision, recall, and mAP0.5, respectively.

2. Performance Comparison with Other Models: Comparative experiments were conducted using SSD, YOLOv8n, YOLOv11n, and the proposed YOLOv11-TYW model. Results show that YOLOv11-TYW achieves the highest performance across all metrics, with a precision of 88.4%, recall of 89.1%, and mAP0.5 of 91.2%. Moreover, it maintains competitive efficiency with 8.9 GFLOPs, 4.5M parameters, and a real-time processing speed of 108.21 FPS, demonstrating its overall superiority in both accuracy and efficiency.

3. Robustness Under Complex Conditions: Multi-scenario experiments were conducted under diverse conditions, including sunny and cloudy weather, mutual flower occlusion, and occlusion by branches. The YOLOv11-TYW model consistently produced bounding boxes closer to the actual flower regions, with fewer false positives and missed detections compared to YOLOv11n. These results demonstrate that the proposed model performs robustly in real-world orchard environments and offers reliable detection for densely clustered kiwifruit flowers.

In summary, the proposed YOLOv11-TYW model exhibits excellent performance in detecting kiwifruit flowers under complex environmental conditions and provides a strong technical foundation for intelligent and automated pollination systems.

ACKNOWLEDGEMENT

This study was funded by the Shaanxi Province Key Industry Innovation Chain (Group) Project (2024NC-ZDCYL-05-10).

REFERENCES

[1] Bai K., Zhang Y., Su D. (2025). Peanut leaf disease detection method based on improved YOLO v8n (基于改进 YOLO v8n 的花生叶片病害检测方法). *Journal of Agricultural Machinery*, Vol.56, pp.518-526., Hubei/China.

[2] Gao Z., Wang A., Liu Y. (2017). Research on Intelligent Sorting System for Fresh Tea Based on Convolutional Neural Network (基于卷积神经网络的鲜茶叶智能分选系统研究). *Journal of Agricultural Machinery*, Vol.48, pp.53-58, Anhui/China.

[3] Gong W., Yang Z., Li K. (2023) A method for detecting kiwifruit flowers in natural environments based on improved YOLOv5s (基于改进 YOLOv5s 的自然环境下猕猴桃花朵检测方法). *Journal of Agricultural Engineering*, Vol.39, pp.177-185, Shanxi/China.

[4] Han T., Yu S., Ma L. (2025). Research on the Foreign Object and Defect Detection Model of Photovoltaic Panels Based on Improved YOLOv11n (基于改进 YOLOv11n 的光伏板异物与缺陷检测模型研究). *Computer Engineering and Applications,* Vol.61, pp.1-17, Anhui/China.

[5] Jiang B., Luo R, Mao. (2018). Acquisition of Localization Confidence for Accurate Object Detection. *Computer Vision–ECCV*. pp.816-832. Cham/Switzerland.

[6] Li K., Chen Y., Liu J. (2022). A review of deep learning based object detection algorithms (基于深度学习的目标检测算法综述). *Computer Engineering,* Vol.48, pp.1-12, Shanxi/China.

[7] Liu J., Li S. (2025). A method for detecting the flowering period of densely distributed kiwifruit based on improved YOLOv8n (基于改进 YOLOv8n 的密集分布猕猴桃花期检测方法). *Journal of Agricultural Engineering,* Vol.41, pp.172-181, Shanxi/China.

[8] Long M., Ou Y., Chun J., Liu H. (2018). Oil tea disease image recognition based on convolutional neural network and transfer learning (基于卷积神经网络与迁移学习的油茶病害图像识别). *Journal of Agricultural Engineering*, Vol.34, pp.194-201, Jiangxi/China.

[9] Misra D., Nalamada T., Arasanipalai A.U. (2021). Rotate to attend: convolutional triplet attention module. *2021 IEEE Winter Conference on Applications of Computer Vision*, pp.3138-3147. Waikoloa/USA.

[10] Nicholson C.C., Egan P.A. (2020). Natural hazard threat to pollinators and pollination. *Global Change biology*, Vol.26, pp.380-391, Mississippi/USA.

[11] Redmon J., Farhadi A. (2017). YOLO9000: Better, faster, stronger. *Computer Vision and Pattern Recognition*. pp.6517-6525. Honolulu/USA.

[12] Shang Y., Zhang Q., Song H. (2022). Application of YOLOv5s based deep learning in natural scene apple flower detection (基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用). *Journal of Agricultural Engineering*, Vol.38, pp222-229, Shanxi/China.

[13] Sun J., Tan W., Mao H., (2017). Identification of Multiple Plant Leaf Diseases Based on Improved Convolutional Neural Networks (基于改进卷积神经网络的多种植物叶片病害识别). *Journal of Agricultural Engineering,* Vol.33, pp.209-215, Jiangsu/China.

[14] Sun J., Wang D., Xiao X. (2025). Study of pollen deposition performance of an airflow-assisted targeted pollinating device for kiwi fruit flower. *Biosystems Engineering*, Vol.251, pp.31-47, Shanxi/China.

[15] Wang A., Chen H., Lin Z. (2024). RepViT: revisiting mobile CNN from ViT perspective. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition.* pp.15909-15920, Anhui/China.

[16] Zhang J., Mo Z., Huang S. (2014). Development of World Kiwi Industry and Analysis of China's Kiwi Trade and International Competitiveness since the 21st Century (21 世纪以来世界猕猴桃产业发展以及中国猕猴桃贸易与国际竞争力分析). *China Agricultural Journal*, Vol. 30, pp.48-55, Jiangsu/China.

[17] Zhou F., Jin L., Dong J. (2017). A Review of Research on Convolutional Neural Networks (卷积神经网络研究综述). *Chinese Journal of Computer Science*, Vol.40, pp.1229-1251, Jiangsu/China.

[18] Zhao Y., Rao Y., Dong S. (2020). A review of deep learning object detection methods (深度学习目标检测方法综述). *Chinese Journal of Image and Graphics*, Vol.25, pp.629-654, Shanxi/China.

[19] Zhang X., Wei H., Kong Degang, (2025). Pear Blossom Recognition Method Based on Improved YOLOv7 in Natural Environment (自然环境下基于改进 YOLOv7 的梨花识别方法). *Journal of Agricultural Engineering*, Vol.41, pp 224-232, Shanxi/China.

[20] Zhang X., Hu G., Li P. (2024). Lightweight safflower recognition method based on improved YOLOv8n (基于改进 YOLOv8n 的轻量化红花识别方法). *Journal of Agricultural Engineering*, Vol.40, pp.163-170, Shanxi/China.