RESEARCH ON REAL-TIME CORN PEST DETECTION METHOD BASED ON CSPPC LIGHTWEIGHT MODULE AND Wise-IoU
/
基于 CSPPC 轻量化模块与 Wise-IoU 的玉米虫害实时检测方法研究

Qiuyan LIANG [1], Haiyang YU [1], Aidi XU [2], Mengyuan JIA [1], Zihan ZHAO[1], Jia CHI [*3]
[1] School of Mechanical Engineering, Jiamusi University, Jiamusi, Heilongjiang 154007, China
[2] Academy of Agricultural Machinery Engineering Sciences, Heilongjiang Province, Harbin, Heilongjiang 150081, China
[*3] Logistics Management Office, Jiamusi University, Jiamusi, Heilongjiang 154007, China
Tel: +8615246498069; E-mail: chijia@jmsu.edu.cn
Corresponding author: Jia CHI

DOI: https://doi.org/10.35633/inmateh-77-57

***Keywords:** Corn pest detection; Lightweight model; CSPPC module; Wise-IoU loss; Mobile deployment*

ABSTRACT
To address the issues of large number of parameters and low deployment efficiency on mobile devices in the existing YOLOv8-DSFF model for corn pest detection, this study proposes an improved object detection model that integrates the CSPPC lightweight module and the Wise-IoUv3 loss function. The optimized model reduces the number of parameters by 85.6%, achieves an mAP@0.5 of 90.8%, reaches 204 FPS inference speed on PC and 42 FPS on mobile devices. This provides a practical low-power solution for real-time field monitoring of corn pests.

摘要
此前玉米虫害识别研究中，YOLOv8-DSFF 模型虽较其他检测模型优势显著，但存在参数量大、移动端部署效率低的问题。为此，研究提出融合 CSPPC 轻量化模块与 Wise-IoUv3 的目标识别模型改进方案。通过改进模型参数量降低 85. 6%；模型 mAP@0. 5 达 90. 8%；PC 端推理速度 204FPS；移动端帧率 42FPS，可为田间实时监测提供低功耗方案。

INTRODUCTION
Corn is an important crop in China, and its yield and quality are of great significance for national food security and farmers' income *(Chen, 2013)*. However, pests have always been one of the key factors affecting corn growth and yield. According to statistics, pests can cause a 10%-15% reduction in corn yield, resulting in huge losses to agricultural production (*Wang et al., 2025*). Traditional manual detection methods are inefficient and difficult to meet the real-time monitoring needs of large-scale farmland. Therefore, developing an efficient and accurate real-time corn pest detection method is of great practical significance (*Wu et al., 2024*).

In recent years, deep learning technology has made significant progress in the field of object detection. The YOLO (You Only Look Once) series models have been widely used in various object detection tasks due to their high efficiency and accuracy (*Chu et al, 2025*). *Gao et al., (2025)*, proposed AGRI-YOLO, a lightweight corn weed detection model based on YOLO v11n. Optimized with DWConv, ADown, and LADH, it achieves 82.8% mAP50 (similar to baseline), with 46.6% fewer parameters, 49.2% lower GFLOPs, fit for edge devices. *Nguyen et al., (2025),* proposed αSiLU, an improved activation function for YOLO models, integrating scaling factor α into standard SiLU to boost feature extraction. With α=1.05 on tomato/cucumber datasets, YOLOv11n's mAP@50 rose by 1.1%/0.2%, with minimal inference speed impact, suitable for real agriculture. *Kamat et al., (2025),* benchmarked four models (YOLOv5, YOLOv6, YOLOv7, SSD-MobileNetv1) for multi-class fruit ripeness detection on strawberries and avocados to reduce post-harvest losses. Using a publicly available, naturally captured annotated dataset and 5-fold cross-validation, YOLOv6 achieved the highest mean accuracy (99.5%) and a good balance with real-time speed (85.2 FPS), proving most reliable for smart sorting. *Hao et al., (2025),* proposed BCS_YOLO based on YOLOv11n for corn leaf pest detection, adding SPCGA, HLFFE, LAE modules. It achieved 78.4% precision, 82.0% mAP@50,3. 0%-4.6% higher than baseline, outperforming mainstream models. *Ganapathy et al., (2025),* evaluated four YOLO models for guava defect detection. YOLOv11n, with 2.6M params and 6.3 GFLOPs, achieved 98.0% mAP50-95 and 255 FPS, outperforming others in lightweight and efficiency for resource-constrained scenarios.

Alkhawaldeh et al., (2025), used YOLOv3/YOLOv4 to address difficult early plant disease recognition. YOLOv4 achieved 98% accuracy, 98% mAP, 29s detection time with lower complexity, outperforming YOLOv3.

In summary, YOLO models are generally applicable to object recognition tasks in the agricultural field. However, the existing YOLOv8-DSFF model has a large parameter count (reaching 12.69 million), leading to problems such as slow recognition speed and target missing detection when deployed on mobile devices. Existing lightweight models (e.g., MobileNet-SSD) reduce computational complexity through depthwise separable convolution (*Kamath, 2024*), but suffer from insufficient small-target detection accuracy; ShuffleNet-YOLO optimizes feature fusion using channel shuffle, yet its parameter count remains relatively high *(Yu et al., 2024)*. In contrast, YOLOv8-DSFF has advantages in detection accuracy, but its computational complexity restricts deployment on mobile devices, limiting its promotion in practical field applications.

To address the above issues, this study proposes a real-time corn pest detection method based on the CSPPC lightweight module and the Wise-IoU loss function (*Liao et al., 2025*). By introducing the CSPPC lightweight module to replace the original C2f convolution module, the model's parameter count is significantly reduced (by 85.6%) *(Guo et al., 2025)*. Meanwhile, the Wise-IoU loss function is adopted to optimize the model's small-target detection performance, improving the small-target detection accuracy by 3.2%. This method can significantly enhance the model's real-time performance while ensuring detection accuracy, making it suitable for deployment on mobile devices and providing visual support for field robot detection (*Song et al., 2025*).

MATERIALS AND METHODS

YOLOv8-DSFF Object Recognition Model

Based on the original YOLOv8 model, the YOLOv8-DSFF object recognition model optimizes three key modules: feature extraction, feature fusion, and detection head. The specific optimizations are as follows: Backbone: Replaces the last C2f module with DAttention. By dynamically adjusting the position and weight of attention sampling points, it accurately focuses on the key features of small-target corn pests. Neck: Designs the C2f_SCConv module to replace all C2f modules. Through spatial reorganization and channel reorganization, SCConv reduces false detections caused by the similar colors of pests and corn leaves. Head: Replaces the original decoupled head with ASFF (Adaptive Spatial Feature Fusion). ASFF adaptively adjusts the weights of feature maps at different scales, enhancing adaptability to pests of different sizes. Through the above improvements to the YOLOv8 model, a new YOLOv8-DSFF model is obtained.

On the self-built pest dataset, the mAP@0.5 (mean Average Precision at IoU=0.5) of the YOLOv8-DSFF object recognition model reaches 93.8%, increasing by 5.9 percentage points compared with the original YOLOv8 (87.9%). The AP (Average Precision) values of all 5 pest categories are improved: among them, the AP of small-target aphids rises from 78.9% to 86.6%, and the AP of corn borers increases from 85.4% to 95.0%.

Model Parameter Metrics

FLOPs (Floating-Point Operations): Refers to the number of floating-point operations required to train a single image, and is used to measure the model's computational complexity.

Model Parameter Count: Measured by the total number of weights and biases, and is directly related to the model's storage and transmission costs.

Inference Speed (FPS, Frames Per Second): Refers to the number of images processed by the model per second, and reflects the real-time detection capability (*Zhang et al., 2025*).

Parameter Analysis

The comparison of specific parameters and characteristics among the YOLOv8, YOLOv8-DSFF, YOLOv8-Datt model with DAttention introduced alone, YOLOv8-SCConv model with SCConv introduced alone, and YOLOv8-ASFF model with the ASFF detection head is shown in detail in Table 1.

Table 1

Comparison of Parameter Analysis Among Various Models

Model	Parameters	GPLOPS	FPS	Core Characteristics
YOLOv8	3011628	8.2	4.8	Basic model with moderate balance
YOLOv8+Datt	3595039	10.1	6.1	Enhances feature extraction capability and has optimal performance-cost balance
YOLOv8+SCConv	2501919	6.5	5.4	Reduces parameters, but internal logic offsets speed advantage

Model	Parameters	GPLOPS	FPS	Core Characteristics
YOLOv8+ASFF	4379905	10.3	5.5	Parameter count and computational complexity increase significantly
YOLOv8+DSFF	12693761	37.1	10.7	Optimal accuracy but large parameter count and computational complexity

From the comprehensive comparison, it can be concluded that YOLOv8+Datt is the optimal lightweight base model. The reasons are as follows: First, YOLOv8+Datt retains the feature extraction capability of DAttention for small targets, laying the foundation for "accuracy preservation" after subsequent lightweight modification; second, its parameter count is only 3.60 M, far lower than the 12.69 M of YOLOv8-DSFF, offering great potential for lightweight modification; third, its model inference speed is 6.1 ms, close to the real-time requirements of mobile devices, and is expected to be further reduced to less than 5 ms after lightweight modification.

CSPPC Lightweight Module

While mainstream depthwise separable convolutions and group convolutions can reduce FLOPs (Floating-point Operations), they tend to lead to a sharp drop in detection accuracy. Based on the DualConv concept, the CSPPC module achieves channel-wise feature fusion by replacing the inverted residual structure and adding convolutional layers in the PConv (partial convolution) stage, thereby balancing "reducing computational complexity" and "maintaining accuracy".

PConv and PW-Conv

Partial Convolution (PConv) leverages the high similarity of feature maps between channels, performing convolution only on partial channels to extract spatial features while leaving the other channels unchanged *(Yan et al., 2025)*. Its FLOPs are lower than those of conventional convolution but higher than those of depthwise convolution and group convolution; under a specific ratio, the FLOPs of PConv are only 1/4 of those of conventional convolution.

Pointwise Convolution (PW-Conv) and PConv form a T-shaped convolution, which focuses more on the center of the feature map (consistent with the law of significant position distribution of pre-trained ResNet18 filters). Further FLOPs reduction can be achieved after decomposition *(Yan et al., 2025)*.

When PW-Conv is appended after PConv, their effective receptive fields form a T-shaped convolution—compared with conventional convolution that processes regions uniformly, this T-shaped convolution focuses more on the central position. By querying the histogram of significant positions of pre-trained ResNet18 filters, it is confirmed that the central position is most frequently the significant position, which is consistent with the characteristics of the T-shaped convolution.

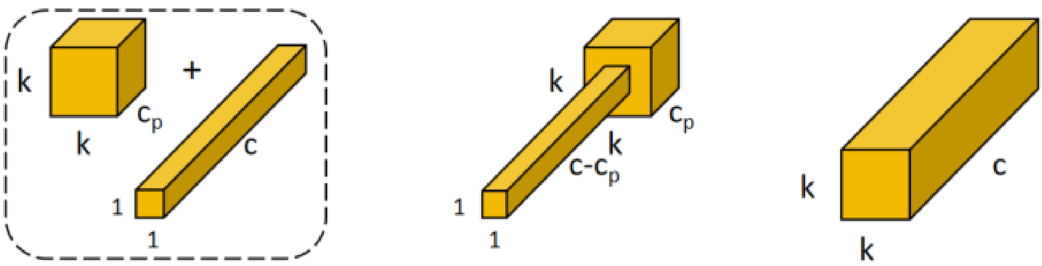


Fig. 1 - T-shaped Convolution in Partial Convolution (PConv)

Under the same input and output conditions, the FLOPs of the T-shaped convolution are higher than the sum of those of PConv (Partial Convolution) and PW-Conv (Pointwise Convolution), and they satisfy a specific numerical relationship. Decomposing the T-shaped convolution into PConv and PW-Conv can leverage the redundancy between filters, thereby achieving further FLOPs reduction.

Structural Optimization Based on the FasterNet Backbone Network

The CSPPC module is designed based on the FasterNet general backbone network. The architectural feature of FasterNet lies in the combination of hierarchical feature extraction and efficient module stacking.

Specifically, the overall architecture of a new FasterNet general backbone network adapted to CSPPC, which is composed of PConv and PWConv, is shown in Figure 2.

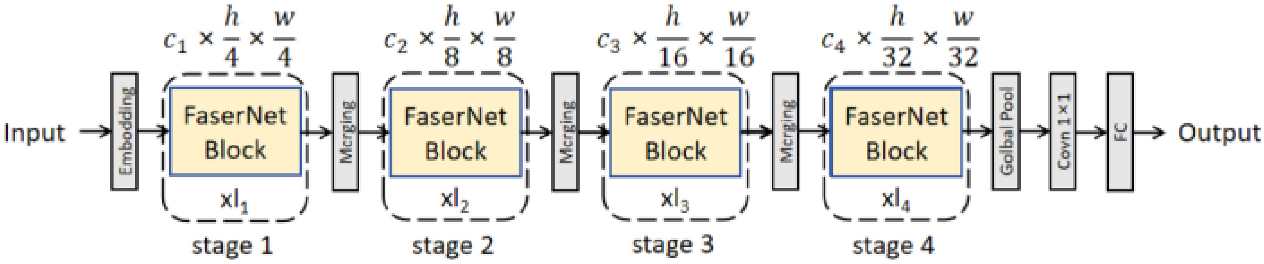


Fig. 2 - Structural Diagram of the FasterNet Backbone Network

This architecture consists of 4 stages (stage1-stage4). Before each stage, an "embedding layer" (a standard convolution with a stride of 4, used for input downsampling and channel expansion) or a "merging layer" (a standard convolution with a stride of 2, used for downsampling between stages) is arranged. Each stage stacks multiple FasterNet modules, where each module is composed of 1PConv(Partial Convolution)+2PW-Conv (Pointwise Convolution) and adopts an inverted residual structure.

On this basis, the CSPPC module optimizes and transforms the original structure by replacing the inverted residual with a residual structure, enhancing feature concatenation, and adjusting the allocation of stage modules. The structural diagram of the CSPPC module is shown in Figure 3.

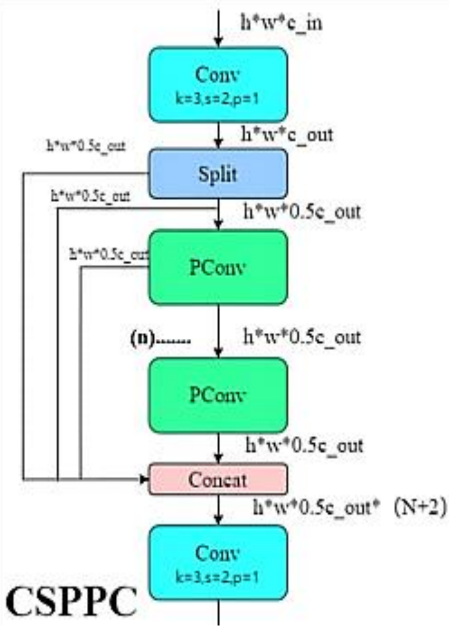


Fig. 3 - Structural Diagram of the CSPPC Module

Bounding Box Loss Function Wise-IoU

In the lightweight process, module replacement may lead to a decline in small-target localization accuracy. Therefore, the Wise-IoU loss function is introduced to address the defects of the original IoU (Intersection over Union) in small-target detection. Intersection over Union (IoU) struggles to accurately measure small-resolution targets in object detection; thus, Wise-IoU (WIoU) is proposed. Compared with the traditional IoU loss, which suffers from gradient vanishing in small-target detection, Wise-IoU evaluates anchor box quality by introducing an outlier degree (*Yang et al., 2025*). When there is a large aspect ratio deviation between the predicted box and the ground truth box (e.g., small targets such as corn aphids), the dynamic non-monotonic Focusing Mechanism (FM) enhances the gradient penalty for low-quality anchor boxes while reducing the weight of high-quality anchor boxes, preventing the model from falling into a local optimum. Compared with GIoU and DIoU, Wise-IoU maintains computational efficiency while achieving a more significant improvement in small-target localization accuracy.

The Wise-IoU loss consists of two components: classification loss and regression loss. Among them, the regression loss is constructed based on IoU (Intersection over Union) and FM(OD) (Focusing Mechanism with Outlier Degree), and its formula is as follows:

$$LWIoU = Lcls(p, p*) + \lambda \cdot [1 - IoU + FM(OD)] \quad (1)$$

where: $L_{cls}(p, p^{*})$ denotes the classification loss: the cross-entropy loss is adopted, where p represents the pest category probability predicted by the model, and $p*$ is the ground-truth category label—both are used to optimize the accuracy of category judgment. λ is the regression loss weight, which balances the training priority of classification and regression tasks. $1-IoU$ serves as the basic regression loss, ensuring the maximization of the Intersection over Union (IoU) between the predicted box and the ground-truth box. $FM(OD)$ is the dynamic penalty term, which enhances the penalty for outlier anchor boxes of small targets to improve localization accuracy.

Lightweight Convolution Module Based on YOLOv8-CSPPC

To address the issues of large parameter count in the ASFF (Adaptive Spatial Feature Fusion) and high computational complexity in the C2f_SCConv of YOLOv8-DSFF, the optimization solutions are as follows: retain the DAttention (Dilated Attention) mechanism to preserve the small-target feature extraction capability; replace the original C2f module with the CSPPC module to reduce parameters and computational complexity; adopt "focus-adjusted feature fusion" to enhance the corn pest feature fusion in the Neck section, offsetting the accuracy loss caused by lightweighting; and replace the loss function with Wise-IoU to reduce the model regression error (*Su et al., 2025*).

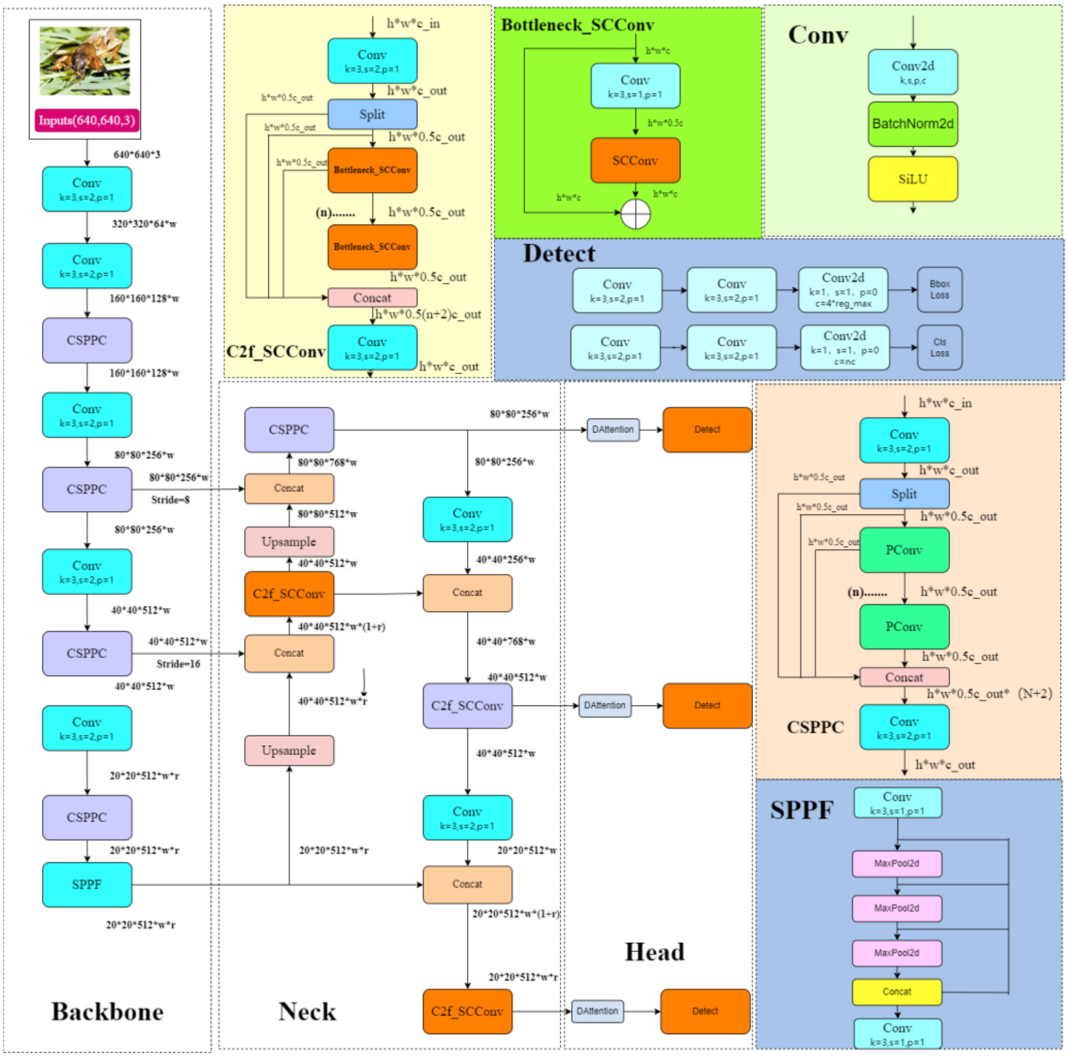


Fig. 4 - Lightweight Model Based on CSPPC

Corn Pest Dataset

To support the training of corn pest detection models, a corn pest dataset was constructed (Clarke et al., 2024) The dataset images were collected in two categories: first, for Helicoverpa armigera (cotton bollworm), Ostrinia furnacalis (Asian corn borer), Agriotes spp. (wireworms), and Gryllotalpa spp. (mole crickets), images were captured in a self-built indoor simulated environment based on their activity characteristics; second, for aphids, relevant images were selected from the open-source platform Kaggle.

Fig. 5 - Simulated Photography Based on the Living Habits of Pests

Fig. 6 - Corn Aphid Dataset from the Kaggle Open-Source Platform

After completing the collection of raw images for the dataset, the sizes of the collected corn pest images were uniformly processed to ensure that the image dimensions meet the input size requirements of the YOLOv8 model (i.e., 640 pixels × 640 pixels). In addition, to address issues such as brightness variations and occlusions that may occur during the subsequent actual deployment of the model, this study expanded the dataset by adjusting image brightness and adding color block occlusions. The specific effect of the dataset expansion is shown in Figure 7.

a) Original Image b) Brightness Adjustment c) Occlusion Addition

Fig. 7 - Effects of Image Enhancement

After the collected images underwent two preprocessing steps—normalization and data augmentation—the final total number of samples reached 2,480. The number of samples for each category is as follows: 410 for Helicoverpa armigera (cotton bollworm), 570 for Ostrinia furnacalis (Asian corn borer), 450 for Agriotes spp. (wireworms), 430 for Gryllotalpa spp. (mole crickets), and 620 for aphids. This effectively alleviates data imbalance. After annotation, the dataset was randomly split into a training set, validation set, and test set at a ratio of 8:1:1.

RESULTS

Performance Comparison Experiment.

A comparison experiment between the lightweight model, YOLOv8, and YOLOv8-DSFF was conducted under the following experimental environment: Hardware configuration - Windows 10 operating system, Intel Core i7-11800HQ processor, 8 GB RAM, and an RTX 4060 (8 GB) graphics card. Software environment - Based on Python 3.8, with the PyTorch 1.9.0 deep learning framework and CUDA 11.3.

Table 2

Comparison Table of Model Sizes

	Model	Total Parameters/10k	mAP@0. 5	GFLOPs	Detection Time/ms	FPS
1	YOLOv8	268.5	87.9	6.8 G	5.0	200
2	YOLOv8-DSFF	1269.3	93.8	37.1 G	11.2	89
3	YOLOv8-CSPPC	180.3	90.8	5.0 G	4.3	204
4	YOLOv5s-lite	270.5	88.5	6.3 G	4.9	185

As shown in Table 2, although YOLOv8-DSFF increases the mAP@0.5 (mean Average Precision@0.5) to 93.8%, its parameter count (12.693 million), GFLOPs (37.1 G), and detection time (11.2 ms) increase significantly, making it difficult to meet the requirements of mobile devices.

Through optimization with the CSPPC lightweight module and Wise-IoU loss function, YOLOv8-CSPPC reduces the parameter count to 1.803 million (a decrease of 85.6% compared with YOLOv8), achieves 5.0 G GFLOPs and a detection time of 4.3 ms, and reaches an mAP@0.5 of 90.8% (an increase of 2.9% compared with YOLOv8). Additionally, its FPS (Frames Per Second) is 204, which is 10.3% higher than that of YOLOv5s-lite. This realizes a balance between lightweight performance and accuracy.

Ablation Experiment

Ablation experiments were conducted using AP (Average Precision), Precision (P), Recall (R), and mAP (mean Average Precision) as metrics to evaluate three optimization strategies: the DAttention mechanism, the C2f_SCConv convolution module (SCConv + C2f), and the ASFF (Adaptive Spatial Feature Fusion) adaptive feature fusion detection head. The experimental results are shown in Table 3:

Table 3

Summary Table of Ablation Experiment Results

	Model	P(%)	R(%)	AP					mAP
				mo	w	cot	corn	ap	
1	YOLOv8	82.9	66.6	95.8	87.7	91.5	85.4	78.9	87.9
2	YOLOv8+DAtt	91.0	69.7	96.0	85.9	93.3	88.7	79.9	88.8
3	YOLOv8+SCConv	90.0	69.1	94.8	87.2	93.2	88.5	80.9	88.9
4	YOLOv8+ASFF	85.4	68.0	95.9	85.5	97.9	90.4	76.8	89.3
5	YOLOv8+DA+SCC	84.9	68.7	95.2	87.9	96.1	89.9	80.9	90.0
6	YOLOv8+DA+ASFF	91.3	69.3	95.8	88.9	94.6	91.0	81.7	90.4
7	YOLOv8+SCC+ASFF	89.3	68.1	96.5	88.8	96.3	89.1	81.1	90.4
8	YOLOv8-DSFF	93.8	81.9	97.1	92.2	98.0	95.0	86.6	93.8

For the experiments (Experiments 2, 3, and 4) involving the addition of a single improvement strategy, the specific setups are as follows:

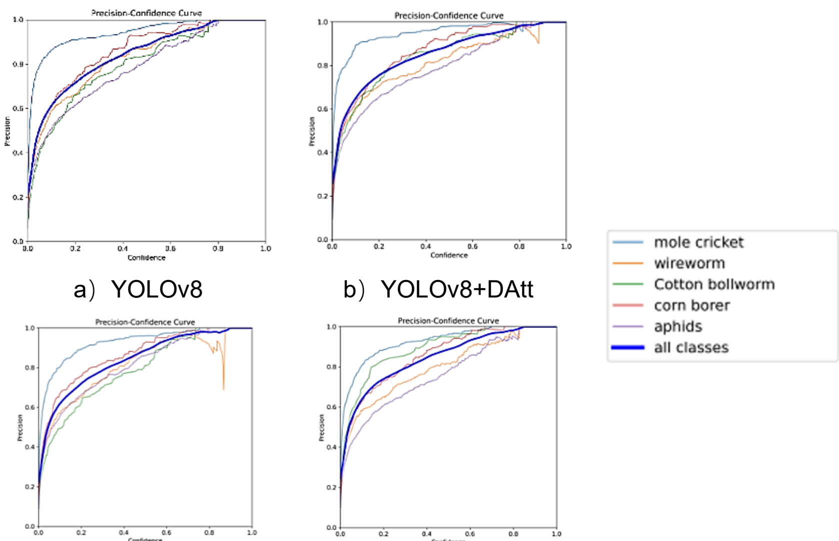
- In Experiment 2: The DAttention mechanism was added after the last C2f module in the Backbone.
- In Experiment 3: A new C2f_SCConv convolution module (composed of spatial-channel reorganization convolution) was used in the Neck section to replace the original C2f module of the model.
- In Experiment 4: The ASFF detection head was introduced in the Head section.

Compared with Experiment 1 (baseline):

- The precision of Experiment 2 increased from 82.9% to 91.0%, an improvement of 8.1%.
- The precision of Experiment 3 increased from 82.9% to 90.0%, an improvement of 7.1%.
- The precision of Experiment 4 increased from 82.9% to 85.4%, an improvement of 2.5%.

When comparing Experiments 2, 3, and 4, the order of precision (P) from highest to lowest is: DAttention > C2f_SCConv > ASFF.

Experiments 1–4 show that all models with a single improvement strategy achieve higher precision compared to the original YOLOv8. The precision values of Experiments 1–4 are illustrated in Figure 8.



a) YOLOv8 b) YOLOv8+DAtt

c)YOLOv8+SCConv d) YOLOv8+ASFF

Fig. 8 - Precision (P) Curve for Single Improvement Strategies

Experiments 5, 6, and 7 were designed with random combinations of two improvement strategies, with specific setups as follows:

- Experiment 5: Added the DAttention mechanism + replaced the original module with the C2f_SCConv convolution module.
- Experiment 6: Added the DAttention mechanism + adopted the ASFF (Adaptive Spatial Feature Fusion) adaptive detection head.
- Experiment 7: Replaced the original module with the C2f_SCConv convolution module + adopted the ASFF adaptive detection head.

Performance Comparison with the Baseline (Experiment 1)

Compared with Experiment 1 (baseline model):

- The precision of Experiment 5 increased from 82.9% to 84.9%, an improvement of 2%.
- The precision of Experiment 6 increased from 82.9% to 91.3%, an improvement of 8.4%.
- The precision of Experiment 7 increased from 82.9% to 89.3%, an improvement of 6.4%.

In-Depth Analysis of Combined Strategies

1. Effective Combination: DAttention + ASFF (Experiment 6)

The performance improvement of Experiment 6 (DAttention + ASFF), with an mAP of 90.4%, outperformed the single-strategy experiments (e. g. , Experiments 2 and 4). This is because:

- The DAttention mechanism enhances the model's feature representation capability, enabling it to capture fine-grained features of pests.
- The ASFF fuses multi-scale features via adaptive weights, optimizing the detection of pest targets of varying sizes. The synergy between these two strategies effectively improves overall detection performance.

2. Ineffective Combination: DAttention + C2f_SCConv (Experiment 5)

Experiment 5 showed a precision decline compared to the effective combined strategy. This may be attributed to computational redundancy between the spatial-channel reorganization operation of C2f_SCConv and the feature weighting mechanism of DAttention. Such redundancy could lead to overfitting, thereby limiting precision improvement.

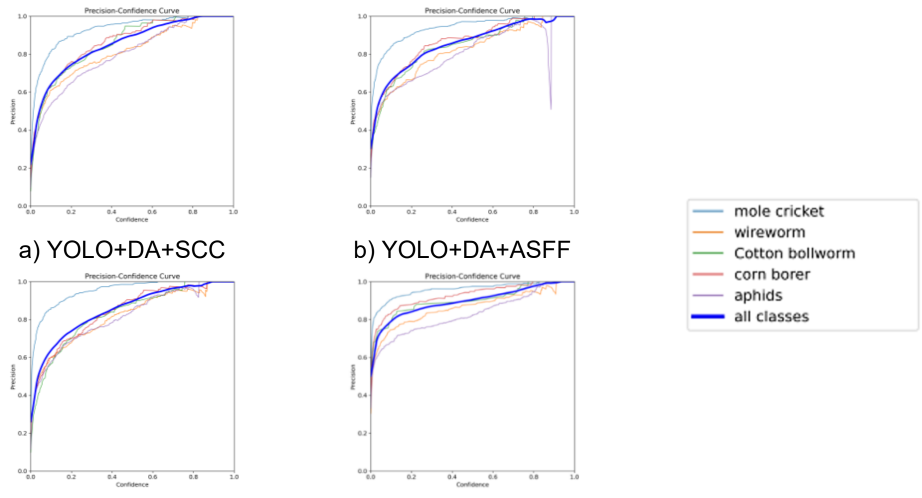
3. Comparison with Single-Strategy Experiments

While all two-strategy combinations (Experiments 5–7) achieved higher precision than the baseline (Experiment 1), their performance varied when compared to the corresponding single-strategy experiments:

- Experiment 5 (DAttention + C2f_SCConv) vs. single strategies: Precision decreased by 6.1% compared to Experiment 2 (DAttention alone) and by 5.1% compared to Experiment 3 (C2f_SCConv alone).
- Experiment 6 (DAttention + ASFF) vs. single strategies: Precision increased by 0.3% compared to Experiment 2 (DAttention alone) and by 5.9% compared to Experiment 4 (ASFF alone).
- Experiment 7 (C2f_SCConv + ASFF) vs. single strategies: Precision decreased by 0.7% compared to Experiment 3 (C2f_SCConv alone) but increased by 3.9% compared to Experiment 4 (ASFF alone).

Key Conclusion from Experiments 4–6

Although all two-strategy combinations outperformed the original model, some combinations showed lower precision than their corresponding single-strategy counterparts. The precision values of Experiments 5, 6, and 7 are illustrated in Figure 9(a), (b), and (c), respectively.



a) YOLO+DA+SCC b) YOLO+DA+ASFF

c) YOLO+SCC+ASFF d) YOLO-DSFP

Fig. 9 - Precision (P) Curves for Random Combinations of Two Improvement Strategies

Experiments 1–7, the YOLOv8-DSFF model shows the following precision improvements:

- Compared with the original model (Experiment 1), its precision increased by 10.9%.
- Compared with the single-improvement-strategy experiments (Experiments 2–4), its precision increased by 2.8%, 3.8%, and 8.4% respectively.
- Compared with the random two-strategy combination experiments (Experiments 5–7), its precision increased by 8.9%, 2.5%, and 4.5% respectively.

The precision of YOLOv8-DSFF is illustrated in Figure 9(d). The results indicate that the YOLOv8-DSFF model achieves relatively high precision. However, Precision (P) cannot reflect the accuracy of each individual pest category; therefore, further evaluation using mAP (mean Average Precision) is required to draw a comprehensive conclusion.

AP Comparison Analysis Across Pest Categories. This section analyzes the Average Precision (AP) of five pest categories—mole cricket, wireworm, cotton bollworm, corn borer, and aphids—focusing on the performance of single improvement strategies.

AP Performance of Single Improvement Strategies (vs. Baseline Experiment 1)

Experiments 2, 3, and 4 (each with one single improvement strategy) are compared against Experiment 1 (the original YOLOv8 model), with results as follows:

1. Experiment 2 vs. Experiment 1

- Category-specific AP changes:
  Decreases: mole cricket (-0.2%), wireworm (-1.8%);
  Increases: cotton bollworm (+1.8%), corn borer (+3.3%), aphids (+1%).
- Overall mAP change: Increased from 87.9% to 88.8% (an improvement of 0.9%).

2. Experiment 3 vs. Experiment 1

- Category-specific AP changes:
  Decreases: mole cricket (-1%), wireworm (-0.5%);
  Increases: cotton bollworm (+1.7%), corn borer (+3.1%), aphids (+2%).
- Overall mAP change: Increased from 87.9% to 88.9% (an improvement of 1%).

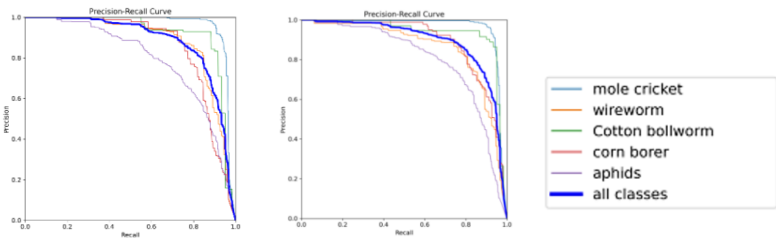
3. Experiment 4 vs. Experiment 1

- Category-specific AP changes:
  Decreases: wireworm (-2.2%), aphids (-2.1%);
  Increases: mole cricket (+0. 1%), cotton bollworm (+6.4%), corn borer (+5%).
- Overall mAP change: Increased from 87.9% to 89.3% (an improvement of 1.4%).

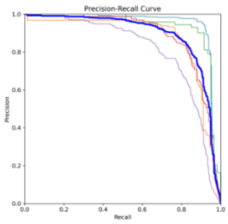
Key Findings from Single Improvement Strategy Analysis (Experiments 1–4)

- Overall mAP improvement: All three single improvement strategies outperformed the original YOLOv8 model in terms of overall mAP, confirming their effectiveness in enhancing general detection performance.
- Category-specific AP declines: Despite overall improvements, each strategy caused AP decreases for certain individual pest categories.
- Consistent decline in wireworm AP: Notably, all three strategies led to a decrease in wireworm AP-this suggests potential challenges in adapting these strategies to wireworm detection (e.g., wireworm’s small size or low contrast with corn plants may conflict with the strategies’ feature extraction logic).

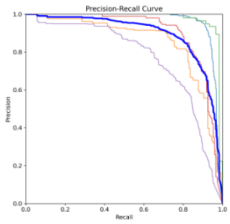
The mAP values of Experiments 1, 2, 3, and 4 are illustrated in Figure 10.



a） YOLOv8　　b） YOLO+DAtt



c) YOLO+SCConv



d) YOLO+ASFF

Fig. 10 - mAP Curve for Single Improvement Strategies

Analysis of mAP for Random Combinations of Two Improvement Strategies. Experiments 5, 6, and 7 were designed to test random combinations of two improvement strategies. Below is their mAP (mean Average Precision) performance analysis, including comparisons with the original YOLOv8 model and the single-strategy experiments (Experiments 2–4):

Comparison with the Original YOLOv8 Model

When compared to the original YOLOv8 model (baseline), all two-strategy combinations showed mAP improvements:

- Experiment 5: mAP increased by 2.1%.
- Experiment 6: mAP increased by 2.5%.
- Experiment 7: mAP increased by 2.5%.

Comparison with Single-Strategy Experiments (Experiments 2–4)

1. Experiment 5 (DAttention + C2f_SCConv)

Experiment 5 combines the DAttention mechanism and C2f_SCConv convolution module. Its performance relative to the corresponding single-strategy experiments is as follows:

Compared with Experiment 2 (DAttention alone):
Overall mAP increased by 1.2%.
For category-specific AP:
Decrease: mole cricket (-0.2%);
Increases: wireworm (+2%), cotton bollworm (+0.2%), corn borer (+4.6%), aphids (+1.0%).
Compared with Experiment 3 (C2f_SCConv alone):
AP of all 5 pest categories increased (no specific declines reported).

2. Experiment 6 (DAttention + ASFF)

Experiment 6 combines the DAttention mechanism and ASFF (Adaptive Spatial Feature Fusion) adaptive detection head. Its performance relative to the corresponding single-strategy experiments is as follows:

- Compared with Experiment 2 (DAttention alone):
  For category-specific AP:
  Decrease: mole cricket (no specific value reported);
  Increases: wireworm (+3.0%), cotton bollworm (+1.3%), corn borer (+2.3%), aphids (+1.8%).
- Compared with Experiment 4 (ASFF alone):
  For category-specific AP:
  Decrease: mole cricket (no specific value reported);
  Increases: wireworm (+3.4%), cotton bollworm (+1.4%), corn borer (+0.6%), aphids (+4.9%).

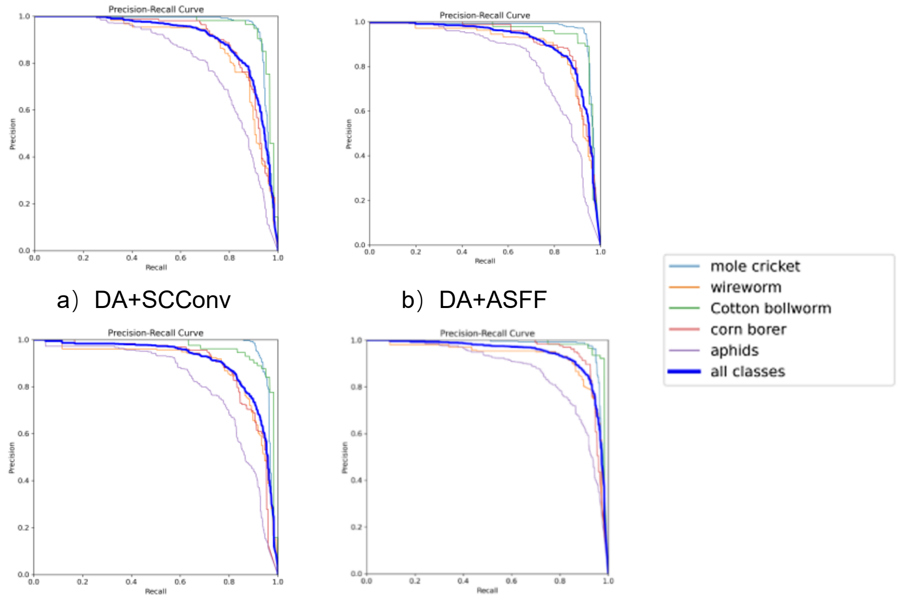
3. Experiment 7 (C2f_SCConv + ASFF)

Experiment 7 combines the C2f_SCConv convolution module and ASFF adaptive detection head. Its performance relative to the corresponding single-strategy experiments is as follows:

- Compared with Experiment 3 (C2f_SCConv alone):
  AP of all 5 pest categories increased, with gains of +1.7% (mole cricket), +1.6% (wireworm), +3.1% (cotton bollworm), +0.6% (corn borer), and +0.2% (aphids).

- Compared with Experiment 4 (ASFF alone):
  For category-specific AP:
  Decreases: cotton bollworm (no specific value reported), corn borer (no specific value reported); Increases: mole cricket (+0.6%), wireworm (+3.3%), aphids (+4.3%).

While random combinations of two improvement strategies achieved overall mAP gains compared to single-strategy experiments, they still resulted in AP declines for certain individual pest categories. This indicates that even combined strategies cannot guarantee AP improvements for all categories.
The mAP values of Experiments 5, 6, and 7 are illustrated in Figure 11(a), (b), and (c), respectively.



a) DA+SCConv b) DA+ASFF
c) SCC+ASFF d)YOLOv8-DSFF

Fig. 11 - mAP Curves for Random Combinations of Two Improvement Strategies

Experiment 8 evaluates the YOLOv8-DSFF model, with its performance compared against Experiments 1–7 (covering the original model, single improvement strategies, and random two-strategy combinations). The results are as follows:

1. Comparison with Experiment 1 (Original YOLOv8 Model)

- Overall mAP: Increased by 5.9% compared to Experiment 1.
- Category-specific AP: All five pest categories showed AP improvements, with gains of 1.3%, 4.5%, 6. 5%, 6.3%, and 6.7% respectively (corresponding to mole cricket, wireworm, cotton bollworm, corn borer, and aphids).

2. Comparison with Single Improvement Strategy Experiments (Experiments 2–4)

Compared to Experiments 2–4 (each with one single improvement strategy):

- Overall mAP: The YOLOv8-DSFF model (Experiment 8) achieved mAP increases of 5%, 4.9%, and 4.5% respectively.
- Category-specific AP: All pest categories in Experiment 8 showed higher AP than those in Experiments 2–4 (no category-specific declines were observed).

3. Comparison with Random Two-Strategy Combination Experiments (Experiments 5–7)

Compared to Experiments 5–7 (random combinations of two improvement strategies):

- Overall mAP: The YOLOv8-DSFF model (Experiment 8) achieved mAP increases of 3.8%, 3.4%, and 3.4% respectively.
- The mAP value of the YOLOv8-DSFF model (Experiment 8) is illustrated in Figure 11(d).

To summarize, after integrating the three strategies, the Precision reaches 93.8% (a 10.9% increase compared to the baseline), and the mAP (mean Average Precision) reaches 93.8% (a 5.9% increase compared to the baseline). The AP (Average Precision) of all 5 pest categories has improved: the AP of aphids rises from 78.9% to 86.6%, and the AP of corn borers rises from 85.4% to 95.0%. There is no precision decline at the category level, which verifies the effectiveness of the multi-strategy synergy. The recognition effect of the model is shown in Figure 12.

a) Cotton Bollworm b) Wireworm c) Mole Cricket d) Aphids

Fig. 12 - Recognition effect of the model

Mobile Deployment of Lightweight Models

Optimization and Format Conversion of the YOLOv8-CSPPC Model

To deploy the YOLOv8-CSPPC model on mobile devices via NCNN, model quantization and format conversion are essential, as detailed below:

1. Convert. pt Format to ONNX Format

The original YOLOv8-CSPPC model is typically saved in PyTorch's .pt format. Converting it to the ONNX (Open Neural Network Exchange) format improves cross-platform compatibility and optimizes inference speed. The conversion steps are:

- Use the ultralytics library to load the pre-trained .pt model.
- Call the export method and specify the output format as ONNX.
- Verify the converted ONNX model using ONNXRuntime to ensure its functionality and correctness (e.g., checking if inference results match the original .pt model).

2. Quantize FP32 Model to INT8 Format

The original YOLOv8-CSPPC model uses FP32 (32-bit floating-point) precision, which results in large model size and high computational complexity—making it incompatible with resource-constrained Android devices. To address this, NCNN's ncnnoptimize tool is used to quantize the model to INT8 (8-bit integer) precision. This quantization:

- Reduces model storage requirements (e.g., cutting the model size by approximately 75% compared to FP32).
- Lowers computational demands, significantly improving inference speed on CPU-only mobile hardware.
- Reduces power consumption, which is critical for prolonged use of mobile or battery-powered devices (e.g., field pest detection with smartphones).

Mobile Deployment and Testing on Android Devices

Deployment Steps: Import the NCNN-compatible model files (the .param file containing the network structure and the .bin file containing the quantized weights) into the target Android app project. Enable USB Debugging on the Android device to facilitate app installation and testing.

Testing and Results: When the app is run on the Android device, it provides the following real-time information: Live video feed from the device's camera; Overlaid detection results (bounding boxes around detected pests); Metadata including the deployed model name (YOLOv8-CSPPC), inference frame rate (FPS), and target confidence scores.

The detection performance is visualized in Figure 13.

Fig. 13 - Mobile Device Detection Test

CONCLUSIONS

This paper proposes a real-time corn pest detection method based on the CSPPC lightweight module and the Wise-IoU loss function. By introducing the CSPPC lightweight module to replace the original C2f convolution module, the model's parameter count is significantly reduced by 85.6%. At the same time, the Wise-IoU loss function is used to optimize the model's small target detection performance, increasing the small target detection accuracy by 3.2%. Experimental results show that the proposed YOLOv8-CSPPC model outperforms the YOLOv8 and YOLOv8-DSFF models in metrics such as parameter count, GFLOPs, and detection time. Moreover, its average detection accuracy reaches 90.8%, which is 1.9% higher than that of the YOLOv8 model and only 3% lower than that of the YOLOv8-DSFF model. In addition, after model quantization implemented through the NCNN framework, the model size is reduced by 75%, and the model is successfully deployed on mobile devices, providing an effective solution for real-time monitoring of corn pests in the field.

ACKNOWLEDGEMENT

This work was supported by the Excellent Young Teachers Basic Research Support Program of Heilongjiang Province (Grant No. YQJH2023218), the Basic Scientific Research Business Expenses Project of Colleges and Universities in Heilongjiang Province (Grant No. 2023-KYYWF-0592), and the Young Innovative Talent Training Program Project of Jiamusi University (Grant No. JMSUQP2023011).

REFERENCES

[1] Alhwaiti Yousef, Khan Muntazir, Asim Muhammad, Siddiqi Muhammad Hameed, Ishaq Muhammad, Alruwaili Madallah. (2025). Leveraging YOLO deep learning models to enhance plant disease identification. *Scientific Reports*, Vol.15, No.1, pp. 7969-7969, London, UK.

[2] Bohari Zul Hasrizal, Mohd Zaid Hazim Hafizullah, Mohd Nasir Mohamad Na'im, Sulaima Mohamad Fani, Ahmad Emy Zairah, Abdullah Ahmad Zaidi, Isa Muzamir (2025). Development of Real-Time Crop Sorting System Based on Deep Learning and ESP32-CAM Integration. *Journal of Power and Energy Engineering*, Vol. 13, No. 08, pp. 49-60, United States.

[3] Chen, Y.H. (2013). Forecast on the development prospect of agricultural mechanization in China （中国农业机械化发展前景预估）. *Hunan Agricultural Machinery*, Vol. 40, No.09, pp.10-11, Changsha, China.

[4] Chu, T., Zha, H., Wang, Y., Yao Z., Wang X., Wu C., Liao J. (2025). MaizeStar-YOLO: Precise Detection and Localization of Seedling-Stage Maize. *Agronomy*, Vol. 15, No. 8, p. 1788, Basel, Switzerland.

[5] Clarke, A., Yates, D., Blanchard, C., Islam M.Z., Ford R., Rehman S., Walsh R. (2024). The Effect of Dataset Construction and Data Pre-Processing on the eXtreme Gradient Boosting Algorithm Applied to Head Rice Yield Prediction in Australia. *Computers and Electronics in Agriculture*, Vol. 219, p. 108716, Amsterdam, Netherlands.

[6] Ganapathy Madasamy Raja, Periasamy Selvaraju, Pugazhendi Pathmanaban, Antony Christy Grace Manuvel (2025). YOLOv11n for precision agriculture: lightweight and efficient detection of guava defects across diverse conditions. *Journal of the Science of Food and Agriculture*, Vol. 105, No. 11, pp. 6182-6195, United States.

[7] Guo Xiaoxia, Li Dongpeng, Lan Chaofeng, Han Chuang. (2025). Small Target Detection Algorithm Based on CSF-YOLOv8. Signal, *Image and Video Processing*, Vol.19, No.11, p. 916, New York, United States.

[8] Hao Shengnan, Gao Erjian, Ji Zhanlin, Ganchev Ivan (2025). BCSYOLO: Research on Corn Leaf Disease and Pest Detection Based on YOLOv11n. *Applied Sciences*, Vol.15, No.15, pp. 8231-8231, Basel, Switzerland.

[9] Kamat Pooja, Gite Shilpa, Chandekar Harsh, Dlima Lisanne, Pradhan Biswajeet (2025). Multi-class fruit ripeness detection using YOLO and SSD object detection models. *Discover Applied Sciences*, Vol.7, No.9, pp. 931-931, Cham, Switzerland.

[10] Kamath, V., Renuka A. (2024). Investigation of MobileNet-SSD on Human Follower Robot for Stand-Alone Object Detection and Tracking Using Raspberry Pi. *Cogent Engineering*, Vol.11, No.1, London, UK. https://doi.org/10.1080/23311916.2024.2333208

[11] Liao Dan, Bi Rengui, Zheng Yubi, Hua Cheng, Huang Liangqing, Tian Xiaowen, Liao Bolin. (2025). LCW-YOLO: An Explainable Computer Vision Model for Small Object Detection in Drone Images. *Applied Sciences*, Basel, Switzerland, Vol.15, No.17, p. 9730, Basel, Switzerland.

[12] Peng, G., Wang, K., Ma, J., Cui B., Wang D. (2025). AGRI-YOLO: A Lightweight Model for Corn Weed Detection with Enhanced YOLO v11n. *Agriculture*, Vol.15, No.18, p.1971 Basel, Switzerland.

[13] Song, P., Li, Z. D., Yang, M., Cui Jiale, Feng Hui, Zhai Ruifang, Yang Wanneng (2025). Research status and prospects of crop phenotyping robots （作物表型机器人研究现状与展望）. *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 56, No.03, pp.1-17, Beijing, China.

[14] Su, B., Zhu, Y., Lin, Y. (2025). Pest-YOLOv8: Enhanced Detection for Small and Complex Agricultural Pests Using Triple Attention and Wise-IoU. *Journal of Plant Diseases and Protection*, Vol. 132, No.5, p.135, Berlin, Germany.

[15] Wang, Q., Liu, Y., Zheng, Q., Tao R., Liu Y. (2025). SMC-YOLO: A High-Precision Maize Insect Pest-Detection Method. *Agronomy*, Vol.15, No.1, p.195, Basel, Switzerland.

[16] Wu, W., Zhang, J., Zhou, G., Zhang Y, Wang J., Hu L. (2024). ESG-YOLO: A Method for Detecting Male Tassels and Assessing Density of Maize in the Field. *Agronomy*, Vol.14, No.2, Basel, Switzerland.

[17] Yan Pingping, Qi Xiangming, Jiang Liang (2025). LI-YOLOv8: Lightweight small target detection algorithm for remote sensing images that combines GSConv and PConv. *PloS One*, Vol. 20, No. 5, e0321026, San Francisco, USA.

[18] Yang, H.X., Chen, Y., Pei, G.Q., Qian Xueying, Li Peiyao; Zhu Caiying, Xia Qian, Liu Zigao; Wu Wendou (2025). A grading method for Yunnan small-grain coffee green beans based on the lightweight YOLOv8-FasterBlock model （基于轻量化 YOLOv8-FasterBlock 模型的云南小粒咖啡生豆分级方法）. *Food Science*, Vol. 46, No. 04, pp. 268-277, Beijing, China.

[19] Yu Yan Nan, Xiong Chun Lin, Yan Ji Chi, Mo Yong Bin, Dou Shi Qing, Wu Zuo Hua, Yang Rong Feng. (2024). Citrus Pest Identification Model Based on Improved ShuffleNet. *Applied Sciences*, Vol. 14, No.11, Basel, Switzerland.

[20] Zhang, Y., Liu, J., Li, S., Feng S., Sun Z., Cui Y., Liu H., Zhang J., Cheein Auat F., Guo W. (2025). ESM-YOLOv11: A Lightweight Deep Learning Framework for Real-Time Peanut Leaf Spot Disease Detection and Precision Severity Quantification in Field Conditions. *Computers and Electronics in Agriculture*, Vol. 238, p. 110801, Amsterdam, Netherlands.