DETECTION METHOD OF COTTON COMMON PESTS AND DISEASES BASED ON IMPROVED YOLOv5S
/
基于改进 YOLOV5S 的棉花常见病虫害检测方法

Yulong WANG[1,2,3,#], Fengkui ZHANG[1,2,3,#], Rina YANGDAO[1,2,3], Ruohong HE[1,2,3], Jikui ZHU[1,2,3], Ping LI[1,2,3,*]
[1)] College of Mechanical and Electrical Engineering, Tarim University, Alar, Xinjiang, 843300, China
[2)] Xinjiang Production and Construction Corps (XPCC) Key Laboratory of Utilization and Equipment of Special Agricultural and Forestry Products in Southern Xinjiang, Xinjiang, China
[3)] Modern Agricultural Engineering Key Laboratory at Universities of Education Department of Xinjiang Uygur Autonomous Region, Xinjiang, China
[#]These authors contributed equally to this work and should be considered co-first authors.
[*]*Corresponding Author. Email: lpdyy716@163.com*
DOI: https://doi.org/10.35633/inmateh-76-18

Keywords: *Cotton, pest and disease, YOLOv5s, lightweight, target detection*

ABSTRACT
To address the low recognition accuracy and slow detection speed of cotton leaf pests and diseases in natural environments, a detection method based on an improved YOLOv5s model was proposed. The enhanced model integrates the Ghost module and the C3Faster module to increase inference speed and reduce model complexity, achieving lightweight performance without significantly compromising accuracy. To counteract the tendency of common cotton pest and disease features to be lost in complex natural scenes, a Coordinate Attention (CA) mechanism was introduced to improve the network's recognition and localization capabilities. The parameters, FLOPs, and weight file size of the improved model were reduced to 65.5%, 66.2%, and 67.1% of those of the original YOLOv5s model, respectively. On a self-built dataset, the improved YOLOv5s model achieved a mean average precision (mAP) improvement of 10.5%, 0.2%, and 0.4% compared to YOLOv4, YOLOv5s, and YOLOv7, respectively. The model was deployed on a Jetson Orin NX development board with CUDA acceleration, achieving a real-time detection speed of 76.3 frames per second.

中文摘要
针对自然环境中棉花叶片病虫害识别精度低、识别速度慢的问题，提出了一种基于改进YOLOv5s的检测方法。该方法基于YOLOv5s模型，通过结合Ghost模块和C3Faster模块来提高推理速度，实现模型轻量化，而不会显著影响准确性。由于棉花常见病虫害的特征在自然环境中很容易丢失，因此添加了CA注意机制模块，以帮助轻量级网络提高识别和定位能力。改进后的模型参数、FLOP和权重文件大小分别为YOLOv5s的65.5%、66.2%和67.1%。将自建数据集与YOLOv4、YOLOv5s和YOLOv7进行比较，结果表明，改进的YOLOv5模型的平均准确率分别提高了10.5%、0.2%和0.4%。改进后的模型部署在Jetson Orin NX开发板上，并使用CUDA加速。加速模型检测速度为76.3帧/秒。

INTRODUCTION
Cotton is the seed fiber of malvaceous cotton plants in subtropical regions and is one of the most important economic crops in the world. China's cotton planting area and yield account for 9.82% and 23.76% of the world, respectively, with Xinjiang having the largest cotton planting area and yield. In 2024, China's cotton planting area reached 2,838.3 thousand hectares, with a total output of 6.164 million tons. Xinjiang accounted for 2,447.9 thousand hectares of the planting area and 5.686 million tons of the total yield, representing 86.2% and 92.2% of the national totals, respectively. The yield and quality of cotton are affected mainly by pests and diseases. The common pests and diseases of cotton in Xinjiang mainly include yellow wilt disease, aphids, and brown spots *(Chen et al., 2018; Lou et al., 2018; Zhao et al., 2023).* Currently, cotton pests and diseases in Xinjiang mainly rely on the application of plant-protection machinery. During operation, the precision of spraying is poor, making it difficult to accurately and effectively prevent different types of pests and diseases.

Yulong WANG, Master Student; Fengkui ZHANG, lecturer; Rina YANGDAO, Master Student; Ruohong HE, Master Student; Jikui ZHU, lecturer; Ping LI, Professor.

As the computer technology progresses, prevention and control of cotton pests and diseases have evolved from being initially determined by experienced experts to using computer technology and deep learning algorithms to determine pest and disease varieties and control measures. To apply the target detection technology to plant protection machinery, the first step is to solve the problems of a large model, slow recognition speed, and low recognition accuracy.

The YOLO algorithm has been widely used in the detection of crop pests and diseases due to its fast inference speed and high accuracy. YOLO is highly suitable for real-time agricultural applications by directly predicting object classes and positions on feature maps (*Ganesan and Chinnappan, 2022; Amarasingam et al., 2022; Apacionado and Ahamed, 2023; Wang et al., 2023a; Cruz et al., 2022*). In recent years, various YOLO-based models have been applied to crops such as rice, corn, tomato, and cotton (*Masykur et al., 2023; Chen and Yang, 2023; Azath et al., 2021; Susa et al., 2022; Sangaiah et al., 2024*). To improve the detection accuracy in complex field conditions, several enhancements have been proposed. *Shi et al. (2023)* introduced the SimAM attention mechanism into YOLOv5s, improving the accuracy except the model size. *Liu et al. (2023)* upgraded YOLOv7 with deformable convolutions, SENet, and Focal-EIOU loss, enhancing the performance at the cost of increased complexity. *Li and Wang (2020)* used image pyramids in YOLOv3 to boost multi-scale detection without reducing model weight. *Zhang et al. (2023)* applied CBAM to YOLOv7 for cotton disease detection to increase the accuracy without reducing FLOPs. *Wang et al. (2023b)* proposed YOLOv4-G using GhostNet and Ghost modules to improve the accuracy, while the memory usage remained high.

To address the problems of large amounts of calculation and difficult feature extraction in the detection model, an improved model based on YOLOv5s was proposed. The lightweight is convenient for the model to be deployed on mobile devices, and the attention mechanism is added to improve the ability to extract important features of the image, providing a theoretical basis for the automation prevention and control of cotton common pests and diseases.

MATERIALS AND METHODS

Image acquisition

Common pests and diseases, such as cotton yellow wilt disease, aphids, and brown spots were selected as the research objects. The cotton pest and disease image dataset were collected from June to September 2023 at the cotton experimental station of Tarim University, using an iPhone14 camera. To meet the diversity of the natural environment, images of different weather conditions were collected at different times, including sunny days, windy days, morning, noon, etc. A total of 2322 cotton pest and diseases were collected. To improve the robustness of model training and avoid overfitting, the collected pictures were horizontally flipped and randomly cropped to expand the dataset. Finally, 4232 pictures of cotton pests and diseases were obtained. The expanded data were screened and pictures with excessive blur, inconsistent features, and excessive differences in actual shooting were deleted. A total of 4000 pictures were selected as the image dataset for cotton pests and diseases. There were 797 pictures of yellow wilt disease, 1065 pictures of aphids, 1422 pictures of brown spots, and 716 pictures of healthy leaves. Some of the collected images are shown in Fig. 1.

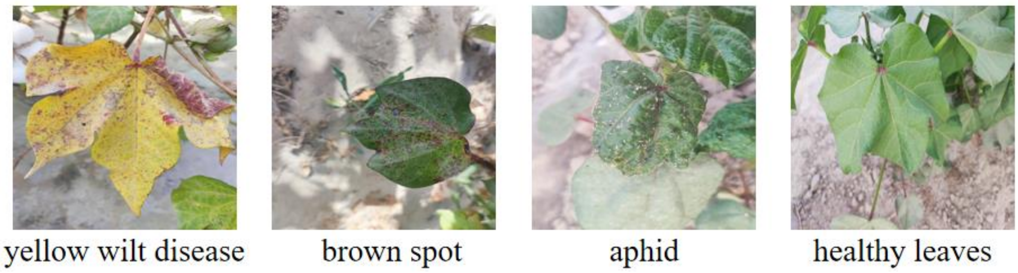


Fig. 1 - Pictures of cotton common pests and diseases

Dataset producing

The cotton pest and disease dataset were labeled with Libeling. Libeling is an open-source data labeling tool that can label tow formats: VOC and YOLO. These 4000 images in the dataset were manually labeled as yellow wilt disease, aphid, brown spot, and healthy leaves. The storage format is a VOC tag format and the file format is an xml file. The code was used to convert the xml format annotation file into a txt format annotation file, which was randomly divided into a training set and a verification set at a proportion of 80%. There were 3203 pictures in the training set and 797 in the verification set.

YOLOv5s algorithm network model

YOLOv5s is a target detection network model in the YOLOv5 series with small model parameters. Because of its small model volume and fast inference speed, it is particularly suitable for real-time target detection tasks deployed on mobile devices or edge computing devices. The model consists of four parts: the input end, Backbone Network, Feature fusion Network Neck and Head Network. In this study, enhancements to the YOLOv5s model are presented, and the architecture of the improved network is illustrated in Fig. 2.

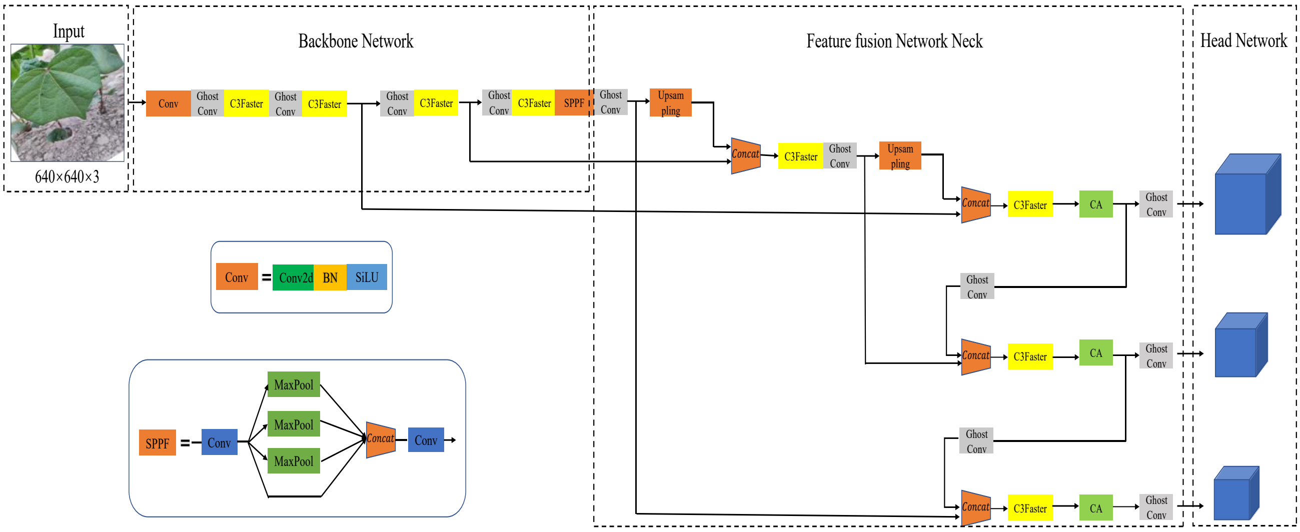


Fig. 2 - Improved YOLOv5s network model structure

Note: The Conv module consists of Conv2d + BN + SiLU, where BN denotes Batch Normalization and SiLU is the activation function. Concat represents tensor concatenation, SPPF stands for Spatial Pyramid Pooling-Fast, and Maxpool denotes the Max Pooling operation. Ghost Conv is the convolution module within the Ghost module used in this study, highlighted with a gray background in the figure. C3Faster refers to the C3Faster module employed in this study, highlighted with a yellow background in the figure. CA represents the CA attention mechanism module used in this study, highlighted with a green background in the figure.

The input end is primarily used for adaptive anchor frame calculation, adaptive image scaling, and mosaic data augmentation *(Passos et al., 2023)*. Mosaic data enhancement is an extension of YOLOv4, which can increase data diversity and model robustness and improve small target detection performance by splicing, flipping, and color gamut changes on four random images. The main structure of the backbone includes Conv, C3, and SPPF modules. After the image enters the neural network through the input end, the backbone network extracts feature and continuously narrows the feature map. The role of the feature fusion Network Neck is to obtain relatively shallow features from the Backbone Network and then stitch them together with deep semantic features. The Head Network contains three prediction branches, which extract feature information to predict targets of different sizes and obtain the category, confidence, and location information of the predicted targets *(Lan et al., 2024).*

Introducing lightweight modules Ghost and C3Faster

In practical applications, it is often necessary to deploy the model on mobile devices to achieve real-time detection. Mobile devices do not possess the computing power of computers. At this point, it is necessary to reduce the calculation time and improve the calculation speed of the model to successfully deploy the model. Based on the YOLOv5s model, all convolutions except the first ordinary convolution are replaced by Ghost Conv, and all C3 modules in the original model are replaced by C3Faster to realize the lightweight of the model and improve the calculation speed while ensuring accuracy.

The Ghost module *(Zeiler and Fergus, 2014)* is a novel plug-and-play component. Since Ghost convolution constitutes the core of the Ghost module, this paper exclusively employs Ghost convolution and refers to the module accordingly. The Ghost module is distinguished by its ability to generate more feature maps using fewer parameters. Initially, intrinsic feature maps are produced through conventional convolution. Subsequently, Ghost feature maps are generated via linear transformation *(Li et al., 2019).* Finally, the intrinsic feature maps and the Ghost feature maps are concatenated to form the final output. The specific structures are shown in Fig. 3.

FasterNet is used to improve the feature expression ability and receptive field coverage while maintaining light weight and high speed *(Shi et al., 2024)*. The network consists of four stages, and its architecture is illustrated in Fig. 4. Based on the YOLOv5s model, the C3Faster module is proposed. It replaces all C3 modules of the original model with C3Faster to reduce the redundancy of calculations and memory access, thereby improving the efficiency of the neural network. It also improves the efficiency of the floating-point operation per second of the model through partial convolution PConv, thereby achieving high-speed operation on mobile devices while ensuring accuracy.

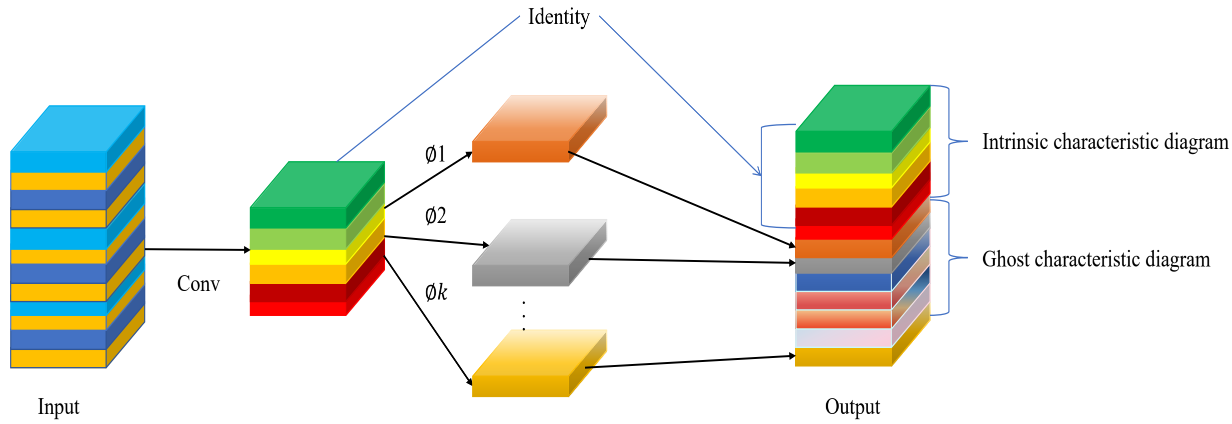


Fig. 3 - Ghost module

Note: Conv represents standard convolution, Ø denotes linear transformation operations, and Identity signifies the identity mapping.

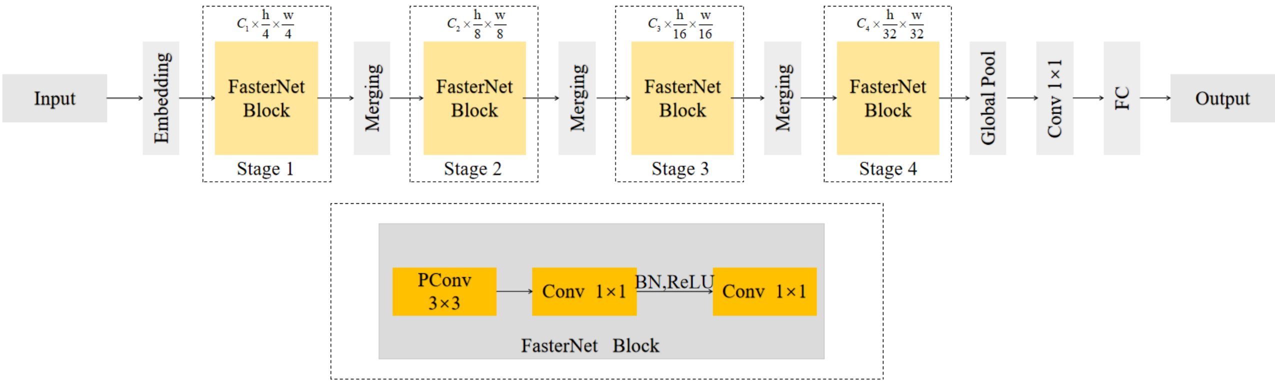


Fig. 4 - FasterNet network architecture

Note: C_x denotes the number of image channels; h represents image height; w represents image width; FasterNet Block refers to a specific module; PConv stands for partial convolution; Conv denotes standard convolution; BN represents Batch Normalization; and ReLU signifies the activation function.

Introducing CA attention mechanism module

In the natural environment, cotton leaves grow densely, and the background is complex. It is difficult for the YOLOv5 model to extract important features of pests and diseases, and the detection results will be misjudged or unrecognized. To improve the extraction effect of the YOLOv5s model on small features, in the neck part of the feature fusion network, a CA (Hou et al., 2021) attention mechanism module was added after other C3 2_1 modules to help the lightweight network rise points. A flowchart of the CA attention mechanism module is shown in Fig. 5(a).

To capture remote spatial interaction with precise location information, global average pooling is decomposed. The specific formula is as follows:

$$z_c^h(h) = \frac{1}{W} = \sum_{0 \le i < W} x_c(h,i) \tag{1}$$

$$z_c^w(w) = \frac{1}{H} = \sum_{0 \le j < H} x_c(j,w) \tag{2}$$

where z_c^h is the height feature diagram representing the output of channel *c*, z_c^w is the width feature diagram representing the output of channel *c*, $x_c(h,i)$ is the input of channel *c* along the *h* direction, and $x_c(j,w)$ is the input of channel *c* along the *w* direction. An input feature diagram with a size of C×H×W is collected in the X- and Y-directions, respectively, and those with a size of C×H×1 and C×1×W are generated, respectively, as shown in Fig. 5(b).

After obtaining the two-feature diagrams, the generated C×H×1 and C×1×W feature diagrams were transformed, spliced, and calculated as follows:

$$f = \delta(F_1([Z^h, Z^w])) \tag{3}$$

where δ is a nonlinear activation function and F_1 is a convolution transformation function. After splicing z^h and z^w, the feature map shown in Fig. 5(b) was generated. F_1 operation (using a 1×1 convolution kernel for dimensionality reduction) and activation operation are performed to generate feature diagram f.

A feature diagram f is generated, and the split operation is performed along the spatial dimension, which is divided into f^h and f^w. A 1×1 convolution was used to increase the dimension, and then the sigmoid activation function was combined to obtain the final attention vectors g^h and g^w. The formula used is as follows:

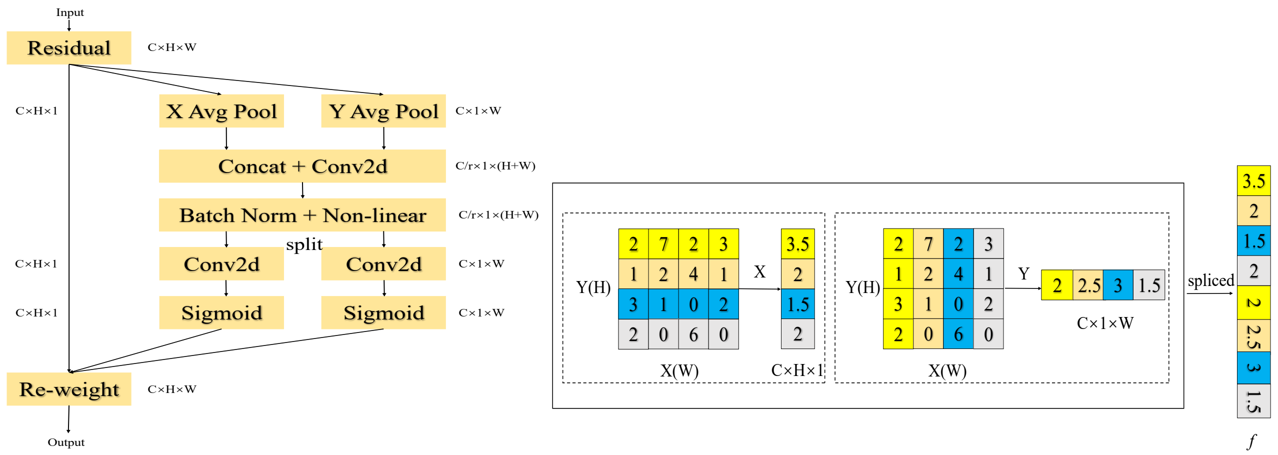
$$g^h = \sigma(F_h(f^h)) \tag{4}$$

$$g^W = \sigma(F_W(f^W)) \tag{5}$$

where σ is the sigmoid activation function, and F_h and F_w are the dimension-up operations. The output formula of the final CA attention mechanism module is as follows:

$$y_c(i,j) = x_c(i,j) \times g_c^h(i) \times g_c^w(j) \tag{6}$$

where $y_c(i,j)$ is the output of channel c, $x_c(i,j)$ is the input of channel c, $g_c^h(i)$ is the attention vector of channel c along the h direction, and $g_c^w(j)$ is the attention vector of channel c along the w direction.



a. CA attention mechanism module flow chart b. C×H×1 and C×1×W characteristic diagram and characteristic diagram *f*

Fig. 5 - CA attention mechanism module

Note: In Figure 5-a, C denotes the number of channels, H represents image height, W represents image width, X Avg Pool refers to horizontal average pooling, Y Avg Pool refers to vertical average pooling, Conv2d denotes the 2D convolution operation, Non-linear represents the non-linear activation function, Sigmoid is the activation function, and Re-weight stands for the weight file. In Figure 5-b, spliced refers to the concatenation operation.

Training environment deployment and parameter adjustment

The hardware environment in this experiment was a CPU 13th Gen Intel (R) Core (TM) i9-13900KF, 3.00GHz, running memory 32.0GB, GPU Nvidia GeForce RTX 4090, operating system Windows 11-64 bit. Deep learning frameworks such as PyCharm2021, Pytorch1.11, and Python3.8, were used to optimize, train, and evaluate the model.

The training parameters of the model were an image input size of 640 × 640, training of 500 rounds, batch size of 16, and maximum working core of 16. The optimizer used was the stochastic gradient descent method, the initial learning rate was 0.001, the learning rate momentum was 0.937, and the weight attenuation coefficient was 0.0005.

Evaluation index

The precision, P, recall rate, R, and mean average precision, mAP, in the training results were used to measure the recognition performance of the model. Parameters, weight file, and FLOPs were used to measure the complexity of the model. Precision P represents the proportion of correctly detected samples to all detected samples. The formula is as follows:

$$P = \frac{TP}{(TP + FP)} \tag{7}$$

The recall rate R denotes the proportion of correctly detected samples to all samples that are expected to be detected and is calculated as follows:

$$R = \frac{TP}{(TP + FN)} \tag{8}$$

The average accuracy mean mAP is the average of the APs of all classes. The average accuracy AP is the area of the P–R curve as a measure of the performance of the multiclass target detection model. The formula is as follows:

$$mAP = \frac{\sum_{i=1}^{N} \int_0^1 P(R)dR}{N} \times 100\% \tag{9}$$

where TPrepresents the number of true positives (samples correctly detected by the model), FP represents the number of false positives (samples incorrectly detected by the model), FN represents the number of false negatives (samples that should have been detected but were missed), and N represents the number of detection categories. The constructed dataset included three types of cotton pests and diseases, along with healthy leaves as a separate category, resulting in a total of four classes ($N = 4$).

RESULTS AND ANALYSIS

Performance comparison of YOLOv5s-FasterCA model ablation test

To clarify the influence of the Ghost, C3Faster, and CA attention mechanism modules on the performance of the YOLOv5s model, ablation tests were conducted using a self-made dataset of common pests and diseases in cotton. The test results are listed in Table 1. It can be seen from Table 1 that in the original network, using the Ghost module and C3Faster module instead of the ordinary convolution module and C3 module can significantly reduce the model parameters, calculation, and weight file, with little reduction in the precision P, recall rate R, and mean average precision mAP. From the experimental results, it can also be seen that only when the CA attention mechanism module is added are the precision P and mean average precision mAP of the original network model increased by 0.7% and 0.4%, respectively. The recall rate R is slightly decreased, and the parameters, floating-point calculation, and weight file are also slightly increased.

Table 1

YOLOv5s model ablation test

Ghost module C3Faster module	CA attention mechanism	Precision *P* / %	Recall rate *R* / %	Mean average precision *mAP* / %	Parameters / ×10⁶M	Weight file size / MB	FLOPs / G
×	×	88.7	92.4	93.4	7.03	13.7	16.0
√	×	88.3	91.8	92.7	4.58	9.14	10.5
×	√	89.4	92.1	93.8	7.06	13.8	16.1
√	√	89.5	91	93.6	4.61	9.20	10.6

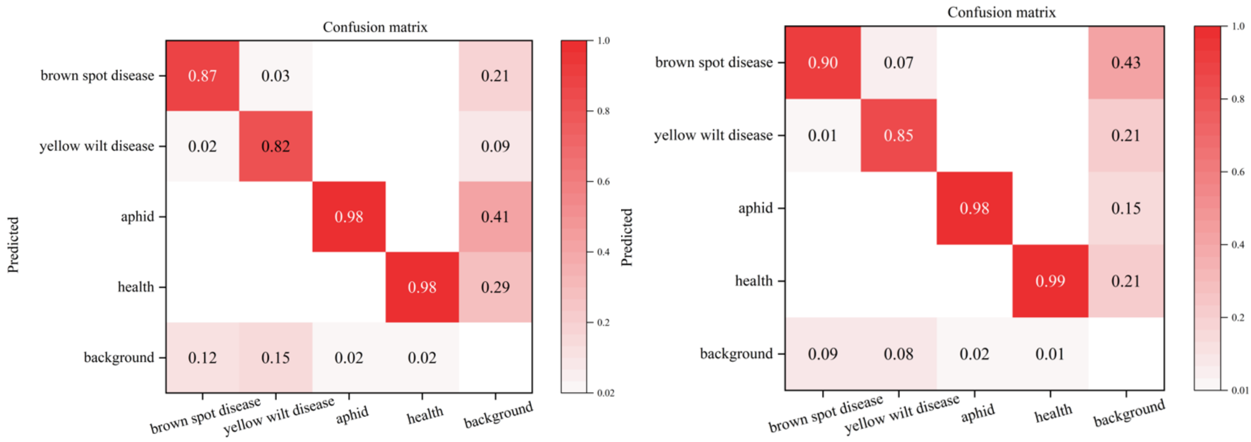
Note: √ indicates the inclusion of the module, while × denotes its exclusion.

Through the ablation test results, under the premise of realizing the lightweight model and ensuring recognition accuracy, the YOLOv5s-FasterCA model is obtained by adding the Ghost, C3Faster, and CA attention mechanism modules to the YOLOv5s model. In the detection and recognition of common cotton pests and diseases, the model parameters were 4.61 × 10^6 M; the FLOPs were 10.6G, and the weight file was 9.20 MB. Compared with the original model, the above parameters were reduced by 34.5%, 33.8%, and 32.9%, respectively, and the accuracy and average accuracy were improved by 0.8% and 0.2%, respectively. This indicates that the YOLOv5s-FasterCA model can improve the recognition accuracy of common cotton pests and diseases while meeting the lightweight requirements.

Performance comparison between YOLOv5s-FasterCA model and original model

To further verify the performance of the improved model, the YOLOv5s original model and YOLOv5s-FasterCA models were used to train and test the dataset of common cotton pests. The model performance was measured by a confusion matrix diagram, P _ curve diagram, R _ curve diagram, and PR _ curve diagram.

From Fig. 6(b), the improved model has a 3% increase in the probability of detecting brown spot disease, a 1% decrease in the probability of misjudging yellow wilt disease, and a 3% increase in the probability of detecting yellow wilt disease compared to Fig. 6(a). There was no change in the probability of aphid detection or a 1% increase in the probability of detecting healthy leaves. It can be seen from the confusion matrix that compared with the original model that the improved model can improve the precision and reduce the probability of misjudgment, indicating that the improved model is better than the original model.

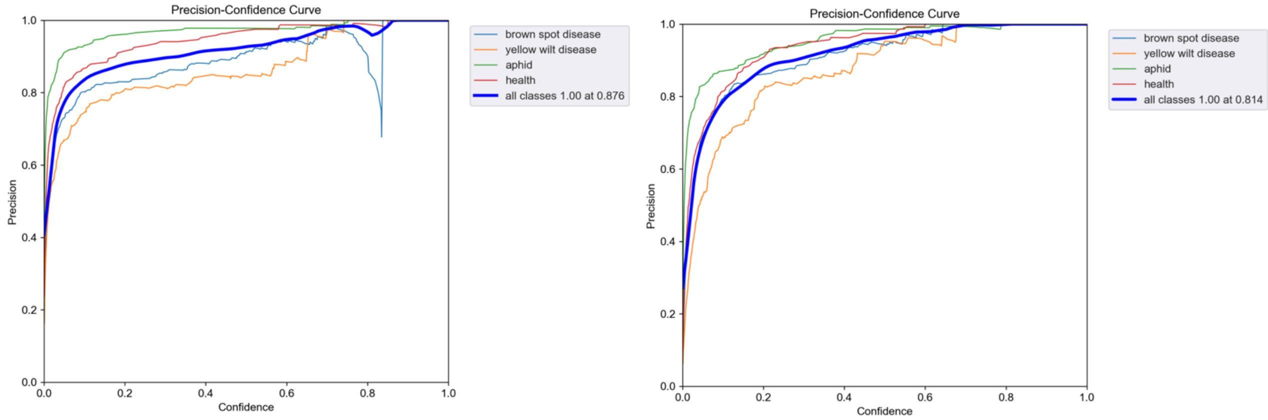


(a) YOLOv5s (b) YOLOv5s-FasterCA

Fig. 6 - Confusion_matrix

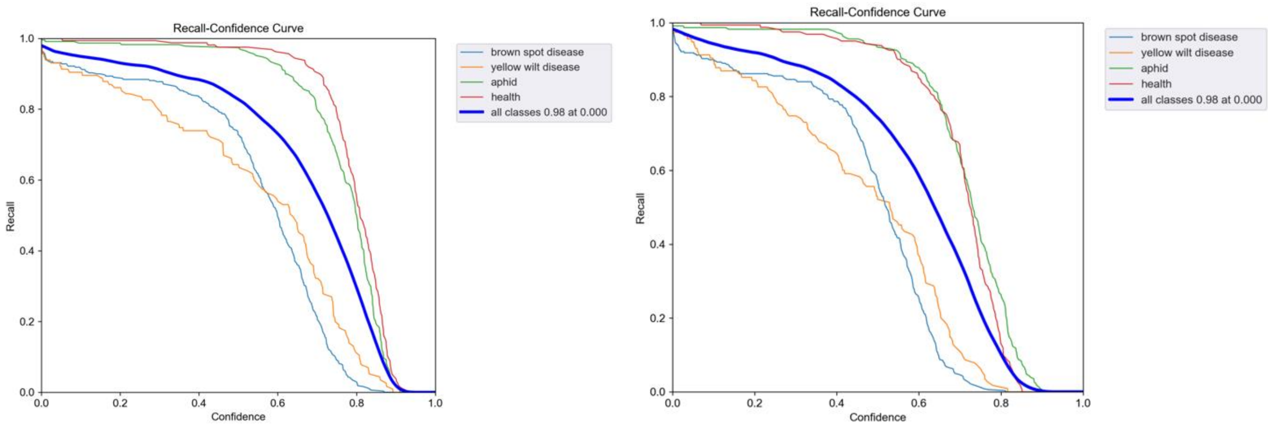
To compare the accuracy and confidence of the improved and original models, the P _ curve graph was used to display the changes in the recognition accuracy of the model for each category under different confidence thresholds.

As shown in Fig. 7(b), the highest confidence of the improved model was 0.814, which was 0.062 lower than that of the original model (Fig. 7(a)). This indicates that the improved model can achieve optimal accuracy when the confidence level is 0.814. Compared with the original model, it can reduce the possibility of missed detection when the confidence threshold is low.



(a) YOLOv5s (b) YOLOv5s-FasterCA

Fig. 7 - P _ curve comparison diagram



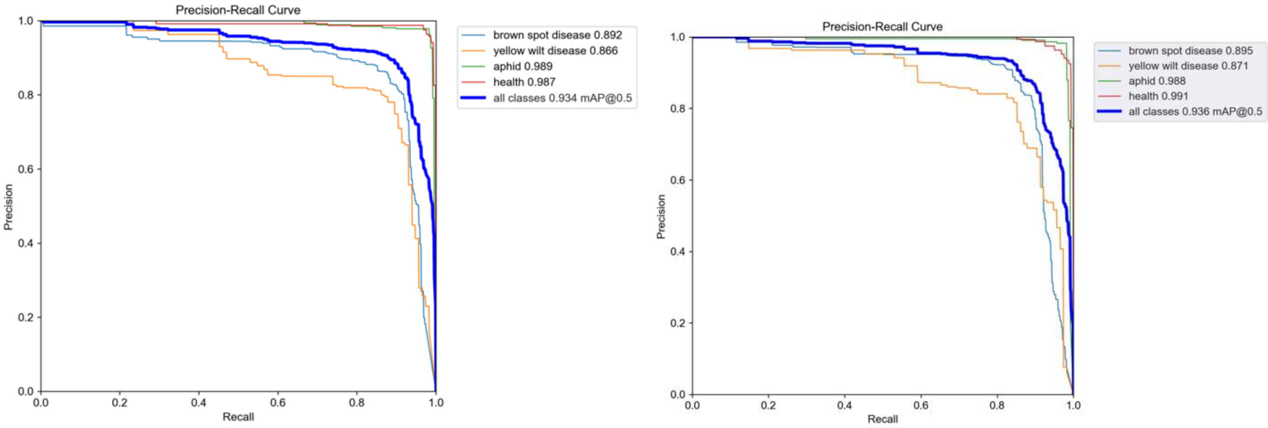
(a) YOLOv5s (b) YOLOv5s-FasterCA

Fig. 8 - R _ curve comparison diagram

To compare the precision and recall rates of the improved model and the original model, the R _ curve diagram was used to display the probability of recall of each category when the confidence was a certain value.

It can be seen from Fig. 8 that when confidence is small, category detection is more comprehensive. Through comparison, it was found that when the confidence level was 0.98, the recall rates of the improved model and the original model were 0.

To compare the accuracy and recall rate of the improved and original models, the PR_curve diagram shows the changes in accuracy with the recall rate. From Fig. 9(b), compared with Fig. 9(a), the improved model has a 0.3% higher accuracy in detecting brown spot disease, a 0.5% lower accuracy in detecting yellow wilt disease, a 0.1% lower accuracy in detecting aphids, a 0.4% higher accuracy in detecting healthy leaves, and a 0.2% improvement in accuracy for all types of mAP.



(a) YOLOv5s (b) YOLOv5s-FasterCA

Fig. 9 - PR _ curve comparison diagram

Performance comparison between YOLOv5s-FasterCA model and different models

The improved YOLOv5s-FasterCA model was compared with YOLOv4, YOLOv5s, and YOLOv7 on a self-built dataset. All models were trained using the same dataset and identical training parameters, and the results are shown in Table 2.

The mean average precision of the YOLOv5s-FasterCA model was 10.5%, 0.2%, and 0.4% higher than that of the other three models. The detection time of a single image was 1ms slower than that of the original model (which can be ignored). This indicates that the detection accuracy and recognition speed of the model meet the requirements for the real-time detection of common cotton pests and diseases.

Table 2

Performance comparison of different models

Modules	Precision P/%	Recall rate R/%	Mean average precision mAP/%	Parameters/×10^{6}M	Weight file size /MB	FLOPs/G	Detect/ms
YOLOv4	78.8	79.42	83.1	64.33	249.7	60.4	37
YOLOv5s	88.7	92.4	93.4	7.03	13.7	16.0	7
YOLOv7	90.2	92.6	93.2	70.83	135	188.9	41
YOLOv5s-FasterCA	89.5	91	93.6	4.61	9.20	10.6	8

YOLOv5s

YOLOv5s-FasterCA

Fig. 10 - Comparison of model detection effects in natural environment

Identification of common pests and diseases of cotton in natural environment

It can be seen from Table 2 that the model parameters, weight files, and FLOPs calculations of YOLOv4 and YOLOv7 are too large to meet the lightweight requirements of the model. Only the YOLOv5 and the improved YOLOv5s-FasterCA models were deployed on the Jetson Orin NX development board to detect common cotton pests and diseases under the same natural environment in the test dataset. The results are shown in Fig. 10. The detection results indicate that the YOLOv5s model may fail to identify certain pest and disease types under complex background. In contrast, the improved YOLOv5s-FasterCA model can significantly reduce false detections and missed detections, demonstrating better performance in accurately identifying common cotton pests and diseases, even in challenging environments. Furthermore, after acceleration with CUDA, the YOLOv5s-FasterCA model achieved a detection speed of 76.3 FPS on the Jetson Orin NX development board, meeting the requirements for real-time detection.

CONCLUSIONS

(1) An improved target detection model, YOLOv5s-FasterCA, was proposed by integrating Ghost and C3Faster modules to replace the standard convolution and C3 modules in the original YOLOv5s. The Coordinate Attention (CA) mechanism was introduced. The improved model achieved a precision of 89.5% and a mean average precision (mAP) of 93.6%, 0.8% and 0.5% higher than the original YOLOv5s respectively. In addition, the parameter count, FLOPs, and weight file size of the model were significantly reduced, meeting the requirements for lightweight deployment.

(2) The comparative experiments showed that the mAP of YOLOv5s-FasterCA was 10.5%, 0.2%, and 0.4% higher than those of YOLOv4, YOLOv5s, and YOLOv7, respectively. Although the improvement in single-frame inference time compared to YOLOv5s was minimal (reduced by approximately 1ms), the enhanced model demonstrated excellent detection performance in complex natural environments. It can effectively reduce false and missed detections of common cotton pests and diseases under real field conditions.

(3) The recognition accuracy and robustness of the YOLOv5s-FasterCA model are improved in unstructured field environments. With CUDA acceleration, it achieved a real-time detection speed of 76.3 frames per second on the Jetson Orin NX development board, making it suitable for mobile deployment scenarios with limited computing resources. However, when encountering unseen or unmarked types of pests and diseases, the model may exhibit decreased recognition performance or misclassify them as known categories. Future work will improve generalization ability by expanding datasets and implementing anomaly detection strategies.

ACKNOWLEDGEMENT

This study was financially supported by the Bingtuan Science and Technology Program (Grant No. 2022DB001). The authors are grateful to anonymous reviewers for their comments.

REFERENCES

[1] Amarasingam, N., Gonzalez, F., Salgadoe, A.S.A., Sandino, J., & Powell, K. (2022). Detection of White Leaf Disease in Sugarcane Crops Using UAV-Derived RGB Imagery with Existing Deep Learning Models. *Remote Sensing*. 14(23): 6137. https://doi.org/10.3390/rs14236137

[2] Apacionado, B.V., & Ahamed, T. (2023). Sooty Mold Detection on Citrus Tree Canopy Using Deep Learning Algorithms. *Sensors.* 23(20): 8519. https://doi.org/10.3390/s2320851

[3] Azath, M., Zekiwos, M., & Bruck, A. (2021). Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis. *Journal of Electrical and Computer Engineering.* 2021: 9981437. https://doi.org/10.1155/2021/9981437

[4] Chen, B., Wang, G., Liu, J.D., Ma, Z.H., Wang, J., & Li. T.N. (201)8. Extraction of Photosynthetic Parameters of Cotton Leaves under Disease Stress by Hyperspectral Remote Sensing. *Spectroscopy and Spectral Analysis.* 38(06): 1834-1838. https://doi.org/10.3964/iissn 1000-0593(2018)06-1834-05

[5] Chen, H.L., & Yang, W. (2023). The Technology of Kiwifruit Pest and Disease Identification Based on Deep Learning. *2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), Guangzhou, China, IEEE*. 38-42. Doi:10.1109/ICHCI58871.2023.10277869

[6] Cruz, M., Mafra, S., Teixeira, E., & Figueiredo, F. (2022). Smart Strawberry Farming Using Edge Computing and IoT. *Sensors*. 22(15): 5866. https://doi.org/10.3390/s22155866

[7] Ganesan, G., & Chinnappan. J. (2022). Hybridization of ResNet with YOLO classifier for automated paddy leaf disease recognition: An optimized model. *Journal of Field Robotics*. 39(7): 1085-1109. https://doi.org/10.1002/rob.22089

[8] Hou, Q.B., Zhou, D.H., Feng, J.S. (2021). Coordinate attention for efficient mobile network design. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, IEEE*. 13713-13722. https://doi.org/10.1109/CVPR46437.2021.01350

[9] Lan, Y.B., Shun, B.S., Zhang, L.C., & Zhao, D.N. (2024). Identifying diseases and pests in ginger leaf under natural scenes using improved YOLOv5s *(基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别). Transactions of the Chinese Society of Agricultural Engineering*. 40(01): 210-216. Doi:10.11975/j.issn.1002-6819.202310124

[10] Li, G., Yun, I.Y., Kim, J.H., & Kim, J.K. (2019). Dabnet: Depth-wise asymmetric bottleneck for real-time semantic segmentation. arXiv preprint arXiv:1907.11357. https://doi.org/10.48550/arXiv.1907.11357

[11] Li, J., & Wang, X.W. (2020). Tomato Diseases and Pests Detection Based on Improved YoloV3 Convolutional Neural Network. *Frontiers in Plant Science*. 11: 898. https://doi.org/10.3389/fpls.2020.00898

[12] Liu, S.Y., Hu, B., & Zhao, C. (2023). Detection and identification of cucumber leaf diseases based improved YOLOv7*(基于改进 YOLOv7 的黄瓜叶片病虫害检测与识别). Transactions of the Chinese Society of Agricultural Engineering*. 39(15): 163-171. Doi:10.11975/j.issn.1002-6819.202305042 *(In Chinese)*

[13] Lou, Z.X., Xin, F., Han, X.Q., Lan, Y.B., Duan, T.Z., & Fu, W. (2018). Effect of Unmanned Aerial Vehicle Flight Height on Droplet Distribution, Drift and Control of Cotton Aphids and Spider Mites. *Agronomy.* 8(9): 187. https://doi.org/10.3390/agronomy8090187

[14] Masykur, F., Adi, K., & Nurhayati, O.D. (2023). Approach and Analysis of Yolov4 Algorithm for Rice Diseases Detection at Different Drone Image Acquisition Distances. *Tem Journal.* 12(2): 928-935. https://doi.org/10.18421/tem122-39

[15] Passos, W. L., Barreto, C. D. S., Araujo, G. M., Haque, U., Netto, S. L., & da Silva, E. A. (2023). Toward improved surveillance of Aedes aegypti breeding grounds through artificially augmented data. *Engineering Applications of Artificial Intelligence.* 123(Part C): 106488. https://doi.org/10.1016/j.engappai.2023.106488

[16] Sangaiah, A.K., Yu, F.N., Lin, Y.B., Shen, W.C., & Sharma, A. (2024). UAV T-YOLO-Rice: *An Enhanced Tiny Yolo Networks for Rice Leaves Diseases Detection in Paddy Agronomy. IEEE Transactions on Network Science and Engineering.* 11(6): 5201-5216. DIO:10.1109/TNSE.2024.3350640

[17] Shi, J., Lin, S.S., Luo, J.G., Yang, L.L., Zhang, Y.J., & Gu, L.C. (2023). Study on a detection method for crop diseases and insect pests based on YOLO v5s improved model *(基于 YOLO v5s 改进模型的玉米作物病虫害检测方法). Jiangsu Agricultural Sciences.* 51(24): 175-183. DOI:10.15889/j.issn.1002-1302.2023.24.024 *(In Chinese)*

[18] Shi, L., Lei, J.K., Wang, J., Yang, C.K., Liu, Z.H., Xi, L., & Xiong, S.F. (2024). Lightweight Wheat Growth Stage Identification Model Based on Improved FasterNet *(基于改进 FasterNet 的轻量化小麦生育期识别模型)*. Transactions of the Chinese Society for Agricultural Machinery. 55(5): 226-234. *(In Chinese)*

[19] Susa, J.A.B., Nombrefia, W.C., Abustan, A.S., Macalisang, J., & Maaliw, R.R. (2022). Deep Learning Technique Detection for Cotton and Leaf Classification Using the YOLO Algorithm. *2022 International Conference on Smart Information Systems and Technologies (SIST), Sultan, Kazakhstan, IEEE.* 1-6. DIO:10.1109/SIST54437.2022.9945757

[20] Wang, W.X., Liu, Z.Q., Gao, P., Liao. F., Li, Q., & Xie, J.X. (2023b). Detection of Litchi Diseases and Insect Pests Based on Improved YOLOv4 Model *(基于改进 YOLO v4 的荔枝病虫害检测模型). Transactions of the Chinese Society for Agricultural Machinery*. 54(05): 227-235. Doi:10.6041/j.issn.1000-1298.2023.05.023 *(In Chinese)*

[21] Wang, Y.K., Xu, R.J., Bai, D., & Lin H.F. (2023a). Integrated Learning-Based Pest and Disease Detection Method for Tea Leaves. *Forests.* 14(5): 1012. https://doi.org/10.3390/f14051012

[22] Zeiler, M.D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. arXiv:1311.2901. https://doi.org/10.48550/arXiv.1311.2901

[23] Zhang, N.N., Zhang, X., Bai, T.C., Shang, P., Wang, W.H., & Li, L. (2023). Identification Method of Cotton Leaf Pests and Diseases in Natural Environment Based on CBAM-YOLO v7 *(基于 CBAM-YOLO v7 的自然环境下棉叶病虫害识别方法). Transactions of the Chinese Society for Agricultural Machinery*. 54(S1): 239-244. Doi:10.6041/j.issn.1000-1298.2023.S1.025 *(In Chinese)*

[24] Zhao, H.X., Shao, M.Y., Pan, P., Wang, Z.A., Mu, Q., He, Z.K., & Zhang, J.H. (2023). A Training Dataset for Deep Neural Network Model Recognition of Common Cotton Diseases *(一种面向深度神经网络模型的棉花常见病害训练数据集). Journal of Agricultural Big Data*. 5(04): 47-55. DOI: 10.19788/j.issn.2096-6369.230405 *(In Chinese)*