# Recognition of tea buds based on an improved YOLOv7 model

Mengxue Song[1], Ce Liu[1*], Liqing Chen[1], Lichao Liu[1], Jingming Ning[1,2], Chuanyang Yu[1]
(1. *College of Engineering, Anhui Agricultural University, Hefei 230036, China*;
2. *State Key Laboratory of Tea Biology and Resource Utilization, Hefei 230036, China*)

**Abstract:** The traditional recognition algorithm is prone to miss detection targets in the complex tea garden environment, and it is difficult to satisfy the requirement for tea bud recognition accuracy and efficiency. In this study, the YOLOv7 model was developed to improve tea bud recognition accuracy for some extreme tea garden scenarios. In the improved model, a lightweight MobileNetV3 network is adopted to replace the original backbone network, which reduces the size of the model and improves detection efficiency. The convolutional block attention module is introduced to enhance the attention to the features of small and occluded tea buds, suppressing the interference of the complex tea garden environment on tea bud recognition and strengthening the feature extraction capability of the recognition model. Moreover, to further improve recognition accuracy for dense and occlusive scenarios, the soft non-maximum suppression strategy is integrated into the recognition model. Experimental results show that the improved YOLOv7 model has the precision, recall, and mean average precision (mAP) values of 88.3%, 87.4%, and 88.5%, respectively. Compared with the Faster R-CNN, SSD, and original YOLOv7 algorithms, the mAP of the improved YOLOv7 model is increased by 7.4, 7.9, and 3.9 percentage points, respectively, and its recognition speed is also promoted by 94.9%, 46.2%, and 16.9%. The proposed model can rapidly and accurately identify the tea buds in multiple complex tea garden scenarios - such as dense distribution, being close to the background color, and mutual occlusion - with high generalization and robustness, which can provide theoretical and technical support for the recognition of tea-picking robots.

**Keywords:** tea bud recognition, YOLOv7, lightweight MobileNetV3 network, CBAM
**DOI:** 10.25165/j.ijabe.20241706.9095

**Citation:** Song M X, Liu C, Chen L Q, Liu L C, Ning J M, Yu C Y. Recognition of tea buds based on an improved YOLOv7 model. Int J Agric & Biol Eng, 2024; 17(6): 238–244.

## 1 Introduction

Tea is one of the most popular beverages worldwide, consumed by two-thirds of the world population for its rich nutritional value and distinctive flavor. As tea's place of origin, China plays an important role in tea production and exports[1]. However, current tea bud picking is still mainly based on manual labor, which is time-consuming and labor-intensive. Moreover, with the increasing proportion of the elderly, the agricultural labor force has shrunk significantly. Therefore, there is an increasingly urgent demand for intelligent tea bud picking, which puts forward higher requirements for the recognition of tea buds[2,3].

For the purpose of realizing the recognition of tea buds, many scholars have conducted in-depth research and exploration on tea bud recognition methods and have provided many recognition algorithms. Zhang et al.[4] developed the watershed algorithm to avoid the recognition omission for the highlighted surface of tea buds, improving the segmentation accuracy of tea buds and old leaves. Zhang et al.[5] applied the improved B-G algorithm and Otsu algorithm to distinguish the RGB images of tea buds from the tea-tree canopy. Shao et al.[6] employed the improved K-means clustering algorithm to segment the image of tea leaves by converting the RGB color model to a suitable color model, which provides technical support for the research and development of intelligent picking robots. The above-mentioned recognition methods are mainly based on color, shape, and texture features, which may lead to low recognition accuracy and efficiency for the illumination and occlusion scenarios of the complex tea garden environment[7-10].

Recently, inspired by the theory of deep learning, tea bud recognition models based on deep learning have been widely studied. Li et al.[11] modified the YOLOv3 network by introducing the SPP module and then compressed the model to achieve high-efficiency detection of tea buds. Wang et al.[12] explored the influence of different attention mechanisms on the recognition accuracy and verified the superiority of the convolutional block attention module (CBAM) in improving the recognition ability of densely distributed tea buds. Then Li et al.[13] investigated a high-precision and lightweight target detection model based on an improved YOLOv4 model by altering the backbone, adding the CBAM, and replacing the loss function to enhance the feature extraction capability and generalization ability, while reducing the size of the recognition model. The YOLO series algorithms have been widely applied for tea bud recognition[14,15], and have also been verified to be of good generalization and robustness[16-18]. Nevertheless, considering the extremely complex growth environment and diverse growth posture of tea buds, if the YOLO series algorithms are directly applied to tea bud recognition without corresponding improvements, it may be difficult to guarantee

**Received date:** 2024-05-25 **Accepted date:** 2024-11-01
**Biographies: Mengxue Song**, Undergraduate, research interest: image processing, Email: 2055349002@stu.ahau.edu.cn; **Liqing Chen**, PhD, Professor, research interest: intelligent agricultural machinery design theory and technology, Email: lqchen@ahau.edu.cn; **Lichao Liu**, PhD, Associate Professor, research interest: agricultural robot, Email: llchao@ahau.edu.cn; **Jingming Ning**, PhD, Professor, research interest: Tea Processing, Email: ningjm@ahau.edu.cn; **Chuanyang Yu**, PhD candidate, research interest: agricultural robot, Email: chuanyang@ahjzu.edu.cn.
***Corresponding author: Ce Liu**, PhD, research interest: intelligent agricultural machinery equipment design. Mailing address: agricultural intelligent equipment, Hefei 230036, China, Tel: +86-18756099526, Email: liuce@ahau.edu.cn.

recognition accuracy.

To deal with the aforementioned problems, this study further develops the YOLOv7 model for the recognition of tea buds. Compared to the original YOLOv7 model, the improved YOLOv7 model replaces the CSPDarknet53 backbone network with the lightweight MobileNetV3 network to reduce the model size. Moreover, the CBAM is introduced into the Neck network, and the original non-maximum suppression strategy (NMS) is substituted with the soft non-maximum suppression strategy (Soft-NMS), which can enhance the model sensitivity and suppress the influence of complex tea garden environments on tea bud recognition, thus realizing the rapid and accurate recognition of tea buds.

## 2 Materials and methods

### 2.1 Data acquisition

In this study, the tea bud image dataset was collected in the National High-Tech Agricultural Park of Anhui Agricultural University in early April 2023. To guarantee the diversity of the dataset and enhance the model's adaptability, 2600 original images from different tea cultivars were acquired from various angles, weather conditions, and times of day. The image dataset covers various scenarios of different tea garden environments, such as uneven illumination, bud overlap, and occlusion of leaves and branches, as shown in Figure 1.

### 2.2 Dataset production

To enhance the generalization ability and avoid over-fitting of the tea bud recognition model, the image dataset was expanded to 9000 images by means of translation transformation, vertical mirroring, rotation, brightness enhancement, and so on, as shown in Figure 2. The dataset was labeled with the LabelImg application and then kept in XML format.


Figure 1 Examples of tea bud image dataset for different cultivars and scenarios

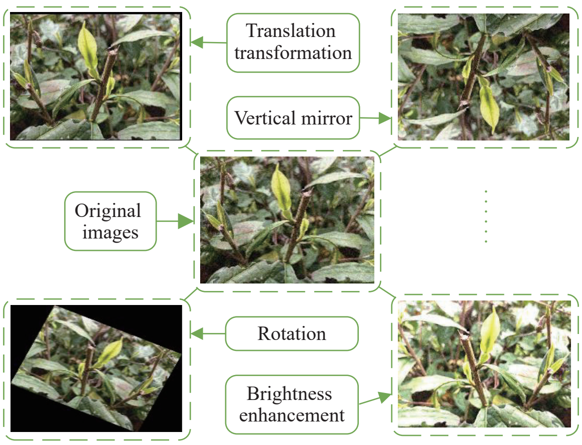



Figure 2 Examples for augmentation of image dataset

### 2.3 Original YOLOv7 recognition model

The YOLOv7 algorithm has been successfully applied for the recognition of grapes[19], apples[20], vehicles[21], and so on. In terms of recognition speed and accuracy, the YOLOv7 can operate within a frame rate range of 5 to 160 frames per second, thereby allowing for the processing of more images in a shorter amount of time. In the YOLOv7 model, the CSPDarknet53 is adopted as its backbone network, improving the expression capabilities and recognition performance of the model. The efficient layer aggregation networks module and SPPCSPC module are introduced to the neck part, which enhances the learning ability and recognition accuracy while maintaining relatively low computational parameters. Moreover, the RepConv module that utilizes a convolutional layer with multiple branches is adopted for model training, which effectively reduces the computational complexity and memory consumption.

The YOLOv7 model may be suitable for tea bud recognition due to its superior detection performance and capability. However, in view of the influence of complex tea garden environments, the YOLOv7 model cannot be directly employed without modification because of its deficiencies regarding dense and occlusive scenarios and recognition efficiency.

### 2.4 Improved YOLOv7 model

To guarantee the high efficiency and accuracy of tea bud recognition in complex tea garden environments, the YOLOv7 model is improved by a lightweight design with the MobileNetV3 network structure. Then the CBAM module is introduced to enhance the recognition ability of small tea buds in the complex tea garden environment. Furthermore, the NMS is substituted with the Soft-NMS for the circumstances of mutual occlusion of tea buds. The network structure of the improved YOLOv7 model is shown in Figure 3.

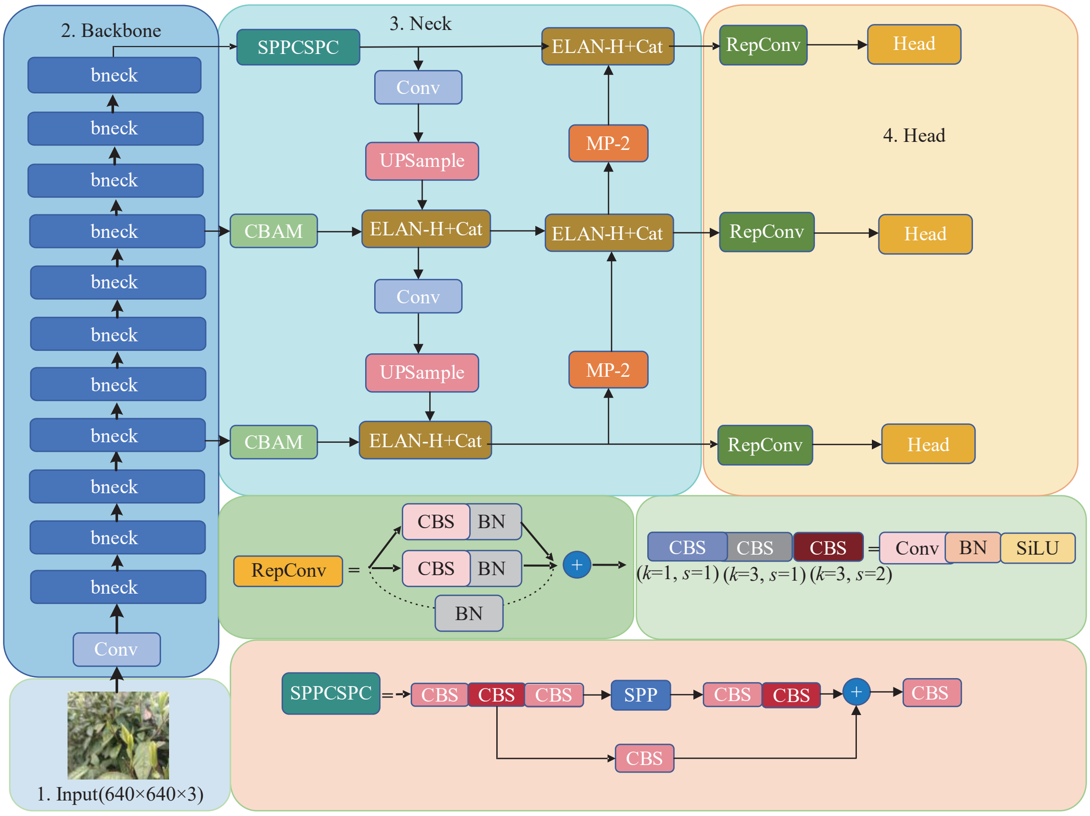



Figure 3 Representation of network structure of improved YOLOv7 model

2.4.1 MobileNetV3 network structure

To reduce the model parameters and satisfy the real-time requirement of tea bud recognition, a lightweight design of original YOLOv7 model is conducted with the MobileNetV3 network, whose structure is shown in Figure 4.

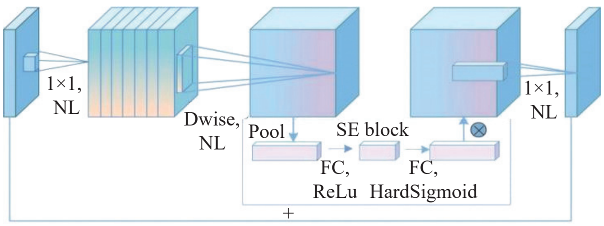



Figure 4 MobileNetV3 network structure

The MobileNetV3 network is a typical lightweight network that inherits the depthwise separable convolution block of MobileNetV1 and inverse residual structure of MobileNetV2. Moreover, it is also developed by introducing the SE attention module and updating the activation function. The MobileNetV3 network is thus composed of the depthwise separable convolution, inverse residual structure, and SE attention module. In the depthwise separable convolution, the standard convolution is decomposed into depthwise convolution and pointwise convolution modules. Firstly, the feature images are derived by depthwise convolution, whose number is equal to the number of output channels, and the number of output channels can be adjusted by pointwise convolution operation. By this means, feature extraction and feature combination are separated, which results in considerable reduction of the model parameters. Then the inverse residual structure conducts the feature extraction process in a higher dimension to retain more feature information. The SE attention mechanism is a lightweight channel attention module, which can selectively assign the weight of the network structure to amplify valuable feature channels.

Compared with the original CSPDarknet53 network, the MobileNetV3 reduces the model size and enhances the feature capture capability, which can effectively reduce the computational complexity and required storage space of the recognition model[22-24].

2.4.2 Convolutional block attention module

Considering that tea buds are relatively small and close to the background color[25], this paper further optimizes the YOLOv7 model by introducing the CBAM to improve tea bud recognition accuracy and robustness. Figure 5 shows the structure of CBAM which is composed of the Channel Attention Module (CAM) and Spatial Attention Module (SAM)[26]. It can concentrate on the desired sectors and pay more attention to the proportion of channels and pixels.

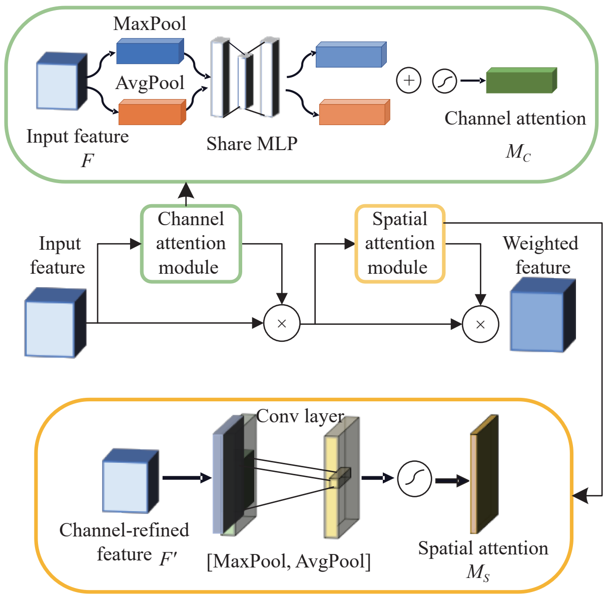



Figure 5 Schematic representation of the CBAM structure

In the CAM, the feature space information is aggregated through maximum pooling and average pooling to obtain two 1×1×$C$ feature maps. They are then fed into shared multilayer perceptron for processing. Subsequently, the results are added element by element and calculated by sigmoid activation function to acquire the weight $M_C$. Finally, the pixel-wise multiplication operation is executed to derive the channel-refined feature $F'$, which serves as the input for the subsequent SAM. The specific process is shown in Equations (1) and (2).

$$M_C(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

$$F' = M_C(F) \otimes F \quad (2)$$

where, $F$ denotes the input feature map; $\sigma$ is the sigmoid activation function; ⊗ is the multiplication by element; *MLP* is the multilayer perceptron; *AvgPool*($F$) is the average pooling; *MaxPool*($F$) is the maximum pooling.

The SAM performs the global maximum pooling and global average pooling operations on the $F'$. At the same time, it undergoes a convolution operation with a convolution kernel of 7×7 to obtain weight matrix $M_S$ through the sigmoid activation function. The output of the CBAM module is derived by multiplying the input and output values of the SAM. The calculation process can be expressed as Equations (3) and (4).

$$M_S(F) = \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)])) \quad (3)$$

$$F'' = M_S(F') \otimes F' \quad (4)$$

where, $f^{7\times7}$ is the 7×7 convolution operation; $F''$ is the feature map after adding the spatial attention mechanism; $M_S$ is the weight matrix after the SAM processing.

2.4.3 Soft non-maximum suppression

With the integration of MobileNetV3 and CBAM, the model size of YOLOv7 is significantly reduced and tea bud recognition accuracy is also improved. However, it should be noted that the YOLOv7 model shows low recognition accuracy in cases of mutual occlusion of tea buds. Therefore, the Soft-NMS is introduced to the YOLOv7 model to replace the NMS algorithm.

Considering that the tea buds are often densely distributed and highly overlapped, the traditional NMS algorithm uses the hard threshold judgment method to determine whether adjacent bounding boxes are retained and then resets the detection box to zero, which may result in high false rejection. To solve this problem, the Soft-NMS is selected as the postprocessing for the recognition model. In the Soft-NMS algorithm, the variance voting method is employed for weighted judgment of each box to obtain the optimal result, while considering all the confidence degrees and coincidence degrees. Its fractional reset function is[27]:

$$s_i = \begin{cases} s_i, & iou(M, b_i) < N_t \\ s_i(1 - iou(M, b_i)), & iou(M, b_i) \geq N_t \end{cases} \quad (5)$$

where, $s_i$ represents the confidence of current detection frame; *iou* denotes the correlation function; $M$ is the detection frame with the highest confidence; $b_i$ represents the adjacent detection frame in the domain; $N_t$ denotes the threshold of the detection frame.

## 3 Results and discussion

To quantitatively assess the performance of the proposed tea bud recognition model, the evaluation indices of Precision ($P$), Recall ($R$), Average Precision ($AP$), Mean Average Precision ($mAP$), and Frames Per Second ($FPS$) are defined as follows:

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

where, $TP$ denotes the number of tea buds that are correctly detected; $FP$ represents the number of tea buds that are wrongly detected; and $FN$ is the number of missed tea buds.

The $AP$ represents the area surrounded by the $P$-$R$ curve and coordinate axis. The calculation is as follows:

$$AP = \int_0^1 P(r)\mathrm{d}r \quad (8)$$

The $mAP$ indicates the average value of $AP$ for multiple recognition tasks, which can be expressed as:

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP \quad (9)$$

where, $n$ is the number of recognition tasks. Since this study only focuses on the recognition of tea buds, the values of $mAP$ and $AP$ are the same.

The experiment configuration environment is listed in Table 1.

**Table 1 Experimental configuration**

| Configuration | Parameters |
|---|---|
| Operating system | Windows10 |
| CPU | Intel(R) Xeon(R) CPU E5-2678 v3@2.50 GHz |
| GPU | NVIDIA Geforce RTX 2080 |
| Memory | 11 G |
| Accelerated environment | CUDA 10.1, CUDNN 7.6.5.32 |
| Library | PyTorch |
| Compilation language | Python3.8 |
| Software platform | Pycharm |

### 3.1 Ablation experiment of improved YOLOv7 model

To validate the effectiveness of the added modules on the performance promotion of tea bud recognition, four groups of ablation experiments were conducted on the dataset. The experimental results are listed in Table 2. It can be seen that compared with the original YOLOv7 model, the proposed improvement strategies are proved to significantly promote tea bud recognition ability. Compared with the original YOLOv7 model, the $P$ of the improved YOLOv7 model increased from 84.8% to 88.3%, an increase of 3.5%. The $R$ of the proposed model is 2.4% higher than the original YOLOv7 model, reaching 87.4%. The $mAP$ of the proposed model has improved from 84.6% to 88.5%, marking an increase of 3.9%. Moreover, the recognition speed of the improved YOLOv7 model is 16.9% higher than that of the original YOLOv7 model. This remarkable advancement validates the efficacy of the improved YOLOv7 model in promoting the accuracy and efficiency of tea bud recognition.

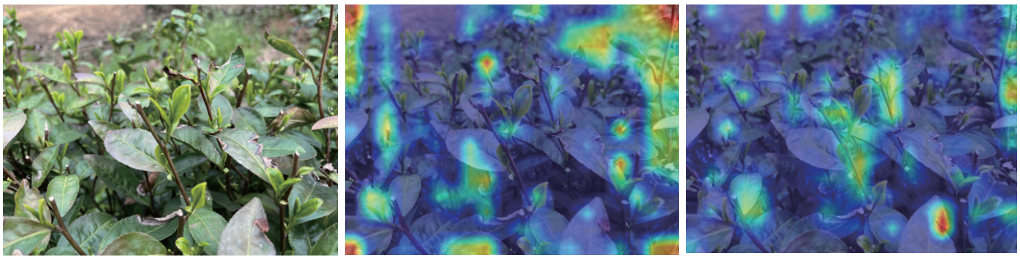
By replacing the CSPDarknet53 backbone network with MobileNetV3 backbone network, the $FPS$ is increased from 65 to 78, marking a 20% boost in speed. Moreover, the $P$, $R$, and $mAP$ of the model are increased by 1.6%, 1.4%, and 2.0%, respectively, proving that the MobileNetV3 network structure can improve tea bud recognition accuracy and speed.

On this basis, the CBAM is introduced to the YOLOv7M model. To present a visual analysis of the module, Gradient-weighted Class Activation Mapping is employed to visualize the model heat map, as depicted in Figure 6. The brightness of the region corresponds to its weight in the tea bud recognition. The results suggest that the CBAM can give more attention to the tea buds that are relatively small and close to the background color. The *P*, *R*, and *mAP* of the YOLOv7MC model turn to 87.4%, 86.9%, and 88.3%, respectively, which validates that the CBAM can enhance the feature extraction ability and attention to the small tea bud targets. However, it also should be noted that the speed has slightly decreased.

**Table 2 Comparison of evaluation indices of detection models**

| Models | *P*/% | *R*/% | *mAP*/% | FPS |
|---|---|---|---|---|
| Original YOLOv7 | 84.8 | 85.0 | 84.6 | 65 |
| YOLOv7M | 86.4 | 86.4 | 86.6 | 78 |
| YOLOv7MC | 87.4 | 86.9 | 88.3 | 76 |
| YOLOv7MCS | 88.3 | 87.4 | 88.5 | 76 |

Notes: YOLOv7M refers to the substitution of backbone network with MobileNetV3 network structure for the original YOLOv7 model, and YOLOv7MC refers to the addition of CBAM on the basis of YOLOv7M. YOLOv7MCS denotes the introduction of Soft-NMS to the YOLOv7MC.



a. Original image b. With CBAM c. Without CBAM

Figure 6 Illustration of model heat maps

Moreover, with the addition of Soft-NMS to the YOLOv7MC model, the *P*, *R*, and *mAP* are respectively increased by 0.9%, 0.5%, and 0.2%, with the recognition speed having no change, which confirms the performance of Soft-NMS for the dense distribution and highly overlapped circumstances of tea buds. The module's effect is represented in Figure 7.

The performance curves intuitively indicate that compared with the original YOLOv7 model, the recognition performance of the improved YOLOv7 model has been significantly promoted, as shown in Figure 8. The ablation experiment results verify that the proposed model can effectively improve tea bud recognition accuracy and speed for the complex tea garden environment.

### 3.2 Recognition results and discussion

To demonstrate the performance and superiority of the improved YOLOv7 model, the current mainstream models of Faster R-CNN, SSD, and original YOLOv7 model were adopted to make a comprehensive comparison for the complex tea garden environment. The tea bud images of different scenarios were randomly selected from the dataset to conduct the recognition experiment, including the single tea bud, backlight, direct light, and obscured tea buds, as depicted in Figure 9.

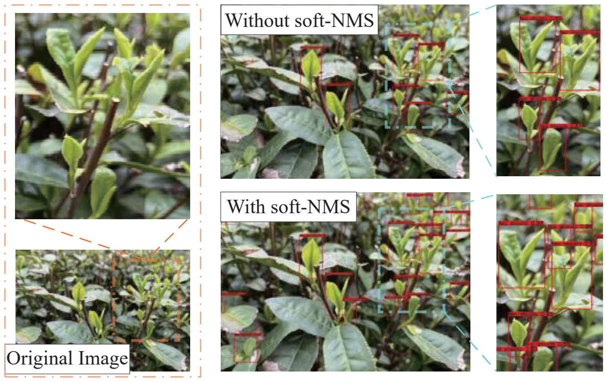
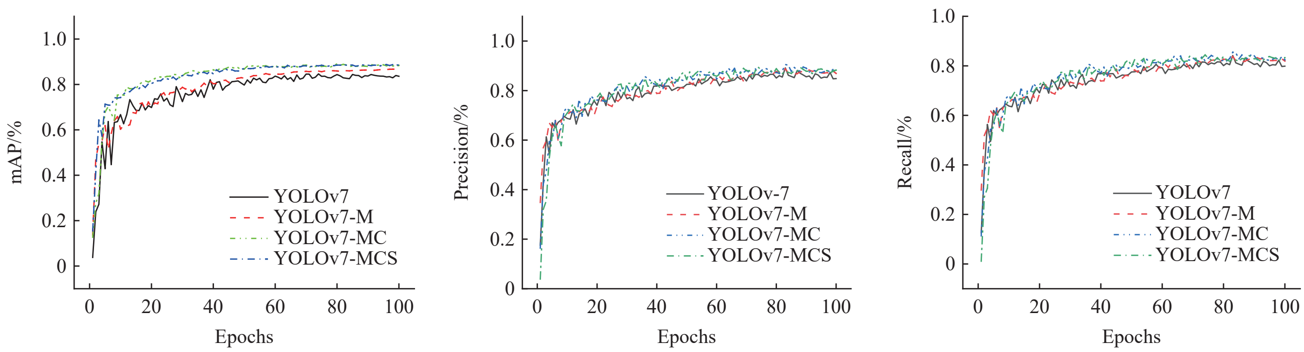



Figure 7 Representation of the effect of Soft-NMS

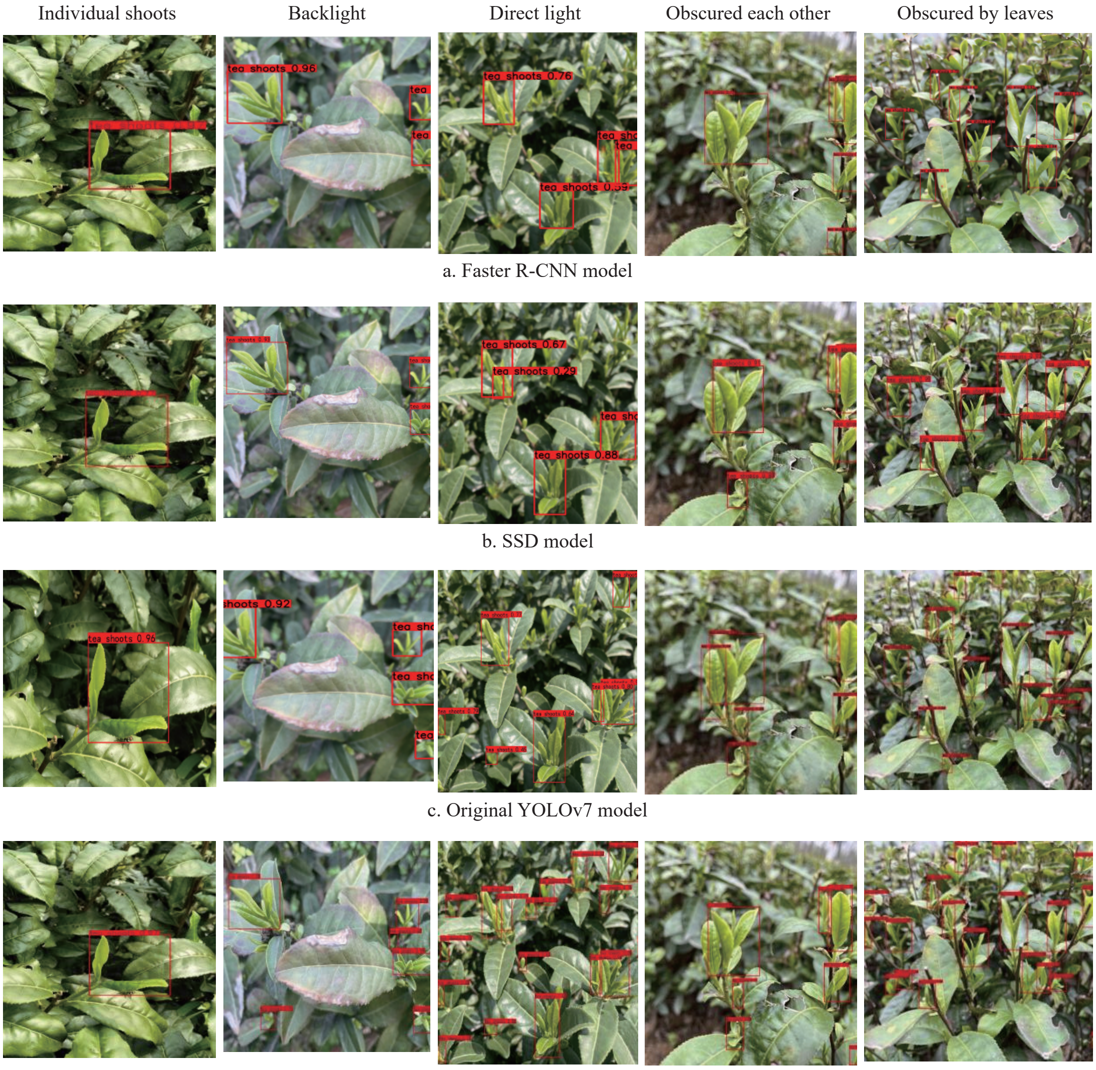



a.YOLOv7 mAP curve contrast diagram b.YOLOv7 Precision curve contrast diagram c. YOLOv7 Recall curve contrast diagram

Figure 8 Comparison of performance curves of the improved algorithms

It can be found that the four models exhibit an accurate identification for the circumstance of a single tea bud. However, with increasing number of tea buds, the Faster R-CNN, SSD, and original YOLOv7 models tend to miss the targets, especially when faced with the scenarios of backlight, direct light, and occlusion. On the contrary, the improved YOLOv7 model in this study can successfully recognize the tea buds in the above-mentioned scenarios, which proves that with the introduction of the lightweight MobileNetV3 network, CBAM, and Soft-NMS, the proposed model can achieve better tea bud recognition accuracy in the complex tea garden environment. The experimental result is listed in Table 3.

It can be found that compared with the existing mainstream models, the improved YOLOv7 model shows better performance on *P*, *R*, *mAP*, *FPS*, and model size. The *mAP* of the improved YOLOv7 model is higher than those of Faster R-CNN, SSD, and original YOLOv7 models by 7.4, 7.9, and 3.9 percentage points respectively, which demonstrates that the introduction of CBAM and Soft-NMS are able to improve the recognition accuracy. Also it can be seen that the improved YOLOv7 model has the smallest model size and exhibits a great advantage in recognition speed, which is 94.9%, 46.2%, and 16.9% higher than those of the Faster R-CNN, SSD, and original YOLOv7 models.




a. Faster R-CNN model

b. SSD model

c. Original YOLOv7 model

d. Improved YOLOv7 model

Figure 9 Representation of recognition results of current mainstream models

**Table 3 Comparison of evaluation indices of different recognition models**

| Model | *P*/% | *R*/% | *mAP*/% | *FPS* | Size of model/M |
|---|---|---|---|---|---|
| Faster R-CNN | 80.2 | 83.4 | 81.1 | 39 | 108 |
| SSD | 80.5 | 81.5 | 80.6 | 52 | 91 |
| Original YOLOv7 | 84.8 | 85.0 | 84.6 | 65 | 72 |
| Improved YOLOv7 | 88.3 | 87.4 | 88.5 | 76 | 48 |

As demonstrated above, it is clear that the YOLOv7 model proposed in this study can improve tea bud recognition accuracy and efficiency for the complex tea garden environment, which can provide theoretical and technical support for the application of tea bud picking robots.

## 4 Conclusions

In this study, the YOLOv7 model was further developed to improve the tea bud recognition accuracy and efficiency for the complex tea garden environment. In this model, the lightweight MobileNetV3 network was adopted to reduce model size and improve recognition speed. To increase the attention to tea buds for the densely distributed and highly overlapped scenarios, the CBAM and Soft-NMS were introduced. Then the effectiveness and superiority of the improved model were demonstrated by experimental verification. The specific conclusions are as follows:

1) The ablation experiment results indicate that the improved YOLOv7 model exhibits a better recognition accuracy with $P$=88.3%, $R$=87.4%, and $mAP$=88.5%, respectively, clearly outperforming the original YOLOv7 model, whose $P$, $R$, and $mAP$ are 84.8%, 85.0%, and 84.6%.

2) The experimental results verify that the improved YOLOv7 model can achieve better recognition performance in the complex tea garden environment, such as the dense distribution and occlusion scenarios, and the recognition speed is also 16.9% higher than that of the original YOLOv7 model.

3) Compared with the current mainstream recognition models of Faster R-CNN, SSD, and original YOLOv7, the $mAP$ of the improved YOLOv7 model was improved by 7.4, 7.9, and 3.9 percentage points, respectively.

This study is only for the recognition model of tea buds without grading (tea bud, a bud and one leaf, a bud and two leaves). In the future we will further optimize the model for grade picking and continue to improve tea bud recognition accuracy and speed.

## Acknowledgements

This research was financially supported by State Key Laboratory of Tea Biology and Resource Utilization (Grant No. SKLTOF20230123).

## [References]

[1] Chen T C, Li H X, Chen J Z, Zeng Z H, Han C Y, Wu W B. Detection network for multi-size and multi-target tea bud leaves in the field of view via improved YOLOv7. Computers and Electronics in Agriculture, 2024; 218: 108700.

[2] Lu J Q, Luo H X, Yu C R, Xiao L, Huang J W, Wu H W, et al. Tea bud DG: A lightweight tea bud detection model based on dynamic detection head and adaptive loss function. Computers and Electronics in Agriculture, 2024; 227: 109522.

[3] Xie D, Chen L, Liu L, Chen L, Wang H. Actuators and sensors for application in agricultural robots: A review. Machines, 2022; 10(10): 913.

[4] Zhang L, Zou L, Wu C Y, Jia J M, Chen J N. Method of famous tea sprout identification and segmentation based on improved watershed algorithm. Computers and Electronics in Agriculture, 2021; 184: 106108.

[5] Zhang L, Zhang H D, Chen Y D, Dai S H, Li X M, Imou K, Liu Z H, et al. Real-time monitoring of optimum timing for harvesting fresh tea leaves based on machine vision. Int J Agric & Biol Eng, 2019; 12(1): 6–9.

[6] Shao P D, Wu M H, Wang X W, Zhou J, Liu S, Wang Y S. Research on the tea bud recognition based on improved k-means algorithm. MATEC Web of Conferences, 2018; 232: 03050.

[7] Chen C L, Lu J Z, Zhou M C, Yi J, Liao M, Gao Z M. A YOLOv3-based computer vision system for identification of tea buds and the picking point. Computers and Electronics in Agriculture, 2022; 198: 107116.

[8] Hu H M, Kaizu Y, Zhang H D, Xu Y W, Imou K, Li M, et al. Recognition and localization of strawberries from 3D binocular cameras for a strawberry picking robot using coupled YOLO/Mask R-CNN. Int J Agric & Biol Eng, 2022; 15(6): 175–179,

[9] Liu B, Wei S S, Zhang F, Guo N W, Fan H Y, Yao W. Tomato leaf disease recognition based on multi-task distillation learning. Frontiers in Plant Science, 2024; 14: 1330527.

[10] Li F M, Zheng Y H, Liu S, Sun F B, Bai H R. A multi-objective apple leaf disease detection algorithm based on improved TPH-YOLOV5. Applied Fruit Science, 2024; 66(2): 339–415.

[11] Li Y T, He L Y, Jia J M, Chen J N, Lyu J, Wu C A Y. High-efficiency tea shoot detection method via a compressed deep learning model. Int J Agric & Biol Eng, 2022; 15(3): 159–166.

[12] Wang Y W, Xiao M H, Wang S, Jiang Q, Wang X C, Zhang Y N. Detection of famous tea buds based on improved YOLOv7 network. Agriculture-Basel, 2023; 13(6): 1190.

[13] Li J, Li J H, Zhao X, Su X H, Wu W B. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4. Computers and Electronics in Agriculture, 2023; 211: 107955.

[14] Gui J S, Wu D W, Xu H R, Chen J N, Tong J H. Tea bud detection based on multi-scale convolutional block attention module. Journal of Food Process Engineering, 2024; 47(2): e14556.

[15] Aqel D, Al-Zubi S, Mughaid A, Jararweh Y. Extreme learning machine for plant diseases classification: a sustainable approach for smart agriculture. Cluster Computing-The Journal of Networks Software Tools and Applications, 2021; 25(3): 2303–2303.

[16] Zhang X W, Xuan C Z, Hou Z F. Recognition model for coated red clover seeds using YOLOv5s optimized with an attention module. Int J Agric & Biol Eng, 2023; 16(6): 207–214.

[17] Gu Z A, He D Q, Huang J D, Chen J Q, Wu X H, Huang B C, et al. Simultaneous detection of fruits and fruiting stems in mango using improved YOLOv8 model deployed by edge device. Computers and Electronics in Agriculture, 2024; 227: 109512.

[18] Zhao L L, Zhu M L. MS-YOLOv7: YOLOv7 based on multi-scale for object detection on UAV aerial photography. Drones, 2023; 7(3): 188.

[19] Du W S, Jia Z H, Sui S S, Liu P. Table grape inflorescence detection and clamping point localisation based on channel pruned YOLOV7-TP. Biosystems Engineering, 2023; 235: 100–115.

[20] Ma H R, Li Y W, Zhang X Y, Li Y Y, Li Z Q, Zhang R Q, et al. Target Detection for coloring and ripening potted dwarf apple fruits based on improved YOLOv7-RSES. Applied Sciences-Basel, 2024; 14(11): 4523.

[21] Zhang Y, Sun Y P, Wang Z, Jiang Y. YOLOv7-RAR for urban vehicle detection. Sensors, 2023; 23(4): 1801.

[22] Howard A, Sandler M, Chen B, Wang W, Chen L C, Tan M, et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019; pp.1314–1324.

[23] Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, et al. MobileNets: efficient convolutional neural networks for mobile vision applications. 2017; arXiv: 1704.04861. https://arxiv.org/pdf/1704.04861

[24] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L C. MobileNetV2: inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018; pp.4510-4520. https://arxiv.org/abs/1704.04861

[25] Woo S, Park J, Lee J Y, Kweon I S. CBAM: convolutional block attention module. Computer Vision - ECCV. Cham: Springer International Publishing, 2018; 3-19.

[26] Khan R U, Khattak H, Wong W S, AlSalman H, Mosleh M A A, Rahman S M M. Intelligent malaysian sign language translation system using convolutional-based attention module with residual network. Computational Intelligence and Neuroscience, 2022; 2021: 9023010.

[27] Bodla N, Singh B, Chellapp.R, Davis L S. Soft-NMS—improving object detection with one line of code. Proceedings of the IEEE International Conference on Computer Vision, 2017; pp.5561–5569.