RESEARCH ARTICLE

Tobacco top flowering period recognition and detection model based on improved YOLOv4 and Mask R-CNN network

Weihua Qin[1], Longfei Wang[3], Xiangguo Cheng[1], Chaofei Yang[1], Shengli Chen[1, *], Lianhao Li[2], Chenhui Zhu[2]

[1]Sanmenxia Branch of Henan Tobacco Company, Sanmenxia, Henan, China. [2]Henan Agricultural University, Zhengzhou, Henan, China. [3]Henan Tobacco Company, Xuchang, Henan, China.

Received: December 11, 2023; **accepted:** February 7, 2024.

Tobacco is an important crop. During the flowering period of tobacco cultivation, the top part of plant contains high levels of water and impurities, which affect the taste of tobacco. Therefore, topping operations are carried out. Currently, deep learning is widely used in modern agricultural image recognition, pest and disease management, and other scenarios. However, due to the cost and technology limitations, its application in agricultural production is relatively low. This study proposed a tobacco flowering period recognition technology based on deep learning. An attention mechanism was introduced into the backbone residual module of deep learning to collect tobacco flower characteristic data at different stages to realize tobacco flower identification. Meanwhile, considering the large computational load of the tobacco flower detection model, a cascade strategy and transmembrane segment splitting were introduced to optimize the repeated features in backpropagation and improve the segmentation effect of the model. In the detection model test, compared with the convolutional and the unimproved models, the proposed model improved the recognition effect of tobacco flowers by 45.65% and 28.65%, respectively. In the multi-model comparison, the accuracy of the proposed model in dark light scenes was 0.853, which was better than other models. In the segmentation model test, the proposed model performed best in terms of confidence, time consumption, and number of frames. Improving the segmentation model during training took 3.5 seconds and detected 5 flowers, which was significantly better than the improved tobacco flower detection model and taking less time. The proposed tobacco flower detection model showed good application effects in tobacco flower detection scenarios, took less time, and had high detection accuracy, meeting the production requirements of modern farmers. The research provided important technical references for the management and cultivation of tobacco varieties.

Keywords: YOLOv4; tobacco flower; attention mechanism; features; cascade strategy; transmembrane segment splitting.

[*]**Corresponding author:** Shengli Chen, Henan Tobacco Company Sanmenxia City Company, Sanmenxia 472000, Henan, China. Email: shengli202311@163.com.

Introduction

Tobacco is an important crop. During the flowering period of tobacco cultivation, the top of tobacco is the reproductive site of tobacco plants, containing tobacco seeds and flowers. When manufacturing tobacco products, the top of tobacco is not suitable for making tobacco leaves because they contain high levels of moisture and impurities, which can affect the quality and taste of the tobacco leaves. Therefore, removing the top of tobacco can improve the quality and utilization of tobacco leaves [1]. At present, tobacco cultivation is mainly carried out through manual removal of tobacco flowers, which is time-consuming, difficult to operate, and can easily lead to tobacco virus infection and other problems. In recent years, with the continuous development of deep learning technology, many intelligent technologies have been applied to the tobacco flower planting process. The recognition of tobacco flower images based on deep learning, tobacco pesticide spraying, pest and disease control, and temperature control of planting environments have all significantly improved the planting effect of tobacco [2]. However, at present, deep learning has not been fully applied to modern agricultural planting processes because of high funding costs, high technical requirements, and harsh application conditions, all of which make the application of deep technology in modern agriculture relatively low [3].

Many scientists try to develop intelligent agriculture based on deep learning for agricultural production and management. Siddique *et al*. proposed a self-supervised training model to improve the model's sensitivity to flower plants through automatic pseudo-labeling. To improve the model's effect on plant prediction, the training data was expanded and converted into pseudo labels through semantic prediction, thereby improving the model's overall identification and inspection effect on plants. By applying the proposed technology to actual scenarios, the technology could effectively detect plant flower data and outperform comparison models [4]. Rahim *et al*. conducted research on agricultural fruit cultivation and found that, to increase crop yields, more fruits and flowers needed to be removed during the plant growing season. To facilitate crop deflowering operations, a flower detection technology based on convolutional network was proposed. The detection effect was improved through color threshold processing to adapt to the crop growth environment. The results showed that the proposed technology had excellent performance in both precision and recall with lower errors [5]. Omer *et al*. found that flowers had similar shapes and were difficult to effectively classify. To solve the problem of flower classification, a different flower classification technology based on convolution was proposed to identify flowers and extract flower features through convolution and improve the flower classification effect of the model through database training by using Kurdistan flower data set. The results showed that the technology had good classification effects with better recognition and classification effects on flowers compared with other classification models [6].

The You Only Look Once (Yolo) v4 algorithm has been widely used in crop fruit and flower identification, providing important technical support for crop management. Yuan *et al*. investigated existing target detection technology and found that target detection technology had important applications in the field of agricultural production. They applied advanced YOLOv4 detection technology to effectively identify and detect apple flowers and analyzed the effect of this technology on flower recognition and classification. The data with 100% to 5% manual injection quality was used for training in the study. Through training, the proposed technology could adapt well to different environments and effectively achieve the classification and recognition of apple flowers [7]. Gai *et al*. found that it was difficult to detect the flowers and fruits of cherries during the growth period and to ensure the quality of the fruits. Therefore, they proposed a cherry fruit identification and diagnosis technology based on the YOLO-V4 framework to improve cherry variety management. Considering the small size of cherries, to improve the detection accuracy, a feature extraction layer network was added to improve the recognition of cherry features. Comparing with the similar detection models, the proposed technology showed excellent results in detection efficiency and detection quality, and significantly improved the management quality of cherries [8]. Mithra *et al*. used computer image technology to detect the pollination of melon flowers and improve the natural pollination effect of melon crops based on the YOLO architecture. The method transferred learning to realize flower recognition and pollination judgment, and optimized and improved the feature extraction effect of the YOLO architecture. The overall recognition effect of the proposed technology was higher, and the diagnostic accuracy was also better than that of related detection technologies [9]. He *et al*. studied automated harvesting technology to improve the strawberry picking effect by building a strawberry flower and fruit detection technology based on the YOLOv4 framework. YOLOv4 was used to identify strawberry flower and fruit characteristics, and the strawberry canopy data set was used to train the detection model, which improved the model's detection effect on strawberry growth. The detection error of the proposed model was lower and better than related models when applying the proposed technology to specific strawberry planting scenarios [10].

The application of machine learning models in the field of agricultural management provides important data support for the scientific management of crops. In tobacco cultivation, excessive nutrients will be consumed during the top flowering stage of tobacco, and the tobacco will be removed during the top flowering stage to ensure the necessary effects on tobacco growth and improve the quality of tobacco growth. The traditional manual detopping process is cumbersome, and the identification of the flowering period of tobacco tops is poor, which affects the growth of tobacco [11]. Faced with the difficulty of artificial tobacco flower decapitation, an improved method for intelligent tobacco flower decapitation recognition model using deep learning technology was proposed in this study based on YOLOv4 and introduced an attention mechanism in its backbone residual module. By collecting feature data of tobacco flowers at different stages, this model could accurately identify tobacco flowers and achieved intelligent topping operation, which would provide important technical references for the management and cultivation of tobacco varieties.

Materials and methods

Collection and processing of tobacco flower data

The tobacco plants used in this study were from the rural cooperative tobacco planting base in Longgang Village, Dongchuan Town, Yao'an County, Chuxiong Yi Autonomous Prefecture, Yunnan, China, which covers 186 acres with an output value of more than 500 billion Chinese yuan. The tobacco at the top flowering stage in a particular area of the base was selected for data collection. A total of 55,000 pieces of tobacco flower growth stage images including flower bud, flower, flower stalk, and other images were collected using Canon EOS R50 camera (Canon, Tokyo, Japan). After image screening and processing using VisualDL, a visualization analysis tool of PaddlePaddle based on Yolov4 (https://github.com/AlexeyAB/darknet), 125,000 pieces of tobacco flower data at each stage were obtained and used as model training materials.

The tobacco growth image information in standard tobacco plantations was collected using a Sony ILCE-6000L camera (Sony, Tokyo, Japan) with image size of 20,000 × 2,999 pixels. To ensure the effect of tobacco flowering period detection in the later period, it was necessary to consider the impact of different lighting environments and humidity environments on image collection [12]. Therefore, growth pictures of tobacco flowers under different lighting environments were collected in three time periods including morning, noon, and night at 6 am, 12 pm, and 7 pm, respectively. Through artificial means, more than 4,000 image data of tobacco flowers in flowers, blooming stages, and different lighting environments were collected to ensure sufficient training image data of tobacco flowers and improve the robustness and generalization ability of the model. The corresponding enhancement processing on the collected images was carried out including brightness, grayscale, contrast, noise, image contour adjustment, and other operations on the image. Through the above enhancement processing, the number of tobacco flower images was increased. Among them, the brightness adjustment was used to simulate the detection environment of fireworks in high light and dark light environments. Multiple different levels of lighting environments were set, and 5 levels of brightness were set from low to high to reflect the model more accurately under different brightness detection effects [13]. The gray scale of the fireworks photos was adjusted according to the three stages of early, middle, and late. The image information was simplified through gray scale adjustment and the learning effect of the model on the image information was enhanced. In addition, contrast, noise, and contour adjustments were performed on the image data. After completing the enhancement processing of tobacco flower image data, the image was to be sample labeled [14]. Sample annotation was an important step in building a tobacco flower model and would directly affect the model's detection effect on tobacco flowers. In data annotation, the first step was to determine the annotation target by determining the target or specific area to be annotated in the image such as the boundary box of the object, the position of key points, *etc.*, and using the annotation tool for content processing, which included labeling the image with the selected annotation tool, and drawing or marking the corresponding information in the image according to the predetermined annotation target. The data set was then sorted out. The marked image and the corresponding marked information were sorted into a specific format, which was usually saved as a pair of data. The data set was divided into a training set, verification set, and test set.

Tobacco flower annotation mainly used the labeling tool for manual annotation. After the image data was annotated, a two-dimensional pixel coordinate file containing the target object would be formed. The software parsed the XML format file and generated the corresponding TXT format file, which would contain image category information, path, and two-digit pixel coordinates. According to the characteristics of tobacco flowers, they were divided into four types of flowering stages in the sample labeling including bud, semi-flowering, full flowering (top flowering), and late flowering. The annotation process included determining the tobacco flower growth stage to which the image belonged to. Integer or one-hot encoding was usually used to represent the category label to ensure that each image was correctly annotated with the corresponding category information. Through the above operations, the collection and processing of tobacco flower sample data were completed.

Construction of tobacco top flowering period detection model based on improved YOLOv4

To effectively identify the growth status of tobacco flowers, the study used Yolov4 to build a tobacco flower top flowering period identification and detection model. In fireworks detection, the Yolov4 model had the advantages of high recognition accuracy and fast detection. To improve the feature recognition effect of the Yolov4 model, an attention mechanism was introduced into the model and the target feature information of tobacco flowers at different growth stages was incorporated, thereby improving the overall detection effect of the model [15]. The framework of the tobacco top flowering period detection model based on the improved model was shown in Figure 1. In the construction of the improved Yolov4 tobacco flower detection model, let N represented the number of samples, M represented the dimension of the feature, and K represented the number of categories of tobacco flowers. X_i indicated the input data of the i th sample, and Y_i indicated the tobacco flower category label corresponding to the i th sample. The prepared sample data needed to be preprocessed. The normalization processing of the data was shown in Equation (1).

$$\hat{X}_i = \frac{X_i - X_{i\min}}{X_{i\max} - X_{i\min}} \tag{1}$$

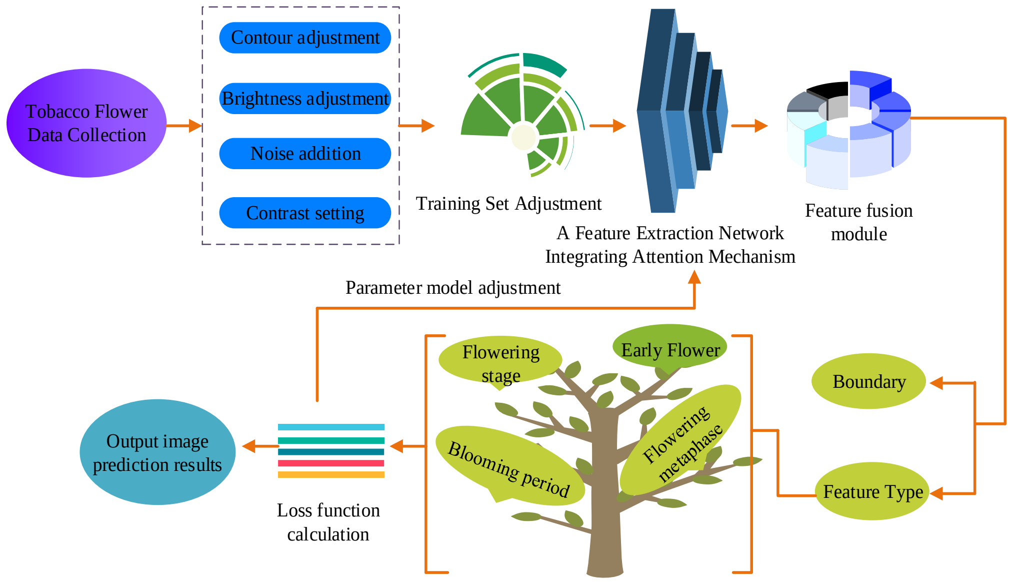


Figure 1. Framework of tobacco flower detection model based on improved Yolov4.

where $\hat{X}_i$ was the input data after preprocessing. After preprocessing the data, the YOLOv4 backbone network, the cross-stage partial residual module (CSPRes), was constructed. The study introduced the attention mechanism into this module to fuse the target feature information of tobacco flowers at different growth stages. Specifically, in the CSPRes module, the attention mechanism was used to weight the convolutional features to enhance attention to important features [16]. The process of attention weighting was shown in Equation (2).

$$\mathrm{Att}_i = \mathrm{Attention}(\mathrm{Conv}_i) \tag{2}$$

Where Attention (Conv_i) was the attention symbol. Att_i was the attention weight. Conv_i was determined by Equation (3).

$$\mathrm{Conv}_i^{\mathrm{att}} = \mathrm{Conv}_i \times \mathrm{Att}_i \tag{3}$$

where Conv_i was the convolution feature i of the first layer. $\mathrm{Conv}_i^{\mathrm{att}}$ represented the convolution feature after attention weighting. Then the region proposal network (RPN) was constructed to generate candidate target frames and classify and regress each candidate frame as shown below.

$$\mathrm{RPN}(\mathrm{Conv}_i^{\mathrm{att}}) = \{\mathrm{Box}_i, \mathrm{Score}_i\} \tag{4}$$

where Box_i represented the position information of the i th candidate box and Score_i represented the confidence of the i th candidate box. Next, non-maximum suppression was required to remove redundant candidate frames and retain the most representative target frames as follows.

$$\mathrm{NMS}(\{\mathrm{Box}_i, \mathrm{Score}_i\}) = \mathrm{Box}_{\mathrm{final}} \tag{5}$$

where $\mathrm{Box}_{\mathrm{final}}$ was the final target box. Further, the target frame needed to be classified and regressed. The research used a fully connected layer to fuse the target frame and its corresponding features and performed corresponding classification and regression prediction. The operation process was shown in Equation (6).

$$\mathrm{FC}(\mathrm{Box}_{\mathrm{final}}, \mathrm{Conv}_i^{\mathrm{att}}) = \{\mathrm{Class}_i, \mathrm{Reg}_i\} \tag{6}$$

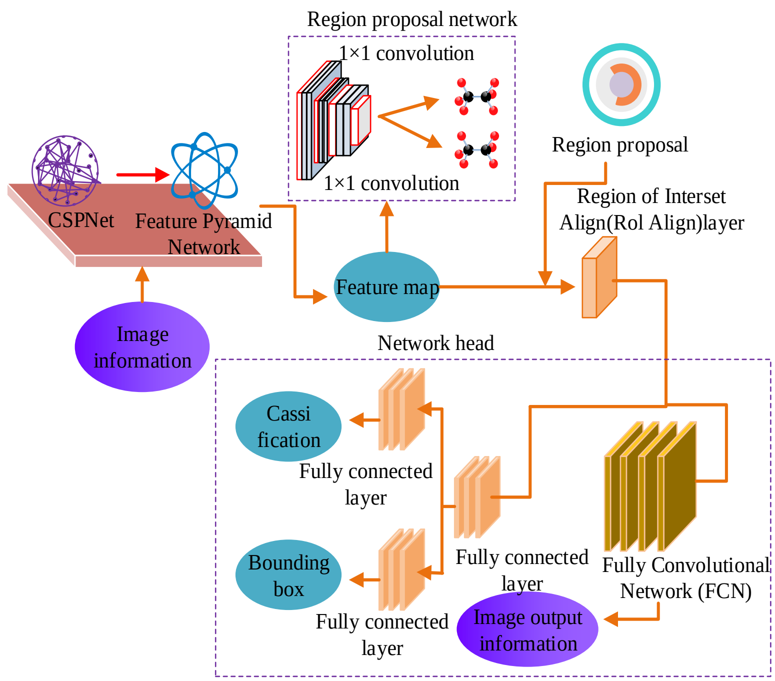


Figure 2. Mask-CNN segmentation model structure framework diagram.

where Class_i was the category prediction of the i th target frame and Reg_i was the position regression prediction of the i th target frame. To measure the accuracy of target box classification and regression, the loss function of the calculation model was introduced, and the cross-entropy loss function and the smooth L1 loss function were used to calculate the classification loss and regression loss. as shown in Equation (7).

$$\text{Loss} = \text{CrossEntropyLoss}(\text{Class}_i, Y_i) + \text{SmoothL1Loss}(\text{Reg}_i, Y_i) \quad (7)$$

where Y_i was the tobacco flower category label of the sample (i). CrossEntropy was the crossover loss. SmoothL1 was the smooth L1 loss. The last step was to optimize and train the model. The use of gradient descent method to update the parameters of the model was studied.

Construction of tobacco flower segmentation model based on improved Mask-CNN

The tobacco flower detection model based on the improved YOLOv4 faced problems such as large computational workload and time-consuming detection in practical applications, which led to a decline in the detection effect of the model after actual deployment. In response to the above problems, A tobacco flower segmentation method that improved Mask region convolutional neural network (Mask-CNN) was proposed, which fused the residuals in Mask -CNN with the Cross Stage Partial Network (CSPNet) and used the cascade strategy and cross-stage splitting to filter the repeated feature data of back propagation to increase the computational complexity of the tobacco flower detection model and improve the detection effect of the model. The structural framework of the improved Mask-CNN tobacco flower segmentation model was shown in Figure 2. To improve the tobacco top flower detection model based on YOLOv4, the research assumed that the recognition image had a grid of $S \times S$. Each grid was responsible for detecting B targets, and each target ($5+C$) was represented by parameters, where C was the number of target categories [17]. The center coordinate parameters generated by each grid were determined by Equation (8).

$$\begin{cases} y = \sigma(t)(\sigma(x) + c) \\ x = \sigma(t_x) + c_x \end{cases} \tag{8}$$

where (x, y) represented the center coordinate of the target. $\sigma(t)$ was the center coordinate parameter. t_x was the offset of the target x-axis. The width and height of the target were shown in Equation (9).

$$\begin{cases} w = p_w e^{t_w} \\ h = p_h e^{t_h} \end{cases} \tag{9}$$

where (w, h) represented the width and height of the target. p_w and p_h were the width and height of the prior box. The probability of the target category was then determined by Equation (10).

$$c = \Pr(C_1) \ldots \Pr(C_c) \tag{10}$$

Where $\Pr(C) = p$ was the probability of the existence of one of the targets, and the target category probability consisted of multiple targets. Mask-CNN was a method for target segmentation that added a segmentation network based on YOLOv4. Mask-CNN was composed of two sub-networks, a Region Proposal Network (RPN) for target region extraction and classification, and a Fully Convolutional Network (FCN) for target segmentation [18]. The Mask-CNN segmentation network used the bounding box information of the target to generate a mask of the target, and further extracted the detailed information of the target. The output of the model included the object's category, location information, confidence, and segmentation mask [19]. The parameters output by the model were shown in Equation (11).

$$\begin{cases} \Delta t = (t_x, t_y, t_w, t_h) \\ \Delta p = (p, p_1, \ldots, p_C) \\ \Delta m = (m_1, \ldots, m_{M \times M}) \end{cases} \tag{11}$$

where (t_x, t_y, t_w, t_h) was the offset of the target. Δt, Δp, Δm were the target offset variable, the target existence probability variable, and the target mask variable, respectively. $M \times M$ was the size of the mask. CSPNet was a network used for image classification. Its main idea was to divide the input feature map into two sub-feature maps with one for fine-grained feature extraction and the other for coarse-grained feature extraction. The output of the model included the feature vector of the target. The eigenvector output was shown in Equation (12).

$$\mathrm{f} = [f_1, f_2, \ldots, f_N] \tag{12}$$

where N was the dimension of the feature vector. To improve the computational load and detection effect of the tobacco flower detection model, the research integrated the residual of Mask-CNN with CSPNet and used the cascade strategy and cross-stage splitting to filter the repeated feature data of back propagation. The fused feature vector was shown in Equation (13).

$$\mathrm{f}' = [f_1', f_2', \ldots, f_{N'}'] \tag{13}$$

where N was the dimension of the fused feature vector. The fusion process could be described as the following mathematical model.

$$\mathrm{f}' = \mathrm{f} + \mathrm{f}_{\mathrm{CSPNet}} + \mathrm{f}_{\mathrm{ResNet}} \tag{14}$$

where $\mathrm{f}_{\mathrm{CSPNet}}$ was the output feature vector of CSPNet. $\mathrm{f}_{\mathrm{ResNet}}$ was the output feature vector of Mask-CNN. To implement the cascade strategy and cross-stage splitting, the study introduced a weight α and defined the fused feature vector as shown in Equation (15).

$$\mathrm{f}' = (1-\alpha)(\mathrm{f} + \mathrm{f}_{\mathrm{CSPNet}}) + \alpha \mathrm{f}_{\mathrm{ResNet}} \tag{15}$$

By adjusting the value of the weights α, the cascading strategy and the degree of splitting across stages could be controlled. By introducing the segmentation network of Mask-CNN and the cascade strategy and cross-stage splitting of CSPNet, the study improved the computational load of the tobacco flower detection model and the detection effect.

Experimental tool parameter setting

To test the performance of the proposed tobacco flower detection model, experimental tests were conducted on the WINDOWS 10, 64-bit platform using Intel i7 16-core processor, 64 G running memory, and Nvidia RTX4070 graphics card. The installation environment was Python 3.7 (https://www.python.org/) and Cuda 10.0 (https://developer.nvidia.com/cuda-zone). The initial parameters of the improved YOLOv4 model were set as shown in Table 1.

Table 1. Model initial parameters.

Parameter indicator	Numerical value
Enter image parameters	608 × 608
Number of training steps for backbone network	500
Model learning rate	0.001
IoU threshold parameters	0.5

Results

Tobacco flower detection model test based on improved YOLOv4

Introduce precision (p), recall (R), and average precision (mAP) were set as the model testing evaluation indicators. During training, Convolutional Neural Networks (CNN) and Yolov4 were used as test benchmarks. The strong light and complex scenes were selected to test the detection effects of different models on tobacco flowers (Figure 3). Comprehensive comparison of the detection effects of the three models showed that the CNN model had a general effect on identifying tobacco flowers in strong light scenes and could not identify flower stems in complex scenes. The traditional Yolov4 model was better than the CNN model in recognition, but it still could not identify flower stems in complex scenes. The proposed improved Yolov4 model performed well in both scenarios, especially in complex scenes where flower stems could be identified. Compared with the CNN model and the traditional Yolov4 model, the improved Yolov4 model had a recognition effect improved by 45.65% and 28.65%, respectively. In strong light scenes, the detection effects of the three models were significantly different (Figure 4). Among them, CNN performed the worst and tended to converge after 200 iterations with the recall value of 0.813. Yolov4 tended to converge after iteration 182 with the recall value of 0.848. The best performance was the improved Yolov4, which tended to converge after 160 iterations with the recall value of 0.914. In the comparison of complex scenes, compared with CNN and Yolov4, the improved Yolov4 was significantly better at identifying tobacco flowers with the converge after 161 iterations and the recall value of 0.907. The multi-model loss performance comparison results showed that, in strong light scenes, the improved Yolov4 model was significantly better than the other two models in terms of convergence and loss performance (Figure 5). Among them, the improved Yolov4 tended to converge after 80 iterations with the loss value of 0.022. In complex scenarios, the detection effects of the three models all declined, but the proposed model had the lowest loss with a loss improvement of 25.64% and 17.65% compared to CNN and Yolov4, respectively.

The comprehensive performance comparison results of different models were shown in Table 2. The model recognition effect was obvious in different scenarios. In low light scenarios, it further tested the overall performance of the model. The accuracy of the improved Yolov4 model was 0.853, and the accuracy of CNN and YOLOv4 were 0.724 and 0.218, respectively. The improved Yolov4 model performed better than the others. Also comparing the model frame values in dark light scenes, the improved Yolov4, CNN, and YOLOv4 were 49.6, 26.3, and 31.2, respectively. The proposed improved Yolov4 showed better detection results.

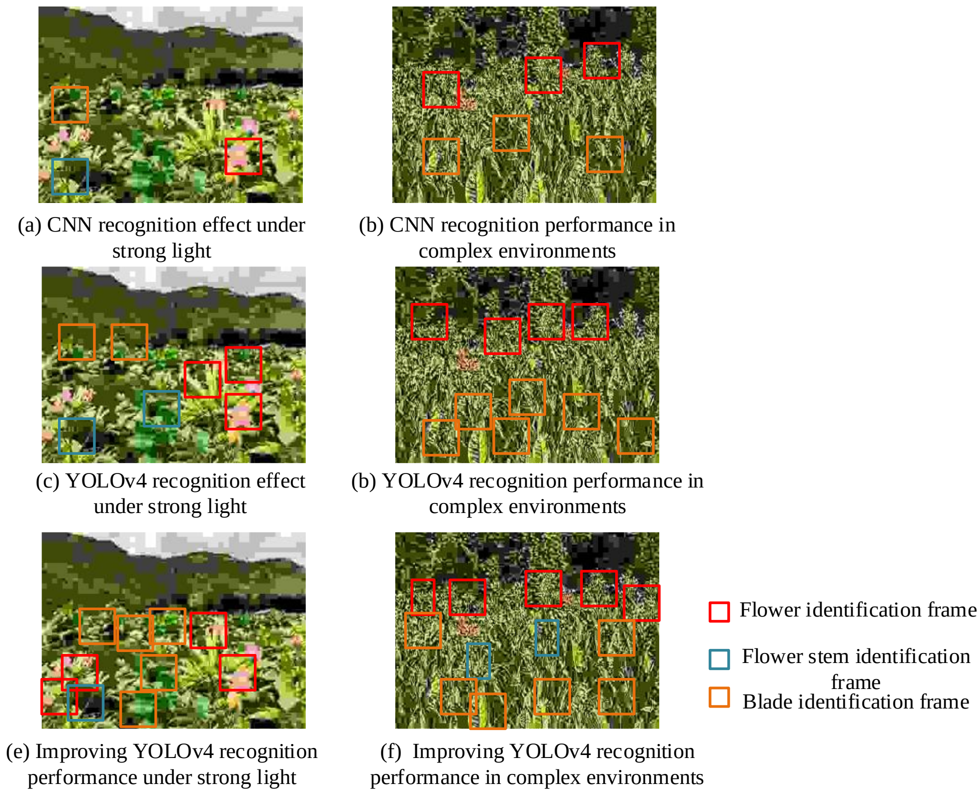


Figure 3. Comparison of recognition effects of different models for fireworks peak flowering period.

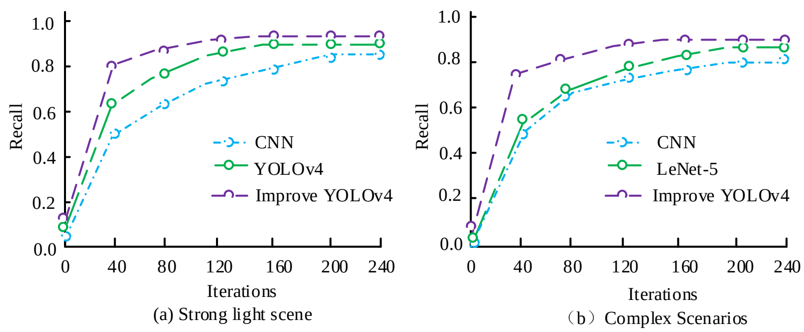


Figure 4. Comparison results of recall rates among different models.

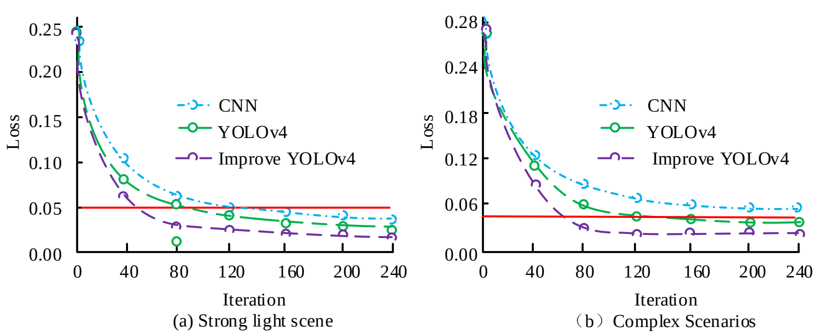
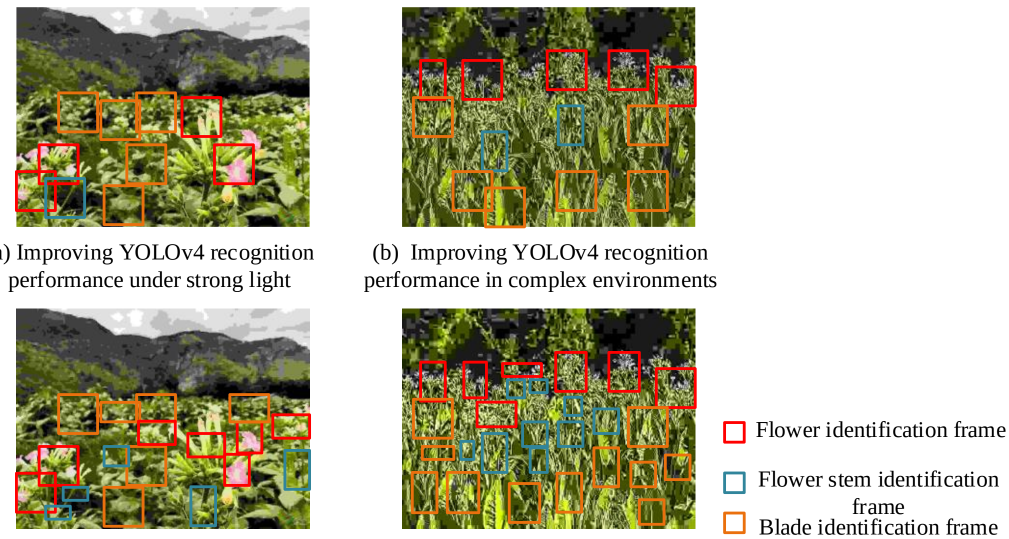


Figure 5. Comparison results of losses among different models.

Table 2. Comprehensive performance comparison of multiple models.

Scenario Type	Detection Model	Precision	mAP	Recall	FPS
Strong light environment	CNN	0.733	0.823	0.813	28.7
	YOLOv4	0.821	0.836	0.848	32.5
	Improve YOLOv4	0.864	0.935	0.914	50.6
Complex Scenarios	CNN	0.724	0.811	0.802	27.6
	YOLOv4	0.812	0.842	0.837	32.5
	Improve YOLOv4	0.850	0.914	0.907	49.7
Dim light environment	CNN	0.724	0.801	0.792	26.3
	YOLOv4	0.818	0.844	0.833	31.2
	Improve YOLOv4	0.853	0.923	0.906	49.6



(a) Improving YOLOv4 recognition performance under strong light

(b) Improving YOLOv4 recognition performance in complex environments

(c) Improving Mask-CNN recognition performance under strong light

(d) Improving Mask-CNN recognition performance in complex environments

Figure 6. Comparison of two improved tobacco flower detection models for tobacco flower recognition.

Tobacco flower detection model test based on improved Mask-CNN

Under the same experimental scenario, the performance of the improved Mask-CNN model was tested. The minimum image dimension was set to 384, while the maximum dimension was set to 768. Anchor frame specifications included 8 × 6, 16 × 6, 64 × 6, and other specifications. The model learning rate was set to 0.1 and the weight attenuation coefficient was 0.0001. The tobacco flower detection results of the two improved models were shown in Figure 6. Compared with the improved Yolov4, the improved Mask-CNN model could efficiently segment images and improve the overall detection effect of the model. The improved Mask-CNN model could recognize significantly more tobacco flowers, leaves, and stems, and the recognition effect was also more accurate. The dark light scene was selected to compare the comprehensive performance of the two models with the improved YOLOv4 and improved Mask-CNN as 92.65 and 96.75, respectively (Figure 7). Three stages of tobacco flowers were selected for comparison (Table 3). The improved Mask-CNN model demonstrated better performance than that in YOLOv4 in confidence, detection accuracy, time consumption, and frame number comparison. The confidence level of YOLOv4 improved during the bud stage was 0.82, while

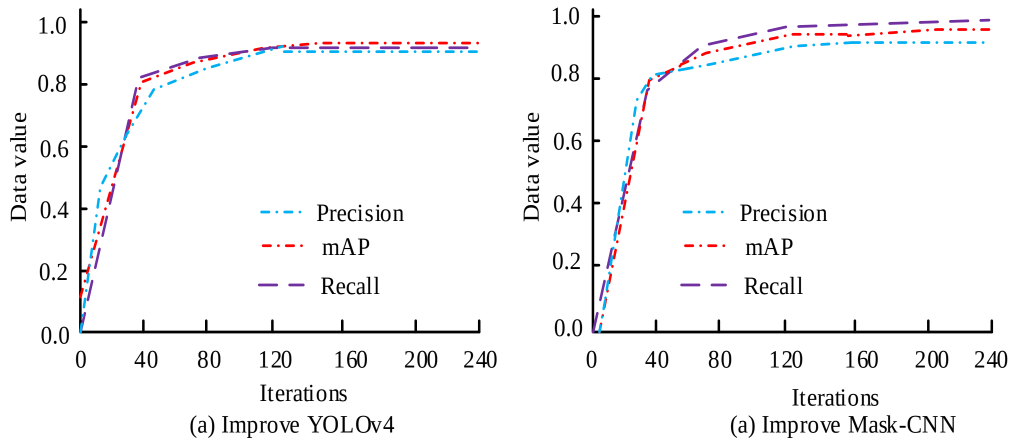


Figure 7. Comparison of the comprehensive performance of two improved tobacco flower detection models.

Table 3. Comparison of tobacco flower detection effects between two models.

Model	Stage	Confidence	Average Confidence	Detection accuracy (%)	Time consumption (s)	FPS	Number of identified items
Improved YOLOv4	Bud stage	0.82	0.820	90.32	4.6	49.5	3
	Apical flowering period	0.82 0.85 0.89	0.853	92.52	5.5	50.2	4
	Terminal flowering stage	0.91 0.92		92.56	3.5	50.3	4
Improved Mask-CNN	Bud stage	0.93	0.933	96.54	3.5	51.5	5
	Apical flowering period	0.92 0.95 0.98	0.950	97.64	3.4	51.4	6
	Terminal flowering stage	0.95 0.94	0.945	98.54	2.9	51.6	5

the improved Mask CNN was 0.93. Comparing the time consumption of the two models in the bud stage, the improved YOLOv4 detection time was 4.6 s, and the number of flowers detected was 3, while the improved Mask-CNN took 3.5 s, and the number of flowers detected was 5. The improved Mask-CNN model showed better tobacco flower detection performance and was more in line with tobacco agricultural production requirements.

Discussion

Tobacco needs to be decapitated in time at the flowering stage to ensure the nutrients needed for tobacco growth and development and ensure the quality of tobacco. The traditional manual topping operation is time-consuming, difficult to operate, and easy to cause tobacco virus infection and other problems. The introduction of deep learning technology, through the construction of intelligent tobacco flower top florescence recognition model, can accurately identify tobacco flowers and realize automatic topping operation. To improve the effect of tobacco flower topping, an intelligent tobacco flower detection method was proposed. Firstly, YOLOv4 was used to construct the tobacco blossom detection model. The local residual module and attention mechanism were introduced to improve the analysis effect of the model on features. Considering the time-consuming and computational complexity of the tobacco flower detection model, MASK-CNN and CSPNet were integrated to reduce the repeated characteristics of the model propagation process and improve the detection effect of the model on tobacco flowers through cascading strategy and transmembrane segment splitting. In the test of the improved YOLOv4 model, the detection effects of various models on tobacco flowers were compared. The improved YOLOv4 model could accurately identify tobacco flowers in complex scenes, which were 45.65% and 28.65% higher than that of CNN and YOLOv4, respectively. Comparing the comprehensive performance of different models, the maximum frame number of the improved YOLOv4 model in the dark scene was 49.6, and the recall rate and accuracy rate were the best. In the test of the improved MASK-CNN model, three important stages of tobacco flowers were selected for the test. At the flower bud stage, the confidence of the improved MASK-CNN model was higher one with the average confidence of 0.933, while the improved YOLOv4 was 0.820. The detection accuracy and time-consuming of the two methods were compared, and the improved MASK-CNN was excellent. The results showed that the proposed technology had excellent ability in the actual tobacco flower top florescence recognition. Compared with similar products, the model had higher accuracy and stronger stability than the others. The results indicated that the importance of deep learning technology in tobacco flower topping recognition was to improve the quality and utilization rate of tobacco leaves. The top of tobacco usually contains high moisture and impurities, which is not suitable for making tobacco, and will affect the quality and taste of tobacco. Through the deep learning technology to realize the accurate identification of tobacco flowering period, it can help farmers automatically remove the top of tobacco and improve the quality and utilization rate of tobacco, which has important economic significance and competitive advantage for the tobacco planting industry. Moreover, the impact of deep learning technology in tobacco flower topping recognition is to improve planting efficiency and reduce labor costs. The traditional manual roof removal operation needs a lot of manpower and time, and the operation is difficult. The use of deep learning technology for automatic flower topping operation can greatly save manpower and time costs and improve planting efficiency. Farmers can spend more time and energy on other tobacco planting links to improve the overall planting efficiency. The application of deep learning technology also brings the possibility of technological progress and innovative development for the tobacco planting industry. With the continuous development and optimization of deep learning algorithm, the accuracy and stability of tobacco flower topping recognition technology will be further improved. Meanwhile, deep learning technology can also be combined with other modern agricultural technologies such as the Internet of things, artificial intelligence, big data, *etc*. to build an intelligent tobacco planting system, achieve comprehensive and accurate management and monitoring, and improve the scientific and technological content and competitiveness of the tobacco industry.

References

1. Khan MA, Ali M, Shah M, Mahmood T, Ahmad M, Jhanjhi NZ, *et al*. 2021. Machine learning-based detection and classification of walnut fungi diseases. Intell Autom Soft Comput. 30(3):771-785.
2. Guo D, Liu J, Wang X. 2021. On development of multi-resolution detector for tomato disease diagnosis. JIFS. 41(6):6461-6471.
3. Li Z, Li F, Zhu L. 2020. Vegetable recognition and classification based on improved VGG deep learning network model. INT J COMPUT INT SYS. 13(1):559-564.

4. Siddique A, Tabb A, Medeiros H. 2022. Self-supervised learning for panoptic segmentation of multiple fruit flower species. IEEE Robot Autom Let. 7(4):12387-12394.
5. Rahim UF, Mineno H. 2020. Tomato flower detection and counting in greenhouses using faster region-based convolutional neural network. Journal of Image and Graphics. 8(4):107-113.
6. Omer SM, Hasan RM, Anwer BN. 2020. An image dataset construction for flower recognition using convolutional neural network. Science Journal of University of Zakho. 8(3):112-117.
7. Yuan W, Choi D, Bolkas D, Heinemann, PH, He L. 2022. Sensitivity examination of YOLOv4 regarding test image distortion and training dataset attribute for apple flower bud classification. Int J Remote Sens. 43(8):3106-3130.
8. Gai R, Chen N, Yuan H. 2023. A detection algorithm for cherry fruits based on the improved YOLO-v4 model. Neural Comput Appl. 35(19):13895-13906.
9. Mithra S, Nagamalleswari TYJ. 2023. Cucurbitaceous family flower inferencing using deep transfer learning approaches: CuCuFlower UAV imagery data. Soft Comput. 27(12):8345-8356.
10. Na MH, Cho W, Kim SK. 2020. A construction of web application platform for detection and identification of various diseases in tomato plants using a deep learning algorithm. Journal of Korean Society for Quality Management. 48(4):581-596.
11. Ni J, Zhou Z, Zhao Y, Han Z, Zhao LG. 2023. Tomato leaf disease recognition based on improved convolutional neural network with attention mechanism. Plant Pathol. 72(7):1335-1344.
12. Kim MK. 2020. Tomato crop disease classification using an ensemble approach based on a deep neural network. Journal of Korea Multimedia Society. 23(10):1250-1257.
13. Tripathi MK, Maktedar DD. 2020. A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey. Inf. Process. Agric. 7(2):183-203.
14. Yuan W. 2023. Accuracy comparison of YOLOv7 and YOLOv4 regarding image annotation quality for apple flower bud classification. Agri Engineering. 5(1):413-424.
15. He Z, Karkee M, Zhang Q. 2022. Detecting and localizing strawberry centers for robotic harvesting in field environment. IFAC-Papers Online. 55(32):30-35.
16. Lawal OM. 2021. Development of tomato detection model for robotic platform using deep learning. Multimed Tools Appl. 80(17):26751-26772.
17. Nugroho DP, Widiyanto S, Wardani DT. 2022. Comparison of deep learning-based object classification methods for detecting tomato ripeness. Int J Fuzzy Log Intell Syst. 22(3):223-232.
18. Yang Y, Song X. 2022. Research on face intelligent perception technology integrating deep learning under different illumination intensities. JCCE. 1(1):32-36.
19. Chen Z. 2022. Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm. JCCE. 1(3):103-108.