Current status and prospects of the visual detection and positioning technology for intelligent picking of famous tea

Yujie Zhou[1,4], Leiying He[1,2], Jianneng Chen[1,2*], Jiangming Jia[1,2], Chuanyu Wu[3,2], Yatao Li[1,2], Yikun Wei[1]

(1. *School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China*;
2. *Key Laboratory of Agricultural Intelligent Perception and Robotics of Zhejiang Province, Hangzhou 310018, China*;
3. *Zhejiang Ocean University, Zhoushan 316022, China*;
4. *School of Electronic Information, Zhejiang Business Technology Institute, Ningbo 315012, China*)

Abstract: The mechanization of famous tea harvesting is an essential way to develop China's tea industry. This paper centers on the detection and positioning technologies in famous tea harvesting, systematically reviewing research progress in these domains. In tea detection, traditional methods rely on color space selection and image segmentation, exhibiting limitations such as insufficient accuracy and poor generalization capability. Conversely, deep learning algorithms demonstrate superior detection accuracy and robustness. Current research focuses on enhancing detection accuracy, inference speed, and multi-variety recognition. In picking positioning, depth information measurement technology utilizing RGB-D cameras provides foundational support. Positioning methods have evolved from traditional visual processing techniques to deep learning and point cloud approaches, seeking to overcome challenges including occlusion and irregular growth patterns. Notwithstanding notable technological advancements, existing methods confront three primary limitations: difficulties in adapting to diverse growth stage characteristics, reliance on large-scale annotated datasets, and inadequate occlusion handling. Future research ought to concentrate on three directions: developing highly universal tea bud detection models, refining model training techniques for small-sample scenarios, and improving tea-picking point positioning accuracy under occluded conditions. This review aims to furnish critical references for advancing high-end intelligent tea-picking machinery, thereby facilitating the tea industry's mechanization and intelligentization.

Keywords: famous tea, detection, position, deep learning, intelligent picking
DOI: 10.25165/j.ijabe.20251806.9245

Citation: Zhou Y J, He L Y, Chen J N, Jia J M, Wu C Y, Li Y T, et al. Current status and prospects of the visual detection and positioning technology for intelligent picking of famous tea. Int J Agric & Biol Eng, 2025; 18(6): 1–11.

1 Introduction

Tea is a popular drink all over the world, and plays an important part in China's agricultural economy. In 2023, the total output of dry tea in China was 3.3395 Mt, with a total output value of 329.668 billion CNY[1], among which famous tea accounted for about 75%, while the cost of famous tea accounted for about 60% of the selling price of fresh leaves, according to statistics[2]. With expanding production and increasing demand, labor shortage during the tea-picking season increased, which has been a crucial factor restricting the development of the tea industry in China.

China has systematically researched and explored mechanized tea-picking technology since the 1950s[3]. After the 1980s, China stepped into the industrialization of tea-picking machines, during which different models were successfully launched[4]. At present, most machines on the market are of the cutting type, which operate by driving a cutter bar in a reciprocating motion to shear the tea shoots, thereby achieving efficient bulk harvesting. Compared with manual picking, the adoption of these machines has significantly improved picking efficiency and drastically reduced reliance on seasonal labor, thereby lowering overall production costs. However, this method cannot guarantee the quality of picking, so these machines are primarily suitable for the harvesting of bulk tea. For the famous tea, which demands careful selection of tea buds, selective manual picking remains essential[5].

The core of intelligent picking technology for famous tea lies in accurately detecting the tea bud under variable lighting and complex backgrounds, precisely locating picking points based on morphological features and stem-branch separation, and efficiently picking in various natural tea fields with minimal damage to the buds and surrounding leaves[6]. With the application of deep learning in the field of tea detection, effective identification and localization of most tea targets can now be achieved even in unstructured natural environments[7,8]. However, there are still some challenges to be solved, including occlusion by adjacent leaves, varying maturity levels of tea buds, adaptation to different weather conditions, and improving real-time processing speed.

In the past decades, the mechanized picking technology for famous tea has consistently remained a key research focus within the field of agricultural harvesting, which involves interdisciplinary

Received date: 2024-07-25 **Accepted date:** 2025-10-10
Biographies: Yujie Zhou, PhD candidate, research interest: agricultural engineering, Email: zyj_408@163.com; **Leiying He**, Associate professor, research interest: machine vision and manipulator control, Email: heleiying@163.com; **Jiangming Jia**, Associate professor, research interest: agricultural mechanization engineering, Email: Jarky@zstu.edu.cn; **Chuanyu Wu**, Professor, research interest: agricultural machinery equipment and technology, Email: cywu@zstu.edu.cn; **Yatao Li**, Lecturer, research interest: machinevision and agricultural robot, Email: ytli@zstu.edu.cn; **Yikun Wei**, Professor, research interest: machine learning, deep learning, Email: yikunwei@zstu.edu.cn.
***Corresponding author: Jianneng Chen**, Professor, research interest: agricultural machinery equipment and technology. Faculty of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China. Email: jiannengchen@zstu.edu.cn.

integration. Currently, several leading research institutions and universities have successfully developed prototype machines, and these prototypes have undergone preliminary picking experiments in actual tea gardens, as shown in Figure 1.

Figure 1 Tea-picking machines for famous tea

Focusing on the critical technologies for the intelligent picking of famous tea, this paper summarizes and analyzes the current research progress in detection and positioning. It points out existing technical limitations and practical challenges, such as accuracy under variable lighting conditions and adaptability to complex growth environments. Furthermore, the study offers insightful prospects for future development. This work provides valuable references for the research and development of picking machines for famous tea, and also suggests promising research directions for enhancing visual detection and positioning technologies in agricultural picking machines.

2 Target detection for famous tea

Famous teas have various appearances in natural environments (Figure 2). Firstly, visual characteristics like brightness, contrast, and color of tea can be significantly influenced by different lighting conditions. Secondly, it is difficult to distinguish the tea buds and their backgrounds with similar colors. Finally, feature information may be lost or covered due to the inevitable occlusion of tea leaves.

a. Different lighting conditions

b. Target be occluded

c. Occlusion

Figure 2 Tea buds in natural environments

There are some challenges in the visual detection of tea buds in various natural environments, and more difficulties are confronted in mechanized picking, which requires accurate identification and fast picking. Besides, the detection algorithm must meet the real-time request to guarantee the efficiency of picking, and a single algorithm should be able to identify multiple tea varieties. Accordingly, with the development of computing, visual detection of tea buds has recently become a research focus, mainly including traditional vision and deep learning.

2.1 Tea bud detection based on traditional vision

Tea bud detection based on traditional vision generally distinguishes the targets from backgrounds through image segmentation based on characteristics such as color, shape, and texture. There are three main processes of image segmentation. Firstly, the input image needs to be processed to eliminate environmental noise and improve the image quality. Secondly, the appropriate color space will be selected, and the specific algorithm will amplify characteristic disparities between the target and background. Finally, a segmentation algorithm separates the target pixels from the background pixels.

2.1.1 Color space selection

The color space needs to be selected to realize the segmentation of tea buds based on color. In early research, segmentation is generally based on RGB color space. Yang et al.[9] extracted the G (RGB) for segmentation, verifying the feasibility of using the visual method for detecting tea buds. However, the limited information carried by a single channel makes it difficult to distinguish between tender and old leaves. The RGB color space is significantly impacted by lighting, so the difference in components in RGB is often used for image segmentation. Wu et al.[10] analyzed the detection effect of different component combinations in RGB color space, which showed that G-B (RGB) could distinguish tea buds from old leaves and suppress the background effectively.

The detection method based on RGB color space has poor robustness in natural environments[11], and other color spaces, such as HSI, HSV, and Lab, which are less affected by lighting conditions, will also be applied to tea bud detection. Zhao et al.[12] compared the segmentation effect in HSV and HSI color spaces, and found that the characteristics of the tea buds were relatively prominent in HSI color space. Furthermore, combining components between different color spaces is an effective method for segmentation. Zhang et al.[13] selected the G-B (RGB) and the b (Lab) for segmentation to eliminate the correlation in RGB color space. Xia et al.[14] combined five components using SLIC super-pixel to achieve tea bud detection, and experiments showed that the average accuracy was 16.6% higher than that based on the G-B component.

The color-based detection method can achieve better results with less occlusion. However, in dense environments, it is necessary to predict by combining the morphological and textural features of the tea leaves.

2.1.2 Image segmentation

The purpose of image segmentation is to separate the background from the tea buds effectively. Currently, the mainstream tea bud segmentation method is mainly based on color segmentation, and sometimes, the segmentation results are post-processed by morphology.

Color-based segmentation method separates the target and background by selecting the appropriate threshold interval, such as the iterative threshold method[15], OSTU method, Bayesian method[16], *K*-means clustering method, and support vector machine

(SVM) method. Both the *K*-means clustering method and the SVM method achieve image segmentation by dividing pixels into target and background, and the OSTU method obtains the segmentation threshold by calculating the maximum variance between the classes of the histogram of the gray image. Compared with the three segmentation methods (Table 1), the *K*-means method exhibits satisfactory overall recognition accuracy, but it has poor real-time performance when dealing with high-resolution images, and it is also relatively sensitive to isolated points[17]. Furthermore, the parameter *K* has a particular impact on the results. The OSTU method is less affected by image size, but its accuracy is relatively low, thus requiring the incorporation of morphological methods. The SVM method can integrate features such as color, texture, and shape for segmentation, but its detection performance relies heavily on image preprocessing and data transformation due to its lack of interpretability for high-dimensional data.

Table 1 Comparison of segmentation method

Method	*K*-means	OSTU	SVM
Precision	+++	+	+
Robustness	+	+++	++
Limitation	Sensitive to outlier data; Need to preset *K*; Size-sensitive speed	Low precision; Image hole; Incomplete segmentation	Poor universality
References	[13, 14, 17, 18, 21]	[10, 15, 22, 23]	[19, 20]

Color-based segmentation methods ignore the morphological features of tea leaves, resulting in issues such as internal holes and incomplete segmentation[24]. The morphological algorithms are generally used as a follow-up to color-based segmentation to obtain a complete tea bud. Zhang et al.[25] proposed an improved watershed segmentation algorithm to reduce the over-segmentation rate for low differentiation between tea buds and old leaves. Jiang et al.[26] improved the watershed algorithm by introducing the BM3D denoising algorithm and grayscale stretching method, further enhancing the distinguishability of tea buds and background by increasing the contrast.

In summary, the research has proved the feasibility of traditional vision-based detection for tea bud detection, and it can adapt to changes in the environment to a certain extent. However, it still cannot meet the application requirements in accuracy and generalization. Therefore, there have not been many reports about intelligent tea-picking robots before 2020. To speed up the process of mechanized picking, we need a more extensive method of tea bud identification.

2.2 Tea bud detection based on deep learning

Deep learning is a data-driven method that can precisely learn the complex features of tea buds through training on many tea samples. Compared to traditional visual detection methods, the detection method based on deep learning can capture and represent features in higher dimensions[27], so it can compensate for the shortcomings of traditional algorithms that rely solely on color, shape, or texture features in terms of recognition robustness. Deep learning methods have shown potential in detection and segmentation due to their excellent detection accuracy and robustness[28]. The detection process based on deep learning is shown in Figure 3. The detection algorithms based on deep learning can be divided into two categories: one-stage algorithm and two-stage algorithm.

In two-stage algorithms, the proposal boxes are first generated based on feature extraction from the image and then regressed to get the tea bud targets[29]. Zhu et al.[30,31] explored the application of the Faster R-CNN (Figure 4) in detecting tea buds under complex backgrounds. Experiments showed that the deep learning method was significantly better than the traditional one in detection accuracy and speed, and it also had good performance in complex backgrounds. Wang et al.[32] proposed a tea detection model based on Mask R-CNN which uses resnet50 and feature pyramid network (FPN) to extract features. The experimental results showed that the AP reached 93.95% and the AR reached 92.48%.

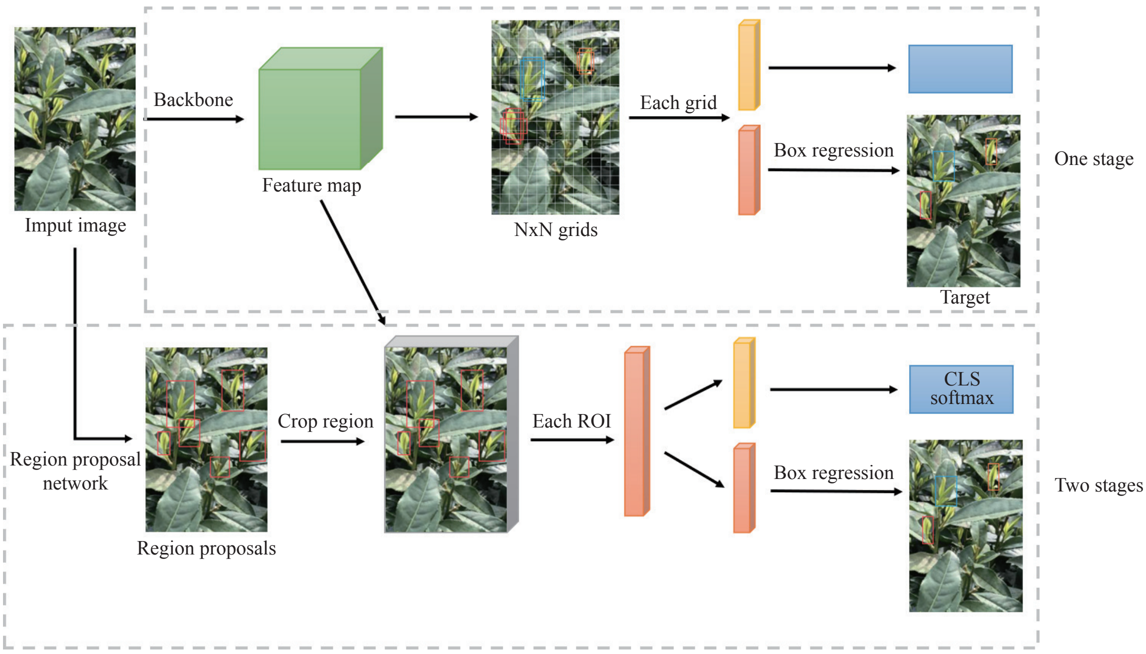

Figure 3 Flowchart of deep learning detection

Although two-stage detection methods have higher precision, their algorithms are time-consuming. Another category is one-stage detection algorithms, which integrate the detection process and directly provide detection results from RGB images. YOLO[33] and

SSD[34] are representative detection networks. Chen et al.[35] developed a detection method based on SSD and achieved an AP of 83.9% in the experiment. Li et al.[36] implemented a tea bud detection algorithm utilizing a lightweight compressed YOLOv3 and conducted experiments in the natural environment, achieving mAP of 85.16%.

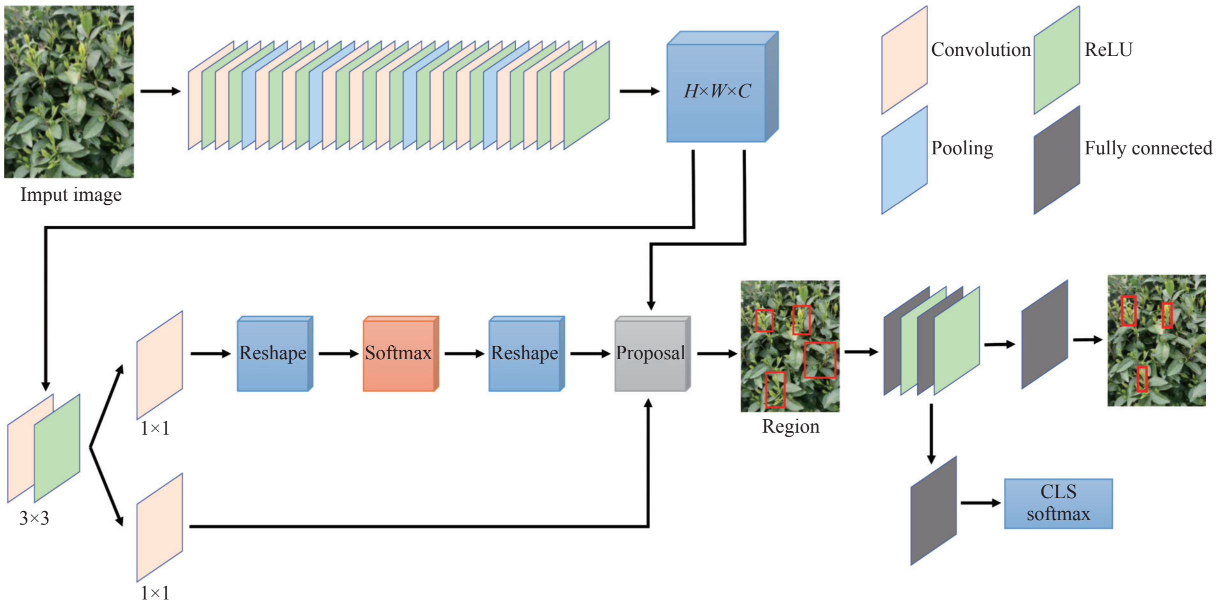

Figure 4 Faster-RCNN architecture diagram

With the profound application of deep learning in tea bud detection, significant progress has been made in the research and development of tea-picking robots. Yet, there is still a considerable gap before practical commercial application[37]. The mechanized picking of famous tea poses higher requirements on detection algorithms, especially in three key aspects: detection precision, inference speed, and multi-varietal recognition.

2.2.1 Methods for improving detection precision

The detection accuracy of famous tea directly impacts the precision of subsequent tea picking. Current research primarily focuses on two core dimensions: backbone networks and neck networks. The backbone network is a critical component of feature extraction. Its design focuses on improving the sensitivity and recognition ability of the model to small tea, strengthening the semantic representation ability of high-dimensional features, and lightweight design of the model under the premise of maintaining high accuracy to meet the deployment requirements of edge devices. The neck network performs multi-level feature fusion and semantic enhancement, involving cross-scale feature integration mechanisms, construction of context semantic perception module, strategies for overlap and occlusion, recognition structure with strong generalization for different varieties of tea images, and lightweight convolution design for real-time detection. Both of them determine the accuracy, robustness, and practical application value of the tea recognition model.

The backbone is responsible for feature extraction from the input image. Xu et al.[38] employed the Faster R-CNN model based on VGG-16, ResNet-50, and ResNet-101 as the backbone to train samples of tea bud respectively. The experimental results showed that the VGG-16 had an excellent recognition effect as compared to the others. Xu et al.[39] integrated the DenseNet201 as the backbone of YOLOv3 model, presenting a two-level fusion network detection method with a variable universe. The DenseNet201 realizes a novel dense connection mechanism that concatenates features between layers. This method improves detection speed and ensures detection precision, achieving a mAP of 95.71%.

A densely connected feature extraction network is conducive to the extraction of high-dimensional semantic information, but it will lead to a decline in the generalization ability of the model when the network is too dense[40]. With the outstanding performance of the attention mechanism in image processing[41], researchers have embarked on integrating it into tea bud detection. Li et al.[42] proposed a detection model based on YOLOx, which used the transformer as the backbone to enhance the overall detection precision and achieved a 5.73% higher accuracy than the original YOLOx. Unlike the Convolutional Neural Network (CNN), the attention mechanism considers pixels' information, enabling the model to capture the dependencies between pixels and better understand the image. Gui et al.[43] introduced a convolutional block attention module (CBAM, Figure 5) into the YOLOv3-SPP model. The CBAM combines channel attention and spatial attention mechanism, which enables the network to capture the details of the image and further improve the effectiveness of feature extraction for small targets in dense scenes.

The neck is responsible for processing the features extracted by the backbone, and integrating them to enhance their expressive ability and robustness. Yang et al.[44] utilized an image pyramid structure in the neck to fuse different scales of tea features. Fang et al.[45] adopted a bidirectional feature pyramid network (BiFPN) to integrate the feature map of various scales, which greatly reduced the missing rate for tiny buds. Gui et al.[46] proposed a lightweight tea bud detection model based on YOLOv5 with a weighted feature fusion mechanism in the neck to fuse the low-level and high-level features efficiently.

Deepening the backbone network can improve the detection precision for small tea targets. However, an excessively deep backbone network may lead to overfitting issues[47]. Although CNN boasts advantages such as local perception and parameter sharing, it may fall short in capturing global feature correlations, which results in a low detection rate. The attention mechanism can adaptively adjust the region of interest according to the input data, which often performs better in tea detection. However, this mechanism has a significant computational complexity, and it is necessary to consider the lightweight design in applications[48]. The neck network is responsible for fusing the features output by the backbone network. Low-level features encapsulate positional and detailed information[49], whereas high-level features encode semantic information. Integrating these features through diverse fusion mechanisms enhances the precision and robustness of detection tasks.

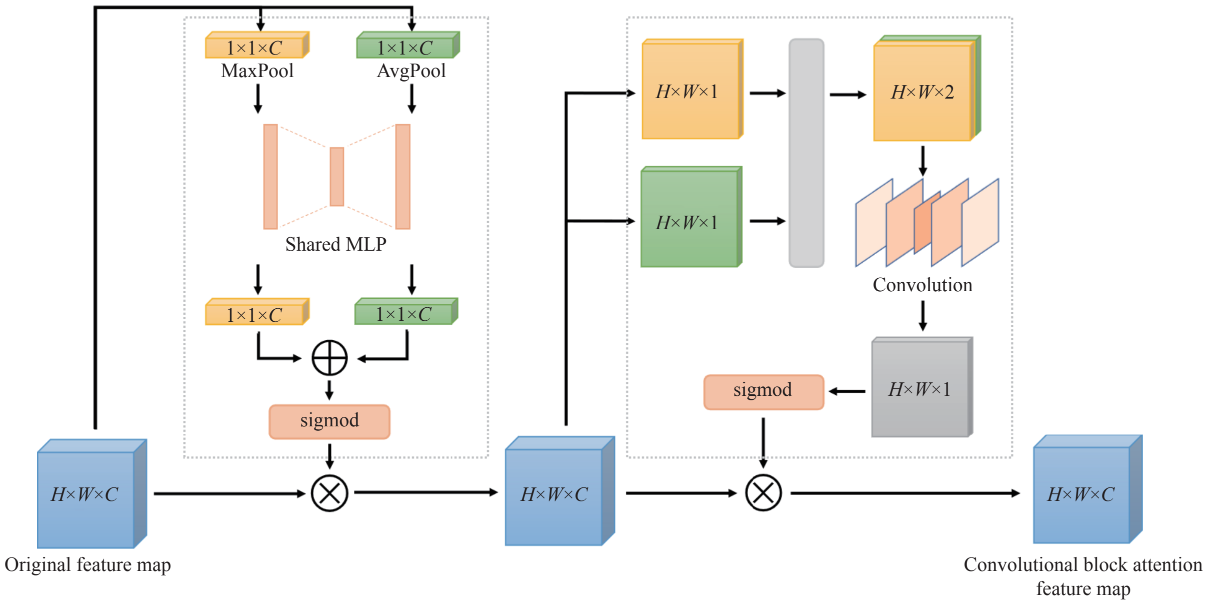

Figure 5 CBAM unit architecture diagram

2.2.2 Methods for enhancing inference speed

Tea bud detection methods based on deep learning exhibit outstanding performance[50]. However, in the feature extraction process, the number of model parameters also increases sharply with the gradual increase in the number of network layers, which hinders the application of the model on mobile terminals, thus restricting the development of tea-picking robots[51]. To reduce the model size and inference time, the research is mainly carried out by designing the lightweight model and performing pruning operations on the trained model[52].

At present, the mainstream idea is to design a model with fewer parameters, such as GhostNet[53], MobileNet[54], and ShuffleNet[55], generating more features with a smaller parameter size. Lin et al.[56] utilized a lightweight YOLOx-S model to detect the tea buds. Li et al.[57] proposed a lightweight model based on YOLOv4, which replaces the backbone with GhostNet, reducing the computational complexity. Compared with the original YOLOv4, the proposed model's computational complexity and the number of parameters were reduced by 89.11% and 82.36%. Huang et al.[58] improved the backbone and neck of the YOLOv4 using the Ghost Module (Figure 6), thereby reducing the model's memory by 80%. Then, the optimized model was deployed on the RK3568 development board, with an average inference time of 67.1 ms and a mAP of 70.2%.

Another way to reduce the weight is to compress the trained model, and the mainstream methods are pruning and sparseness[59]. The pruning operation reduces the model's size by removing unnecessary weights or biases from the neural network; the sparseness sets the redundant parameters in the neural network to zero, reducing the computational complexity of the model. Li et al.[60] utilized channel and layer pruning algorithms to compress the YOLOv3 and maintain precision by fine-tuning the model. The model was deployed on the Jetson Xavier NX platform, and the experiments showed that the inference time of the model was reduced by 59%, while the precision was only reduced by 0.40%.

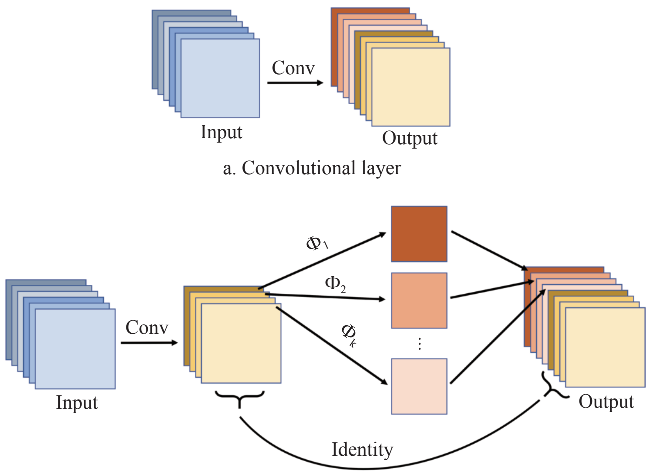

Figure 6 Comparison between Ghost module and convolution module

The implementation of lightweight design substantially reduces model size and decreases inference latency. However, this compression frequently induces a degradation in the model's representational capacity. Consequently, achieving an optimal equilibrium between computational efficiency and predictive accuracy is imperative in practical applications,and further fine-tuning of the compressed model is typically required to approximate the original accuracy level.

2.2.3 Methods for multi-varietal recognition

Detection models for famous tea are primarily based on specific harvest times and a single variety of tea. However, there are thousands of types of tea in the world[61]. Therefore, enhancing the generalization of the model and enabling it to possess the ability to detect multiple types of tea at the same time will be of great significance.

Because of the similar characteristics of tea varieties, the general model often needs many samples to distinguish two kinds of tea when training. To realize multi-species recognition with limited samples, Yu et al.[62] introduced the ECA-Net (Figure 7) into the YOLOx, enhancing the representation capability for similar features by allowing the network to adjust its focus on different feature channels dynamically. Zhao et al.[63] proposed a model suitable for multiple varieties of tea by introducing the ECA-Net into YOLOv7. The proposed model was compared with the YOLOv7 and YOLOv7+CBAM in three datasets, which showed that the proposed model had a higher precision than other models in each variety. The ability to represent similar features can be enhanced by designing a dedicated backbone. However, this model needs to be trained on a large dataset, which may lead to overfitting of the model[64].

The fusion of multi-modal data can also increase the recognition performance of the model for multiple species. Wu et al.[65] designed an end-to-end RGB-D detection model that integrates depth information into the YOLOv7 model. The Cross-modal Spatial Attention Fusion Module (CSFM) was intended to co-fuse depth features with RGB features in a unidirectional manner, which can learn the morphological characteristics of tea buds and help distinguish tea varieties. The multi-modal fusion method was improved in multi-variety recognition, but it also needs a high cost in terms of sample labeling.

In summary, the method based on deep learning can effectively improve the accuracy and robustness of tea bud detection, which leads the intelligent tea-picking robots to develop rapidly. Deep learning technology has broad application potential in agricultural robots, but it is still insufficient for tea bud detection. At present, the detection lacks effective methods in the face of occlusion, which is also an issue that the follow-up research needs to solve.

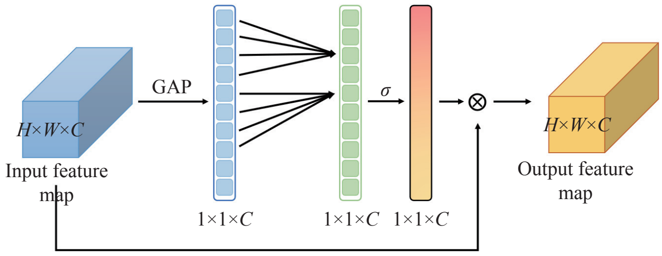

Figure 7 ECA-Net architecture diagram

3 Picking points positioning for famous tea

Positioning picking points is a crucial step in the mechanized picking of famous tea. The picking points are inferred by combining tea bud ROI and depth map, and then the coordinate is transmitted to the robotic arm to guide the picking manipulator to the designated position. There are two kinds of technical paths to achieve picking points positioning (Figure 8).

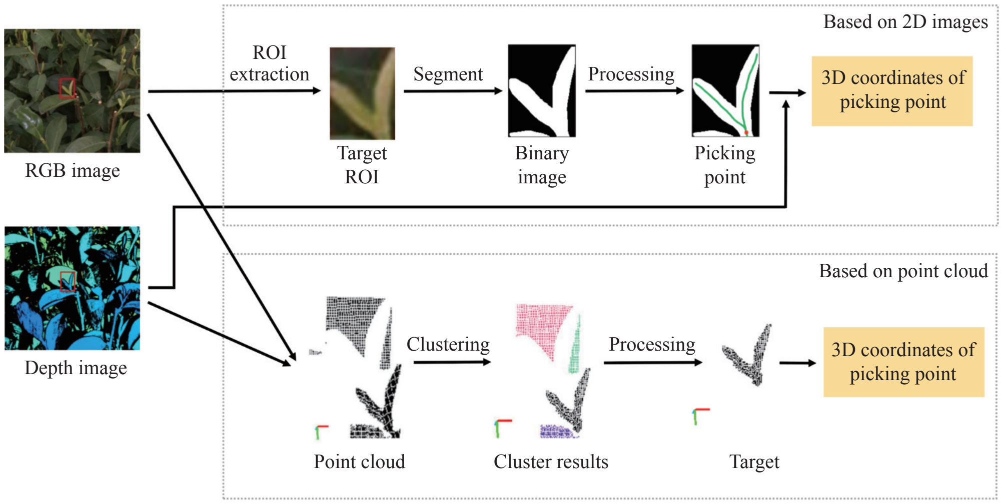

Figure 8 Technical path of picking points positioning

(1) Infer the most suitable picking points in the RGB image, then calculate the spatial coordinates of the picking points combined with the depth map.

(2) Obtain the point cloud by fusing the RGB image and the depth map, then calculate the picking points directly using the point cloud processing algorithm.

Famous tea is mainly picked according to the standards of "one bud one leaf" or "one bud two leaves"[66,67], and the picking point is generally selected about 3 mm below the intersection of bud and leaf. Tea buds often face occlusion or overlap in natural environments, making the feature unable to be fully captured, resulting in positioning errors and affecting the success rate of picking[68]. Therefore, the first difficulty to overcome in positioning picking points is occlusion. According to the principles of positioning methods, they can be categorized into image-based and point cloud-based positioning methods.

3.1 Depth information measurement technology

An essential challenge for agricultural harvesting systems is locating the target, and the first step is accurately measuring the depth map. At present, the popular techniques for measuring the depth map encompass Laser Range Finding[69,70], Ultrasonic Ranging[71,72], Motion Estimation[73], and RGB-D Camera-based Ranging[74,75]. The laser ranging method and ultrasonic ranging method have high hardware requirements and can measure only one point, which is not suitable for tea picking. The motion estimation method has large precision errors. In contrast, the RGB-D camera has been widely used in agricultural picking because of its precision. According to the principle, the RGB-D camera is mainly divided into the binocular camera, structured light camera, and TOF (Time-of-Flight) camera[76], and the comparison of characteristics is listed in Table 2.

The binocular camera identifies corresponding pixels in two images, and then the depth map is determined through the geometric relationship[77] (Figure 9). It is difficult to extract effective features in the case of insufficient illumination and lack of texture, and it is often unable to generate a dense depth map.

Table 2 Comparison of different types of RGB-D cameras

Type	Binocular camera	Structured Light camera	TOF camera
Advantage	Simple structure; Low cost; Light insensitive	High precision and resolution; Texture insensitive	Light insensitive; Texture insensitive; Far distance
Distance	++	+	+++
Precision	++	+++	+
Limitation	Texture sensitive; High computational complexity; Cannot work at night	Light sensitive; Precision reduction in long-distance	Reflections sensitive; Low precision at edge

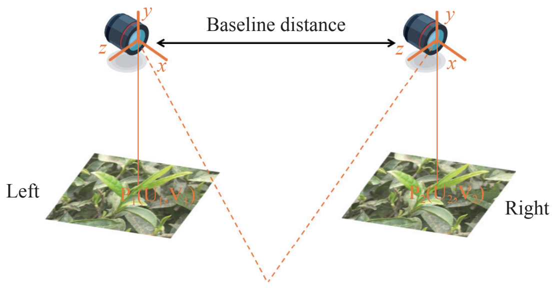

Figure 9 Schematic diagram of binocular camera

The structured light camera projects encoded light onto objects, and the depth map is calculated by analyzing reflected light patterns. This method is sensitive to light, and its performance drops sharply under strong light[78]. The TOF camera achieves 3D positioning by continuously emitting light pulses to the object, receiving the reflected signals, and measuring the round-trip time of the light pulses to calculate the depth map. Compared to other RGB-D cameras, TOF cameras perform 3D positioning faster and have higher precision, but TOF cameras generally have low resolution and are expensive. With the development of depth measurement technology, RGB-D cameras have become an important component of machine vision systems, widely used to recognize and position tea buds.

3.2 Positioning methods based on image

Due to the limitation of hardware, early positioning methods generally determine the picking points through RGB images. Image-based positioning methods can be divided into traditional vision processing and deep learning.

3.2.1 Positioning by traditional vision processing

The traditional vision processing methods mainly utilize morphological knowledge to select picking points, which can be divided into minimum bounding rectangle (MBR) algorithm and skeleton-based algorithm.

The MBR algorithm determines the picking points by selecting the minimum outer rectangle of tea buds. Pei et al.[79] extracted the bounding rectangle of tea buds and took the center of the rectangle as the picking point. However, the picking points can not be positioned precisely only using the MBR algorithm. Lu et al.[80] used the centroid of tea buds and the external rectangle to calculate the picking points. Compared with the MBR algorithm, the skeleton-based algorithm can separate the tea stem and determine the picking points. Long[81] extracted the skeleton of the tea bud by using the Canny operator, and the lowest point of the skeleton was used as the picking point. Zhang et al.[82] refined the tea bud skeleton by using the expansion and corrosion, and then used the Shi-Tomasi algorithm to calculate the picking points.

The above methods provide an effective way to obtain picking points, but they still have significant limitations. All of them have specific requirements for the growth form of tea buds[83]. Compared with the MBR algorithm, the skeleton-based algorithm has better accuracy and robustness, but it cannot work when the bud is occluded.

3.2.2 Positioning by deep learning

In deep learning-based methods, the picking points are considered in the designing stage of the model and selected directly according to the image. Chen et al.[84] added a fully convolutional network (FCN) to the regression head of the Faster R-CNN model to predict the picking points. Shuai et al.[85] introduced the self-attention mechanism in the YOLOv5 to improve the selection effect of picking points in a large field of view.

In addition to the direct output of picking point coordinates, the method based on deep learning can also output the appropriate picking area, and then the picking points are selected by the traditional image processing method. Zhang et al.[86] proposed the MDY7-3PTB model for tea bud detection and picking point positioning (Figure 10). The recommended picking area was detected through the proposed model, with the center of the picking area as the picking point.

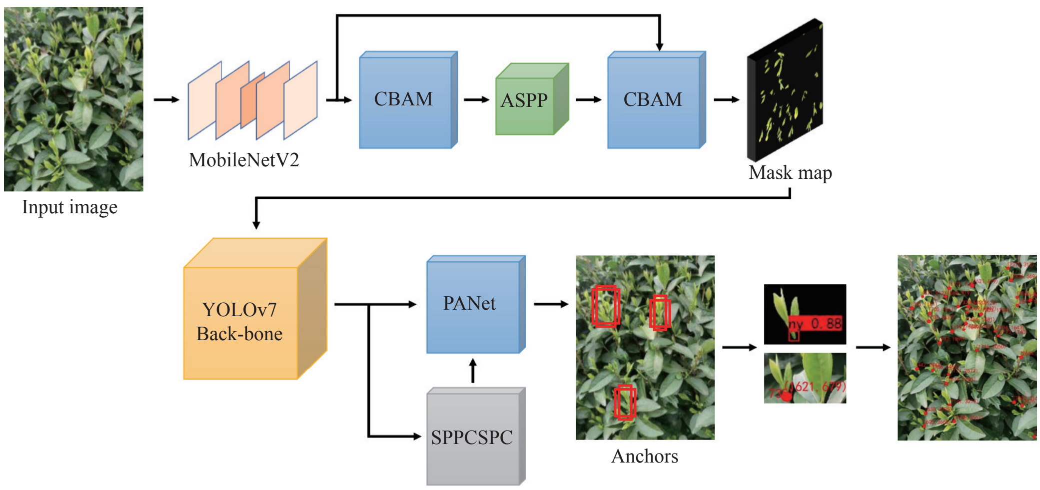

Figure 10 MDY7-3PTB architecture diagram

Meng et al.[87] proposed a model combining tea bud detection and picking area segmentation. The improved YOLOx model was used to obtain the tea bud. On this basis, the improved PSP-Net was used to segment the picking area, and the appropriate picking points were calculated using the centroid method.

The comparison of image-based positioning methods is listed in Table 3. Compared with traditional methods, the deep learning models show better performance in dealing with the influence of factors such as lighting change, perspective change, and occlusion, which enables them to maintain precision in complex environments[88].

Table 3 Comparison of image-based positioning methods

Method	Advantages	Limitations	Evaluation
MBR	Low cost; Simple	Poor precision; Poor universality	+
Skeleton-based	Direct; Good robustness	Occlusion; Light sensitive	++
Deep learning with points	Good universality	Extra labeling; Poor robustness	++
Deep learning with area	High robustness; High precision	Extra labeling	+++

However, deep learning methods typically require extensive annotated data, particularly for specific picking regions or points, which increases the labor costs. Moreover, the model's performance isconstrained by the quality of dataset. The poor samples may impair the model's generalization and robustness, ultimately leading to reduced effectiveness in practical applications[89].

3.3 Positioning methods based on point cloud

It is difficult to achieve accurate positioning only by relying on RGB images due to the irregular growth and occlusion of tea buds. Therefore, to improve the positioning accuracy, it is necessary to combine the depth information. Point cloud is the primary form of three-dimensional information, which contains a wealth of spatial relationships. In the 3D positioning method, the depth information is usually converted into the point cloud, and the picking points are identified based on the point cloud[90].

Point cloud positioning should first extract the target point cloud from the ROI[91]. Because of the disorder of point cloud data, clustering is usually used to extract it. Li et al.[92] extracted the point cloud of tea buds by Euclidean clustering and then determined the coordinates of the picking point according to the growth characteristics of tea buds. Zhu et al.[93] applied DBSCAN clustering to extract the point cloud of tea buds, and then the picking point was selected using the minimum bounding box of the point cloud. Because of the density of the point cloud, it often takes a lot of time to calculate. Chen et al.[94] proposed an optimal pose-vertices search method (OPVSM) to find the best vertex of the tea bud from the point cloud, the voxel method was used to de-sample the point cloud, effectively improving the efficiency of the algorithm. The algorithm flow is shown in Figure 11.

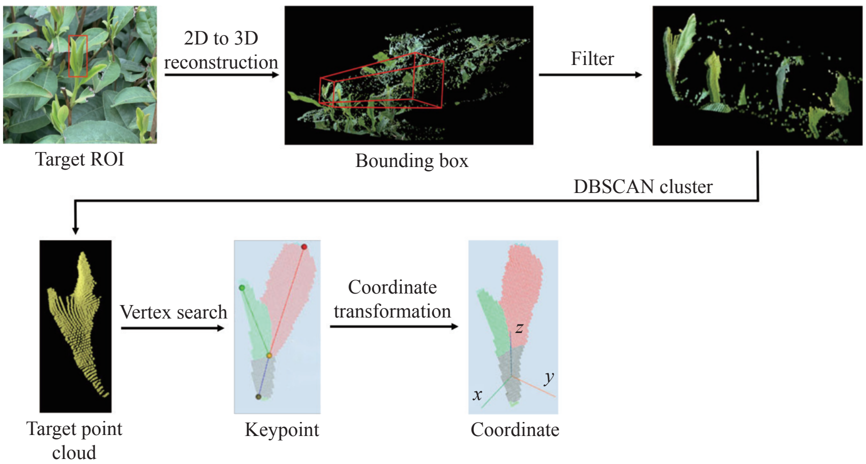

Figure 11 Flowchart of OPVSM algorithm

In summary, the picking point positioning of tea buds is gradually transitioning from 2D positioning to 3D positioning[95]. This method is only realized by traditional point cloud clustering, and there is no case realized by deep learning technology. When processing the point cloud, a large number of calculations are usually involved, which results in that the positioning speed is closely related to the density of the point cloud. At the same time, the density of the point cloud also affects the positioning accuracy. A balance should be achieved between accuracy and speed when designing a model.

4 Prospect of detection and positioning technology

Although the current research on the detection and positioning of famous tea has made a lot of progress, there are still significant limitations in the application. Compared with the traditional visual method, the method based on deep learning has a good potential for application, but there are still many technical difficulties to overcome. First is the morphology of tea bud changes in the tea growth cycle, which leads to large fluctuations in the detection effect[96]. Secondly, the deep learning-based method needs many tea samples, requiring a lot of time to label them. Finally, with the increased complexity of the model, the detection efficiency will inevitably be affected. The current algorithms of positioning the picking point seem to be simplified and fail to fully consider the spatial characteristics of tea buds, which leads to the accuracy being reduced when occluded. Moreover, the deep learning-based method lacks an effective solution for occlusion.

To achieve efficient, accurate and intelligent picking of famous tea with all-weather and complete growth cycle, the intelligent picking technology of famous tea will focus on the following aspects:

4.1 Tea bud detection with high generalization

Deep learning-based detection has dramatically improved compared to traditional visual detection. However, the current detection network can only detect a specific variety of tea. Moreover, the morphology of famous tea changes significantly during the growth cycle, which leads to the instability of the network detection in the growth cycle. To solve this problem, a particular backbone network can be designed to strengthen the distinction between similar features so that the model can simultaneously detect different varieties of tea. In addition, a detection network with time features can be designed by introducing the time series network to combine the static and dynamic features.

4.2 Model training for small samples

The deep learning-based model requires many samples to achieve effective detection results, which requires much work. As one of the solutions, semi-supervised training can be introduced to use a small number of labeled samples and a large number of unlabeled samples for training to improve the learning effect. In addition, migration learning can also be used for small-sample training. A small number of tea samples can fine-tune the existing model, and the model will be quickly trained to detect a new variety of tea.

4.3 Picking point positioning in occlusion

Compared with RGB images, the point cloud contains the spatial relationship of tea buds. The morphological characteristics of tea buds can be learned by using deep learning technology for point clouds to eliminate the influence of light, background, and other factors[97]. In the face of occlusion, three-dimensional reconstruction of tea ridge can be carried out by multi-view image fusion technology to obtain complete point cloud information and provide accurate environmental data for picking point positioning and obstacle avoidance of picking movement.

5 Conclusions

The mechanized picking of famous tea is of great significance to China's tea industry, and the detection and positioning technology of famous tea is an important part of realizing mechanized picking. In this paper, the detection and positioning technology of famous tea is reviewed, and its development status and future trends are discussed. Thanks to deep learning technology, the detection and positioning of famous tea have significantly improved in recent years, but the algorithm still has limitations.

Given the current research, this paper points out three directions for future research: 1) highly generalized tea bud detection technology, 2) model training technology for small samples, and 3) picking point positioning technology in an occlusion environment, to provide a reference for the follow-up study of famous tea detection and positioning, thus promoting the development of mechanized tea picking.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (Grants No. U23A20175, 52305289, and 32472009) and China Agriculture Research System of MOF and MARA (CARS-19).

[References]

[1] Mei Y, Liang X. Analysis of China's tea production, sales, import and export situation in 2023. China Tea, 2024; 46(4): 18–26.

[2] Liu W, Xiong R. Research on the key technology of intelligent tea picking machine-measurement of depth and height of tea gathering in visual processing system. in Proc. Biotech Middle East 2017, Dubai, United Arab Emirates, 2017; pp.146–156. DOI: 10.25236/isafb.2019.025

[3] Quan Q A. The development of tea picking machinery and the development of the mechanization of tea picking in China. Chinese Tea, 2018; 40(10): 4–9.

[4] Han Y, Xiao H R, Song Z Y, Chen Q M, Ding W Q, Mei S. Design and experiments of 4CJ-1200 self-propelled tea plucking machine. Int J Agric & Biol Eng, 2021; 14(6): 75–84.

[5] Wei Y, Wen Y Q, Huang X L, Ma P H, Wang L, Pan Y, et al. The dawn of intelligent technologies in tea industry. Trends in Food Science & Technology, 2024; 144: 104337.

[6] Zhou Y J, Wu Q, He L Y, Zhao R M, Jia J M, Chen J N, et al. Design and experiment of intelligent picking robot for famous tea. Journal of Mechanical Engineering, 2022; 58(19): 12–23.

[7] Li Y T, Zhou Y J, Wang S Q, Chen J N, He L Y, Jia J M, et al. Experimental study on high-quality tea plucking by robot. Journal of Tea Science, 2024; 1: 75–83. DOI: 10.3969/j.issn.1000-369X.2024.01.007 (in Chinese)

[8] Wang Z H, Xun Y, Wang Y K, Yang Q H. Review of smart robots for fruit and vegetable picking in agriculture. Int J Agric & Biol Eng, 2022; 15(1): 33–54.

[9] Yang F Z, Yang L, Tian Y N, Yang Q. Recognition of the tea sprout based on color and shape features. Transactions of the CSAM, 2009; 40(S1): 119–123. (in Chinese)

[10] Wu X M, Zhang F G, Lu J T. Research on recognition of tea tender leaf based on image color information. Journal of Tea Science, 2013; 33(6): 584–589.

[11] Wei J J, Chen Y, Jin X J, Zheng J Q, Shi Y Z, Zhang H. Researches on tender tea shoots identification under natural conditions. Journal of Tea Science, 2012; 32(5): 377–381. (in Chinese)

[12] Zhao B J, Dong W, Sun W Z, Liu Y. Research on tea bud identification technology based on HSI/HSV color transformation. in Proc. 6th Int. Conf. Inf. Sci. Control Eng., Shanghai, China, 2019; pp.511–515. DOI: 10.1109/ICISCE48695.2019.00108

[13] Zhang K, Lu J. Study on automatic segmentation of tea sprouts under natural conditions. Journal of Heilongjiang Bayi Agricultural University, 2016; 28(2): 100–104.

[14] Xia H K, Fang M R, Huang T, Lu J. Study on the method of image segmentation of tea sprouts based on SLIC super-pixel. Journal of Xichang University (Natural Science Edition), 2019; 33(4): 4.

[15] Tang X, Wu X M, Zhang F G, Gu J M. Contrastive research on tender tea recognition based on multiple threshold segmentation methods. Agricultural Equipment&Technology, 2013; 39(6): 10–14.

[16] Zhang L, Zhang H D, Chen Y D, Dai S H, Li X M, 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.

[17] Wu X M, Tang X, Zhang F G, Gu J M. Tea buds image identification based on Lab color model and *k*-means clustering. Journal of Chinese Agricultural Mechanization, 2015; 36(5): 161–164.

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

[19] Chen M T. Recognition and location of high-quality tea buds based on computer vision. Qingdao University of Science and Technology, 2019.

[20] Li W, Chen R, Gao Y. Automatic recognition of tea bud image based on support vector machine. International Conference on Advanced Hybrid Information Processing, 2020; 348: 279–290.

[21] Vu H, Le T-L, Tran T-H, Nguyen T-T. A vision-based method for automatizing tea shoots detection. International Conference on Image Processing, Melbourne: 2013; pp.3775–3779 DOI: 10.1109/icip.2013.6738778

[22] Huang H J, Wu M H, Wang X W, Liu S, Zhou J. Image recognition of tender leaves based on improved watershed algorithm in tea. Guizhou Agricultural Science, 2018; 46(4): 136–138.

[23] Fang Q, Xie Z Q, Tang Z, Chen H R. Related study based on OSTU watershed algorithm and new squeeze and excitation networks for segmentation and level classification of tea buds. Neural Processing Letters, 2021; 53(3): 2261–2275.

[24] Karunasena G, Priyankara H. Tea bud leaf identification by using machine learning and image processing techniques. International Journal of Scientific Engineering Research, 2020; 11(8): 624–628.

[25] 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.

[26] Jiang H T, He B X, Zhang Y F. Research on the method of recognizing and positioning the shoots of the tea picking manipulator. Machinery & Electronics, 2021; 39(7): 60–64, 69.

[27] Sun X X, Mou S M, Xu Y, Cao Z H, Su T. Detection algorithm of tea tender buds under complex background based on deep learning. Journal of Hebei University, 2019; 39(2): 211–216.

[28] Kamilaris A, Prenafeta-Boldu F X. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 2018; 147(1): 70–90.

[29] Ren S Q, He K M, Ross G, Sun J. Faster R-CNN: Towards real-time object

detection with region proposal networks. Transactions on Pattern Analysis & Machine Intelligence, 2017; 39(6): 1137–1149.

[30] Zhu H C, Li X, Meng Y, Yang H B, Xu Z, Li Z H. Tea bud detection based on Faster R-CNN network. Transactions of the CSAM, 2022; 53(5): 217–224.

[31] Xu B Y, Gao Y F. Tea-buds multi-dimensional recognition with Faster-RCNN deep learning method and its performance analysis. Agricultural Equipment & Vehicle Engineering, 2023; 61(2): 19–24.

[32] Wang T, Zhang K M, Zhang W, Wang R Q, Wan S M, Rao Y, et al. Tea picking point detection and location based on Mask-RCNN. Information Processing in Agriculture, 2023; 10(2): 267–265.

[33] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. Conference on Computer Vision and Pattern Recognition, Las Vegas: 2016; pp.779–788. DOI: 10.1109/cvpr.2016.91

[34] Berg A C, Fu C Y, Szegedy C, Anguelov D, Erhan D, Reed S. SSD: Single shot MultiBox detector. European Conference on Computer Vision, Amsterdam: 2016; pp.21–37. DOI: 10.1007/978-3-319-46448-0_2

[35] Chen B, Yan J, Wang K. Fresh tea sprouts detection via image enhancement and fusion SSD. Journal of Control Science and Engineering, 2021; 1: 1–11.

[36] Li Y T, Wu S K, He L Y, Tong J H, Zhao R M, Jiang J M, et al. Development and field evaluation of a robotic harvesting system for plucking high-quality tea. Computers and Electronics in Agriculture, 2023; 206: 107659.

[37] Zhao X Y, He Y X, Zhang H T, Ding Z T, Zhou C A, Zhang K X. A quality grade classification method for fresh tea leaves based on an improved YOLOv8x-SPPCSPC-CBAM model. Scientifc Reports, 2024; 14: 4166.

[38] Xu G J, Zhang Y, Lai X Y. Recognition approaches of tea bud image based on faster R-CNN depth network. Journal of Optoelectronics Laser, 2020; 31(11): 1131–1139.

[39] Xu W K, Zhao L G, Li J, Shang S Q, Ding X P, Wang T W. Detection and classification of tea buds based on deep learning. Computers and Electronics in Agriculture, 2022; 192: 106547.

[40] Liu F, Wang S D, Pang S C, Han Z. Detection and recognition of tea buds by integrating deep learning and image-processing algorithm. Journal of Food Measurement and Characterization, 2024; 18(4): 2744–2761.

[41] Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, et al. Swin Transformer: Hierarchical vision transformer using shifted windows. International Conference on Computer Vision, Montreal: 2021; pp.9992–10002. DOI: 10.48550/arXiv.2103.14030

[42] Li X T, Liu R X, Li Y X, Li Z L, Yan P, Yu M, et al. An improved model based on YOLOX for detection of tea sprouts in natural environment. Evolving Systems, 2024; 15(5): 1665–1679.

[43] 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): 14556.

[44] Yang H L, Chen L, Chen M T, Ma Z B, Deng F, Li M Z, et al. Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model. IEEE Access, 2019; 7: 7180998–181011.

[45] Fang M R, Lu J, Ruan J Y, Bian L, Wu C Y, Yao Q. Tea buds detection model using improved YOLOv4-tiny. Journal of Tea Science, 2022; 42(4): 549–560.

[46] Gui Z Y, Chen J N, Li Y, Chen Z W, Wu C Y. A lightweight tea bud detection model based on Yolov5. Computers and Electronics in Agriculture, 2023; 205: 107636.

[47] He K, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. Conference on Computer Vision and Pattern Recognition, Las Vegas: 2016; pp.770–778. DOI: 10.1109/CVPR.2016.90

[48] Mao T Y, Zhu J J, Tie J. Research based on recognition method for tea buds based on anchor free detection network. Journal of South-Central Minzu University (Natural Science Edition), 2023; 42(4): 489–496.

[49] Zhang Z, Lu Y Z, Zhao Y Q, Pan Q M, Jin K, Xu G, et al. TS-YOLO: An all-day and lightweight tea canopy shoots detection model. Agronomy, 2023; 13(5): 1411.

[50] Wang G J, Wang Z, Zhao Y, Zhang Y Z. Tea bud recognition based on machine learning. Proceedings of the 41st Chinese Control Conference, Hefei: IEEE, 2022; pp.6533–6537. DOI: 10.23919/ccc55666.2022.9902610

[51] Liu Z, Li J G, Shen Z Q, Huang G, Yan S M, Zhang C S. Learning efficient convolutional networks through network slimming. International Conference on Computer Vision, Venice: 2017; pp.2736–2744. DOI: 10.1109/ICCV.2017.298

[52] He Y H, Zhang X Y, Sun J. Channel pruning for accelerating very deep neural networks. International Conference on Computer Vision, Venice: 2017; pp.1389–1397. DOI: 10.1109/iccv.2017.155

[53] Han K, Wang Y H, Tian Q, Guo J Y, Xu C J, Chang X. GhostNet: More features from cheap operations. Conference on Computer Vision and Pattern Recognition, 2020; pp.1577–1586. DOI: 10.1109/cvpr42600.2020.00165

[54] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: Inverted residuals and linear bottlenecks. Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018; pp.4510–4520. DOI: 10.1109/CVPR.2018.00474

[55] Zhang X Y, Zhou X Y, Ma M X, Sun J. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018; pp.6848–6856. DOI: 10.48550/arXiv.1707.01083

[56] Lin G C, Xiong J T, Zhao R M, Li X M, Hu H N, Zhu L X, et al. Efficient detection and picking sequence planning of tea buds in a high-density canopy. Computers and Electronics in Agriculture, 2023; 213: 108213.

[57] 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.

[58] Huang J C, Tang A, Chen G M, Zhang D, Gao F Z, Chen T. Mobile recognition solution of tea buds based on compact-YOLO v4 algorithm. Transactions of the CSAM, 2023; 54(03): 282–290.

[59] Xie S, Sun H W. Tea-YOLOv8s: A tea bud detection model based on deep learning and computer vision. Sensors, 2023; 23(14): 6576.

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

[61] Yang M D, Yuan W H, Xu G J. YOLOX target detection model can identify and classify several types of tea buds with similar characteristics. Scientific Reports, 2024; 14: 5548.

[62] Yu T J, Chen J N, Chen Z W, Li Y T, Tong J H, Du X Q. DMT: A model detecting multispecies of tea buds in multi-seasons. Int J Agric & Biol Eng, 2024; 17(1): 199–208.

[63] Zhao R M, Liao C, Yu T J, Chen J N, Li Y T, Lin G C, et al. IMVTS: A detection model for multi-varieties of famous tea sprouts based on deep learning. Horticulturae, 2023; 819(9): 1–15.

[64] Wu Y X, Chen J N, Wu S K, Li H, He L Y, Zhao R M, et al. An improved YOLOv7 network using RGB-D multi-modal feature fusion for tea shoots detection. Computers and Electronics in Agriculture, 2024; 216: 108541.

[65] Wu Y X, Chen J N, He L Y, Gui J S, Jia J M. An RGB-D object detection model with high-generalization ability applied to tea harvesting robot for outdoor cross-variety tea shoots detection. Journal of Field Robotics, 2024; 41(4): 1167–1186.

[66] Wang M J, Li Y, Meng H W, Chen Z W, Gui Z Y, Li Y P, et al. Small target tea bud detection based on improved YOLOv5 in complex background. Front. Plant Sci, 2024; 15: 1393138.

[67] Zhang Z, Lu Y Z, Yang M Y, Wang G Q, Zhao Y Q, Hu Y G. Optimal training strategy for high-performance detection model of multi-cultivar tea shoots based on deep learning methods. Scientia Horticulturae, 2024; 328: 112949.

[68] Liu W H, Pan Z Y, Liu W J, Shao Q Q, Hu J, Wang W M, et al. Deep learning for picking point detection in dense cluster. Asian Control Conference, Gold Coast: 2017; pp.1644–1649. DOI: 10.1109/ascc.2017.8287420

[69] Bulanon D M, Kataoka T. A Fruit detection system and an end effector for robotic harvesting of Fuji apples. Agricultural Engineering International: CIGR Journal, 2010; 12(1): 203–210.

[70] Han K S, Kim S C, Lee Y B, Kim S C, Hwang H. Strawberry harvesting robot for bench-type cultivation. Journal of Biosystems Engineering, 2012; 37(1): 65–74.

[71] Wang M, Li J P, Yu Q C, Ji D M, Zhu S M. Positioning method of axillary bud removal point for cherry tomato. Transactions of the CSAM, 2016; 47(9): 23–28.

[72] Li G L, Ji C Y, Gu B X, Xu W R, Dong M. Kinematics analysis and experiment of apple harvesting robot manipulator with multiple end-effectors. Transactions of the CSAM, 2016; 47(12): 14–21, 29.

[73] Barth R, Hemming J, Van Henten E J. Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation. Biosystems Engineering, 2016; 146: 71–84.

[74] Lin G C, Tang Y C, Zou X J, Li J H, Xiong J T. In-field citrus detection

and localization based on RGB-D image analysis. Biosystems Engineering, 2019; 186: 34–44.

[75] Li J, Tang L. Crop recognition under weedy conditions based on 3D imaging for robotic weed control. Journal of Field Robotics, 2017; 35(4): 596–611.

[76] Fu L S, Gao F, Wu J Z, Li R, Karkee M, Zhang Q. Application of consumer RGB-D cameras for fruit detection and localization in field: A critical review. Computers and Electronics in Agriculture, 2020; 177: 105687.

[77] Rao U. Design of automatic cotton-picking robot with machine vision using image processing algorithms. International Conference on Control, Automation, Robotics and Embedded Systems, Jabalpur: 2013; pp.1–5. DOI: 10.1109/care.2013.6733700

[78] Geng J. Structured-light 3D surface imaging: a tutorial. Advances in Optics and Photonics, 2011; 3(2): 128–160.

[79] Pei W, Wang X L. The two-dimension coordinates extraction of tea shoots picking based on image information. Acta Agriculturae Zhejiangensis, 2016; 28(3): 522–527.

[80] Lu J Z, Yang Z M, Sun Q, Gao Z M, Ma W. A machine vision-based method for tea buds segmentation and picking point location used on a cloud platform. Agronomy, 2023; 13(6): 1537

[81] Long Z, Jiang Q, Wang J, Zhu H L, Li B, Wen F J. Research on method of tea flushes vision recognition and picking point localization. Transducer and Microsystem Technologies, 2022; 41(2): 39–41, 45. (in Chinese)

[82] Zhang L, Zou L, Wu C Y, Chen J N, Chen H P. Locating famous tea's picking point based on Shi-Tomasi algorithm. Computers, Materials & Continua, 2021; 69(1): 1109–1122.

[83] Guo S D, Yoon S C, Li L, Wang W, Zhuang H, Wei C J, et al. Recognition and Positioning of Fresh Tea Buds Using YOLOv4-lighted+ICBAM Model and RGB-D Sensing. Agriculture, 2023; 13: 518.

[84] Chen Y T, Chen S F. Localizing plucking points of tea leaves using deep convolutional neural networks. Computers and Electronics in Agriculture, 2020; 171: 105298.

[85] Shuai L Y, Mu J, Jiang X Q, Chen P, Zhang B D, Li H D, et al. An improved YOLOv5-based method for multi-species tea shoot detection and picking point location in complex backgrounds. Biosystems Engineering, 2023; 231: 117–132.

[86] Zhang F Y, Sun H W, Xie S, Dong C W, Li Y, Xu Y T, et al. A tea bud segmentation, detection and picking point localization based on the MDY7-3PTB model. Front. Plant Sci, 2023; 14: 1199473.

[87] Meng J Q, Wang Y X, Zhang J M, Tong S Y, Chen C C, Zhang C X, et al. Tea bud and picking point detection based on deep learning. Forests, 2023; 14(6): 1188.

[88] Yan C Y, Chen Z H, Li Z L, Liu R X, Li Y X, Xiao H, et al. Tea sprout picking point identification based on improved DeepLabV3+. Agriculture, 2022; 12(10): 1594.

[89] Yan L J, Wu K H, Lin J, Xu X G, Zhang J C, Zhao X H, et al. Identification and picking point positioning of tender tea shoots based on MR3P-TS model. Frontiers in Plant Science, 2022; 13: 962391.

[90] Wu W C, Hu Y G, Lu Y Z. Parametric surface modelling for tea leaf point cloud based on non-uniform rational basis spline technique. Sensors, 2021; 21(4): 1304–1304.

[91] 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.

[92] Li Y T, He L Y, Jia J M, Lu J, Chen J N, Qiao X, et al. In-field tea shoot detection and 3D localization using an RGB-D camera. Computers and Electronics in Agriculture, 2021; 185: 106149.

[93] Zhu L X, Zhang Z H, Lin G C, Chen P L, Li X M, Zhang S A. Detection and localization of tea bud based on improved YOLOv5s and 3D point cloud processing. Agronomy, 2023; 13(9): 2412

[94] Chen Z W, Chen J N, Li Y, Gui Z Y, Yu T J. Tea bud detection and 3D pose estimation in the field with a depth camera based on improved YOLOv5 and the optimal pose-vertices search method. Agriculture, 2023; 13: 1405.

[95] 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.

[96] Li R J, Qin W B, He Y T, Li Y D, Ji R B, Wu Y H, et al. Method for the classification of tea diseases via weighted sampling and hierarchical classification learning. Int J Agric & Biol Eng, 2024; 17(3): 211–221.

[97] Cao Q, Zhao C J, Xu Z, Jiang P, Yang H B, Meng X Y, et al. Review of the application of in-situ sensing techniques to address the tea growth characteristics from leaf to field. Int J Agric & Biol Eng, 2024; 17(1): 1–11.