processes

MDPI

Article

Identification and Posture Evaluation of Effective Tea Buds Based on Improved YOLOv8n

Pan Wang [1,2], Tingting He [1,2], Luxin Xie [3,*], Wenyu Yi [1,2,*], Lei Zhao [1,2], Chunxia Wang [1,2], Jiani Wang [1,2], Zhiye Bai [1,2] and Song Mei [4,*]

[1] Sichuan Academy of Agricultural Machinery Sciences, 5 Niusha Road, Chengdu 610066, China; scnjy_wp@163.com (P.W.); njyhett@scsaas.cn (T.H.); zhaolei@scsaas.cn (L.Z.); wangchxgo@163.com (C.W.); scnjy_wjn@163.com (J.W.); baizhiye_wangshen@163.com (Z.B.)
[2] Key Laboratory of Agricultural Equipment Technology for Hilly and Mountainous Areas, Ministry of Agriculture and Rural Affairs, 5 Niusha Road, Chengdu 610066, China
[3] Chongqing Key Laboratory of Intelligent Agricultural Equipment for Hilly and Mountainous Areas, College of Engineering and Technology, Southwest University, 2 Tiansheng Road, Chongqing 400715, China
[4] Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China
* Correspondence: xieluxin@swu.edu.cn (L.X.); ywycd@163.com (W.Y.); meisong@caas.cn (S.M.)

Abstract

Aiming at the low qualification rate and high damage caused by the lack of identification, localization, and posture estimation of tea buds in the mechanical harvesting process of famous tea, a framework of lightweight detection + PCA-skeleton fusion posture estimation was proposed. Based on the YOLOv8n model, the StarNet backbone network was introduced to enable lightweight detection, and the ASF-YOLO multi-scale attention module was embedded to improve the feature fusion ability. Based on the detection frame, the GrabCut-Watershed fusion segmentation was employed to obtain the bud mask. Combined with PCA and skeleton extraction algorithms, the main direction deviations of bent buds and clasped leaves were solved by Bézier curve fitting, and the morphology–posture dual-factor scoring model was thereby constructed to realize the picking ranking. Compared with the original YOLOv8n model, the results showed that the detection accuracy and mAP50 of the Improved model decreased to 85.6% and 90.5%, respectively, and the recall rate increased to 81.7%. Meanwhile, the calculation load of the improved model was reduced by 23.6%, reaching 6.8 GFLOPs, indicating a significant improvement in lightweight. The morphology–posture dual-factor scoring model achieved a score of 0.88 for a single bud in vertical direction ($\theta \approx 90^{\circ}$), a score of approximately 0.66–0.71 for buds with partially unfolded leaves and slightly bent buds, and a score of 0.48–0.53 for severely bent and overlapped buds. The results of this study have the potential to guide the picking robotic arms to preferentially pick tea buds with high adaptability and provide a reliable visual solution for low-loss and high-efficiency mechanized harvesting of famous tea in complex tea gardens.

Keywords: YOLOv8n; principal component analysis; tea bud identification; posture evaluation

check for updates

Academic Editor: Massimo Caruso

Received: 25 September 2025
Revised: 31 October 2025
Accepted: 5 November 2025
Published: 11 November 2025

Citation: Wang, P.; He, T.; Xie, L.; Yi, W.; Zhao, L.; Wang, C.; Wang, J.; Bai, Z.; Mei, S. Identification and Posture Evaluation of Effective Tea Buds Based on Improved YOLOv8n. *Processes* **2025**, *13*, 3658. https://doi.org/10.3390/pr13113658

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Tea (*Camellia sinensis* (L.) O. Kuntze) is one of the most popular beverage crops in the world, and China, the birthplace of over a thousand varieties of tea, is acknowledged as one of the origins of tea culture. In 2023, the global tea plantation area was approximately

4.83 million ha, with China occupying a tea plantation area of 2.96 million ha and an annual output of 3.55 million tons, ranking first in the world [1]. Research related to the tea industry has always been the focus of attention, especially regarding the mechanized harvesting technology of tea leaves [2].

Traditional tea harvesting relies on manual picking, which accounts for about 60% of the total labor costs of the entire tea production chain, making it a time-consuming and labor-intensive industry [3]. The rapid urban modernization contributes to the steady migration of rural labor to cities, making it harder to recruit tea pickers and driving up the cost of tea harvesting. Problems such as delayed or incorrect harvesting can lengthen the subsequent tea sorting procedures and reduce the tea quality simultaneously, causing economic losses. As a result, conventional manual harvesting is becoming a serious constraint for the expanding tea industry. It is now widely acknowledged that mechanization can be one of the most effective solutions to the increasing labor shortage in the tea industry. Research on the mechanical harvesting technique of tea leaves has been conducted all over the world to achieve efficient and high-quality tea harvesting [4–6]. It is well known that the key premise for improving the tea quality of mechanized harvesting is to realize the accurate identification and posture evaluation of tea buds.

The identification, detection, and localization of the tea buds are critical to achieving precise and automatic harvesting [7,8]. As a result, research into intelligent identification and detection of tea buds has both theoretical and practical significance [9–11]. Traditional tea bud identification relied on image processing methods, which used color transformations such as Otsu [12,13] and Watershed algorithms [14] to enhance the contrast between tea buds and background. However, image processing methods were significantly affected by lighting and background (such as old leaves and weeds with a similar color to tea buds), and the identification accuracy was limited [15]. With the rapid improvement and application of machine vision and artificial intelligence technology in recent years, research on tea bud detection using machine vision and deep learning has been extensively studied for its capability to improve target detecting ability in fields [16–18]. Deep learning automatically extracts local image features (such as edges, corners, and textures) through convolution operations, followed by downsampling to compress data volume and improve model robustness while retaining core features. Finally, it uses fully connected layers to map the learned features to the label space, achieving accurate end-to-end recognition and regression. The YOLO series detection algorithms and DeepLabV3 series segmentation algorithms have become the focus of attention. Researchers usually work based on PyCharm environment, using core frameworks such as PyTorch, TensorFlow, and Keras. Model optimization generally proceeds along two main directions: (1) reducing the complexity of the model by replacing the lightweight backbone network (such as GhostNet [19], MobileNetV2 [20], and VanillaNet [21]) and improving the model neck network [22]. However, these methods resulted in a decrease in detection accuracy. (2) improving the accuracy of feature extraction for targets of different sizes by embedding an attention mechanism in networks [23,24] and optimizing feature extraction networks [25,26]. In terms of picking point location, recent studies mostly focus on key point labeling (such as the lowest point of buds [27,28]), direct semantic segmentation of tea stems [29], and determining the position of 4 mm below the tea buds as the picking point based on the manual picking experience [30]. Aiming at the challenges that depth information cannot be obtained from 2D images, and it is difficult to locate picking points under complex lighting and dense canopies, Li et al. [31] integrated RGB and depth images to generate a 3D point cloud of the tea bud area. The tea bud point cloud was extracted by using Euclidean clustering, and the minimum enclosing cylinder was calculated. Combined with the growth characteristics of tea buds, the bottom of the cylinder was finally determined as the picking

point. Zhang [32] used the principal component analysis method (PCA) to fit the minimum circumscribed cuboid of the 3D point cloud of tea buds; the center of mass and eigenvector of the cuboid were calculated, and the center point coordinates of the bottom of the cuboid were determined as the picking point.

However, the efficiency and accuracy of mechanical picking depend on the accurate identification and posture evaluation of tea buds, which are not fully addressed in existing studies. Furthermore, given the limited computing power of field equipment, there is an urgent need for lightweight models. Therefore, this study focused on the morphology of a single bud and one bud with one leaf, and carried out two main research objectives: (1) High-precision and lightweight identification of effective buds based on an improved YOLOv8n model. A lightweight YOLOv8n–StarNet–ASF detection model was developed to reduce the computational load and improve the feature fusion ability. (2) A coordinate + posture dual guidance mechanism was constructed through the integration of PCA and a skeleton extraction algorithm. Thus, the picking consistency and the quality of harvested tea could be improved. The schematic diagram of the research workflow is shown in Figure 1.

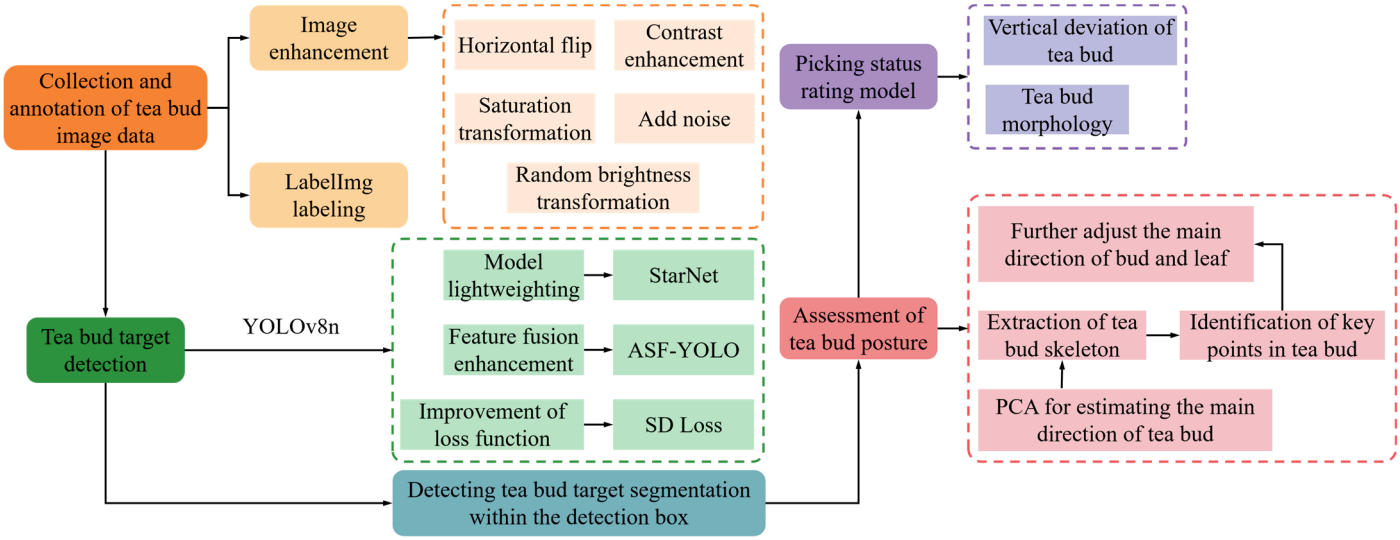


Figure 1. Schematic diagram of this study.

2. Dataset Construction

2.1. Data Acquisition

To ensure the authenticity and diversity of the data, field image collection was conducted from mid-to-late March to early April 2025 (the critical harvesting period of famous tea) at three representative locations: Wenjun Tea Park in Chengdu City, Qingchengdao Tea Sightseeing Garden in Dujiangyan City, and Qianliyun Export Tea Base in Leshan City. The images were captured using a Huawei Mate70 mobile phone (resolution 4032 × 3024) (Huawei, Shenzhen, China) and a Nikon D7200 camera (resolution 6000 × 4000) (Nikon, Tokyo, Japan), covering cloudy, sunny, rainy, and other natural environments. The shooting targets included a single bud, one bud with one leaf (Figure 2), and the shooting angles ranged from 0° to 90° with vertical direction, and the height ranged from 10 to 25 cm relative to the surface of the tea canopy. In order to improve the accuracy of the training model, the collected images were screened to remove unclear images with blurry quality and obvious reflection on the tea bud, and finally, more than 1200 original images without preprocessing were retained.

(**a**) Close shot (**b**) Distant shot (**c**) High brightness (**d**) Low brightness

Figure 2. Tea buds under different conditions.

2.2. Dataset Annotation

Based on the picking standard and growth characteristics of famous tea, the tea buds that meet the following characteristics were defined as effective buds: (1) The shape of buds and leaves was clear, and the contour was clearly recognizable. (2) For a single bud, the stem between the bud and the first leaf was clearly visible. For one bud with one leaf, the stem between the first leaf and the second leaf was clearly visible when the first leaf was not fully expanded, as shown in Figure 3.

(**a**) Single bud (**b**) One bud with one leaf

Figure 3. Morphology of effective buds.

The dataset was expanded using a series of image enhancement techniques, including horizontal flipping, random brightness adjustment, contrast enhancement (±20%), saturation transformation, and the addition of both Gaussian noise (with a standard deviation of 0.1) and salt-and-pepper noise. The random geometric transformation was used to simulate the shooting effect of different angles, and the noise was added to simulate the effect of camera instability. There were 5142 images in the dataset after image enhancement, which were divided into training, validation, and test sets in a ratio of 8:1:1. The LabelImg annotation tool was then used to label the effective buds, and the annotation files were saved as .txt files.

3. Identification and Posture Evaluation of Effective Buds

3.1. Identification of Effective Buds Based on Improved YOLOv8n

As a one-stage object detection algorithm, YOLOv8n has been systematically optimized in terms of model architecture and training mechanism. In the feature extraction section, the model adopts the C2f module as the backbone and core component of the neck. This structure draws on the ELAN idea and effectively enhances the fluidity of gradient feedback by splitting the feature map—with some parts deepening the features through the Bottleneck module, and some directly participating in fusion—and combining multi-level feature concatenation, thereby alleviating the problem of gradient vanishing and improving the feature representation ability of the model. Further introduction of the PAN–FPN structure into the neck enhances the model's adaptability to targets of different sizes through bidirectional cross-scale feature fusion, especially improving the detection performance of small targets in complex scenes.

In the design of the detection head, YOLOv8n adopts a free box and decoupling head mechanism, separating the classification task and bounding box regression task into two independent branches, enabling the model to focus on semantic recognition and position regression, respectively, significantly improving training efficiency and inference accuracy. In terms of the loss function, the model introduces DFL (Distribution Focal Loss) to model the bounding box position as a discrete probability distribution and uses the Focal Loss mechanism to enhance the learning of key positions, improving the localization robustness of the model in complex situations such as target occlusion or edge blur. In addition, YOLOv8n adopts a Task Signed Assign as the sample allocation strategy, which combines classification confidence and IoU localization quality to dynamically select high-quality positive samples, ensuring consistency between classification and regression objectives, and further improving the convergence speed of the training process and the overall performance of the final model. The network structure is shown in Figure 4.

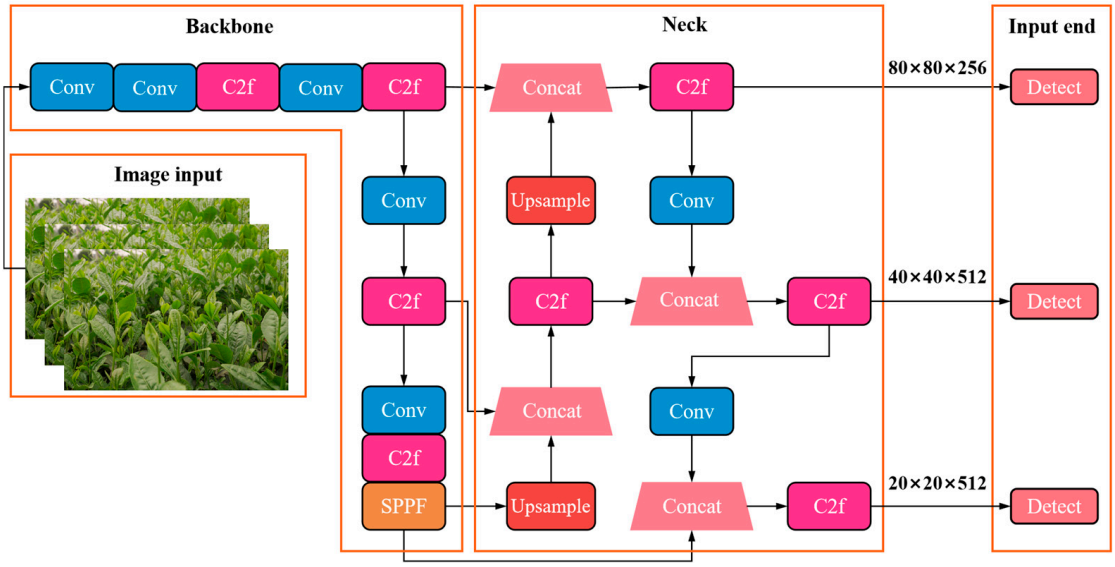


Figure 4. Network structure of YOLOv8n.

3.1.1. Model Lightweight

To address the computational limitation of field equipment, StarNet was adopted to replace the YOLOv8n backbone network. Although YOLOv8 backbone introduced lightweight C2f modules, its overall structure still relied heavily on standard convolutions, feature concatenation, and complex cross-layer connections, resulting in high FLOPs and parameter count, making it difficult to operate efficiently in low computing scenarios. In

contrast, based on the principle of star operation, StarNet generates high-dimensional features in a low-dimensional space calculation and performs high-dimensional nonlinear mapping without increasing network width. StarNet adopts a four-level hierarchical architecture, extracting features via a convolutional layer downsampling and an improved module, and adds depthwise separable convolution at the end of each block. The activation function is replaced with ReLU6 to balance efficiency and nonlinear feature extraction. The network width is doubled step by step with a fixed channel expansion factor of 4, and the layer normalization is replaced by batch normalization (Figure 5). The lightweight of StarNet significantly reduces the computational load while retaining the ability to detect small targets.

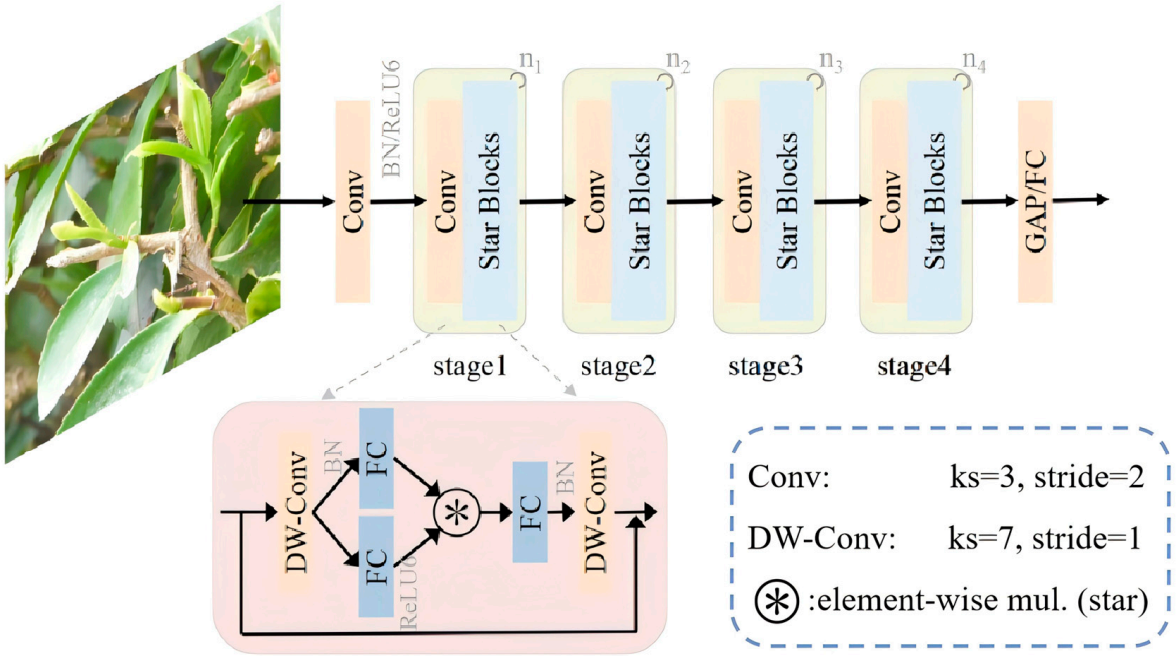
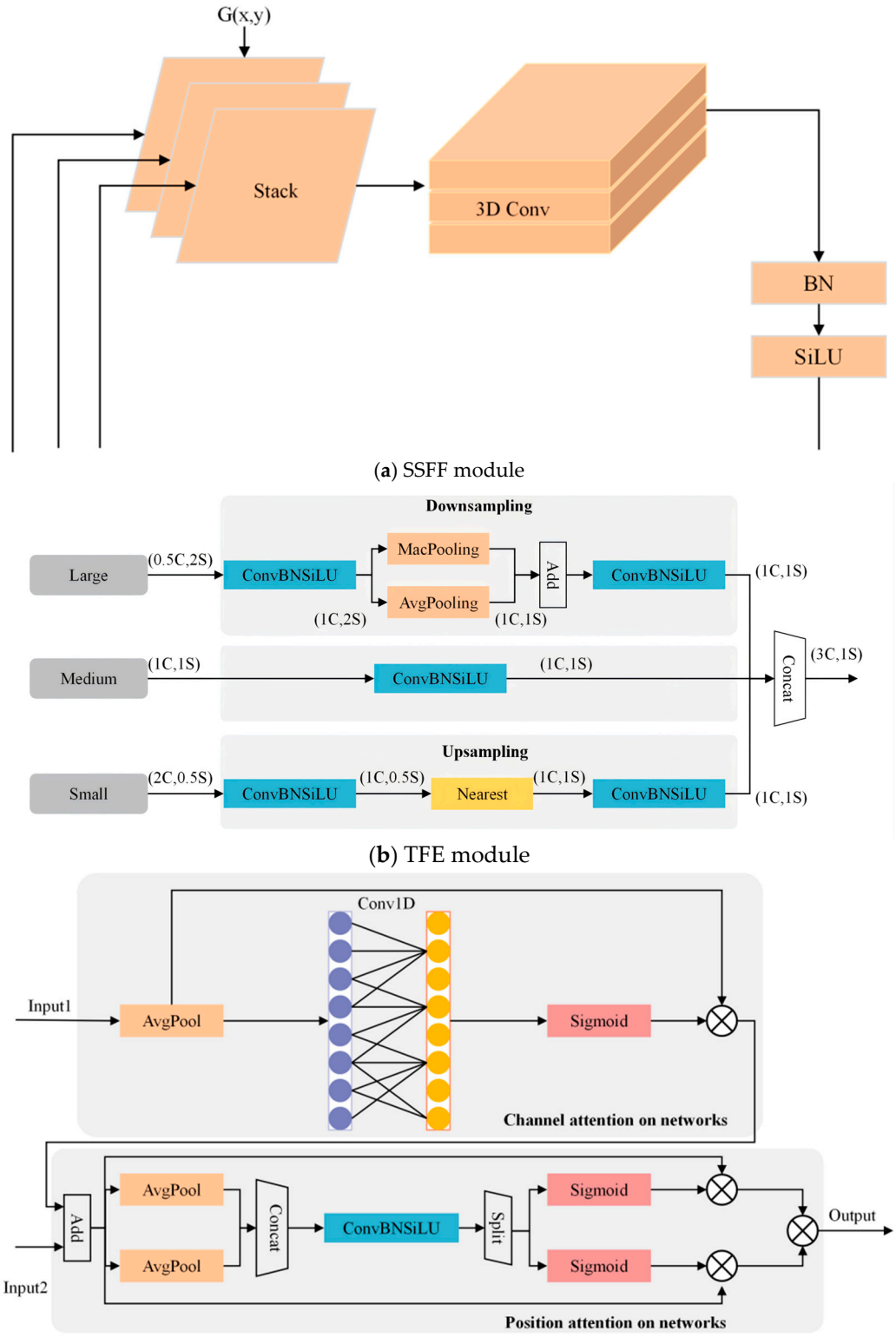


Figure 5. Network structure of StarNet.

3.1.2. Feature Fusion for Target Detection of Dense Tea Buds

Tea buds grow densely and are small in size. Therefore, the original PAN–FPN structure of YOLOv8n has insufficient feature fusion in high-density and multi-overlapping scenes. Therefore, the ASF-YOLO module is introduced in the neck, which includes SSFF, TFE, and CPAM modules. The structure is shown in Figure 6. The SSFF module enhances robustness to targets of different sizes and directions through normalization of the multi-scale feature map, upsampling, stacking, and the combination of features via 3D convolution. The TFE module concatenates the spatial dimensions of large-, medium-, and small-scale feature maps to capture the spatial information of small targets. The CPAM module integrates the feature information of the SSFF and TFE modules and adaptively focuses on the key channels and spatial positions through channel attention and position attention networks, thereby improving the detection accuracy.



(**a**) SSFF module

(**b**) TFE module

(**c**) CPAM module

Figure 6. Network structure of ASF-YOLO.

3.1.3. Improvement of Loss Function

The original regression loss (DFL + CIOU) of YOLOv8n ignores the influence of target size to position sensitivity, which leads to missed or false detection of small targets. In this study, the regression loss function of the original model was replaced with SD Loss which can be dynamically adjusted according to the target size, thereby improving the model stability and detection accuracy. SD Loss dynamically adjusts the influence coefficients

of Sloss and Lloss based on the target scale. For BBox labels, the influence coefficient is determined by β_B:

$$\beta_B = min(\frac{B_{gt}}{B_{gtmax}} \times R_{OC} \times \delta, \delta) \quad (1)$$

where B_{gt} is the area of the current target box, B_{gtmar} = 81, R_{OC} is the size ratio of the original image and the current feature map, and δ is an adjustable parameter. Consequently, $\beta_{\mathcal{L}_{BS}} = 1 - \delta + \beta_B$ and $\beta_{\mathcal{L}_{BL}} = 1 + \delta - \beta_B$, and the final scale-based dynamic loss (SDB loss) of Bbox is given by:

$$\mathcal{L}_{SDB} = \beta_{\mathcal{L}_{BS}} \times \mathcal{L}_{BS} + \beta_{\mathcal{L}_{BL}} \times \mathcal{L}_{BL} \quad (2)$$

where $\mathcal{L}_{BS}$ and $\mathcal{L}_{BL}$ are the scale loss and location loss functions of Bbox, respectively. $\beta_{\mathcal{L}_{BS}}$ and $\beta_{\mathcal{L}_{BL}}$ are the influencing factors of $\mathcal{L}_{BS}$ and $\mathcal{L}_{BL}$, respectively. When the area of the target Bbox is greater than 81, $\mathcal{L}_{SDB}$ degenerates into CioU loss.

For mask labels, the influence coefficient is determined by β_M:

$$\beta_M = min(\frac{M_{gt}}{M_{gtmax}} \times R_{OC} \times \delta, \delta) \quad (3)$$

The scale-based dynamic loss (SDM loss) of the mask is:

$$\mathcal{L}_{SDM} = \beta_{\mathcal{L}_{MS}} \times \mathcal{L}_{MS} + \beta_{\mathcal{L}_{ML}} \times \mathcal{L}_{ML} \quad (4)$$

where $\beta_{\mathcal{L}_{MS}} = 1 + \beta_M$ and $\beta_{\mathcal{L}_{ML}} = 1 - \beta_M$, $\mathcal{L}_{MS}$ and $\mathcal{L}_{ML}$ are the scale loss and location loss functions of the mask, respectively. $\beta_{\mathcal{L}_{MS}}$ and $\beta_{\mathcal{L}_{ML}}$ are the influencing factors of $\mathcal{L}_{MS}$ and $\mathcal{L}_{ML}$, respectively.

In this way, the attention weight of the smaller target on Sloss in Bbox label is reduced, while the influence of Sloss is enhanced in mask labels, thereby reducing the influence of inaccurate labels on the stability of the loss function.

3.2. Evaluation of Tea Bud Position and Posture

It is necessary to revise the deviation in growth direction caused by bent buds and clasped leaves (leaves that wrap around the stem) after the detection frame is located, so as to provide spatial posture guidance for the picking robotic arm. Previously, the skeleton extraction algorithm was often used to detect the inflection points, bending points, and direction points of tea buds by combining it with edge detection [33] and the Shi–Tomasi [34] algorithms, and determining picking points by combining it with morphological features. However, in natural scenes such as uneven lighting, overlapping, and complex backgrounds, the skeleton extraction algorithm is prone to causing issues such as fractures and burrs, poor adaptability to changes in bud posture, and unstable intersection extraction. Furthermore, for the extraction of 3D spatial information of tea buds, PCA may deviate from the actual growth direction when the buds are bent, overlapped, and incomplete, resulting in deviations in the picking angle [35]. Skeleton extraction struggles in complex scenes, and PCA also has limitations in estimating the actual growth direction of bent or overlapping buds. Thus, integrating the two methods is necessary to achieve more accurate posture evaluation. This study focused on improving the detection area of effective buds in the YOLOv8n model and obtaining posture features by combining PCA and skeleton extraction, thereby providing dual guidance of coordinate + posture for picking robotic arms.

3.2.1. Segmentation of Tea Buds

Based on the bounding boxes of effective buds detected by YOLOv8n, the GrabCut segmentation algorithm was used to separate the buds from the background. The small holes were filled, and the contours were smoothed by using a morphological closing operation. The connected domain analysis was utilized to eliminate noise areas, and the largest connected domain was then retained as the target.

3.2.2. Posture Estimation of Tea Buds

The posture of tea buds was quantified based on the segmentation mask. In addition, a staged posture estimation framework was proposed, and the main directions of tea buds were obtained based on PCA. The detailed posture correction was carried out through skeleton analysis, and the problem of direction deviation in scenes with bent buds and clasped leaves was solved.

The covariance matrix C for calculating the bud mask via PCA is represented as:

$$C = \frac{1}{N}\sum_{i=1}^{N}(p_i - \mu)(p_i - \mu)^T \tag{5}$$

where p_i is the pixel coordinate of the mask, $p_i = (x_i, y_i)$, μ *is* the center of mass, and N is the total number of pixels of the target. The eigenvector v corresponding to the maximum eigenvalue of C can be expressed as $v = [v_x, v_y]^{\mathrm{T}}$; thus, the direction angle θ of the main axis is defined as:

$$\theta_{\mathrm{PCA}} = \arctan2(v_y, v_x)\cdot\frac{180}{\pi} \tag{6}$$

where $\theta \in [0^\circ, 180^\circ]$, θ is the angle between the stem and the horizontal direction.

When the main direction estimated by PCA is disturbed by bent buds and clasped leaves ($\Delta\theta > 10^\circ$, a threshold determined via pre-experiments: deviations exceeding 10° would cause the picking robotic arm to clamp incorrectly, increasing tea bud damage), it is corrected using skeleton analysis. The stem path θ_{stem} is traced when a branch point exists, and the bending direction θ_{curve} is fitted by a Bézier curve when there is no branch point. The initial skeleton of the effective buds is extracted using the morphological thinning algorithm, and the eight-neighbor connectivity number $C(p)$ of the skeleton point is calculated to realize the identification of the key points of tea buds:

$$C(P) = \sum_{k=1}^{8}|N_k - N_{k+1}|(N_k \in \{0,1\}) \tag{7}$$

$C(p) = 1$ indicates that the bud endpoint (E_{i}) has been identified, $C(p) \geq 3$ indicates that the branch point (P_{j}) has been identified, and there is a clasped leaf. Three structures, including bud tip E_1, leaf tip E_2, and stem tip E_3, are identified through topological analysis, and the bifurcation point P_m is set as the branch point closest to the effective bud. Then the stem path ($\vec{d}_{stem}$ and θ_{stem}) is traced and generated along the skeleton from P_m to E_3.

$$\vec{d}_{stem} = \frac{1}{K-1}\sum_{k=1}^{K-1}(s_{k+1} - s_k) \tag{8}$$

$$\theta_{stem} = \arctan2(d_y, d_x)\cdot\frac{180}{\pi} \tag{9}$$

If there is no point with $C(p) \geq 3$ in the identification results of the key points, it indicates that the tea bud is a single bud. However, the tea bud may be bent. Therefore, the curvature ρ is defined to determine whether the bud is bent:

$$\rho = \frac{L}{D} \tag{10}$$

where L is the arc length of the bud stem and D is the Euclidean distance between the two endpoints of the bud stem. $\rho \geq 1$, the bending degree of the bud increases with the increase in ρ. For the main direction extraction of the bent tea buds, the secondary Bézier fitting is performed on the bud skeleton path, and the tangential direction at the farthest point from the bud base of the Bézier curve is used as the growth direction of tea buds.

$$B(t) = (1-t)^2 P_1 + 2t(1-t)P_2 + t^2 P_3 \ (t \in \{0,1\}) \tag{11}$$

$$\theta_{curve} = \arctan2(\frac{dy}{dt}, \frac{dx}{dt}) \tag{12}$$

where $B(t)$ is the parameter equation of the Bézier curve. P_1 and P_3 are the starting and ending points of the stem, respectively. P_2 is the intermediate control point fitted by the least squares method. In addition, the angles between the initial direction θ_{PCA} and θ_{stem}, and between θ_{PCA} and θ_{curve} are denoted as $\Delta\theta$. If $\Delta\theta > 10^\circ$, it is considered that the main direction of PCA is affected by tea leaves, which needs to be corrected:

$$\theta_{\mathrm{final}} = \begin{cases} \theta_{\mathrm{stem}} & \text{branch point exists} \cap |\Delta\theta| > 10^\circ \\ \theta_{\mathrm{curve}} & \text{no branch points} \cap \ \rho \ > 1.5 \cap |\Delta\theta| > 15^\circ \\ \theta_{\mathrm{PCA}} & \text{otherwise} \end{cases} \tag{13}$$

3.2.3. Evaluation of Tea Bud Picking

The posture identification of tea buds requires a comprehensive evaluation of the bud morphology, bud posture, and the picking adaptability of picking robotic arms. In this study, an evaluation model of tea bud picking based on morphological and postural characteristics of tea bud was proposed: (1) Morphological characteristic score: The score coefficient α = 1 for a bud with straight stem and no leaf covering; the score coefficient α = 0.8 for one bud with one partially unfolded leaf and a visible stem; and the score coefficient α = 0.6 for bent and overlapped buds. (2) Deviation calculation of main growth direction: The angle between the main growth direction and the horizontal direction of the bud is defined as θ, and the deviation score coefficient β is defined by Equation (14). β = 1 when the bud is vertically upward (θ = 90°), and the picking adaptability for robotic arms is the best.

$$\beta = \begin{cases} 1 - \frac{|\theta - 90^\circ|}{90^\circ}, & |\theta - 90^\circ| \ < \ 90^\circ \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

The morphological characteristic coefficient α is set based on the agricultural picking standards, and the deviation score coefficient β is derived from the optimal angle (90°) of the vertical clamping of the picking robotic arm. The picking score is $\alpha \times \beta$, the detected buds are sorted by the picking score, and the tea buds with high scores are preferentially picked, which provides a basis for the path planning of the picking robotic arm.

4. Results and Discussion

The optimizing, training, and testing of the algorithms used in this study were carried out on a computer with an Intel (R) Xeon (R) W-2245 processor (Intel Corp., Santa Clara, CA, USA), NVIDIA Quadro RTX 4000 CPU (NVIDIA, Santa Clara, CA, USA), and 8 GB of running memory. The software environment used was Windows 10 operating system, based on Python 3.9, using PyCharm (Community Edition 2021.2) as the development environment, and relying on PyTorch (2.7.1+cu126) deep learning framework and CUDA 12.6 computing architecture for algorithm optimization.

Based on the requirement of real-time detection and lightweight performance in a field environment, the YOLOv8n model was selected for identifying effective tea buds, and precision, recall, F1-score, *mAP*50, and *mAP*50-95 were selected as the evaluation metrics.

$$P = \frac{TP}{TP + FP} \tag{15}$$

$$R = \frac{TP}{TP + FN} \tag{16}$$

$$F1 = \frac{2 \times P \times R}{P + R} \tag{17}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N}\int_{0}^{1} P(r)dr \tag{18}$$

where TP is the number of positive samples correctly identified, FP is the number of negative samples mistakenly identified as positive, and FN is the number of positive samples mistakenly identified as negative, P is the precision rate, and R is the recall rate.

AP (Average Precision) is the area of the region enclosed by the precision–recall (*P*–*R*) curve and the coordinate axes. *mAP*50 is the mean average precision at IoU (Intersection over Union) = 0.5, *mAP*50-95 is the mean average precision at IoU = 0.50:0.05:0.95.

4.1. Identification Results of Effective Buds

The sample training, verifying, and testing were conducted under the same conditions. The improved YOLOv8n algorithm was used to train the constructed tea image dataset on a GPU. The parameter settings included: an initial learning rate of 0.01, a batch size of 12, training epochs of 500, a variant of SGD with a momentum of 0.9, and a weight decay (L2 regularization) coefficient of 5×10^{-4} during updates. The input image resolution was set to 640 × 640.

4.1.1. Ablation Test

StarNet_s050, ASF-YOLO, BiFPN, and SEAM modules were gradually added to the original YOLOv8n model through ablation tests. The effectiveness of each module in improving the accuracy of tea bud identification was evaluated in combination with indicators, and the results are shown in Table 1.

Table 1. Comparison of ablation tests.

Model	StarNet_s050	ASF-YOLO	BiFPN	SEAM	Precision Rate/%	Recall Rate/%	F1-Score/%	*mAP*/%		Calculation Load/GFLOPs
								*mAP*50	*mAP*50-95	
1	×	×	×	×	87.3	80.5	83.8	91.8	67.9	8.9
2	√	×	×	×	85.0	83.2	84.1	89.9	63.8	6.5
3	√	√	×	×	85.6	81.7	83.6	90.5	65.2	6.0
4	√	×	√	×	84.9	84.4	84.65	88.9	64.6	6.5
5	√	×	×	√	83.2	82.6	82.89	89.6	64.3	6.7

The results indicate that the computational load of the original YOLOv8n model (Model 1) was 8.9 GFLOPs. Although its *mAP*50 was 91.8%, and *mAP*50-95 was only 67.9%, indicating poor adaptability to targets with different degrees of overlap, the model cannot meet the lightweight requirement for real-time detection in the field. After replacing the backbone of YOLOv8n with the StarNet_s050 network (Model 2), the calculation load decreased to 6.5 GFLOPs (a 27% reduction compared with the original YOLOv8n's 8.9 GFLOPs), which significantly improved the model's lightweight performance. The recall rate increased to 83.2%, indicating that the star operation of StarNet enhances the ability to capture small targets. However, the precision rate (85.0%), *mAP*50 (89.9%), and *mAP*50-95 (63.8%) of the model decreased, indicating that the completeness of feature extraction was

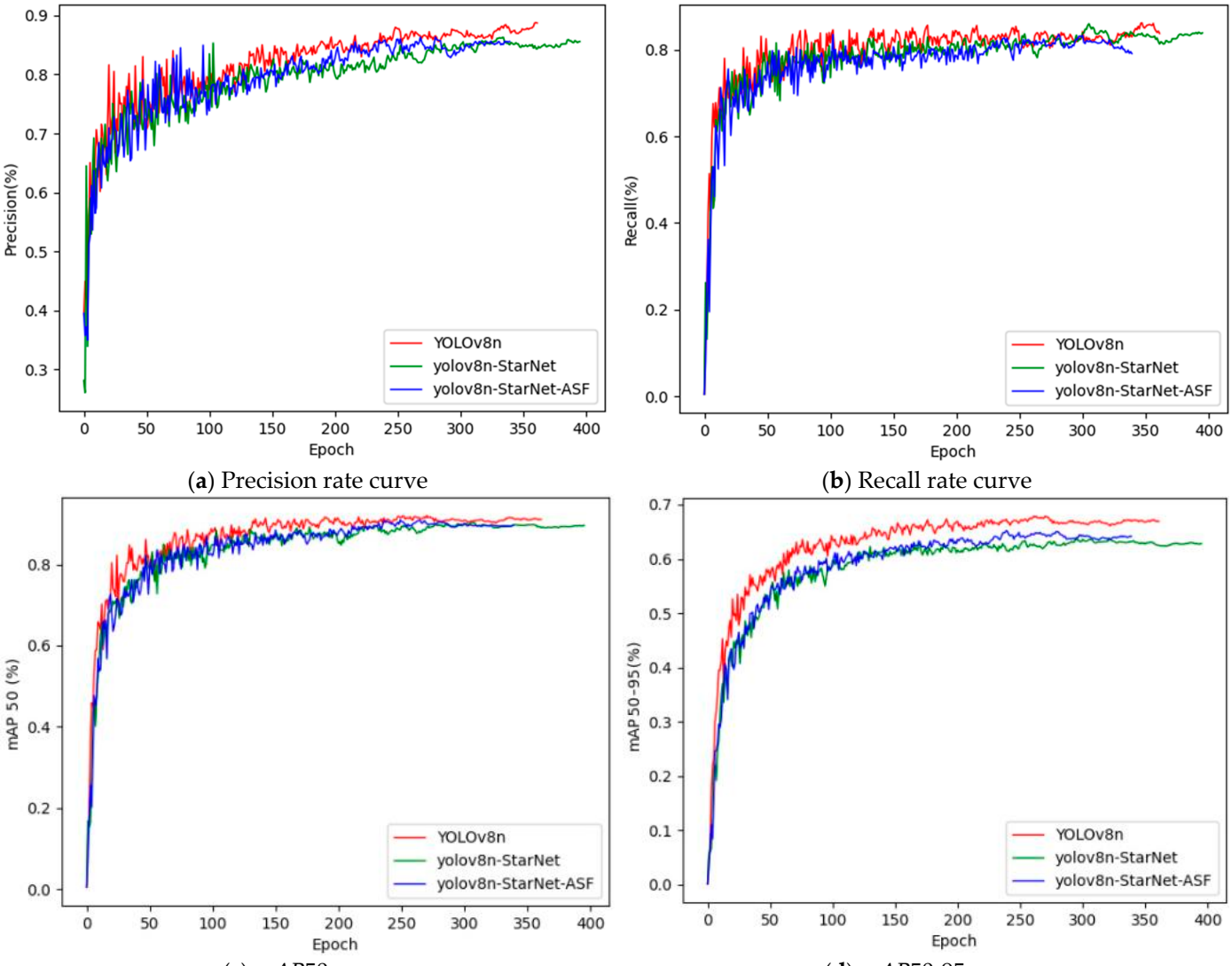
insufficient. After introducing the ASF-YOLO module on the basis of Model 2 (Model 3), the calculation load increased to 6.8 GFLOPs (lower than that of the original model). However, the precision rate, *mAP*50, and *mAP*50-95 increased to 85.6%, 90.5% and 65.2%, respectively, which indicated that the ASF-YOLO module effectively enhanced the feature fusion ability for small targets through synergy of SSFF, TFE, and CPAM, and made up for the feature lack of StarNet. Furthermore, in order to verify the influence of different feature fusion structures on the performance improvement of YOLOv8n model, on the basis of replacing the Backbone with StarNet_s050 network, BiFPN (Bidirectional Feature Pyramid Network) [36] and SEAM (Separated Enhancement Attention Module) [37] were introduced into the neck network to compare with the ASF-YOLO module, so as to select the most suitable feature fusion module for tea bud identification in dense tea fields. Although the F1-score of Model 4 with BiFPN (84.65%) was slightly higher than that of Model 3, its *mAP*50 (88.9%) and *mAP*50-95 (64.6%) were both lower. Combined with the performance of Model 5 (with SEAM), these results confirm that the ASF-YOLO module has better optimization effects on feature fusion in dense tea fields. Model 3 effectively balances the detection accuracy and adaptability to complex scenes while remaining lightweight (reducing computational load by 23.6%) and is the most suitable model for real-time identification of effective buds in the field.

4.1.2. Detection Results of Effective Buds

Figure 7 shows the dynamic changes in the key indicators of YOLOv8n, YOLOv8n–StarNet, and YOLOv8n–StarNet–ASF during training. From the precision curve (Figure 7a), it can be seen that the original YOLOv8n curve has a high starting point and significant fluctuations, indicating that the model was insufficient in suppressing false detection for dense and small targets. After employing the StarNet network, the model precision decreased to 85.0% and the curve became smoother, indicating that the stability of the model had improved while the false detection increased slightly. The precision rose to 85.6% and the fluctuation decreased after introducing ASF-YOLO, indicating that the module effectively compensates for the precision loss caused by StarNet. In addition, the original model had the lowest recall rate (80.5%), and its curve showed a weak upward trend and significantly missed detection (Figure 7b). The recall rate of the model increased to 83.2% after replacing the backbone with StarNet, indicating that the star operation enhances the activation of small targets. There was a slight decrease when ASF was introduced. However, it is still higher than the original network, and the curve is more stable, indicating that the model maintains stability while improving the detection rate. In terms of the average precision, the original YOLOv8n showed better performance on *mAP*50 and *mAP*50-95 curves (Figure 7c,d), accompanied by significant fluctuations. The *mAP*50 and *mAP*50-95 curves showed a slight decrease after introducing StarNet. However, the *mAP*50 and *mAP*50-95 curves increased to 90.5% and 65.2%, respectively, after integrating the ASF-YOLO module, and the curves converged smoothly, which verified the improvement of the ASF module on target consistency under different IoU thresholds. In conclusion, it can be seen that the YOLOv8n–StarNet–ASF model achieves the balance among lightweight, recall rate and robustness in complex scenes, and is more suitable for real-time detection of tea buds in fields.

Figure 8 shows the robustness of the improved model in four typical scenes of single target, multi-target, overexposure, and rainy conditions. The detection box of the original YOLOv8n for a single bud had a slight deviation, with an IoU of about 0.82. The IoU of YOLOv8n–StarNet was about 0.85. However, the locating of the bud tip deviated slightly due to feature simplification. The detection box of YOLOv8n–StarNet–ASF almost coincided with the real box (IoU = 0.93). When the density of tea buds increased, the original

YOLOv8n showed a higher missed identification rate, as its PAN–FPN structure was prone to feature confusion under high-density conditions–specifically, small and overlapping buds were easily omitted or mislocalized. Although the missed identification rate decreased after adding StarNet, the small tea buds were still not detected. The missed identification rate of the YOLOv8n–StarNet–ASF model for dense targets was significantly decreased. When the contrast between buds and the background decreased under strong lighting, the original YOLOv8n and YOLOv8n–StarNet had a false detection each due to the lack of color features. YOLOv8n–StarNet–ASF accurately identifies effective tea buds and reflective old leaves and reduces the false identification rate. When rainfall leads to image blurring and the surface of the tea bud reflects light, the original YOLOv8n mistakenly identifies the first leaf with a similar color as a tender bud. Although YOLOv8n–StarNet accurately distinguished effective tea buds from the first leaf, it can be seen from the heat map that the detection box of the model is deviated, and the complete tea bud could not be identified. In contrast, YOLOv8n–StarNet–ASF not only accurately identified effective buds in the image, but also significantly improved the stability of the detection box.



(**a**) Precision rate curve

(**b**) Recall rate curve

(**c**) *mAP*50 curve

(**d**) *mAP*50-95 curve

Figure 7. Comparison of key performance indicators during training.

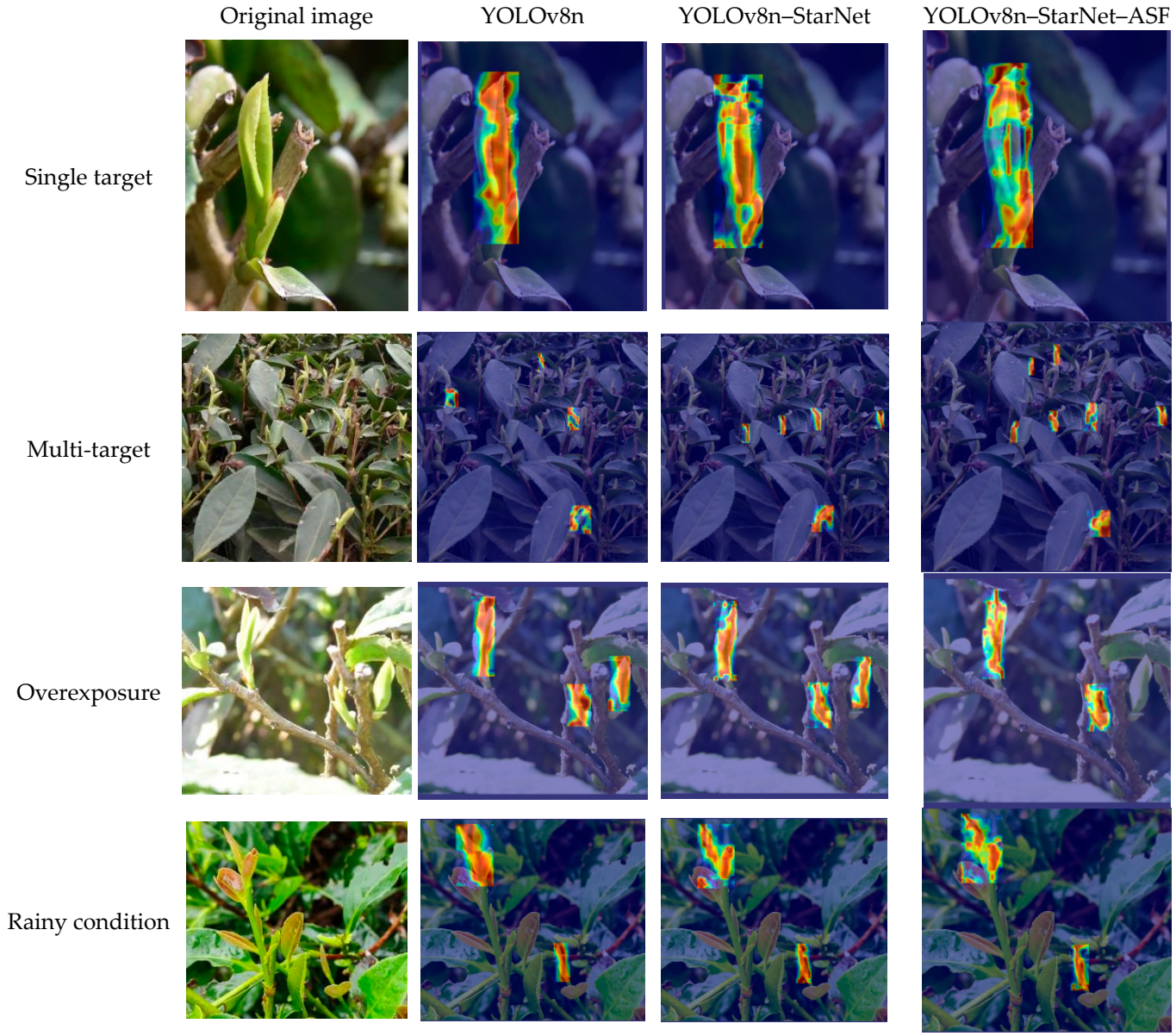
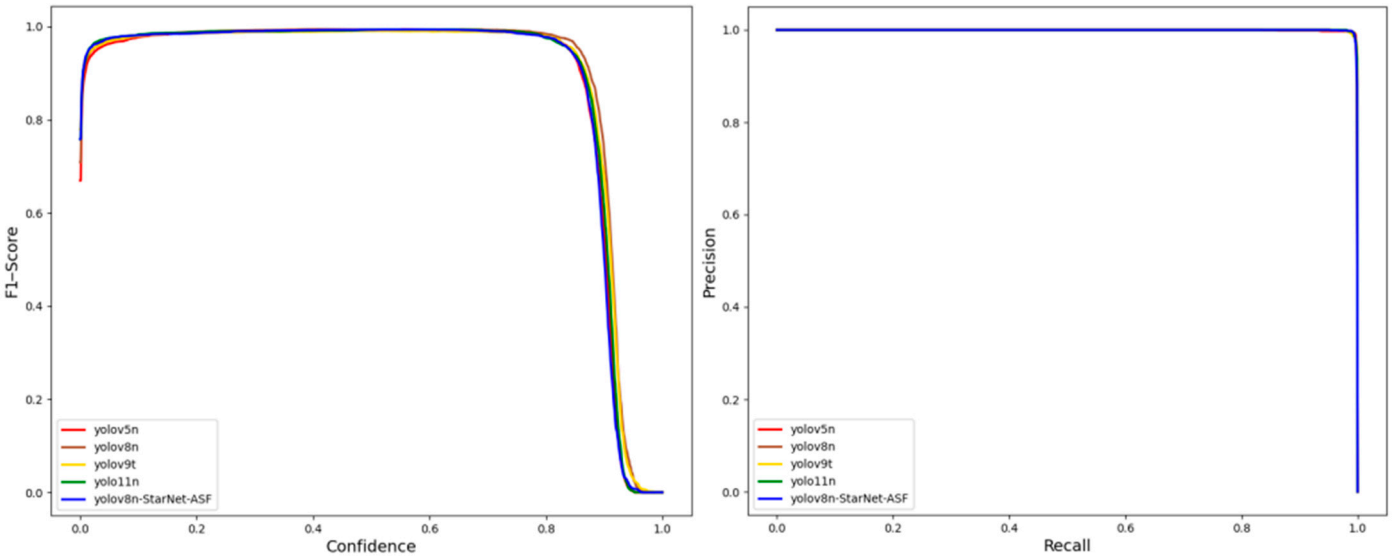


Figure 8. Heatmaps of the feature extraction layer for different models.

To verify the comprehensive performance of the improved YOLOv8n–StarNet–ASF model in tea bud detection, the mainstream lightweight YOLO series models (YOLOv5n, original YOLOv8n, YOLOv9t, and YOLO11n) were selected as the control group, and 1047 images of effective buds were used as the benchmarks. The comparison results of the F1-score curve and the P–R curve are shown in Figure 9. The F1-score comprehensively reflects the model's ability to achieve a balance between reducing false detection and minimizing missed detection. The fluctuation of the curve indicates the stability of the model's performance at different confidence thresholds. From the F1-score curve in Figure 9a, it can be seen that the original YOLOv8n model maintains the highest F1-score in the confidence range, benefiting from its mature feature fusion architecture. Although the F1-curve of the improved YOLOv8n–StarNet–ASF model is slightly lower than that of the original YOLOv8n, the overall smoothness of the curve is better, indicating that the performance stability is superior to that of the original model when dealing with the complex backgrounds in tea gardens (such as interference from old leaves and complex lighting). Compared with YOLOv5n and YOLOv9t, the F1-curve of the improved model shows better performance in

the medium-to-high confidence range (0.6–0.9), reflecting the balanced capability of the improved model between precise detection and broad scenario adaptation.



(**a**) F1-score curve (**b**) P–R curve

Figure 9. Comparison of F1-score curves and P–R curves for tea bud recognition among different YOLO models.

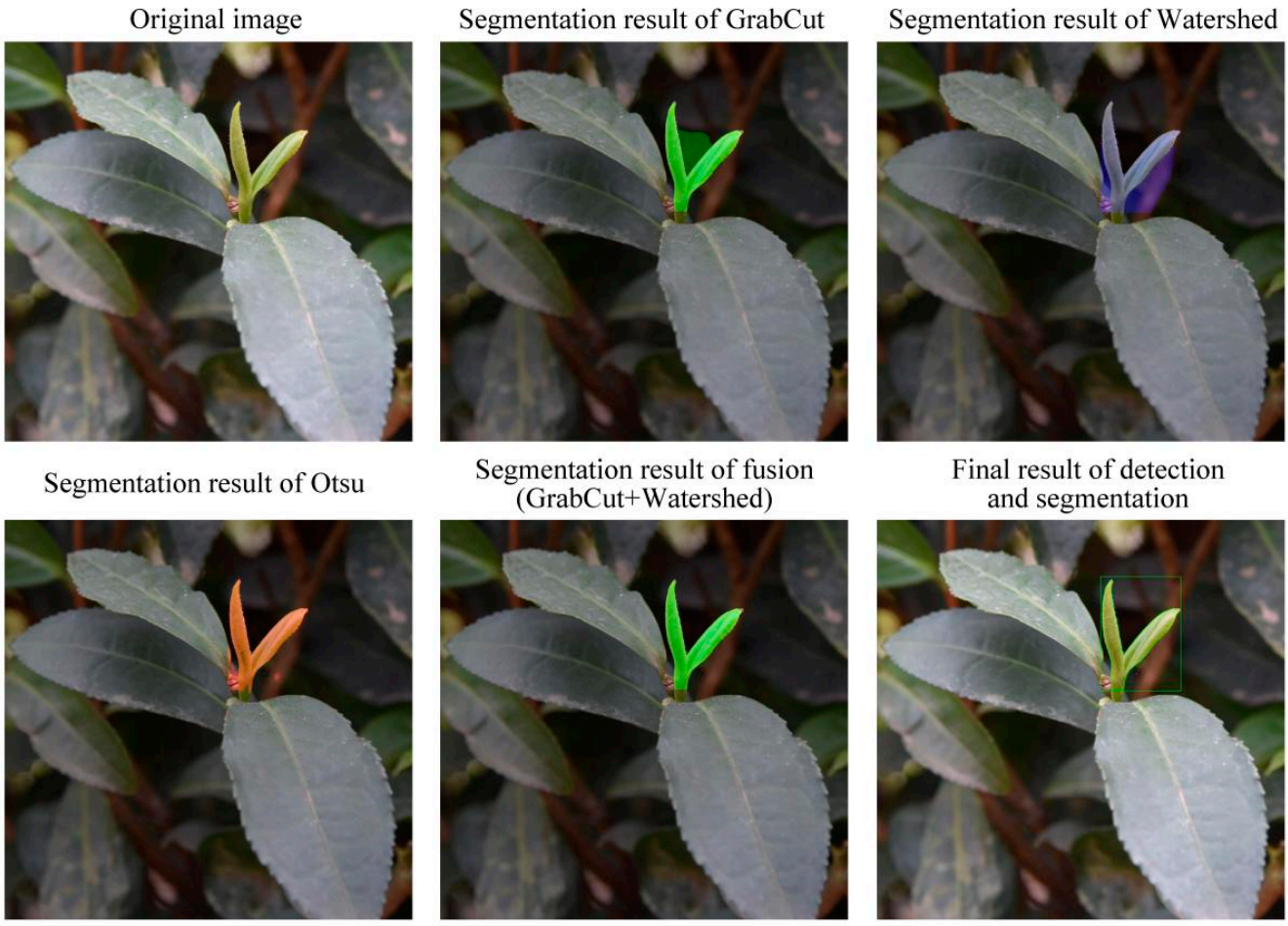
The area under the P–R curve represents the model's comprehensive detection capability for targets, and a curve closer to the upper-right corner indicates better model performance. As shown in Figure 9b, the shape of the P–R curve of the improved YOLOv8n–StarNet–ASF model is highly consistent with the lightweight models YOLOv5n, YOLO11n, YOLOv8n, and YOLOv9t. This indicates that through the low-dimensional to high-dimensional feature mapping of the StarNet backbone and the multi-scale attention fusion of ASF-YOLO, the improved model achieves accurate identification of small-sized and blurry tea buds while maintaining lightweight performance, thereby compensating for the precision loss typically seen in traditional lightweight models.

4.2. Assessment Results of Tea Bud Posture

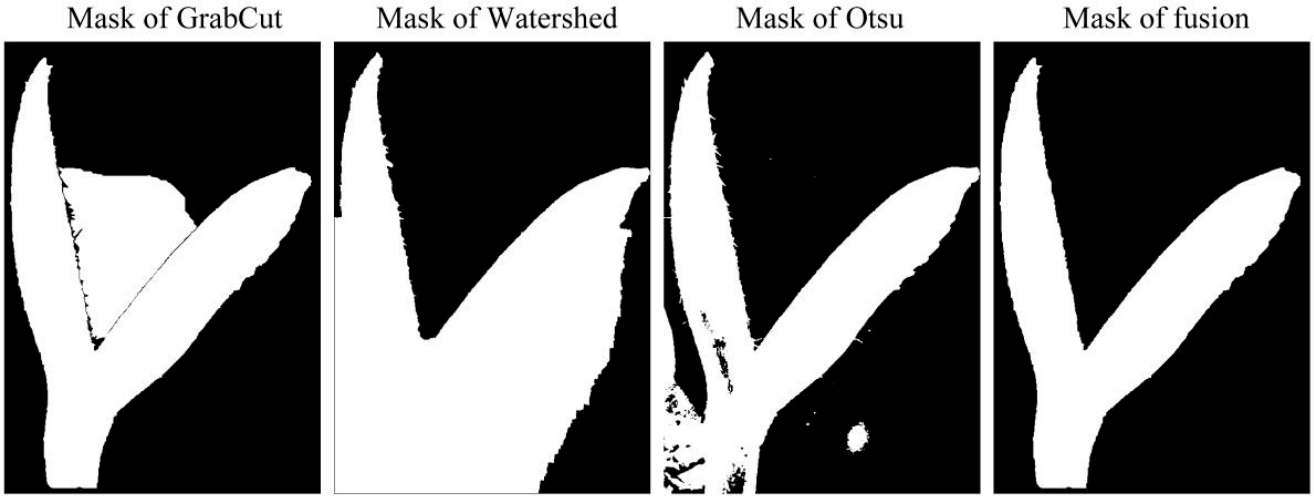
4.2.1. Target Extraction of Tea Buds

The basis of extracting the morphological and postural features of tea buds through PCA and skeleton extraction is the accurate segmentation of tea buds. In this study, the GrabCut algorithm was used to segment the mask of the tea bud, and the performance of the Watershed, Otsu, and other segmentation algorithms was compared with that of GrabCut. The results are shown in Figure 10. Figure 10a shows that, based on the detection results of the improved YOLOv8 model, there is still interference with the tea buds, such as old leaves and stems with similar colors in the detection box. Both the Otsu and Watershed algorithms struggle to effectively filter out the old leaves and stems, and the segmentation results contain obvious background regions. The GrabCut algorithm relies on the constraint of the initial detection box and focuses on the target area, which reduces the interference of similar background regions. However, there are still some incorrect segmentations of dark background regions. From the segmentation mask (Figure 10b), it can be seen that the mask edges generated by the GrabCut algorithm are smoother and have a higher degree of overlap with the actual contour of tea buds. However, the mask edges of the Otsu and Watershed algorithms are rough and contain some burrs and fractures, making it difficult to accurately reflect the morphological features of tea buds. After integrating GrabCut and

Watershed algorithms, the non-target areas are filtered out, and the complete contour of tea buds is retained, providing high-quality masks for subsequent posture evaluation.



(**a**) Comparison of segmentation results of GrabCut, Watershed, and Otsu algorithms



(**b**) Segmentation masks of GrabCut, Watershed, and Otsu algorithms

Figure 10. Segmentation results of tea buds.

4.2.2. Estimation of the Main Direction of Tea Buds Based on PCA

PCA was used to estimate the main direction of different tea buds, and the results are presented in Figure 11. The single bud in Figure 11a has a regular morphology, a complete contour, and no covering or interference (e.g., bent buds or clasped leaves), and its pixels are concentrated along the stem's growth direction. The main growth direction of

tea buds could be accurately identified, and the estimated results were highly consistent with the actual growth posture. In the case of one bud with one leaf (Figure 11b), the morphological features of the leaves interfere with the extraction of the main axis from the overall data distribution, and the structural differences between the bud and the clasped leaf could not be distinguished. As a result, the eigenvector corresponding to the maximum eigenvalue calculated by PCA deviates from the actual growth direction of the tea bud. In the absence of obvious overlapping among multiple tea buds (Figure 11c), PCA independently calculates the main direction through the segmented mask of each individual bud, avoiding interference between different buds, and it can still distinguish the growth direction of each bud. Furthermore, for the bent tea buds in Figure 11c, the pixel distribution shows a nonlinear extension. The main axis direction calculated by PCA is not along the actual growth direction of the bent stem, which leads to an incorrect clamping angle of the picking robotic arm and increases the risk of damage.

(**a**) Single bud (**b**) One bud with one leaf (**c**) Bent and multiple tea buds

Figure 11. Estimation results of the main direction of tea buds.

In conclusion, PCA has the advantages of high efficiency and stability for single buds with regular morphology, no bent buds, no clasped leaves, and no significant interference. However, PCA is a linear dimension reduction method, which is difficult to handle interference from bent buds, clasped leaves, irregular morphology, and complex backgrounds, and tends to cause deviation of the main direction from the actual growth posture. Therefore, this study proposed an integrated skeleton extraction algorithm to compensate for the shortcomings of PCA, which improved the accuracy of posture estimation by calculating parameters such as bending degree and local curvature in stages.

4.2.3. Estimation of Tea Bud Posture Based on Skeleton Extraction

Figure 12 shows the results of estimating the main direction of tea buds through skeleton extraction in scenes involving single buds, one bud with one leaf, multiple targets, and bent buds. It can be seen that the skeleton extraction algorithm accurately captures the growth structure of tea buds through topological structure analysis and curve fitting, which provides a morphological basis for the calculation of growth direction. For a single bud (Figure 12a), the skeleton extraction algorithm completely extracted the stem direction, formed a clear path, and accurately identified the endpoints. The main direction was consistent with the estimation result of PCA. For one bud with one leaf (Figure 12b), the skeleton extraction algorithm showed better performance than PCA, as it could identify three structures (the bud tip, leaf tip, stem tip, and branch point) and separate the stem through topological analysis, thereby avoiding the interference of leaves. For bent and multiple tea buds (Figure 12c), the skeleton extraction algorithm completely retains the bent shape. The nonlinear stem was fitted by a Bézier curve, and the distal tangent direction could

accurately align with the actual growth direction. In summary, the skeleton extraction algorithm effectively compensates for the shortcomings of the direction estimation deficiency of PCA. The calculation accuracy of the main direction was improved by approximately 20% and the robustness of the algorithm in complex scenes was significantly improved.

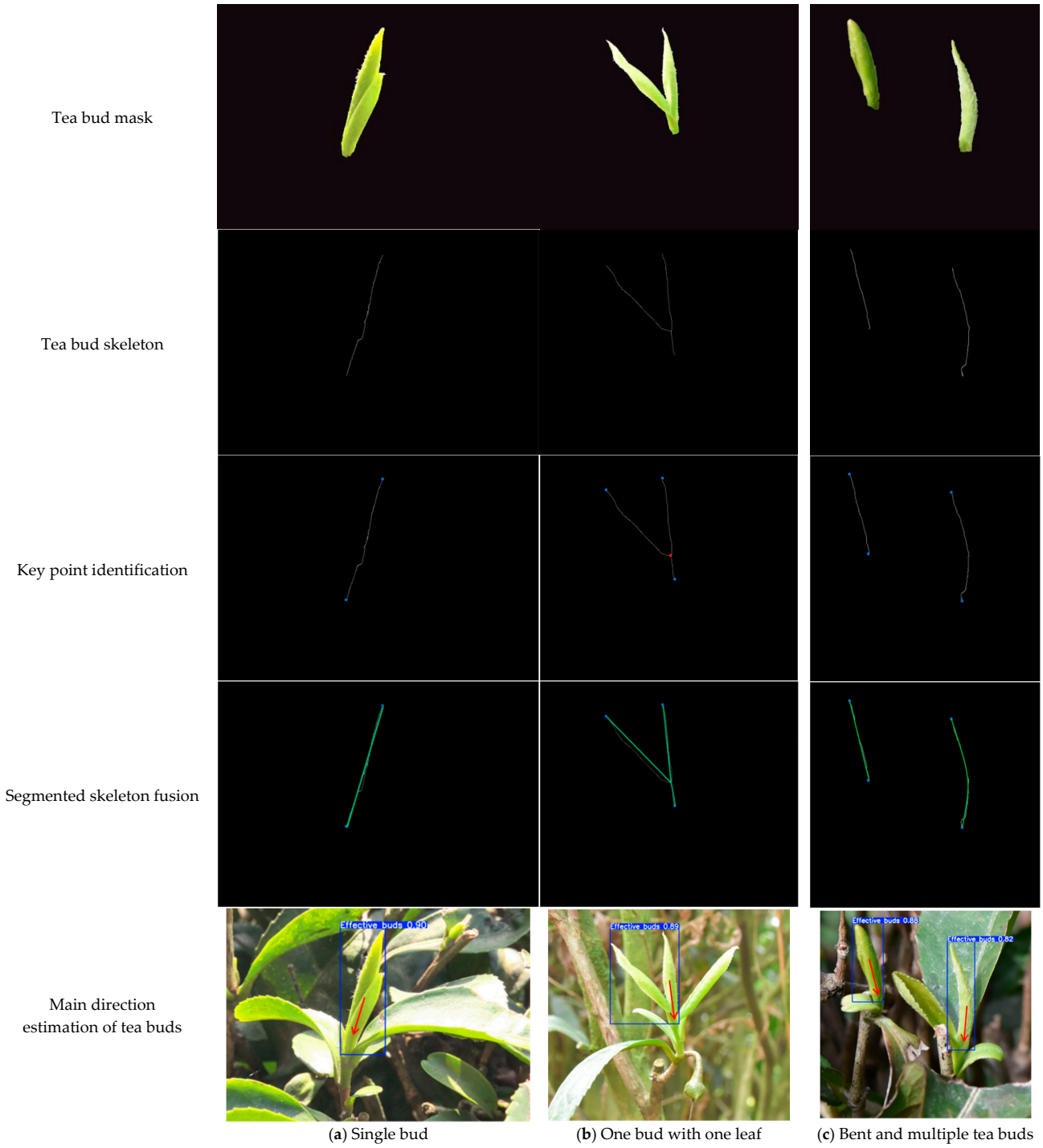
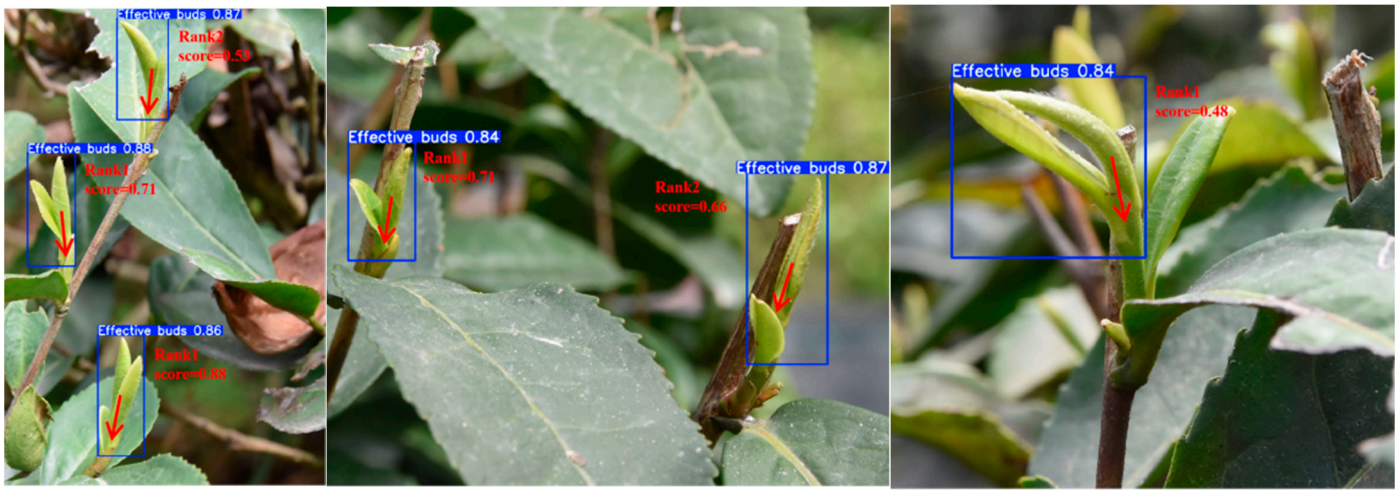


Figure 12. Results of skeleton extraction and endpoint detection.

4.2.4. Evaluation of Tea Bud Scoring Model

To verify the effectiveness of the tea bud picking model proposed in this study, the picking scores ($A \times B$) and picking priorities of effective buds in different scenes were ranked. The results are shown in Figure 13. For a single bud that is vertically upward ($\theta \approx 90°$), straight, and has no covering (Figure 13a), the picking robotic arm could perform

vertical clamping and picking operations. Path planning is simpler, and the picking priority is the highest. For partially unfolded tender leaves and slightly bent buds ($\theta \approx 60$–80°) (Figure 13b), the clamping angle of the picking robotic arm needs to be adjusted, and the complexity of path planning increases; the picking priority is secondary. For tea buds that are severely bent or overlapped (Figure 13c), the path planning is more complicated. The picking success rate is low, and the tender leaves are easily damaged; thus, the picking priority is the lowest. In summary, the scoring model based on the morphology and posture of tea buds quantifies the picking feasibility and provides effective guidance for picking robotic arms. In practical applications, it is necessary to combine environmental conditions (such as lighting intensity and wind speed, etc.) and the spatial information of tea buds to achieve 3-Degree-of-Freedom (3-DoF) dynamic posture evaluation, thereby improving the qualification rate of mechanical harvesting.



(**a**) Single bud (**b**) Partially unfolded tender leaves (**c**) Severely bent buds

Figure 13. Ranking results of picking scores of effective buds in different scenarios.

5. Conclusions

To meet the demand for mechanical picking of famous tea, this study developed a lightweight YOLOv8n–StarNet–ASF detection model. The model reduced the computational load by 23.6% while limiting the performance degradation to less than 2%. In addition, the improved model maintained robustness in complex tea garden environments. Results of field experiments showed that the precision rate (P), recall rate (R), $mAP50$, and $mAP50$-95 of the improved model were 85.6%, 81.7%, 90.5% and 65.2%, respectively. The staged posture estimation framework integrating PCA and skeleton extraction solved the direction deviation in scenes of bent buds and clasped leaves through stem tracking and Bézier fitting. Combined with the morphology–posture dual-factor scoring model, the ranking of tea buds for picking was realized, which accurately determined the picking priority in different scenes and provided coordinate + posture dual guidance for picking robotic arms. In the next step, the 3D point cloud data will be fused to realize the real-time estimation of depth information and dynamic posture, so as to further improve the picking adaptability of robotic arms under dense canopies.

Author Contributions: Conceptualization, P.W., W.Y., and S.M.; Methodology, P.W., W.Y., and S.M.; Data curation, T.H.; Funding acquisition, L.X.; Investigation, L.Z.; Resources, C.W.; Software, J.W.; Validation, Z.B.; Writing—original draft, T.H.; Writing—review and editing, L.X. All authors have read and agreed to the published version of the manuscript.

Funding: The authors sincerely appreciate the careful and precise reviews by the anonymous reviewers and editors. This work was supported by the open project of the Key Laboratory of Agricultural Equipment Technology for Hilly and Mountainous Areas, Ministry of Agriculture and Rural Affairs, China (2024QSNZ05), and the 2025 Modern Agricultural Science and Technology and Product Industrialization Demonstration Project (2025-14), Sichuan Academy of Agricultural Sciences.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. FAOSTAT. *Crops Statistics*; United Nations FAO: Rome, Italy, 2025. Available online: http://www.fao.org/faostat/en/#data/QC (accessed on 15 August 2025).
2. Lu, J.; Yang, Z.; Sun, Q.; Gao, Z.; Ma, W. A machine vision-based method for tea buds segmentation and picking point location used on a cloud platform. *Agronomy* **2023**, *13*, 1537. [CrossRef]
3. Wang, M.; Gu, J.; Wang, H.; Hu, T.; Fang, X.; Pan, Z. Method for identifying tea buds based on improved YOLOv5s model. *Trans. Chin. Soc. Agric. Eng.* **2023**, *39*, 150–157.
4. Tang, Y.; Han, W.; Hu, A.; Wang, W. Design and experiment of intelligentized tea-plucking machine for human riding based on machine vision. *Trans. Chin. Soc. Agric. Mach.* **2016**, *47*, 15–20.
5. Lin, Y.K.; Chen, S.F.; Kuo, Y.F.; Liu, T.L.; Lee, S.Y. Developing a guiding and growth status monitoring system for riding-type tea plucking machine using fully convolutional networks. *Comput. Electron. Agric.* **2021**, *191*, 106540. [CrossRef]
6. Zhou, Y.; Wu, Q.; He, L.; Zhao, R.; Jia, J.; Chen, J.; Wu, C. Design and experiment of intelligent picking robot for famous tea. *J. Mech. Eng.* **2022**, *58*, 12–23.
7. Yang, Z.; Ma, W.; Lu, J.; Tian, Z.; Peng, K. The application status and trends of machine vision in tea production. *Appl. Sci.* **2023**, *13*, 10744. [CrossRef]
8. Zhang, Z.; Lu, Y.; Yang, M.; Wang, G.; Zhao, Y.; Hu, Y. Optimal training strategy for high-performance detection model of multi-cultivar tea shoots based on deep learning methods. *Sci. Hortic.* **2024**, *328*, 112949. [CrossRef]
9. Xu, W.; Zhao, L.; Li, J.; Shang, S.; Ding, X.; Wang, T. Detection and classification of tea buds based on deep learning. *Comput. Electron. Agric.* **2022**, *192*, 106547. [CrossRef]
10. Zhang, F.; Sun, H.; Xie, S.; Dong, C.; Li, Y.; Xu, Y.; Zhang, Z.; Chen, F. A tea bud segmentation, detection and picking point localization based on the MDY7-3PTB model. *Front. Plant Sci.* **2023**, *14*, 1199473. [CrossRef]
11. Li, J.; Li, J.; Zhao, X.; Su, X.; Wu, W. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4. *Comput. Electron. Agric.* **2023**, *211*, 107955. [CrossRef]
12. Wang, Y. Study on Automatic Bud Recognition and Buds Classification System of JunMei Tea in Fujian Province. Master's Thesis, Fujian Agriculture and Forestry University, Fuzhou, China, 2020.
13. Long, Z.; Wang, J.; Li, G.; Zeng, C.; He, Y.; Li, B. An Algorithm for Identifying and Locating Tea Buds. China Patent CN111784764A, 16 October 2020.
14. Jiang, H.; He, B.; Zhang, Y. Research on the Method of Recognizing and Positioning the Shoots of the Tea Picking Manipulator. *Mach. Electron.* **2021**, *39*, 60–64+69.
15. Wang, W.; Xiao, H.; Chen, Q.; Song, Z.; Han, Y.; Ding, W. Research progress analysis of tea intelligent recognition and detection technology based on image processing. *J. Chin. Agric. Mech.* **2020**, *41*, 178–184.
16. Liu, F.; Wang, S.; Pang, S.; Han, Z. Detection and recognition of tea buds by integrating deep learning and image-processing algorithm. *Food Meas.* **2024**, *18*, 2744–2761. [CrossRef]
17. Wu, Y.; Chen, J.; He, L.; Gui, J.; Jia, J. An RGB-D object detection model with high-generalization ability applied to tea harvesting robot for outdoor cross-variety tea shoots detection. *J. Field Robot.* **2024**, *41*, 1167–1186. [CrossRef]
18. Wei, Y.; Wen, Y.; Huang, X.; Ma, P.; Wang, L.; Pan, Y.; Lv, Y.; Wang, H.; Zhang, L.; Wang, K.; et al. The dawn of intelligent technologies in tea industry. *Trends Food Sci. Technol.* **2024**, *144*, 104337. [CrossRef]
19. Zhu, M. *Research and Application of Recognition and Localization Algorithm for Tea Buds Based on Embedded Platform*; Central China Normal University: Wuhan, China, 2024.
20. Zhang, F. *Research and System Development of Tea Bud Recognition and Location Based on Computer Vision*; Hangzhou Dianzi University: Hangzhou, China, 2023.
21. Shi, W.; Yuan, W.; Yang, M.; Xu, G. A lightweight model for identifying the stalks of tea buds based on the improved YOLOv8n-seg. *Jiangsu J. Agric. Sci.* **2025**, *41*, 75–86.
22. Yang, D.; Huang, Z.; Zheng, C.; Chen, H.; Jiang, X. Detecting tea shoots using improved YOLOv8n. *Trans. Chin. Soc. Agric. Eng.* **2024**, *40*, 165–173.

23. Li, H.; Gao, Y.; Xiong, G.; Li, Y.; Yang, Y. Extracting tea bud contour and location of picking points in large scene using case segmentation. *Trans. Chin. Soc. Agric. Eng.* **2024**, *40*, 135–142.
24. Yu, T.; Chen, J.; Chen, Z.; Li, Y.; Tong, J.; Du, X. DMT: A model detecting multispecies of tea buds in multi-seasons. *Int. J. Agric. Biol. Eng.* **2024**, *17*, 199–208. [CrossRef]
25. Hu, C.; Tan, L.; Wang, W.; Song, M. Lightweight tea shoot picking point recognition model based on improved DeepLabV3+. *Smart Agric.* **2024**, *6*, 119–127.
26. Gu, J.; Wang, M.; Wang, H.; Hu, T.; Zhang, W.; Fang, X. Construction Method of Improved YOLOv5 Target Detection Model and Method for Identifying Tea Buds and Locating Picking Point. China Patent CN114882222A, 9 August 2022.
27. Cheng, Y.; Li, Y.; Zhang, R.; Gui, Z.; Dong, C.; Ma, R. Locating tea bud keypoints by keypoint detection method based on convolutional neural network. *Sustainability* **2023**, *15*, 6898. [CrossRef]
28. Guo, S.; Yoon, S.; Li, L.; Wang, W.; Zhuang, H.; Wei, C.; Liu, Y.; Li, Y. Recognition and positioning of fresh tea buds using YOLOv4-lighted + ICBAM model and RGB-D sensing. *Agriculture* **2023**, *13*, 518. [CrossRef]
29. Wei, T.; Zhang, J.; Wang, J.; Zhou, Q. Study of tea buds recognition and detection based on improved YOLOv7 model. *J. Intell. Agric. Mech.* **2024**, *5*, 42–50.
30. Luo, K.; Zhang, X.; Cao, C.; Wu, Z.; Qin, K.; Wang, C.; Li, W.; Chen, L.; Chen, W. Continuous identification of the tea shoot tip and accurate positioning of picking points for a harvesting from standard plantations. *Front. Plant Sci.* **2023**, *14*, 1211279. [CrossRef]
31. Li, Y.; He, L.; Jia, J.; Lv, J.; Chen, J.; Qiao, X.; Wu, C. In-field tea shoot detection and 3D localization using an RGB-D camera. *Comput. Electron. Agric.* **2021**, *185*, 106149. [CrossRef]
32. Zhang, Z. *Research into the Essential Technique of Intelligent Harvesting of the Famous Tea*; Zhongkai University of Agriculture and Engineering: Guangzhou, China, 2024.
33. Long, Z.; Jiang, Q.; Wang, J.; Zhu, H.; Li, B.; Wen, F. Research on method of tea flushes vision recognition and picking point localization. *Transducer Microsyst. Technol.* **2022**, *41*, 39–41+45.
34. Zou, L.; Zhang, L.; Wu, C.; Chen, J. A Method for Obtaining Location Information of Picking Points of Famous Tea Based on Machine Vision. China Patent CN112861654A, 2024.
35. Zhu, L.; Zhang, Z.; Lin, G.; Zhang, S.; Chen, J.; Chen, P.; Guo, X.; Lai, Y.; Deng, W.; Wang, M.; et al. A Secondary Locating Method for Tender Bud Picking of Famous Tea. China Patent CN116138036B, 2 April 2024.
36. Arısoy, M.V.; Uysal, İ. BiFPN-enhanced SwinDAT-based cherry variety classification with YOLOv8. *Sci. Rep.* **2025**, *15*, 5427. [CrossRef]
37. Guan, S.; Lin, Y.; Lin, G.; Su, P.; Huang, S.; Meng, X.; Liu, P.; Yan, J. Real-Time Detection and Counting of Wheat Spikes Based on Improved YOLOv10. *Agronomy* **2024**, *14*, 1936. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.