agronomy

MDPI

Article

TS-YOLO: An All-Day and Lightweight Tea Canopy Shoots Detection Model

Zhi Zhang [1], Yongzong Lu [1], Yiqiu Zhao [1], Qingmin Pan [1], Kuang Jin [1], Gang Xu [2] and Yongguang Hu [1,*]

[1] School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China; zzhi@stmail.ujs.edu.cn (Z.Z.); yzlu@ujs.edu.cn (Y.L.); zhaoyiqiu1997@163.com (Y.Z.); 2112116011@stmail.ujs.edu.cn (Q.P.); kjin@stmail.ujs.edu.cn (K.J.)
[2] Jurong Maoshan Tea Farm, Jurong 212443, China; g62433520@163.com
* Correspondence: deerhu@ujs.edu.cn

Abstract: Accurate and rapid detection of tea shoots within the tea canopy is essential for achieving the automatic picking of famous tea. The current detection models suffer from two main issues: low inference speed and difficulty in deployment on movable platforms, which constrain the development of intelligent tea picking equipment. Furthermore, the detection of tea canopy shoots is currently limited to natural daylight conditions, with no reported studies on detecting tea shoots under artificial light during the nighttime. Developing an all-day tea picking platform would significantly improve the efficiency of tea picking. In view of these problems, the research objective was to propose an all-day lightweight detection model for tea canopy shoots (TS-YOLO) based on YOLOv4. Firstly, image datasets of tea canopy shoots sample were collected under low light (6:30–7:30 and 18:30–19:30), medium light (8:00–9:00 and 17:00–18:00), high light (11:00–15:00), and artificial light at night. Then, the feature extraction network of YOLOv4 and the standard convolution of the entire network were replaced with the lightweight neural network MobilenetV3 and the depth-wise separable convolution. Finally, to compensate for the lack of feature extraction ability in the lightweight neural network, a deformable convolutional layer and coordinate attention modules were added to the network. The results showed that the improved model size was 11.78 M, 18.30% of that of YOLOv4, and the detection speed was improved by 11.68 FPS. The detection accuracy, recall, and AP of tea canopy shoots under different light conditions were 85.35%, 78.42%, and 82.12%, respectively, which were 1.08%, 12.52%, and 8.20% higher than MobileNetV3-YOLOv4, respectively. The developed lightweight model could effectively and rapidly detect tea canopy shoots under all-day light conditions, which provides the potential to develop an all-day intelligent tea picking platform.

Keywords: tea canopy shoots; all-day light conditions; YOLOv4; lightweight network

Citation: Zhang, Z.; Lu, Y.; Zhao, Y.; Pan, Q.; Jin, K.; Xu, G.; Hu, Y. TS-YOLO: An All-Day and Lightweight Tea Canopy Shoots Detection Model. *Agronomy* **2023**, *13*, 1411. https://doi.org/10.3390/agronomy13051411

Academic Editors: Baohua Zhang and Yongliang Qiao

Received: 3 May 2023
Revised: 15 May 2023
Accepted: 17 May 2023
Published: 19 May 2023

1. Introduction

Tea is the second most consumed beverage in the world [1,2]. While it is beneficial to human beings, harvesting tea is often a major challenge for farmers. Currently, there are two primary methods for harvesting tea, which are hand-picking (manual) and mechanical harvesting. Famous tea picking is highly time-sensitive, and the main problem with the hand-picking process is the time delay due to its time-consuming and labor-intensive nature [3]. Although the mechanical harvesting method partly improves labor productivity, its "one-size-fits-all" cutting operation greatly reduces the economic value of tea products [4], especially Chinese famous tea production, which is limited as nearly all the tea shoots are manually picked.

With the development of agricultural harvesting robots, developing intelligent famous tea picking platforms is a vital trend to promote the famous tea industry. Accurate and rapid detection of tea canopy shoots in complex field environments is one of the crucial technologies for intelligent picking platforms. Computer vision technology has been

widely applied in target detection of various fruits and vegetables, such as apple [5], tomato [6], strawberry [7], kiwifruit [8], and grape [9]. The primary techniques used for tea shoot detection involve traditional image processing and deep learning methods. Traditional image processing methods typically rely on differences in color, texture, and shape between the target foreground and background to extract the detection target [10,11]. Wu et al. proposed a method to detect tea shoots based on image G and G-B component information, and to automatically extract segmentation thresholds through maximum variance [3]. Yang et al. used the G component as a color feature to segment the background and tea shoots with the double thresholds, and detected the edges of tea leaves based on shape features [12]. Zhang et al. employed the process of improved G-B algorithm graying, median filtering, OTSU binarization processing, morphological processing, and edge smoothing to extract the tea fresh leaves shape from the RGB images of the tea canopy [13]. Karunasena et al. developed a cascade classifier based on the histogram of oriented gradients features and support vector machine to detect tea shoots [14]. Zhang et al. constructed G-B◎components to enhance the distinction between tea shoots and background in images by a segmented linear transformation, and then detected tea shoots based on the watershed segmentation algorithm [15]. The effectiveness of image feature extraction is crucial for the detection performance of the above-mentioned methods, but it is often compromised by the complex and variable light conditions of the tea field environment.

The rapid advancement of deep learning techniques has led to the deployment of numerous deep learning models for recognition and detection tasks of agricultural robots in unstructured environments [16]. These models are designed to leverage the ability of automatic feature extraction to enhance detection performance and improve robustness [17]. Zhu et al. constructed a tea shoots detection model based on the Faster RCNN and evaluated the model detection performance under different shoot types. That model had the highest detection accuracy for one bud and one leave/two leaves with an *AP* of 76% [18]. Xu et al. compared the detection performance of Faster RCNN and SSD models with VGG16, ResNet50, and ResNet101 as feature extraction networks for tea shoots, and found that the Faster RCNN with VGG16 as its feature extraction network had the better detection performance with the precision of 85.14%, recall of 78.90%, and a mAP of 82.17% [19]. Lv et al. compared several detection models based on the same dataset, and their results revealed that YOLOv5+CSPDarknet53 outperformed SSD+VGG16, Faster RCNN+VGG16, YOLOv3+Darknet53, and YOLOv4+CSPDarknet53 for the detection of tea shoots, with precision and recall of 88.2% and 82.1%, respectively [20]. Yang et al. proposed an improved YOLOv3 model for the detection of tea shoots by adding an image pyramid structure and residual block structure, and the average detection accuracy was found to be over 90% [21]. Xu et al. proposed a two-level fusion model for tea bud◎detection with an accuracy of 71.32%. The detection process used YOLOv3 to extract the tea shoot regions from the input images, followed by classification of the extracted regions using DenseNet201 [22]. Using deep learning methods for detecting tea shoots have be shown to demonstrate a significantly better performance compared to traditional image processing methods, thanks to their excellent feature-extracting ability. As the depth of the network layers and the number of model parameters increase, it becomes increasingly challenging to deploy deep learning models on movable and embedded devices with limited computing power. This limitation poses a challenge to the development of intelligent tea picking equipment that requires real-time and on-site tea shoots detection. Furthermore, previous research mainly focused on the detection of tea shoots under natural light conditions, and to our knowledge, there are no reports of detection under artificial light conditions at night. Since nighttime takes up one-third of the whole day, the efficiency of the all-day work will be significantly improved with continuous and effective harvesting at night [23]. Tea harvesting is time-sensitive, and tea shoots must be picked at the right time to ensure the best quality of tea. Enabling all-day picking, including at night, can significantly increase the efficiency of the harvest and the income of tea farmers.

The current detection models have slow inference speed and are not easily deployable on movable platforms, which hinders the development of intelligent tea picking equipment. Furthermore, the detection of tea canopy shoots is currently limited to natural daylight conditions, with no reported studies on detecting tea shoots under artificial lighting during the nighttime. Developing an all-day tea picking platform would therefore significantly improve the efficiency of tea picking. Considering these issues, the research objective for our study was to propose an all-day lightweight detection model for tea canopy shoots (TS-YOLO) based on YOLOv4. The main contributions of this study were:

(1) To collect an image dataset of tea canopy shoots samples under natural light and artificial light at night, and to establish and annotate an all-day light conditions image dataset of tea canopy shoots;
(2) To reduce the model size and increase the inference speed, with the feature extraction network of YOLOv4 and the standard convolution of entire network being replaced by the lightweight neural network and the depth-wise separable convolution;
(3) A deformable convolutional layer and coordinate attention modules were introduced into the network to compensate for the shortage of the lightweight neural network on feature extraction ability.

We constructed an image dataset of tea canopy shoots under natural daylight and artificial light conditions at night in tea plantations, and proposed our TS-YOLO model which combines YOLOv4, MobileNetV3, deformable convolutional, and coordinate attention modules. Our model can efficiently and accurately detect tea canopy shoots under natural daylight and artificial light conditions, making it an innovative all-day application.

2. Materials and Methods

2.1. Image Data Acquisition

All tea canopy shoots images were collected in April 2022 at Jiangsu Yinchunbiya Tea filed located in Danyang, China (latitude 32°01′35″ N, longitude 119°40′21″ E) (Figure 1). The image acquisition devices utilized were a digital camera (Canon Power Shot SX30 IS) and a smartphone (iPhone 8). The sampled tea variety was Zhongcha 108, which has a strong tenderness and resistance to adversity, and is a common variety in the middle and lower parts of the Yangtze River regions. For diversity enrichment of the image dataset, images were acquired with different shooting angles and heights under different weather (sunny and cloudy) and light (low light at 7:30–8:30 and 17:30–18:30, medium light at 9:00–10:00 and 16:00–17:00, high light at 11:00–15:00, and artificial light of LED at night) conditions. A total of 2417 images were acquired with the image resolutions of 4320 pixels × 3240 pixels and 4032 pixels × 3024 pixels, respectively. The dataset could improve the models robustness and applicability in the case study of tea canopy shoots, and particularly, it would help to develop an efficient all-day picking platform for famous tea by collecting images under artificial light at night (Figure 2).

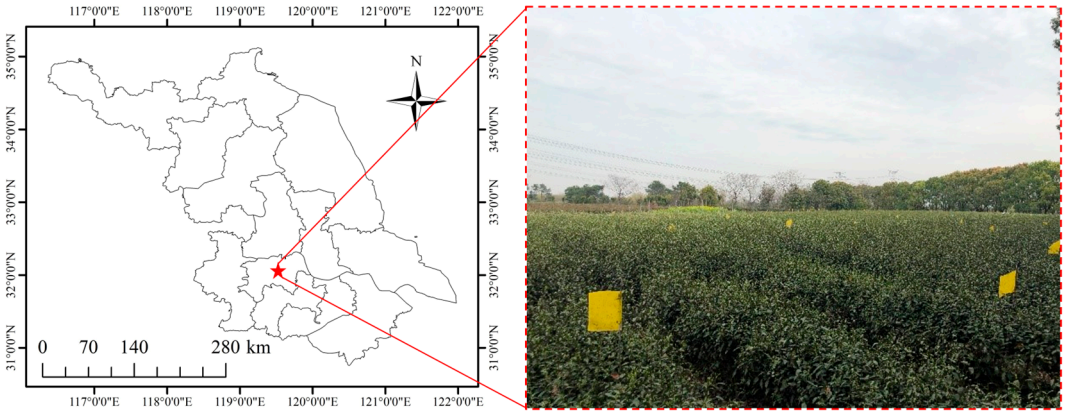

Figure 1. Acquisition site of the tea canopy shoot images.

(**a**) (**b**) (**c**) (**d**)

Figure 2. Tea canopy shoots under different light conditions. (**a**) Low light; (**b**) medium light; (**c**) high light; and (**d**) artificial light.

2.2. Images Annotation and Dataset Production

LabelImg was used to annotate one bud and one leaf (as "TS") of the tea canopy shoots, and tea shoots that were more than 75% occluded or blurred were not annotated (Figure 3), and the annotation information obtained was saved in XML format [24]. The training set, validation set, and testing set were randomly divided in a ratio of 6:2:2. For enhancing the richness of the experimental dataset and improving the generalization ability of the model, the dataset was expanded via rotating, mirroring, adding noise, and randomly changing the brightness and contrast (Figure 4). Data augmentation was performed for each training image in the dataset using random combinations of the above five methods. Meanwhile, the corresponding annotation file of each image was transformed. The final number of images in the training set, validation set, and testing set are 4347, 484, and 484, respectively. The division of the image dataset under different light conditions was shown in Table 1.

Figure 3. Annotation of the tea canopy shoots.

(**a**) (**b**) (**c**)

(**d**) (**e**) (**f**)

Figure 4. Data augmentation. (**a**) Original image; (**b**) rotate; (**c**) mirror; (**d**) Gaussian noise; (**e**) brightness variation; and (**f**) contrast enhancement.

Table 1. Statistics of datasets under different light conditions.

Light Conditions	Original			Data Enhancement		
	Training	Validation	Testing	Training	Validation	Testing
Low	370	124	124	1110	124	124
Moderate	362	121	121	1086	121	121
Intense	361	120	120	1083	120	120
Artificial	357	119	119	1068	119	119
Sum	1449	484	484	4347	484	484

2.3. YOLOv4

The YOLO series unifies the tasks of object classification and bounding box regression as a single regression problem to object detection. Its essential idea was using the whole image as the model input and directly regresses the position and class of prediction boxes in the output layer [25]. The input image was divided into $S \times S$ grids, and one target prediction was achieved through the grid if the target center fell into it (Figure 5). The YOLOv4 was proposed based on YOLOv3, and the main network structure includes the backbone network, neck network, and head network [26]. CSPDarknet53 is the backbone feature extraction network that uses the CSP (cross-stage-partial-connections) structure to divide the feature map of the base layer into two parts, and then merge them through cross-stage hierarchy to reduce the computation while also maintaining accuracy [27]. In the neck network, the feature maps after pooling operations with different size kernels were concatenated together through SPP (spatial pyramid pooling network), which could extract spatial features in different sizes and improve the robustness of the model on spatial layout and object deformation [28]. Then, the acquired feature was enhanced by PAN (path aggregation networks), which added a bottom-up path to FPN to effectively keep the lower layers of localization information [29]. Compared with two-stage detection models such as Fast RCNN and Faster RCNN, YOLOv4 exhibited a significantly faster inference time while utilizing a simpler structure (Figure 6).

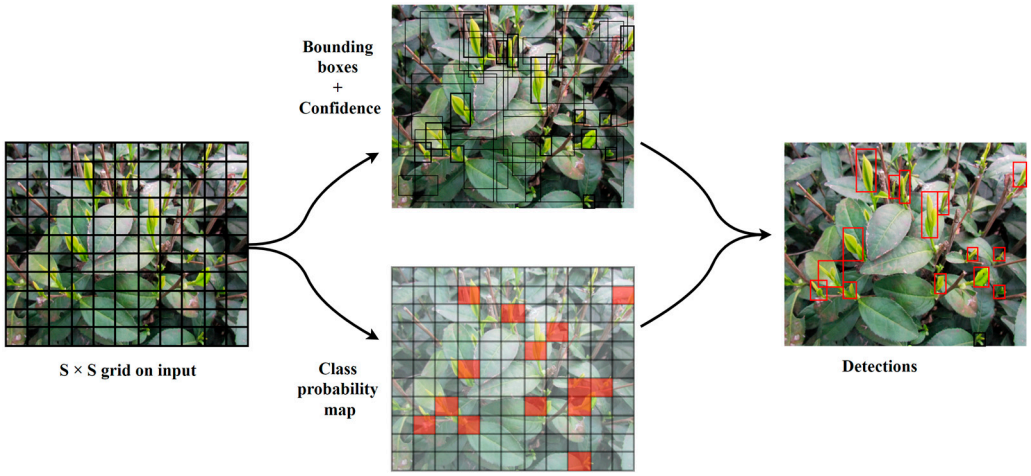

Figure 5. The detection principle of YOLO.

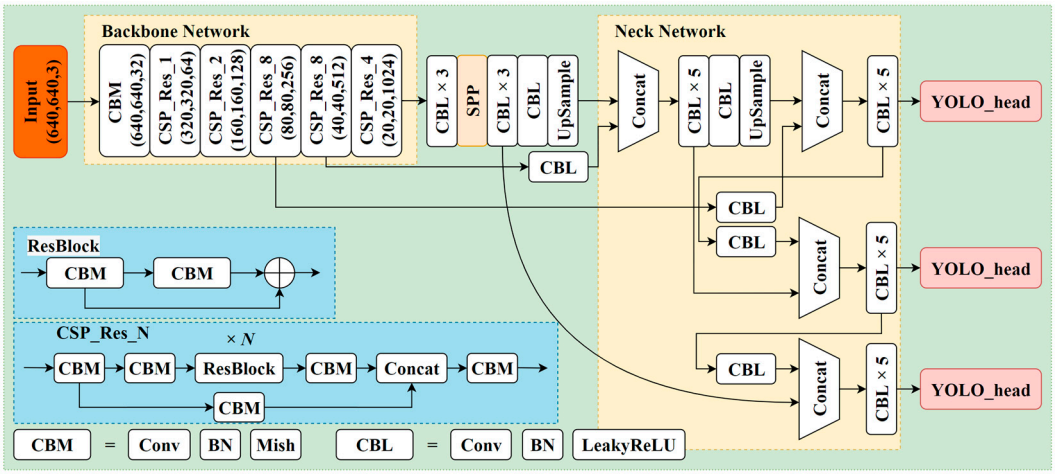

Figure 6. Structure of YOLOv4.

2.4. Improved YOLOv4 Model (TS-YOLO)

To achieve accuracy and rapid detection of tea canopy shoots in an unstructured environment, we proposed an improved target detection model (TS-YOLO) in this paper (Figure 7). Firstly, a lightweight neural network MobileNetV3 was used as the feature extraction network to reduce the model size and improve the inference speed, which extracts features from the input images, and obtains prediction feature layers of different sizes after down sampling. Then, a deformable convolutional layer and coordinate attention modules were added to the network to compensate for the shortage of the lightweight neural network on feature extraction ability. Finally, to further reduce the model size and improve the inference speed of the model, the standard convolution of the whole network was replaced by depth-wise separable convolution (DepC). The training loss of the improved model follows the loss function of YOLO, which includes confidence loss, location loss, and classification loss.

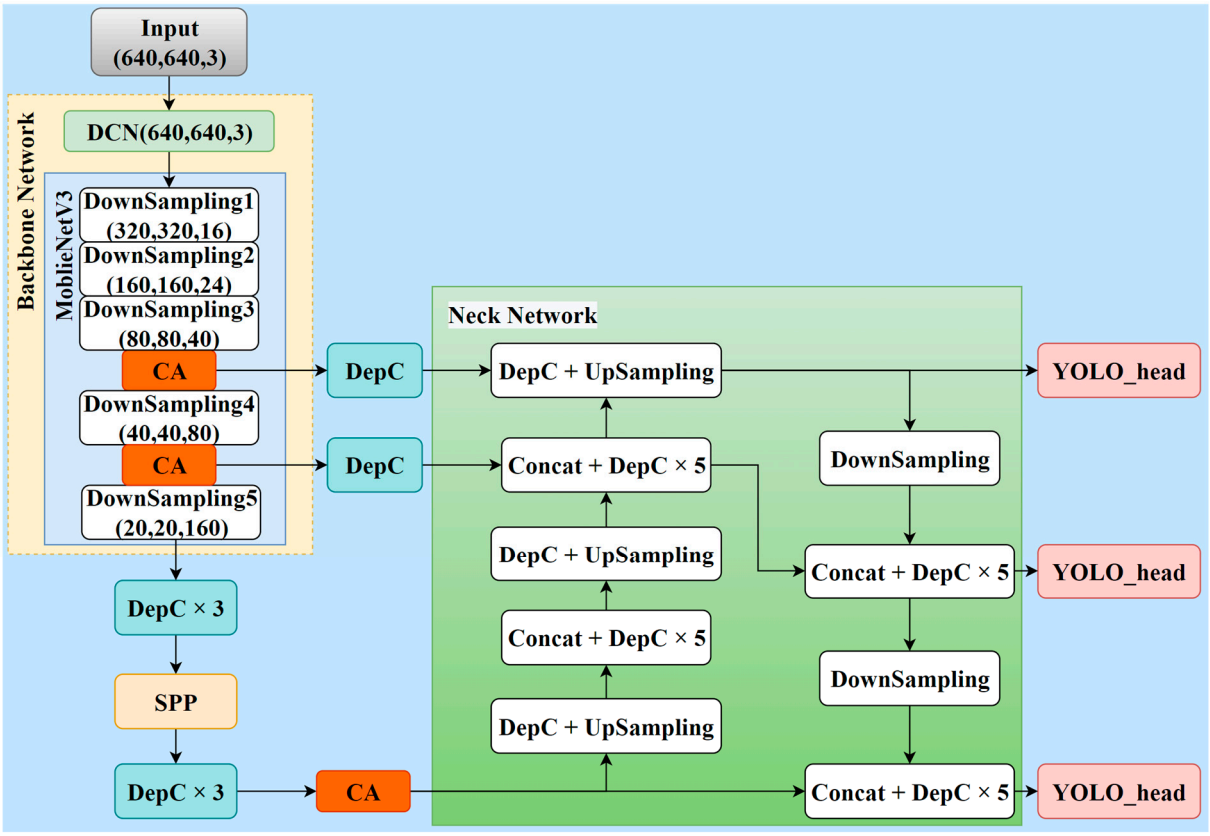

Figure 7. Overall structure of TS-YOLO.

2.4.1. MobilenetV3

Although convolutional neural networks have a high detection performance, deepening network layers and increasing complexity can lead to more parameters and a slower inference speed, making them less suitable for real-time target detection on mobile platforms. For mobile devices with resource constraints, lightweight convolutional neural networks have significant advantages in terms of inference speed and the number of parameters. MobileNet is a lightweight deep neural network based on depth-wise separable convolution proposed by the Google team in 2017, which greatly reduces the model parameters and operations with a slightly decreased accuracy compared with traditional convolutional neural networks [30]. Depth-wise separable convolution divides standard convolution into depth-wise convolution and pointwise convolution. MobileNetV3 combined the depth-wise separable convolution of V1, the inverted residual and linear bottlenecks of V2 [31], and introduced the SE attention module, using NAS (neural architecture search) technology to search the configuration and parameters of the network, which further improves the performance and inference speed of the model [32]. The base module of MobileNetV3 is shown in Figure 8.

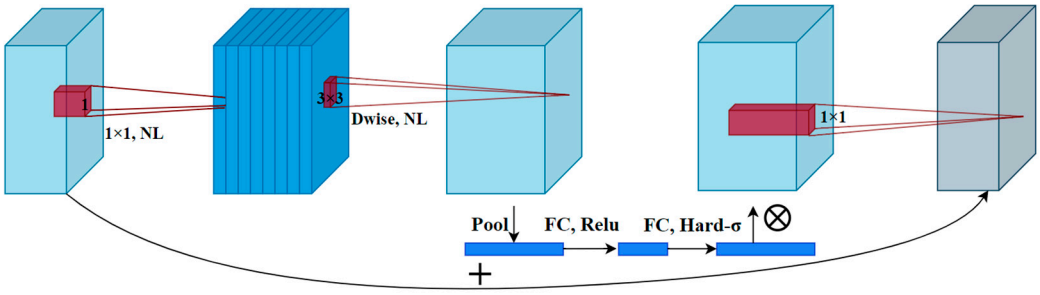

Figure 8. The MobileNetV3 base module.

2.4.2. Deformable Convolution

The traditional convolution kernel is typically of a fixed-size square structure with a poor adaptability and generalization ability to irregular targets. During the sampling process, the valid features outside of the sampling region are either ignored or incorrectly divided into other sampling regions, while the invalid features are not ignored. Deformable convolution is an excellent approach to solving this problem [33]. By adapting the sampling regions, deformable convolution can learn more valid features that might have been ignored by traditional convolution, resulting in improved model robustness. Deformable convolution introduces additional learnable parameters, termed offset parameters, to each element of the convolution kernel, which control the sampling locations of the input feature map and allow for more flexible sampling. During training, deformable convolution can be extended to a larger range, and allows the sampling region to adaptively deform according to the actual shape of the detected target, thus adapting to geometric deformations such as the shape and size of objects (Figure 9). The specific procedures of deformable convolution are as follows:

(1) Extraction of features from input feature maps using traditional convolution kernels;
(2) Applying another convolution layer to the feature map obtained in the first step, obtaining deformable convolution offsets with 2N number of channels;
(3) During training, convolutional kernels for generating output features and convolutional kernels for generating offsets are learned simultaneously through backpropagation, where offsets are learned by interpolation algorithms.

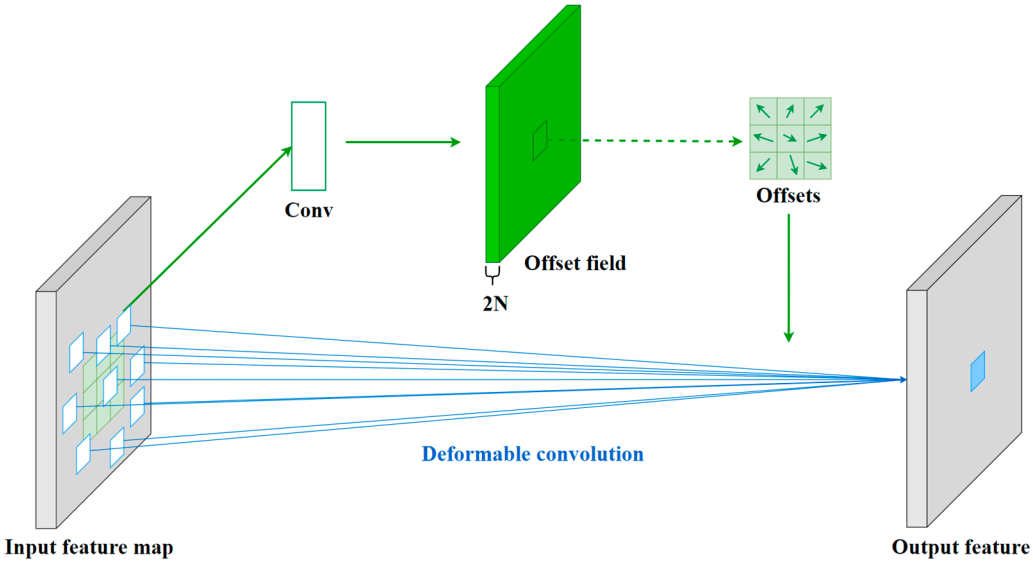

Figure 9. Schematic diagram of the deformable convolution.

2.4.3. Coordinate Attention Module

In the tea canopy, tea shoots are easily disturbed by the illumination intensity and shading of the leaves, branches, and trunks, which leads to missed and false detection. The attention modules enable the model to focus on relevant local information, and enhance its concentration on the tea shoots region, resulting in an improved detection performance. The most common attention modules in computer vision are the SE module (squeeze-and-excitation attention module) [34], the CBAM (convolutional block attention module) [35], and the BAM (bottleneck attention module) [36], etc. The SE module only considers internal channel information and ignores the important position information and spatial structure in detection tasks. The BAM and CBAM can collect local location information via global pooling operations, but they strip away spatial attention and channel attention. The coordinate attention module (CA) can maintain channel information

while acquiring more distant position information [37]. It has two steps: coordinate information embedding and coordinate attention generation. Coordinate information embedding aggregates features along with two spatial directions, generating a pair of direction-aware attention maps. Coordinate attention generation produces attention maps with the global field of perception and precise location information. The structure of CA is shown in Figure 10.

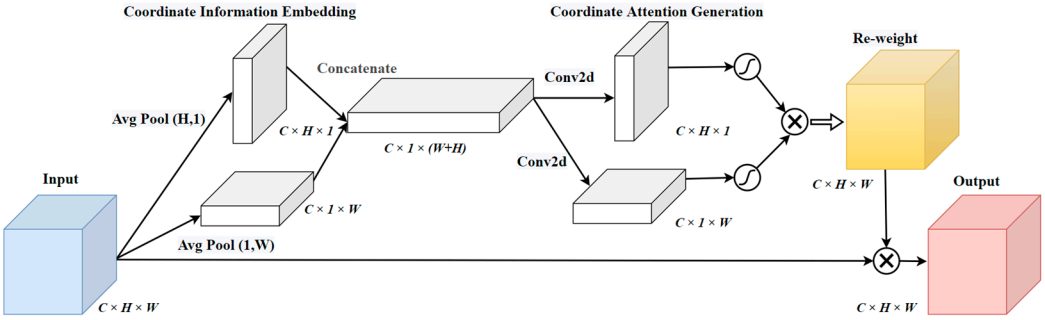

Figure 10. Schematic diagram of the coordinate attention module.

3. Results and Analysis

The hardware environment used for the experiment is shown in Table 2. The standard stochastic gradient descent was used to train the models. The momentum set was 0.937, the initial learning rate was 0.001, and the weight decay was 0.0005. Considering the calculation speed of model training, the input image size was set to 640×640, the batch size was 4, and a total of 100 epochs were performed.

Table 2. Experimental configuration.

Configuration	Parameter
CPU	AMD 3700X
GPU	Nvidia GeForce RTX 3080TI
Operating system	Windows10
Accelerated environment	CUDA 11.6, cuDNN 8.3.2
Library	PyTorch 1.13.1

3.1. Evaluation of Model Performance

The performance of the trained models in detecting the tea canopy shoots was evaluated using common target detection metrics, including precision (*P*), recall (*R*), average precision (*AP*), model size, and frame per second (FPS). The equations for the relevant metrics are as follows:

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

$$AP = \int_0^1 P(R)dR \tag{3}$$

where *TP* is the sample accurately predicted as tea shoot by the model, *FP* is the sample falsely predicted as tea shoot by the model, *FN* is the sample wrongly judged as background, and *AP* is the area under the *P*-*R* curve. The precision evaluates the percentage of

objects in the returned list which are correctly detected, and the recall evaluates the percentage of correctly detected objects in total.

3.2. Performance Effect of Different Modules on the Model

3.2.1. Anchor Boxes Optimization and Data Augmentation

The YOLO series network utilizes anchor boxes as a prior box to aid in predicting the boundaries of the targets, and the appropriate size of the anchor boxes can further enhance the performance of target detection. In this paper, based on the annotation information of the training dataset, $1 - IoU$ was used as the clustering distance, and the size of the anchor boxes was calculated by the k-means algorithm. Nine groups of anchor boxes {(13,21), (15,35), (24,26), (21,52), (35,40), (31,71), (57,58), (47,106), and (85,154)} were obtained, and the average IoU was 0.74, which was found to be 0.12 higher than the default size of the anchor boxes in YOLOv4. To improve the generalization ability of the model based on limited datasets and mitigate overfitting, data augmentation was used to enable the model learning more robust features [38]. The performance impact on these models after the process of anchors boxes optimization (AO) and data augmentation (DA) is shown in Table 3.

Table 3. Model performance after AO and DA.

Parameters	*AP* (%)	*R* (%)	*P* (%)
YOLOv4	64.53	57.14	80.22
YOLOv4 + AO	65.16	57.15	82.41
YOLOv4 + DA	84.05	77.08	87.17
YOLOv4 + AO + DA	84.61	78.08	87.69

After the process of AO, the *P* was improved by 2.19%, while *AP* and *R* did not change significantly. After the process of DA, the *AP*, *R*, and *P* were all improved by 19.52%, 19.19%, and 6.95%, respectively. After the process of combining AO and DA, the *AP*, *R*, and *P* were all improved by 20.08%, 20.91%, and 7.47%, respectively.

3.2.2. Lightweight Convolutional Neural Networks

To improve the portability of the model and increase the inference speed based on AO and DA, lightweight neural networks were used as the feature extraction network, and depth-wise separable convolution was applied in replacing standard convolution in the neck network. Five kinds of lightweight neural networks, which were ShuffleNetV2 [39], MobileNetV2, MobileNetV3, EfficientNetV2 [40], and GhostNet [41], were compared and analyzed (Table 4).

Table 4. Performance of different lightweight models.

Feature Extraction Network	*AP* (%)	*R* (%)	*P* (%)	Model Size (M)	FPS
CSPDarknet53	84.61	78.08	87.69	64.36	37.18
ShuffleNetV2	67.33	60.83	83.34	9.89	50.6
MobileNetV2	64.11	53.81	84.06	10.80	54.39
MobileNetV3	73.92	65.90	84.27	11.73	49.34
EfficientNetV2	70.56	63.78	84.75	28.84	28.17
GhostNet	72.13	68.73	83.90	11.43	41.36

After replacing the original feature extraction network with the lightweight neural network, MobileNetV3 was found to have the highest *AP* value of 73.92%, which was 10.69% lower than that of CSPDarknet53. For EfficientNetV2, it was found to have the highest *P* value of 84.75%, which was 2.94% lower than that of CSPDarknet53. GhostNet had the highest *R*

value of 68.73%, which was 9.35% lower compared to CSPDarknet53. ShuffleNetV2 had the smallest model size of 9.89 M, which was 84.63% lower compared to CSPDarknet53. MobileNetV2 had the highest FPS of 54.39, which improved by 46.29% compared to CSPDarknet53. Although the model size of EfficientNetV2 was found to be significantly lower than that of CSPDarknet53, FPS decreased rather than increased, unlike the results published in other studies [42,43], which may be caused by the compatibility of the experimental hardware platform with the model inference process. The training loss value of different models all plummeted at the 70th epoch, which may be caused by the change of learning rate during model training. With the combination of the validation loss value, all the models converged off after the 70th epoch in the different models (Figure 11). To balance the model size, inference speed, and detection performance, MobileNetV3 was chosen as the feature extraction network in this paper. Similar to previous research [44–46], using MobileNetV3 as the feature extraction network had achieved a better detection performance.

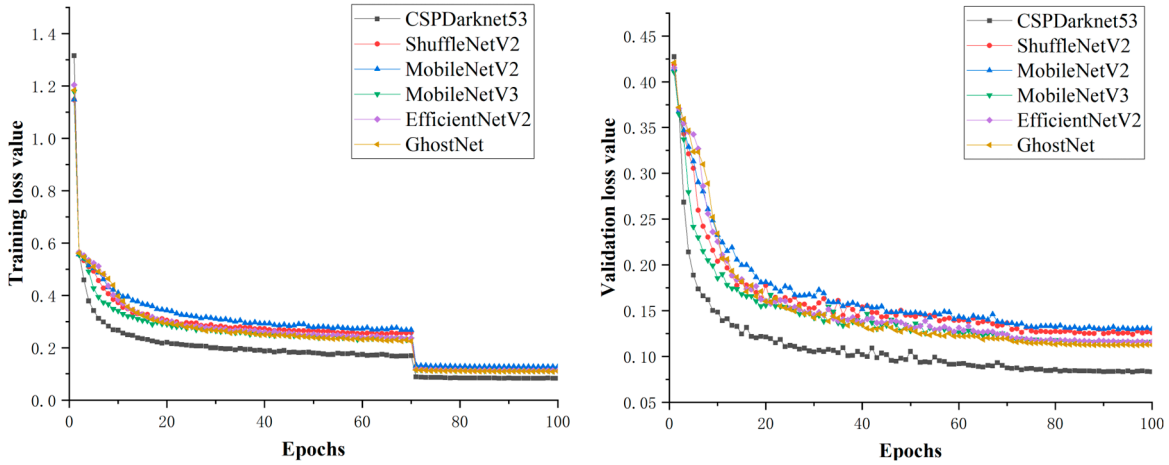

Figure 11. Loss change in the training process for different models.

3.2.3. Ablation Experiments

To evaluate the effectiveness of the proposed improved model on detection performance, we validated the model performance using different modules based on YOLOv4 (Table 5). MobileNetV3 was used as the feature extraction network to develop the model 1; A deformable convolutional layer was added in model 1 to develop the model 2; SE, CBAM, and CA attention modules were added in model 2 to establish the models 3, 4, and 5, respectively.

Table 5. Ablation experiment results.

Models	*AP* (%)	*R*(%)	*P* (%)	Model Size (M)	FPS
Model 1	73.92	65.90	84.27	11.73	49.34
Model 2	75.38	73.86	84.93	11.73	49.37
Model 3	78.77	75.26	85.20	11.76	49.34
Model 4	81.05	77.26	85.72	11.80	48.56
Model 5	82.12	78.42	85.35	11.78	48.86

The backbone structure of the lightweight network was relatively simple, and the detection performance of the tea tree canopy shoots of different morphologies and sizes was yet to be improved. To improve the detection performance of the model, a deformable convolutional layer and attention modules were added to improve the models ability to extract complex features. As shown in Table 4, when a deformable convolutional layer was added, the *R* value was significantly improved by 7.96%, compared to model 1 with almost no change of model size and inference speed. When the attention modules were introduced, the detection performance of the model was further improved. Among them,

model 3 with the added SE modules had improved *AP*, *R*, and *P* by 3.39%, 1.4%, and 0.27%, respectively, compared to model 2. Model 4 with the added CBAM modules had improved *AP*, *R*, and *P* by 5.67%, 3.4%, and 0.79%, respectively, compared to model 2. Model 5 with the added CA modules had improved *AP*, *P*, and *R* by 6.74%, 4.56%, and 0.42%, respectively compared to model 2. Heat map visualization of the detection process of the tea canopy shoots by Grad-CAM for the model adding attention modules was shown in Figure 12. After adding CA, the focus range of the model became broader and more focused compared to SE and CBAM. Thus, when CA was introduced, it effectively improved the detection performance of the model for the tea canopy shoots.

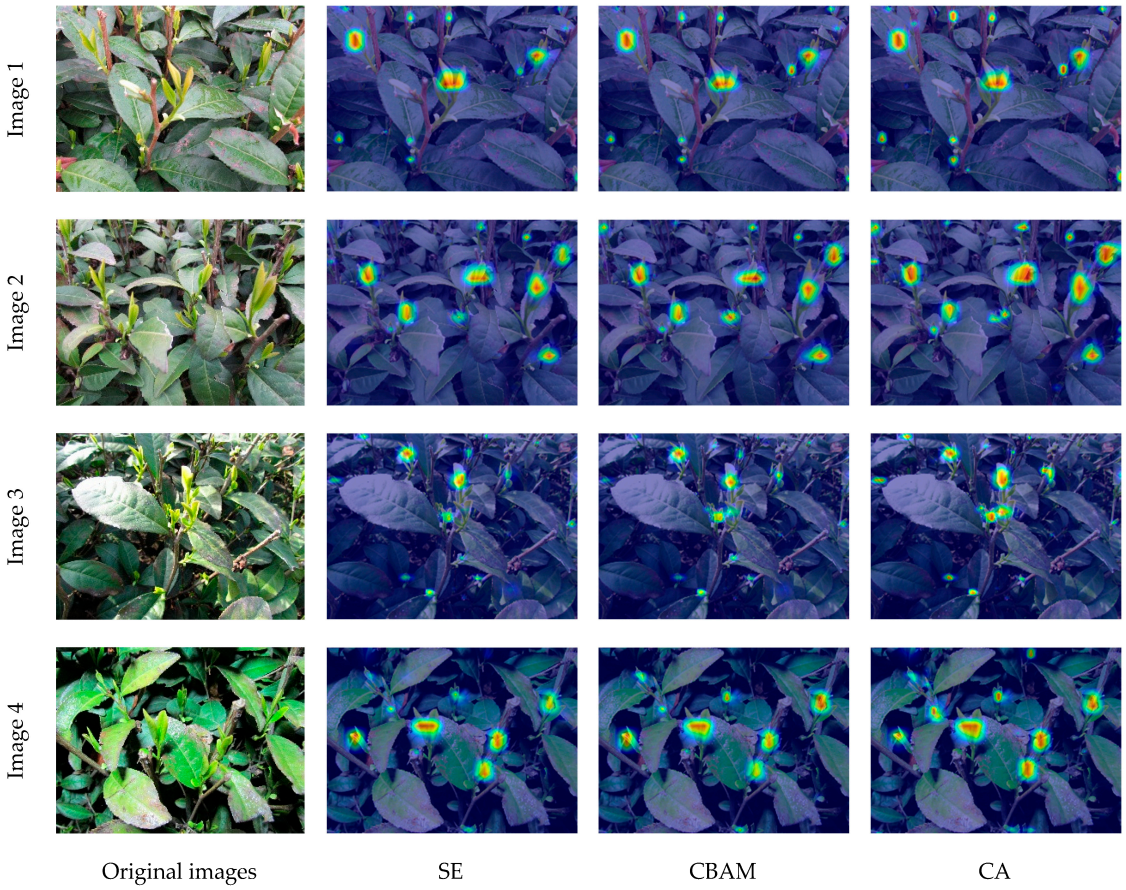

Figure 12. Visualization results of heat map with adding different attention modules.

3.3. Detection Performance under Different Light Conditions

The complex and variable light conditions in the field environment are crucial factors that affect the accuracy of target detection tasks, and the tea canopy exhibits diverse characteristics that vary under different lighting conditions. As illustrated in Figure 2, in low light conditions, the tea canopy shoots exhibited a bright yellow marginal part, with clearly visible light green veins on the leaves. Moreover, the branches and old leaves of the tea trees displayed a greater degree of color difference from the shoots, and dew can be observed on the surface of old leaves. Under medium light conditions, the color differentiation between the tea shoots and old leaves was reduced, and the color of tea shoots became greener. However, the contours of the tea shoots remained clearly defined, making it possible to detect them accurately. Under high light conditions, the high intensity of the light can cause reflection on the surface of old leaves and tea shoots, which can make it challenging to detect and distinguish them from the surrounding environment. Moisture condensation on the surface of tea leaves can occur due to high environmental humidity at night, while the reflection phenomenon on the surfaces of tea leaves and shoots can be caused by high light exposure. The non-uniformity of light

intensity can cause shadows to appear under high light and artificial light conditions, which can further complicate the detection of tea canopy shoots. Table 6 presents the detection performance of the model for tea canopy shoots under various light conditions.

Table 6. Detection performance under different light conditions.

Light Conditions	*AP* (%)	*R* (%)	*P* (%)
Low	82.94	78.31	85.82
Medium	83.44	78.96	85.93
High	82.73	77.74	85.70
Artificial	82.68	77.58	85.87

Under medium light conditions, the model's detection performance was the best, with *AP*, *P*, and *R* of 83.44%, 78.96%, and 85.93%, respectively. The model's detection performance was the worst under artificial light conditions at night, as indicated by the lowest *AP*, *P*, and *R* values of 82.68%, 77.58%, and 85.87%, respectively. Despite several variations in the detection performance of the model under different light conditions, the differences observed were relatively small. Therefore, it can be inferred that the model exhibits a good robustness in detecting tea canopy shoots throughout the day, regardless of variations in the natural or artificial lighting conditions.

3.4. Comparative Experiments of the Different Detection Models

In this paper, different object detection models were compared with proposed TS-YOLO, such as Faster RCNN, SSD, YOLOv3, YOLOv4, M-YOLOv4 (MobileNetV3-YOLOv4), and YOLOv5, and experimental results are shown in Table 7.

Table 7. Detection results of tea shoots by the different detection models.

Models	*AP* (%)	*R* (%)	*P* (%)	Model Size (M)	FPS
Faster RCNN	68.24	76.02	48.98	138.31	42.23
SSD	78.60	61.41	83.74	34.17	58.95
YOLOv3	80.19	60.52	86.33	61.95	61.11
M-YOLOv4	73.92	65.90	84.27	11.73	49.34
YOLOv4	84.61	78.08	87.69	64.36	37.18
YOLOv5	79.29	71.72	85.94	21.19	50.28
TS-YOLO	82.12	78.42	85.35	11.78	48.86

Based on the results, the two-stage detection model Faster RCNN exhibited significantly lower *AP* and *P* values compared to the other models. Faster R-CNN does not incorporate image feature pyramid, which may therefore limit its ability to accurately detect objects of different scales and sizes. The image feature pyramid is a commonly used technique in object detection models, which involves extracting multi-scale features from different layers of the network. These features are then used to detect objects of varying sizes and scales. Compared with YOLOv4, the proposed TS-YOLO *AP* and *P* values decreased by 2.49% and 2.34%, respectively, but the model size was reduced by 81.70% and inference speed was increased by 31.41%. Compared with M-YOLOv4, the *AP*, R, and *P* values of TS-YOLO increased by 8.20%, 12.52%, and 1.08%, respectively. Compared with YOLOv5 (the selected YOLOv5m, which has a similar size to the proposed model), the *AP* and *R* values of TS-YOLO increased by 2.83% and 6.70%, while the model size was reduced by 44.40%, respectively. The comparison results revealed that there is a trade-off between the complexity of the network structure and the model detection performance. *AP* is a comprehensive evaluation index of model precision and recall, while FPS measures the model's inference speed. However, there is currently no evaluation index that considers both the detection performance, and the inference speed of these object detection models. In practical applications, it is necessary to comprehensively consider

the detection performance and inference speed of the model in conjunction with the computing performance of the picking platform. On high-performance computing platforms, *AP* can be given more weight since it has little impact on the real-time detection performance. However, on platforms with limited computing resources, both *AP* and the inference speed of the model should be considered to meet the requirements of real-time detection. TS-YOLO uses a trade-off strategy to balance the detection performance and the inference speed. By reducing the model size and optimizing the network architecture, it can achieve a faster inference speed while maintaining a certain level of detection performance. In the future, we aim to focus on improving the model by implementing high-accuracy strategies to minimize the loss of detection performance. The results of these different models for the detection of tea canopy shoots are as shown in Figure 13.

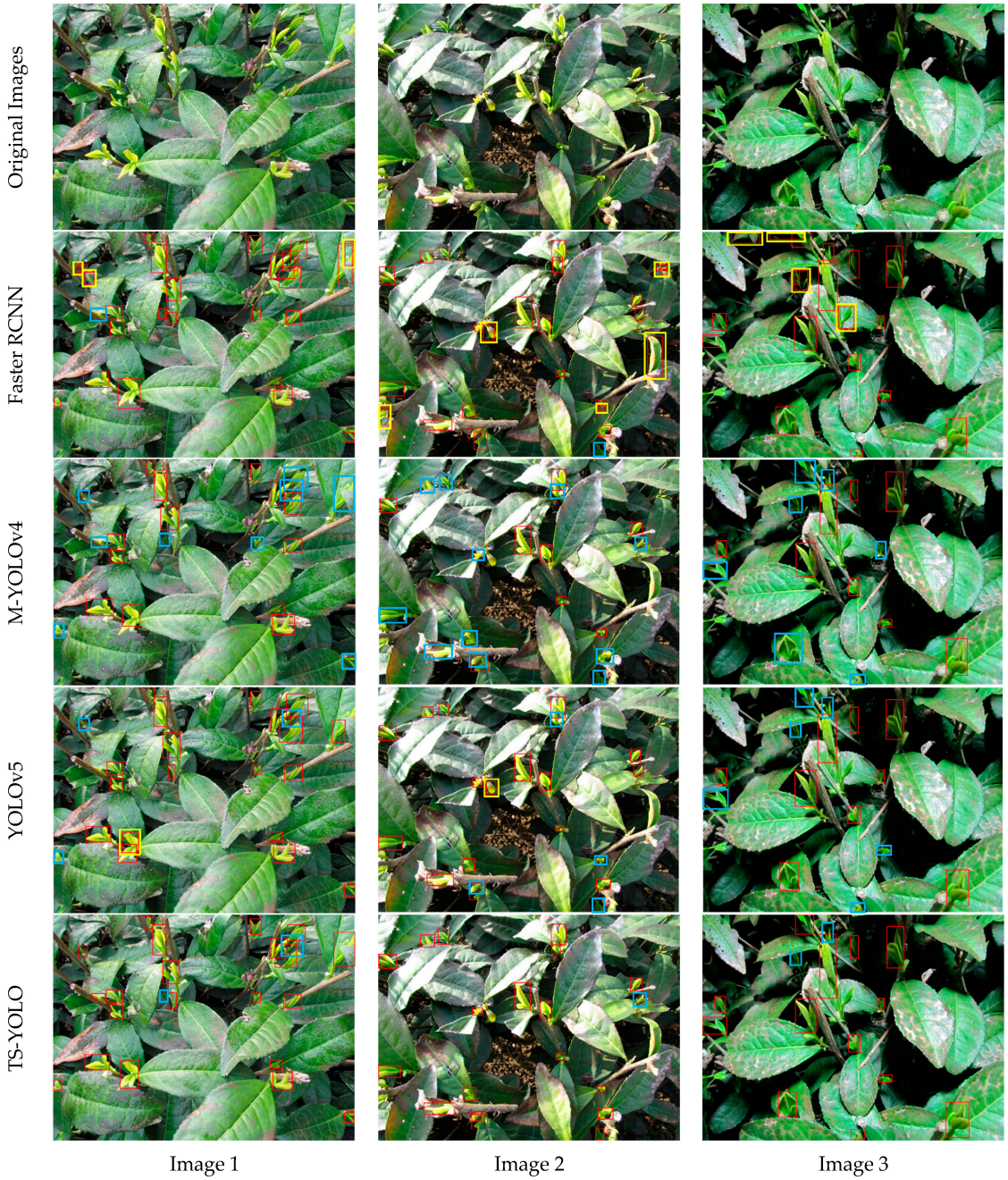

Figure 13. Detection results of the different models. Yellow boxes are false detections, and blue boxes are missed detection.

4. Discussion

The results of this study compared to other studies are summarized in Table 8. Yang et al. [12], Wu et al. [47], Karunasena et al. [14], and Zhang et al. [15], used traditional image processing methods for the detection of tea shoots. When using traditional image processing methods for target detection, the feature characters used for the description are artificially designed, and the method performs well for detection performance when the image is clear, uniformly illuminated, and minimally occluded. In the practical tea field, however, these conditions are often not met. Among the deep learning methods, Zhu et al. [18], Wang et al. [48], Li et al. [49], Wang et al. [50], and Chen et al. [51] used Faster RCNN, Mask RCNN, YOLOV3, YOLOv5, and so on, to detect the tea shoots, respectively. Although its detection results are better and the robustness to complex field environments are higher, the large model size and slow inference speed are not suitable to be deployed on movable platforms for the real-time detection of tea canopy shoots. With respect to model light-weighting, it is mainly achieved by using lightweight modules and model compression. Gui et al. used ghost convolution to replace the standard convolution and added the bottleneck attention module to the backbone feature extraction network [52]. Huang et al. replaced the feature extraction network with GohstNet and replaced the standard convolution in the neck network with ghost convolution [53]. Cao et al. introduced the GhostNet module and coordinated attention module in the feature extraction network and replaced PAN with BiFPN [54]. Guo et al. add attention modules and replaced PAN with FPN to achieve a lightweight model [55]. Compared with these related studies, the detection performance of the proposed model in this paper was found to be slightly lower, and its main reasons were probably the following: (1) The dataset used in this paper was acquired under natural and artificial light conditions with more complex light variations; (2) The height and angle of the shots during image capture wee variable, and the morphology of the tea shoots were more diverse compared to the fixed height and angle shots. Thus, for further improving the detection performance of the model for all-day tea canopy shoots, the following approaches will be used for future research: (a) Elimination of the effects of light variations with image enhancement processing; (b) Combination with the tea picking platform, with the suitable height and angle to take images; (c) Multiple detections can be realized by adjusting the position of the picking platform cameras to improve the picking success rate. In conclusion, this study introduces a novel model, TS-YOLO, for detecting tea canopy shoots, and creates an image dataset captured under varying lighting conditions, including under natural daylight and artificial light at night. The proposed model exhibits a high efficiency and accuracy in detecting tea canopy shoots under all-day lighting conditions, which has significant implications for the development of all-day intelligent tea-picking platforms.

Table 8. Detection performance of this paper compared with other papers.

References	Methods/Model	*AP* (%)	*P* (%)	*R* (%)	Accuracy (%)	Model Size (M)	FPS
Yang et al. [12]	Color and shape features	——	——	——	94.0	——	——
Wu et al. [47]	K-means	——	——	——	94.0	——	——
Karunasena et al. [14]	SVM	——	——	——	55.0	——	——
Zhang et al. [15]	Watershed algorithm	——	——	——	95.79	——	——
Zhu et al. [18]	Faster RCNN	76.0	98.0	76.0	——	——	5.0
Wang et al. [48]	Mask RCNN	——	——	94.62	95.53	——	——
Li et al. [49]	YOLOv3	——	93.10	89.30	——	——	——
Wang et al. [50]	YOLOv5	75.80	94.90	75.70	——	8.80	——
Chen et al. [51]	YOLOv3	——	74.51	69.56	——	——	——
Gui et al. [52]	YOLOv5(lightweight)	92.66	88.82	87.99	——	23.85	29.51
Huang et al. [53]	YOLOv4(lightweight)	72.93	51.07	78.67	——	——	32.10

Cao et al. [54]	YOLOv5(lightweight)	——	76.31	88.42	——	10.0	——
Guo et al. [55]	YOLOv4(lightweight)	——	94.19	93.50	——	——	——
Our study	TS-YOLO	82.12	85.35	78.42	——	11.78	48.86

5. Conclusions

The research proposed an all-day lightweight detection model for tea canopy shoots (TS-YOLO) based on YOLOv4, which employed MobileNetV3 as the backbone network for YOLOv4, and replaced the standard convolution with depth-wise separable convolution to achieve the reduction in model size and increase the inference speed. To overcome the detection limitations, a deformable convolutional layer and coordinate attention modules were introduced. Compared with YOLOv4, the TS-YOLO model size was 18.30% of it, and the detection speed was improved by 11.68 FPS. The detection accuracy, recall, and *AP* of tea canopy shoots under different light conditions were 85.35%, 78.42%, and 82.12%, respectively, which were 1.08%, 12.52%, and 8.20% higher than that of MobileNetV3-YOLOv4, respectively.

While this study yielded promising results, there were two limitations that require attention. Firstly, the position, phenotype, and occlusions during the picking process must be considered to determine whether the tea canopy shoot can be harvested. Secondly, to improve the model's applicability across various tea varieties, future research should integrate an intelligent tea picking platform to analyze the harvestability of the detected tea shoots and evaluate the model's effectiveness.

Although there were several minor research limitations, the developed lightweight model has demonstrated its efficacy in detecting tea canopy shoots quickly and effectively, even under all-day light conditions. This breakthrough could pave the way for the development of an all-day intelligent tea picking platform, which could revolutionize the tea industry.

Author Contributions: Conceptualization: Z.Z.; data curation: Z.Z., Y.Z., Q.P. and K.J.; formal analysis: Z.Z.; funding acquisition: Y.H.; investigation: Z.Z.; methodology: Z.Z.; project administration: Z.Z. and Y.H.; resources: Z.Z.; software: Z.Z.; supervision: Z.Z., Y.L. and Y.H.; validation: Z.Z.; visualization: Z.Z.; writing—original draft: Z.Z.; writing—review and editing: Z.Z., Y.L., G.X. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD-2018-87), the China and Jiangsu Postdoctoral Science Foundation (2022M711396 and 2021K614C), the Fund of Key Laboratory of Modern Agricultural Equipment and Technology, Jiangsu University (MAET202119), the Fund of Jiangsu Province and Education Ministry Co-sponsored Synergistic Innovation Center of Modern Agricultural Equipment (XTCX2013), and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (21KJB210019).

Data Availability Statement: Not applicable.

Acknowledgments: The principal author is extremely express our gratitude to the School of Agricultural Engineering, Jiangsu University, for providing the essential instruments without which this work would not have been possible. We would like to thank the anonymous reviewers for their precious attention.

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

References

1. Zhu, J.C.; Niu, Y.; Xiao, Z.B. Characterization of the key aroma compounds in Laoshan green teas by application of odour activity value (OAV), gas chromatography-mass spectrometry-olfactometry (GC-MS-O) and comprehensive two-dimensional gas chromatography mass spectrometry (GC× GC-qMS). *Food Chem.* **2021**, *339*, 128136.
2. Ning, J.; Cao, Q.; Su, H.; Zhu, X.; Wang, K.; Wan, X.; Zhang, Z. Discrimination of six tea categories coming from different origins depending on polyphenols, caffeine, and theanine combined with different discriminant analysis. *Int. J. Food Prop.* **2017**, *20* (Suppl. 2), 1838–1847.
3. Wu, X.M.; Zhang, F.G.; Lv, J.T. Research on recognition of tea tender leaf based on image color information. *J. Tea Sci.* **2013**, *33*, 584–589.

4. Chen, Y.T.; Chen, S.F. Localizing plucking points of tea leaves using deep convolutional neural networks. *Comput. Electron. Agric.* **2020**, *171*, 105298.
5. Wang, D.; He, D. Channel pruned YOLO V5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosyst. Eng.* **2021**, *210*, 271–281.
6. Cardellicchio, A.; Solimani, F.; Dimauro, G.; Petrozza, A.; Summerer, S.; Cellini, F.; Reno, V. Detection of tomato plant phenotyping traits using YOLOv5-based single stage detectors. *Comput. Electron. Agric.* **2023**, *207*, 107757.
7. Fan, Y.; Zhang, S.; Feng, K.; Qian, K.; Wang, Y.; Qin, S. Strawberry maturity recognition algorithm combining dark channel enhancement and YOLOv5. *Sensors* **2022**, *22*, 419.
8. Ma, L.; He, Z.; Zhu, Y.; Jia, L.; Wang, Y.; Ding, X.; Cui, Y. A Method of Grasping Detection for Kiwifruit Harvesting Robot Based on Deep Learning. *Agronomy* **2022**, *12*, 3096.
9. Sozzi, M.; Cantalamessa, S.; Cogato, A.; Kayad, A.; Marinello, F. Automatic bunch detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms. *Agronomy* **2022**, *12*, 319.
10. Bargoti, S.; Underwood, J.P. Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Robot.* **2017**, *34*, 1039–1060.
11. Leemans, V.; Destain, M.F. A real-time grading method of apples based on features extracted from defects. *J. Food Eng.* **2004**, *61*, 83–89.
12. Yang, F.; Yang, L.; Tian, Y.; Yang, Q. Recognition of the tea sprout based on color and shape features. *Trans. Chin. Soc. Agric. Mach.* **2009**, *40*, 19–123.
13. Zhang, L.; Zhang, H.; Chen, Y.; Dai, S.; Li, X.; Kenji, I.; Liu, Z.; Li, M. Real-time monitoring of optimum timing for harvesting fresh tea leaves based on machine vision. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 6–9.
14. Karunasena, G.; Priyankara, H. Tea bud leaf identification by using machine learning and image processing techniques. *Int. J. Sci. Eng. Res.* **2020**, *10*, 624–628.
15. Zhang, L.; Zou, L.; Wu, C.; Jia, J.; Chen, J. Method of famous tea sprout identification and segmentation based on improved watershed algorithm. *Comput. Electron. Agric.* **2021**, *184*, 106108.
16. Tang, Y.; Chen, M.; Wang, C.; Luo, L.; Li, J.; Lian, G.; Zou, X. Recognition and localization methods for vision-based fruit picking robots: A review. *Front. Plant Sci.* **2020**, *11*, 510.
17. Kang, H.; Chen, C. Fast implementation of real-time fruit detection in apple orchards using deep learning. *Comput. Electron. Agric.* **2020**, *168*, 105108.
18. Zhu, H.; Li, X.; Meng, Y.; Yang, H.; Xu, Z.; Li, Z. Tea Bud Detection Based on Faster R-CNN Network. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 217–224.
19. Xu, G.; Zhang, Y.; Lai, X. Recognition approaches of tea bud image based on faster R-CNN depth network. *J. Optoelectron.·Laser* **2020**, *31*, 1131–1139.
20. Jun, L.; Mengrui, F.; Qing, Y. Detection model for tea buds based on region brightness adaptive correction. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 278–285.
21. Yang, H.; Chen, L.; Chen, M.; Ma, Z.; Deng, F.; Li, M.; Li, X. Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model. *IEEE Access* **2019**, *7*, 180998–181011.
22. 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.
23. Liu, X.; Zhao, D.; Jia, W.; Ruan, C.; Tang, S.; Shen, T. A method of segmenting apples at night based on color and position information. *Comput. Electron. Agric.* **2016**, *122*, 118–123.
24. Tzutalin, D. LabelImg. Available online: https://github.com/tzutalin/labelImg (accessed on 21 October 2022).
25. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
26. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
27. Wang, C.Y.; Liao HY, M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 390–391.
28. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916.
29. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
30. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
31. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
32. Howard, A.; Sandler, M.; Chu, G.; Chen, L.C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
33. Dai, J.; Qi, H.; Xiong, Y.; Li, Y.; Zhang, G.; Hu, H.; Wei, Y. Deformable convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 764–773.

34. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
35. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
36. Park, J.; Woo, S.; Lee, J.Y.; Kweon, I.S. Bam: Bottleneck attention module. *arXiv* **2018,** arXiv:1807.06514.
37. Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
38. Wang, L.; Zhao, Y.; Liu, S.; Li, Y.; Chen, S.; Lan, Y. Precision detection of dense plums in orchards using the improved YOLOv4 model. *Front. Plant Sci.* **2022**, *13*, 839269.
39. Ma, N.; Zhang, X.; Zheng, H.T.; Sun, J. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 116–131.
40. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual Event, 18–24 July 2021; pp. 10096–10106.
41. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. Ghostnet: More features from cheap operations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1580–1589.
42. Li, K.; Zhu, J.; Li, N. Lightweight automatic identification and location detection model of farmland pests. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 9937038.
43. Yu, L.; Pu, Y.; Cen, H.; Li, J.; Liu, S.; Nie, J.; Ge, J.; Lv, L.; Li, Y.; Xu, Y.; et al. A Lightweight Neural Network-Based Method for Detecting Estrus Behavior in Ewes. *Agriculture* **2022**, *12*, 1207.
44. Lang, X.; Ren, Z.; Wan, D.; Zhang, Y.; Shu, S. MR-YOLO: An Improved YOLOv5 Network for Detecting Magnetic Ring Surface Defects. *Sensors* **2022**, *22*, 9897.
45. Zeng, T.; Li, S.; Song, Q.; Zhong, F.; Wei, X. Lightweight tomato real-time detection method based on improved YOLO and mobile deployment. *Comput. Electron. Agric.* **2023**, *205*, 107625.
46. Liu, L.; Ke, C.; Lin, H.; Xu, H. Research on pedestrian detection algorithm based on MobileNet-YOLO. *Comput. Intell. Neurosci.* **2022**, *2022*, 8924027.
47. Wu, X.; Tang, X.; Zhang, F.; Gu, J. Tea buds image identification based on lab color model and K-means clustering. *J. Chin. Agric. Mech.* **2015**, *36*, 161–164.
48. Wang, T.; Zhang, K.; Zhang, W.; Wang, R.; Wan, S.; Rao, Y.; Jiang, Z.; Gu, L. Tea picking point detection and location based on Mask-RCNN. *Inf. Process. Agric.* **2021**, *10*, 267–275.
49. 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.
50. Wang, J.; Li, X.; Yang, G.; Wang, F.; Men, S.; Xu, B.; Xu, Z.; Yang, H.; Yan, L. Research on Tea Trees Germination Density Detection Based on Improved YOLOv5. *Forests* **2022**, *13*, 2091.
51. Chen, C.; Lu, J.; Zhou, M.; Yi, J.; Liao, M.; Gao, Z. A YOLOv3-based computer vision system for identification of tea buds and the picking point. *Comput. Electron. Agric.* **2022**, *198*, 107116.
52. Gui, Z.; Chen, J.; Li, Y.; Chen, Z.; Wu, C.; Dong, C. A lightweight tea bud detection model based on Yolov5. *Comput. Electron. Agric.* **2023**, *205*, 107636.
53. Huang, J.; Tang, A.; Chen, G.; Zhang, D.; Gao, F.; Chen, T. Mobile recognition solution of tea buds based on compact-YOLOv4 algorithm. *Trans. Chin. Soc. Agric. Mach.* **2023**. Available online: https://kns.cnki.net/kcms/detail/11.1964.S.20230113.1315.002.html (accessed on 15 March 2023).
54. Cao, M.; Fu, H.; Zhu, J.; Cai, C. Lightweight tea bud recognition network integrating GhostNet and YOLOv5. *Math. Biosci. Eng. MBE* **2022**, *19*, 12897–12914.
55. Guo, S.; Yoon, S.C.; Li, L.; 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.