agronomy

MDPI

Article

A Real-Time Detection Algorithm for Sweet Cherry Fruit Maturity Based on YOLOX in the Natural Environment

Zhiyong Li [1,2,†], Xueqin Jiang [1,2,†], Luyu Shuai [1,2], Boda Zhang [1,2], Yiyu Yang [1,2] and Jiong Mu [1,2,*]

[1] College of Information Engineering, Sichuan Agricultural University, Ya'an 625000, China
[2] Sichuan Key Laboratory of Agricultural Information Engineering, Ya'an 625000, China
* Correspondence: jmu@sicau.edu.cn; Tel.: +86-133-4060-8699
† These authors contributed equally to this work.

Abstract: Fast, accurate, and non-destructive large-scale detection of sweet cherry ripeness is the key to determining the optimal harvesting period and accurate grading by ripeness. Due to the complexity and variability of the orchard environment and the multi-scale, obscured, and even overlapping fruit, there are still problems of low detection accuracy even using the mainstream algorithm YOLOX in the absence of a large amount of tagging data. In this paper, we proposed an improved YOLOX target detection algorithm to quickly and accurately detect sweet cherry ripeness categories in complex environments. Firstly, we took a total of 2400 high-resolution images of immature, semi-ripe, and ripe sweet cherries in an orchard in Hanyuan County, Sichuan Province, including complex environments such as sunny days, cloudy days, branch and leaf shading, fruit overlapping, distant views, and similar colors of green fruits and leaves, and formed a dataset dedicated to sweet cherry ripeness detection by manually labeling 36068 samples, named SweetCherry. On this basis, an improved YOLOX target detection algorithm YOLOX-EIoU-CBAM was proposed, which embedded the Convolutional Block Attention Module (CBAM) between the backbone and neck of the YOLOX model to improve the model's attention to different channels, spaces capability, and replaced the original bounding box loss function of the YOLOX model with Efficient IoU (EIoU) loss to make the regression of the prediction box more accurate. Finally, we validated the feasibility and reliability of the YOLOX-EIoU-CBAM network on the SweetCherry dataset. The experimental results showed that the method in this paper significantly outperforms the traditional Faster R-CNN and SSD300 algorithms in terms of mean Average Precision (mAP), recall, model size, and single-image inference time. Compared with the YOLOX model, the mAP of this method is improved by 4.12%, recall is improved by 4.6%, F-score is improved by 2.34%, while model size and single-image inference time remain basically comparable. The method in this paper can cope well with complex backgrounds such as fruit overlap, branch and leaf occlusion, and can provide a data base and technical reference for other similar target detection problems.

Keywords: orchard environment; YOLOX; target detection; dataset; cherry; CBAM; EIoU

check for updates

Citation: Li, Z.; Jiang, X.; Shuai, L.; Zhang, B.; Yang, Y.; Mu, J. A Real-Time Detection Algorithm for Sweet Cherry Fruit Maturity Based on YOLOX in the Natural Environment. *Agronomy* **2022**, *12*, 2482. https://doi.org/10.3390/agronomy12102482

Academic Editors: Zhanyou Xu, Reka Howard and Lizhi Wang

Received: 22 September 2022
Accepted: 10 October 2022
Published: 12 October 2022

1. Introduction

The sweet cherry is a very old cultivated tree, with a history of more than 2000 years of artificial cultivation, and economic cultivation starting in the 16th century. It has developed into a worldwide fruit tree, and according to FAO data, sweet cherries were grown in 70 countries and regions worldwide in 2014. The sweet cherry grows in different locations, with small and tender individual fruits that are easily broken, and is currently usually picked by hand. Sweet cherries have a short ripening period and need to be picked in a short period, which can lead to fruit decay if it takes too long. With continuous updates in the field of agricultural machinery, sweet cherry picking is gradually becoming fully mechanized. In order to reduce the fragmentation rate during cherry picking and improve the picking efficiency, the design of a target detection method with high real-time detection and high detection accuracy has become a hot and difficult problem of widespread concern in the current academic and industrial fields.

Early studies on sweet cherry fruit detection focused on hand-extracted features based on the shape, texture, and color of fruit images. Rabby et al. [1] extracted the color and shape features of fruits and successfully classified apples and oranges using a Modified Canny Edge Detection (MCED) algorithm. The results showed that the edge detection accuracy of the improved method was better than that of the traditional method, especially in the presence of more noise or dark light conditions. Lu et al. [2] creatively proposed a method to detect immature citrus fruits using only texture and intensity distributions with 82.3% accuracy. However, such methods rely on manual feature extraction, which is a time-consuming and complex process.

With the gradual application of deep learning technology in agriculture, target detection based on deep learning has become a hot direction for intelligent fruit picking. The two main popular algorithms in target detection algorithms are two-stage and one-stage [3]. Two-stage refers to the detection algorithm executed in two steps: first, a candidate frame is selected, and then classification and regression are performed, such as in the Region-CNN (R-CNN) family. Gené-Mola et al. [4] used Fast Region-based Convolutional Network (Faster R-CNN) [5] for apple and kiwi fruit detection in a complex orchard context, achieving 80.8% and 88.4% detection accuracy, respectively. Yu et al. [6] introduced Mask for improving the performance of deep learning in a strawberry picking robot fruit detection Region-based Convolutional Neural Network (Mask-RCNN) [7] to achieve visual localization of strawberry picking points. In general, the R-CNN series has relatively high detection accuracy, but is computationally intensive and runs slowly. Both the earlier and current studies described above are based on RGB images, and thus possess superiorities such as high resolution, ease of use, and low cost.

One-stage refers to direct regression without acquiring candidate regions separately [8], such as the You Only Look Once (YOLO) series [9–13]. The YOLO series significantly improves the speed of model running inference while keeping the detection accuracy largely unchanged and is better able to meet the demand for real-time detection. Among them, YOLOV3 [11] uses Darknet53, which draws on the idea of the residual network [14] as the backbone network, while improving single-label classification into multi-label classification and introducing multi-scale features for object detection by drawing on the Feature Pyramid Network (FPN) [15], which improves the detection accuracy and small object detection. YOLOV4 [12] changed the backbone network to CSPDarknet53 based on the YOLOV3 network structure, and introduced the Path Aggregation Network (PANet) [16] and Spatial Pyramid Pooling (SPP) layer [17] into the neck network. Among them, the structure of the Cross Stage Partial Network (CSPNet) [18] enables richer gradient combinations while improving accuracy and inference speed. Gai et al. [19] proposed an improved YOLOV4 [12] deep learning algorithm to detect cherry fruits by replacing the prior box in the YOLOV4 model with a suitable cherry fruit shape, using DenseNet [20] to replace the Cross Stage Partial Network (CSPDarknet53) [11] structure, and modifying the loss function to Leaky Rectified Linear Unit (Leaky ReLU) [21], but the dataset of this study was not open-source and the model still had a slight misdetection problem for ripe, semi-ripe, and immature cherry fruits.

The YOLOX network [13] combines the advantages of the YOLO series of networks and can significantly improve the inference speed of the model while keeping the detection accuracy basically unchanged to achieve the demand of real-time detection. The backbone network part of the YOLOX model mainly adopts the residual structure [14] for feature extraction and feature fusion, and the neck network part mainly uses the Feature Pyramid Network (FPN) [15] layer for feature fusion of feature maps in three dimensions. How to focus on the relationship between the three dimensions at the same time and automatically learn the important dimensions of different channel features and different spatial features is a problem worth studying [22]. The Convolutional Block Attention Module (CBAM) [22] is a simple and effective attention module proposed by Woo et al. for feedforward convolutional neural networks, which focuses on both spatial focus and channel focus and has been widely used in networks other than YOLOX networks [23,24]. In addition, the YOLOX model uses Intersection over Union (IoU) [25] as the bounding box loss function, and the IoU value is 0 when the prediction box and the real box do not intersect, which leads to a large range of loss functions without gradients. Therefore, the design of the bounding box loss function is also one of the important directions of current research. Yao et al. [26] improved the regression accuracy of the bounding box to some extent after modifying the loss function of YOLOV5 to Complete-IoU (CIoU) [27], but CIoU sometimes prevents the model from optimizing the similarity effectively. For this problem, Zhang et al. [28] proposed Efficient IoU (EIoU) by splitting the influence factor of the aspect ratio of the prediction frame and the real frame based on the penalty term of CIoU.

In view of these limitations, the main objective of this study was to achieve fast, accurate, and non-destructive detection of sweet cherries and their ripeness. For such a purpose, we collected and manually labeled sweet cherries in three categories: immature, semi-ripe, and ripe, and based on this, we proposed a YOLOX-based algorithm for real-time sweet cherry ripeness detection in a natural environment, YOLOX-EIoU-CBAM. This method embedded the Convolutional Block Attention Module (CBAM) between the backbone and the neck of the YOLOX model to improve the attention capability of the model for different channels and spaces; Efficient IoU (EIoU) loss was used to replace the original bounding box loss function of the YOLOX model to make the regression of the prediction box more accurate. Finally, the feasibility and reliability of this paper's method were verified on the SweetCherry dataset.

The subsequent sections are structured as follows: Section 2 describes SweetCherry, a dataset for sweet cherry fruit detection and its ripeness classification, and the details of the YOLOX-EIoU-CBAM target detection algorithm proposed in this study; Section 3 evaluates the performance of the YOLOX-EIoU-CBAM network through experiments; Section 4 shows the discussion results; Section 5 summarizes the work of this study, and points out the shortcomings and prospects of this study.

2. Materials and Methods

2.1. Study Area and Plant Material

Hanyuan County, Ya'an City, Sichuan Province, is one of the areas where sweet cherries are grown in China, and has the reputation of being the "hometown of sweet cherries in China". The planting area of sweet cherries has reached more than 42 square kilometers, with a production of 20,000 tons, and more than 30 varieties of sweet cherries have been introduced and promoted from early to late maturity. Hanyuan has become the largest cultivated area in Sichuan Province and even in the southwest region, and has been listed by the Sichuan Provincial Department of Agriculture as "Sichuan Province advantageous characteristic benefit agriculture sweet cherry base". The map of the location of the sweet cherry orchard is shown in Figure 1.

Sweet cherries are affected by altitude differences and the ripening cycle can last from early May to July. The size of sweet cherry fruit is relatively small, with a grain diameter of about 26–30 mm and few picture-pixel features. This means that the variation in cherry color and volume poses a challenge for automatic recognition by intelligent picking robots [19]. As shown in Table 1, immature sweet cherry fruits are greenish-white and green in color and have a growth cycle from approximately late March to mid-April. Semi-ripe sweet cherry fruits are yellow-green and light red in color and have a growth cycle from approximately mid-April to May. Ripe sweet cherries are red and purplish-red in color and grow from May to June. Most of these immature sweet cherries are greenish in color, similar to the color of green leaves, and there is some difficulty in target detection. Secondly, during the earliest batch of fruit harvesting, we could observe not only semi-ripe cherries but also ripe cherries on one branch, all of which made the picking more difficult for the robot. In addition, the cherry fruits may also overlap each other or be shaded by the trunk or leaves.

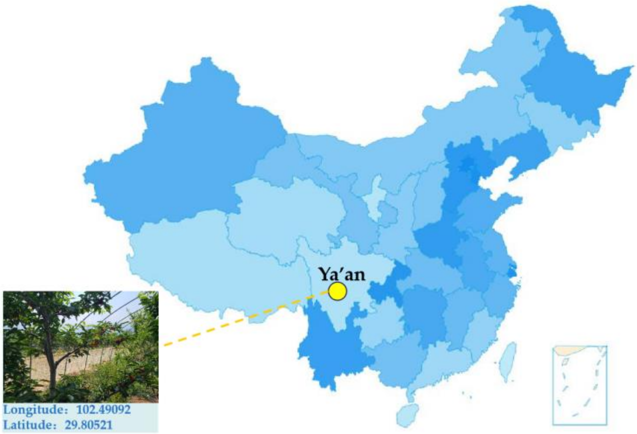
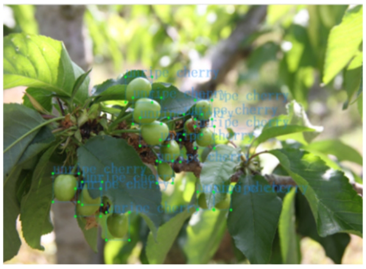
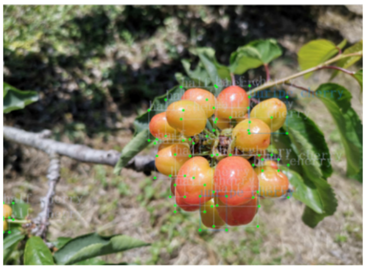
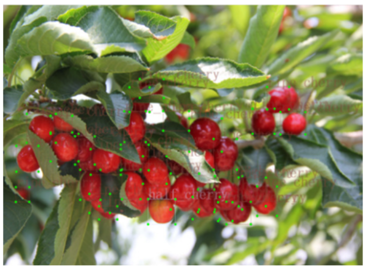


Figure 1. The location of sweet cherry orchard in Hanyuan County, Ya'an, as marked on the map.

Table 1. Information (Category, Label name, Number of pictures, Number of annotations, Number of cherries, Growth cycle, Sample examples) about the SweetCherry dataset.

Category	Label Name	Number of Pictures	Number of Annotations	Number of Cherries	Growth Cycle	Sample Examples
Immature	unripe_cherry	800	800	12979	late March to mid-April	
Semi-ripe	half_cherry	800	800	4395	mid-April to May	
Ripe	ripe_cherry	800	800	18694	May to June	

2.2. Image Acquisition and Classification

We collected high-resolution images of immature, semi-ripe, and ripe sweet cherries of the Red Lantern variety from April 2021 to June 2021 at sweet cherry growing sites in Hanyuan County. To improve the robustness and generalization of the network, a 5184 × 3456-pixel Canon EOS60D SLR camera, a 12-megapixel Apple phone, and a 12-megapixel Samsung phone were used to capture the images. All images were taken in natural light, while some images were taken of sweet cherries in complex environments such as light changes, fruit shading, fruit overlap, fruit close up, and plant far away, as shown in Figure 2. The shooting time was 10–12 a.m. and 1–4 p.m. Ultimately, 800 images of each of the three types of sweet cherries were collected, totaling 2400 images with a maximum resolution of 5184 × 3456 pixels, and stored in JPG format.

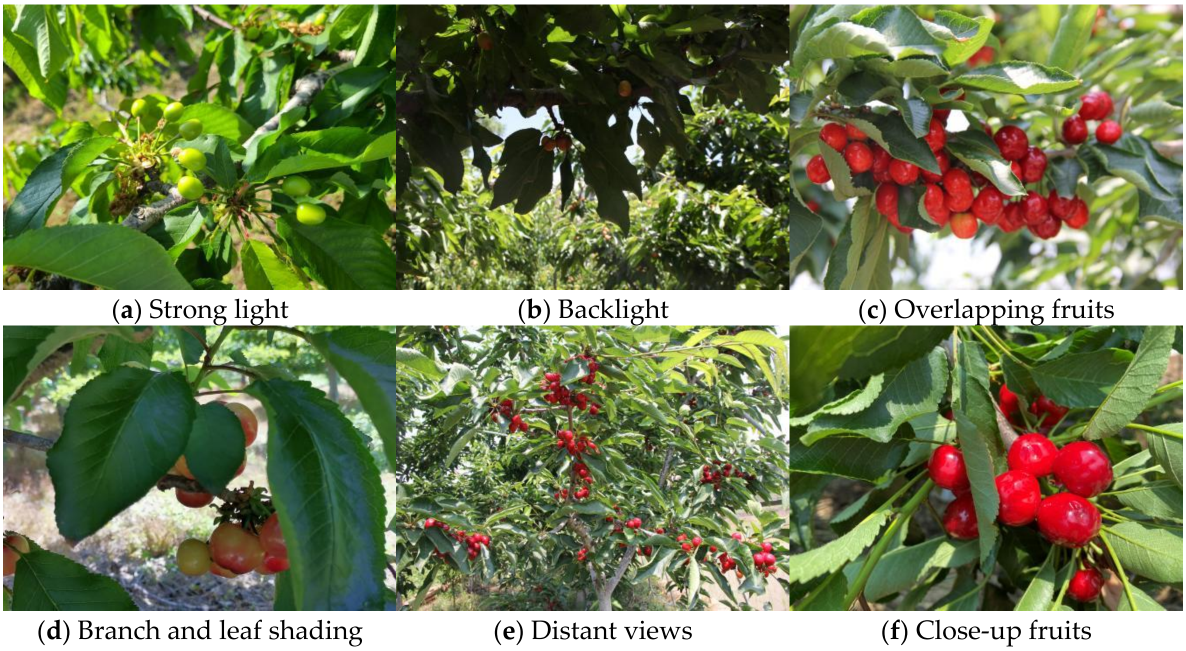

(**a**) Strong light
(**b**) Backlight
(**c**) Overlapping fruits
(**d**) Branch and leaf shading
(**e**) Distant views
(**f**) Close-up fruits

Figure 2. Part of the SweetCherry dataset. (**a**) Strong light. (**b**) Backlight. (**c**) Overlapping fruits. (**d**) Branch and leaf shading. (**e**) Distant views. (**f**) Close-up fruits.

Since sweet cherry fruits at different stages of ripeness have different color characteristics (see Section 2.1 for details), inspired by Gai et al. [19], we manually annotated 2400 SweetCherry datasets using fruit color as the basis for determining sweet cherry ripeness categories. This is an exercise that relies on visual senses and to avoid the influence of subjective bias, and the annotation of the dataset was carried out by only one person under an agronomist and a sweet cherry grower. Specifically, this study classified the ripeness categories of cherry fruit into ripe cherry, semi-ripe cherry, and immature cherry, and used the image annotation software LabelImg to label the SweetCherry dataset as ripe_cherry, half_cherry, and unripe_cherry, as shown in Table 1. The annotation file is in "XML" format, with a total of 2400 annotated images and 36068 sweet cherry fruit labels, including 18649 samples of ripe cherry, 4395 samples of semi-ripe cherry, and 12979 samples of immature cherry. The SweetCherry dataset is open-source at https://www.kaggle.com/datasets/jiangxueqin/sweetcherry (accessed on 9 September 2022).

2.3. *Improved YOLOX Network*

YOLOX [13] used Darknet-53 [11] as the backbone network, and used the Spatial Pyramid Pooling (SPP) [17] layer used by YOLOV4 and YOLOV5, greatly reduced the number of parameters with the Anchor-free idea, and solved the optimal transmission problem with the SimOTA dynamic matching positive sample algorithm. In addition, YOLOX has made effective improvements to YOLOV5 by creating different model structures such as YOLOX-S, YOLOX-M, YOLOX-L, and YOLOX-X. After considering the object of study, as well as the detection accuracy and lightweight requirements of the network, we chose to use the YOLOX-S network as the initial detection model.

2.3.1. The YOLOX-EIoU-CBAM Network

First, we modified the bounding box loss function IoU of the base YOLOX network to Efficient IoU (EIoU) [27] to improve the prediction accuracy of the model, which was named

YOLOX-EIoU. Second, we added a mechanism to focus on spatially important features and important channel features in YOLOX-S, which was named YOLOX-CBAM. As can be seen from Figure 3, we embedded the Convolutional Block Attention Module (CBAM) between the backbone network and the neck network of the initial YOLOX, with the embedding positions on the branch of the second Cross Stage Partial_1 (CSP_1) block leading to the neck network, the branch of the third CSP_1 block leading to the neck network, and the first Cross Stage Partial_2 (CSP_2) block (the structure of CSP_1 and CSP_2 blocks is detailed in the lower left corner of Figure 3) leading to the branch of the neck network.

We named the model replacing the bounding box loss function and resulting from adding the CBAM to YOLOX as YOLOX-EIoU-CBAM. Structurally, the CBAM consists of space and channel attention, both of which mainly use global average pooling to abstract features to a series of point attention weights, and then establish associations of these weights and attach them to the original space or channel features (see Section 2.3.4). In terms of extraction location, extraction at the beginning of the network is too large for the spatial feature map and too small for the number of channels. The extracted channel weights are too generalized without falling to some specific features, the extracted spatial attention is not generalized enough due to the small number of channels, and the spatial attention is sensitive and difficult to learn, which is more likely to cause negative effects. Layers that are too far back with too many channels easily cause overfitting; feature maps are too small to use convolution, and improper operation will instead introduce a large proportion of non-pixel information.

The YOLOX-EIoU-CBAM network consists of five main parts.

Input: Mosaic data enhancement (see Section 2.3.2).

Backbone: The backbone side mainly includes Focus, CSP_1, and Spatial Pyramid Pooling (SPP) structures, and the SiLU function is used in the activation function. Among them, there are two branches of the CSP_1 module, the input features will enter into two branches separately and perform different operations, then the features inputted from two branches will be Concat and the results will be sent to the Convolution + Batch Normalization + SiLU (CBS) block.

Between the backbone network and the neck network: Three Convolutional Block Attention Modules (CBAMs).

Neck: The neck end consists of Feature Pyramid Network (FPN), Path Aggregation Network (PANet), and Cross Stage Partial_2 (CSP_2) structures, which significantly improve the extraction of large cherry features, and the activation function also uses the SiLU function. One of the FPN structures can effectively fuse the feature map, which takes the high-level feature information in a top-down manner and passes the fusion through upsampling to obtain the feature map for making predictions. The PANet takes the output features of the CBAM as input and fuses the deep features with the shallow features after upsampling, and the fused shallow features are then downsampled and then fused with the deep features to further enhance the feature extraction capability of the network.

Head: Decoupled Head is used for the detection head, and Anchor-free, Multi-positives, and SimOTA algorithms are introduced. Decoupled Head can improve the model accuracy and speed up the convergence of the network, and Multi-positives can reduce the neglect of other high-quality predictions.

Compared with the base YOLOX model, YOLOX-EIoU-CBAM can extract features better and obtain better recognition accuracy, which is verified in the experiments in Section 3.

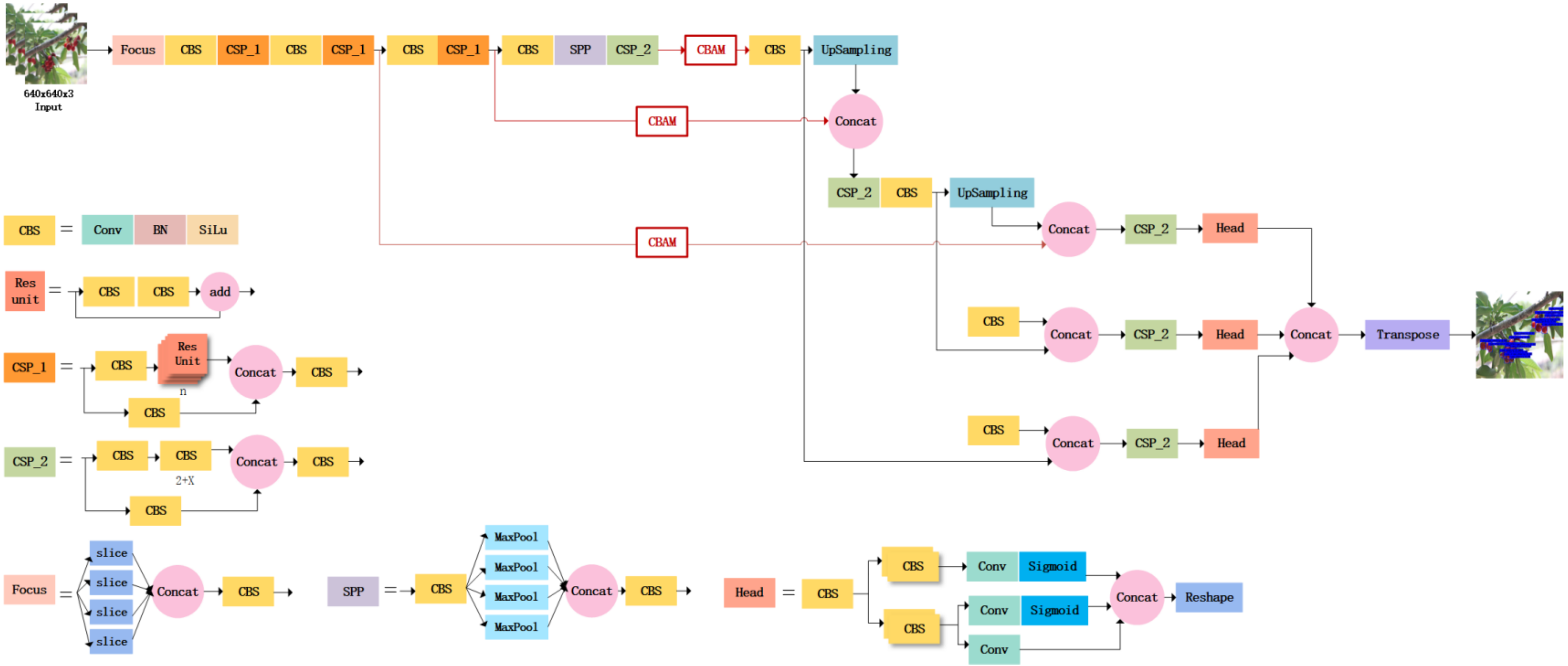

640x640x3
Input
Focus
CBS
CSP_1
CBS
CSP_1
CBS
CSP_1
CBS
SPP
CSP_2
CBAM
CBS
UpSampling
CBAM
Concat
CSP_2
CBS
UpSampling
CBAM
Concat
CSP_2
Head
CBS
Concat
CSP_2
Head
Concat
Transpose
CBS
Concat
CSP_2
Head
CBS = Conv BN SiLu
Res unit = CBS CBS add
CSP_1 = CBS Res Unit n Concat CBS CBS
CSP_2 = CBS CBS 2*X Concat CBS CBS
Focus = slice slice slice slice Concat CBS
SPP = CBS MaxPool MaxPool MaxPool MaxPool Concat CBS
Head = CBS CBS Conv Sigmoid CBS Conv Sigmoid Conv Concat Reshape

Figure 3. The architecture of YOLOX-EIoU-CBAM.

2.3.2. Data Augmentation

The detection algorithm used in this paper is based on a modified model of YOLOX. For the benchmark model YOLOX, it uses the Mosaic data enhancement [19].

For the Mosaic data enhancement method, the object detection context can be enriched and the dataset can be effectively augmented. Since four images from the Batch Normalization (BN) layer are fed into each training session, the batch size value need not be too large.

The Mosaic data enhancement is shown in Figure 4. The detailed implementation process is as follows:

- read four random images from the dataset at a time;
- perform operations such as flip (flip the original image left and right), zoom (scale the size of the original image), and gamut change (change the hue, brightness, saturation of the original image), etc.;
- make combinations of pictures and boxes.

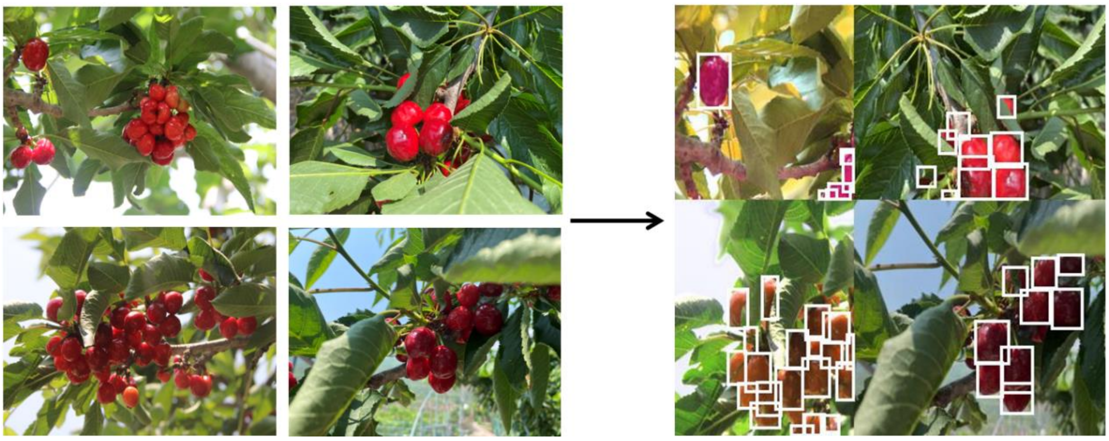


Figure 4. Mosaic image enhancement: Random selection of four cherry images, gamut change, flip, zoom, combined images, combo box.

2.3.3. The Efficient IoU Loss Function

The principle of the Efficient IoU (EIoU) loss [28] bounding box regression loss function is as follows:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2}. \quad (1)$$

As can be seen from Equation (1), the EIoU loss function consists of three components: L_{IoU} represents the loss of overlap between the prediction frame and the real frame, L_{dis} represents the loss of center distance between the prediction frame and the real frame, L_{asp} represents the loss of width and height of the predicted box and the real box. c_w and c_h are the width and height of the smallest enclosing box that covers both boxes. In fact, the EIoU continues the Complete-IoU (CIoU) [27] approach, but it is more in line with the regression mechanism of the bounding box, leading to faster convergence of the model and improved regression accuracy of the prediction frame.

2.3.4. Convolutional Block Attention Module

As shown in Figure 5 below, the Convolutional Block Attention Module (CBAM) [22] consists of two independent sub-modules: the Channel Attention Module (CAM) and the Spatial Attention Module (SAM), which pay attention to the relationships between channels and spaces, respectively. It automatically learns the importance of different channel features and different spatial features. In addition, as a lightweight general-purpose module, the CBAM can be easily integrated into Convolutional Neural Network (CNN) architecture without additional overhead.

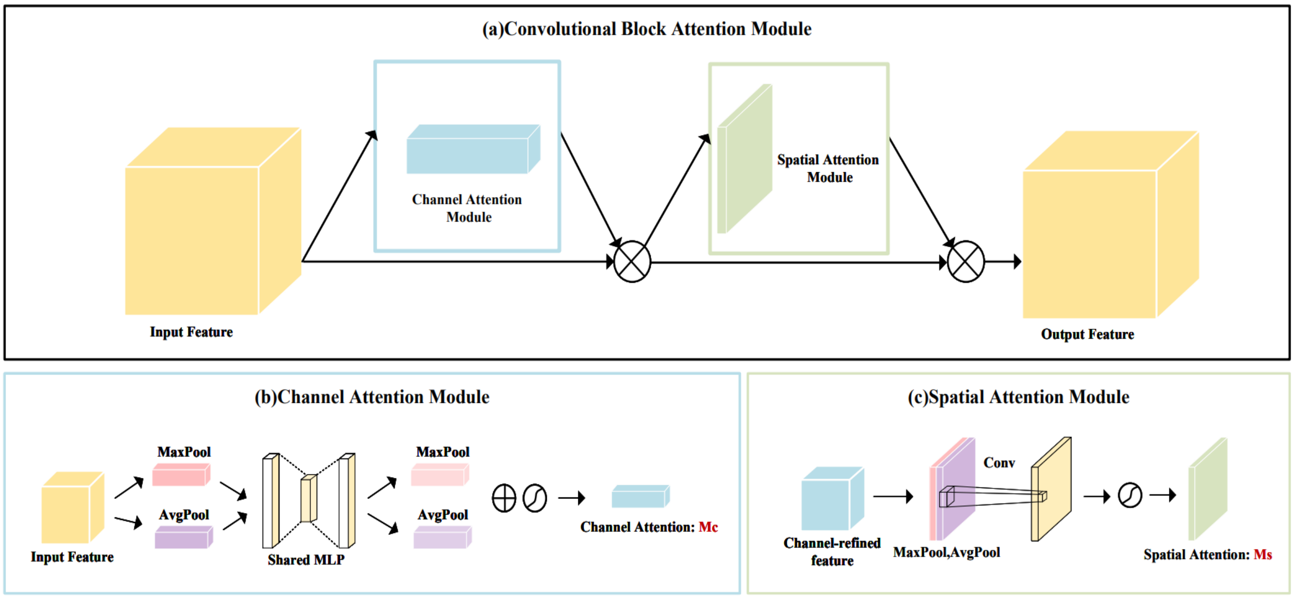


Figure 5. Convolutional block attention module architecture. (**a**) The overall architecture of CBAM. (**b**) The channel attention module. (**c**) The spatial attention module.

2.4. Training Environment and Evaluation Indicators

2.4.1. Transfer Learning

Transfer learning [29] is a popular method in computing that helps us to build more accurate models in less time. In this experiment, we chose transfer learning to initialize the parameters of the YOLOX network to give the model the ability to learn quickly and reduce overfitting to a certain extent. This allows the model to have significant generalization capabilities even in complex environments and to improve it in this paper, a widely used ImageNet dataset [30] was chosen to pretrain the network and obtain the initialization weights.

2.4.2. Cosine Annealing

The cosine annealing learning rate [31] enables the learning rate to be adjusted during the training of the model. As the epoch increases, the learning rate decreases rapidly until the model finds a local optimum and saves the model at that point. After the model is saved, the learning rate reverts to a larger value, escaping the current local optimum and finding a new local optimum. The process is repeated, adjusting the learning rate according to the cycle until the training is complete. The principle of the cosine annealing algorithm is shown in Equation (2).

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})(1 + cos(\frac{T_{cur}}{T_i}\pi)), \tag{2}$$

where η_{max}, η_{min} denote the maximum and minimum learning rates, respectively. T_{cur} refers to the number of iterations after restart and T_i denotes the number of iterations in round i.

2.4.3. Experimental Environment

The framework for deep learning was PyTorch 1.8.1. The experiments were performed using Windows 10 with an Intel Xeon Gold 5218 CPU with a base frequency of

2.30 GHz, 128 GB of RAM, NVIDIA Quadro RTX 5000 graphics, and Compute Unified Device Architecture (CUDA) 10.2.

In this study, the original YOLOX and the modified YOLOX-EIoU-CBAM were trained separately. The hyperparameters were set as follows: the input image size of the model was set to 640 × 640, the maximum learning rate of the model was set to 0.01 and the momentum of the learning rate was set to 0.937, the IoU threshold for mean Average Precision (mAP) was set to 0.5. To improve the training speed, the batch size of the freezing and thawing phases of the backbone network was set to 16 and 8 times, respectively, and the number of iterations was set to 200.

2.4.4. Evaluation Indicators

The evaluation metrics of the model are: precision, recall, mean Average Precision (mAP), and F-score. The expressions are as follows:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{4}$$

$$\text{mAP} = \frac{\sum_1^N AP}{N} = \frac{\sum_1^N \int_0^1 P(R)dR}{N}, \tag{5}$$

$$\text{F-score} = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{6}$$

where True Positive (*TP*) is the number of positively categorized samples, False Positive (*FP*) is the number of negatively categorized samples, and False Negative (*FN*) is the number of positive samples incorrectly categorized. *AP* is the area under the accuracy recall curve (P-R curve) and represents the average accuracy. mAP is the average of *AP* for different categories. *N* is the number of categories of samples tested. In this experiment, the category of sweet cherries needed to be detected, so *N* is 3. F-score can balance the impact of precision and recall.

3. Results

In this section, we evaluate the performance of the YOLOX-EIoU-CBAM using the SweetCherry dataset. Firstly, the SweetCherry dataset was divided into a training set, validation set, and test set according to 8:1:1. Secondly, SSD300, Faster R-CNN, YOLOX, YOLOX-EIoU, YOLOX-CBAM, and YOLOX-EIoU-CBAM networks were used to conduct comparative analysis studies.

3.1. Multi-Indicator Performance Evaluation

Table 2 shows the mean Average Precision (mAP), precision, recall, model size, and inference time of a single image for the six networks on the SweetCherry dataset. From Table 2, it can be seen that the YOLOX-EIoU-CBAM model's mAP, precision, recall, and F-score were almost all significantly better than the other five networks (although SSD300 achieves the best performance on precision). Among them, YOLOX-EIoU-CBAM improves mAP by 4.12%, recall by 4.6%, and F-score by 2.34%, compared to the base YOLOX model. This is because YOLOX-EIoU-CBAM used a more accurate bounding box loss function, EIoU, and added a mechanism to the YOLOX network that focuses on both channel and spatial information of the image, resulting in better extraction of image features and better detection results. When YOLOX-EIoU-CBAM was compared to the Faster R-CNN and SSD model, the mAP is improved by 19.89% and 41.69%, respectively, which proved the effectiveness and feasibility of the YOLOX-EIoU-CBAM network.

Table 2. Multi-metric performance evaluation (mAP, precision, recall, F-score, model size, and inference time) of six models (SSD300, Faster R-CNN, YOLOX, YOLOX-EIoU, YOLOX-CBAM, and YOLOX-EIoU-CBAM) on 240 test samples.

Models	mAP (%)	Precision (%)	Recall (%)	F-Score (%)	Model Size (MB)	Inference Time (S)
SSD300	39.41	86.52	26.29	31.33	100.27	1.41
Faster R-CNN	61.21	43.74	71.35	54.33	522.99	2.21
YOLOX	76.98	86.50	68.70	76.33	34.10	0.59
YOLOX-EIoU	78.20	83.89	70.80	76.33	34.10	0.60
YOLOX-CBAM	79.16	84.52	71.25	77.67	34.75	0.62
YOLOX-EIoU-CBAM	81.10	84.96	73.30	78.67	34.87	0.64

Table 2 also shows the experimental results of YOLOX versus YOLOX-EIoU, with the latter outperforming the former, proving that EIoU is more suitable for the detection of the ripeness category of sweet cherry fruits than the IoU bounding box loss function originally used by YOLOX. In addition, the experimental results of YOLOX and YOLOX-CBAM in Table 2 also showed us that embedding the Convolutional Block Attention Module (CBAM) with excellent performance between the backbone and neck networks of YOLOX can indeed significantly improve the detection accuracy of the existing method.

Finally, the model size and single-image inference time of the proposed method in this paper are comparable to the basic YOLOX, but significantly better than SSD300 and Faster R-CNN, ensuring the real-time detection. Overall, Table 2 effectively verifies that the overall performance of the YOLOX-EIoU-CBAM network outperforms Faster R-CNN, SSD300, YOLOX, YOLOX-EIoU, and YOLOX-CBAM on the SweetCherry dataset, and is a more suitable algorithm for sweet cherry fruit category detection.

3.2. Performance Evaluation of Different Maturity Categories

To compare the detection performance of this paper's model for different categories of sweet cherries in detail, the Average Precision (AP) and F-score of the six network models for different categories of sweet cherries are given in Table 3. It can be seen that the proposed model significantly outperforms the other five networks in terms of AP and F-score for different ripeness categories of sweet cherries. Among them, the YOLOX-EIoU-CBAM model improved the AP values by 2.61%, 5.02%, and 4.74%, and F-score by 1%, 1%, and 5%, respectively, on the unripe, semi-ripe, and ripe categories relative to the base YOLOX model. This is because YOLOX-EIoU-CBAM uses a more accurate bounding box EIoU loss function and can focus on effective information from both spatial and channel dimensions, which leads to better extraction of image features and better detection results.

Table 3. Multi-metric performance evaluation (AP and F-score) of six models (SSD300, Faster R-CNN, YOLOX, YOLOX-EIoU, YOLOX-CBAM, and YOLOX-EIoU-CBAM) on 240 test samples at different stages of maturity (immature, semi-ripe, and ripe).

Models	Average Precision (AP) (%)			F-Score (%)		
	Immature	Semi-Ripe	Ripe	Immature	Semi-Ripe	Ripe
SSD300	44.06	42.86	42.86	33	40	21
Faster R-CNN	70.21	70.21	50.71	57	57	49
YOLOX	81.57	71.64	77.73	81	73	75
YOLOX-EIoU	82.25	71.46	80.89	80	71	78
YOLOX-CBAM	82.87	73.40	81.22	81	73	79
YOLOX-EIoU-CBAM	84.18	76.66	82.47	82	74	80

YOLOX-EIoU showed a certain degree of improvement in AP and F-score on the immature and ripe sweet cherry relative to the base YOLOX model, but a slight decrease in the detection of semi-ripe sweet cherries was observed. We analyzed that this may be due to the detection difficulty of this class of data samples itself, as its fruit color features are between the other two classes, and EIoU ignored part of the feature information, affecting the classification in the fast convergence process and thus affecting the detection accuracy of semi-ripe sweet cherries. This problem was effectively solved by our embedding the Convolutional Block Attention Module (CBAM), and the AP and F-score of YOLOX-EIoU-CBAM are improved by 5.2% and 3%, respectively, on semi-ripe sweet cherries compared to YOLOX-EIoU. It can also be seen that after embedding the CBAM between the backbone and neck networks of YOLOX alone, YOLOX-CBAM showed significant improvements in AP and F-score for sweet cherries in all three ripeness categories, further verifying that the CBAM is more likely to focus on important features of small target samples and effectively improves the detection effectiveness of the model. Overall, Table 3 effectively verifies that the YOLOX-EIoU-CBAM network outperforms Faster R-CNN, SSD300, YOLOX, YOLOX-EIoU, and YOLOX-CBAM for the detection of sweet cherry samples in the three ripeness categories, and is a more suitable target detection algorithm for the classification of sweet cherries and their ripeness.

Figure 6 shows the confusion matrix for the three ripeness categories of sweet cherries, where each row of the matrix represents the actual category, each column represents the predicted category, and the diagonal line represents the number of each category that was correctly predicted. From this, we can see that the number of sweet cherries predicted is higher for the ripe and unripe categories, while the number of semi-ripe categories predicted accurately is relatively small. The confusion matrix showed that, firstly, more cherries in the immature category were judged to be ripe, and by looking at the images where such false inferences occurred, we found that most of the targets were heavily shaded or in dark light conditions with "black" fruits and were therefore misjudged as ripe. Secondly, more cherries in the semi-ripe category were incorrectly predicted as unripe, which we suggest may be due to the similarity of color characteristics of unripe sweet cherries and semi-ripe sweet cherries, which further increases the difficulty of recognition and classification under variable light. Finally, some ripe cherries were predicted as the semi-ripe category, which we suggest may also be because some light red ripe sweet cherries showed similar color and shape features to the slightly red semi-ripe cherries under bright light. Overall, the three ripeness categories of cherries were correctly predicted with relatively high accuracy, which showed that our model is suitable for detecting sweet cherries and their ripeness categories in complex environments.

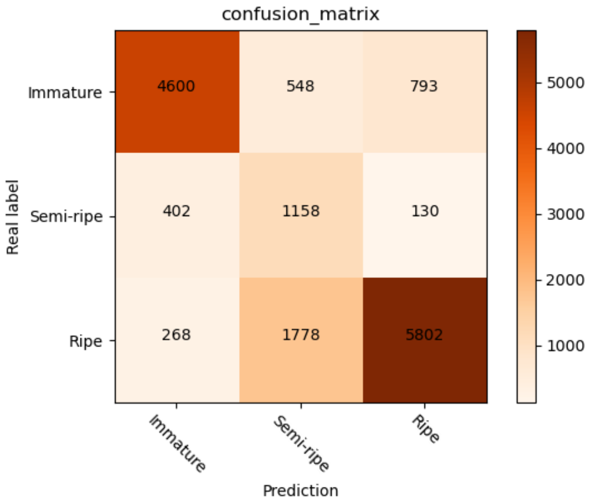
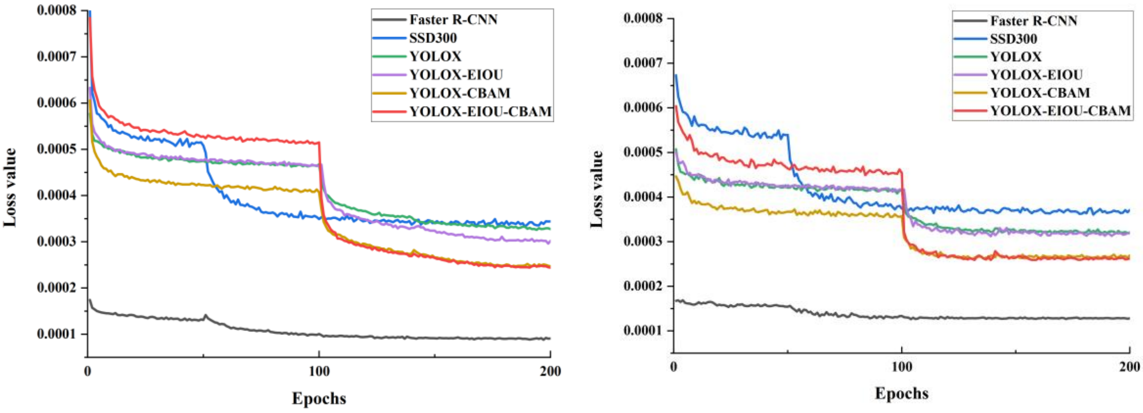


Figure 6. Confusion matrix of YOLOX-EIoU-CBAM network.

3.3. Loss Function Evaluation

Since the loss function can accurately reflect the convergence of the model during the training process of the network, in this paper, the loss function curves of six network models are compared and analyzed, as shown in Figure 7. Figure 7a shows the loss curve of the training set, and Figure 7b shows the loss curve of the validation set. It can be seen that the training loss curve and validation loss curve of the improved YOLOX-EIoU-CBAM model both showed a decreasing trend as the number of iterations increases and the training and validation losses were close to each other when the model iterated 192 times. When the model reached 197 iterations, the training loss showed a slightly decreasing trend, while the validation loss remained stable, but overall, they were close to each other, indicating a slight degree of overfitting in the proposed method. In general, compared with Faster R-CNN, SSD, YOLOX, YOLOX-EIoU, YOLOX-CBAM, and the base YOLOX model, the YOLOX-EIoU-CBAM model had a faster and more accurate reduction in loss values during training and validation.

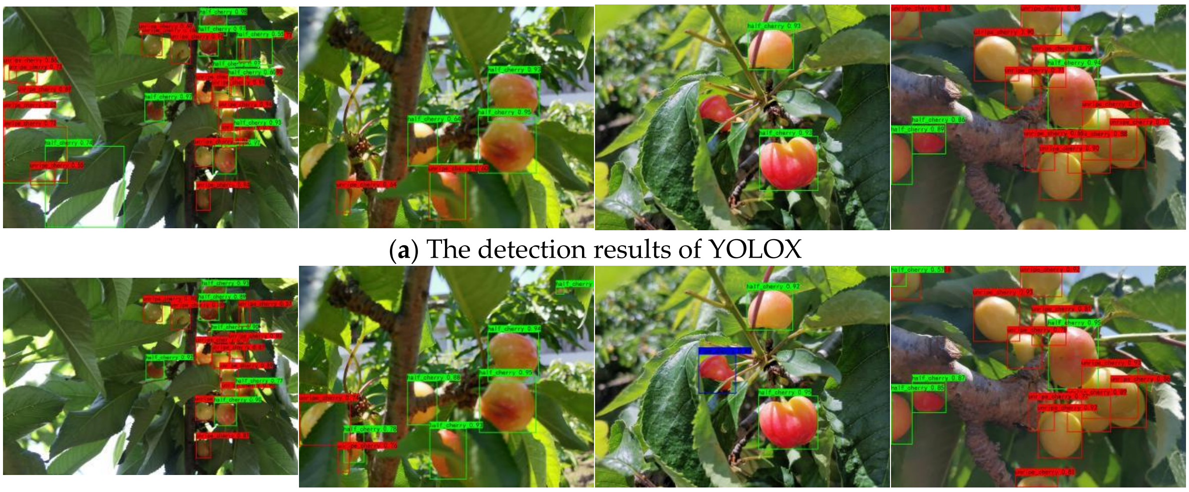


(**a**) Train loss (**b**) Validation loss

Figure 7. The training loss curve and the validation loss curve for six models (Faster R-CNN, SSD300, YOLOX, YOLOX-EIoU, YOLOX-CBAM, YOLOX-EIoU-CBAM). (**a**) Train loss curve. (**b**) Validation loss curve.

3.4. Image Detection Effect Demonstration

The results of the base YOLOX and the improved YOLOX-EIoU-CBAM for the detection of sweet cherry fruit are shown in Figure 8. As can be seen from the figure, the original YOLOX model had certain problems of missing and misdetection for sample pictures in complex orchard backgrounds. For example, in the first picture in Figure 8a below, the YOLOX model identified the leaves as fruits. In the second image in Figure 8a, the YOLOX model did not identify the sweet cherry ripeness category obscured by the leaves in the lower left corner of the image. In the third image of Figure 8a, the YOLOX model did not identify the fruit obscured by the leaves in the middle of the picture. In the last picture of Figure 8a, the YOLOX model did not accurately identify the overlapped fruit in the upper left corner of the picture and missed the fruit located at the lower edge position in the middle of the picture. In contrast, the improved YOLOX-EIoU-CBAM effectively improves the model's ability to detect the ripeness of sweet cherry fruits by modifying the bounding box loss function and embedding CBAM layers in the backbone and neck networks of the YOLO model (see the first and second pictures of Figure 8b), and is able to accurately detect fruits in the case of overlapping fruits (see the fourth picture of Figure 8b), and shows friendly detection ability for fruits in the corners of the pictures (see the fourth panel of Figure 8b).

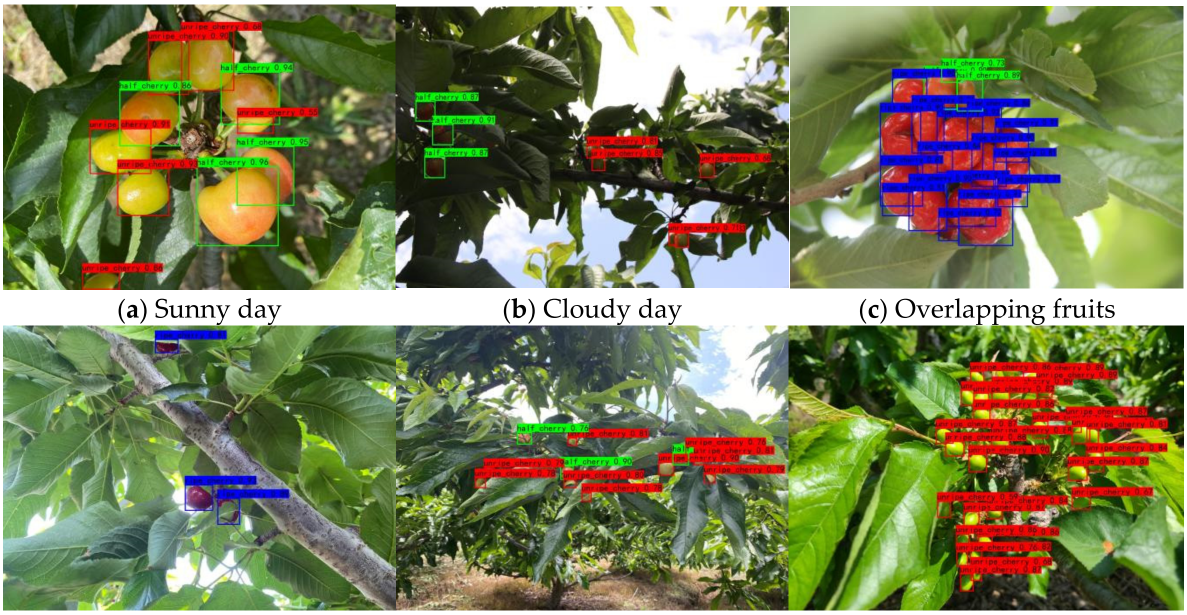


(**a**) The detection results of YOLOX

(**b**) The detection results of YOLOX-EIoU-CBAM

Figure 8. Comparison of the detection results of YOLOX and YOLOX-EIoU-CBAM. (**a**) The detection results of YOLOX. (**b**) The detection results of YOLOX-EIoU-CBAM.

Figure 9 shows the recognition effect of YOLOX and YOLOX-EIoU-CBAM models for sweet cherry ripeness categories under sunny days, cloudy days, branch and leaf shading, fruit overlapping, distant view, and similar color of green fruit and leaves, respectively. It can be seen that the YOLOX-EIoU-CBAM model can detect the sweet cherry category in complex environments in real time and accurately, and the detection time is only 0.64 s (list in Table 2), which can be applied to real agricultural production environments.



(**a**) Sunny day (**b**) Cloudy day (**c**) Overlapping fruits

(**d**) Branch and leaf shading (**e**) Distant views (**f**) Similar colors

Figure 9. Effectiveness of YOLOX-EIoU-CBAM for the detection of sweet cherry fruit ripeness categories in complex orchard environments ((**a**) sunny day. (**b**) Cloudy day. (**c**) Overlapping fruits. (**d**) Branch and leaf shading. (**e**) Distant views. (**f**) Similar colors).

4. Discussion

For the complex orchard environment, factors such as strong light, backlight, branch and leaf shading, fruit overlap, and distant view all affect the detection effect of the model [19]. To improve the model's ability to cope with the above problems, we specifically collected pictures of sweet cherries in complex environments such as light changes, fruit occlusion, fruit overlap, fruit close up, and plant distant view, and adjusted the model structure of YOLOX to improve the model's detection of sweet cherry fruits with complex backgrounds.

The current advanced deep learning target detection algorithm, YOLOX, has assembled a large number of tricks, but it is still prone to the problem of wrong and missed detection when dealing with some complex backgrounds, as shown in Figure 8a. In target detection, Bounding Box Regression (BBR) is a key step in determining the performance of object localization [32]. However, we found that the BBR loss function of YOLOX has two main drawbacks: (1) the Intersection over Union (IoU) loss function cannot effectively describe the target of BBR, which leads to slow convergence and inaccurate regression results [33]. (2) The IoU ignores the imbalance problem in BBR, and a large number of anchor frames with less overlap with the target frame plays the largest role in the optimization of BBR [27]. To mitigate the resulting adverse effects, we tried to directly replace the original loss function in the YOLOX detection network by using EIoU which had excellent convergence speed and focused on the problem of imbalance between difficult and easy samples. The experimental results showed that YOLOX-EIoU can achieve better results compared to YOLOX using IoU as the loss function, with an improvement of 1.22% and 2.1% in the mAP and recall metrics, respectively. Compared with IoU loss, EIoU can have significant advantages in both convergence speed and localization accuracy. The above results confirm the feasibility and reliability of using EIoU to directly replace the original loss function in the YOLOX detection network.

The neck network of YOLOX is constructed between the backbone and the head for bringing together different feature maps [12]. How to focus on the relationship between different features at the same time here and the importance of automatically learning different channel features and different spatial features is another problem to be solved in this study [24]. Therefore, in this study, we try to add the lightweight general-purpose module CBAM between the YOLOX backbone network and the neck network. The experimental results showed that YOLOX-CBAM can achieve better results compared to YOLOX which only focuses on channel features [13]. YOLOX-EIoU-CBAM can also achieve better results compared to YOLOX-EIoU. This is mainly because, given the intermediate feature map of the trunk output, the CBAM sequentially infers the attention map along two independent dimensions (channel and space) and then multiplies the attention map with the input feature map to perform adaptive feature optimization [23,24]. The above results confirm the feasibility and reliability of adding the lightweight CBAM to the YOLOX backbone network and the neck network. In addition, there is still room for improvement in the lightweighting of the YOLOX-EIoU-CBAM model.

5. Conclusions

In this paper, we proposed an algorithm based on improved YOLOX sweet cherry fruit ripeness detection in real time in a natural environment. Firstly, we manually photographed and labeled a dataset, SweetCherry, containing three categories of immature sweet cherries, semi-ripe sweet cherries, and ripe sweet cherries from April to June 2021 in a sweet cherry orchard in Hanyuan County, Sichuan Province. The dataset consisted of 2400 images and labeled files and 36,068 cherry samples. Secondly, in order to solve the problem of missed and false detection of YOLOX on the SweetCherry dataset, we made improvements to the YOLOX model: first, the IoU loss function was modified to EIoU to improve the regression accuracy of the prediction frame. Second, we added a Convolutional Block Attention Module (CBAM) between the backbone network and neck network of the YOLOX model to improve the sensitivity of the network to spatially important features and important channel features, to cope with the detection difficulties such as green fruit and leaf color similarity, color features of semi-ripe sweet cherries between immature and ripe sweet cherries, and optimize the complex orchard environment such as fruit overlap, fruit shading, distant view, cloudy sky, and strong light. Considering the same target detection task, the YOLOX-based real-time sweet cherry fruit ripeness detection algorithm in the natural environment can be extended to a variety of fruits, such as green dates and lychees.

Since this study mainly dealt with the detection and ripeness classification of one variety of sweet cherries, in the future, we will consider extending the method of this paper to other varieties of sweet cherries with different colors and different appearance characteristics. In addition, we will use more informative sensors for image acquisition, such as multi-spectral cameras. Additionally, RGB image processing techniques are considered for 3D modeling of sweet cherries.

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

Funding: This research was funded by research on intelligent monitoring and early warning technology for major rice pests and diseases of Sichuan Provincial Department of Science and Technology, grant number 2022NSFSC0172.

Data Availability Statement: The data in this study are available on request from the corresponding author.

Acknowledgments: Thanks to Jia Yao for his advice throughout the study. Thanks to Jingfan Chen and Yongxiu Li for their help in collecting the dataset. Thanks to all partners in AI Studio for their support.

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

References

1. Rabby, M.K.M.; Chowdhury, B.; Kim, J.H. A Modified Canny Edge Detection Algorithm for Fruit Detection Classification. In Proceedings of the ICECE 2018—10th International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, 20–22 December 2018; pp. 237–240. [CrossRef]
2. Lu, J.; Lee, W.S.; Gan, H.; Hu, X. Immature Citrus Fruit Detection Based on Local Binary Pattern Feature and Hierarchical Contour Analysis. *Biosyst. Eng.* **2018**, *171*, 78–90. [CrossRef]
3. Du, L.; Zhang, R.; Wang, X. Overview of Two-Stage Object Detection Algorithms. *J. Phys. Conf. Ser.* **2020**, *1544*, 12033. [CrossRef]
4. Gené-Mola, J.; Vilaplana, V.; Rosell-Polo, J.R.; Morros, J.R.; Ruiz-Hidalgo, J.; Gregorio, E. Multi-Modal Deep Learning for Fuji Apple Detection Using RGB-D Cameras and Their Radiometric Capabilities. *Comput. Electron. Agric.* **2019**, *162*, 689–698. [CrossRef]
5. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]
6. Yu, Y.; Zhang, K.; Yang, L.; Zhang, D. Fruit Detection for Strawberry Harvesting Robot in Non-Structural Environment Based on Mask-RCNN. *Comput. Electron. Agric.* **2019**, *163*, 104846. [CrossRef]
7. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *42*, 386–397. [CrossRef] [PubMed]
8. Bharati, P.; Pramanik, A. Deep Learning Techniques—R-CNN to Mask R-CNN: A Survey. *Adv. Intell. Syst. Comput.* **2020**, *999*, 657–668. [CrossRef]
9. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016, Las Vegas, NV, USA, 27–30 December 2016; pp. 779–788. [CrossRef]
10. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2016, Honolulu, Hawaii, 21–26 July 2017; pp. 6517–6525. [CrossRef]
11. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767. [CrossRef]
12. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934. [CrossRef]
13. Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. YOLOX: Exceeding YOLO Series in 2021. *arXiv* **2021**, arXiv:2107.08430. [CrossRef]
14. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016, Las Vegas, NV, USA, 27 June–30 June 2016; pp. 770–778. [CrossRef]
15. Zhao, Y.; Han, R.; Rao, Y. A New Feature Pyramid Network for Object Detection. In Proceedings of the 2019 International Conference on Virtual Reality and Intelligent Systems, ICVRIS 2019 2019, Jishou, China, 14–15 September 2019; pp. 428–431. [CrossRef]
16. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. PANet: Path Aggregation Network for Instance Segmentation. *arXiv* **2018**, arXiv:1803.01534.

17. 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. [CrossRef]
18. Wang, C.Y.; Mark Liao, H.Y.; 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 Computer Society Conference on Computer Vision and Pattern Recognition Workshops 2020, Seattle, WA, USA, 16–18 June 2020; pp. 1571–1580. [CrossRef]
19. Gai, R.; Chen, N.; Yuan, H. A Detection Algorithm for Cherry Fruits Based on the Improved YOLO-v4 Model. *Neural Comput. Appl.* **2021**, *2021*, 1–12. [CrossRef]
20. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
21. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models. In Proceedings of the ICML Workshop on Deep Learning for Audio, Speech, and Language Processing 2013, Atlanta, GA, USA, 16 June 2013; Volume 28.
22. Woo, S.; Park, J.; Lee, J.; 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.
23. Yang, L.; Yan, J.; Li, H.; Cao, X.; Ge, B.; Qi, Z.; Yan, X. Real-Time Classification of Invasive Plant Seeds Based on Improved YOLOv5 with Attention Mechanism. *Diversity* **2022**, *14*, 254. [CrossRef]
24. Gong, H.; Mu, T.; Li, Q.; Dai, H.; Li, C.; He, Z.; Wang, W.; Han, F.; Tuniyazi, A.; Li, H.; et al. Swin-Transformer-Enabled YOLOv5 with Attention Mechanism for Small Object Detection on Satellite Images. *Remote Sens.* **2022**, *14*, 861. [CrossRef]
25. Yu, J.; Jiang, Y.; Wang, Z.; Cao, Z.; Huang, T. UnitBox: An Advanced Object Detection Network. In Proceedings of the MM 2016—2016 ACM Multimedia Conference 2016, Amsterdam, The Netherlands, 15–19 October 2016; pp. 516–520. [CrossRef]
26. Yao, J.; Qi, J.; Zhang, J.; Shao, H.; Yang, J.; Li, X. A Real-Time Detection Algorithm for Kiwifruit Defects Based on Yolov5. *Electronics* **2021**, *10*, 7274. [CrossRef]
27. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. In Proceedings of the AAAI 2020—34th AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 12993–13000. [CrossRef]
28. Zhang, Y.F.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and Efficient IOU Loss for Accurate Bounding Box Regression. *Neurocomputing* **2022**, *506*, 146–157. [CrossRef]
29. Weiss, K.; Khoshgoftaar, T.M.; Wang, D.D. A Survey of Transfer Learning. *J. Big Data* **2016**, *3*, 1–40. [CrossRef]
30. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
31. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. In Proceedings of the 5th International Conference on Learning Representations, ICLR 2017—Conference Track Proceedings, Toulon, France, 24–26 April 2017. [CrossRef]
32. Qian, X.; Lin, S.; Cheng, G.; Yao, X.; Ren, H.; Wang, W. Object Detection in Remote Sensing Images Based on Improved Bounding Box Regression and Multi-Level Features Fusion. *Remote Sens.* **2020**, *12*, 143. [CrossRef]
33. Rezatofighi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2019, Long Beach, CA, USA, 15–20 June 2019; pp. 658–666. [CrossRef]