processes

MDPI

Article

# Detection of Cotton Seed Damage Based on Improved YOLOv5

**Zhicheng Liu [1,2], Long Wang [1,2,*], Zhiyuan Liu [1], Xufeng Wang [1,2], Can Hu [1,2] and Jianfei Xing [1,2]**

[1] College of Mechanical and Electrical Engineering, Tarim University, Alar 843300, China; 10757222241@stumail.taru.edu.cn (Z.L.); liuz030914@outlook.com (Z.L.); wxf@taru.edu.cn (X.W.); 120140004@taru.edu.cn (C.H.); 120200012@taru.edu.cn (J.X.)
[2] Modern Agricultural Engineering Key Laboratory at Universities of Education Department of Xinjiang Uygur Autonomous Region, Tarim University, Alar 843300, China
* Correspondence: 120140002@taru.edu.cn

**Abstract:** The quality of cotton seed is of great significance to the production of cotton in the cotton industry. In order to reduce the workload of the manual sorting of cotton seeds and improve the quality of cotton seed sorting, this paper proposed an image-detection method of cotton seed damage based on an improved YOLOv5 algorithm. Images of cotton seeds with different degrees of damage were collected in the same environment. Cotton seeds of three different damage degrees, namely, undamaged, slightly damaged, and seriously damaged, were selected as the research objects. Labeling software was used to mark the images of these cotton seeds and the marked images were input into the improved YOLOv5s detection algorithm for appearance-based damage identification. The algorithm added the lightweight upsampling operator CARAFE to the original YOLOv5s detection algorithm and also improved the loss function. The experimental results showed that the mAP_0.5 value of the improved algorithm reached 99.5% and the recall rate reached 99.3% when the uncoated cotton seeds were detected. When detecting coated cotton seeds, the mAP_0.5 value of the improved algorithm reached 99.2% and the recall rate reached 98.9%. Compared with the traditional appearance-based damage detection approach, the improved YOLOv5s proposed in this paper improved the recognition accuracy and processing speed, and exhibited a better adaptability and generalization ability. Therefore, the proposed method can provide a reference for the appearance detection of crop seeds.

**Keywords:** crop seed sorting; machine vision; deep learning; algorithm; appearance inspection

check for updates

**Citation:** Liu, Z.; Wang, L.; Liu, Z.; Wang, X.; Hu, C.; Xing, J. Detection of Cotton Seed Damage Based on Improved YOLOv5. *Processes* **2023**, *11*, 2682. https://doi.org/10.3390/pr11092682

Academic Editors: Wentao Ma, Siyuan Peng, Xinghua Liu and Jiandong Duan

Received: 9 August 2023
Revised: 2 September 2023
Accepted: 5 September 2023
Published: 7 September 2023



## 1. Introduction

Xinjiang is currently the main cotton-producing area in China due to its unique geographical environment and light conditions. In 2022, the total output of cotton in Xinjiang reached 5,391,000 tons [1], accounting for the majority of the national cotton output. Seeds are an indispensable part of agricultural production and play an irreplaceable role in the development of modern agriculture [2]. In particular, cotton seed is the basis of cotton production. Damage to cotton seeds is typically caused by the mechanical pressure generated during picking and processing. The damage will affect the quality of cotton seeds, and the damaged cotton seeds will consequently affect the emergence and growth of cotton, thus impacting the quality of cotton. Therefore, damaged cotton seeds should be identified and removed prior to sowing.

At present, the main methods of cotton appearance detection include both manual and machine detection [3]. Artificial observations and the evaluation of cotton seed surfaces depend on professional and technical personnel. Such manual methods are simple and easy, but the accuracy is greatly affected by human factors. In contrast, machine detection is more reliable and accurate. Thus, the application of dielectric sorting, color sorting, and machine-vision-recognition technology to the sorting of delinted cotton seeds is becoming increasingly more popular [4]. Kan et al. conducted an optimization test on the dielectric separation parameters of delinted cotton seeds. Through the optimization test, a mathematical model was established to express the relationship between the test factors and the cotton seed index. The influence of the influencing factors of dielectric separation on the cotton seed index was studied by mathematical model [5]. Yu Shuhua et al. developed a cotton seed color-sorting detection system through a series of analyses and designs to perform the cotton damage seed sorting of red seeds, black seeds, and intact seeds [6]. Peng et al. used the principle of the BackPropagation (BP) neural network to construct a vigor prediction model, with conductivity as the characteristic output vector parameter to measure the activity of cotton seeds. The authors obtained a high accuracy, facilitating research for subsequent seed grading [7]. Machine detection is generally realized by image feature extraction and computer-vision technology. For the identification of cotton seed damage, images of cotton seeds are collected with a digital camera, and the damage to the cotton seed surface is automatically detected and identified by computer processing and analyzing the image data [8]. Machine detection can not only improve the detection efficiency but also has higher accuracy than manual detection.

In recent years, with the rapid development of computer-vision technology, automatic detection methods based on image processing and machine learning have become increasingly advanced. Aiming at the problem of crop seed detection, researchers have proposed numerous solutions based on digital image processing, deep learning, and other methods. Among them, target-detection technology is a popular approach, particularly the detection algorithms based on You Only Look Once (YOLO) models that have achieved promising results [9]. Many scholars in China have also employed the YOLO algorithm to detect crop seeds. Wang Qiaohua and Gu Wei improved YOLOv4 to realize the damage identification of double-sided cotton seeds, achieving detection accuracy and recall rates of 95.33% and 96.31%, respectively, and a missed detection rate of 0 [10]. Li Haorui tested the purity of corn seeds by improving the YOLO model, and was able to reach a 99.5% accuracy in the detection of corn seed purity [11]. Fan Xiaofei adopted the improved YOLOv4 to detect the appearance quality of corn seeds, with the average F1 and mAP values of the improved model reaching 93.09% and 98.02%, respectively. The average detection time of each image was 1.85 s, and the number of model parameters was compressed to 20% of the original model [12]. Compared with the previous version of YOLO, YOLOv5 adopted a backbone network architecture based on Focus–Attention Transformation (FAT), which is a simple and effective feature pyramid network. Moreover, YOLOv5 introduced the channel attention (CA) and spatial attention (SA) mechanisms to further improve the quality of the feature representation and enhance the detection ability of various targets [13–15]. In addition, YOLOv5 optimized the anchor box and replaced it with a more flexible anchor-based box training strategy to obtain better detection results. YOLOv5 also used adaptive convolution instead of traditional fixed convolution, thereby reducing the number of parameters in the network [16–18]. To achieve better detection results, it is necessary to improve the YOLOv5 algorithm. At present, improved methods of YOLOv5 include the addition of a lightweight upsampling operator, an attention mechanism, a reverse residual module, and improving the loss function [19,20].

In this paper, we improved the YOLOv5 algorithm in terms of the lightweight feature and high performance of the model. The improved algorithm was used to detect the appearance damage of cotton seeds. The upsampling operator Content-Aware Reassembly of Features (CARAFE) was added to reduce the computational calculations, and thus, the model was lightweight. In addition, the improved loss function was used to accelerate the convergence speed and improve the performance to meet the needs of cotton seed damage detection in complex environments. Therefore, the purpose of this study was to use the improved YOLOv5 algorithm to quickly detect the damage of cotton seed appearance from the collected cotton seed images.

## 2. Materials and Methods

### 2.1. Material Selection

The quality of cotton depends on the quality of cotton seed, and cotton seed is the product of cotton production and processing. The process of cotton production and processing is as follows: after the mature cotton is harvested by manual or mechanical harvesting, it is processed by mechanical equipment such as cotton stripping machine to separate the cottonseed from the cotton fiber; next, the cotton seed is sorted to remove the damaged cotton seed and retain the high-quality cotton seed; after sorting, the cotton seed is coated; and finally, the coated cotton seed is packaged and stored. The working flow chart is shown in Figure 1.


**Figure 1.** Cotton-processing flow chart.

In the process of cotton-seed processing, the appearance damage detection of cotton seeds should be carried out twice. For the first time, the cotton seeds should be sorted after the delinting process of cotton seeds. The damaged cotton seeds caused by mechanical pressure in the process of picking and processing should be picked out by sorting, and the intact cotton seeds should be left for coating treatment. The intact seeds should be selected so that the coated seeds can have higher germination rate. Cotton-seed-coating treatment is the process of coating a layer of protective material on the surface of the seed. This material is usually composed of polymers, nutrients, fungicides, and other specific additives. Coating treatment can prevent and control pests and diseases, improve seed germination rate and survival rate, and increase crop yield and quality. The second cotton seed appearance damage detection is that after the cotton-seed-coating treatment, in order to obtain the seed with uniform coating, people often choose mechanical coating when coating, and mechanical coating can not avoid the damage to the cotton seed. Therefore, after the coating, the cotton seed should be sorted again to package and store the cotton seed, which further improves the quality of the cotton seed and facilitates the subsequent sowing operation. The cotton seed picture is shown in Figure 2. Therefore, in order to better select the damaged cotton seeds during the processing of cotton seeds and improve the quality of cotton seeds, this paper will select uncoated cotton seeds and coated cotton seeds for appearance detection.

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

**Figure 2.** Cotton seed images. (**a**) Uncoated cotton seed; (**b**) coated cotton seed.

*2.2. Data Acquisition and Processing*

Cotton seeds will be subjected to different external forces in the process of production and processing, which will cause different degrees of damage. According to the degree of damage, the appearance of cotton seeds can be divided into three types: intact, slightly damaged, severely damaged. The appearance of uncoated cotton seeds is shown in Figure 3, and the appearance of coated cotton seeds is shown in Figure 4. This study was divided into two groups. The first group of objects was Xinjiang cotton variety Xinluzao 21, which was not coated. A total of 1600 cotton seeds were randomly selected, including 540 intact cotton seeds, 530 slightly damaged cotton seeds, and 530 seriously damaged cotton seeds. The second group was Xinluzao 21, a Xinjiang cotton seed treated with coating. A total of 1600 coated cotton seeds were randomly selected, including 540 intact cotton seeds, 530 slightly damaged cotton seeds, and 530 seriously damaged cotton seeds.

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

**Figure 3.** Appearance of uncoated cotton seed. (**a**) Undamaged; (**b**) slight-damaged; (**c**) seriously damaged.

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

**Figure 4.** Appearance of coated cotton seed. (**a**) Undamaged; (**b**) slight-damaged; (**c**) seriously damaged.

2.2.1. Data Collection and Annotation

The images used for the appearance-based damage detection of the cotton were collected using an image acquisition device composed of a carrier platform, light source, light source controller, camera, lens, and computer. Figure 5 presents a schematic diagram of the acquisition device. When collecting the image of cotton seeds, because the size of cotton seeds was small and the damage characteristics of cotton seeds should be better seen during detection, the annular light source was selected when building the acquisition device. The annular light source was a high-brightness divergent illumination. Therefore, to make the defects and special structures on the surface of the detected object better displayed under its illumination, it was necessary to keep the annular light source above the cotton seed and parallel to the carrier platform. Due to the determination of the position of the light source, to better use the camera to take the characteristics of the cotton seed surface, it was necessary to make the camera's placement angle parallel to the light source. For the cotton seed image acquisition process, the cotton seed was placed directly below the camera lens. Each cotton seed was ensured to be in the same position. The annular light source was used to uniformly illuminate the surface of the cotton seed, adjust the focal length and aperture of the camera, and maintain the cotton seed at a certain distance from the camera lens. The cotton seed image was collected immediately and the obtained cotton seed image was stored on the computer.

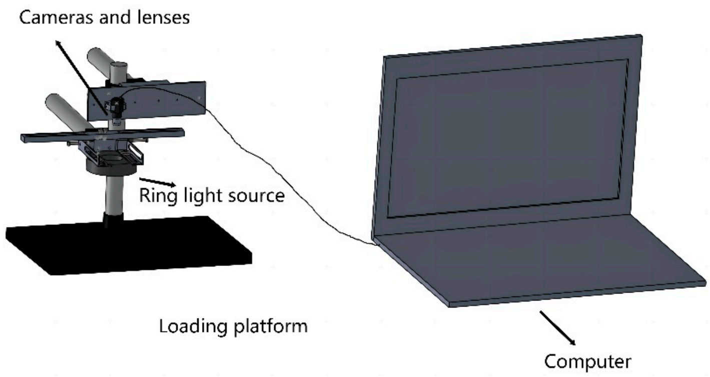


**Figure 5.** Schematic diagram of the acquisition device.

Prior to the damage detection, the collected cotton seed images were marked using labeling. The cotton seed was placed in the center of a rectangular box to reduce the interference caused by the redundant background. The position of the box was determined by the coordinates of the two diagonals. Cotton seeds with different appearances were marked with different color boxes and denoted as undamaged, slightly damaged, and seriously damaged. The color and name tag of different marker boxes allowed us to see the identified results more directly. Figure 6 presents the labeling process.

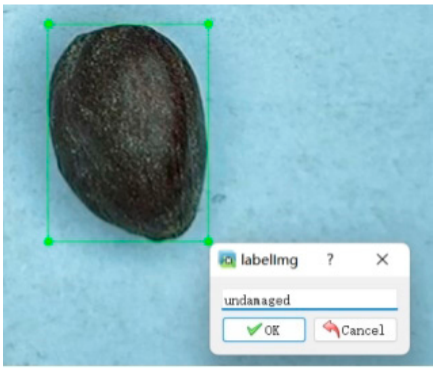


**Figure 6.** Labeling process.

2.2.2. Data Enhancement

In order to better identify the damage of cotton seeds and improve the accuracy and accuracy of cotton seed identification, this paper used two methods of adding noise and brightness transformation for data processing. The brightness transformation was used to change the color value of each channel of the pixel. There are many types of noise, but the most commonly used types in practice are Gaussian noise and salt and pepper noise. Gaussian noise, also known as white noise, is a random noise that obeys normal distribution. Salt and pepper noise is also known as impulse noise. It randomly changes some pixel values, which makes some pixels white and some pixels black on the binary image. Compared with other types of noise, these two kinds of noise are simpler, which makes the training data enhancement simpler. Because they are white noise and impulse noise, they can simulate the noise in the actual scene. Gaussian noise can simulate the random noise in the actual scene, and salt and pepper noise can simulate the sudden noise in the actual scene, so as to improve the robustness of the model. Therefore, Gaussian noise and salt and pepper noise were selected in this paper when adding noise. Through data enhancement, the diversity of training data was increased, so that the cotton seed detection algorithm can better adapt to the object recognition task under different scenarios and changing conditions. By performing a series of processing on the original image, a more diverse image was generated, so that the algorithm can learn richer features and context information. The image after data processing is shown in Figure 7.

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

**Figure 7.** Data enhancement. (**a**) Gaussian noise; (**b**) salt and pepper noise; (**c**) brightness transformation.

After data-enhancement processing, the number of cotton seed images in each group was 2400, of which 800 images were selected as the validation set for the final evaluation of algorithm performance. The remaining 1600 images were used as training sets for algorithm training. The purpose of data enhancement is to increase the number of images in the data set. With the increase in the number of images in the data set, the accuracy, robustness, and generalization ability of the YOLOv5 target detection algorithm were effectively improved, so that it can better learn the characteristics of the object and accurately identify the appearance damage of cotton seeds in a complex environment.

### 2.3. Overview of YOLOv5 Algorithm

YOLOv5 is a deep-learning-based single target detection algorithm that improves on previous versions (e.g., YOLOv3 and YOLOv4) and adopts a new architecture. The YOLOv5 model consists of three components, namely, the backbone network, head network, and loss function. YOLOv5 can be divided into the following four models according to the number of network layers and feature map channels (in ascending order): YOLOv5s; YOLOv5m; YOLOv5l; and YOLOv5x. In these four models, YOLOv5s has a smaller model size, less calculation, and less storage space than other versions, and so, YOLOv5s is faster than other models. At the same time, YOLOv5s is suitable for dealing with smaller target scales, while the size of cotton seeds is relatively small and more suitable for YOLOv5s. Therefore, the model used in this paper was YOLOv5s. Figure 8 depicts the network structure diagram of YOLOv5.

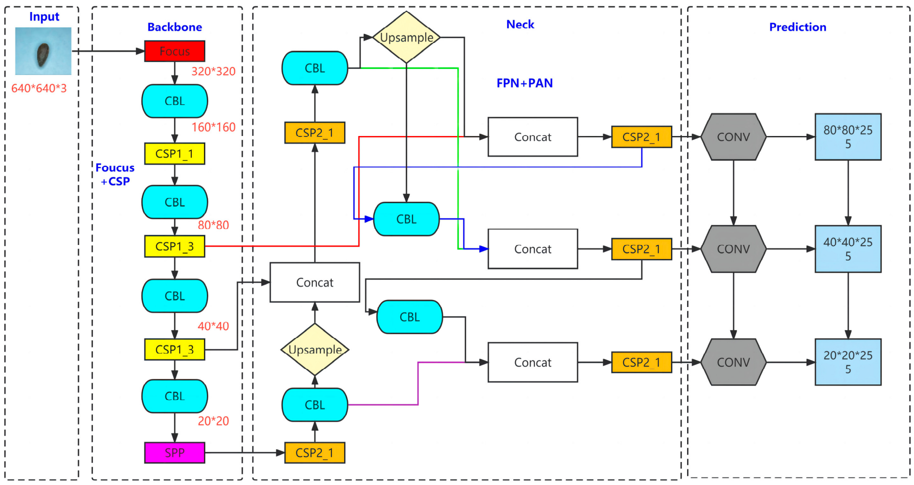


**Figure 8.** YOLOv5 network structure diagram.

The three key components of the YOLOv5s network structure are the backbone, neck, and head. The backbone uses the FAT (Focus–Attention Transformation) structure, which includes three convolution blocks, two repeat modules, and one SPP (Spatial Pyramid Pooling) module. The backbone's task is to extract features from the input image and gradually enhance the expression ability of the feature map. In particular, the backbone uses convolution and pooling operations to extract features from input images of different sizes. Based on the backbone, a set of simple and effective feature pyramid modules are added in the neck component. These modules can adaptively focus on the features of different regions, thereby improving the quality and accuracy of the feature representation. Each branch of the feature pyramid module has a different receptive field size to adapt to different scale object detection tasks [21]. Finally, YOLOv5s uses the predicted head to output the target detection results (the object category, border position, and confidence). In this process, three-layer convolution is used for the feature extraction and synthesis [22,23].

## 3. Improvements in the YOLOv5 Algorithm

In order to improve the accuracy of the cotton-seed identification and reduce the missed and false detection rates, this paper improved the original YOLOv5 algorithm. More specifically, we added the lightweight upsampling operator CARAFE and an improved loss function to improve the detection of the appearance damage of cotton seeds.

### 3.1. Addition of Lightweight Upsampling Operator CARAFE

In YOLOv5, the upsampling operation refers to expanding the size of the original image or feature map to a certain size. Bilinear interpolation is typically employed for the upsampling, that is, the interpolating function is usually adopted. The bilinear interpolation method only determines the upsampling kernel by the spatial position of the pixel point and does not use the semantic information of the feature map. It can be regarded as a 'uniform' upsampling, and the sensing domain is usually very small [24,25]. Therefore, to further improve the performance of YOLOv5, this paper added the lightweight upsampling operator CARAFE to the YOLOv5s model. This can effectively increase the resolution and quality of the feature map, improve the performance of the model, and reduce the computational calculations compared to the traditional upsampling operator. CARAFE consists of an up-sampling kernel prediction module and a content-aware reorganization module. Figure 9 presents its frame diagram.

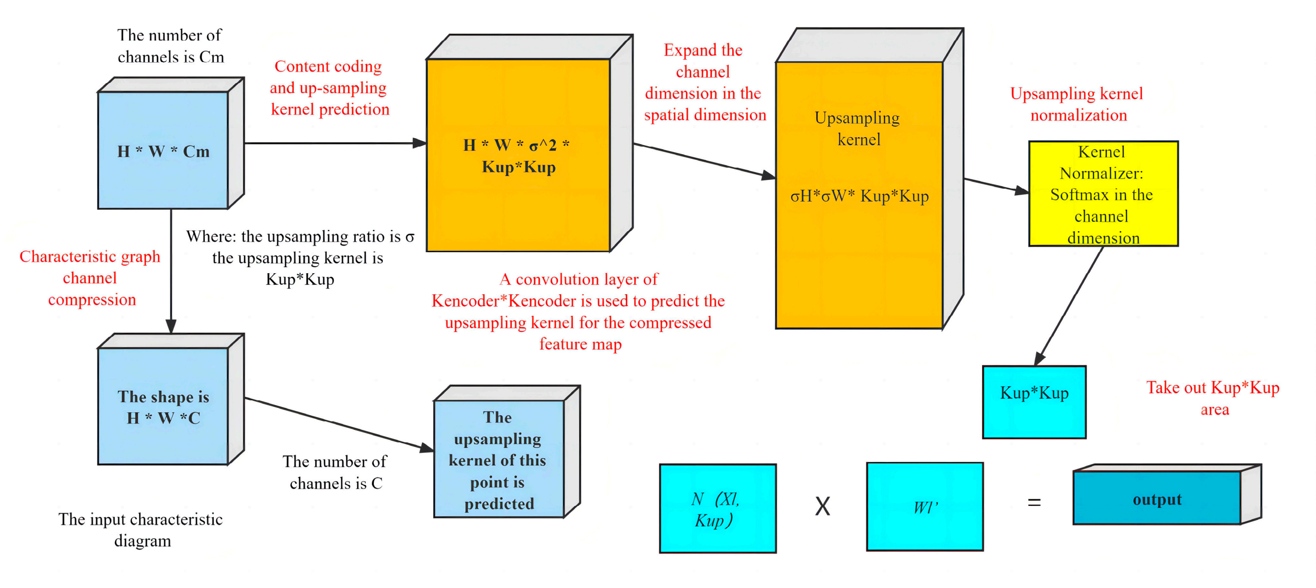


**Figure 9.** CARAFE frame diagram.

The upsampling kernel prediction module predicts a learnable recombination kernel based on the content of each target position, thereby achieving adaptive convolution operations. For this process, the feature map was input into the module and the number of feature map channels were compressed. The receptive field of each target after block was obtained and the local attention mechanism was used to predict the recombination kernel of each target position. The feature recombination module employed the predicted recombination check to recombine the features, thus obtaining more discriminative features. In particular, the adaptive convolution operation was performed on the input feature map using the reconstructed kernel predicted in the previous step, and the feature representation of each target position was obtained. The obtained feature representation was then fused with the original feature and was processed by the convolution layer. The processed feature is the final output [26,27].

Compared with the upsampling of bilinear interpolation, CARAFE has a larger receptive field and can aggregate context information in a large receiving domain. CARAFE has different content awareness for different samples, that is, CARAFE supports instance-specific content-aware processing, which can dynamically generate an adaptive kernel. The FPN (Feature Pyramid Network) architecture with CARAFE improves on the upsampling compared to the original FPN framework, and its performance is also significantly better than its operator. The addition of CARAFE to the network can effectively improve the model performance. Moreover, CARAFE is lightweight and can, thus, maintain the lightweight characteristic of the model.

### *3.2. Improvement of the Loss Function*

The original loss function of YOLOv5 uses CIoU (Complete-IoU), the improvement of the loss function in this study was to change CIoU to EIoU (Efficient-IoU). CIoU is an index to measure the distance between two bounding boxes. It is an improved algorithm of the traditional IoU. CIoU initially calculates IoU (Intersection over Union) and subsequently adds a correction factor, which accounts for variables such as the distance between the center points of the two bounding boxes and the aspect ratio. This allows CIoU to better adapt to objects with different aspect ratios and proportions [28,29]. The specific formula is as follows:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha V \tag{1}$$

$$V = \frac{4}{\pi^2}\left(\arctan\frac{w^2}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \tag{2}$$

$$\alpha = \begin{cases} 0, \text{ if } IoU < 0.5 \\ \frac{V}{(1-IoU)+V}, \text{ if } IoU \geq 0.5 \end{cases} \tag{3}$$

where V is the normalization of the difference between the length–width ratio of the prediction box and the real box; b is the parameter of the prediction center coordinate; $b^{gt}$ is the parameter of the center of the real target bounding box; c is the diagonal length of the minimum circumscribed rectangle of the two rectangles; $\alpha$ is the balance factor used to weigh the loss caused by the length–width ratio and the loss caused by the IoU component; $\rho$ is the Euclidean distance between b and $b^{gt}$; and w, h and $w^{gt}$, $h^{gt}$ are the height and width of the prediction frame and the height and width of the real frame, respectively.

CIoU increases the loss of the detection frame scale and the loss of length and width on the basis of DIoU, so that the prediction box of CIoU is more in line with the real box, However, it is still associated with several shortcomings. For example, although CIoU considers the overlap area, center point distance, and aspect ratio of the bounding box regression, it reflects the difference in the aspect ratio through V in Formula (1) rather than the true difference between the width and height and their confidence. Thus, CIoU can hinder the effective optimization similarity of the model. Therefore, based on CIoU, scholars further split and optimized the aspect ratio, developing EIoU, which includes focal focusing technology to optimize the quality of the anchor frame [30]. EIoU is described as follows:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \frac{\rho^2(w,w^{gt})}{C_w^2} + \frac{\rho^2(h,h^{gt})}{C_h^2} \tag{4}$$

where $C_w$ and $C_h$ are the width and height of the smallest bounding box covering both boxes. EIoU consists of three components, namely, overlap loss, center distance loss, and width–height loss. The first two components follow the method of CIoU, and the width–height loss directly minimizes the difference between the width and height of the prediction box and real box. Thus, the convergence speed is faster than that of CIoU. In order to overcome the problem of unbalanced training samples in the BBox regression, Focal Loss is added in EIOU. This also further improves the accuracy. Moreover, EIoU also solves the problem whereby the width and height of CIoU cannot simultaneously increase or decrease. In summary, EIoU accelerates the convergence and enhances the regression accuracy based on CIoU, and also improves the imbalance of training samples by adding Focal Loss.

## 4. Experimental Results and Analysis

### *4.1. Experimental Environment and Experimental Methods*

The Windows 11 operating system was adopted as the experimental platform, with a 12th Gen Intel (R) Core (TM) i9-12900 KF processor (Lenovo Beijing Limited, Beijing, China) and PyTorch as the deep learning framework. Pycharm2022, Python3.9.7, and anaconda3 were employed as the IDE environment, programming language, and package manager, respectively. In addition, the input size of the image was 640 × 640, the batch size was 16, the initial learning rate was 0.01, and the number of iterations was 200.

### *4.2. Model Evaluation Indicators*

Numerous indictors can be used to represent the detection results of YOLOv5. Among these, accuracy and recall are two important indicators used to evaluate the performance of the YOLOv5 object-detection system. The accuracy rate refers to how many of the objects detected by the object detection system are correct, and the recall rate refers to the probability that the real object is detected [31]. In general, the higher the correct rate, the higher the recall rate, and the better the performance of the system. In this paper, the intersection over union (IoU), mean average precision (mAP), precision ($R_p$), recall ($R_r$), and accuracy ($R_a$) are used as the evaluation indexes of the model performance with the following formula:

$$\mathrm{IoU} = \frac{\mathrm{A} \cap \mathrm{B}}{\mathrm{A} \cup \mathrm{B}} \tag{5}$$

$$\mathrm{R_p} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{6}$$

$$\mathrm{R_r} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{7}$$

$$\mathrm{R_a} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}} \tag{8}$$

where A is the prediction box and B is the real box; TP is a classifier used to classify the correct positive cases into positive cases, and FP is used as a classifier to classify the wrong negative cases into positive cases; FN is a classifier used to classify the wrong classification of positive examples into negative examples; and TN is a classifier used to classify the correct classification of negative cases-predicting negative cases. Furthermore, the mAP refers to the average of all AP types in all images. The higher the mAP, the higher the prediction accuracy of the model.

### *4.3. Comparison of YOLOv5 Algorithm before and after Improvement*

#### 4.3.1. Ablation Experiment Test

In this cotton seed damage detection test, the YOLOv5 algorithm was improved in two modules. To verify the optimization effect of each improvement on the original algorithm, an ablation test was designed based on the YOLOv5s algorithm according to these two modules. There were four groups of experiments, which were the original algorithm, the improved EIOU on the original algorithm, the addition of CARAFE on the original algorithm, and the improved two-point algorithm in this paper. Table 1 reports the test results.

**Table 1.** Ablation test results. '√' indicates an improvement in the module, while '×' indicates no improvement in the module.

| Method | CARAFE | EIOU | mAP_0.5/% | Average Test Time/ms | Model Size/MB |
|---|---|---|---|---|---|
| YOLOv5s | × | × | 98.3 | 42 | 7.62 |
| YOLOv5s + EIOU | × | √ | 98.9 | 41 | 5.23 |
| YOLOv5s + CARAFE | √ | × | 99.2 | 38 | 3.78 |
| Proposed algorithm | √ | √ | 99.5 | 36 | 3.36 |

Table 1 reveals improvements to the original algorithm in several components. The addition of the lightweight upsampling operator CARAFE enhanced the feature representation ability by performing a nonlinear transformation on the feature map without increasing the network parameters and computational complexity. It also highlighted the features and improved the object-detection accuracy. As a result, the average detection time and model size were reduced by 4 ms and 3.84 MB, respectively, compared with the original algorithm. The model performance was also enhanced after the improvement of the loss function. Compared with the addition of the lightweight upsampling operator CARAFE, its average detection time and model size were reduced by 1 ms and 2.39 MB, respectively. Table 1 demonstrates that the improved Yolov5 algorithm proposed in this paper also changed the loss function to EIOU, while adding the lightweight upsampling operator CARAFE, and, thus, has the optimal performance for the appearance-based damage detection of cotton. In particular, its mAP_0.5 value reached 99.5%, which exceeds that of the original YOLOv5s by 1.2%. The average detection time and model size were also reduced by 6 ms and 4.26 MB, respectively.

#### 4.3.2. Comparison of Results before and after Improvement

In this paper, the original YOLOv5 algorithm was improved by adding the lightweight upsampling operator CARAFE. In order to better verify the improvement of the proposed YOLOv5 algorithm compared to the original YOLOv5 algorithm, a total of 1600 images were used to test the cotton seed appearance-based detection performance of the models. Under the same number of iterations, the performance indicators of the improved and original YOLOv5 algorithms were analyzed. Table 2 reports the results.

**Table 2.** Performance index comparison of uncoated cotton seed.

| Algorithm | Recall/% | AP/% | | | mAP_0.5/% |
|---|---|---|---|---|---|
| | | Undamaged | Slightly Damaged | Seriously Damaged | |
| YOLOv5s | 98.4 | 98.6 | 97.7 | 98.6 | 98.3 |
| Improve YOLOv5s | 99.3 | 99.8 | 99.6 | 99.1 | 99.5 |

Table 2 reveals that the original YOLOv5 algorithm was improved based on our proposed changes to the algorithm, with the recall rate reaching 99.3%. The average accuracy AP of undamaged, slightly damaged, and severely damaged cotton seeds was 99.8%, 99.6%, and 99.1%, respectively, and the mAP_0.5 value reached 99.5%.

In order to improve the survival rate and germination rate of cotton seeds, people usually coat cotton seeds. In order to detect the appearance damage of cotton seeds more comprehensively and accurately and to explore whether the coating treatment will have an impact on the detection of the appearance damage of cotton seeds, this paper used the appearance damage detection algorithm of uncoated cotton seeds to detect the appearance damage of coated cotton seeds. The same data-enhancement processing was used to obtain

2400 pictures of coated cotton seeds, and 800 pictures were selected as the verification set, and the remaining 1600 pictures were used as the training set. The original YOLOv5 algorithm was used for detection and the improved YOLOv5 algorithm was used for detection. We observed the results of coated cotton seed detection before and after the algorithm improvement and analyzed them under the same number of iterations. Table 3 reports the results.

**Table 3.** Performance index comparison of coated cotton seed.

| Algorithm | Recall/% | AP/% | | | mAP_0.5/% |
|---|---|---|---|---|---|
| | | Undamaged | Slightly Damaged | Seriously Damaged | |
| YOLOv5s | 98.2 | 98.4 | 98.8 | 98.9 | 98.7 |
| Improve YOLOv5s | 98.9 | 99.4 | 98.8 | 99.4 | 99.2 |

It can be seen from Table 3 that the average accuracy AP of the improved YOLOv5 algorithm for the appearance damage detection of coated cotton seeds was also very high. The average accuracy AP of undamaged, slightly damaged, and severely damaged cotton seeds was 99.4%, 98.8%, and 99.4%.

In addition, in order to more intuitively see the changes before and after the improvement under the same number of iterations, the four indicators of mAP_0.5, positioning loss box_loss, classification loss cls_loss, and confidence loss obj_loss, were compared. The results are shown in Figure 10, where epochs is the number of iterations, that is, all training data sets have been trained during training, mAP_0.5 is to set IoU to 0.5, we calculated the AP of all pictures in each category, and then averaged all categories; box_loss was used to measure the difference between the box predicted by the model and the real box. The smaller the value of box_loss, the more accurate the box is. Cls_loss was used to determine whether the model can accurately identify the objects in the image and classify them into the correct category. The smaller the value of cls_loss, the more accurate the classification; obj_loss was to calculate the confidence of the network. The smaller the value of obj_loss is, the more accurate the target detection is.

Based on the indicator mAP_0.5, the improved algorithm was better than the original YOLOv55s in the appearance damage detection of uncoated cotton seeds, and the number of iterations increased significantly in the range of 25–50, and the curves of the two algorithms fluctuated greatly (Figure 10a). When the number of iterations was in the range of 50–200, the curves of the two algorithms gradually became stable and fluctuated less. Figure 10 also shows that the three losses of location loss, box_loss, classification loss cls_loss, and confidence loss obj_loss, were lower for the proposed algorithm compared to the original algorithm. This indicates that the improved algorithm has a smaller loss value and better convergence. There was an obvious improvement in the positioning loss box_loss curve of the proposed model compared to that of the original algorithm, with the former reaching a value less than 0.06, and subsequently converging rapidly (Figure 10b). This indicates that the improved algorithm had better prediction accuracy for the target position compared to the original algorithm. The curve of the classification loss cls_loss of the improved algorithm and the classification loss cls_loss of the original algorithm fluctuated gently (Figure 10c). The initial value of the classification loss cls_loss was the same for both algorithms, yet as the number of iterations increased, the classification loss cls_loss of the improved algorithm gradually became smaller than that of the original algorithm. However, the difference between the values of the two algorithms was small. This indicates that the improved algorithm increased the prediction accuracy of the target category, but the increase was not very obvious. The loss curve of the first 25 iterations of the confidence loss obj_loss converged rapidly, and the initial value was less than 0.04 for the proposed algorithm (Figure 10d). This demonstrates the improved recognition ability and detection accuracy of the damage characteristics of cotton seeds for the proposed algorithm. In summary, the improved YOLOv5 algorithm achieved a certain improvement in the appearance damage detection of uncoated cotton seeds.

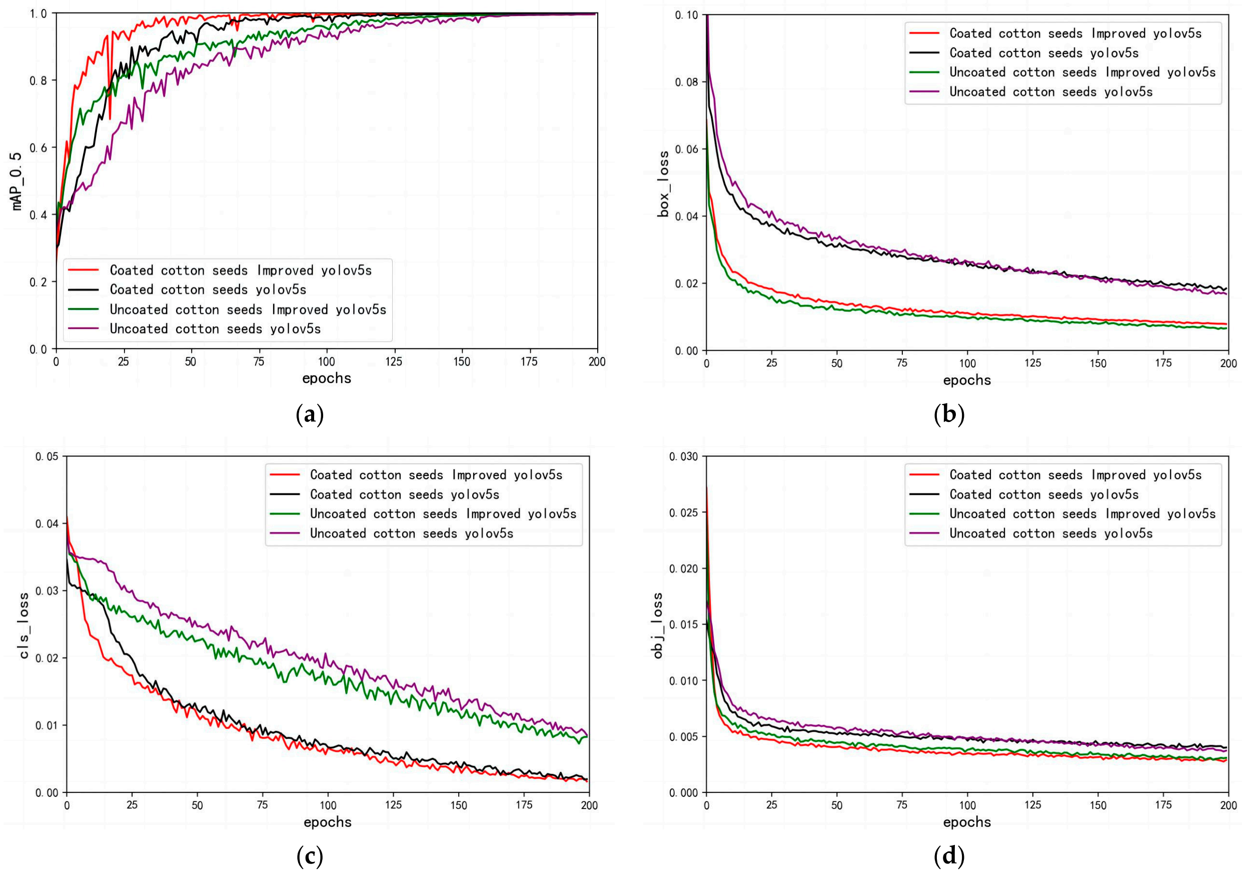


**Figure 10.** Comparison of indexes: (**a**) mAP_0.5; (**b**) box_loss; (**c**) cls_loss; (**d**) obj_loss.

It can be seen from Figure 10 that the improved algorithm in the appearance damage detection of coated cotton seeds was also improved compared with the original YOLOv5s. Based on the indicator mAP_0.5, from Figure 10a, it is evident that the number of iterations significantly increased in the range of 0–25 for both algorithms, with considerable fluctuations in their curves. In the range of 25–75 iterations, the fluctuations decreased compared to the 0–25 range, gradually stabilizing. Between 75 and 200 iterations, both curves tended to stabilize. Additionally, from Figure 10a, it can be observed that the improved algorithm's curve was lower than the original algorithm's curve near 25, 75, and 100 iterations, indicating some impact on the extraction of cotton seed appearance damage features after packaging. Regarding the box_loss metric for localization loss, Figure 10b shows that the improved algorithm's box_loss curve was more significantly enhanced compared to the original algorithm, and it converged faster, indicating improved accuracy in predicting target locations. From Figure 10c, in the early iterations, the improved algorithm's cls_loss curve was higher than the original algorithm's cls_loss curve, but in the later iterations, the difference between the two curves decreased, suggesting a moderate improvement in predicting target categories with the improved algorithm. As seen in Figure 10d, during the initial iterations, the improved algorithm's obj_loss curve was higher than the original algorithm's obj_loss curve, but it converged faster with larger fluctuations. However, in the later iterations, the improved algorithm's curve was lower than the original algorithm's, indicating better recognition ability for cotton seed appearance damage features and higher accuracy in detecting wrapped cotton seed appearance damage. In summary, the improved Yolov5 algorithm also improved the appearance damage detection of coated cotton seeds, but it is not as good as uncoated cotton seeds in the index of mAP_0.5.

### 4.4. Visual Analysis

In order to be able to more intuitively observe the detection effect of the original and the improved algorithms for the detection of cotton seed damage, cotton seed images were randomly selected from the test set and a comparison was performed in the same environment. Figures 11 and 12 present the test results of the original and improved algorithms.


**Figure 11.** Test results of original algorithm.

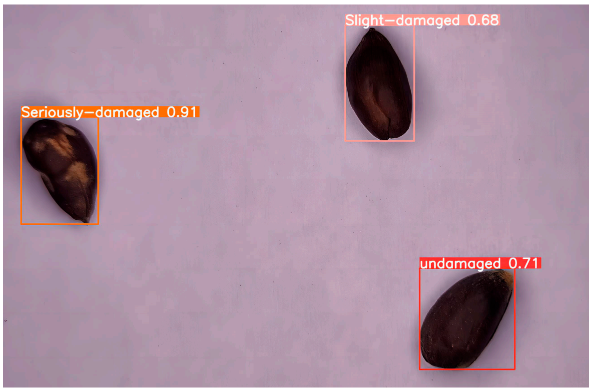


**Figure 12.** Test results of improved algorithm.

Figure 11 shows a false detection in the detection of cotton seed damage for the original algorithm. This may be attributed to the low number of layers of YOLOv5s, an insufficient receptive field, and a large amount of calculations during the feature fusion. In Figure 12, the improved algorithm can well identify the damage of cotton seeds, as well as cotton seeds with different damage conditions. Compared with the original algorithm, the detection accuracy of cotton seed appearance-based damage detection was greatly improved. Thus, the improved algorithm improved the problem of false detection in the detection of cotton damage and enhanced the detection accuracy, thus resulting in a better cotton damage detection effect.

In order to observe whether the coating has an effect on the appearance damage detection of cotton seeds, the images of coated cotton seeds were randomly selected from the test set, and tested and compared in the same test environment. Through the appearance damage detection of coated cotton seeds, it was found that the original algorithm had the phenomenon of false detection for the appearance damage detection of cotton seeds and the phenomenon that one cotton seed corresponded to multiple labels. The detection accuracy was not high, which may be due to the fact that the number of layers of YOLOv5s model was too small, the receptive field was insufficient, and the network structure had a large amount of calculation when the features were fused, resulting in the loss of target information, which led to the occurrence of false detection. The improved algorithm improved the recognition accuracy of the appearance damage of coated cotton seeds, and can well identify intact coated cotton seeds and severely damaged coated cotton seeds, but can not well identify slightly damaged coated cotton seeds. This may be due to the fact that the characteristics of slight damage were not obvious enough to be affected by coating, so it can not well identify slightly damaged coated cotton seeds. In addition, it may also be due to the training of a single cotton seed, while identifying multiple cotton seeds at the same time will be biased, and the accuracy of identifying a single cotton seed is higher. The detection of a single-coated cotton seed is shown in Figure 13; however, when identifying uncoated cotton seeds, the improved algorithm can identify multiple cotton seeds at the same time, indicating that the coating had some influence on the identification of the appearance damage of cotton seeds.


**Figure 13.** Detection diagram of single-coated cotton seed. (**a**) Undamaged; (**b**) slightly damaged; (**c**) seriously damaged.

## 5. Discussion and Conclusions

### *5.1. Discussion*

At present, target detection is widely used, and the corresponding target detection technology has made continuous progress. However, the accurate and rapid detection and identification of targets still requires improvements. In recent years, with the development of the YOLO algorithm, many domestic scholars have used it to detect crop seeds. Wang Qiaohua and Gu Wei improved YOLOv4 and used the improved YOLOv4 to realize the damage identification of double-sided group cotton seeds. The experimental results show that the improved YOLOv4 has a detection accuracy of 95.33%, a recall rate of 96.31%, and a missed detection rate of 0 for damaged and intact cotton seeds. The detection effect is better than the original YOLOv4 network [10]. Li Haorui detected the purity of corn seeds by improving the YOLO model. By comparing the advantages and disadvantages of the genetic algorithm and particle swarm optimization algorithm, the genetic algorithm was used to improve particle swarm optimization algorithm, and the optimization of deep learning network based on YOLO was realized. The final detection accuracy reached 99.5% [11]. Fan Xiaofei combined with four-channel multi-spectral images using improved YOLOv4 to study the appearance–quality detection of corn seeds. The experimental results showed that the average F1 value and mAP of the improved model reached 93.09% and 98.02% [12]. Through the research of many domestic scholars, it was found that the YOLO algorithm has achieved good results in the detection of crop seeds, and the detection accuracy at this stage is above 90%. It can be seen that the use of YOLOv5 for cotton seed appearance damage detection has certain reference significance. The advantage of this study was that the lightweight upsampling operator CARAFE and the improved loss function were added to the original YOLOv5 algorithm, which made the algorithm not only lightweight but also have high performance.

Our test results revealed that the detection of cotton seeds with the proposed model was affected by the image-acquisition environment. The quality of cotton seed images collected in under environments was also different. To effectively realize the damage detection of cotton seeds, it was necessary to ensure that the acquisition environment is optimal. However, in the actual scene, the acquisition environment was uncertain. Therefore, the extraction of cotton seed feature information in the detection of cotton seed damage must be strengthened in order perform the detection of different quality cotton seed images under different sampling environments. The mAP_0.5, AP, and recall rate values of the proposed model were very high. This may be due to the optimal cotton seed image acquisition environment in this paper. The mAP_0.5, AP, and recall rate values were obtained in a theoretical state. Therefore, in practical applications, it is also necessary to comprehensively evaluate and adjust the performance of the model in combination with specific application scenarios. The improved YOLOv5 algorithm proposed here can act as a reference for cotton seed damage detection, yet it needs to be improved and optimized with specific application scenarios.

### *5.2. Conclusions*

This paper proposed an algorithm based on the improved YOLOv5 target detection algorithm to improve the accuracy of cotton seed appearance-based damage detection, reduce the false detection and missed detection rates, and solve the problem of unclear cotton seed damage identification. By adding the lightweight upsampling operator CARAFE, the resolution and quality of the feature map were markedly increased, and the performance and loss function of the model were improved. CIoU was replaced with EIoU, which not only sped up the convergence speed, but also improved the imbalance of the training samples. Based on the aforementioned two improved YOLOv5 algorithms, all tested performance index of the cotton seed appearance-based damage detection were improved. In particular, the mAP_0.5, recall rate, average detection time, and model size reached 99.5%, 99.3%, 36 ms, and 3.36 MB, respectively. The experimental results verified the higher accuracy and faster detection speed of the improved YOLOv5s model compared to the original version. At the same time, the coated cotton seeds were also tested and analyzed. The test results showed that the mAP_0.5 value of the improved algorithm reached 99.2%, and the recall rate reached 98.9%. Through visual analysis, it was found that the coating treatment had an impact on the appearance damage detection of cotton seeds. In the future, we will continue to optimize the algorithm to cope with the increasingly complex and changeable practical application scenarios and the detection of different cotton seeds, and provide reference for the appearance detection of crop seeds in the future.

**Author Contributions:** Writing—Original draft preparation, Z.L. (Zhicheng Liu); writing—review and editing, L.W.; data curation, Z.L. (Zhiyuan Liu); resources, X.W.; visualization, C.H.; supervision, J.X.; project administration, L.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Bintuan Science and Technology Program (Grant No. 2022CB001-06) and the President Fund of the Tarim University (TDZKYS202302).

**Institutional Review Board Statement:** Not applicable.

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

**Acknowledgments:** The authors are grateful for the support of the Bingtuan Science and Technology Program and to anonymous reviewers for their comments.

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

## References

1. State Statistical Bureau. *China Statistical Yearbook*; China Statistics Press: Beijing, China, 2022.
2. He, J. Research on optimization and innovation of Xinjiang cotton logistics system from the perspective of science and technology empowerment. *China Cott.* **2022**, *49*, 44–46.
3. Zhu, J.; Li, J.; Kan, Z.; Zhang, R. The Simple Analysis of the Current Situation and the Development Tendency of Sorting Cotton Seed which Escaping the Fabric in Xinjiang Production and Construction Corps. *J. Agric. Mech. Res.* **2008**, *1*, 216–219.
4. Wang, J.; Lu, Y.; Yang, H.; Gao, Z.; Liu, S. Present situation of delinted cotton seed selection technology. *Jiangsu Agric. Sci.* **2014**, *42*, 5–7.
5. Kan, Z.; Gu, C.; Wang, L.; Fu, W. Optimization of parameters for delinted cottonseeds dielectric selection. *Trans. Chin. Soc. Agric. Eng.* **2010**, *26*, 114–119.

6. Yu, S.; Liu, Y.; Wang, S.; Du, H. Design and Simulation Analysis of Dual CCD Detection System for Delinted Cottonseeds Color Sorter. *Trans. Chin. Soc. Agric. Mach.* **2015**, *46*, 55–61.
7. Peng, S.; You, J.; Li, J. Software Design of the Testing System for the Dynamic Test of the Cotton Seed. *J. Agric. Mech Res.* **2017**, *39*, 66–71.
8. Weng, H.; Ye, Q. Applications of Feature Extraction and Enhancement Algorithm in Image Processing. *J. Chongqing Univ. Technol. Nat. Sci.* **2016**, *30*, 119–122.
9. Wu, Y.; Xu, W.; Zhang, D. A Study on the Extraction and Application of Image Contour Feature Codes. *J. Residuals Sci. Technol.* **2016**, *13*, 5–12.
10. Wang, Q.; Gu, W.; Cai, P.; Zhang, H. Detection Method of Double Side Breakage of Population Cotton Seed Based on Improved YOLO v4. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 389–397.
11. Fan, X.; Wang, B.; Liu, J.; Zhou, Y.; Zhang, J.; Suo, X. Corn Seed Appearance Quality Estimation Based on Improved YOLO v4. *Trans. Chin. Soc. Agric. Mach.* **2022**, *53*, 226–233.
12. Li, H. *Research on the Application of Improvedimage Processing System Based on YOLOmodel in Seed Purity Inspection*; Sichuan Agricultural University: Ya'an, China, 2020.
13. Li, S.; Zhang, S.; Xue, J.; Sun, H.; Ren, R. A Fast Neural Network Based on Attention Mechanisms for Detecting Field Flat Jujube. *Agriculture* **2022**, *12*, 717. [CrossRef]
14. Li, Z.; Tian, X.; Liu, X.; Liu, Y.; Shi, X. A Two-Stage Industrial Defect Detection Framework Based on Improved-YOLOv5 and Optimized-Inception-ResnetV2 Models. *Appl. Sci.* **2022**, *12*, 834. [CrossRef]
15. Ahmad, T.; Cavazza, M.; Matsuo, Y.; Prendinger, H. Detecting Human Actions in Drone Images Using YoloV5 and Stochastic Gradient Boosting. *Sensors* **2022**, *22*, 7020. [CrossRef] [PubMed]
16. Yang, X.; Zhao, J.; Zhao, L.; Zhang, H.; Li, L.; Ji, Z.; Ganchev, I. Detection of River Floating Garbage Based on Improved YOLOv5. *Mathematics* **2022**, *10*, 4366. [CrossRef]
17. Zhou, J.; Jiang, P.; Zou, A.; Chen, X.; Hu, W. Ship Target Detection Algorithm Based on Improved YOLOv5. *J. Mar. Sci. Eng.* **2021**, *9*, 908. [CrossRef]
18. Ahmad, I.; Yang, Y.; Yue, Y.; Ye, C.; Hassan, M.; Cheng, X.; Wu, Y.; Zhang, Y. Deep Learning Based Detector YOLOv5 for Identifying Insect Pests. *Appl. Sci.* **2022**, *12*, 10167. [CrossRef]
19. Wang, H.; Shang, S.; Wang, D.; He, X.; Feng, K.; Zhu, H. Plant Disease Detection and Classification Method Based on the Optimized Lightweight YOLOv5 Model. *Agriculture* **2022**, *12*, 931. [CrossRef]
20. Obsie, E.Y.; Qu, H.; Zhang, Y.; Annis, S.; Drummond, F. YOLOv5s-CA: An Improved YOLOv5 Based on the Attention Mechanism for Mummy Berry Disease Detection. *Agriculture* **2022**, *13*, 78. [CrossRef]
21. Khan, Z.N.; Ahmad, J. Attention induced multi-head convolutional neural network for human activity recognition. *Appl. Soft Comput.* **2021**, *110*, 107671. [CrossRef]
22. Du, Q.; Huang, L.; Tian, L.; Huang, D.; Jin, S.; Li, M. Recognition of Passengers' Abnormal Behavior on Escalator Based on Video Monitoring. *J. South China Univ. Technol. Nat. Sci.* **2020**, *48*, 10–21.
23. Qu, Z.; Gao, L.; Wang, S.; Yin, H.; Yi, T. An improved YOLOv5 method for large objects detection with multi-scale feature cross-layer fusion network. *Image Vision Comput.* **2022**, *125*, 104518. [CrossRef]
24. Zhou, Y.; Chen, S. Research on Lightweight Improvement of Sonar Image Classification Network. *J. Phys. Conf. Ser.* **2021**, *1883*, 012140. [CrossRef]
25. Li, Z. Road Aerial Object Detection Based on Improved YOLOv5. *J. Phys. Conf. Ser.* **2022**, *2171*, 012039. [CrossRef]
26. Wu, S.; Li, Z.; Li, S.; Liu, Q.; Wu, W. Static Gesture Recognition Algorithm Based on Improved YOLOv5s. *Electronics* **2023**, *13*, 596. [CrossRef]
27. Wang, J.; Chen, K.; Xu, R.; Liu, Z.; Loy, C.C.; Lin, D. CARAFE++: Unified Content-Aware ReAssembly of FEatures. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 4674–4687. [CrossRef]
28. Zhao, J.; Zhang, X.; Yan, J.; Qiu, X.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W. A Wheat Spike Detection Method in UAV Images Based on Improved YOLOv5. *Remote Sens.* **2021**, *13*, 3095. [CrossRef]
29. Sun, Z.; Yang, H.; Zhang, Z.; Liu, J.; Zhang, X. An Improved YOLOv5-Based Tapping Trajectory Detection Method for Natural Rubber Trees. *Agriculture* **2022**, *12*, 1309. [CrossRef]
30. Shen, Y.; Zhang, F.; Liu, D.; Pu, W.; Zhang, Q. Manhattan-distance IOU loss for fast and accurate bounding box regression and object detection. *Neurocomputing* **2022**, *500*, 99–114. [CrossRef]
31. Jiao, H. Intelligent Research Based on Deep Learning Recognition Method in Vehicle-Road Cooperative Information Interaction System. *Comput. Intell. Neurosci.* **2022**, *2022*, 4921211. [CrossRef]