Seedling-YOLO: High-Efficiency Target Detection Algorithm for Field Broccoli Seedling Transplanting Quality Based on YOLOv7-Tiny

Tengfei Zhang [1,2], Jinhao Zhou [1,2], Wei Liu [1,2], Rencai Yue [1,2], Mengjiao Yao [1,2], Jiawei Shi [1,2] and Jianping Hu [1,2,*]

1   School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China;
    2112016001@stmail.ujs.edu.cn (T.Z.); 2112316042@stmail.ujs.edu.cn (J.Z.); mario_liu@ujs.edu.cn (W.L.);
    2111916008@stmail.ujs.edu.cn (R.Y.); 2221916042@stmail.ujs.edu.cn (M.Y.); 2112116016@stmail.ujs.edu.cn (J.S.)
2   Jiangsu Provincial Key Laboratory of Hi-Tech Research for Intelligent Agricultural Equipment, Jiangsu
    University, Zhenjiang 212013, China
*   Correspondence: hujp@ujs.edu.cn; Tel.: +86-511-88797338

Abstract: The rapid and accurate detection of broccoli seedling planting quality is crucial for the implementation of robotic intelligent field management. However, existing algorithms often face issues of false detections and missed detections when identifying the categories of broccoli planting quality. For instance, the similarity between the features of broccoli root balls and soil, along with the potential for being obscured by leaves, leads to false detections of "exposed seedlings". Additionally, features left by the end effector resemble the background, making the detection of the "missed hills" category challenging. Moreover, existing algorithms require substantial computational resources and memory. To address these challenges, we developed Seedling-YOLO, a deep-learning model dedicated to the visual detection of broccoli planting quality. Initially, we designed a new module, the Efficient Layer Aggregation Networks-Pconv (ELAN_P), utilizing partial convolution (Pconv). This module serves as the backbone feature extraction network, effectively reducing redundant calculations. Furthermore, the model incorporates the Content-aware ReAssembly of Features (CARAFE) and Coordinate Attention (CA), enhancing its focus on the long-range spatial information of challenging-to-detect samples. Experimental results demonstrate that our Seedling-YOLO model outperforms YOLOv4-tiny, YOLOv5s, YOLOv7-tiny, and YOLOv7 in terms of speed and precision, particularly in detecting 'exposed seedlings' and 'missed hills'-key categories impacting yield, with Average Precision (*AP*) values of 94.2% and 92.2%, respectively. The model achieved a mean Average Precision of 0.5 (*mAP@0.5*) of 94.3% and a frame rate of 29.7 frames per second (FPS). In field tests conducted with double-row vegetable ridges at a plant spacing of 0.4 m and robot speed of 0.6 m/s, Seedling-YOLO exhibited optimal efficiency and precision. It achieved an actual detection precision of 93% and a detection efficiency of 180 plants/min, meeting the requirements for real-time and precise detection. This model can be deployed on seedling replenishment robots, providing a visual solution for robots, thereby enhancing vegetable yield.

Keywords: missed hills; exposed seedling; partial convolution; seedling replenishment robot

1. Introduction

Vegetables are essential in daily diets. Data from the United Nations' Food and Agriculture Organization show that China leads globally, with 52.25% of the world's vegetable planting area and 58.31% of its total production [1]. With the increasing adoption of smart agricultural technologies, traditional methods of vegetable production are evolving. The use of transplanting machines, in particular, has greatly improved the efficiency of vegetable cultivation [2,3]. However, when using transplanting machines for vegetable cultivation, instances of substandard planting quality arise, including issues like excessive planting depth (covered seedlings), inadequate depth (exposed seedlings), and missed

hills [4]. Factors contributing to substandard planting quality include mechanical design [5–7], agronomy [8,9], and various environmental aspects related to the field. Vavrina, et al. [10] evaluated the impact of transplanting depth on tomato and bell pepper yields, revealing that transplanting up to the first true leaf or cotyledon results in greater yields than transplanting to the top of the stem. As shown in Figure 1, currently, the process of detecting and replanting seedlings with substandard planting quality primarily relies on manual labor. This method is marked by inconsistent standards and demands a significant amount of work. A major challenge in transitioning from manual to mechanized replanting is the development of effective target detection algorithms [11]. The speed and precision of these algorithms are critical, as they directly influence the efficiency of the robots and the yield of field-grown vegetables. Therefore, this study initially categorizes the conditions of seedlings that impact yield and aims to develop fast and accurate detection algorithms for these specific categories.

Figure 1. Manual Replanting of Broccoli.

The current prevalent technologies for field detection include machine vision [12,13], ultrasonic sensor detection [14], and 3D Light Detection and Ranging (LiDAR) detection [15,16]. Ultrasonic sensors and 3D LiDAR can detect the presence of vegetable seedlings within an area, yet they face difficulties in accurately distinguishing the planting quality of these seedlings. Machine vision technology, known for its capability to capture comprehensive, precise, and intelligent information, demonstrates significant potential in target detection of broccoli seedling planting quality [17,18].

Currently, deep learning is widely applied in the field of agricultural detection [19]. Scholars worldwide focus mainly on using deep learning to detect missing seedlings in seedling planting quality assessments, with less emphasis on detecting planting depth. Lin, et al. [20] developed a detection model for field peanut seedlings, combining an improved YOLOv5s with DeepSort, and utilized drones for seedling emergence detection. Although efficient, this model fails to locate non-emerged seedlings and cannot assess planting depth quality. Cui, et al. [21] enhanced the YOLOv5s by adjusting its detection head structure and incorporating a transformer, developing a rice missing seedling detection and counting model with a precision of 93.2%. Wu, et al. [22] improved YOLOv5s by replacing its Neck network with the Slim-Neck network, developing a sugarcane field missing seedling detection model and proposing a method for predicting replanting locations. However, this model tends to miss detecting small sugarcane seedlings, presenting limitations for the detection of the "Covered seedling" category in our task. Zhang, et al. [23] replaced the

upsampling module in the neck network of YOLOv5s with the Content-aware ReAssembly of Features (CARAFE) module, enhancing the performance in detecting small targets.

For the detection of complex multi-target tasks such as broccoli planting quality assessment, deep learning models need to achieve high precision across each category. Zhao, et al. [24] developed a deep learning model for grading vegetable seedlings, utilizing ShuffleNet Version 2 (ShuffleNet-V2) as the backbone network for feature extraction and integrating the Efficient Channel Attention (ECA) attention mechanism. This model achieved high precision in categorizing seedlings as weak, damaged, or strong, with a precision rate of 94.23%. Attention mechanisms enable network models to focus on relevant areas within local information. Commonly used attention mechanisms include Squeeze-and-Excitation (SE) [25,26], Convolutional Block Attention Module (CBAM) [27], and Coordinate Attention (CA) [28]. SE focuses solely on channel information, overlooking spatial information, whereas CBAM employs global pooling operations to capture local spatial information. CA, on the other hand, maintains channel information while concentrating on long-range spatial information in feature maps. Zhu, et al. [29] integrated the CA mechanism with YOLOX-s, enhancing the network's focus on regions of interest and effectively improving the detection precision for corn silk obscured by leaves.

To address challenges such as lower algorithm recognition rates and weak robustness in natural environments, Sun, et al. [30] focused on the detection of broccoli seedlings. They proposed a method based on the Faster Region-based Convolutional Neural Network (Faster R-CNN) model, achieving a recognition precision of 91.73% with an average detection time of 249 ms. While two-stage detection models like Faster R-CNN [31] offer higher precision, they also have slower image processing times. In contrast, one-stage detection models, such as those in the YOLO series [32], bypass the candidate region selection stage and directly treat object detection as a regression task, facilitating end-to-end detection. In 2022, the novel YOLOv7 architecture was introduced, outperforming all known object detectors within a performance range of 5 to 160 fps [33]. Among its variants, YOLOv7-tiny maintains the cascade-based model scaling strategy of YOLOv7 and features improvements in the Efficient Long-Range Aggregation Network (ELAN) [34]. YOLOv7-tiny employs a more compact network architecture and an optimized training strategy. By reducing model parameters and computational requirements, it offers a viable solution for target detection in computationally constrained environments.

For mobile deployment in field environments, two primary methods are typically used to reduce network model weights: (1) Utilizing lightweight architectures with fewer parameters, such as MobileNet [35], ShuffleNet [36,37], and GhostConv [38], which decrease the parameter count while minimizing performance loss. In precision agriculture, attention-based lightweight models are often used in network models that require high accuracy but fewer parameters [39]. (2) Implementing techniques like sparse training and model pruning to further reduce the model's parameters and computational demand. In addressing the precise identification and localization of cabbage, Zhai, et al. [40] evaluated Faster R-CNN, Single Shot MultiBox Detector (SSD), and YOLOv5. They opted for YOLOv5s as the base model and implemented lightweight modifications using MobileNet V3s. This model achieved a recognition precision of 93.14% with an image processing time of 54.09 ms, marking a 26.98% reduction in processing time compared to the base model. However, the model demonstrated reduced precision in detecting small cabbages with missing leaves, particularly noticeable post-transplantation. Moreover, these studies have primarily concentrated on inter-class classification, with the nuanced task of intra-class fine-grained detection still presenting a significant challenge. Ref. [41] enhanced YOLOv3-tiny with Path Aggregation Network (PANet) and Spatial Attention Module (SAM) for hierarchical tomato seedling detection, effectively distinguishing no-seedlings, weak, and healthy seedlings. However, research in this area is predominantly performed in stable conditions.

In summary, current research primarily focuses on detecting missing seedlings, mainly for assessing crop yield, with limited attention given to the detection of seedlings with improper planting depth. This issue has emerged due to the transition from semi-automatic to fully automatic transplanting machines, where the manual process of picking and placing seedlings has been replaced by mechanical arms, leading to instances of substandard planting quality, a new and common phenomenon. Ensuring vegetable yield necessitates the detection of these poorly planted seedlings. Existing algorithms require significant computational resources and memory, and they face limitations in recognizing categories with similar features and small targets. For example, the similarity between the features of broccoli root balls and soil clods, the ease with which root balls can be obscured by leaves, the resemblance of "missed hill" features to the background, and the small size of features in the "Covered Seedling" category, present significant challenges. To address these challenges, our contributions are as follows: (1) We proposed a method for target detection and classification specifically for broccoli planting quality, categorizing the planting quality into "qualified seedlings", "exposed seedlings", "covered seedlings", and "missed hills", and created a dataset for this research, thereby contributing exploratory work to the field of vegetable planting quality detection. (2) We developed the Seedling-YOLO deep learning model for identifying substandard broccoli planting quality in the field. (3) We introduced the ELAN_P module for the backbone network, which reduces model parameters without sacrificing precision. Furthermore, by integrating CARAFE and CA, we addressed issues of false and missed detections, especially prevalent in "exposed seedlings", "covered seedlings", and "missed hills".

2. Materials and Methods

2.1. Broccoli Seedling Planting Quality Related

The classification of broccoli seedlings in this study is based on their planting status, which plays a crucial role in their subsequent growth and development. We referred to the transplanter performance experiment and the study of the impact of vegetable planting depth on yield [10,42]. As depicted in Figure 2, we classified broccoli seedling planting quality into four types: (1) Qualified seedlings: Properly planted with adequate depth covering the root ball without reaching the cotyledons or first true leaves; (2) Exposed seedlings: The root ball of the seedling is exposed on the ground; (3) Covered seedlings: Planted too deep, with the depth exceeding the first true leaf or the top of the seedling stem; (4) Missed hills: Absence of seedlings in the designated planting locations within the established inter-plant spacing. In this study, we developed a broccoli planting quality target detection algorithm based on four defined classification situations. The development process and key steps of this algorithm are illustrated in Figure 3. Initially, we conducted broccoli planting experiments using three different types of transplanting machines. Following this, we collected image data regarding the quality of broccoli planting and built a dataset through data augmentation techniques and the use of Labelimg software (version 1.8.6; tzutalin, 2021). Moreover, we developed the Seedling-YOLO model and compared it with four currently high-performing algorithms. Finally, we evaluated the model's performance using validation sets and field experiments, discussing the advantages, limitations, future directions, and the potential application of the model in mechanized replanting.
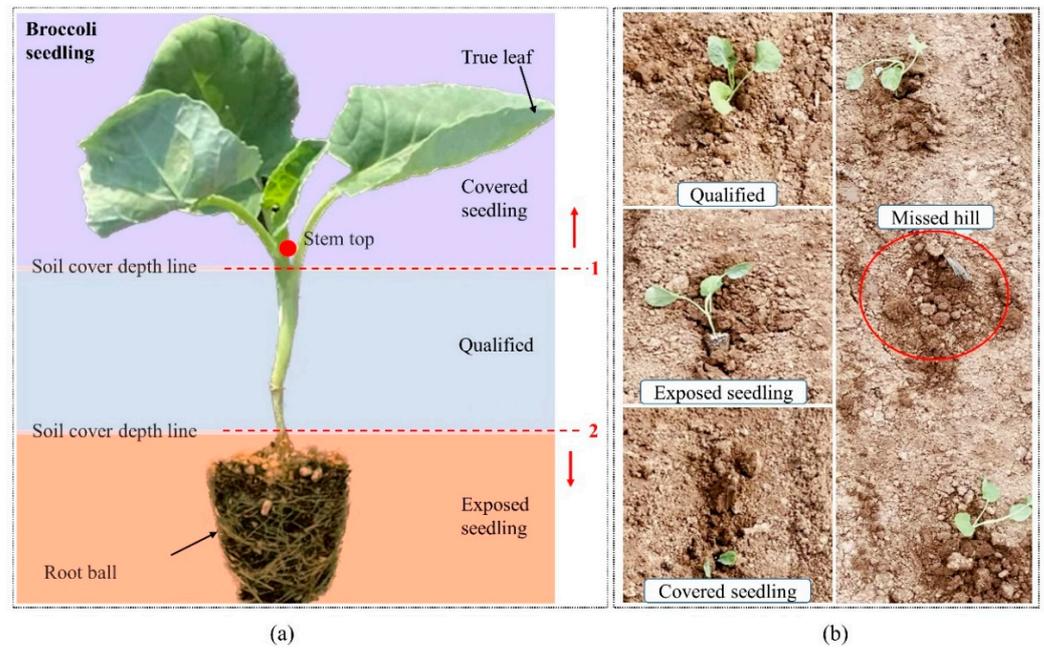
Figure 2. Broccoli seedling planting quality classification standards: (**a**) Diagram of soil covering depth for broccoli seedling planting; (**b**) Examples of different planting qualities of broccoli seedlings. Note: (**a**) Lines 1 and 2 in the figure represent the critical lines of soil coverage depth. Arrows above dashed line 1 indicate that the soil covering depth exceeds critical line 1, representing a Covered seedling. Arrows below dashed line 2 indicate that the soil covering depth is below critical line 2, representing an Exposed seedling.
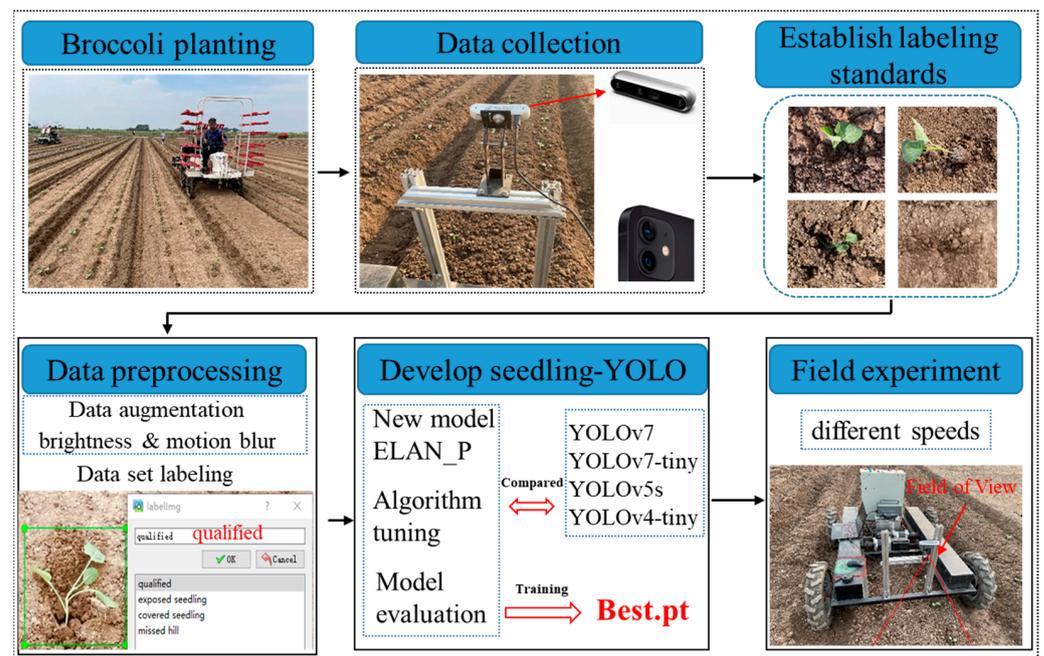

Figure 3. Flowchart of this study.

2.2. Dataset Construction and Image Preprocessing

From 21 August to 26 August 2022, data were collected at the demonstration base for whole-process mechanization of broccoli production in Xiangshui County, Jiangsu Province, China. These data originated from broccoli plants transplanted using three different types of machines: the Yanmar (2ZQ-2) vegetable transplanting machine (Yanmar Co., Ltd., Osaka,

Japan), the Jiangsu University (2ZBA-2) automatic transplanting machine [43] (Jiangsu University, Zhenjiang, China), and the AMEC (2ZS-2)vegetable transplanting machine [2] (AMEC, Changzhou, China). The seedlings were derived from the local broccoli seedling base with an age of 28–32 days, an average height of 12.4 cm, and 3–5 leaves. During the image acquisition phase, we utilized a vision platform equipped with the Intel RealSense D455 (Intel Corporation, Santa Clara, CA, USA) for video recording. The RGB resolution is 1920 × 1080 pixels. To minimize interference from non-target background in the field of view, the camera was fixed at a height of 0.6 m. Given the requirement for the developed detection model to adapt to the seedling detection needs of a replanting robot, moving through various angles and moments in time, we set the camera's installation angles at 45° and 90° for data collection. Additionally, we utilized an iPhone 12 (Apple Inc., Cupertino, CA, USA) to collect single-plant images of broccoli seedling planting quality, aiming to capture the fine-grained features of the broccoli. The main camera resolution of the iPhone 12 is 4032 × 3024 pixels. This dataset comprises images and video clips of broccoli taken under various lighting conditions. Figure 4 displays a portion of this dataset.
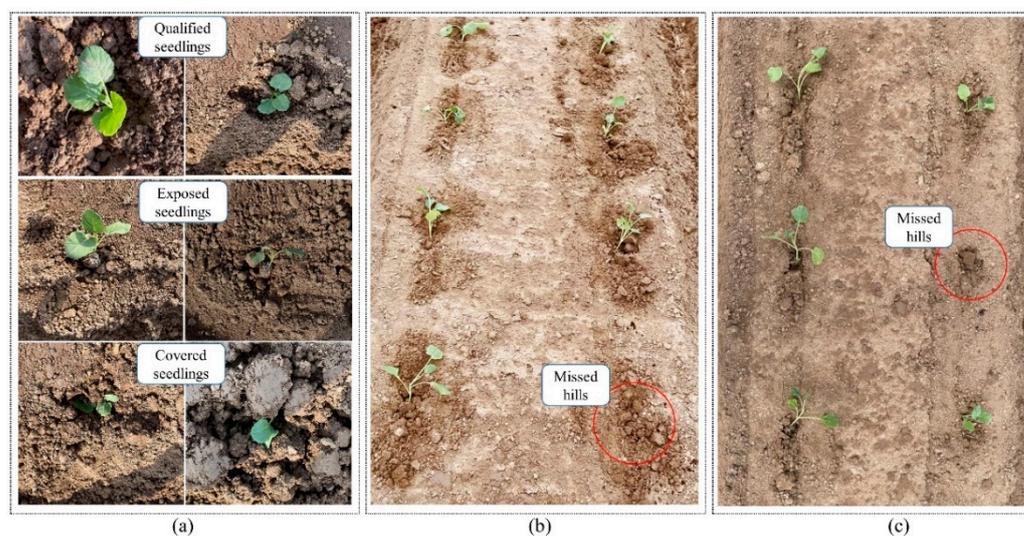
Figure 4. Example images of broccoli dataset: (**a**) Close-up View; (**b**) 45° Angle View; (**c**) 90° Angle View.

Data augmentation addresses sample imbalance and enhances the diversity of training samples. It compels the model to learn more robust features and significantly improves the model's generalization capabilities. We used the ImgAug 3.2 software for image augmentation available at: https://github.com/Fafa-DL/Image-Augmentation (URL, accessed on 08 March 2023). To enrich the dataset and prevent model overfitting, offline augmentation is employed through brightness adjustment and motion blur addition, resulting in a total of 6000 images in the dataset. Labeling is performed using the LabelImg 1.8.6 software available at: https://github.com/tzutalin/labelImg (URL, accessed on 09 March 2023). The labeled objects are categorized into four classes, adhering to the COCO dataset format. The dataset is partitioned into training, testing, and validation sets in an 8:1:1 ratio. Specific information is shown in Table 1, the training set is used for network parameter training, the testing set assesses the model's generalization error, and the validation set optimizes hyperparameters utilized during training, thereby enhancing the model's performance.

Table 1. Dataset breakdown for training, validation, and testing with class labels and image counts.

Set	Target Box				Number of Images
	Qualified	Exposed	Covered	Missed	
Train	3600	3300	2940	3150	4800
Validation	460	416	362	395	600
Test	475	428	342	382	600
Total	4535	4144	3644	3927	6000

2.3. Improvement of YOLOv7-Tiny

The improved network structure, including the ELAN_P module in the backbone network for efficient feature extraction and the inclusion of CA in the Neck network, is depicted in Figure 5. Additionally, the CARAFE operator is utilized for upsampling in the model.
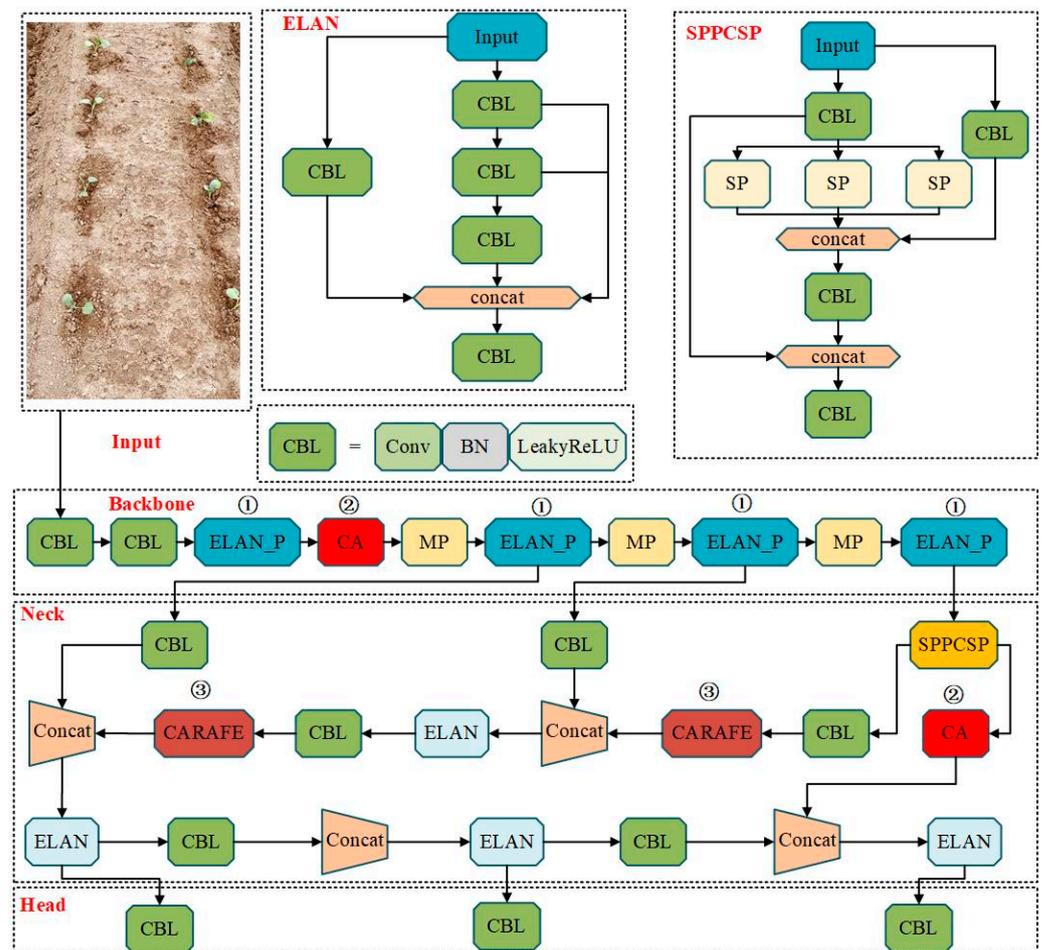
Figure 5. Network structure of the Seedling-YOLO. Note: ①–③ in the figure are the specific locations of the improved modules.

2.4. Efficient ELAN-P Block

Partial convolution involves applying convolutional operations only to valid pixels while disregarding or masking out invalid or missing pixels [44]. In image tasks, conventional convolutions treat missing pixels as zero or entirely ignore them. In contrast, partial convolution dynamically determines the contribution of each valid pixel to the output, taking into account the absence or corruption of pixels, thereby addressing this issue. The working principle of partial convolution is shown in Figure 6. For the convolution opera-

tion of size $H \times W \times C$ broccoli seedling image, it only needs to apply regular convolution on a part of the input channel for spatial feature extraction and keep the rest of the channels unchanged. It only uses the C_p channel for spatial feature extraction. Therefore, the FLOPs of Pconv are as Formula (1).

$$FLOPs = W \times H \times C_p^2 \times k^2 \tag{1}$$

For a typical r = C_p/C = 1/4, the FLOPs of PConv are only 1/16 of those of ordinary convolutions.
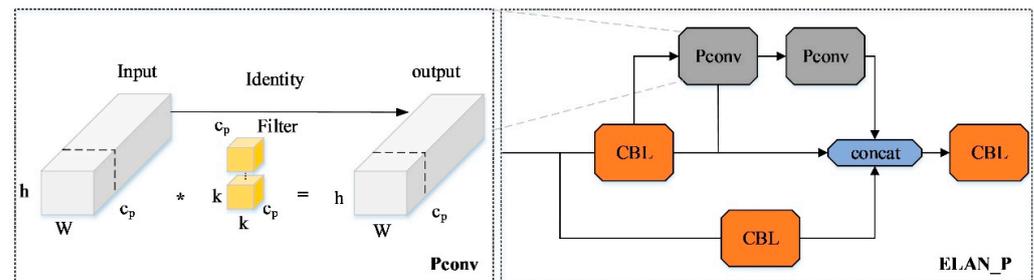
Figure 6. The structure of the Pconv and ELAN_P. Note: In the figure, the asterisk (*) represents the convolution operation, and the dashed lines indicate the replacement of the standard CBL block with a Pconv to enhance feature extraction and reduce the number of parameters.

By integrating Pconv into the ELAN module, our model efficiently processes valid pixels in the feature map and addresses missing pixels with advanced padding or restoration techniques. This integration, particularly within the ELAN_P module, improves the handling of incomplete data, thereby boosting the model's robustness and stability. The ELAN_P module's use of partial convolution optimization reduces computational demands and memory usage.

2.5. Content-Aware ReAssembly of Features

Conventional upsampling techniques, such as nearest-neighbor and bilinear interpolation, often fall short in complex tasks like detecting broccoli planting quality, as they do not utilize the semantic context within feature maps. To overcome this, we adopted the CARAFE operator, which adaptively performs upsampling by leveraging spatial information, thus enhancing the detail and texture preservation in upsampled images. Figure 7 depicts the CARAFE network architecture [45]. The CARAFE module consists of two key components: The Upsampling Kernel Prediction Module and the Feature Recombination Module. The first component generates content-aware recombination kernels, optimizing the process through channel compression and convolution operations to balance performance and efficiency. The resulting feature map is restructured and normalized for effective use in upsampling. In the Feature Recombination Module, features within local regions of the input are reorganized using recombination kernels. This process involves weighted summation based on the kernel and the position within the feature map, allowing for precise and context-aware reassembly of features. The combined efforts of these modules result in more accurate and detailed feature representation in the upsampling process, crucial for tasks like detecting the quality of broccoli planting.
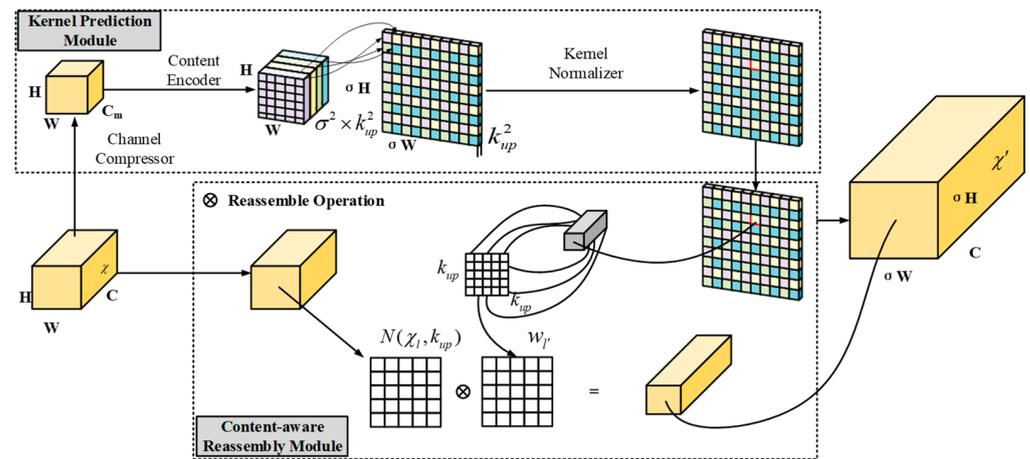
Figure 7. CARAFE network structure.

2.6. Integrating the CA Attention Mechanism

Location information is the key to capturing target features in visual tasks, and CA brings significant advantages to this task [28]. CA emplaces spatial information of seedlings into channel attention and is able to capture long-range dependencies in one spatial direction while retaining precise location information in the other. Its structure diagram is shown in Figure 8. The input feature map uses pooling kernels of size (*H*, 1) and (1, *W*) to encode each channel along the horizontal coordinate direction and vertical coordinate direction. The output expressions of the *C* channel with height *H* and width *W* are obtained as (2) and (3).

$$z_c^h(h) = \frac{1}{W} \sum_{0 \le i < w} x_c(h, i) \tag{2}$$

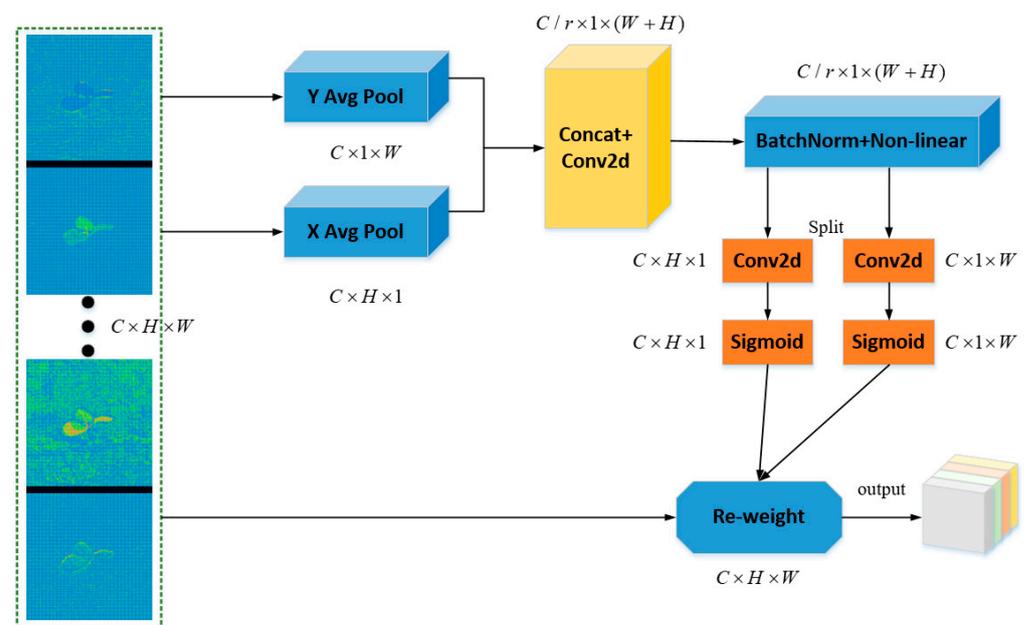$$z_c^w(w) = \frac{1}{H} \sum_{0 \le i < H} x_c(j, w) \tag{3}$$

Figure 8. Coordinate attention network structure.

In Equation (2), $z_c^h(h)$ represents the output of the c-th channel with a height of h, while $x_c(h, i)$ signifies the feature vector of the i-th row. Similarly, in Formula (3) $z_c^w(w)$ denotes the output of the c-th channel with a width of w, and $x_c(j, w)$ stands for the feature vector of the j-th column. Employing a one-dimensional global pooling operation aggregates features along the two spatial directions. This facilitates the attention module to capture spatially relevant relationships along one direction while preserving positional information from the other direction. In the attention generation phase, two feature images are first concatenated. Then, a 1×1 convolution is used to reduce the number of channels, followed by a nonlinear activation. The resulting output is split along the spatial dimension into horizontal and vertical attention tensors. In this process, two sets of 1×1 convolutions are employed to increase the channel count of the image, and nonlinear activation is applied using the Sigmoid function. Finally, the generated attention images are element-wise multiplied with the input feature image to implement the application of CA.

2.7. Model Training and Evaluation

2.7.1. Model Training

In this study, the experimental hardware platform consisted of an Intel i7-13700KF CPU running the Windows 10 operating system, along with an RTX 4080 GPU and 32 GB of RAM. The code was written using the PyTorch 1.13.0 deep learning framework and developed using Python 3.8.15. To achieve optimal performance, hyperparameter tuning was conducted, and the specific hyperparameter settings are presented in Table 2.

Table 2. Hyperparameter settings for network training.

Parameters	Values
Lr0	0.01
Momentum	0.937
Weight decay	0.0005
Epochs	300
Batch size	16
Pre-trained weight	YOLOv7-tiny.pt

In order to improve the o.

2.7.2. Model Evaluation

For the evaluation of broccoli seedling planting quality, precision (P) and recall (R) were used to measure the precision and completeness of model detection. P represents the proportion of all samples judged to be positive by the model, which are truly positive. Its calculation formula is (4):

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

(4)

Among them, TP represents true positive cases and FP represents false positive cases. R measures the model's ability to detect positive samples. Its calculation formula is (5):

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

(5)

Among them, FN stands for false negative example. In addition, the average precision (AP) and mean average precision (mAP) are used to comprehensively evaluate the performance of the model on different categories. AP measures the precision of the model on a single class (6), while mAP takes the average AP of all classes. Its calculation formula is (7):

$$AP = \int_0^1 P(R)dR$$

(6)

$$mAP = \frac{1}{C}\sum_{i=1}^{C} AP_i$$

(7)

mAP@0.5 means that when the IOU value and ground truth value of the detection frame are greater than 0.5, the sample is considered positive. As a comprehensive evaluation index of model precision. The number of parameters(#param.), FLOPs, and FPS of the model give insights into its size, computational load, and real-time processing ability, all vital for designing resource-efficient planting robots. By analyzing these metrics together, we were able to gain insight and evaluate the overall performance and efficiency of the model.

3. Results

### 3.1. Training Loss Function Analysis

To ensure a fair comparison, both models were trained under identical environmental and parameter conditions. Figure 9 presents the loss function graphs for Seedling-YOLO and the baseline model. It is observed that Seedling-YOLO's loss on the training set decreases more rapidly compared to the baseline model, though the overall convergence pattern is similar. On the validation set, the baseline model's loss value shows a slightly quicker decrease than the improved model in the initial 150 epochs. However, post 150 epochs, the improved model demonstrates a more significant reduction in loss value. Ultimately, the loss value stabilizes after 250 iterations. The model effectively learns image features and converges to an optimal solution.
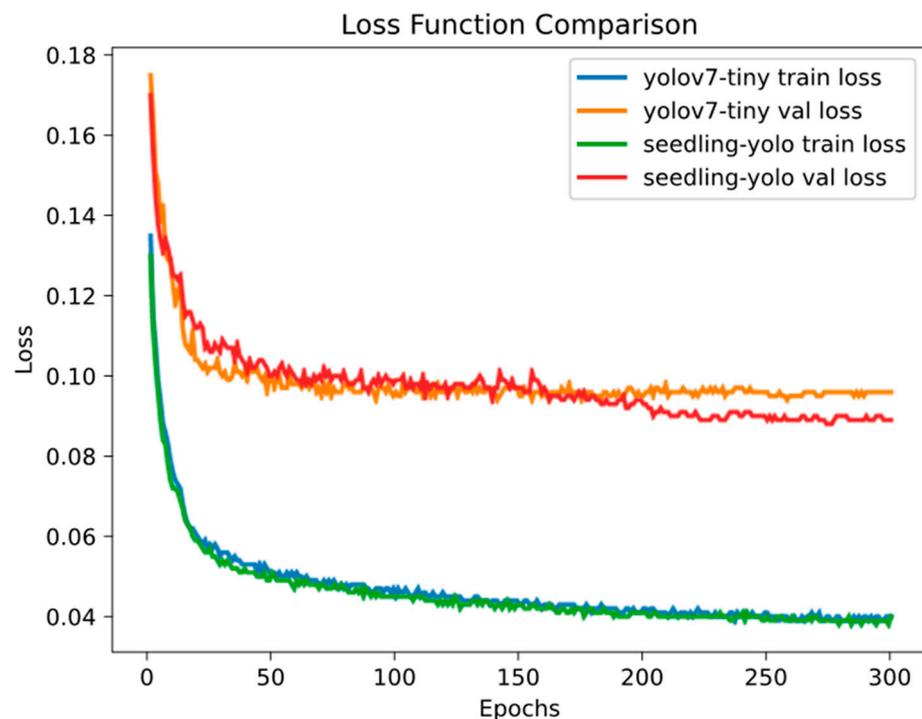
Figure 9. Comparison of improved model and baseline loss function curve.

3.2. Model Performance and Comparison with State-of-the-Art

It can be seen from Figure 10 that the improved model has the best detection effect on qualified seedlings and exposed seedlings, with *AP* values of 95.9% and 95.4%, respectively. In addition, better results were obtained for covered seedlings and missed hills, with *AP* values of 93.7% and 92.2%, respectively. The improved model achieved the highest *mAP@0.5*, reaching 94.5%.
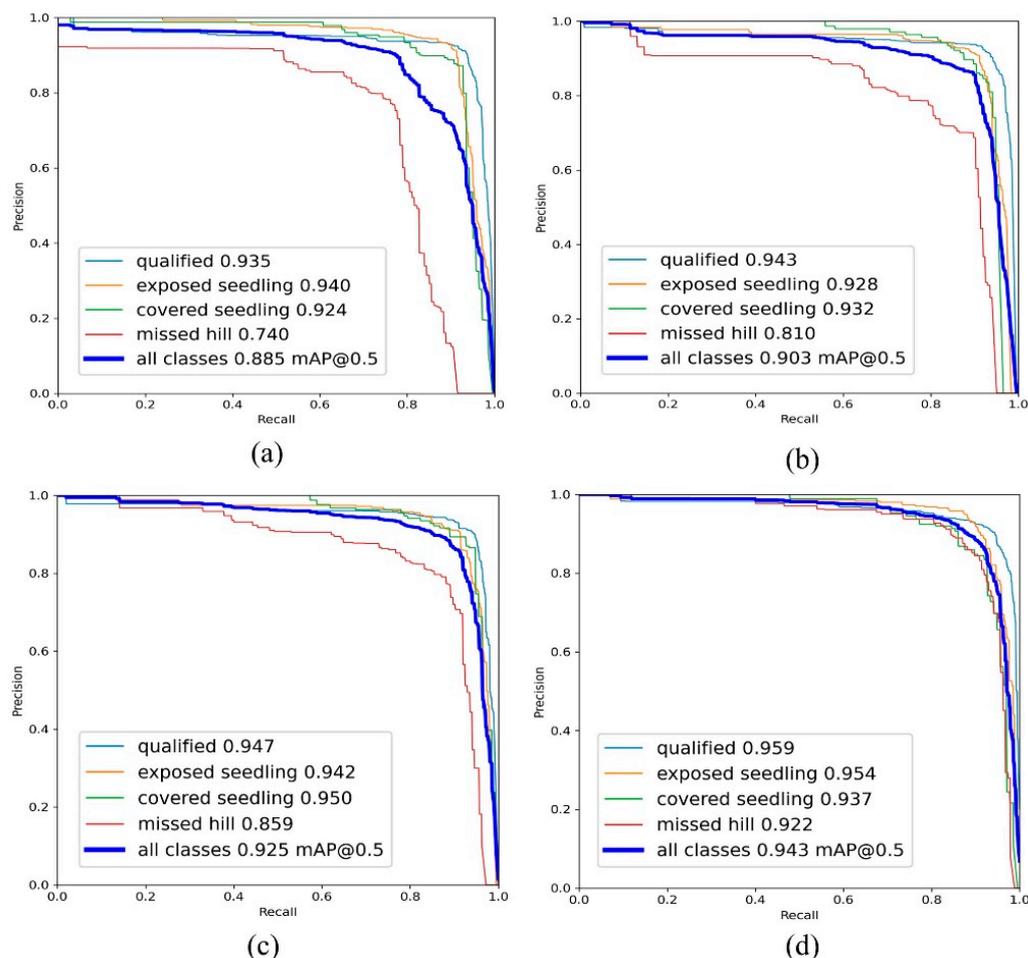
Figure 10. Different models P-R curves: (**a**) YOLOv5s; (**b**) YOLOv7-tiny; (**c**) YOLOv7; (**d**) Seedling-YOLO.

As indicated by Table 3, YOLOv7 achieved the highest *P* of 87.8%, *R* of 91.1%, and *mAP@0.5* of 92.5% when compared with YOLOv7-tiny, YOLOv5s, and YOLOv4-tiny. However, YOLOv7 also has the largest number of parameters and FLOPs, as well as the lowest FPS, suggesting that while YOLOv7 offers superior detection performance, it is also the slowest in detection speed. Among these four existing algorithms, YOLOv7-tiny presents the best balance of accuracy and speed performance, which is why we chose to enhance it. The improved model, Seedling-YOLO, has 4.98 M parameters and 11.6 G FLOPs, the lowest among all models, denoting a more lightweight architecture. Additionally, with the highest FPS of 29.7, it also proves to be the fastest. Moreover, Seedling-YOLO achieves the best detection performance with a *P* of 91.3%, *R* of 92.1%, and *mAP@0.5* of 94.3%. As shown in Table 4, our Seedling-YOLO model, compared to YOLOv7-tiny, demonstrates improvements in both precision and recall rates for the "exposed seedlings" and "missed hills" categories. The most significant improvement is observed in the "missed hills" category, with an increase of 10.9% in precision and a 9% increase in recall rate. The *AP* value for "missed hills" saw the largest increase, reaching 11.2%.

Table 3. Comparison of the experimental findings produced by various algorithms.

| Networks	#Param.	FLOPs	*P*	*R*	*mAP@0.5*	FPS
YOLOv7	37.2 M	105.2 G	87.8%	91.1%	92.5%	24.9
YOLOv7-tiny	6.23 M	13.9 G	87.5%	90.1%	90.3%	26.5
YOLOv5s	7.27 M	17.2 G	87.2%	87.4%	88.5%	25.8
YOLOv4-tiny	6.07 M	13.2 G	87.1%	86.9%	87.7%	26.7
Seedling-YOLO	4.98 M	11.6 G	91.3%	92.1%	94.3%	29.7

Table 4. Seedling-YOLO vs. YOLOv7-tiny: precision and recall in four broccoli planting quality categories.

| Set	YOLOv7-Tiny			Seedling-YOLO		
	P	*R*	*AP*	*P*	*R*	*AP*
Qualified seedling	91.3%	94.1%	94.3%	91.5%	93.9%	95.9%
Exposed seedling	89.3%	92.4%	92.8%	91.8%	92.6%	95.4%
Covered seedling	89.7%	93.1%	93.2%	91.3%	92.1%	93.7%
Missed hill	79.7%	80.8%	81.0%	90.6%	89.8%	92.2%

In order to test the detection effect of the Seedling-YOLO algorithm, a verification test is carried out on the test set. Based on the consideration of the recall rate and experimental comparison effect, the confidence level is set to 0.5. Figure 11 displays a selection of detection results from the improved model, illustrating successful detection across all four challenging categories. Figure 12 offers a comparative analysis of detection between the improved model and the baseline. In Figure 12A,B, it is evident that YOLOv7-tiny incorrectly identifies "exposed seedlings" as qualified seedlings and misinterprets the background seedling leaves as "covered seedlings." In Figure 12C, the model mistakenly recognizes the background as "missed hills," and in Figure 12D, the similarity between seedling root balls and soil clods, coupled with obstructions, leads to the model's inability to accurately classify the specific category. Additionally, the model exhibits low precision in detecting small target seedlings. In contrast, the improved model correctly identifies these challenges, and its overall score and regression performance surpass that of the original model. Using Grad-CAM [46] as a visualization tool, it is evident from Figure 13 that our CA mechanism is able to locate the object of interest more precisely compared to the base network.

(a)　　　　　　　　　　　(b)

(c)　　　　　　　　　　　(d)
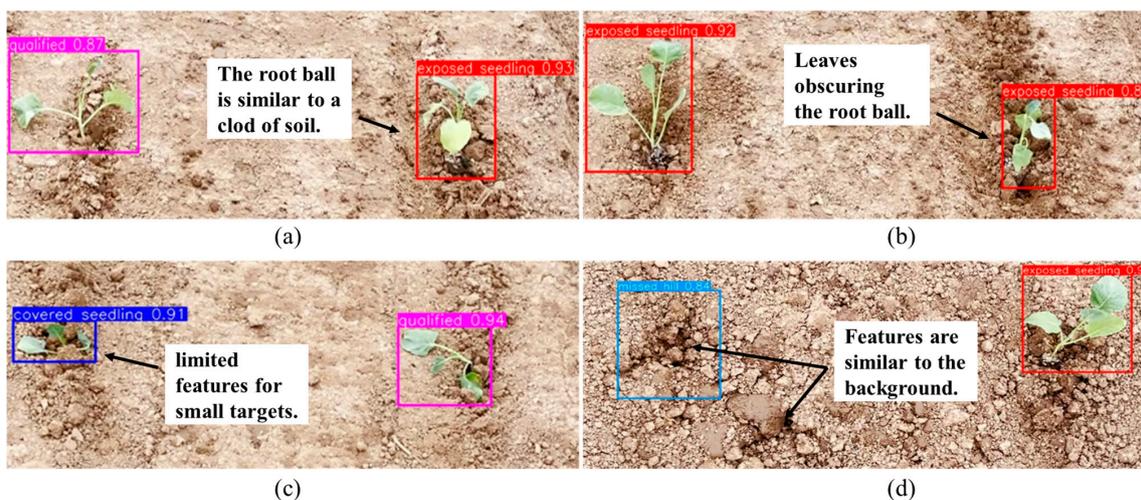
Figure 11. Improved model detection performance: (**a**) Detection performance for root balls similar to soil clods; (**b**) Detection effectiveness for obscured root balls; (**c**) Detection performance for "covered seedlings"; (**d**) Detection effectiveness for background similarity. Note: The figures (**a**–**d**) show the detection performance on various types of challenging detection samples.
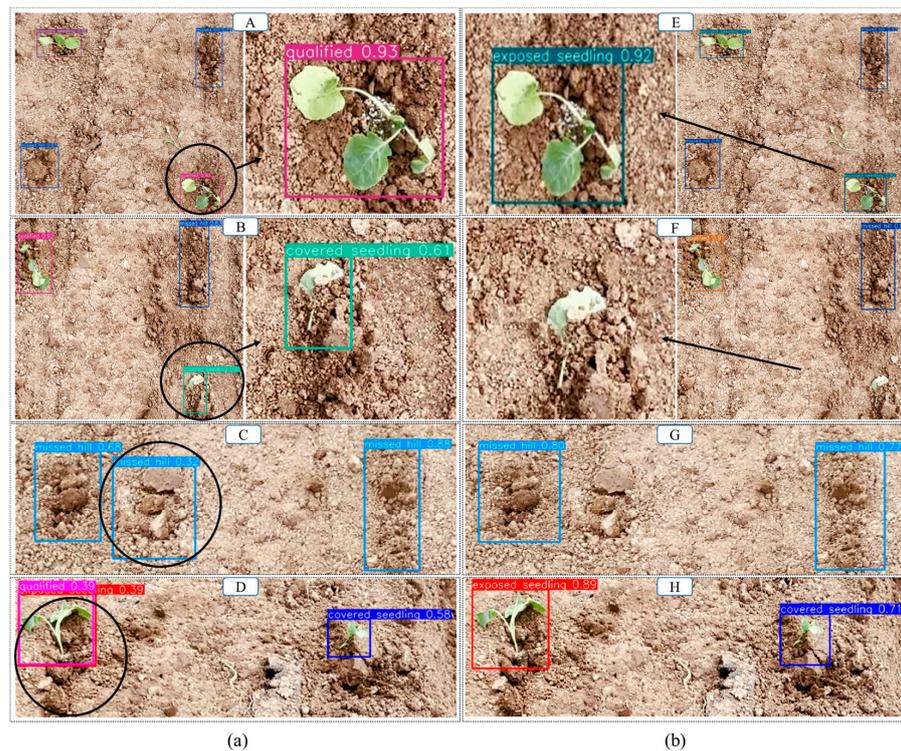
Figure 12. Comparative analysis of detection performance between Seedling-YOLO model and YOLOv7-tiny: (**a**) Detection Effectiveness of YOLOv7-tiny; (**b**) Detection Effectiveness of Seedling-YOLO. Note: (**A**–**H**) in the figure show the comparison of detection performance on various challenging detection samples. Black circles (○) in the diagram represent false detections.
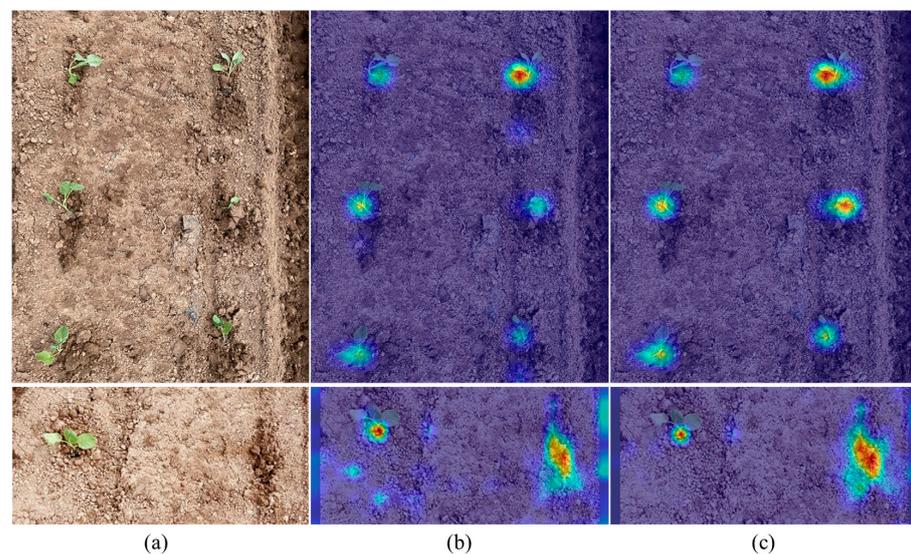
Figure 13. Add CA attention and compare with baseline heat map: (**a**) Initial image; (**b**) YOLOv7-tiny; (**c**) YOLOv7-tiny-CA. Note: Blue represents areas of lower feature activation. Red and its gradients indicate the importance of the areas of interest to the model, where the deeper red signifies the areas that the model focuses on more intensely.

3.3. Ablation Experiment

In order to study the contribution of ELAN_P, CARAFE and CA to Seedling-YOLO, the ablation experiment was conducted for verification, and the data were uniformly processed using the YOLO.py script. The experimental results are shown in Table 5. Among them, the architecture information of the eight models is as follows: M1 is the original YOLOv7-tiny;

M2 is the YOLOv7-tiny backbone network using ELAN_P; M3 incorporates CARAFE into YOLOv7-tiny; M4 integrates CA into YOLOv7-tiny; M5 combines YOLOv7-tiny with both ELAN_P and CARAFE; M6 adds ELAN_P and CA to YOLOv7-tiny; M7 fuses CARAFE and CA into YOLOv7-tiny; M8 integrates all three modules—ELAN_P, CARAFE, and CA—into YOLOv7-tiny. It can be seen from the table that the model parameters and GFLOPs decreased by 20.8% and 27.3%, respectively, after the M2 adopted the ELAN_P module integrated with Pconv, and the *mAP@0.5* increased by 2.4%. M3 and M4 significantly improve the detection precision with a small increase in the amount of parameters. When ELAN_P integrates CARAFE or CA, it plays a positive role in promoting precision, and M5 and M6 have increased by 3.6% and 2.9% respectively. CARAFE decreased by 0.3% compared with M4 after CA fusion compared with M7. In the end, the fusion of the three modules of YOLOv7-tiny played a positive role in the precision of the model. Compared with the basic model, the parameters of the improved M8 were reduced by 20%, FLOPs were reduced by 16.5%, and *mAP@0.5* reached its highest of 94.3%.

Table 5. Comparisons of ablation experiments.

Model	ELAN_P	CARAFE	CA	#Param.	FLOPs	*P*	*R*	*mAP@0.5*
M1	×	×	×	6.23 M	13.9 G	87.5%	90.1%	90.3%
M2	√	×	×	4.93 M	10.1 G	89.5%	90.9%	92.7%
M3	×	√	×	6.33 M	14.0 G	89.4%	91.8%	93.2%
M4	×	×	√	6.2 M	13.9 G	90.1%	91.5%	93.9%
M5	√	√	×	5.04 M	10.9 G	90.4%	91.2%	93.9%
M6	√	×	√	4.94 M	10.8 G	89.8%	89.1%	93.2%
M7	×	√	√	6.35 M	14.0 G	90.3%	88.4%	93.6%
M8	√	√	√	4.98 M	11.6 G	91.3%	92.1%	94.3%

"×" and "√" indicate that the improved method was not and was applied in the model, respectively.

3.4. Experimental Results at Different Speeds

In order to verify the effectiveness and efficiency of the quality object detection model for broccoli seedling transplanting at different speeds, we developed a vision platform for field experiments. As shown in Figure 14, the experimental environment is a field with a wide view and clear weather. The distance between broccoli plants is 0.4 m. In this experiment, the speed of the vehicle is controlled by adjusting the motor speed, and the speed can be adjusted from 0.2 to 1.2 m/s. We compared the detection performance at different speeds. To evaluate the model's effectiveness on hard-to-recognize samples, we also selected more complex ridge surfaces for validation, as shown in Figure 15. When the speed is 0.3 m/s, the model has a high recognition precision and recall rate, and it also has a good recognition effect on complex ridges. As the speed increases to 0.6 m/s, the score decreases slightly. When speed continues to increase to 0.7 m/s, missed detection and false detection of each category begin to occur, and the detection frame score further decreases. The circles in the figure are false detections, and the triangles are missed detections. In the experiment, we conducted statistics based on the number of missed and false detections. The model's detection performance on a total of 200 objects, comprising 100 from normal ridges and 100 from complex ridge surfaces, was evaluated using precision and recall as the key metrics. The detection results are shown in Table 6. The experimental results show that the Seedling-YOLO proposed in this paper can achieve a detection precision of more than 95% when the motion speed of broccoli is lower than 0.4 m/s. The recognition precision can be maintained above 93% within the speed of 0.6 m/s. At this point, the model exhibited good recognition performance on both normal ridges and complex ridge surfaces. When the speed increased to 0.7 m/s, due to the increase in motion blur, the feature extraction of the model was affected, and the precision began to decline. Resulting in a precision rate of 84.5% and a recall rate of 89.6%. When the speed continued to increase, the precision and recall dropped sharply. At this time, the model was unable to recognize normally. When the spacing between broccoli plants is 0.4 m, our algorithm achieves an efficiency of 180 plants per minute for detecting double-row vegetable ridges. The detection speed meets the planting speed of existing high-speed transplanting machines [5]. Our proposed

Seedling-YOLO satisfies the recognition speed and precision requirements of replanting robots, providing visual recognition technology support for these robots.

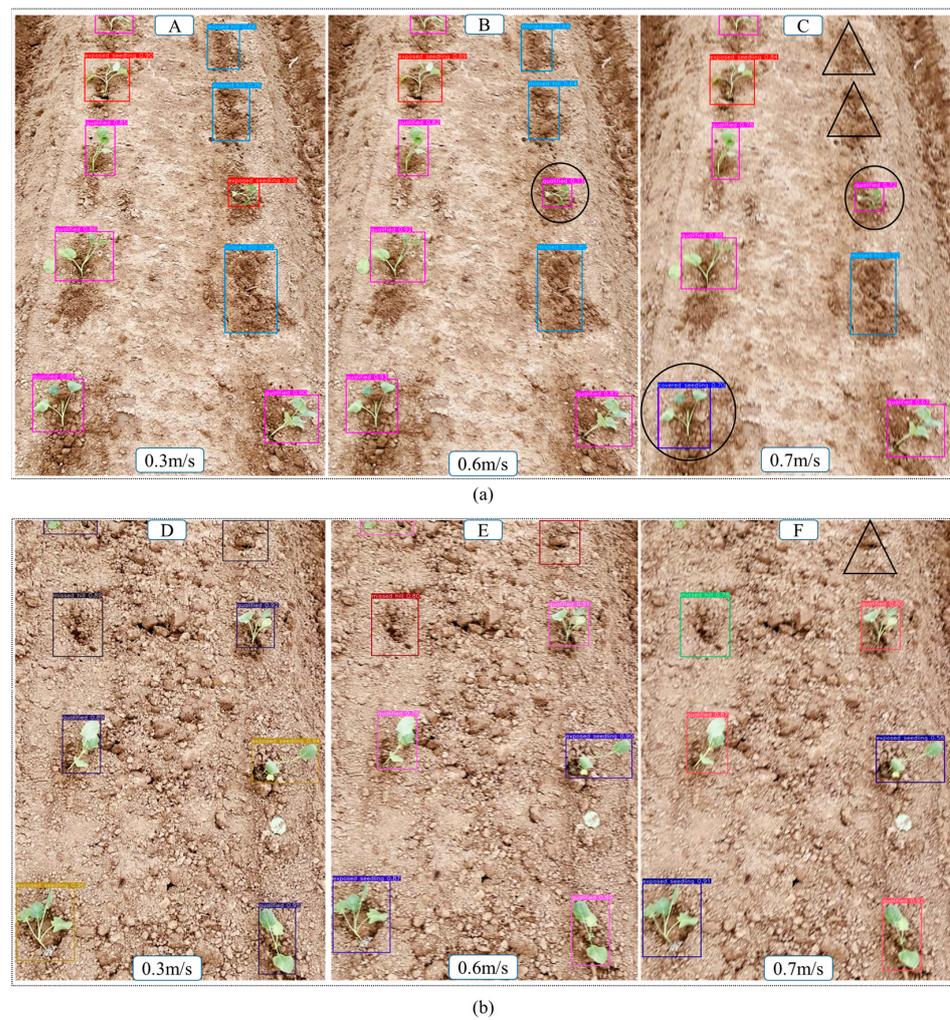**Figure 14.** Broccoli planting quality target detection platform.

Figure 15. Comparison of the detection effect of the improved model at different speeds: (**a**) Detection performance on normal ridge surfaces; (**b**) Detection performance on complex ridge surfaces. Note: In diagrams (**A**–**F**), black circles (○) signify false detections, and black triangles (△) indicate missed detections.

Table 6. Performance of the model at different speeds.

Speed	Number of Targets	Number of Missed Detections	Number of False Classification	P	R
0.3 m/s	200	2	3	97.5%	98.3%
0.4 m/s	200	2	6	96%	98.6%
0.6 m/s	200	3	11	93%	98.0%
0.7 m/s	200	12	19	84.5%	89.6%
1 m/s	200	35	41	62%	65.7%

4. Discussion

Currently, the most advanced models are typically those that excel on public datasets. However, for specific tasks, these models often require customization and development tailored to particular recognition challenges. The detection of vegetable planting quality is a new challenge brought about by the development of fully automatic transplanting machines. In this study, we found that the features of broccoli root balls are similar to soil clods and can easily be obscured by leaves, as shown in Figure 12A,D, leading to false detections in the most advanced models. In the "missed hills" category, features left by the end effector resemble the background, making recognition difficult, as evidenced by false detections in YOLOv7-tiny in Figure 12C. In the detection of "covered seedlings", as shown in Figure 12B,D, the small target size leads to false detections and low precision in existing algorithms. To address these challenges, we developed Seedling-YOLO, which integrates YOLOv7-tiny with CARAFE, CA, and our proposed ELAN_P module. As shown in Table 5, the ablation experiments reveal that M3 and M4 significantly increase the model's accuracy and recall rate. From the heatmap in Figure 13, it is evident that the model pays more attention to areas of interest, enhancing focus on small targets and reducing interference caused by background similarities. According to M5-M8, the introduction of the CARAFE and CA operators does not significantly increase the overhead of the model. This integration effectively resolves issues related to feature similarity, occlusion, and small target size, resulting in a notable improvement in detection, as illustrated in Figures 11 and 12E–H. Typically, developing fast detection algorithms sacrifices model precision, but as shown in Table 5, Seedling-YOLO reduced the model's parameters without losing precision using ELAN_P, making it more suitable for deployment on resource-constrained devices. Compared to previous research [20,21], which focused on the detection of emergence rate and missing seedlings for yield assessment, [22] developed a sugarcane seedling replanting model with a replanting location prediction method. Although successful in predicting missing seedlings, it tended to miss small seedlings. In contrast, Seedling-YOLO can directly locate missed plantings and detect exposed and covered seedlings. This study primarily developed a high-precision, real-time detection model for replanting robots, as shown in Table 4. High precision improves the precision of replanting, reducing the cases of incorrect replanting, thereby enhancing the efficiency of replanting robots. A high recall rate ensures the robot identifies and replants more areas that actually require it, which is crucial for ensuring overall crop yield. Additionally, the proposed model has been applied to a visual chassis for recognition verification at different speeds, achieving over 90% precision at speeds up to 0.6 m/s, as shown in Figure 15. As the speed increases, due to motion blur, false and missed detections begin to occur, similar to the conclusions drawn in the literature [40,47].

Regarding the applicability and limitations of the model, Seedling-YOLO can assist replanting robots in identifying substandard plantings. It also calculates the distance between the robot and the seedling using the coordinates of the bounding box, guiding the motor to the location requiring replanting. Furthermore, the bounding box coordinates aid in directing the robot's end effector for precise replanting positioning. Additionally, the model is suitable for completing replanting tasks within a few days following transplantation by a transplanting machine. This not only ensures consistent growth between replanted and field-grown seedlings but also prevents the deterioration of mound surfaces due to weather

or human factors, which could reduce the precision of missing seedling detection. Given the model's adaptability to unstructured environments, it is expected to excel even further in stable, controlled settings. However, our model is currently specific to broccoli planting quality detection. In the future, we plan to collect data on more vegetable varieties, enabling the model to be used for quality detection in a broader range of vegetable plantings. In the future, we can incorporate deblurring algorithms to enable the model to adapt to faster walking speeds of replanting robots. Additionally, by adding layers for small targets, we can further enhance the model's precision in detecting small objects.

5. Conclusions

In this study, we successfully designed Seedling-YOLO, an efficient object detection algorithm for the planting quality of broccoli seedlings. This model efficiently handles real-time detection of diverse planting conditions, including qualified, exposed, covered seedlings, and missed hills, which are commonly problematic in field environments due to false and missed detections by existing algorithms.

Leveraging YOLOv7-tiny, we redesigned the ELAN module by incorporating Pconv, significantly reducing the model's parameter, and thereby streamlining the backbone feature extraction process. Further enhancements were achieved by integrating the CARAFE operator, which uses a larger receptive field for upsampling to boost model precision. Additionally, we introduced CA in the backbone and neck shallow layers, focusing the model more on critical areas when capturing features.

The architecture of Seedling-YOLO has shown substantial improvements in terms of precision and speed. Experimental validation confirmed that the model can effectively classify four types of broccoli seedling planting qualities. Notably, the *AP* for detecting missing seedlings increased by 11.2%. Compared to the original model, Seedling-YOLO's parameters were reduced by 20%, and FLOPs by 16%, with an *mAP@0.5* of 94.3%, and an FPS of 29.7. This streamlined, more accurate model is suitable for deployment on standard hardware, achieving a detection precision of 93% at a speed of 0.6 m/s and a recognition efficiency of 180 plants/min in dual-row vegetable ridges with a plant spacing of 0.4 m. These capabilities fulfill high-speed planting requirements and provide robust technical support for field vegetable seedling supplementation.

In future work, we plan to broaden the application of the model to include various other vegetable seedlings, aiming to increase the versatility of the seedling recognition system for diverse agricultural environments. Additionally, we plan to explore the development of advanced seedling picking and planting devices that integrate our visual recognition technology, potentially revolutionizing the mechanization of seedling planting and replanting operations. These expansions could significantly contribute to the global efforts in precision agriculture, aiming to improve crop yields, optimize resource use, and ensure food security.

References

1. Wang, H.; He, J.; Aziz, N.; Wang, Y. Spatial Distribution and Driving Forces of the Vegetable Industry in China. *Land* **2022**, *11*, 981. [CrossRef]
2. Yu, G.; Lei, W.; Liang, S.; Xiong, Z.; Ye, B. Advancement of mechanized transplanting technology and equipments for field crops. *Trans. CSAE* **2022**, *53*, 1–20.
3. Jin, Y.; Liu, J.; Xu, Z.; Yuan, S.; Li, P.; Wang, J. Development status and trend of agricultural robot technology. *Int. J. Agric. Biol. Eng.* **2021**, *14*, 1–19. [CrossRef]
4. Cui, Z.; Guan, C.; Xu, T.; Fu, J.; Chen, Y.; Yang, Y.; Gao, Q. Design and experiment of transplanting machine for cabbage substrate block seedlings. *INMATEH Agric. Eng.* **2021**, *64*, 375–384. [CrossRef]
5. Ji, J.; Yang, L.; Jin, X.; Ma, H.; Pang, J.; Huang, R.; Du, M. Design of intelligent transplanting system for vegetable pot seedling based on PLC control. *J. Intell. Fuzzy Syst.* **2019**, *37*, 4847–4857. [CrossRef]
6. Zhao, S.; Liu, J.; Jin, Y.; Bai, Z.; Liu, J.; Zhou, X. Design and Testing of an Intelligent Multi-Functional Seedling Transplanting System. *Agronomy* **2022**, *12*, 2683. [CrossRef]
7. Ma, G.; Mao, H.; Han, L.; Liu, Y.; Gao, F. Reciprocating mechanism for whole row automatic seedling picking and dropping on a transplanter. *Appl. Eng. Agric.* **2020**, *36*, 751–766. [CrossRef]
8. Tong, J.; Shi, H.; Wu, C.; Jiang, H.; Yang, T. Skewness correction and quality evaluation of plug seedling images based on Canny operator and Hough transform. *Comput. Electron. Agric.* **2018**, *155*, 461–472. [CrossRef]
9. Han, L.; Mo, M.; Gao, Y.; Ma, H.; Xiang, D.; Ma, G.; Mao, H. Effects of new compounds into substrates on seedling qualities for efficient transplanting. *Agronomy* **2022**, *12*, 983. [CrossRef]
10. Vavrina, C.S.; Shuler, K.D.; Gilreath, P.R. Evaluating the impact of transplanting depth on bell pepper growth and yield. *HortScience* **1994**, *29*, 1133–1135. [CrossRef]
11. Liu, C.; Gong, l.; Fan, J.; Li, Y. Current status and development trends of agricultural robots. *Trans. CSAE* **2022**, *53*, 1–22.
12. Bini, D.; Pamela, D.; Prince, S. Machine Vision and Machine Learning for Intelligent Agrobots: A review. In Proceedings of the 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), Coimbatore, India, 5–6 March 2020; pp. 12–16. [CrossRef]
13. Mavridou, E.; Vrochidou, E.; Papakostas, G.A.; Pachidis, T.; Kaburlasos, V.G. Machine vision systems in precision agriculture for crop farming. *J. Imaging* **2019**, *5*, 89. [CrossRef]
14. Colaço, A.F.; Molin, J.P.; Rosell-Polo, J.R.; Escolà, A. Application of light detection and ranging and ultrasonic sensors to high-throughput phenotyping and precision horticulture: Current status and challenges. *Hort. Res.* **2018**, *5*, 35. [CrossRef]
15. Andújar, D.; Rueda-Ayala, V.; Moreno, H.; Rosell-Polo, J.R.; Escolá, A.; Valero, C.; Gerhards, R.; Fernández-Quintanilla, C.; Dorado, J.; Griepentrog, H.-W. Discriminating Crop, Weeds and Soil Surface with a Terrestrial LIDAR Sensor. *Sensors* **2013**, *13*, 14662–14675. [CrossRef]
16. Micheletto, M.J.; Chesñevar, C.I.; Santos, R. Methods and Applications of 3D Ground Crop Analysis Using LiDAR Technology: A Survey. *Sensors* **2023**, *23*, 7212. [CrossRef]
17. Zou, K.; Ge, L.; Zhang, C.; Yuan, T.; Li, W. Broccoli seedling segmentation based on support vector machine combined with color texture features. *IEEE Access* **2019**, *7*, 168565–168574. [CrossRef]
18. Ge, L.; Yang, Z.; Sun, Z.; Zhang, G.; Zhang, M.; Zhang, K.; Zhang, C.; Tan, Y.; Li, W. A method for broccoli seedling recognition in natural environment based on binocular stereo vision and Gaussian mixture model. *Sensors* **2019**, *19*, 1132. [CrossRef]
19. Shahi, T.B.; Xu, C.-Y.; Neupane, A.; Guo, W. Recent Advances in Crop Disease Detection Using UAV and Deep Learning Techniques. *Remote Sens.* **2023**, *15*, 2450. [CrossRef]
20. Lin, Y.; Chen, T.; Liu, S.; Cai, Y.; Shi, H.; Zheng, D.; Lan, Y.; Yue, X.; Zhang, L. Quick and accurate monitoring peanut seedlings emergence rate through UAV video and deep learning. *Comput. Electron. Agric.* **2022**, *197*, 106938. [CrossRef]
21. Cui, J.; Zheng, H.; Zeng, Z.; Yang, Y.; Ma, R.; Tian, Y.; Tan, J.; Feng, X.; Qi, L. Real-time missing seedling counting in paddy fields based on lightweight network and tracking-by-detection algorithm. *Comput. Electron. Agric.* **2023**, *212*, 108045. [CrossRef]
22. Wu, T.; Zhang, Q.; Wu, J.; Liu, Q.; Su, J.; Li, H. An improved YOLOv5s model for effectively predict sugarcane seed replenishment positions verified by a field re-seeding robot. *Comput. Electron. Agric.* **2023**, *214*, 108280. [CrossRef]
23. Zhang, C.; Liu, J.; Li, H.; Chen, H.; Xu, Z.; Ou, Z. Weed Detection Method Based on Lightweight and Contextual Information Fusion. *Appl. Sci.* **2023**, *13*, 13074. [CrossRef]
24. Zhao, S.; Lei, X.; Liu, J.; Jin, Y.; Bai, Z.; Yi, Z.; Liu, J. Transient multi-indicator detection for seedling sorting in high-speed transplanting based on a lightweight model. *Comput. Electron. Agric.* **2023**, *211*, 107996. [CrossRef]
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
26. Zhang, J.-L.; Su, W.-H.; Zhang, H.-Y.; Peng, Y. SE-YOLOv5x: An Optimized Model Based on Transfer Learning and Visual Attention Mechanism for Identifying and Localizing Weeds and Vegetables. *Agronomy* **2022**, *12*, 2061. [CrossRef]
27. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
28. Hou, Q.; Zhou, D.; Feng, J. Coordinate attention for efficient mobile network design. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.

29. Zhu, D.; Wen, R.; Xiong, J. Lightweight corn silk detection network incorporating with coordinate attention mechanism. *Trans. CSAE* **2023**, *39*, 145–153.

30. Sun, Z.; Zhang, C.-L.; Ge, L.-Z.; Zhang, M.; Li, W.; Tan, Y. Image detection method for broccoli seedlings in field based on faster R-CNN. *Trans. Chin. Soc. Agric. Mach.* **2019**, *50*, 216–221.

31. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 2969239–2969250. [CrossRef]

32. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

33. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023; pp. 7464–7475.

34. Zhang, X.; Zeng, H.; Guo, S.; Zhang, L. Efficient long-range attention network for image super-resolution. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; pp. 649–667.

35. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861. [CrossRef]

36. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6848–6856.

37. Ji, W.; Pan, Y.; Xu, B.; Wang, J. A real-time apple targets detection method for picking robot based on ShufflenetV2-YOLOX. *Agriculture* **2022**, *12*, 856. [CrossRef]

38. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. Ghostnet: More features from cheap operations. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 1580–1589.

39. Shahi, T.B.; Sitaula, C.; Neupane, A.; Guo, W. Fruit classification using attention-based MobileNetV2 for industrial applications. *PLoS ONE* **2022**, *17*, e0264586. [CrossRef]

40. Zhai, C.; Fu, H.; Zheng, K.; Zheng, S.; Wu, H.; Zhao, X. Establishment and experimental verification of deep learning model for on-line recognition of field cabbage. *Trans. Chin. Soc. Agric.* **2022**, *53*, 293–303.

41. Zhang, X.; Jing, M.; Yuan, Y. Tomato seedling classification detection using improved YOLOv3-Tiny. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 221–229.

42. Wang, Y.; He, Z.; Wang, J.; Wu, C.; Yu, G.; Tang, H. Experiment on transplanting performance of automatic vegetable pot seedling transplanter for dry land. *Trans. CSAE* **2018**, *34*, 19–25.

43. Yao, M.; Hu, J.; Liu, W.; Yue, R.; Zhu, W.; Zhang, Z. Positioning control method for the seedling tray of automatic transplanters based on interval analysis. *Trans. CSAE* **2023**, *39*, 27–36.

44. Liu, G.; Dundar, A.; Shih, K.J.; Wang, T.-C.; Reda, F.A.; Sapra, K.; Yu, Z.; Yang, X.; Tao, A.; Catanzaro, B. Partial convolution for padding, inpainting, and image synthesis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 6096–6110. [CrossRef]

45. Wang, J.; Chen, K.; Xu, R.; Liu, Z.; Loy, C.C.; Lin, D. Carafe: Content-aware reassembly of features. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 3007–3016.

46. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.

47. Liu, J.; Abbas, I.; Noor, R.S. Development of deep learning-based variable rate agrochemical spraying system for targeted weeds control in strawberry crop. *Agronomy* **2021**, *11*, 1480. [CrossRef]