

Article

A SPH-YOLOv5x-Based Automatic System for Intra-Row Weed Control in Lettuce

Bo Jiang, Jian-Lin Zhang, Wen-Hao Su *  and Rui Hu 

College of Engineering, China Agricultural University, Haidian, Beijing 100083, China; jib6011978@gmail.com (B.J.); zjl878607028@163.com (J.-L.Z.); hurui777@outlook.com (R.H.)

* Correspondence: wenhao.su@cau.edu.cn

Abstract: Weeds have a serious impact on lettuce cultivation. Weeding is an efficient way to increase lettuce yields. Due to the increasing costs of labor and the harm of herbicides to the environment, there is an increasing need to develop a mechanical weeding robot to remove weeds. Accurate weed recognition and crop localization are prerequisites for automatic weeding in precision agriculture. In this study, an intra-row weeding system is developed based on a vision system and open/close weeding knives. This vision system combines the improved you only look once v5 (YOLOv5) identification model and the lettuce–weed localization method. Compared with models including YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5n, and YOLOv5x, the optimized SPH-YOLOv5x model exhibited the best performance in identifying, with precision, recall, *F1*-score, and mean average precision (*mAP*) value of 95%, 93.32%, 94.1% and 96%, respectively. The proposed weed control system successfully removed the intra-row weeds with 80.25% accuracy at 3.28 km/h. This study demonstrates the robustness and efficacy of the automatic system for intra-row weed control in lettuce.

Keywords: weeds remove; mechanical weeding robot; deep learning; YOLOv5 model



Citation: Jiang, B.; Zhang, J.-L.; Su, W.-H.; Hu, R. A SPH-YOLOv5x-Based Automatic System for Intra-Row Weed Control in Lettuce. *Agronomy* **2023**, *13*, 2915. <https://doi.org/10.3390/agronomy13122915>

Academic Editor: Andrea Peruzzi

Received: 10 October 2023

Revised: 15 November 2023

Accepted: 23 November 2023

Published: 27 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Lettuce is essential in human diets and is widely cultivated around the world [1]. Weed control is one of the most important factors affecting the production of lettuce. Intra-row weed removal is more challenging than inter-row weed removal due to the closer proximity of the intra-row weeds to the crop line [2]. Manual weeding can accurately remove weeds in the rows, but the expensive labor costs and low efficiency make it impossible to satisfy the demands of large-scale production [3]. Chemical herbicides are widely used because of their high efficiency in removing weeds from fields on a large scale. However, long-term use of herbicides results in increased weed resistance and environmental damage [4–6]. Mechanical weed control is efficient and environmentally friendly. However, most of the current mechanical weed control is not effective in removing weeds from rows due to the lack of an accurate real-time plant location technology [7,8]. Therefore, it is important to develop an intelligent, fast, and real-time lettuce field identification and positioning technology for lettuce field weeding machinery.

Computer vision technology has demonstrated great potential for the rapid classification and localization of crops and weeds with the increase in computer computing power and camera performance [9]. Iowa State University (USA) used a stereo camera to extract depth information from the scene and identified the center of corn plants in a laboratory environment through eight basic steps, including image segmentation and skeletonization [10]. Li et al. [11] successfully extracted the grain skeleton and eventually identified individual corn plants with an accuracy of 96.7% by using computer vision to process depth images. However, the algorithm estimated the location of the corn with only 70% accuracy and the skeleton extraction algorithm was not effective in separating the crop from the weed in the case of overlapping crop and weed. In addition, many scholars

achieved crop and weed classification by extracting the color, texture, shape, and other features of crops using machine learning techniques [12–14]. Nonetheless, relying solely on manually selected features fails to exploit the full potential of other distinguishing characteristics inherent in crops and weeds. Deep learning techniques have a brilliant ability to extract complex features from large amounts of data, making it a remarkable achievement in image classification tasks [15,16]. By employing deep learning methodologies for crop and weed classification, researchers have achieved remarkable outcomes in terms of classification results and accuracy, as demonstrated in various scholarly studies [17–19]. For instance, Tao et al. [20] proposed a hybrid convolutional neural networks–support vector machines (CNN-SVM) classifier for detecting weeds in winter rape fields and achieved an accuracy of 92.1%. Mohd Anul Haq et al. [21] proposed a CNN LVQ model to classify weed from soybean in a dataset of UAV image and showed an overall accuracy of 99.44%. Completing the classification task is not enough, because the ultimate goal is to rapidly obtain the precise location of crops and weeds. YOLO (you only look once) is an advanced object detection algorithm proposed by Joseph Redmon and others, which offers higher real-time performance compared to traditional object detection algorithms by simultaneously detecting multiple objects in an image within a short period [22]. A vertical rotating intra-row robot based on YOLOv3 developed by Quan et al. [23] successfully detect the maize from weeds with a crop detection rate of 98.50%, and a weed detection rate of 90.9%. Chen et al. [24] proposed a YOLO-sesame model based on YOLOv4, using an attention mechanism to detect weeds from sesame. The mean accuracy of the sesame crop and weeds was 96.16% with a speed of 27.17 ms/f and the F1-score of crops and weeds were 0.91 and 0.92. Wang et al. [25] improved a YOLOv5s model to develop a precision spray system in corn fields. The accuracy of the improved model is 92.6%, which was 3% better than the original model with a speed of 3.2 ms/f. The above literature demonstrates the enormous potential of the YOLO model in the field of rapid crop and weed detection and localization. The ultimate goal of building a deep learning crop weed recognition model is to be able to assist field weeding machinery in real time for intra-row weed removal. In addition to accurate crop weed identification and localization technology, fast and efficient weeding devices are particularly important. The weeder system that accompanies the intelligent localization recognition model is crucial. Therefore, in addition to proposing a novel deep learning weed recognition model, this paper designs a weeding system to validate the model based on prior research.

In this paper, an inter-row weeding system is developed based on a vision system and open/close weeding knives. The novelty of this study is the development of an integrated system for weed identification and lettuce localization. The specific contribution can be summarized as follows: (1) building a self-attention prediction head combined YOLOv5x (SPH-YOLOv5x) deep learning model to distinguish crops–weeds under complex backgrounds; (2) proposing an effective vision system for the identification of lettuce stem emerging points that combines the improved YOLOv5 model and the lettuce–weed localization method; (3) developing an intra-row weeding device based on a vision system and open/close weeding knives.

2. Materials and Methods

2.1. Dataset Preparation

Lettuce and weed images were collected in Weifang, Shandong, China (36.755, 119.211; Sunny Day), forming the original dataset, which includes 275 images. This dataset encompasses images of five distinct weed species: 52 *Veronica polita* Fries (VP) weeds, 51 *Avena fatua* L. (AF) weeds, 54 *Malachium aquaticum* L. (MA) weeds, 87 *Plantago asiatica* L. (PA) weeds, and 23 *Sonchus wightianus* DC. (SW) weeds, as is visually represented in Figure 1. To meet the data requirements for machine learning model training, data augmentation techniques were employed. Besides using flip and rotate, the original images of both weeds and lettuce underwent enhancement, involving chroma adjustment and brightness adjustment. Through comparison, it was found that when the brightness and chroma are

adjusted to between 0.5 and 0.7 of the original images, there can be obvious differences between the original images after adjustment, and too many shadows and highlights will not be generated in the image process; that is, the original image is randomly selected [26]. Adjusting the chroma and brightness of the image from 0.5 to 0.7 for data enhancement can show the difference from the original image while still retaining the details of the original image. In the case of lettuce plants, 100 images were randomly chosen for data augmentation, while the remaining 175 lettuce images were designated as the test set. Following data augmentation, the new dataset included 312 VP images, 306 AF images, 324 MA images, 408 PA images, 138 SW images, and 430 lettuce images. For model training, 116 plant images were allocated to the training set, with 372 images for testing purposes. The effect of the data enhancement is shown in Figure 2.

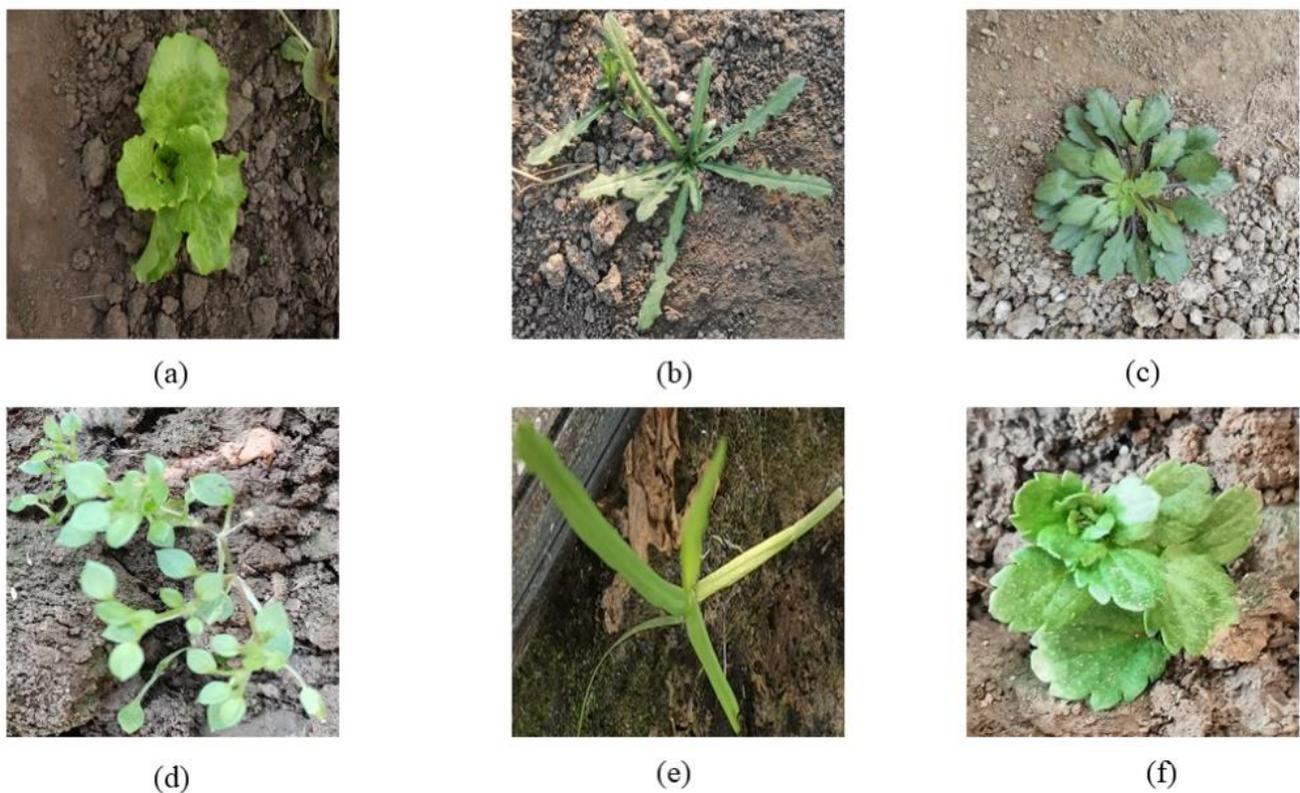


Figure 1. Examples of lettuce and weed images. (a) An example of lettuce, (b) an example of SW, (c) an example of PA, (d) an example of MA, (e) an example of AF, and (f) an example of VP.

2.2. Model Training Conditions

The vision system mainly consists of an industrial camera (Gaoshe, Shenzhen, China, UBS 2.0) and a computer. The training images are acquired by the camera in real time and then processed by the computer. The parameters of the camera include a maximum resolution of 4500×3500 pixels and a frame rate of 30 frames per second. The height of the camera from the ground is about 500 mm, and the horizontal distance between the camera and the weed knife blade is 50 mm. The acquisition range of each camera was 400 square centimeters, depending on the needs of the vision system. The computer was equipped with an NVIDIA GTX 960 graphics card. The data training process was carried out on a server with a workstation equipped with an NVIDIA GeForce RTX 3090 GPU (NVIDIA, Santa Clara City, California, US, 24 GB RAM), an Intel (R) Xeon (R) Platinum 8156 CPU (Intel, Santa Clara City, California, US), and 20 GB RAM (Kingston, Fremont, California, US). The parameters for training were as follows: learning rate 0.001, batch size 16, and epoch number 150.

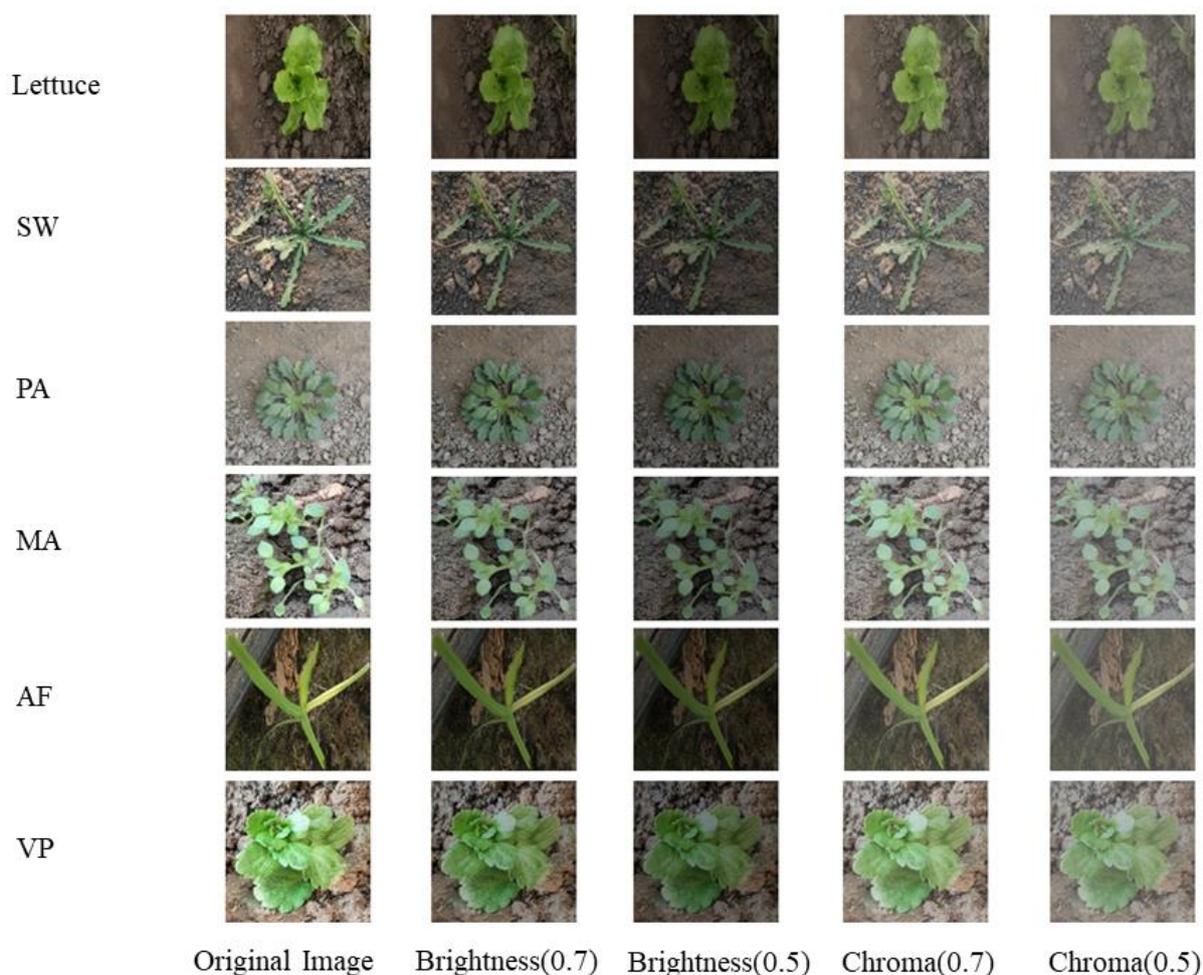


Figure 2. Data augmentation effects chart.

2.3. SPH-YOLOv5x Model

Among the seven versions in the YOLO series, YOLOv5 is the most widely used for object detection [27]. YOLOv5 series networks have been widely used in the field of agriculture for weed and crop detection [28–30]. YOLOv5 comprises a backbone architecture known as CSP-Darknet53, a neck structure called the path aggregation network (PANet), and a prediction head referred to as YOLO [31]. YOLOv5x improves detection performance and robustness compared to YOLOv5 by employing a larger model, multi-scale inference, an enhanced model backbone network, and training strategies, striking a better balance between speed and accuracy [26]. SPH-YOLOv5x is a one-stage object detection model: the same as the original YOLOv5 model. The original YOLOv5 was modified to achieve the identification and location of weeds and lettuce more precisely.

2.3.1. Detection Neck and Head

The neck is improved to better employ the features extracted by the backbone network. It can reprocess the feature maps extracted by the backbone network at different stages. The neck is a crucial link in the target detection network. Usually, a detection neck consists of several top-down paths and several down-top paths. The original fast spatial pyramid pooling (SPPF) module is replaced by the SPPFCSPC in the neck network.

SPPFCSPC is a combination of two modules: SPPF and SPPCSPC. SPPF is a module in the YOLOv5 network. SPPCSPC is a module in the YOLOv7 network [32]. The modules SPPF and SPPCSPC are proposed to solve the two problems of image distortion and to avoid repetition of proposed features. The combination of the two modules can better solve

these two problems and lead to improved performance of the model. The structure of the three modules is shown in Figure 3.

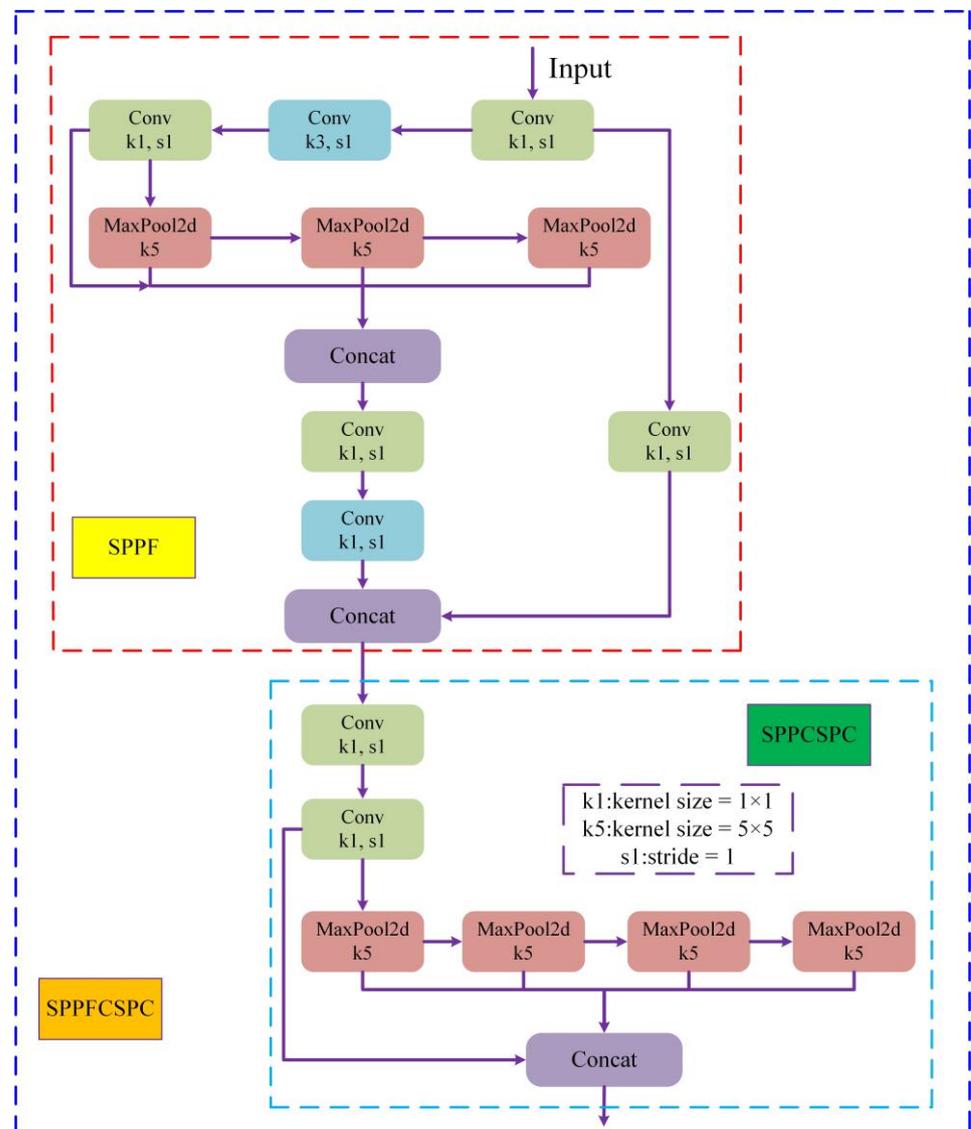


Figure 3. The structure diagram of SPPCSPC and SPPFCSPC.

The backbone network focusing on classification is not equipped to handle object positioning. Therefore, the head module takes charge of discerning both the object's category and its spatial coordinates using the extracted feature maps from the backbone. It can be found that it contains many extremely small instances in the weed/lettuce dataset. Thus, one more prediction head is added for tiny objects detection. Combined with the other three prediction heads, the four-head structure can easily find the small weeds. The prediction head (head No. 1) is added for a low-level, high-resolution feature map, which is more sensitive to tiny objects. After adding an additional detection head, although the computation and memory cost increase, the performance of tiny objects detection improves greatly.

2.3.2. Backbone

The convolutional block attention module (CBAM) [33] represents a potent attention mechanism, offering a lightweight solution compatible with various prominent CNN architectures. CBAM is trainable in an end-to-end fashion and exhibits a two-fold inference process on feature maps: it sequentially derives attention maps along two distinct dimen-

sions; namely, channel and spatial. These derived attention maps are then employed to dynamically refine the input feature map through element-wise multiplication.

The structure of the CBAM module is shown in Figure 4. According to the experiment in the paper written by Zhu et al. [34], after integrating CBAM into different models on various detection datasets, the performance of the model improves greatly. By using CBAM, the attention area can be extracted to help SPH-YOLOv5x resist confusing information. The CBAM module sequentially infers a 1D channel attention map $M_c \in \mathbb{R}^{C \times 1 \times 1}$ and a 2D spatial attention map $M_s \in \mathbb{R}^{1 \times H \times W}$ as shown in Figure 4 [33]. The overall attention process can be summarized as follows:

$$F' = M_c(F) \otimes F \tag{1}$$

$$F'' = M_s(F') \otimes F' \tag{2}$$

where \otimes represents element-wise multiplication. During the multiplication, attention values are broadcasted along the spatial dimension. F'' is the final refined output. Figure 4 shows the computation process of each attention map. Next, the details of each attention module were described.

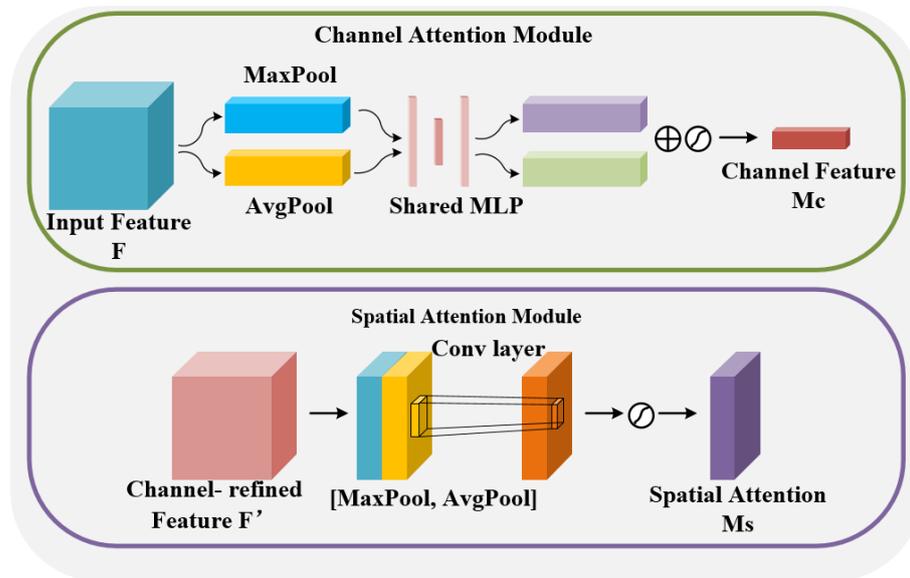


Figure 4. The structure of the channel attention module and the spatial attention module.

In the channel attention module, a channel attention map was produced by exploiting the inter-channel relationship of features. As each channel of a feature map acts as a feature detector, channel attention is meaningful in focusing on “what” given an input image. The spatial dimension of the input feature map was squeezed to compute the channel attention. For aggregating spatial information, the max-pooled and average-pooled features simultaneously. Both descriptors are forwarded to a shared network to produce channel attention map $M_c \in \mathbb{R}^{C \times 1 \times 1}$. The shared network is composed of a multi-layer perceptron (MLP) with one hidden layer. In short, the channel attention module is computed as

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \tag{3}$$

where σ represents the sigmoid function, $W_0 \in \mathbb{R}^{C/r \times C}$, and $W_1 \in \mathbb{R}^{C/r \times C}$. Note that the MLP weights (W_0 and W_1) are shared for both inputs and the ReLU activation function is followed by W_0 .

In the spatial attention module, a spatial attention map was generated by utilizing the inter-spatial relationship of features. Distinguish from channel attention, spatial attention is complementary to the channel attention according to its focus on the informative region. For computing the spatial attention, max-pooling and average-pooling operations are applied along the channel axis. And the two operations are concatenated to generate an effective feature descriptor. It was proved that applying pooling operations can effectively highlight informative regions [33]. In short, the channel attention module is computed as

$$M_S(F) = \sigma\left(f^{7 \times 7}([AvgPool(F); MaxPool(F)])\right) = \sigma\left(f^{7 \times 7}\left(\left[F_{avg}^S; F_{max}^S\right]\right)\right) \quad (4)$$

The framework of SPH-YOLOv5x is illustrated in Figure 5.

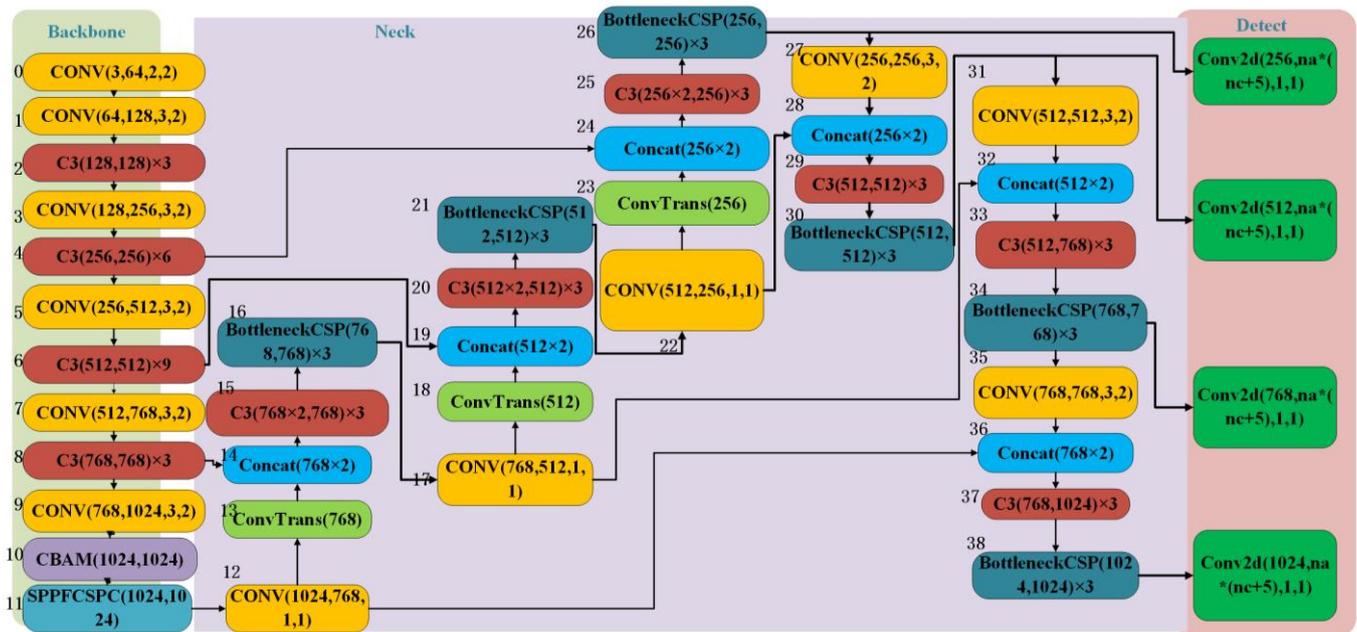


Figure 5. The framework of the SPH-YOLOv5x model.

2.3.3. Localization Algorithm

As shown in Figure 6, the localization system is implemented by building an added localization module based on the improved SPH-YOLOv5x target detection model. The localization method has been mentioned in Zhang’s paper [26]. Specifically, it is implemented based on the bounding enclosing box and HSV color space generated by the deep learning model prediction. The center of the lettuce rootstock is located by locating the center of the bounding box. Its final output will be in the form of a “txt” file of plant species and their rootstock coordinates at 30 frames per second. The positioning system can not only specialize in locating crops in combination with open and closed weed knives for indiscriminate weed removal within the crop row, but they can also be positioned to focus on locating weeds between plants by precision spraying or hammering. In this article, the positioning accuracy and weed control effectiveness of the system will be verified in combination with the open and closed weed knife.

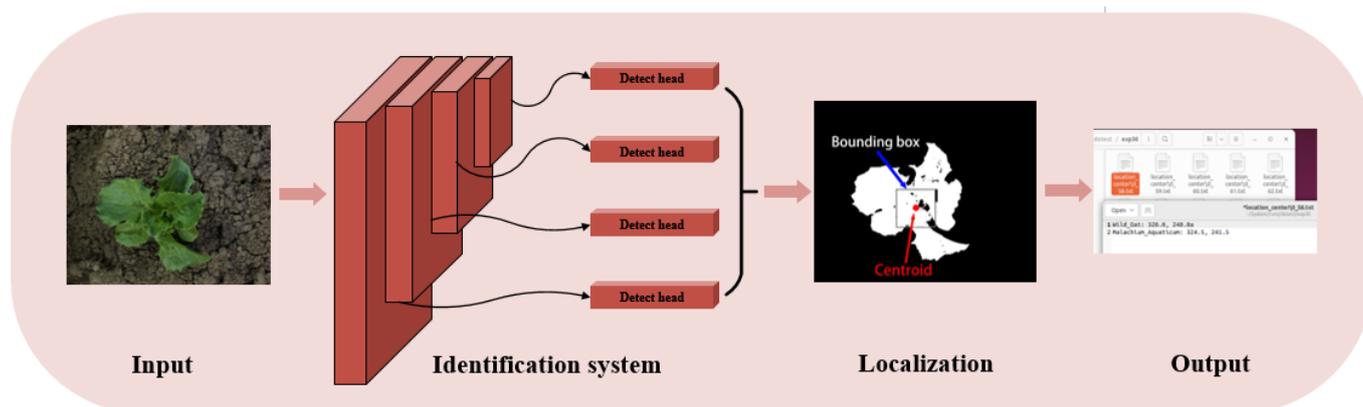


Figure 6. The structural diagram for weed–lettuce identification and localization.

2.4. Model Evaluation Methods

The primary metrics employed encompass the loss function, precision, recall, *F1*-score, mean average precision (*mAP*), and the confusion matrix. The loss function gauges the data fitting effectiveness, precision assesses the positive category prediction accuracy, recall signifies the positive category recognition rate, while the *F1*-score harmonizes the precision and recall to balance the trade-off. *mAP* comprehensively evaluates the target detection performance, and the confusion matrix furnishes detailed insights into the classification performance. The equations of recall, precision, *F1*-score, and *mAP* were as follows:

$$\text{Recall} = \frac{tp}{tp + fn} \quad (5)$$

$$\text{Precision} = \frac{tp}{tp + fp} \quad (6)$$

$$\text{F1 - score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

$$mAP = \frac{\sum_{n=1}^N AP(n)}{N} \quad (8)$$

2.5. Intelligent Intra-Row Weeding System

2.5.1. Mechanical Weeding Device

An intra-row weed control device was designed by combining target detection and an intra-row mechanical weed knife. The design principle is similar to the dilution knife in the vegetable diluter developed by Eversman [35], but the accurate identification and localization of intra-row weeds are based on an optimized target detection model. The mechanical weeding equipment part is shown in Figure 7. The intra-row weeding knife consists of two weeding blades (shown in Figure 7e). It is shaped as an isosceles triangle with a low triangle width of 7 cm and a triangle height of 3.2 cm. The forward-pointing side of the weeding blade is sharpened to form a cutting edge. The two intra-row weeding knife blades are fixed to the mechanical arms of the weeding device. The forward direction of the intra-row weeding device is the direction in which the tips of the triangular blades point. In addition, the bottom surface of the triangular blade is parallel to the soil surface and remains below approximately 2.5 cm at all times during operation. The mechanical arm to which the weeding blade is attached has a pivot point at a distance of 40 cm from the soil surface. This fulcrum point allows the weeding knife blade to move in a direction across the direction of travel of the weeding device. The driving force of the intra-row weeding knife is achieved by a double-acting cylinder (model SC40×50, JuXiang, Shenzhen, China). This cylinder is connected between the frame of the weeding unit and the robotic arm. The cylinder has a travel distance of 40 cm (0.5 in.) and a bore diameter of 1.6 cm (1.5 in.).

The movement of the knife is controlled by an electronically actuated solenoid control valve (model 4V210-08 DC24V, JuXiang, Shenzhen, China) that supplies air pressure (0.7 MPa) to the cylinder.

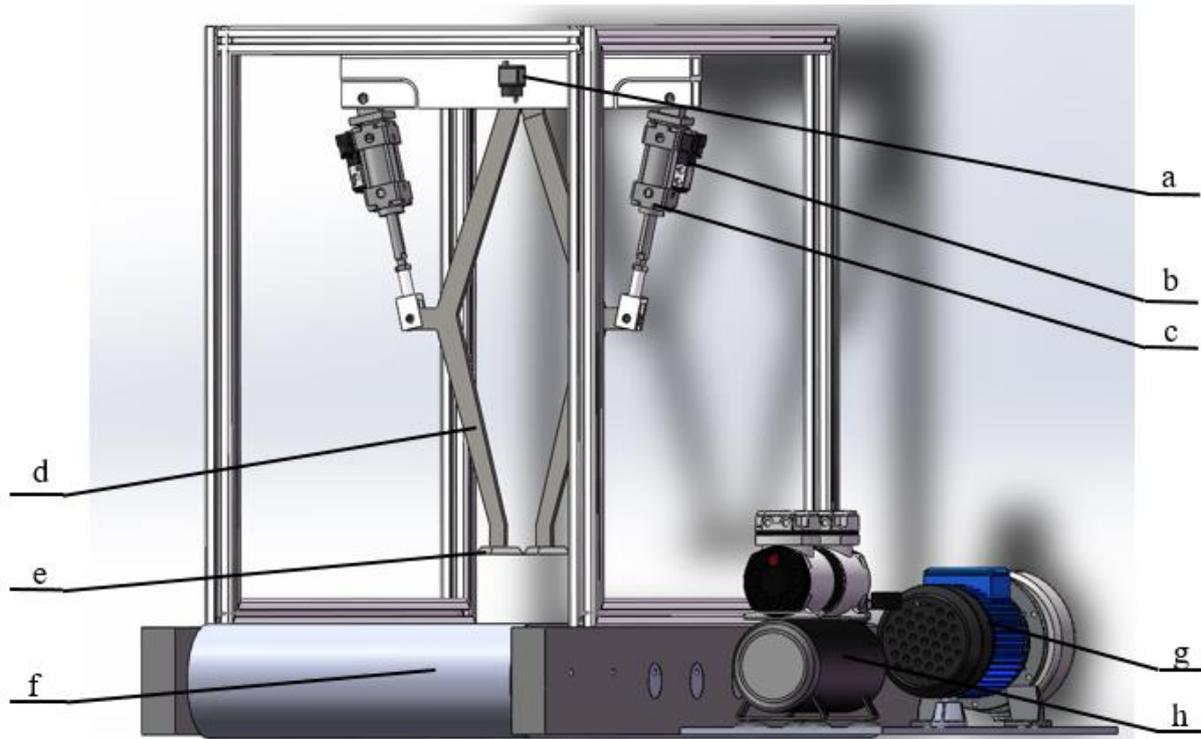


Figure 7. Open-closed intra-row weed control device: (a) industrial camera; (b) solenoid valve; (c) pneumatic cylinder; (d) mechanical arm; (e) weed cutting blade; (f) conveyor belt; (g) air compressor/pneumatic pump; (h) electric motor.

Figure 8 illustrates the working process of the intra-row weeding device. In the crop field, three areas are artificially distinguished: area A is the inter-row area, area B is the intra-row area, and area C is the crop safety area. Within the crop row, the weeding knife blade (about 7 cm wide) was used to control intra-row weeds. Figure 8b shows the sequence of the three positions in which the weeding knife blade is moved from left to right. In position ① and position ③, the weeding blade is positioned in the “closed” position in the intra-row area driven by the cylinder, and the two blades advance in parallel. As the blade approaches the lettuce plant, in position ②, the cylinder separates the blade into the inter-row area along the purple dotted line. This leaves the safety zone C unbroken. After the knife blade has passed the lettuce, the cylinder re-drives the blade back into the intra-row area. This process is repeated for each lettuce crop. When not in the crop area, the weeding knife is in a closed state and penetrates deep into the soil. As the machine (conveyor belt) travels, all weeds in the area are removed by their roots.

The system requires accurate vision recognition tools to ensure that it opens in the “open” position. The accuracy of the vision system’s positioning directly affects the effectiveness of intra-row weeding equipment, and obtaining accurate information on the distance between seedlings can reduce the probability of equipment injury. In the early stages, machine vision sensing technology could acquire the position of crops but often failed to identify their specific types. By incorporating machine learning into machine vision technology, machines have gained the capability to recognize and differentiate various crop types. However, when confronted with numerous weeds, the effectiveness of crop recognition and localization cannot be guaranteed. Deep learning-based target detection methods have gradually become popular in the agricultural field. In facing the intra-row weeding task in agriculture, the deep learning target detection technique also provides

a new solution and has achieved great value in the initial exploration. Therefore, in this paper, the designed intra-row weeding equipment uses deep learning target detection as a recognition tool.

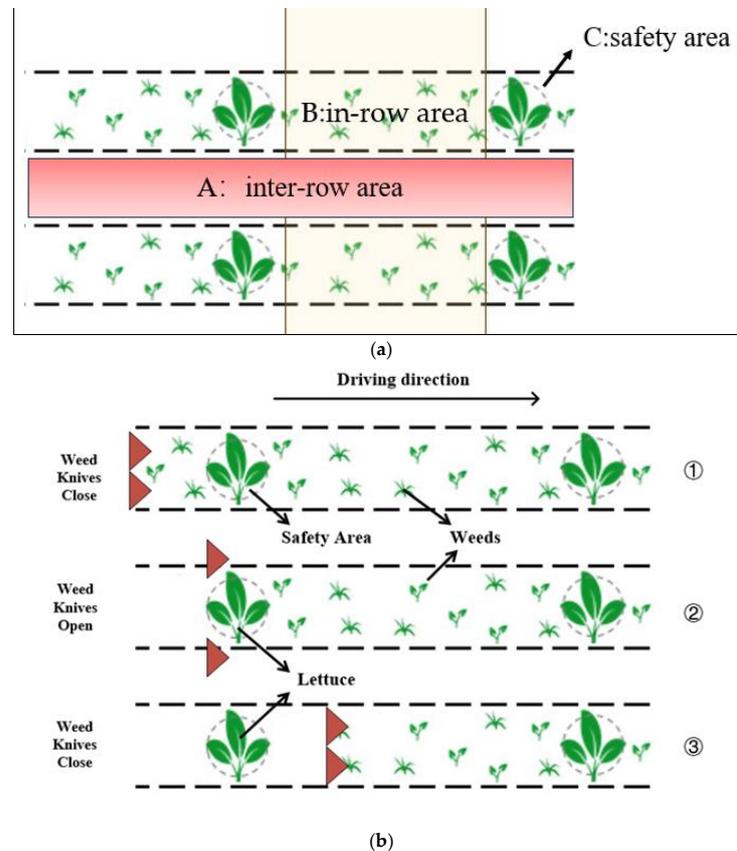


Figure 8. Schematic of weed distribution: (a) working area schematic diagram; (b) schematic diagram of weed cutting blade working principle.

2.5.2. Real-Time Control System

The hardware system of the intra-row weeding device is built as shown in Figure 9a. A controller (Arduino Uno R3, Arduino S.r.l., Via Andrea Pollini 11, 20159 Milan, Italy) was used for the lower computer control of the weeding gear. Arduino, an open-source single-chip microcontroller, uses an Atmel AVR microcontroller. This microcontroller uses an open-source software and hardware platform and it is built on a simple output/input interface board. The 220 v power supply is used to supply power to the laptop, adjustable power supply, and air compressor. The camera is connected to the laptop via USB. The camera acquires information about crops and weeds in real time and transmits it with a video stream to a laptop deployed with the SPH-YOLOv5x deep learning model. The detection program (detect.py) of SPH-YOLOv5x analyzes the video transmitted by the camera to obtain information about the type of crops and weeds and the coordinates of the crops. The positioning algorithm calculates the center point of the crop based on the crop coordinate information transmitted by the detection program. This information is converted into signals and passed to the controller. The microcontroller triggers the solenoid valve to control the pneumatic cylinder to provide the power required by the weed knife system to complete the full coverage of weed removal in the row avoiding the crop. The power of the solenoid valve is provided by the adjustable power supply, and the power of the air pressure cylinder is provided by the air compressor. Throughout the entire weeding process, from the initiation of the system, the lettuce identification program operates on the computer, actively monitoring and processing real-time information about lettuce weeds captured by the conveyor belt's camera. The program calculates the precise location data of the identified lettuce and promptly transmits

this information to the Arduino development board. Subsequently, the Arduino development board assumes control over the operation of the weeding knife, regulating its opening and closing functions. The entire process from identification to weeding by the weeding knife is completed continuously without interruption. Thus, the intra-row weeding device completes a complete mechanical weeding process.

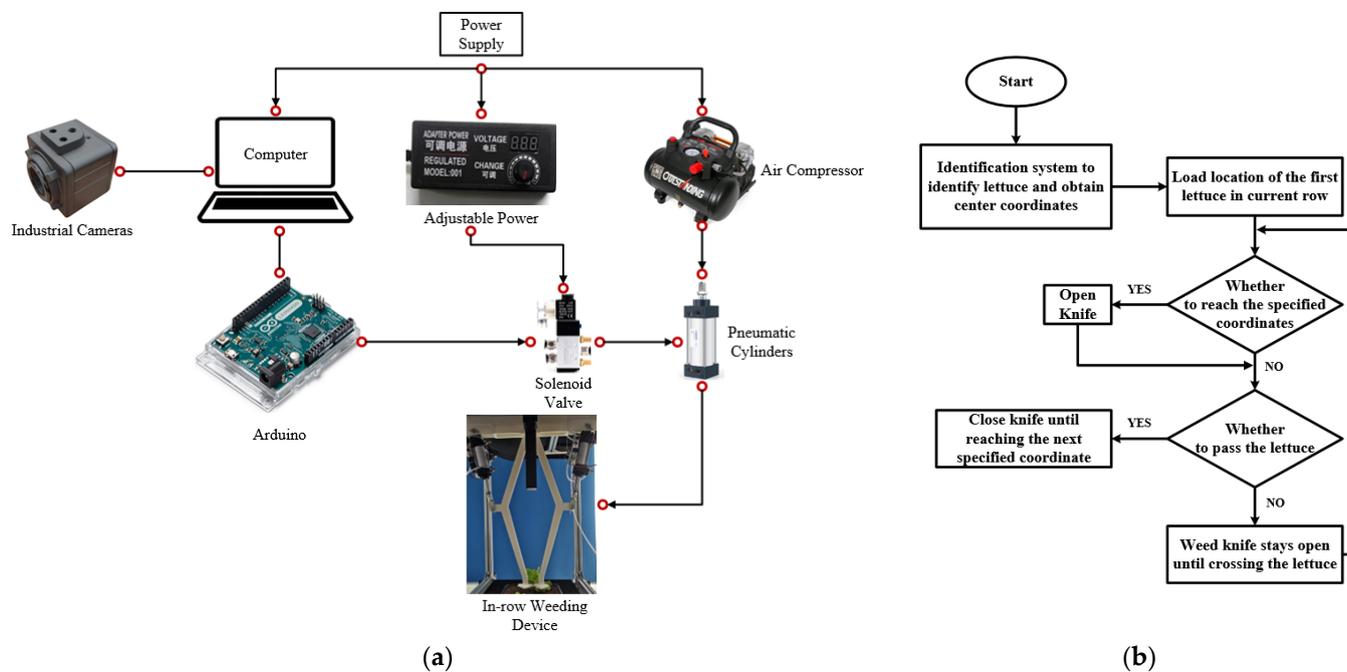


Figure 9. Intra-row weeding system: (a) components of the intra-row weeding system; (b) control program flow chart of the intra-row weeding system.

A real-time weed knife control system for intra-row weeding was developed. A flow chart of its control algorithm is shown in Figure 9b. On the test platform, the conveyor belt travels at a speed of 3.24 km/h, simulating a weed wagon walking in the field. During this process, the camera records video in real time and saves it to a local computer. The computer calculates the location of the weed knife from the crop by capturing the location of the crop and the tag information. When the weed knife is about to enter the crop safety zone, the computer transmits a signal to the Arduino microcontroller through the serial port. At this point, the Arduino microcontroller controls the cylinder to open the weed knife to bypass the crop. The process from the detection of the crop to the opening of the weed knife is almost instantaneous.

2.6. Method of Conveyor Belt Experiment

In the conveyor belt experiment, we mainly verified the weeding effect of the weeding system under different weed densities and different light conditions. The control of lighting conditions is mainly controlled by LED light strips, and different densities of weeds are planted between the lettuce to simulate different densities of weeds. The light strips that are turned on are considered to represent good lighting conditions. The positioning of weed density is based on the research of Rekha Raja et al. [36]; that is, light weed density is defined as 10 or fewer weeds per square meter; moderate weed density is 11–100 weeds per square meter; and heavy weed density is defined as more than 100 weeds per square meter. Figure 10 shows the experimental equipment and real experimental scenes.

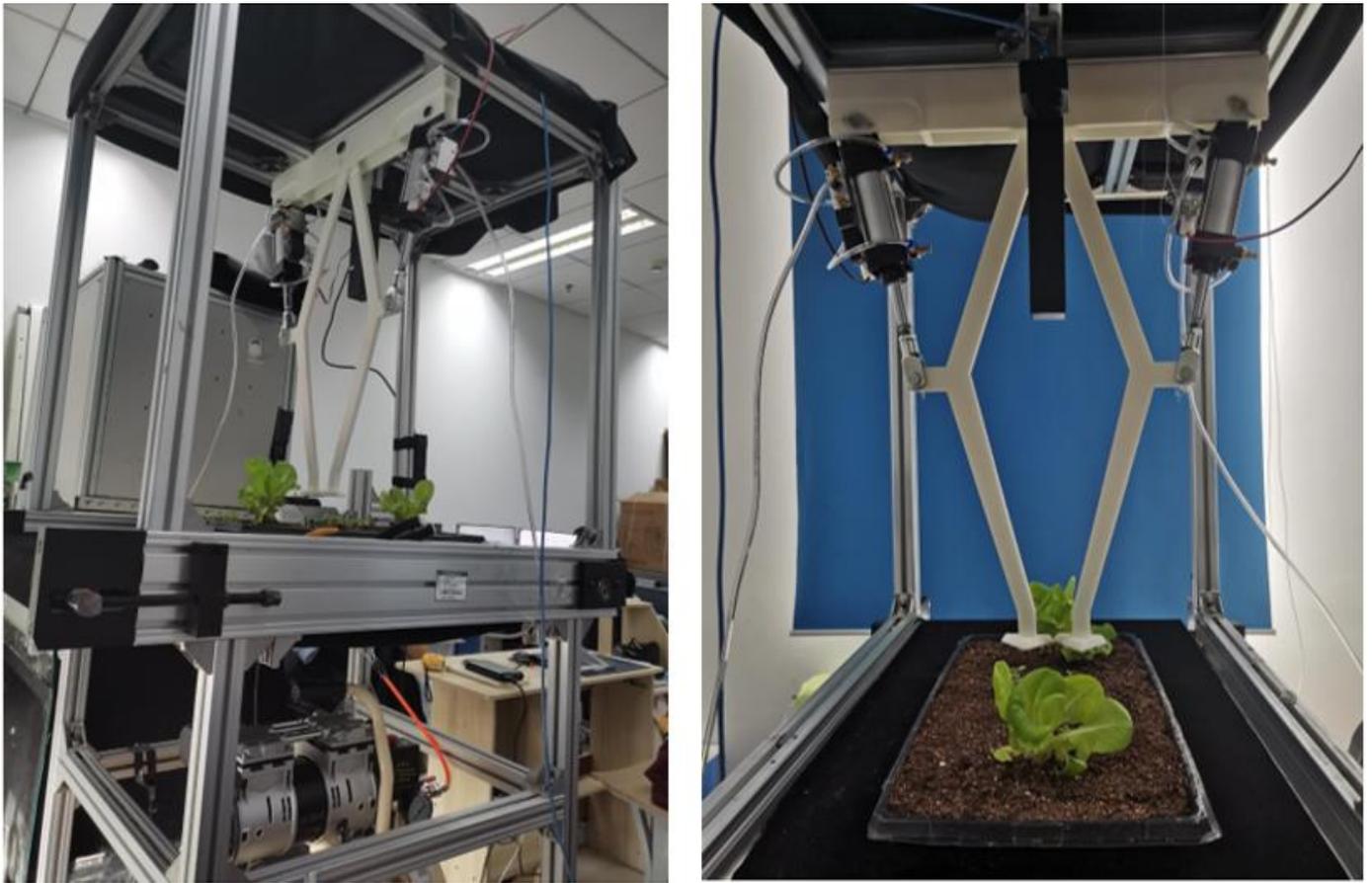


Figure 10. The final image of the mechanical weeding device entity.

3. Results

3.1. Training of Optimized YOLOv5 Model

The optimized YOLOv5 model was compared with five other YOLOv5 model sizes. The variation values during the training of the six models was shown in Figure 11 and Table 1. Figure 11a shows the training loss curve with the training period in weed–lettuce identification. All models decreased rapidly within 40 cycles and eventually stabilized as the training period increased. Among the six models trained, the proposed SPH-YOLOv5x model obtained the lowest loss values. Figure 11b shows the loss curves of the six models in the validation phase. Although the proposed SPH-YOLOv5x model fits slightly slower than the other models until at least 20 rounds, as with the training loss curves, the SPH model eventually obtains the lowest validation loss value. The proposed SPH-YOLOv5x model achieved the lowest loss values on both the training and validation sets, demonstrating its strong learning capabilities and generalization ability in weed classification.

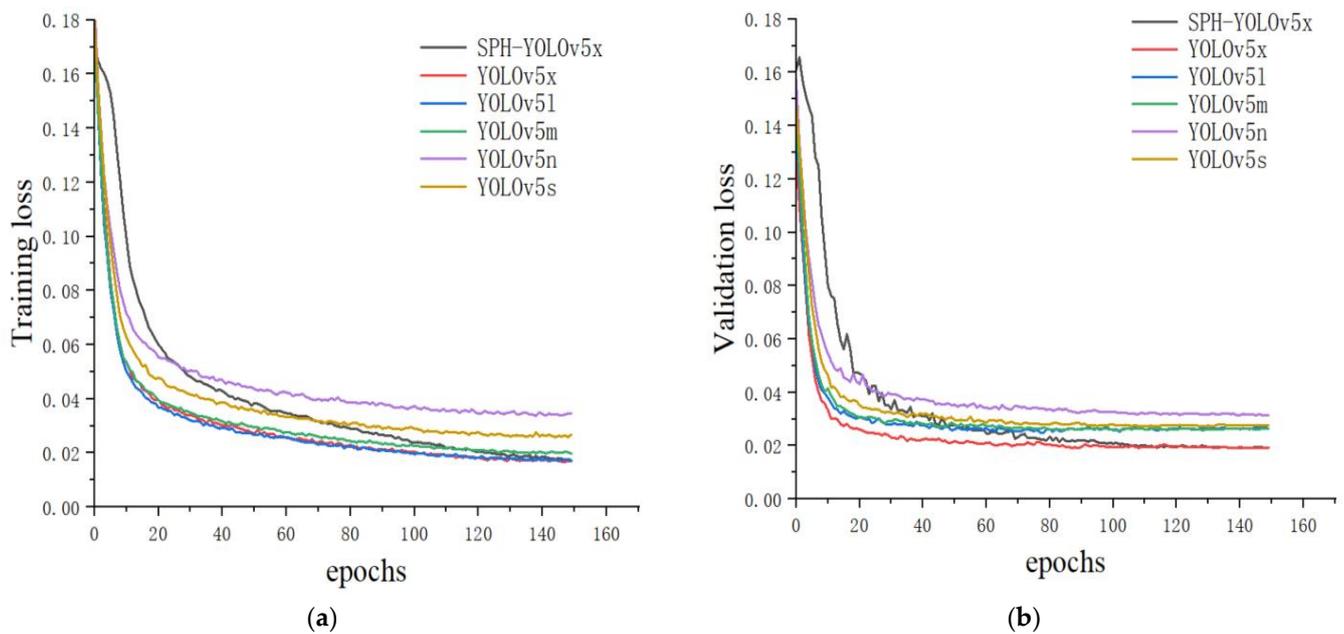


Figure 11. Comparison of training loss curves for six YOLO models: (a) train loss curves of the six models; (b) validation loss curves of the six models. An epoch represents one complete iteration of training, signifying one full pass through the training dataset for model parameter updates and learning.

Table 1. Training and validation loss values of YOLO models.

Model	Training Loss	Validation Loss
SPH-YOLOv5x	0.01738	0.01937
YOLOv5l	0.01713	0.02648
YOLOv5m	0.0197	0.02651
YOLOv5n	0.03453	0.03142
YOLOv5s	0.02668	0.0277
YOLOv5x	0.04272	0.08988

3.2. Classification and Detection of Optimized Model

In this paper, a total of six models, including the proposed SPH model, were trained to test their classification and localization abilities among lettuce weeds. Table 2 shows the scores of each model, where the proposed models achieved 95%, 93.2%, 96%, and 94.1% for precision, recall, mAP , and $F1$ -scores, respectively. It can be seen that the proposed model has higher precision and mAP than the other models, except YOLOv5X, which proves the reliability of the model.

Table 2. Lettuce and weeds classification results of YOLO models.

Model	Precision (%)	Recall (%)	$mAP@_{0.5\%}$ (%)	$F1$ -Score (%)
SPH-YOLOv5x	0.950	0.932	0.96	0.941
YOLOv5l	0.944	0.946	0.947	0.945
YOLOv5m	0.919	0.946	0.945	0.932
YOLOv5n	0.925	0.938	0.942	0.931
YOLOv5s	0.925	0.934	0.938	0.929
YOLOv5x	0.952	0.935	0.943	0.943

As shown in Table 3, the classification performance of the SPH-YOLOv5x model was analyzed for lettuce crops with five weeds: *Veronica polita* Fries, *Avena fatua* L., *Malachium aquaticum* L., *Plantago asiatica* L., and *Sonchus wightianus* DC. (VP, AF, MA, PA, SW). The accuracy of the model in correctly identifying lettuce was 92.9% for the identification of weeds. The model had the highest classification accuracy of 98.7% for PA and the lowest

recognition ability of 89% for SW. Even so, the model’s ability to classify lettuce and weeds is still at a high level. By observing the confusion matrix shown in Figure 12, it is evident that the proposed model achieves good identification and classification results for both lettuce and various weeds.

Table 3. Lettuce and weeds classification results of the YOLO models.

Plant Species	Precision (%)	Recall (%)	<i>mAP</i> @0.5% (%)	F1-Score (%)
lettuce	0.878	0.878	0.929	0.878
VP	0.991	1	0.995	0.967
AF	0.971	1	0.991	0.909
MA	0.888	0.875	0.933	0.876
PA	0.973	0.976	0.987	0.94
SW	0.889	0.85	0.89	0.861

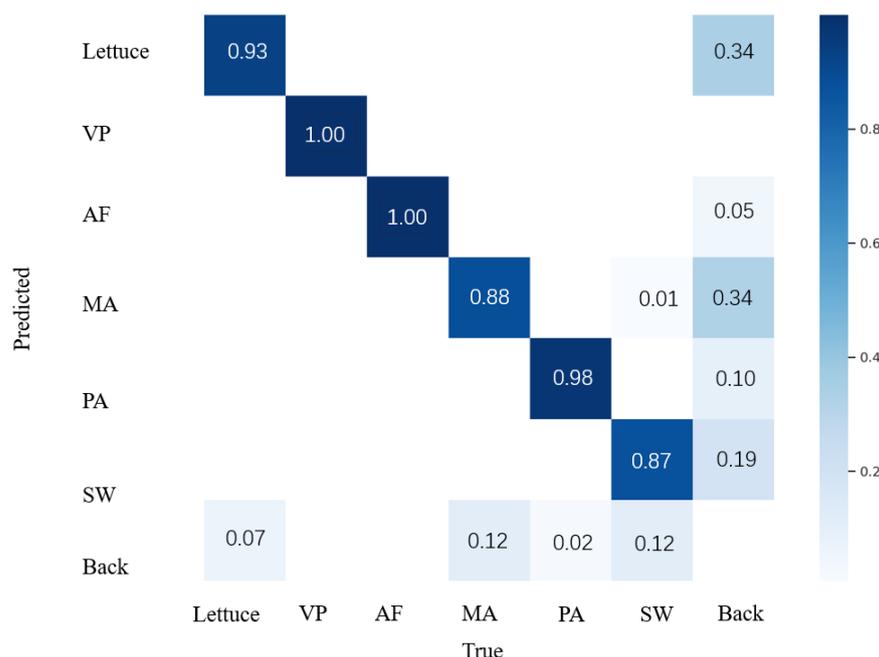


Figure 12. Confusion matrix of the trained SPH-YOLOv5x model.

3.3. Results of the Conveyor Belt Experiment

In the experimental design, we conducted a total of three trials to simulate the distribution of lettuce weeds under different light and irrigation conditions. The specific operational parameters for each of the three trials are detailed in Table 4. Figure 13 illustrates the positioning results of the three trials. During the experimental process, the experimenters assessed the positioning accuracy by observing whether the weed knife could accurately avoid the lettuce crop, and they recorded their observations accordingly. The main errors observed during the experiment were two types: missed identification and misidentification, with both falling under the category of recognition errors. The formula for the success rate of lettuce crop positioning was

$$P = \left(1 - \frac{N_m + N_l}{N}\right) \times 100\% \tag{9}$$

where *P* is the localization success rate of lettuce; *N_m* is the number of lettuce misidentifications; *N_l* is the number of lettuces missed identifications; and *N* is the total number of lettuce trials.

Table 4. Validation results of weed control devices under different irrigation and light conditions.

Experimental Batch	Luminous Conditions	Plant Number	Incorrect/Missed	Correct Detection	Success
First Batch	Good	239	49	239	82.99
Second Batch	Inferior	252	101	252	71.39
Third Batch	Good	279	44	279	86.38

**Figure 13.** Results of the localization of lettuce.

The average positioning accuracy of the three trials was found to be 80.25%, as shown in Table 4. The experimental results indicate that the first and third trials exhibited notably higher positioning accuracy compared to the second trial, which is mainly attributed to variations in lighting conditions. The second trial experienced inferior lighting conditions, which significantly impacted the recognition and positioning system. The observed experimental results did not show any significant impact of irrigation conditions on the system's recognition performance.

Based on the experiments conducted using the SPH-YOLOv5x model for weed detection, the positioning system shows promising prospects. With an average positioning accuracy of 80.25% under the current conditions (3.28 km/h), there is considerable scope for further development by optimizing both hardware and software aspects. This detection and positioning system, along with intra-row mechanical weeding equipment, is expected to contribute to increased organic agriculture yields in the future.

3.4. Efficiency of the Weed Moving System

The pre-arranged distribution of light, moderate, and heavy weed density conditions was subjected to weeding tests on the experimental platform. The weeding results are shown in Figure 14a–c. Based on the weeding outcomes, the cutting edge of the weeding knife effectively covered the majority of the intra-row area. Figure 14a–c corresponds to the light, moderate, and heavy weed density conditions, respectively.

As illustrated in Figure 15a, the weeding knife was able to remove or bury the weeds effectively in the case of light weed density. In Figure 15b, under moderate weed density, the weeding knife successfully pulled out almost all weed roots from the soil. However, in Figure 15c, under heavy weed density, the reciprocating weeding knife indiscriminately pushed weeds from the front crop to the rear crop, causing them to accumulate around the latter. While the reciprocating weeding knife showed good performance in high-density weed removal between crop rows, its effectiveness diminished when the weeds surrounded the crops. This reduced performance was mainly due to the high-density weed distribution, which often resulted in the weeds encircling the lettuce crops. Consequently, the model experienced significant localization errors during the recognition process, greatly affecting weed clearance and occasionally leading to accidental crop damage.



Figure 14. Low-density, middle-density, high-density weeds distribution diagram.

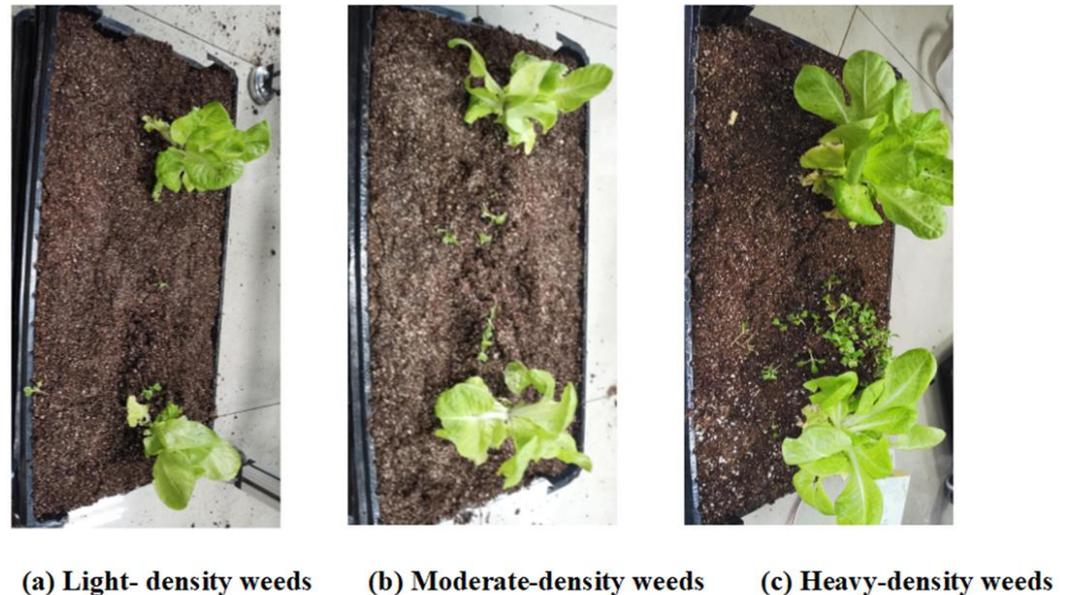


Figure 15. Low density, medium density, and high density weed control effect.

4. Discussion

In this study, we successfully designed and operated a cost-effective intra-row weeding device that incorporates a vision system to determine the opening/closing events of the weeding blades for efficient removal of intra-row weeds. The vision system was developed based on an optimized YOLOv5 model that was specifically tailored for the identification and localization of lettuce weeds. Compared to traditional machine vision methods [37], the optimized YOLOv5 approach proved to be more suitable for complex field environments. With the assistance of the vision system, the intra-row weeding device performed effectively

on the conveyor belt. The results from laboratory experiments demonstrated that the intra-row weeding device was both feasible and efficient. However, the impact of increasing the working speed on the weed control efficiency and seedling injury rate of the proposed weeding system has not yet been validated.

The vision system yielded a location accuracy as high as 80.25% and extremely good weed control at the speed of 3.2 km/h of the conveyor belt with a 30 FPS camera. Table 5 shows the results of related research about the intelligent intra-row cultivator based on different techniques in recent years. Specifically, computer vision can perform repetitive actions with greater reliability over the course of an entire work shift without fatigue factors. The burden on the human supervisor is greatly reduced. But this is achieved with the aid of agronomy. Mechanical–thermal weeding tools equipped on the tractors achieved a weeding effect of 90% with no record of major crop losses in the RHEA project, which employs agricultural robots and related high-tech equipment in addition to transforming existing small tractors on the market so that they can independently complete agricultural tasks [38]; however, this system relies on the collaborative efforts of multiple intelligent robots to accomplish the task, resulting in comparatively higher costs. Bawden et al. [39] designed a weed control robot called AgBotII which removes weeds by spray or mechanical means based on the information captured by the visual system [19]. The misclassification rate of the system is 7.7% which means weeds may be retained and crops removed. Wu designed an automatic classification weeding robot that combined the spray and stamp methods. The weeding robot achieved good recognition and weeding results on both flat and uneven ground. However, the robot was not evaluated at speeds higher than 2 m/s due to the huge tangential force on the stamper caused by the driving vehicle [40]; Quan et al. [23] developed an intelligent intra-row robot based on YOLO v3 using a rotating hoe. The robotic system successfully classified and removed the weeds with an accuracy of 86.13% in the conveyor belt experiment with the speed under 0.5 m/s. Raja et al. [41] developed a real-time weed sprayer machine based on the crop signal with machine vision. At a speed of 3.2 km/h with the conveyor belt, 98% of the weeds were successfully sprayed among the 83.7% detected weeds. However, this system requires an additional crop spraying system, which adds additional work procedures. The greatest advantage of the Institute’s proposed weeding device over some of the smart weeding equipment mentioned above is its low cost and simplicity and efficiency.

Table 5. Research progress of intelligent weeding equipment in recent years.

Reference	System Name	Technology	Crop Name
Gonzalez-de-Santos et al. [38]	Co-robot	Sensing technique	Tomato
Bawden et al. [39]	NaN	Machine vision	Cabbage
Sujaritha et al. [42]	NaN	Machine vision	Bok choy, celery, lettuce, and radicchio
Wu et al. [40]	NaN	Ultrasonic sensor	
Quan et al. [23]	SLIC Super-pixel algorithm	ConvNet	Soybean
Raja et al. [41]	SSWM system	Deep learning	Corn and soybean
Proposed method	SPH-YOLOv5x system	SPH-YOLOv5x	Lettuce

The model proposed in this study demonstrates higher accuracy in crop recognition compared to the aforementioned weeding system. Furthermore, the model performed well according to the benchmark of YOLO object detectors for weed detection in different turf grass scenarios proposed by Sportelli et al. [43] and in the cotton production system proposed by Dang et al. [44]. However, it is essential to acknowledge that during the experimental validation phase, the proposed model did not achieve the same level of performance as it did during training. The reasons for the positioning accuracy error can be attributed to both software and hardware factors. On the software side, the main issue lies in the communication between the laptop and the Arduino microcontroller. Regarding hardware, the accuracy is closely related to the performance of both the laptop and the

microcontroller. Although the conveyor experimental system's crop recognition accuracy did not reach 95%, as was achieved during model training, this does not necessarily imply poor practical applicability of the model. The actual recognition performance is significantly influenced by hardware conditions, which can be improved through technical enhancements. Additionally, the model has not been validated in more complex scenarios, such as at higher conveyor speeds and real field experiments. These validations will be conducted after further optimization of the software and hardware to better assess and verify the feasibility of the proposed weeding system. The proposed mechanical weeding system integrated with deep learning technology has been experimented with and validated for its feasibility and efficiency. However, it is important to note that the system assumes a flat ground surface, and its weeding efficiency cannot be guaranteed when encountering uneven terrain. For future design considerations, it is essential to address the working mechanism of the weeding system when dealing with uneven surfaces. Although the primary focus of the mechanical weeding system proposed in this article is the identification and positioning of lettuce, without utilizing specific weed information, the model's capability to identify and locate weeds still holds significance in the realm of automated weeding. The information pertaining to weed identification and positioning serves as a fundamental basis for the precise deployment of more advanced weeding components, such as spray weeding or laser weeding. This implies that, while the article's primary emphasis is on lettuce, the knowledge gained about weed location can provide valuable support and a foundation for subsequent, more sophisticated weeding operations.

5. Conclusions

In this study, an advanced intra-row weeding mechanism has been successfully formulated by leveraging a deep learning-based vision system, aiming to achieve precise and automated weeding processes within the realm of agriculture. The intelligent intra-row weeding system comprises a recognition and localization system, along with a mechanical weeding device. Within the vision system, the SPH-YOLOv5X YOLOv5x model was presented, which is a customized iteration of the YOLOv5x model, designed explicitly for the purpose of identifying and pinpointing field lettuce weeds. The improved YOLOv5 model demonstrated remarkable recognition performance, achieving accuracy, recall, *F1*-score, and *mAP* values of 95%, 93.2%, 94.1%, and 96%, respectively. Through conveyor belt simulations, the proposed model effectively facilitated the mechanical weeding device to achieve crop discrimination and weed removal with an impressive accuracy of 80.25% at a speed of 3.28 km/h. The knowledge generated from this research is expected to significantly contribute to the development of automatic weeding robots and offer innovative solutions for precise automated weeding in modern agriculture.

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

Funding: This research was funded by the National Natural Science Foundation of China, grant number 32101610.

Data Availability Statement: Due to privacy, data are only available upon request.

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

References

1. Kim, M.J.; Moon, Y.; Tou, J.C.; Mou, B.; Waterland, N.L. Nutritional value, bioactive compounds and health benefits of lettuce (*Lactuca sativa* L.). *J. Food Compos. Anal.* **2016**, *49*, 19–34. [[CrossRef](#)]
2. Pérez-Ruiz, M.; Slaughter, D.; Gliever, C.; Upadhyaya, S. Automatic GPS-based intra-row weed knife control system for transplanted row crops. *Comput. Electron. Agric.* **2012**, *80*, 41–49. [[CrossRef](#)]

3. Melander, B.; Rasmussen, G. Effects of cultural methods and physical weed control on intrarow weed numbers, manual weeding and marketable yield in direct-sown leek and bulb onion. *Weed Res.* **2001**, *41*, 491–508. [[CrossRef](#)]
4. Dai, X.; Xu, Y.; Zheng, J.; Song, H. Analysis of the variability of pesticide concentration downstream of inline mixers for direct nozzle injection systems. *Biosyst. Eng.* **2019**, *180*, 59–69. [[CrossRef](#)]
5. Perotti, V.E.; Larran, A.S.; Palmieri, V.E.; Martinatto, A.K.; Permingeat, H.R. Herbicide resistant weeds: A call to integrate conventional agricultural practices, molecular biology knowledge and new technologies. *Plant Sci.* **2020**, *290*, 110255. [[CrossRef](#)]
6. Song, J.-S.; Chung, J.-H.; Lee, K.J.; Kwon, J.; Kim, J.-W.; Im, J.-H.; Kim, D.-S. Herbicide-based weed management for soybean production in the Far Eastern region of Russia. *Agronomy* **2020**, *10*, 1823. [[CrossRef](#)]
7. Ronchi, C.; Silva, A.; Korres, N.; Burgos, N.; Duke, S. *Weed Control: Sustainability, Hazards and Risks in Cropping Systems Worldwide*; CRC Press: Boca Raton, FL, USA, 2018.
8. Fennimore, S.A.; Slaughter, D.C.; Siemens, M.C.; Leon, R.G.; Saber, M.N. Technology for automation of weed control in specialty crops. *Weed Technol.* **2016**, *30*, 823–837. [[CrossRef](#)]
9. Tang, J.; Wang, D.; Zhang, Z.; He, L.; Xin, J.; Xu, Y. Weed identification based on K-means feature learning combined with convolutional neural network. *Comput. Electron. Agric.* **2017**, *135*, 63–70. [[CrossRef](#)]
10. Jin, J.; Tang, L. Corn plant sensing using real-time stereo vision. *J. Field Robot.* **2009**, *26*, 591–608. [[CrossRef](#)]
11. Li, W.; Wu, G.; Zhang, F.; Du, Q. Hyperspectral image classification using deep pixel-pair features. *IEEE Trans. Geosci. Remote Sens.* **2016**, *55*, 844–853. [[CrossRef](#)]
12. Burgos-Artizzu, X.P.; Ribeiro, A.; Guijarro, M.; Pajares, G. Real-time image processing for crop/weed discrimination in maize fields. *Comput. Electron. Agric.* **2011**, *75*, 337–346. [[CrossRef](#)]
13. Gée, C.; Bossu, J.; Jones, G.; Truchetet, F. Crop/weed discrimination in perspective agronomic images. *Comput. Electron. Agric.* **2008**, *60*, 49–59. [[CrossRef](#)]
14. Montalvo, M.; Pajares, G.; Guerrero, J.M.; Romeo, J.; Guijarro, M.; Ribeiro, A.; Ruz, J.J.; Cruz, J. Automatic detection of crop rows in maize fields with high weeds pressure. *Expert Syst. Appl.* **2012**, *39*, 11889–11897. [[CrossRef](#)]
15. Gu, J.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; Chen, T. Recent Advances in Convolutional Neural Networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
16. Yu, J.; Sharpe, S.M.; Schumann, A.W.; Boyd, N.S. Deep learning for image-based weed detection in turfgrass. *Eur. J. Agron.* **2019**, *104*, 78–84. [[CrossRef](#)]
17. Moazzam, S.I.; Khan, U.S.; Tiwana, M.I.; Iqbal, J.; Qureshi, W.S.; Shah, S.I. A review of application of deep learning for weeds and crops classification in agriculture. In Proceedings of the 2019 International Conference on Robotics and Automation in Industry (ICRAI), Rawalpindi, Pakistan, 21–22 October 2019; pp. 1–6.
18. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [[CrossRef](#)]
19. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [[CrossRef](#)]
20. Tao, T.; Wei, X. A hybrid CNN–SVM classifier for weed recognition in winter rape field. *Plant Methods* **2022**, *18*, 29. [[CrossRef](#)]
21. Haq, M.A. CNN Based Automated Weed Detection System Using UAV Imagery. *Comput. Syst. Sci. Eng.* **2022**, *42*.
22. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
23. Quan, L.; Jiang, W.; Li, H.; Li, H.; Wang, Q.; Chen, L. Intelligent intra-row robotic weeding system combining deep learning technology with a targeted weeding mode. *Biosyst. Eng.* **2022**, *216*, 13–31. [[CrossRef](#)]
24. Chen, J.; Wang, H.; Zhang, H.; Luo, T.; Wei, D.; Long, T.; Wang, Z. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion. *Comput. Electron. Agric.* **2022**, *202*, 107412. [[CrossRef](#)]
25. Wang, B.; Yan, Y.; Lan, Y.; Wang, M.; Bian, Z. Accurate Detection and Precision Spraying of Corn and Weeds Using the Improved YOLOv5 Model. *IEEE Access* **2023**, *11*, 29868–29882. [[CrossRef](#)]
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. Wang, A.; Peng, T.; Cao, H.; Xu, Y.; Wei, X.; Cui, B. TIA-YOLOv5: An improved YOLOv5 network for real-time detection of crop and weed in the field. *Front. Plant Sci.* **2022**, *13*, 1091655. [[CrossRef](#)] [[PubMed](#)]
28. Wang, Q.; Cheng, M.; Huang, S.; Cai, Z.; Zhang, J.; Yuan, H. A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed *Solanum rostratum* Dunal seedlings. *Comput. Electron. Agric.* **2022**, *199*, 107194. [[CrossRef](#)]
29. Junior, L.C.M.; Ulson, J.A.C. Real time weed detection using computer vision and deep learning. In Proceedings of the 2021 14th IEEE International Conference on Industry Applications (INDUSCON), São Paulo, Brazil, 15–18 August 2021; pp. 1131–1137.
30. Doddamani, P.K.; Revathi, G. Detection of Weed & Crop using YOLO v5 Algorithm. In Proceedings of the 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 16–17 October 2022; pp. 1–5.
31. Jocher, G.; Stoken, A.; Borovec, J.; Chaurasia, A.; Changyu, L.; Hogan, A.; Hajek, J.; Diaconu, L.; Kwon, Y.; Defretin, Y. ultralytics/yolov5: v5. 0-YOLOv5-P6 1280 models, AWS, Supervise. ly and YouTube integrations. *Zenodo* **2021**. [[CrossRef](#)]

32. 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 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 7464–7475.
33. 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.
34. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788.
35. Kepner, R.; Bainer, R.; Barger, E. Selective mechanical or chemical thinning. In *Principles of Farm Machinery*; The Avi Publishing Company, Inc.: Westport, CT, USA, 1978.
36. Raja, R.; Nguyen, T.T.; Slaughter, D.C.; Fennimore, S.A. Real-time weed-crop classification and localisation technique for robotic weed control in lettuce. *Biosyst. Eng.* **2020**, *192*, 257–274. [[CrossRef](#)]
37. Hlaing, S.H.; Khaing, A.S. Weed and crop segmentation and classification using area thresholding. *Int. J. Res. Eng. Technol.* **2014**, *3*, 375–382.
38. Gonzalez-de-Santos, P.; Ribeiro, A.; Fernandez-Quintanilla, C.; Lopez-Granados, F.; Brandstötter, M.; Tomic, S.; Pedrazzi, S.; Peruzzi, A.; Pajares, G.; Kaplanis, G. Fleets of robots for environmentally-safe pest control in agriculture. *Precis. Agric.* **2017**, *18*, 574–614. [[CrossRef](#)]
39. Bawden, O.; Kulk, J.; Russell, R.; McCool, C.; English, A.; Dayoub, F.; Lehnert, C.; Perez, T. Robot for weed species plant-specific management. *J. Field Robot.* **2017**, *34*, 1179–1199. [[CrossRef](#)]
40. Wu, X.; Aravecchia, S.; Lottes, P.; Stachniss, C.; Pradalier, C. Robotic weed control using automated weed and crop classification. *J. Field Robot.* **2020**, *37*, 322–340. [[CrossRef](#)]
41. Raja, R.; Slaughter, D.C.; Fennimore, S.A.; Siemens, M.C. Real-time control of high-resolution micro-jet sprayer integrated with machine vision for precision weed control. *Biosyst. Eng.* **2023**, *228*, 31–48. [[CrossRef](#)]
42. Sujaritha, M.; Annadurai, S.; Satheshkumar, J.; Sharan, S.K.; Mahesh, L. Weed detecting robot in sugarcane fields using fuzzy real time classifier. *Comput. Electron. Agric.* **2017**, *134*, 160–171. [[CrossRef](#)]
43. Sportelli, M.; Apolo-Apolo, O.E.; Fontanelli, M.; Frasconi, C.; Raffaelli, M.; Peruzzi, A.; Perez-Ruiz, M. Evaluation of YOLO Object Detectors for Weed Detection in Different Turfgrass Scenarios. *Appl. Sci.* **2023**, *13*, 8502. [[CrossRef](#)]
44. Dang, F.; Chen, D.; Lu, Y.; Li, Z. YOLOWeeds: A novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. *Comput. Electron. Agric.* **2023**, *205*, 107655. [[CrossRef](#)]

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