Article

Design and Experimental Verification of the YOLOV5 Model Implanted with a Transformer Module for Target-Oriented Spraying in Cabbage Farming

Hao Fu 1,2,†, Xueguan Zhao 3,†, Huarui Wu 3, Shenyu Zheng 1, Kang Zheng 1 and Changyuan Zhai 1,2,*

1 Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
2 College of Mechanical Engineering, Guangxi University, Nanning 530004, China
3 National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China
* Correspondence: zhaicy@nercita.org.cn; Tel.: +86-10-5150-3886
† These authors contributed equally to this work.

Abstract: Due to large line spacing and planting distances, the adoption of continuous and uniform pesticide spraying in vegetable farming can lead to pesticide waste, thus increasing cost and environmental pollution. In this paper, by applying deep learning and online identification methods, control technology for target-oriented spraying is studied with cabbages as the research object. To overcome motion blur and low average precision under strong light conditions during the operation of sprayers, an innovative YOLOV5 model implanted with a transformer module is utilized to achieve accurate online identification for cabbage fields under complex environments. Based on this concept, a new target-oriented spray system is built on an NVIDIA Jetson Xavier NX. Indoor test results show that the average precision is 96.14% and the image processing time is 51.07 ms. When motion blur occurs, the average precision for the target is 90.31%. Then, in a field experiment, when the light intensity is within the range of 3.76–12.34 wlx, the advance opening distance is less than 3.51 cm, the delay closing distance is less than 2.05 cm, and the average identification error for the cabbage diameter is less than 1.45 cm. The experimental results indicate that changes in light intensity have no significant impact on the identification effect. The average precision is 98.65%, and the savings rate reaches 54.04%. In general, the target-oriented spray system designed in this study achieves the expected experimental results and can provide technical support for field target spraying.

Keywords: precision agriculture; precision pesticide spraying; deep learning; target-oriented spray; target identification

1. Introduction

With the increasing global population and demand for food and vegetables, crop output is facing severe strain, especially under the limited scale of arable land. Cabbage is one of the most widely planted vegetables in the world. Yield is mainly affected by diseases and pests. To date, pesticide spraying has been the most effective means of disease and pest control. However, the traditional continuous spraying method causes 60–70% of the chemical solution to be deposited in nontarget areas [1], which causes a series of problems, such as significant pesticide waste, excessive chemical residue, and environmental pollution. Target-oriented spraying is an effective way to solve these problems [2]. Sensors are utilized to obtain target information in real time, and independent nozzle control technologies are combined to achieve precise spraying on a single target [3].

The precondition of target-oriented spraying is to obtain precise target information in real time. Field detection systems include machine vision technology [4–6], ultrasonic sensor detection technology [7], lidar detection technology [8], ultrasonic sensors and lidar, which can be used to determine the positions of targets in the area based on reflected...
ultrasonic waves and beams. However, such a system cannot accurately distinguish crops and weeds in the field. In any case, machine vision technology has some advantages in cabbage identification due to the characteristics of abundant information acquisition, accuracy and intelligence.

In recent years, with the development of computer technology and graphics processing units (GPUs), vision technology based on machine learning algorithms has gradually begun to be used for field crop identification [9–11]. On the basis of manually extracted contours, textures, colors and other feature information of the targets, a machine learning classifier can be adopted to identify targets [12]. However, because the target features are extracted manually in this method, it is difficult to comprehensively evaluate the complex and changeable field environment, resulting in low generalization of the identification method.

Based on a large amount of image data, the target features can be automatically extracted by adopting deep learning. Therefore, this method also has good generalization performance even in complex environments [13]. Currently, many scholars have begun to use deep learning models to identify field crops [14–16]. The mechanism of action is that a model is trained on a large number of images in advance, and the model then automatically extracts the features used to distinguish the target crops, thus improving the precision and robustness of identification [17]. For example, Suh et al. [18] collected weed images in different environments to conduct the deep learning of VGG19. The test results showed that the average precision of the model reached 98.7% after model training. In addition, in a study using the VGG16 model for transfer learning by Ahmad et al. [19], the average precision of the model was 98.90% after the weed-identification model was trained using the pictures of the test set. However, such a model can determine only the target category and cannot output the location information of the target. To enable such a model to output the position information of the target while also identifying the target crop, target-detection algorithms based on deep learning can be used for field crop detection [20,21]. For example, Zhang and Li [22] designed a target-detection model based on YOLO v5s. Under the indoor lighting environment, the average precision of this model is 99.6%. Ying et al. [23] used Mobiletv3-small to replace the backbone in YOLOv4 to design a lightweight weed-identification model that was implemented on a high-performance RTX3090 GPU, with an average precision of 88.46% and an image processing time of 12.65 ms. However, in all of the above studies, only offline processing has been carried out on images acquired under static conditions. The performance of these models has not been verified under harsh field conditions, in which the light intensity is variable and the motion of the equipment during operation causes motion blurring of the captured images.

In an unstructured field environment with variable light intensity, the target images acquired by a camera will exhibit problems of strong highlight areas, fuzzy features, and difficulty in segmenting and identifying crops and weeds. Therefore, it is necessary to design systems and algorithms with high average precision for different light intensities. At the hardware level, some researchers and robotics companies [24–27] have designed active light sources and light shields to provide a stable lighting environment for cameras. However, in this scheme, the large-scale light shield on the machine is large in volume, which causes the photographic equipment to have low field passing ability, and the light shield may also damage the vegetables. Moreover, the power of the light source is high and requires large energy consumption, which also greatly reduces the endurance time of the machine. In terms of algorithms, to reduce the influence of illumination on the average precision, researchers have tried to preprocess the images to be input into the deep learning model in advance. For example, Wang et al. [28] developed three image augmentation methods through color space conversion and color index calculation, which can be used to optimize the input of the deep learning model to improve its robustness under different lighting conditions. In addition, image normalization, size adjustment, contrast augmentation, degradation and calibration (conversion from image coordinates to world coordinates) are common image preprocessing methods [29]. However, some feature information of the targets in the original images may easily be missing after the
images are preprocessed, which will affect the average precision and increase the operation time and memory consumption of the model, thus making it difficult to achieve real-time identification.

When working in the field, vibration or shaking of the camera will cause the blurring of the collected images, which will affect the accuracy of image identification. In general, the number of relevant studies focused on reducing the average precision of crop identification caused by blurred images during the operation is limited. For example, by implanting a weed-identification model into an NVIDIA GeForce GTX 1080 GPU on a spray machine, Liu et al. [30] designed a weed target sprayer. When the operation speed of the sprayer was 3 km/h, the weed-identification precision was 94%. However, when the speed was further increased, motion blur was generated, which reduced the identification precision to 86%. Tan et al. [31] designed a cotton seedling tracking model based on YOLOv4 and the optical flow method, and it was adopted to track a video collected in the early stage in 1920 × 1080 pixels that are processed offline. The average precision of the model for some videos and pictures with motion-blurred targets was 98.84%.

The image processing speed is one of the important factors affecting the precision of target orientation. To improve the processing speed of their models, some researchers [23,30,32] have implemented identification models on NVIDIA GPUs and have achieved relatively ideal modeling speeds. However, under field conditions where no external power supply is available, a high-power NVIDIA GPU will have a short operation time. Moreover, the vibration during machine operation can also cause such a GPU to easily fail. Therefore, some researchers have attempted to lighten their models by using algorithms to implant their identification models into edge computing devices for conducting real-time reasoning. For example, Partel et al. [33] implanted a deep learning model into an NVIDIA Jetson TX2 for detection, and the average precision of spraying on artificial simulated plants was 90%; in comparison, when using real plants for experiments, the precision rate was reduced to 59%. Additionally, de Aguiar et al. [34] accelerated the SSD mobilenet-v2 model with an Intel neural rod, achieving an average precision of 52.98% and an image processing time of 23.14 ms. These scholars used edge computing equipment in crop identification modeling and achieved a short image processing time. However, the average precision has remained too low to meet the needs of actual applications in many cases. Therefore, it is necessary to design a real-time identification model for edge computing devices with high average precision.

In summary, innovation and optimization at the levels of the model algorithm and model training method were carried out in this study. Specifically, based on a deep learning algorithm, a lightweight online identification model for cabbage fields was built to improve the average precision by establishing a transformer module and using an augmentation method for motion-blurred data. Furthermore, based on the cabbage-identification model, a target-oriented spraying system was built, and the performance of this system was tested in cabbage planting areas.

2. Materials and Methods

2.1. Construction of the Cabbage-Identification Model

In this study, the current mainstream YOLOv5 model and their optimized variants with different structures are selected for comparative research to develop an optimal cabbage-recognition model. YOLOv5 is a one-stage target-detection model, which is used to integrate the functions of classification and positioning of cabbages into a single neural network. It uses a C3 structure based on a cross-stage partial network (CSPNet) to extract target features, thereby achieving a good feature extraction effect for large targets without mutual occlusion; however, for the situation in which the cabbages and weeds in the image are blocked by each other, relatively little cabbage feature information can be extracted. In this paper, by adding location coding for cabbage targets into the original network and adopting a multihead attention mechanism, the feature acquisition ability for cabbage targets is improved, and the recognition accuracy of the model is enhanced. The amount of
calculation needed for a single transformer module is lower than that of the C3 structure; consequently, the image processing speed can also be improved to a certain extent. To solve the problem of large computational burden of traditional convolution, depthwise separable convolution is used in place of traditional convolution. The computation formulas for traditional convolution and depthwise separable convolution are shown in Equations (1) and (2). For inputs consisting of a 480 × 288 pixel image, the number of output feature maps is 16, and the convolution kernel size is 3 × 3. The computational complexity of traditional convolution is \(5.972 \times 10^7\), whereas the computational complexity when using depthwise separable convolution is \(3.732 \times 10^6\), only 6.25% of that of traditional convolution. Therefore, this paper uses depthwise separable convolution instead of traditional convolution to further reduce the computational load of the model.

\[
FLOPs = WHCD^2K \tag{1}
\]

\[
FLOPs_{(DWConv)} = WHCD^2 + D^2CK \tag{2}
\]

Here,
- \(FLOPs\) — number of floating-point operations for traditional convolution.
- \(FLOPs_{(DWConv)}\) — number of floating-point operations for depthwise separable convolution.
- \(W\) — number of image pixel columns.
- \(H\) — number of image pixel rows.
- \(C\) — number of image (feature map) channels.
- \(D\) — convolution kernel size.
- \(K\) — number of output feature maps.

2.1.1. Implementation of the Transformer Module

Since the local brightness of the cabbage targets is high, the reflection is significant, and the imaging quality is poor under strong light conditions. This leads to the reduction of cabbage characteristics in the captured images and the reduction of average precision. The transformer module [35] is implemented to solve these problems by enhancing those characteristics in this study. The transformer includes a kind of self-attention mechanism, and its structure is shown in Figure 1. The calculation process is as follows. First, the cabbage feature map input from the upper layer is encoded to facilitate the calculation of the relationships among different pixels. Then, the position coding of the feature map is added to increase the position information for the cabbage target in the feature map. Finally, the re-encoded cabbage feature map is processed with a multihead attention mechanism to improve the ability of the model to capture cabbage targets, thereby improving the average precision. After processing by the multihead attention mechanism, a fully connected layer is used to adjust the number of output channels of the cabbage feature map, and a residual structure is used to connect pixels. The deep and shallow semantic features are fused to reduce the loss of features during calculation. The transformer model has few convolution operations, resulting in a relatively high processing efficiency.

2.1.2. Overall Structure of the Cabbage-Detection Model

The cabbage-identification model constructed in this study is shown in Figure 2, and it is composed of five parts: the input network, backbone feature extraction network, neck network, detection network and output network. First, the image of 1920 pixels × 1080 pixels collected by the network camera is converted to a resolution of 480 pixels × 288 pixels by equal scaling and edge filling. In the convolutional layers, depthwise separable convolution with limited calculations is used to replace the traditional convolution operation to further reduce the number of model calculations. In the last layer of the backbone feature extraction network, a spatial pyramid pooling (SPPF) structure is used to pool the maximum values from cabbage feature maps at three different scales, and the pooled results are combined with the original feature maps to achieve rich feature map expression. This approach is conducive to detecting cabbages with different growth cycles or large size differences.
The neck network part includes a feature pyramid network (FPN) and a path aggregation network (PAN). The FPN is used to extract cabbage feature information at different scales, from shallow to deep, to improve the detection accuracy for small individual cabbages. The PAN is used to fuse the semantic features of the cabbage images obtained from the deep and shallow layers, thus enhancing the localization features of the cabbages and avoiding the loss of feature information after multiple convolution operations. Finally, the detection network outputs three sizes according to the different pixels of the input image. Each head outputs a plurality of cabbage prediction boxes. Finally, the nonmaximum suppression algorithm of the output part obtains the boundary box with the highest confidence of each cabbage for output.

2.1.3. Positioning Method for Cabbages

In this study, real-time identification and positioning of cabbage fields is carried out by video streaming. In light of the problem that a single cabbage may repeatedly appear and be recognized in different frame images, which will cause the positioning information of this cabbage to be repeatedly sent to the execution structure, a cabbage-positioning method based on the Kalman filter and the Hungarian algorithm proposed by researchers in a previous study [36] is adopted in this paper. The positioning process is as follows. First, each cabbage appearing in the field of view of the camera is given a unique and constant ID number to establish connections among the same cabbage in different frames. In addition, a virtual line is established in the field of view. When the connecting line at the front end of the cabbage boundary box on the two images intersects the virtual line, it is determined that the cabbage is on the virtual line. At this time, the model sends the cabbage position information and diameter information to the lower computer. According to the cabbage

![Figure 1. Structural diagram of the transformer.](image-url)
planting mode, the field of view is divided into four parts to detect different columns of cabbage targets, as shown in Figure 3.

Figure 2. Structure of the cabbage detection model.

Figure 3. Schematic diagram of the field-of-view segmentation of cabbages.

2.2. Preprocessing Method

Cabbage images were collected 10 times in the cabbage field (40°10′36″ N and 116°26′44″ E) of the National Precision Agriculture Research and Demonstration Base in Xiaotangshan Town, Beijing, from 1 May 2021 to 10 June 2021, and from September 18 2021 to October 8 2021. The line spacing of the cabbages in the photos was 35 cm, and the planting distance was 40 cm. The acquisition device was a Daheng industrial–experimental camera (Model: MER2-507-23GC; Manufacturer: Beijing Image Vision Technology Branch of China Daheng (Group) Co; city, country: Beijing, China) with a resolution of 2592 pixels × 1944 pixels, a USB 3.0 interface and a vertical down mode for image and video acquisition. Using the camera driver software (software name: Daheng Galaxy SDK; version: 1.12.2107.9211; creator: Beijing Image Vision Technology Branch of China Daheng (Group) Co; city, country: Beijing, China) provided by the Beijing Image Vision Technology Branch of China Daheng (Group) Co., Ltd., cabbage images were acquired based on manual triggering. Most of the images were of cabbages from the
seedling stage to the rosette stage, during which there was no overlap among the cabbages. Under such conditions, it is ideal to adopt target-oriented spraying to reduce pesticide waste. To improve the comprehensiveness of the data set, data collection was carried out under three types of weather conditions—sunny, cloudy and rainy—and during different growth cycles and at different weed densities. A total of 2425 cabbage images were collected. The collected images contained multiple cabbage targets and weed samples, which were present in the field of view of the camera during the process of target-oriented spraying. To address the problem that the vibration and movement of the system during operation cause motion blur in the cabbage images collected by the camera during online cabbage recognition, which leads to a decline in recognition accuracy, motion blur was added to the collected images for data enhancement, as shown in Figure 4. Labelimg software (Software name: Labelimg; Version: v1.8.6; Creator: tzutalin; https://github.com/heartexlabs/labelling/tree/v1.8.6 accessed on 14 July 2022) was used to manually label the augmented data set and save it in the YOLO file format to obtain the boundary box coordinate matrix of the target in each image. Consistent with the format of the COCO data set, the data set was divided into a training set, a test set and a verification set at a ratio of 4:1:1. The training set was used to train the network parameters, the test set was used to evaluate the generalization error of the model after model training, and the validation set was used to optimize the hyperparameters used in the training process to improve the performance of the model.

Figure 4. Motion blur augmentation.

2.3. Model Training

The augmented data set was used to train the established detection model for identifying cabbages. Then, the trained model was compared with the YOLOV5n model, which provides excellent average precision and processing speed in target detection. Data sets without motion-blur augmentation were used first to train the model proposed in this paper, and then 500 images subjected to motion blur processing were used to verify the average precision of the trained model. Videos of the cabbages were collected in the field in such a way that one picture was extracted every 10 frames to verify the effect of motion-blur augmentation on model performance. The hardware specifications and the specific model training process applied in this study are as follows.

2.3.1. Training Platform

The model training platform is a desktop workstation. The workstation is configured with 48 GB of memory, an Intel Core i9-10900k CPU with a main frequency of 3.70 GHz, an NVIDIA RTX 3090 GPU with a video memory capacity of 24 GB and a floating-point speed of 35.6 TFLOPS, CUDA version 11.1.96, cuDNN version 8.0.5, an operating environment based on the Windows 10 (64-bit) operating system, the Python 3.8 programming language, and the PyTorch 1.7 deep learning framework.

2.3.2. Training Strategy

To further improve the richness of the image backgrounds in the data set, the mosaic online data augmentation method [37] was used to enhance the data set. The augmentation process is as follows. First, four cabbage images in each training batch are randomly selected for scaling, color space adjustment, brightness adjustment and angle rotation.
These four processed images are spliced into an image with more abundant background feature information, as shown in Figure 5. Based on the above description, the cabbage-detection model is input for training. According to the results of preliminary experiments, the number of training iterations is set to 300. The algorithm adopted for parameter optimization is the Adam algorithm. The specific parameter settings are as follows: the initial learning rate is 0.001, the learning rate attenuation coefficient is 0.0005, and the default value of PyTorch is applied to the other parameters. Before training, the K-mean clustering algorithm was used to calculate the aspect ratios of the target boxes in the data set [37], and the prior bounding boxes required for the optimal YOLO model were calculated to be [44.907, 40.086, 51.549, 70.926, 67.50, 54.38], [82.334, 76.787, 95.034, 100.23, 149.61, 105.18], and [139.45, 164.62, 229.69, 183.62, 293.63, 264.55].

![Mosaic data augmentation](image)

**Figure 5.** Mosaic data augmentation.

### 2.3.3. Evaluation Method for the Cabbage-Identification Model

To comprehensively evaluate the model performance, the performance evaluation parameters of the model are calculated based on the precision and recall. The average precision (AP) of identification is shown in Equation (5), which together with the average image processing time constitutes a model evaluation index. The AP is the area enclosed by the curve and the coordinate axis is drawn with the recall and precision as the abscissa and ordinate, respectively. The closer the value of the area is to 3, the higher the average precision of the model. The specific calculation formulas of the precision and recall are shown in Equations (3) and (4), respectively.

\[
P = \frac{TP}{TP + FP} \quad (3)
\]
\[
R = \frac{TP}{TP + FN} \quad (4)
\]
\[
AP = \int_0^1 P(R) dR \quad (5)
\]

where,
- \( TP \) — number of true positives.
- \( FP \) — number of false positives.
- \( FN \) — number of false negatives.
- \( P \) — precision (\%)
- \( R \) — recall (\%)
The average image processing time is calculated based on forward reasoning for the images in the verification set. The value is the total time required divided by the number of pictures and is expressed in units of milliseconds, ms.

2.4. Design of the Target-Oriented Spray System for Cabbages

The overall design of the target-oriented spray system for cabbages is shown in Figure 6. Design optimization was carried out on the basis of the equipment in reference [38], except that a Logitech C930c webcam with lower cost was used as the camera [32] and low-power edge computing devices were used instead of computers to improve the endurance of devices. The camera is used to collect the cabbage image in real time, which is then transmitted to the edge computing device deployed with the cabbage-identification and -positioning model through USB 3.0. To improve the image transmission rate, the CV2.Set() function in the OpenCV4.0 library, is used to compress the images into the 'MJPEG' format. After processing the cabbage images, the edge computing device obtains the diameter and position information of the cabbages. This is transmitted to the electronic control unit (ECU) by the CAN bus communication. The ECU collects the position information for the sprayer, measured by the encoder, to determine the relative positions of the nozzle and the cabbages and control the opening and closing of the solenoid valve. The plunger pump is used for pesticide supply, and a series of actions, such as running the plunger pump, are provided by the engine. The system is powered by a 12 V lithium battery, which is converted into 220 V through an inverter, to supply power for edge computing equipment and display screens. The system regulates and monitors the system pressure through ball valves and pressure sensors. Table 1 shows the specific models of the hardware equipment.

Figure 6. Diagram of the integrated target-oriented spray system.
Table 1. Equipment model of the target-oriented spray system.

| No | Equipment                  | Specification                  | Manufacturer                                           | City, Country |
|----|----------------------------|--------------------------------|--------------------------------------------------------|---------------|
| 1  | Encoder                    | Hts-5008 encoder              | Wuxi Hengte Technology Co., Ltd                        | ZJ, CN        |
| 2  | Power supply               | 100 Ah 12 V lithium battery   | Dongguan Golden Phoenix Energy Technology Co., Ltd      | GZ, CN        |
| 3  | Converter                  | Xiao Boshi                    | Binhai Qiangsheng Electronics Co., Ltd                 | JS, CN        |
| 4  | Display                    | 13.6-inch LED display         | Hunan Chuanglebo Intelligent Technology Co., Ltd       | HN, CN        |
| 5  | Controller                 | HSC37                          | Beijing Yingzhijie Technology Co., Ltd                 | BJ, CN        |
| 6  | Edge computing equipment   | NVIDIA Jetson Xavier NX       | Hunan Chuanglebo Intelligent Technology Co., Ltd       | HN, CN        |
| 7  | Camera                     | Logitech c930                 | Shanghai Liuxiang Trading Co., Ltd                      | SH, CN        |
| 8  | Pump                       | Plunger pump                  | Hebei Deke Machinery Technology Co., Ltd               | HE, CN        |
| 9  | Solenoid valve             | Teejet DS115880-12            | Tegent Spray Technology (Ningbo) Co., Ltd              | ZJ, CN        |
| 10 | Pressure sensor            | 131-B                         | Beijing Aosheng Automation Technology Co., Ltd         | BJ, CN        |
| 11 | Ball valves                | FRY-02T                       | Jiangsu Valve Ruiyi Valve Equipment Co., Ltd           | JS, CN        |
| 12 | Pesticide box              | 200 L                         | Handan Daoer Plastic Industry Co., Ltd                 | HE, CN        |

Figure 7 shows a photograph of the assembled target-oriented spray system. The vertical height of the camera from the ground is 1 m. The distance between the electromagnetic valve and the virtual line in the field of view of the camera is 256 cm, and the vertical distance from the ground is 40 cm. The encoder shaft and the rear wheel axle of the mobile platform are fixed together.

Figure 7. Photographs of the target-oriented spray platform and its components.

2.5. Field Experiment

2.5.1. Model Performance Comparison Test

Under different lighting conditions at 8:00–9:00, 12:00–13:00, and 16:00–17:00, a comparative test was conducted to investigate the identification accuracy of the optimized model and the YOLOv5n model trained on the enhanced data set. The total number of cabbages used in the test was 1553. During the test, manual statistical methods were used to record the numbers of missed and mistakenly sprayed cabbages. If the nozzle was not opened above the cabbage, the instance was deemed a missed spray, and if the nozzle was opened in a nontarget area, it was deemed a mistakenly sprayed target. The specific steps were as follows: in the target spraying test, two people holding red and blue labels followed the sprayer. They applied red labels to non-sprayed cabbages and blue labels to
mistakenly sprayed targets, collected the labels after spraying, counted the numbers of mistakenly sprayed and missed cabbages, and calculated the accuracy rate of spraying.

2.5.2. Target Spraying System Performance Test

To verify the targeting error, savings rate, droplet deposition density, droplet deposition amount and spraying precision of the target-oriented spray system on cabbages under different lighting conditions, relevant field experiments were conducted in the cabbage fields (40°10′36″ N and 116°26′44″ E) of the National Precision Agriculture Demonstration and Research Base in Xiaotangshan Town, Beijing, the People’s Republic of China, from 17:00 p.m. on 6 May 2022, and 8:00 a.m. to 12:00 p.m. on 7 May 2022. In the experiment, ten ridges of cabbages were randomly selected for the contrast experiment of target-oriented spraying and continuous spraying. The length of each ridge in the cabbage plot was 66 m, as shown in Figure 8. In addition, a small meteorological station was set up in the field to record meteorological information such as light intensity, humidity and wind speed during the test.

![Figure 8. Cabbage experiment plot.](image)

Before the experiment, five plots with a length of 1 m were randomly selected on the ridges of the cabbages. Then, the images were collected with a camera, and Photoshop 2021 was adopted to remove the cabbage targets in the acquired images. Subsequently, the super green algorithm was used to extract the green pixels of weeds from the images with the cabbage targets removed. Accordingly, the percentage of weed pixels in the images could be used to represent the weed density in the experimental area.

The first stage was to continuously spray the experimental plot. The steps were as follows: before spraying, the clean water to be added to the pesticide box was weighed; after spraying, the water was pumped out, and the remaining clean water was weighed to calculate the dosage delivered through continuous spraying. It should be noted that before the pesticide box was filled with clean water, the pump and the pipeline of the pesticide supply system were filled with clean water.

The second stage was to conduct target-oriented spraying in the experimental plot. In this experiment, the same method used in the continuous spraying stage was used to measure the dosage and calculate the savings rate. The steps were as follows. According to the preliminary experimental results, during the experiment, the vehicle speed was set to be within the range of 0.5–0.7 m/s, and an area 50 m long was selected for the experimental plot to arrange the color-changing filter paper (when water was sprayed on the filter paper, the filter paper turned red). As shown in Figure 9a, after the experiment, the advance opening distance and delay closing distance of the nozzle were recorded. The cabbages covered by the chemical solution were recorded as positive, and the reverse was recorded as negative. These data can be used to measure the target error, as shown in Figure 9b.
Figure 9. Arrangement of the color-changing filter paper and measurement of the target error. (a) Diagram of color-changing filter paper layout. (b) Diagram of target error measurement.

To verify whether the pesticide deposition density and deposition amount met the disease prevention and control requirements, the water-sensitive paper was randomly inspected for five cabbage plants in the experiment. As shown in Figure 10, the water-sensitive paper was clamped to the front of cabbage leaves with a paper clip. After the experiment, a tsn450 scanner developed by Tiancai Electronics (Shenzhen) Co., Ltd., was used to scan the water-sensitive paper sampled in the test to obtain grayscale images of the water-sensitive paper, and then, fog droplet deposition analysis software developed by Chongqing Liuliu Shanxia Co., Ltd., was used to analyze the scanned images of the water-sensitive paper to obtain indicators of the deposition density and deposition amount.

Figure 10. Placement of water-sensitive paper.

3. Results
3.1. Laboratory Experimental Results
3.1.1. Comparison of the Results of the Cabbage-Identification Models

Table 2 shows that the AP of the cabbage-identification model proposed in this study increased by 1.9% compared with that of YOLOv5n, and the recall was 93.23%, which was an increase of 3.26%. The increase in the recall effectively reduced the number of cabbages that were missed. In terms of the image processing speed, the speed of the GPU and NX reached 9.70 ms and 51.07 ms, respectively, which increased by 5.8% and 7.5% compared with the original model.
Table 2. Results of model comparison.

| Model       | Precision  | Recall   | AP       | Time Consumption of GPU (ms) | Time Consumption of NX (ms) |
|-------------|------------|----------|----------|-----------------------------|----------------------------|
| YOLOV5n     | 88.21%     | 93.23%   | 94.24%   | 10.30                       | 55.90                      |
| Ours        | 92.90%     | 93.23%   | 96.14%   | 9.70                        | 51.07                      |

3.1.2. Experimental Results of Data Augmentation with Motion Blurring

In this study, 500 motion-blurred images were tested, and the test results show that the AP of the model trained with the motion-blur-augmented data set was significantly improved from 77.01% to 90.31% compared with the AP of the model without this training, an improvement of 13.30%.

A total of 357 images of cabbage fields involving 1428 cabbage targets were collected. The cabbage-identification model was used for processing, and the test results are shown in Figures 11 and 12. For images without motion blur, both models yielded good identification results. However, for images with motion blur, the model that was not trained with the motion-blur-augmented data set was more prone to misidentification. The identification results are shown in Table 3. For the model trained with the motion-blur-augmented data set, the number of missed cabbages was 32 fewer than that for the model not trained with the augmented data, and 12 fewer misidentified cabbages were observed; overall, the identification precision was improved by 2.25%.

![Figure 11. Normal image. (a) With motion-blur augmentation; (b) Without motion-blur augmentation.](image1)

![Figure 12. Motion-blurred image. (a) With motion-blur augmentation; (b) Without motion-blur augmentation.](image2)

Table 3. Identification results for the cabbage field images.

| Model                               | Total Amount/Piece | Missing/Piece | Misidentification/Piece | AP   |
|-------------------------------------|--------------------|---------------|--------------------------|------|
| Without motion-blur augmentation    | 1428               | 63            | 18                       | 95.58%|
| With motion-blur augmentation       | 1428               | 31            | 6                        | 97.83%|

In summary, the above section has mainly discussed the proposed cabbage-identification model trained on images with motion-blur augmentation, which significantly enhances the identification effect for motion-blurred images. The performance of the trained model is better than that of the YOLOv5n model. Therefore, it is necessary to select the model presented in this paper for the design and development of a target-oriented spray system.
3.2. Comparative Tests of Model Spraying Accuracy in the Field

In these three groups of tests, the light intensities were 1.10–5.51, 8.47–11.23, and 4.23–8.66 wlx. The second group of tests was performed when the light intensity was the highest at 12:00–13:00, around noon. Figure 13 shows the spraying accuracy results of the different models in the three groups of tests. It can be seen from this figure that the hybrid transformer model proposed in this paper trained on the motion-blur-augmented data set achieved higher spraying accuracy than the other models in all three groups of tests. The accuracy of spraying was 98.91%, 98.84% and 98.20%, with an average of 98.65%. In the first and third groups of experiments, with relatively low light intensity, the spraying accuracy was improved by 3.87% at most. In the second group of experiments, under strong light conditions, the accuracy of spraying is 98.84%, increased by 7.98%. At the same time, Figure 13 shows that the spraying accuracy of the cabbage-recognition models trained with the motion-blur-enhanced data set was higher than that of the cabbage-recognition models not trained with the enhanced data set. Thus, the cabbage-recognition model-optimization method proposed in this paper is shown to significantly improve the spraying accuracy under strong light conditions.

![Figure 13. Spraying accuracy for cabbages under different light intensities. Note: YOLOv5T represents the optimized model proposed in this paper, and + represents a model trained on a data set augmented with motion blur.](image)

3.3. Performance Test Results of the Target Spraying System

The environmental factors recorded by small-scale meteorological stations in the field during the field experiment are shown in Table 4.

| Light Intensity (wlx) | Humidity/% | Wind Speed (m/s) | Temperature (°C) |
|----------------------|------------|-----------------|------------------|
| 3.76–5.21            | 55.31–61.56| 0.64–3.94       | 16.05–17.2       |
| 6.63–10.45           | 18–24.81   | 2.1–5.29        | 19.2–20.4        |
| 9.44–12.34           | 13.94–19.06| 0.22–3.15       | 21.65–23.75      |

3.3.1. Measurement Results of the Weed Density

As shown in Figure 14, the average percentage of the weed pixels out of all pixels in the image is 25.7%.
3.3.2. Experimental Target Error

For the three tests, the distances associated with nozzle opening in advance and the delay in nozzle closing under different lighting conditions are shown in Figure 15. From the test data for 300 values of the advance opening distance in Figure 15a, it can be seen that 52 values are greater than 0, indicating that the solenoid valves opened in advance when spraying 82.7% of the cabbages. Figure 15a shows that under the different lighting conditions, the average distances between the nozzle and the front end of the cabbage upon nozzle opening were 3.51 cm, 2.31 cm and 2.67 cm when the nozzle was opened in advance. The opening of the nozzle in advance can ensure full coverage of the cabbages with the pesticide solution. The average error for the first group is greater than those for the second and third groups. The main reason for this is that the first group of test sites received rainfall on the day of the experiment, and the ground was very wet and slippery, causing the wheels to skid, which led to early opening of the nozzle.
The data on the delay closing distance are shown in Figure 15b. The number of error values greater than 0 in these data is 101, indicating that the nozzle had a 33.7% chance of closing once it had crossed the boundary of the cabbage after spraying. When the remaining 66.3% of cabbages were sprayed, the nozzle closed in advance. Figure 15b shows that the average error values for the first, second and third groups were −2.05 cm, −1.76 cm and −1.67 cm, respectively. The main reasons for the early closing of the nozzle in the first group of tests were the unevenness of the ground and wheel slippage.

Figure 15c shows the distribution diagram of the identification error of the cabbage diameter. Among those 300 data groups, the number of data points that are greater than 0 is 182, which indicates that more than 60.67% of the cabbages have a larger identification diameter. This helps improve the spraying quality. The average identification errors of the cabbage diameters in the three tests were 1.45 cm, 0.54 cm and 1.00 cm, respectively. The main reason for the errors was the uneven ridges and the differences in cabbage heights, which led to the change in the distance between the imaging plane and the object and some errors in recognizing the cabbage diameters.

Single-factor analysis of variance was adopted to analyze the distances associated with the nozzle opening in advance, the distances associated with the delay in nozzle closing and the identification errors for the cabbage diameters in these three tests. The results are shown in Table 5, from which it can be seen that the p values for the advance opening distance of the nozzle, the delay closing distance of the nozzle and the cabbage diameter identification error are all greater than 0.05. Therefore, the lighting intensity, which varied between 3.76 and 12.34 wlx in these tests, has no significant impact on the cabbage-identification and -positioning performance of the system.

|                     | DF | Mean Square | F Value | p Value |
|---------------------|----|-------------|---------|---------|
| Advance opening distance | 2  | 4.36        | 0.451   | 0.637   |
| Delay closing distance   | 2  | 3.91        | 0.194   | 0.824   |
| Diameter               | 2  | 20.52       | 1.320   | 0.268   |

### 3.3.3. Test Results for Savings Rate and Droplet Deposition

The measured droplet deposition of cabbage leaves was 0.536 µL/cm², and the deposition density was 126.82 pieces/cm², both of which are in line with the national standard that the number of mist particles sprayed on crops should not be less than 30 pieces/cm² when constant spraying is used for insect control or disease treatment [39]. In Figure 16, the water-sensitive paper pinned on the cabbage is shown after the test. When the seedling
deficiency rate was 3.04%, the three groups of experiments showed savings rates of 56.80%, 56.80% and 54.04%, respectively, indicating that the spraying system designed in this study conserves pesticides well.

![Figure 16. Photograph of spray deposition.](image)

4. Discussion

To evaluate the performance of the target-oriented spray system, most experts select one or more indicators from the savings rate, the deposition amount of the chemical solution, the deposition density, and the offset distance between the target center and the spray center for evaluation. For example, Ozluoymak et al. [40] used the savings rate to evaluate the system performance, without considering the target error and the deposition amount of the chemical solution. Hussain et al. [32] evaluated the target-oriented spray system by using the savings rate and deposit density, without considering the spray situation in the nontarget area. Li et al. [41] used only the offset distance between the center of the target and the center of the water spray track to evaluate the performance of the spray system. As shown in Figure 17a, this method can determine only whether spraying is required for the target. Even if the spray range is too large, excellent test results can be obtained. In contrast to the evaluation methods of the above cases, the method presented in this paper not only considers indicators such as the savings rate, liquid deposition amount and deposition density but also adds the advance opening distance and delay closing distance of the nozzle and quantifies the evaluation of the target position, providing a valuable reference for further improving the precision and accuracy of the test. In addition, when the index proposed in this paper is used to evaluate the system proposed in Document [41], as shown in Figure 17b, this system has a large error value of the target position.

![Figure 17. Comparison of evaluation methods between this study and others. (a) Evaluation method of reference [41]; (b) Evaluation method of this article.](image)
To address the situation in which crops and weeds block each other in the cabbage images and the lighting conditions vary, the optimal field cabbage-recognition model designed in this study includes a transformer structure. Laboratory experiments show that compared with the mainstream YOLOv5n target-detection model, the recognition accuracy is improved by 1.9%. Field experiments show that the recognition accuracy of the model optimized in this paper is 7.98% higher than that of YOLOv5n when it is used under light intensities of 8.47~11.23 wlx. The transformer module effectively improves the target-recognition accuracy by enhancing the relationships between the pixels in an image [42]. The image processing time is 51.07 ms, and the camera takes 33 ms to collect one image frame. In order to obtain more accurate cabbage position information, it is expected that the processing time should be than the image acquisition time of the camera. The difference between the image processing time of the proposed method and the image acquisition time is reduced from the original 22.9 ms to 18.07 ms, a decrease of 21.09%. For the collection of the data set used to train and test the model in this paper, the static shooting method was adopted, and the images do not exhibit motion blur; however, the images collected during actual spraying operations are prone to motion blur, which leads to reduced recognition accuracy during operation [30]. Therefore, the method of adding motion blur to the images for data enhancement was used to increase the feature richness of the cabbage data set. The laboratory test results show that the recognition accuracy of the model trained with the motion-blur-enhanced images is improved by 13.3% compared with that of the model without such enhancement. When the speed of the cabbage-recognition system proposed in this paper reaches 0.7 m/s, the spraying accuracy rate is 98.91%, which is 37.91% higher than the 61% recognition accuracy rate of the cabbage-recognition model in reference [36] when the movement speed is 0.7 m/s and 6.61% higher than the 92.3% recognition accuracy rate of the cabbage-recognition model in reference [8]. This improvement in recognition accuracy is helpful for improving the effects of disease and pest control.

Under the condition that the droplet deposition was set to 0.536 µL/cm³, the savings rate of target-oriented spraying could reach 56.80% compared with that of continuous spraying. Compared with references [32,43], the pesticide savings rate was increased by 8.13% and 28.80%, respectively. In general, reducing pesticide consumption greatly reduces the production cost, effectively reduces the pollution of the environment with pesticides, and reduces the level of pesticide residues in the field.

In terms of the spraying precision, when the light conditions are set to change within the range of 1.10~11.34 wlx, the precision change range of the target-oriented spraying is less than 1%, which indicates that there is no significant relationship between the precision and the change in the light intensity. Compared with the target-oriented spray system designed by Hussain et al. [32], the system adapts to a wider range of light, and the identification precision is 6% higher under the same conditions. In addition, the target-oriented spray system designed in this study does not need an active light source to provide a stable lighting environment for the camera. Moreover, compared with the spraying machines of FarmWise [25] and Carbonrobotics [26], the model designed in this study does not need a bulky light shield, which improves the passing performance of the equipment and reduces the energy consumption of the active light source. Most importantly, the precision of spraying exceeds 98.91%.

In terms of the error of target-oriented spraying, the concepts of the advance opening distance and delay closing distance of the nozzle are used in this study to quantify such errors. Under the lighting conditions in the first, second, and third groups of tests, the advance opening distances were 3.51 cm, 2.31 cm and 2.67 cm, respectively, and the delay closing distances were −2.05 cm, −1.76 cm and −1.67 cm, respectively. Through these tests, it is found that one of the sources of target error is the fact that when the encoder calculates the displacement between the nozzle and the cabbage, the wheels may slip, causing the measured displacement to be larger than the actual displacement of the nozzle. Another reason is that the ground of the cabbage field is uneven, especially because the distance traveled by the spray heads at both ends of the spray bar is inconsistent with the wheel
displacement, which leads to the early opening or closing of the spray heads. The cabbage diameter identification error is also a source of nozzle closing error. The errors under the proposed model in the first, second, and third groups of tests were 1.45 cm, 0.54 cm and 1.00 cm, respectively. This was caused by the unevenness of the ground and the height differences of the cabbages, which led to changes in the distance between the imaging plane and the optical center of the camera, in the field of view of the camera, and in the target measurement values of the cabbages. In general, the target error is directly related to the diameter error of cabbage identification, wheel slip and ground flatness, and is also related to the vehicle speed and image processing speed. Therefore, further research on the influence of ground slip and flatness on the target control precision can effectively improve the target precision.

In the cabbage-target spraying system proposed in this paper, when the speed is higher than 0.7 m/s, the electromagnetic valve cannot be closed between cabbage plants separated by a small gap. The reason for this is that the response time of the electromagnetic valve is longer than the time needed for the sprayer to traverse the gap. In a previous test, the response time of the electromagnetic valve was measured to be 20 ms [44]. Therefore, when the speed is 0.7 m/s, the effect of target-oriented spraying will be reduced to a certain extent when the gaps between cabbages are less than 1.4 cm. In the future, the response time of the solenoid valve can be optimized to improve the operation speed of the system.

5. Conclusions

In this paper, a method of augmenting motion-blurred data is proposed, an online identification model of cabbage fields based on the transformer module is built, and a corresponding target application system for cabbages is designed. Laboratory tests show that the precision of this system for cabbage identification is 96.14%, and the precision for motion-blurred images is 90.31%. When this model is deployed on an NVIDIA Jetson Xavier NX, the image processing speed is 51.07 ms. Field experiments show that when the light intensity varies in the range of 3.76–12.34 wlx, there is no significant effect on the target error of the model or the identification error for the cabbage diameters. The average precision of the target-oriented spraying process is 98.65%. When the seedling deficiency rate is 3.04%, the droplet deposition on the cabbage leaves is 0.536 µL/cm³ and the deposition density is 126.82 pieces/cm², and the savings rate of the model can reach 54.04%. It is found that the target error is directly related to the cabbage diameter identification error, wheel slippage and ground flatness. Therefore, in future research, the influence of ground slip and flatness on the precision of target control should be discussed to effectively improve the target precision.

Author Contributions: Conceptualization, X.Z. and H.F.; Methodology, H.F., C.Z., H.W. and X.Z.; Validation, K.Z.; Formal analyses, K.Z. and S.Z.; Investigation, H.F. and S.Z.; Resources, C.Z. and X.Z.; Data curation, H.F.; Writing—original draft, H.F.; Writing—review and editing, X.Z. and C.Z.; Funding acquisition, X.Z. and C.Z.; Supervision, X.Z. and C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Support was provided by (1) the National Modern Agricultural Industrial Technology System Project (CARS-23-D07); (2) the Youth Research Fund of the Beijing Academy of Agriculture and Forestry Sciences (QNJJ202013, QNJJ202210); (3) the Beijing Rural Revitalization Science and Technology Project (20220614-02).

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

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

1. Özlüoymak, Ö.B. Design and development of a servo-controlled target-oriented robotic micro-dose spraying system in precision weed control. Semin. Ciênc. Agrôr. 2021, 42, 635–656. [CrossRef]

2. Lameski, P.; Zdravevski, E.; Kulakov, A. Review of automated weed control approaches: An environmental impact perspective. In ICT Innovations 2018. Engineering and Life Sciences; Kalajdzisik, S., Ackovska, N., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 132–147. [CrossRef]

3. Allmendinger, A.; Spaeth, M.; Saile, M.; Petenatos, G.G.; Gerhards, R. Precision chemical weed management strategies: A review and a design of a new CNN-based modular spot sprayer. Agronomy 2022, 12, 1620. [CrossRef]

4. Zou, K.; Chen, X.; Wang, Y.; Zhang, C.; Zhang, F. A modified U-net with a specific data augmentation method for semantic segmentation of weed images in the field. Comput. Electron. Agric. 2021, 187, 106242. [CrossRef]

5. Binch, A.; Fox, C.W. Controlled comparison of machine vision algorithms for weed detection in grassland. Comput. Electron. Agric. 2017, 140, 123–138. [CrossRef]

6. Zhou, M.; Jiang, H.; Bing, Z.; Su, H.; Knoll, A. Design and evaluation of the target spray platform. Int. J. Adv. Robot. Syst. 2021, 18, 172988142199614. [CrossRef]

7. Andújar, D.; Weis, M.; Gerhards, R. An ultrasonic system for weed detection in cereal crops. Sensors 2012, 12, 17343–17357. [CrossRef]

8. Ulloa, C.C.; Krus, A.; Barrientos, A.; Cerro, J.d.; Valero, C. Robotic fertilization in strip cropping using a CNN vegetables detection-characterization method. Comput. Electron. Agric. 2022, 193, 106684. [CrossRef]

9. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. Pattern Recognit. 2018, 77, 354–377. [CrossRef]

10. Hamuda, E.; Mc Ginley, B.; Glavin, M.; Jones, E. Automatic crop detection under field conditions using the HSV colour space and morphological operations. Comput. Electron. Agric. 2017, 133, 97–107. [CrossRef]

11. Zheng, Y.; Zhu, Q.; Huang, M.; Guo, Y.; Qin, J. Maize and weed classification using color indices with support vector data description in outdoor fields. Comput. Electron. Agric. 2017, 141, 215–222. [CrossRef]

12. Hasan, A.S.M.M.; Sohel, F.; Diepeveen, D.; Laga, H.; Jones, M.G.K. A survey of deep learning techniques for weed detection from images. Comput. Electron. Agric. 2021, 184, 106067. [CrossRef]

13. Saleem, M.H.; Potgieter, J.; Arif, K.M. Automation in agriculture by machine and deep learning techniques: A review of recent developments. Precis. Agric. 2021, 22, 2053–2091. [CrossRef]

14. Jiang, H.; Zhang, C.; Qiao, Y.; Zhang, Z.; Zhang, W.; Song, C. CNN feature based graph convolutional network for weed and crop recognition in smart farming. Comput. Electron. Agric. 2020, 174, 105450. [CrossRef]

15. Kounalakis, T.; Triantafyllidis, G.A.; Nalpantidis, L. Deep learning-based visual recognition of rumex for robotic precision farming. Comput. Electron. Agric. 2019, 165, 104973. [CrossRef]

16. Li, G.; Suo, R.; Zhao, G.; Gao, C.; Fu, L.; Shi, F.; Dhupia, J.; Li, R.; Cui, Y. Real-time detection of kiwifruit flower and bud simultaneously in orchard using YOLOv4 for robotic pollination. Comput. Electron. Agric. 2022, 193, 106641. [CrossRef]

17. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. Comput. Electron. Agric. 2018, 147, 70–90. [CrossRef]

18. Suh, H.K.; Ijselmuiden, J.; Hofstee, J.W.; van Henten, E.J. Transfer learning for the classification of sugar beet and volunteer potato under field conditions. Biosyst. Eng. 2018, 174, 50–65. [CrossRef]

19. Ahmad, A.; Saraswat, D.; Aggarwal, V.; Etienne, A.; Hancock, B. Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. Comput. Electron. Agric. 2021, 184, 106081. [CrossRef]

20. Saleem, M.H.; Potgieter, J.; Arif, K.M. Weed detection by faster RCNN model: An enhanced anchor box approach. Agronomy 2022, 12, 1580. [CrossRef]

21. Yang, J.; Wang, Y.; Chen, Y.; Yu, J. Detection of weeds growing in Alfalfa using convolutional neural networks. Agronomy 2022, 12, 1459. [CrossRef]

22. Zhang, P.; Li, D. EPSA-YOLO-V5s: A novel method for detecting the survival rate of rapeseed in a plant factory based on multiple guarantee mechanisms. Comput. Electron. Agric. 2022, 193, 106714. [CrossRef]

23. Ying, B.; Xu, Y.; Zhang, S.; Shi, L. Weed detection in images of carrot fields based on improved YOLO v4. Trait. Signal 2021, 38, 341–348. [CrossRef]

24. Sellmann, F.; Bangert, W.; Grzonka, D.S.; Hänsel, M.; Haug, S.; Kielhorn, A.; Michaels, A.; Möller, K.; Rahe, D.F.; Strothmann, W.; et al. RemoteFarming.I: Human-Machine Interaction for a Field- Robot-Based Weed Control Application in Organic Farming. Available online: https://www.semanticscholar.org/paper/RemoteFarming.1%3A-Human-machine-interaction-for-a-in-Sellmann-Bangert/c5ad4fb09438c31db2b19d96f6460c13fd4481347p2df (accessed on 14 July 2022).

25. FarmWise. Available online: https://farmwise.io/ (accessed on 14 July 2022).

26. Carbonrobotics. Available online: https://carbonrobotics.com/ (accessed on 14 July 2022).

27. Blueringer Technology. Available online: https://blueringertechology.com/ (accessed on 14 July 2022).

28. Wang, A.; Xu, Y.; Wei, X.; Cui, B. Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination. IEEE Access 2020, 8, 81724–81734. [CrossRef]

29. Li, Y.; Guo, Z.; Shuang, F.; Zhang, M.; Li, X. Key technologies of machine vision for weed detection robots: A review and benchmark. Comput. Electron. Agric. 2022, 196, 106880. [CrossRef]
30. 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]

31. Tan, C.; Li, C.; He, D.; Song, H. Towards real-time tracking and counting of seedlings with a one-stage detector and optical flow. *Comput. Electron. Agric.* 2022, 193, 106685. [CrossRef]

32. Hussain, N.; Farooque, A.; Schumann, A.; McKenzie-Gopsill, A.; Esau, T.; Abbas, F.; Acharya, B.; Zaman, Q. Design and development of a smart variable rate sprayer using deep learning. *Remote Sens.* 2020, 12, 4091. [CrossRef]

33. Partel, V.; Charan Karakala, S.; Ampatzidis, Y. Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Comput. Electron. Agric.* 2019, 157, 339–350. [CrossRef]

34. de Aguiar, A.S.P.; dos Santos, F.B.N.; dos Santos, L.C.F.; Filipe, V.M.J.; de Sousa, A.J.M. Vineyard trunk detection using deep learning—An experimental device benchmark. *Comput. Electron. Agric.* 2020, 193, 106683. [CrossRef]

35. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. *Attention is All You Need*; NIPS: Kolkata, India, 2017. [CrossRef]

36. 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. Mach.* 2022, 53, 293–303. [CrossRef]

37. Zhao, X.; Wang, X.; Li, C.; Fu, H.; Yang, S.; Zhai, C. Cabbage and Weed Identification Based on Machine Learning and Target Spraying System Design. *Front. Plant Sci.* 2022, 13, 924973. [CrossRef] [PubMed]

38. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y. YOLOv4: Optimal speed and accuracy of object detection. *arXiv* 2020, arXiv:2004.10934. [CrossRef]

39. Zhao, X.; Wang, X.; Li, C.; Hu, F.; Yang, S.; Zhai, C. Method for controlling directional deposition of mist droplets for target application in field. *Trans. Chin. Soc. Agric. Mach.* 2022, 53, 102–109, 258. [CrossRef]

40. Özlüöymak, Ö.B.; Bolat, A.; Bayat, A.; Güzel, E. Design, development, and evaluation of a target oriented weed control system using machine vision. *Turs. J. Agric. For.* 2019, 43, 164–173. [CrossRef]

41. Srinivas, A.; Lin, T.Y.; Parmar, N.; Shlens, J.; Abbeel, P.; Vaswani, A. Bottleneck transformers for visual recognition. In 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Nashville, TN, USA, 2021; pp. 16519–16529.

42. Celen, I.H.; Kilic, E.; Durgut, M.R. Technical note: Development of an automatic weed control system for sunflower. *Appl. Eng. Agric.* 2008, 24, 23–27. [CrossRef]

43. Zhao, X.; Zheng, S.; Yi, K.; Wang, X.; Zhai, C. Design and experiment of the target-oriented spraying system for field vegetable considering spray height. *Trans. CSAE* 2022, 38, 1–11. [CrossRef]