



Article

The Extraction Method of Navigation Line for Cuttage and Film Covering Multi-Functional Machine for Low Tunnels

Yumeng Li ^{1,2,3}, Yanjun Zhu ^{1,2,3} , Shuangshuang Li ^{1,2,3} and Ping Liu ^{1,2,3,*} ¹ Shandong Agricultural Equipment Intelligent Engineering Laboratory, Tai'an 271000, China² Shandong Provincial Key Laboratory of Horticultural, Machinery and Equipment, Tai'an 271000, China³ College of Mechanical and Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China

* Correspondence: liupingsdau@126.com

Abstract: Aiming at the problem of low intelligence in the automatic navigation of the cuttage and film covering multi-functional machine for low tunnels, this study proposed a navigation line extraction method based on the improved YOLOv5s model, which can achieve the accurate extraction of navigation lines based on two planting methods of seedling transplanting and direct seeding. Firstly, we pre-processed the acquired images using inverse perspective transformation. Next, the Coordinate Attention and Ghost modules were applied to improve the YOLOv5s architecture, increasing the detection accuracy and speed of field targets. Finally, we extracted the feature points and fit the navigation lines based on the shape features of the targets using the geometric method. The experimental results showed that, compared with other algorithms, the accuracy of the proposed algorithm could reach more than 96%, the accuracy of navigation line extraction reached 98%, and the average detection time was 51 ms. The proposed method was robust and universal, and it can provide reliable navigation paths for the cuttage and film covering multi-functional machine.

Keywords: navigation line; YOLOv5; image processing; computer vision



Citation: Li, Y.; Zhu, Y.; Li, S.; Liu, P. The Extraction Method of Navigation Line for Cuttage and Film Covering Multi-Functional Machine for Low Tunnels. *Inventions* **2022**, *7*, 113. <https://doi.org/10.3390/inventions7040113>

Academic Editor: Anastasios Doulamis

Received: 12 November 2022

Accepted: 29 November 2022

Published: 2 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The key to realize the intelligent operation of cuttage and film covering multi-functional machines for low tunnels is automatic navigation technology. Machine vision navigation is the focus of research to realize the automatic navigation of cuttage and film covering multi-functional machines for low tunnels due to its high accuracy, low cost, better flexibility and continuity [1–3]. The key of machine vision navigation is the accurate extraction of navigation lines in the complex field environment [4–7]. Currently, the extraction of navigation lines in the field is mainly applied for mature crops, where crop row lines are extracted while identifying crop features, and then navigation lines are fitted [8]. The operation time of cuttage and film covering multi-functional machines for low tunnels is usually during the seed sowing period and the early stage of crop growth. The reference targets for extracting navigation lines are seed pits, field monopolies and seedlings. The target features are similar to farmland features, and it is difficult to extract navigation lines accurately and quickly using existing methods.

Deep learning techniques are capable of mining deep features of collected data in complex farming environments with high robustness. Therefore, deep learning techniques are widely used in various aspects of agriculture [9]. Many deep learning algorithms, such as YOLO [10–12], Faster R-CNN [13], Mask R-CNN [14] and UNet [15], have been applied to tasks such as weed localization [16], pest and disease identification [17] and fruit ripeness detection [18]. In addition, deep learning techniques have been applied to field navigation line extraction due to their advantages of high detection accuracy as well as fast detection speed.

There has been a great deal of research on deep learning techniques in field navigation line extraction to improve the detection accuracy while making the selection of target feature

points more realistic. Zongbin Gao used an optimized pair of YOLOv3 Tiny-3p models for kiwifruit trunk detection, and used a visualization method to analyze the features of kiwifruit trunk images to effectively distinguish feature differences between trunks and water pipes, improving the detection accuracy [19]. Bell et al. used a semantic segmentation method based on convolutional neural networks to segment kiwifruit orchard roads or trees for navigation path extraction [20]. André et al. used a deep learning approach to detect grape trunk and vine trunk features to solve the problem of feature extraction in vineyards [21]. Tan et al. performed the detection of seedling targets and were able to obtain information on the location of seedlings while counting [22], by calculating the centroid of the seedling detection frame to fit the navigation line. Louis et al. simplified the CSP network in the YOLOv4 model for the early stages of maize and bean crops, added an anchor point-based prediction head and implemented stem detection and localization based on the stem localization method to make the feature points more closely match the seed sowing position and improve the reliability of the navigation line fitting [23]. However, the target detection algorithms for field monopoles, seedlings and seed pits still have the problem of low detection accuracy. In addition, the high complexity and many parameters of many models lead to their insufficient real-time performance. Therefore, it is necessary to design a field target detection algorithm that meets the requirements of real-time recognition while ensuring the detection accuracy. In this paper, we propose a navigation line extraction method based on an improved YOLOv5s model for field target recognition and navigation line extraction, taking different planting methods of direct seeding and seedling transplanting as the research objects. This method can provide technical support for the real-time and accurate extraction of navigation lines for the cuttage and film covering multi-functional machine.

2. Materials and Methods

2.1. Image Acquisition

In the study, we use a cuttage and film covering multi-functional machine for low tunnels (Figure 1) as the image acquisition platform. The image acquisition equipment uses an easy-to-disassemble and portable Jereh Microcom DW800_2.9 mm camera, which is installed in the front of the self-propelled trellis machine frame at a distance of 0.8 m from the ground; the horizontal angle between the camera's optical axis and the ground is 35 degrees, and the specific installation details are shown in Figure 2. Image acquisition was performed using VideoCap software, and the hardware was a Lenovo laptop with an Intel Core i5-4200 CPU, 8 G of RAM and Windows 10 64-bit system.



Figure 1. The structure of cuttage and film covering multi-functional machine for low tunnels.



Figure 2. Field images: (a) direct seeding strip sowing method; (b) direct seeding hole sowing method; (c) seedling transplanting method; (d) seedling transplanting method with weeds.

The images were collected at the agronomy experimental station of Shandong Agricultural University with a resolution of 800×600 pixels and a format of JPG, totaling 1000 images. The images were classified into four types of planting methods: direct seeding strip sowing method, as shown in Figure 2a; direct seeding hole sowing method, as shown in Figure 2b; seedling transplanting method without weeds, as shown in Figure 2c; and seedling transplanting method with weeds, as shown in Figure 2d. The distance between the center line of each planting row and plant spacing was 60 cm and the field ridge spacing was 200 cm.

2.2. Image Pre-Processing

For the problem wherein the feature area and scale of the field ridge target vary greatly, the shape distortion caused by the camera distortion and shooting angle, etc., if the target is directly labeled and detected, there are some irrelevant areas in the labeled frame with similar features to the field ridge target, which affects the detection accuracy of the field ridge target. Therefore, inverse perspective transformation is used to process the field image dataset, as shown in Figure 3. In the top view of the obtained field image, the field ridge targets are in a regular rectangular shape, and the shape features of the field ridge targets can be accurately extracted while the irrelevant areas are eliminated during labeling, which eliminates the influence of the above factors and improves the detection accuracy of the field ridge targets. The detection accuracy of other targets is not affected, while the detection accuracy of field ridge targets is improved because the features of other targets, such as seedlings and seed pits, change less before and after the inverse perspective transformation.

The pre-processed images were divided into a training set, validation set and test set according to 5:3:2. The minimum outer rectangular box of the field targets in the images was labeled using Labeling to ensure that the rectangle contained as little background area as possible; then, the field ridge, seed pit, seedling without weed, seedling with weed and weed targets were named “ridge”, “seed”, “plant-1”, “plant-2” and “plant”, respectively.

Due to uncertainties such as shooting angles and weather, resulting in a complex environment for image acquisition, a large number of images containing field ridges, seedlings and seed pit targets are required in the training process in order to improve the accuracy and robustness of the training model, enhance the generalization ability of the model and prevent overfitting, which leads to poor detection. The number of images in the dataset was increased using image augmentation methods such as panning, changing

brightness, rotating angle and mirroring. The image augmentation results are shown in Figure 4.

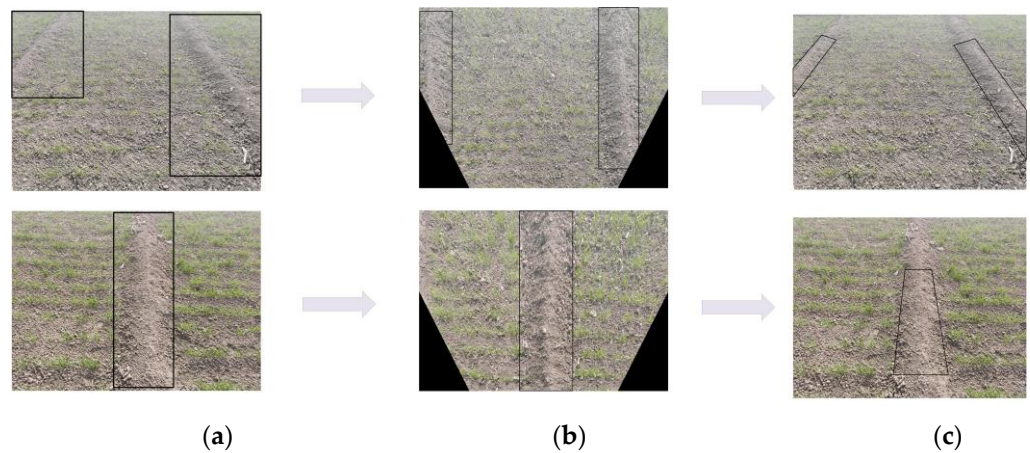


Figure 3. Schematic diagram of inverse perspective transformation image navigation line extraction: (a) original image; (b) inverse perspective transformation; (c) marking results. Note: The black box indicates the shape and position of the field ridge.

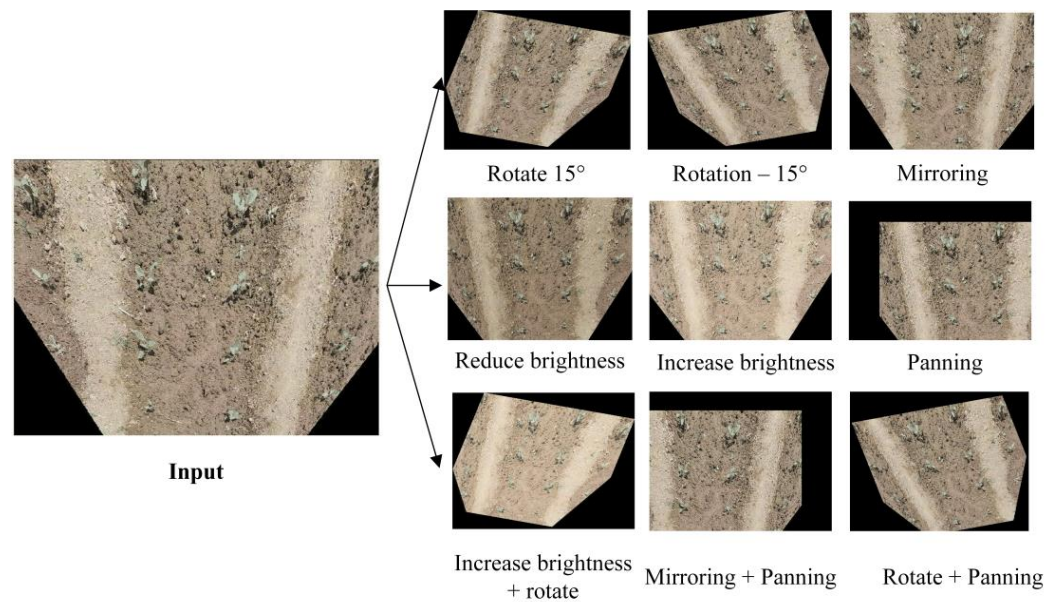


Figure 4. Data amplification images.

2.3. Improved YOLOv5s Model

The convolutional neural network can not only extract low-level features such as texture, shape, contour and color, but also can further extract abstract features with strong classification and detection capabilities. YOLOv5s is the smallest version of the YOLOv5 series, with a size of only 14.50 Mb. On the other hand, the YOLOv5s algorithm uses neural networks to learn the features needed for each type of target adaptively, with high detection accuracy and fast inference, with the fastest detection speed reaching 140 frames/s. The main structure of YOLOv5s consists of the input, backbone, neck network layer and head detection. As the navigation line extraction of the cuttage and film covering multi-functional machine for low tunnels possesses high requirements in terms of the accuracy and real-time nature of the field target detection model, in order to improve the detection accuracy and detection speed of field ridge, seedling and seed pit targets, two improvements are designed for the field target awareness detection model based on the YOLOv5s architecture.

- (1) To solve the problem of false or missed detection of field targets in images due to scale changes and different shapes, the Coordinate Attention (CA) mechanism is introduced into the YOLOv5s architecture to improve the detection accuracy of field targets by learning target features while suppressing non-target features.
- (2) The backbone network of the YOLOv5s architecture is improved by using the Ghost module to reduce the number of parameters and computation while eliminating invalid and duplicate feature maps to achieve an effective increase in detection speed while taking into account the detection accuracy.

The improved YOLOv5s architecture is shown in Figure 5

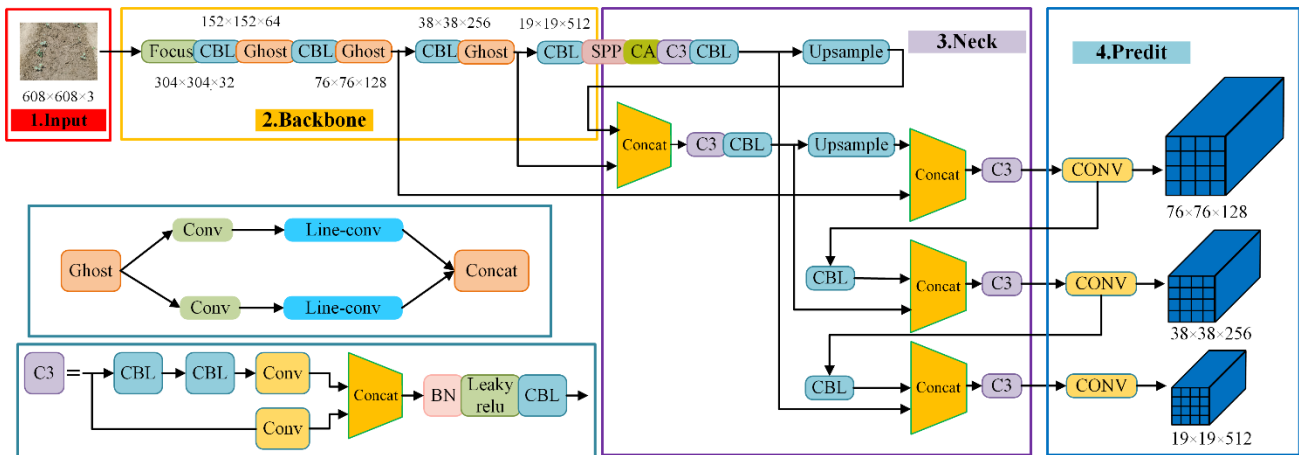


Figure 5. The improved YOLOv5s architecture.

2.3.1. Attention Mechanism

Since field ridge and seed pit targets are susceptible to the influence of a similar environment within the labeled frame, which can lead to false detection and missed detection, in order to further improve the accuracy and model performance of field target detection, the CA mechanism is introduced into the YOLOv5s architecture to learn and assign weights to the importance of each local feature, to achieve the purpose of enhancing the attention of field targets while suppressing the influence of similar environments.

The CA mechanism is able to consider both the inter-channel relationships and location information of features. It can capture long-term dependencies along one spatial direction while retaining accurate location information along another spatial direction, which helps to locate the target of interest more accurately. The CA module is added at the end of the backbone. Firstly, to obtain the aggregated feature maps of the image width and height, the CA module divides the input feature maps into two directions—width and height, respectively—for global average pooling to obtain a pair of direction-aware output feature maps. This path can be calculated based on Equations (1) and (2), as shown below.

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

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \tag{2}$$

where W, H and c are the width, height and dimension of the input feature map, respectively. x is the given input. $z_c^h(h), z_c^w(w)$ are the outputs of the c -channel with height h and width w , respectively.

Then, the feature maps in the width and height directions of the obtained global perceptual field are stitched together, after which they are fed into a convolution module with a shared convolution kernel of 1×1 to reduce their dimensionality to the original C/r . The batch-normalized feature map F_1 is then fed into the sigmoid activation function

to obtain a feature map f in the shape of $1 \times (W + H) \times C/r$. This path can be calculated based on Equation (3).

$$f = \delta(F_1([z^h, z^w])) \tag{3}$$

where $[,]$ is the concatenate operation along the spatial dimension, δ is the nonlinear activation function, f is the intermediate feature mapping that encodes the spatial information in the horizontal and vertical directions, F_1 is the 1×1 convolutional transform function, and r is the scaling rate.

Then, we decompose f into 2 separate tensors along the spatial dimensions $f^h \in R^{C/r \times H}$ and $f^w \in R^{C/r \times W}$. We use the other two 1×1 convolution transforms F_h and F_w to transform f_h and f_w into tensors with the same number of channels as the input X , respectively. The output is as shown in Equation (4).

$$\begin{aligned} g^h &= \sigma(F_h(f^h)) \\ g^w &= \sigma(F_w(f^w)) \end{aligned} \tag{4}$$

where σ is the sigmoid activation function.

In order to reduce the complexity and computational effort of the model, the number of channels is reduced by using an appropriate reduction ratio. The output g^h and g^w is expanded and used as the weight of attention, respectively. Then, the final output Y of the CA module can be written as shown in Equation (5).

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \tag{5}$$

The YOLOv5s architecture, with the introduction of the CA module, is able to accurately identify seedling and seed pit features with small features while detecting field ridges that are easily missed. The structure of the CA module is shown in Figure 6.

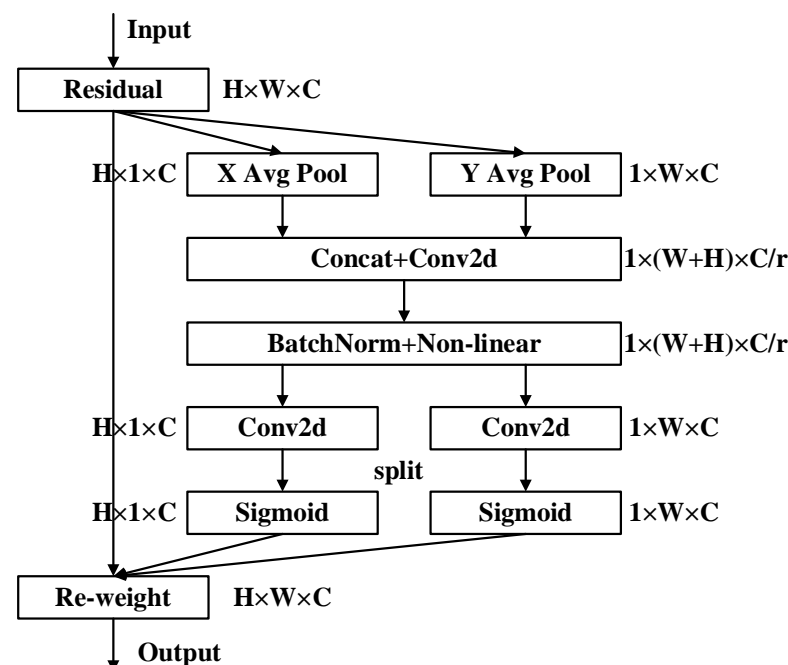


Figure 6. CA module structure diagram.

2.3.2. Backbone Network Optimization

Since the field target detection model of the cuttage and film covering multi-functional machine for low tunnels not only needs to accurately identify field targets in various situations in complex field environments, but also needs to compress the size of the model as much as possible to facilitate deployment in hardware devices, the study optimizes and improves the backbone network of the YOLOv5s architecture. The number of network

weight parameters and their size are reduced to achieve a lightweight and improved design of its field target recognition network while ensuring detection accuracy.

The backbone network of the YOLOv5s architecture contains four bottleneck CSP modules containing multiple convolutional layers. Although the convolutional operation can extract features in the image, the large number of convolutional kernels will lead to an increase in the number of parameters in the recognition model, while the CA module added in the YOLOv5s architecture increases the computational volume and detection time, although it can improve the accuracy of target detection. To ensure the target detection accuracy while reducing the detection time and computation, the backbone network needs to be optimized. The backbone network in the YOLOv5s architecture, when extracting feature maps from the input images, will have a large number of duplicate or feature-obscure feature maps that can be eliminated, as shown in Figure 7. To reduce the processing of such redundant feature maps, a lightweight convolutional network Ghost module that obtains a large number of feature maps through a small number of calculations is used to generate a large number of feature maps by combining feature maps using linear convolution while reducing the number of convolutional operation channels, and Figure 8 shows the structure of the Ghost module schematically. Compared with other networks, it is able to obtain higher accuracy with the same amount of computation, and it requires less computation with similar accuracy in the image classification task.

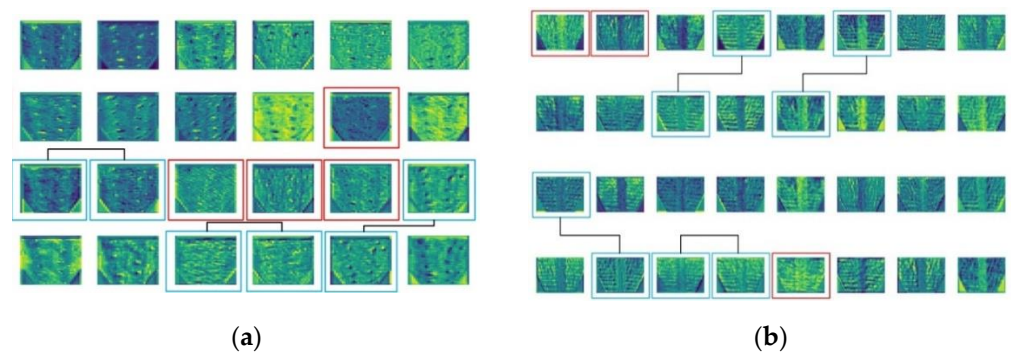


Figure 7. Selected feature maps extracted from field images: (a) selected feature maps for seedling transplanting image extraction; (b) partial feature map extracted from direct seeded images. Note: The red box is part of the obscure features, and the blue box is part of the repeated features.

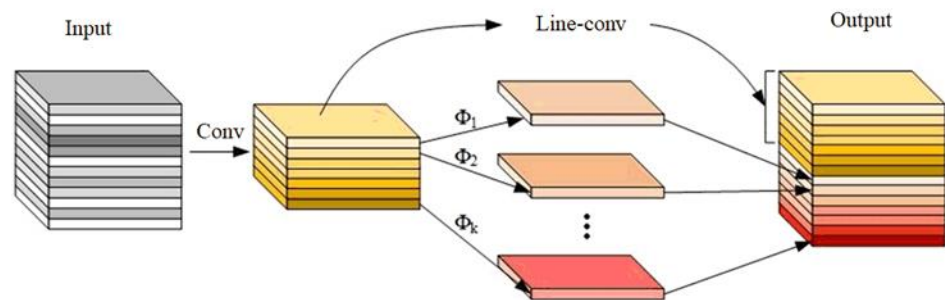


Figure 8. Ghost module structure schematic.

2.3.3. Network Training Hyperparameters

The improved YOLOv5s model is built on Win10 with Python 3.8, Pytorch 1.9.0 and Cuda 11.1. The hardware device GPU is RTX3060, and the CPU is Core i7, with 64 GB RAM. The program code is written in the Python language, using CUDA, Cudnn, OpenCV and other required libraries to realize the training and testing of the field target recognition model.

In the training process, the improved YOLOv5s model is trained by setting the batch size of the model training to 4. The regularization is performed by the BN layer each time to update the weights of the model. The momentum factor (momentum) is set to 0.937 and the weight decay rate (decay) is set to 0.0005. Both the initial vector and IOU (joint intersection)

thresholds are set to 0.01. The number of training steps is set to 1000. After training, the obtained weight files of the recognition model are saved and the performance of the model is evaluated using a test set.

2.4. Navigation Line Fitting Method

After obtaining the target detection results, the planting method is determined using the detection labels. For the direct seeding hole sowing method and seedling transplanting planting method, extracting the location information of seed pits and seedlings' target detection frames, the center of the detection frame is used as the feature point, and the feature points within 50 pixels of the center of the image are grouped; if the number of feature points is greater than 4, the planting row is judged to be odd, and the center line of the planting row is fitted with least squares for the grouped feature points as the navigation line; in other cases, the planting row is judged to be even, and the feature points are fitted. The difference between the feature points and the horizontal coordinates of the poles is calculated, the center line of the planting rows in each group is fitted by the least squares method according to the positive and negative of the difference, and the angle bisector of the two is extracted as the navigation line. For the direct seeding method, if there are odd planting rows in the image, i.e., there is a field ridge in the center of the image, the border line on the left and right sides of the detection frame is used as the characteristic straight line, and the angle bisector of the two is the navigation line; if there are even planting rows in the image, i.e., the field ridges are distributed on both sides of the image, the diagonal line of the detection frame on the left and right sides of the calculated center is used as the characteristic straight line, and the angle bisector of the two is the navigation line. The actual target detection results of the improved YOLOv5s model and the navigation line extraction results are shown in Figure 9.

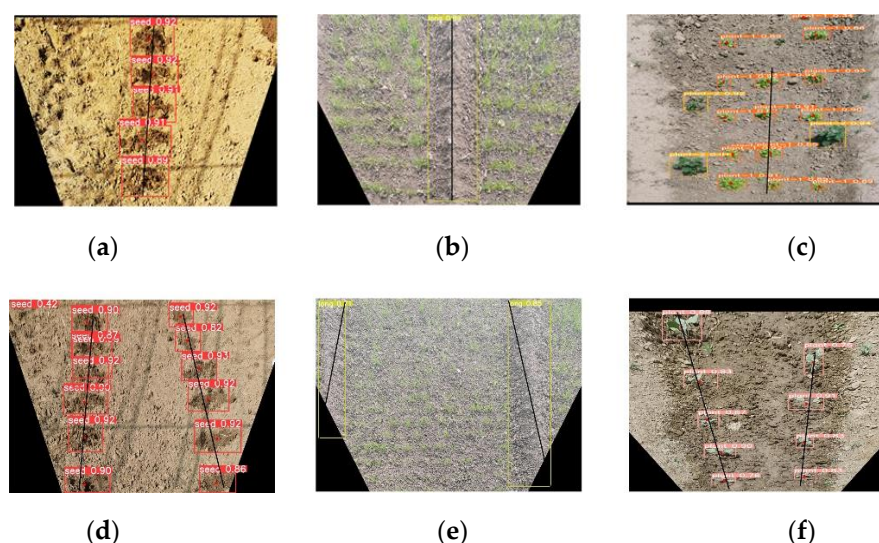


Figure 9. The actual detection effect of YOLOv5s model and the results of navigation line extraction: (a) the results of direct seeding hole sowing method with odd planting rows; (b) the results of direct seeding strip sowing method with odd planting rows; (c) the results of seedling transplanting method with odd planting rows; (d) the results of direct seeding hole sowing method with even planting rows; (e) the results of direct seeding strip sowing method with even planting rows; (f) the results of seedling transplanting method with even planting rows.

3. Results

3.1. Improved YOLOv5s Model Training Results

In the study, the detection performance of the improved YOLOv5s model was evaluated using the mean average precision (mAP) and average detection time. The training results are shown in Figure 10. The detection frame loss value (Box_loss) was 0.02, and the

classified loss value (cls_loss) was 0.001, which indicated that the model could correctly frame the detection targets and also could accurately classify seedlings and weeds in the images. The detection accuracy of seedlings and weeds was 99% and 92%, respectively. For the seed pit target, its detection accuracy was 96%, and the system could accurately detect the small area and irregularly shaped seed pit target from a long distance. The detection accuracy of the field ridge target was only 94%, but in the actual target detection process, only the main location information of field ridges in the center or on both sides of the image is needed. The field ridges, which could not be detected, are at the edge of the image. This does not adversely affect the accuracy of the navigation line extraction. The number of images detected per second was 55, which improved the accuracy of target detection while ensuring the detection speed.

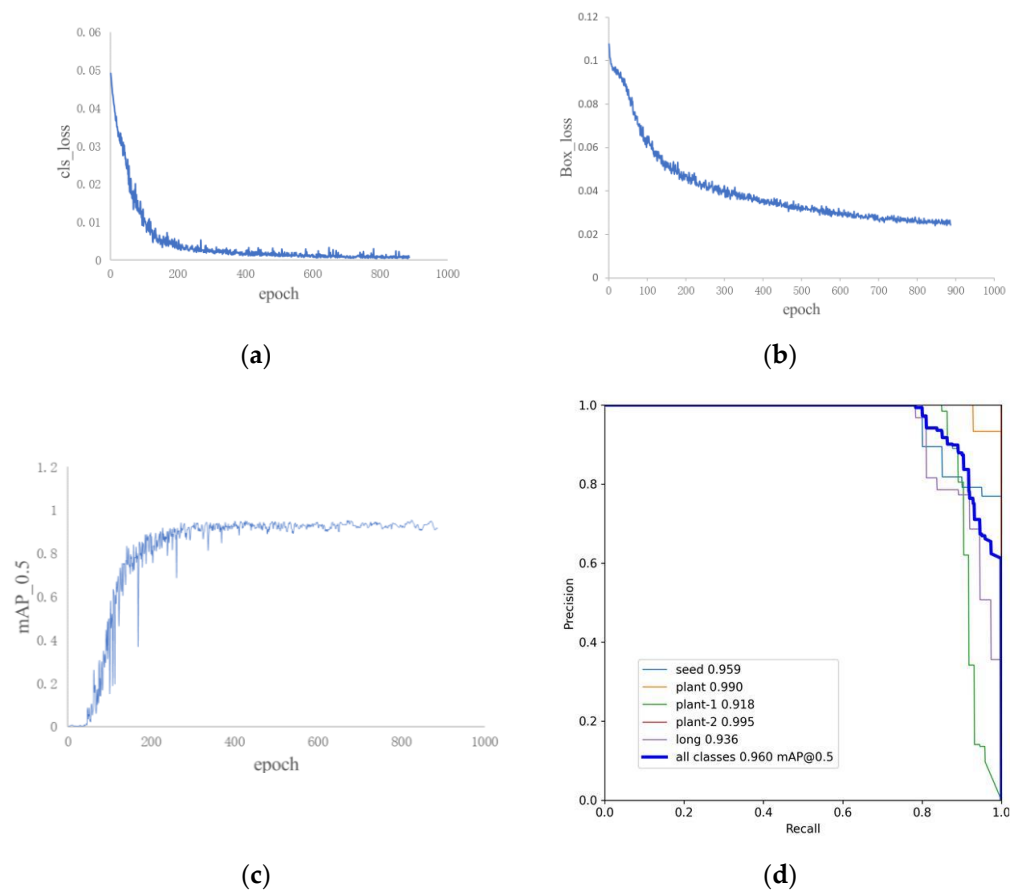


Figure 10. Training results: (a) evaluation of loss function curves for classification (cls_loss curve); (b) evaluation of loss function curves for detection frame positioning (Box_loss curve); (c) mean average precision at IoU value of 0.5 (mAP_0.5 curve); (d) precision–recall curves of each detection target.

3.2. Ablation Experiments

To verify the effectiveness of using different improvement methods, ablation experiments were performed on the improved YOLOv5s model, including the CA mechanism, Ghost module and image pre-processing. During training, the initial hyperparameters of each model were consistent, and the results are shown in Table 1. The detection accuracy of the YOLOv5s model with the CA module increased by 4.7 compared with the unimproved YOLOv5s model, but the detection time increased and the effect of invalid features was effectively suppressed. The detection accuracy of the YOLOv5s model with the inverse perspective transformation processing dataset was improved by 1.1% and the detection speed was reduced by 4 ms. It effectively reduces the influence of non-target regions in the detection frame; in addition, the detection speed is improved by 6 ms after adding the Ghost model, although the detection accuracy is reduced by 1.8%. The final results show that the

improved YOLOv5s has improved the detection accuracy by 5% and detection speed by 3 ms; thus, it can reduce the detection speed while maintaining the detection accuracy.

Table 1. Graph of ablation experiment results.

CA Module	Ghost Module	Inverse Perspective Transformation	mAP/%	Average Detection Time/ms
			91.2	21
✓			95.9	25
	✓		89.4	15
		✓	92.3	17
✓	✓	✓	96.2	18

3.3. Comparison Experiments with Different Networks

To further analyze the recognition performance of the proposed algorithm for field targets, we compared the recognition results using different target detection algorithms. The improved YOLOv5s model was compared with the unimproved YOLOv5s, YOLOX-s, YOLOv3, YOLOv7 and Faster-RCNN. The mAP value and the average recognition speed of the models were used as evaluation metrics, which are shown in Table 2. The improved YOLOv5s recognition model proposed in the study possessed the highest mAP value, which was 5% higher than that of the unimproved YOLOv5s network, and 13.5%, 8.3%, 7.6% and 1.6% higher than the values for the Faster-RCNN, YOLOv3, YOLOX-s and YOLOv7 networks, respectively. This indicated that the proposed algorithm was the most suitable for field target recognition among the five methods. Regarding the recognition speed of the model, the average detection speed of the improved YOLOv5s model was 18 ms, which was 2.4, 2, 1.8, 1.4 and 1.2 times higher than the efficiency of the Faster-RCNN, YOLOv3, YOLOX-s, YOLOv7 and unimproved YOLOv5s networks, respectively, indicating that the model could meet the detection requirements of field targets.

Table 2. Object detection network test results.

Object Detection Network	Map/%	Average Detection Time/ms
Faster-RCNN	82.7	43
YOLOv5s	91.2	21
YOLOv3	87.9	36
YOLOX-s	88.6	32
YOLOv7	94.6	26
Improved YOLOv5s	96.2	18

3.4. Test Results of Navigation Line Extraction Method

The images in the test set were selected as the object of the navigation line extraction test. The heading angle parameters of the navigation line were obtained and compared with the manually observed heading angle, which was greater than 2° and was judged to be an extraction error. The test results of navigation line extraction method are shown in Table 3. After comparing each image, the images could be divided into three categories: (1) if the key target (the center of the image and the targets near the center on both sides) in the image was detected accurately, then the navigation line could be extracted accurately; (2) if the key target in the image was missed, there was little effect on the navigation line extraction when it occurred in the image of the direct seeding hole sowing method and seedling transplanting planting method; navigation lines could be extracted incorrectly or even impossible to extract when it occurred in the image of the direct seeding strip planting method; (3) if the key target in the image was misidentified, the image could not be determined as the planting method and the navigation line could not be extracted. The average processing time was 51 ms, and no weed and seedling misclassification occurred

in the experiment, which effectively proved the accuracy of the navigation line extraction of field ridge targets.

Table 3. Navigation line extraction test results.

Image Type	Accuracy/%	Average Processing Time/ms
direct seeding strip sowing method	96	86
direct seeding hole sowing method	97	47
seedling transplanting method without weeds	99	31
seedling transplanting method with weeds	99	39

4. Conclusions

In the study, a navigation line extraction method based on the improved YOLOv5s model was proposed to achieve the fast and accurate detection of field targets and improve the accuracy of navigation line extraction. The Ghost module was used in the improved YOLOv5s model to optimize the network architecture, which reduced the computation of the feature map extraction step and improved the real-time target detection; the CA module was introduced in the backbone network to improve the target detection accuracy while reducing the occurrence of missed and false detections. In addition, the study used the inverse perspective transformation to pre-process the acquired images, in order to reduce the invalid areas in the labeled frame. The ablation test results showed that the proposed improved YOLOv5s model possessed a slightly lower average detection speed compared with the unimproved YOLOv5s model, but the detection accuracy was improved by 9.7 percentage points. To verify the detection effectiveness of the improved YOLOv5s model field target detection model, the study compared it with Faster-RCNN, YOLOv3, YOLOv7 and YOLOX-s. The results showed that the improved YOLOv5s model possessed the best detection accuracy and detection speed. The morphology-based navigation line fitting method was also proposed to achieve navigation line fitting under different planting methods for different types of field targets by extracting feature points based on the morphological features of the targets, with average accuracy of 98% and an average processing time of 51 ms. However, due to the large similarity between ridge target features and the background, it is still possible to improve the accuracy of target detection. It is planned to reduce the impact of the field image background on image processing in the future, and further improve the accuracy of navigation line extraction.

The proposed method was applicable to the detection of the navigation path of the cuttage and film covering multi-functional machine under different planting methods. Its accuracy rate met the requirements of operation, and it will provide a theoretical basis for realizing the automatic driving of visual navigation for cuttage and film covering multi-functional machines.

Author Contributions: Software, writing—original draft preparation, validation, Y.L.; validation, S.L. and Y.Z.; conceptualization, validation, writing—review and editing, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shandong Provincial Key Research and Development Program (Major Science and Technology Innovation Project)—Boost Plan for Rural Vitalization Science and Technology Innovation (No.2021TZXD001).

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

References

1. Mousazadeh, H. A technical review on navigation systems of agricultural autonomous off-road vehicles. *J. Terramech.* **2013**, *50*, 211–232. [[CrossRef](#)]
2. Han, S.; He, Y.; Fang, H. Recent development in automatic guidance and autonomous vehicle for agriculture: A Review. *J. Zhejiang Univ. Agric. Life Sci. Ed.* **2018**, *44*, 11. [[CrossRef](#)]
3. Ji, C.; Zhou, J. Current Situation of Navigation Technologies for Agricultural Machinery. *J. Agric. Mach.* **2014**, *45*, 44–54. [[CrossRef](#)]

4. Lu, W.; Zeng, M.; Wang, L.; Luo, H.; Mukherjee, S.; Huang, X.; Deng, Y. Navigation algorithm based on the boundary line of tillage soil combined with guided filtering and improved anti-noise morphology. *Sensors* **2019**, *19*, 3918. [[CrossRef](#)] [[PubMed](#)]
5. Yang, W.; Li, T.; Jia, H. Simulation and experiment of machine vision guidance of agriculture vehicles. *J. Agric. Eng.* **2004**, *1*, 160–165. [[CrossRef](#)]
6. Liu, Y.; Gao, G. Research Development of Vision-Based Guidance Directrix Recognition for Agriculture Vehicles. *Agric. Mech. Res.* **2015**, *37*, 7–13. [[CrossRef](#)]
7. 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](#)]
8. Wang, T.; Bin, C.; Zhang, Z.; Li, H.; Zhang, M. Applications of machine vision in agricultural robot navigation: A review. *Comput. Electron. Agric.* **2022**, *198*, 107085. [[CrossRef](#)]
9. Neeru, S.; Redhu, Z.; Zoozeal, T.; Shikha, Y.; Poonam, M. Chapter 37-Artificial intelligence: A way forward for agricultural sciences. In *Bioinformatics in Agriculture*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 641–668. [[CrossRef](#)]
10. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [[CrossRef](#)]
11. Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767. [[CrossRef](#)]
12. Bochkovskiy, A.; Wang, C.; Liao, H. YOLOv4: Opstimal speed and accuracy of object detection. *arXiv* **2004**, arXiv:2004.10934. [[CrossRef](#)]
13. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
14. He, K.; Gkioxari, G.; Dollr, R.; Girshick, R. Mask R-CNN. International Conference on Computer Vision. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2961–2969. [[CrossRef](#)]
15. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Volume 9351. [[CrossRef](#)]
16. Peng, M.; Xia, J.; Peng, F. Efficient recognition of cotton and weed in field based on Faster R-CNN by integrating FPN. *J. Agric. Eng.* **2019**, *35*, 202–209. [[CrossRef](#)]
17. Zhang, S.; Xu, X.; Qi, G.; Shao, Y. Detecting the pest disease of field crops using deformable VGG-16 model. *J. Agric. Eng.* **2021**, *37*, 188–194. [[CrossRef](#)]
18. Long, J.; Zhao, C.; Lin, S.; Guo, W.; Wen, C.; Zhang, Y. Segmentation method of the tomato fruits with different maturities under greenhouse environment based on improved Mask R-CNN. *J. Agric. Eng.* **2021**, *37*, 100–108. [[CrossRef](#)]
19. Gao, Z. *Method for Kiwi Trunk Detection and Navigation Line Fitting Based on Deep Learning*; Northwest Agriculture and Forestry University of Science and Technology: Xianyang, China, 2020. [[CrossRef](#)]
20. Bell, J.; MacDonald, B.A.; Ahn, H.S. Row following in pergola structured orchards. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Korea, 9–14 October 2016; pp. 640–645. [[CrossRef](#)]
21. André, S.; Filipe, B.; Luís, C.; Vitor, M.; Armando, J. Vineyard trunk detection using deep learning—An experimental device benchmark. *Comput. Electron. Agric.* **2020**, *175*, 105535. [[CrossRef](#)]
22. 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*, 106683. [[CrossRef](#)]
23. Lac, L.; Da, C.J.; Donias, M.; Keresztes, B.; Bardet, A. Crop stem detection and tracking for precision hoeing using deep learning. *Comput. Electron. Agric.* **2022**, *192*, 106606. [[CrossRef](#)]