Weed Detection in Images of Carrot Fields Based on Improved YOLO v4

Boyu Ying, Yuancheng Xu, Shuai Zhang, Yinggang Shi*, Li Liu

College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, China

Corresponding Author Email: syg9696@nwafu.edu.cn

ABSTRACT

The accurate weed detection is the premise for precision prevention and control of weeds in fields. Machine vision offers an effective means to detect weeds accurately. For precision detection of various weeds in carrot fields, this paper improves You Only Look Once v4 (YOLO v4) into a lightweight weed detection model called YOLO v4-weeds for the weeds among carrot seedlings. Specifically, the backbone network of the original YOLOv4 was replaced with MobileNetV3-Small. Combined with depth-wise separable convolution and inverted residual structure, a lightweight attention mechanism was introduced to reduce the memory required to process images, making the detection model more efficient. The research results provide a reference for the weed detection, robot weeding, and selective spraying.

1. INTRODUCTION

Carrots are one of the top ten vegetables around the world. The sowing period of carrots lasts from late summer to early autumn in China. The high temperature and abundant rainfall in this period are favorable for weed growth. The weeds impede the normal growth of carrots, making carrots susceptible to diseases and pests. One of the effective ways to increase the yield and improve the quality of carrots is to remove the weeds timely.

China is the largest carrot grower in the world. There is a total of 5 million mu of carrot fields in the country, about 40% of the global planting area of carrots. Each year, China produces more than 14 million tons of carrots, accounting for 33% of the total output of world carrots [1]. There is a huge need for automated weeding equipment. The precision detection of weeds in the field is the premise of adaptive herbicide application and mechanical weeding. However, the similar color and shape between carrot seedlings and field weeds poses a challenge to the detection of weeds in carrot fields.

For the precision detection and recognition of weeds in fields, many Chinese and foreign scholars have resorted to technical means like machine vision, and spectral detection, and devoted much energy into the relevant research [2-4].

With the recent development of deep learning (DL), convolutional neural network (CNN) has been effectively applied in machine vision [5-6]. Major breakthroughs have been achieved in CNN applications to image recognition [7-8], semantic segmentation [9-10] and object detection [11-12]. Thanks to its strong representation ability of image features, the CNN has been increasingly applied to agriculture.

In the field of crop and weed detection, Sun et al. [13] proposed a crop detection method based on the Faster R-CNN, and improved the crop detection accuracy by replacing the original VGG16 with ResNet101. Sun et al. [14] improved the AlexNet into a multiscale feature fusion convolution model, which uses a wider network structure and integrates dilated convolution with global pooling, so as to realize the accurate recognition of different crop seedlings and weeds. Taking corn seedlings and weeds as the targets, Wang et al. [15] constructed a CNN model that accurately detects targets by combining multiscale features with superpixel segmentation.

All the above studies extract features of the targets with a multilayer deep CNN. The precision detection of the targets was achieved by increasing the depth and width of the network. This strategy complicates the model and drags the detection speed, raising a stricter requirement on the configuration of the embedded devices. In real-world environment, the weeding equipment are not very sophisticated. It is a crucial issue to realize real-time, accurate, resource-saving detection in the limited hardware resources.

The fields of carrot seedlings have a complex environment: the seedlings are mixed with weeds; the two types of targets are similar in size, shape, and color; the weeds are generally small. These features add to the difficulty of precision detection of weeds based on images.

You Only Look Once is a classic object detection algorithm. It is known for its fast detection speed and high detection accuracy [16]. As a new version of YOLO, YOLOv4 integrates various optimization strategies, and improves the ability for small object detection [17], providing a suitable tool for precision weed detection. However, CSPDarknet53, the backbone network of YOLOv4, has a high complexity, and a heavy computing load in image processing, requiring a huge storage space. Therefore, YOLOv4 is not quite suitable for real-time detection in embedded devices.

To solve the above problem, this paper improves YOLOv4 into YOLOv4-weeds for weed detection. The backbone network of YOLOv4 was replaced with a lightweight neural network called MobileNetV3-Small. The improvement can effectively reduce the memory required for image processing, and improve the small weed detection efficiency and accuracy in complex environment, making the model more suitable for deployment in embedded devices.

2. WEED DETECTION MODEL YOLOV4-WEEDS

MobileNetv3-Small is a lightweight neural network [18]. This paper replaces CSPDarkNet-53 with MobileNetV3-Small as the backbone network of YOLOv4. The replacement can effectively reduce the memory required for image processing, and speed up the detection of small targets like weeds. MobileNetV3-Small integrates the depth-wise separable convolution of MobileNetV1 [19] and the inverted residual structure with linear bottleneck of MobileNetV2 [20], and introduces a lightweight attention mechanism, thereby reducing the computing load of feature map and accelerating the propagation speed of feature map in the network. The model of the attention mechanism is shown in Figure 1. In the middle of the model lies a squeeze-and-excitation network (SENet) [21], which consists of three parts: the squeeze function, the excitation function, and the scale function. Among them, the squeeze function adds up all the feature value in the channels and takes the average through global average pooling, such that the lower layers of the network can utilize global information; the excitation function obtains the coefficient of each channel, which falls between 0 and 1, by the sigmoid function, and adjusts the coefficient size through training; the scale function multiplies the value on each channel with the corresponding weight, aiming to enhance the attention to key channels and reduce the parameter quantity and calculation load.
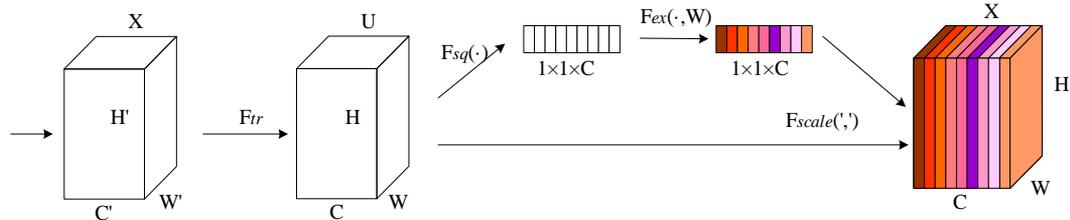

**Figure 1.** Model of the attention mechanism
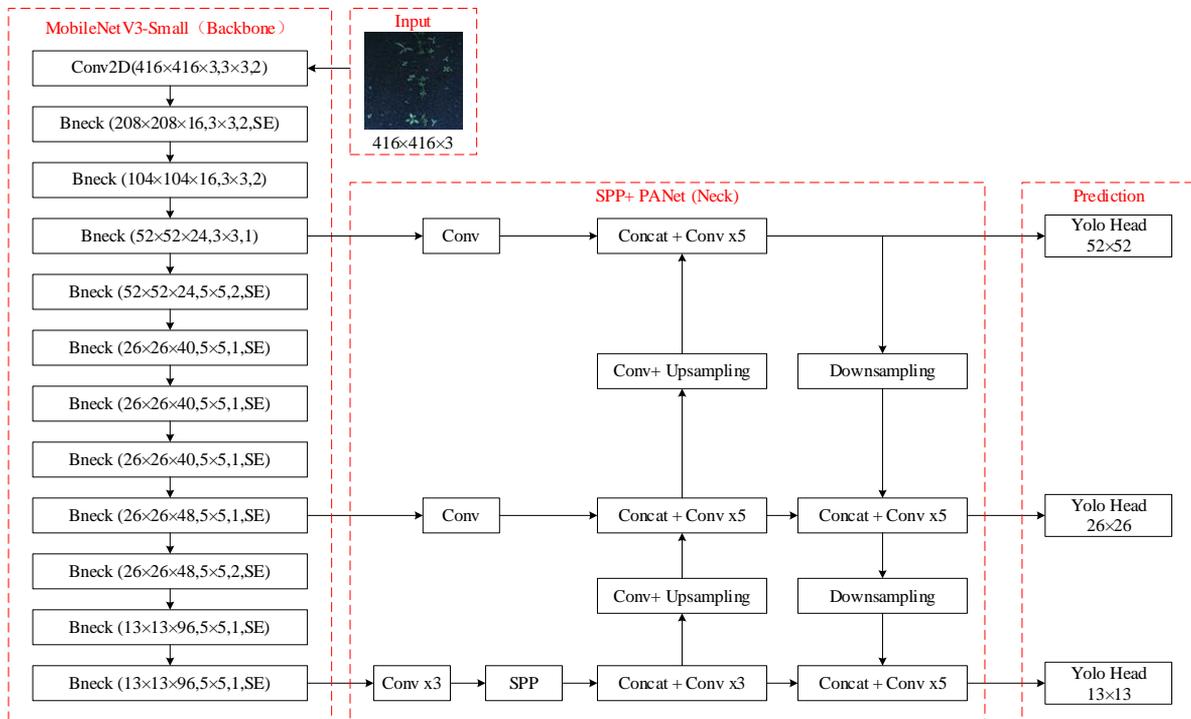
Figure 2. Structure of improved lightweight weed detection model

The structure of the improved lightweight weed detection model YOLOv4-weeds is presented in Figure 2, where the bracketed numbers represent the image resolution, number of channels, convolution kernels, stride, and the presence/absence of the attention mechanism. The backbone module receives a carrot field image of the size 416×416. Then, three effective feature layers are selected from the backbone module, and imported to the neck module. Next, feature extraction is enhanced by the Neck module. After that, the Prediction module obtains three feature maps, including the 52×52 feature map for small object detection, the 26×26 feature map for medium object detection, and the 13×13 feature map for object target detection. Finally, prediction is carried out on the input image, and an image is outputted with labels of different kinds of weeds and carrot seedlings.

In this paper, performance of each detection model is evaluated by four indices: average precision (AP), mean AP (mAP), detection time, and model weight. Let P be precision, R be recall, and AP be the mean precision for the detection of positive samples in each class. Then, the AP value can be calculated by $AP = \int_0^1 P(R)\mathrm{d}R$. mAP is the mean AP for all classes.

3. DATA ACQUISITION AND PREPROCESSING

3.1 Image acquisition

From July to August 2020, the test images were collected from the carrot fields in Fengqiu County, Xinxiang Prefecture, Central China's Henan Province (N: 34°53′; E: 114°14′; elevation: 74m). The carrots are of the variety Yuhong 2#. The carrot seedlings are cultivated by ridge planting, with a ridge spacing of 50-60cm and a plant interval of 10-15cm. Most carrot seedlings are surrounded by weeds.

This paper mainly aims to detect four common field weeds in the carrot fields, namely, crabgrass, plantain, pale persicaria, and cephalanoplos. The features of the four weeds are presented in Figure 3. Crabgrass has slender and long blades. In fact, the blades of crabgrass are narrower than those of any of the other three weeds. Plantain has the widest blades among the four weeds; the egg-shaped blades are very smooth. Pale persicaria has pointed blades with dark purple patches, which are narrower and longer than those of plantain. The blades of cephalanoplos have serrated edges.

The test images were collected with a Canon EOS 5D Mark III single-lens reflex (SLR) camera. The lens is of the model Canon EF 24 ~ 70 mm f/ 2.8L II USM, aperture of F18, and sensitivity of ISO-12800. Each image contains one or several of the four weeds. To facilitate model training, the image resolution was set uniformly to 416 pixels × 416 pixels, and all images were saved in the JPG format.


(1) Crabgrass   (2) Plantain

(3) Pale persicaria   (4) Cephalanoplos
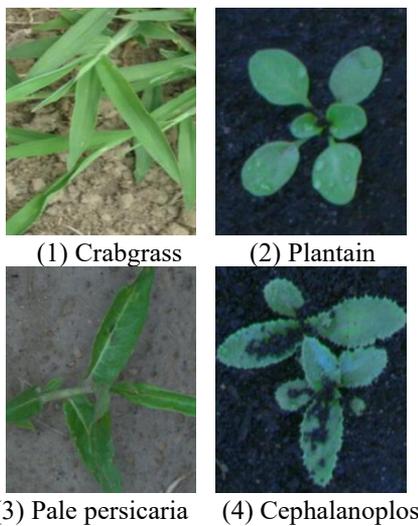
Figure 3. Common weeds in carrot fields

The field images of Yuhong 2# carrot seedlings were collected under three weather conditions: sunny, cloudy, and after rain. To guarantee the image quality, the lens was kept at 40±2cm from the carrot, at an angle of approximately 90° from the ground. In this way, the collected images contain both carrot seedlings and various weeds.

Considering the influence of light intensity on imaging quality, the sunny image collection was carried out in three periods: 100 field images on carrot seedlings during 8:00~10:00, 50 during 14:00~15:00, and 100 during 17:00~18:00; the cloudy image collection was carried out in one period only: 250 images field images on carrot seedlings during 14:00~15:00; the after-rain image collection was carried out in one period: 250 field images on carrot seedlings during the 30min after the rain. In total, 750 original field images on carrot seedlings were collected.

3.2 Image database

Figure 4 show the features of the 750 images collected from the carrot fields with weeds. The features of the images obtained under different weather conditions are presented separately. The images can be divided into two types:

(1) The images with only carrot seedlings, without any weeds;

(2) The images containing both carrot seedlings and one or several weeds.

The sunny images are well lighted, bright, and sharp in color, with a grayish yellow background. The cloudy images are not bright, and dull in color, with a gray background. Due to the wetness of soil, the images taken after rain have a black background, and a high contrast between foreground and background; the carrot seedlings can be easily told apart from the weeds in the images.


(1) Sunny   (2) Cloudy

(3) After rain
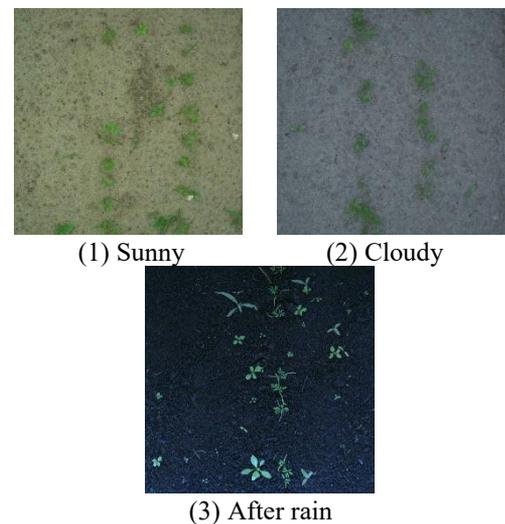
Figure 4. Images collected under different weather conditions

To improve the training efficiency and generalizability of our weed detection model, the 750 images were subject to data augmentation [22] on image processing software. The specific operations include increasing the brightness by 20% or 40%; reducing the brightness by 20% or 40%; increasing the height by +20% or +40%; increasing the contrast by 20% and 40%; reducing the contrast by 20% or 40%; sharpening to the level of 1, 3, 4, or a1. The data augmentation generated 10,500 new images. Table 1 reports the results of data augmentation, where Br, Co, Sp, and Hb represent brightness, contrast, sharpening, and high brightness, respectively.

The new images were combined with the original images into an image database of 11,250 images. From that database, 2,250 images collected under each of the three weather conditions, i.e., a total of 6,750 images (60%), were randomly extracted, and combined into the training set. Among the remaining images, 750 images collected under each of the three weather conditions, i.e., a total of 2,250 images, were randomly extracted, and combined into the validation set. The rest 2,250 images were combined into the test set. The training set, validation set, and test set were adopted to train the weed detection model, to optimize the model parameters, and to evaluate model performance. Table 2 shows how the images in the database are allocated to each set.

Table 1. Results of data augmentation

Parameters	Original	Parameter change/each								Subtotal/each
		+20%	-20%	+40%	-40%	sp01	sp03	sp04	spa1	
Br	100%	750	750	750	750	\	\	\	\	3,000
Co	100%	750	750	750	750	\	\	\	\	3,000
Sp	\	\	\	\	\	750	750	750	750	3,000
Hb	100%	750	\	750	\	\	\	\	\	1,500
Total					10,500					

Table 2. Image allocation

Collection condition	Number of original images/each	Number of images produced by data augmentation/each	Subtotal/each	Training set/each	Validation set/each	Test set/each
Sunny	250	3,500	3,750	2,250	750	750
Cloudy	250	3,500	3,750	2,250	750	750
After-rain	250	3,500	3,750	2,250	750	750
Total	750	10,500	11,250	6,750	2,250	2,250

4. TESTS AND RESULTS ANALYSIS

4.1 Dataset labeling and model training

To improve the weed detection model into a high-performance lightweight model, the premise lies in the precision labeling of test images, and effective training of the model. Taking the carrot seedlings and various weeds in the carrot fields as the targets, the authors labeled the targets on each image with minimum bounding boxes (MBBs). Every MBB should only contain one carrot seedling or one weed, and have as few background pixels as possible. Figure 5 gives an example of image labeling.

The test images were imported to LabelImg in turn. Then, the carrot seedlings and various weeds in the image were labeled as follows: carrot for every carrot seedling, plantain for every plantain, polygonum for every pale persicaria, and cirsium for every cephalanoplos. During the labeling, the software automatically generates an Extensible Markup Language (XML) file, which includes image path, weed type, and coordinates of the area of carrot or weed(s).

The targets with unobvious features or incomplete contours were also labeled with MBBs, aiming to ensure the training reliability, and prevent accidental factors from affecting the detection performance.
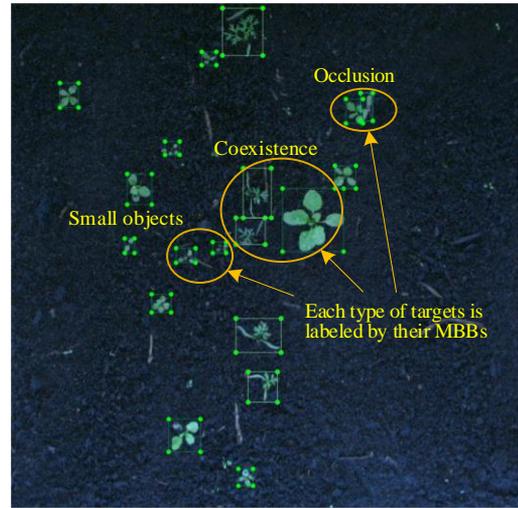

Figure 5. An example of image labeling

Table 3 shows the parameters of the training platform for the improved lightweight weed detection model YOLOv4-weeds. During the model training and testing, the software environment is Ubuntu16.04 and python3.6.12.

Table 3. Parameters of the training platform

Processor	Basic frequency	Memory	Graphics card	Video memory	Hard disk
Intel i7 10700KF	3.8 GHz	64 GB	ZOTAC RTX3090(GPU)	24 GB GDDR6X	4 TB 7200RPM SATA

During model training, the convergence speed of the loss function depends on the learning rate. Only if the learning rate is set properly, could the loss function converge to the global minimum as much as possible. This paper sets the learning rate to 0.001, 0.0001, and 0.00001, in turn. Figure 6 records the train loss and validation loss curves during the iterative model training under each of the three learning rates. It can be seen that, at the learning rate of 0.001, the train loss and validation loss of the model declined the fastest and stabilized the earliest, but the final train loss and validation loss were relatively high. At the learning rate of 0.0001, the train loss and validation loss declined quickly, and converged to the lowest values; in this case, the model achieved the best training effect. At the training rate of 0.00001, the train loss and validation loss declined the slowest and stabilized the latest; the final train

loss and validation loss were relatively high. The loss function values of the model converged well under all three learning rates.

The mAPs of the model under the three learning rates were compared to verify how the learning rate affects the mAP on the validation set. By analyzing the relevant data in Table 4, it can be learned that the model correctly recognized 88.75% of the images in the validation set at the learning rate of 0.0001, which is 5.11% and 2.55% higher than the mAPs at the learning rates of 0.001 and 0.0001, respectively. The detection performance was optimal at the learning rate of 0.0001. This agrees with the analysis result on Figure 6. Therefore, the learning rate was set to 0.0001 to ensure the optimal detection effect of YOLOv4-weeds on carrot seedlings and weeds.
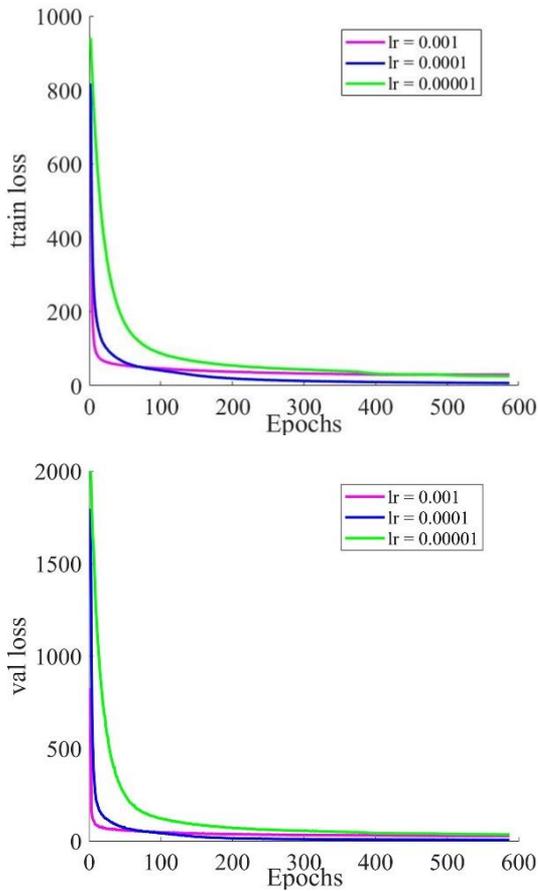
Figure 6. Loss curves of the model under three learning rates

Table 4. Model performance under three learning rates

Learning rate	mAP/%
0.001	83.64
0.0001	88.75
0.00001	86.20

4.2 Detection effect

To verify its detection effectiveness, the improved lightweight weed detection model was tested on the test set. There are two kinds of features in the images of the test set: (1) In some images, the foreground targets are only carrot seedlings, without any weed; (2) In some other images, the foreground targets include both carrot seedlings and various weeds, which are diverse in size and mixed with carrot seedlings. The above two kinds of images are denoted as Type 1 and Type 2, respectively. Then, the trained model YOLOv4-weeds were adopted to process the two kinds of images. Figure 7 provides an example of the detection results. It can be seen that the proposed model YOLOv4-weeds basically accurately detected the carrot seedlings and various weeds in the complex scene of the fields.

To demonstrate its superiority, our model YOLOv4-weeds were compared with several detection models through experiments. For the effectiveness and rigorousness of the experiments, five models, namely, YOLOv4-weeds, YOLOv4, YOLOv4-tiny, YOLOv3, and YOLO v3-tiny, were iteratively trained on the same training set. Then, the five trained models were evaluated against the same test set. Figure 8 presents an example of the detection results of the five models on the same test image.

Type 1

Type 2
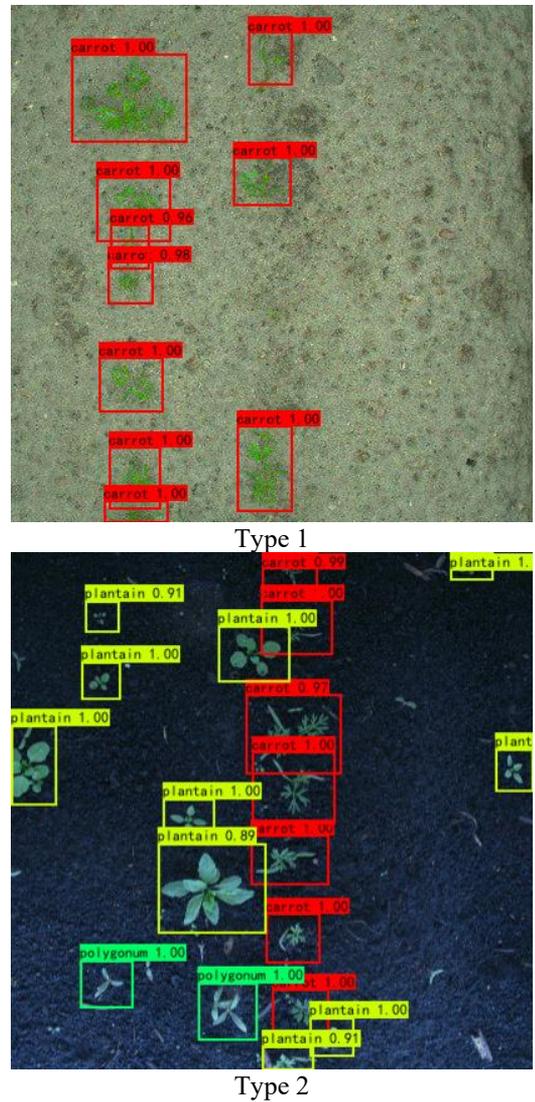
Figure 7. An example of detection results of YOLOv4-weeds

Apparently, all models achieved good detection effects on Type 1 images. The reason is that Type 1 images contain sufficient targets of the same class: carrot seedlings. After training, the models all boast strong robustness on such a balanced dataset, which has no inter-class difference of image features. YOLOv4-weeds had the highest confidence in the detection of carrot seedlings.

For Type 2 images, the proposed model YOLOv4-weeds outperformed the other four models. Our model could effectively detect the targets in the images, and maintain a high confidence. Specifically, YOLOv3 and YOLOv4 mistook the large targets of plantain as cephalanoplos; YOLOv3-tiny and YOLOv4-tiny correctly recognized these targets, but at a low confidence (<0.60). Meanwhile, these targets were correctly detected by YOLOv4-weeds at a high confidence. The poor performance of the other four models is possibly the result of the similarity between large plantain and cephalanoplos. In the training set, there are relatively few large targets of plantain. Facing such a unbalanced dataset, our model is more robust than any of the other four models. As for the coexistence between carrot seedlings and weeds, YOLOv3-tiny and YOLOv4-tiny were prone to miss the carrot seedlings, because their structures are too simple to effectively extract features from complex images.
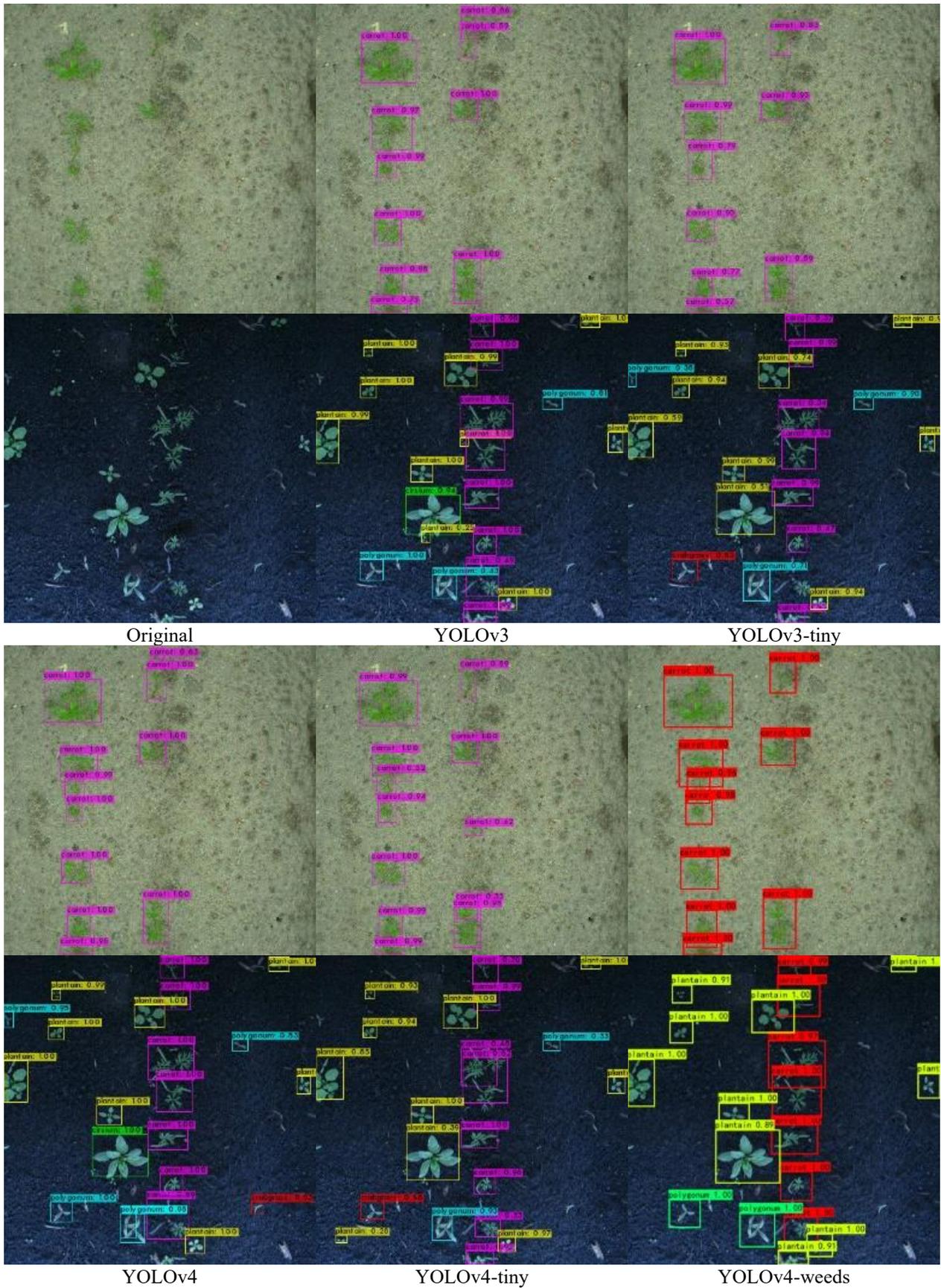
Figure 8. An example of detection results of the five models

In addition, this paper compares the mAP, detection time, mmAP(mean mean average Precision), average detection time, and weight. As shown in Table 5, the mmAPs of YOLOv3, YOLOv3-tiny, YOLOv4, YOLOv4-tiny, and YOLOv4-weeds were 88.91%, 82.11%, 88.95%, 82.71%, and 88.46%, respectively. It is obvious that the proposed model YOLOv4-weeds performed excellently on every metric.

On Type 1 test images, all models achieved a high AP on the foreground targets of carrot seedlings under any weather condition. This is because Type 1 images contain sufficient

targets of the same class; every model is very robust on such images. While the APs were about the same, YOLOv4-weeds consumed a shorter time than YOLOv3 and YOLOv4, suggesting that our model is better at detecting such images than YOLOv3 and YOLOv4.

On Type 2 test images, YOLOv3, YOLOv4, and YOLOv4-weeds achieved high APs, while YOLOv4-tiny and YOLOv3-tiny ended up with low APs. Perhaps the latter two models, with their simple structures, are weak in feature extraction, and poor in the detection of small weeds. By contrast, YOLOv4-weeds took a much shorter detection time than YOLOv3 and YOLOv4, while achieving a relatively high mAP on Type 2 images.

Overall, despite having a greater weight than YOLOv3-tiny and YOLOv4-tiny, YOLOv4-weeds still achieved a high detection precision, which is 6.35% and 5.75% higher than that of YOLOv3-tiny and YOLOv4-tiny, respectively. The model was almost as precise as YOLOv3 and YOLOv4.

From the angle of average detection time, YOLOv4-weeds consumed an average detection time of 12.65ms, which is shorter than the time consumed by YOLOv3 and YOLOv4. The fast detection speed indicates that YOLOv4-weeds could realize real-time detection of carrot seedlings and various seeds in the complex scene of carrot fields, without sacrificing model accuracy.

From the angle of model weight, YOLOv4-weeds had a weight of 159.0MB. Although it was larger than that of YOLOv3-tiny and YOLOv4-tiny, the weight was far smaller than that of YOLOv3 and YOLOv4.

In summary, the proposed model can effectively detect the carrot seedlings and various seeds in the complex scene of carrot fields, and achieve a relatively small weight and short mean detection time. Apart from its high robustness, our model can be easily applied to embedded equipment. It is more suitable for weed detection than the other four models.

Table 5. Detection performance of five models

Network models	mAP/%		Detection time/ms		mmAP /%	Average detection time/ms	Weight /MB
	Type 1	Type 2	Type 1	Type 2			
YOLOv3	89.65	88.16	13.33	14.66	88.91	13.99	246.4
YOLOv3-tiny	86.74	77.47	5.33	4.73	82.11	5.03	34.7
YOLOv4	89.06	88.84	18.66	18.66	88.95	18.66	256.1
YOLOv4-tiny	86.37	79.05	4.81	4.54	82.71	4.68	23.6
YOLOv4-weeds (ours)	89.11	87.80	12.38	12.92	88.46	12.65	159.0

5. CONCLUSIONS

Targeting the various weeds in the growth environment of carrot seedlings, this paper improves YOLOv4 into a lightweight weed detection model YOLOv4-weeds. Through model training and parameter optimization, the model performance was found to be optimal at the learning rate of 0.0001. Comparative experiments show that our model achieved better overall performance than YOLOv4, YOLOv4-tiny, YOLOv3, and YOLOv3-tiny, as evidenced by its mmAP of 88.46%, average detection time of 12.65 ms, and weight of 159.0 MB. This means the proposed model can effectively detect the carrot seedlings and various seeds in the complex scene of carrot field, and be easily applied to embedded equipment. However, the detection effect of our model on a few images is constrained by the fact that: there are relatively few targets of some weed types in the images of our dataset, that is, the training samples are unbalanced, with a inter-class difference of image features; as a result, the features of some targets are not sufficiently extracted during model training. This problem will be solved in future research.

REFERENCES

[1] Wang, F., Wang, G.L., Hou, X.L., Li, M.Y., Xu, Z.S., Xiong, A.S. (2018). The genome sequence of 'Kurodagosun', a major carrot variety in Japan and China, reveals insights into biological research and carrot breeding. Molecular Genetics and Genomics, 293(4): 861-871. https://doi.org/10.1007/s00438-018-1428-3

[2] Wu, X., Aravecchia, S., Lottes, P., Stachniss, C., Pradalier, C. (2020). Robotic weed control using automated weed and crop classification. Journal of Field Robotics, 37(2): 322-340. https://doi.org/10.1002/rob.21938

[3] Sabzi, S., Abbaspour-Gilandeh, Y., Arribas, J.I. (2020). An automatic visible-range video weed detection, segmentation and classification prototype in potato field. Heliyon, 6(5): e03685. https://doi.org/10.1016/j.heliyon.2020.e03685

[4] Louargant, M., Villette, S., Jones, G., Vigneau, N., Paoli, J.N., Gée, C. (2017). Weed detection by UAV: Simulation of the impact of spectral mixing in multispectral images. Precision Agriculture, 18(6): 932-951. http://doi.org/10.1007/s11119-017-9528-3

[5] Fu, L., Majeed, Y., Zhang, X., Karkee, M., Zhang, Q. (2020). Faster R–CNN–based apple detection in dense-foliage fruiting-wall trees using RGB and depth features for robotic harvesting. Biosystems Engineering, 197: 245-256. http://doi.org/10.1016/j.biosystemseng.2020.07.007

[6] Wu, D., Lv, S., Jiang, M., Song, H. (2020). Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments. Computers and Electronics in Agriculture, 178: 105742. https://doi.org/10.1016/j.compag.2020.105742

[7] Wang, J., Li, Y., Feng, H., Ren, L., Du, X., Wu, J. (2020). Common pests image recognition based on deep convolutional neural network. Computers and Electronics in Agriculture, 179: 105834. https://doi.org/10.1016/j.compag.2020.105834

[8] Jiang, J., Xu, C., Cui, Z., Zhang, T., Zheng, W., Yang, J. (2019). Walk-steered convolution for graph classification. IEEE Transactions on Neural Networks and Learning Systems, 31(11): 4553-4566. https://doi.org/10.1109/tnnls.2019.2956095

[9] Du, J., Lu, X., Fan, J., Qin, Y., Yang, X., Guo, X. (2020). Image-based high-throughput detection and phenotype evaluation method for multiple lettuce varieties. Frontiers in Plant Science, 11: 563386. https://doi.org/10.3389/fpls.2020.563386

[10] Barth, R., IJsselmuiden, J., Hemming, J., Van Henten, E. J. (2019). Synthetic bootstrapping of convolutional neural networks for semantic plant part segmentation. Computers and Electronics in Agriculture, 161: 291-304. https://doi.org/10.1016/j.compag.2017.11.040

[11] Champ, J., Mora-Fallas, A., Goëau, H., Mata-Montero, E., Bonnet, P., Joly, A. (2020). Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. Applications in Plant Sciences, 8(7): e11373. https://doi.org/10.1002/aps3.11373

[12] Han, X., Jiang, T., Zhao, Z., Lei, Z. (2020). Research on remote sensing image target recognition based on deep convolution neural network. International Journal of Pattern Recognition and Artificial Intelligence, 34(5): 2054015. https://doi.org/10.1142/S0218001420540154

[13] Sun, Z., Zhang, C., Ge, L., Zhang, M., Li, W., Tan, Y.Z. (2019). Image detection method for broccoli seedlings in field based on Faster R-CNN. Transactions of the Chinese Society for Agricultural Machinery, 50(7): 216-221. https://doi.org/10.6041/j.issn.1000-1298.2019.07.023

[14] Sun, J., He, X., Tan, W., Wu, X., Shen, J., Lu, H. (2018). Recognition of crop seedling and weed recognition based on dilated convolution and global pooling in CNN. Transactions of the Chinese Society of Agricultural Engineering, 34(11): 159-165. https://doi.org/10.11975/j.issn.1002-6819.2018.11.020

[15] Wang, C., Wu, X., Li, Z. (2018). Recognition of maize and weed based on multi-scale hierarchical features extracted by convolutional neural network. Transactions of the Chinese Society of Agricultural Engineering, 34(5): 144-151. https://doi.org/10.11975/j.issn.1002-6819.2018.05.019

[16] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788. https://doi.org/10.1109/CVPR.2016.91

[17] Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[18] Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., Adam, H. (2019). Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1314-1324.

[19] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[20] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510-4520.

[21] Hu, J., Shen, L., Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7132-7141. https://doi.org/10.1109/TPAMI.2019.2913372

[22] Ding, J., Chen, B., Liu, H., Huang, M. (2016). Convolutional neural network with data augmentation for SAR target recognition. IEEE Geoscience and Remote Sensing Letters, 13(3): 364-368. https://doi.org/10.1109/LGRS.2015.2513754