Detection method of potato leaf disease based on YOLOv5s

Jingtao Li,[1] Hao Chen,[1] Guisong Li,[1] Yueqi Liu,[1] Yanli Yang,[2] Xia Liu,[2] Chang Yi Wang[3]

[1]School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming; [2]College of Plant Protection, Yunnan Agricultural University, Kunming; [3]School of Animation and Digital Arts, Communication University of China, Beijing, China

Introduction

Potato leaf diseases are several kinds of diseases that are easy to spread and may break out at any time. Such epidemic diseases often lead to a large number of potato production, resulting in serious economic losses (Huang *et al*., 2016). The occurrence and prevalence of late blight, early blight and anthracnose are greatly affected by climatic factors such as rainfall, field humidity and temperature. Yunnan Province is one of the main potato producing areas in China, and the local climate environment is very easy to cause the outbreak and epidemic (Li *et al*., 2022) of potato diseases. Therefore, the monitoring method of low-altitude long focus shooting by UAV has become a very key means to monitor the infection of a large range of potato plants. Quickly and accurately identifying different types of potato leaf diseases is a prerequisite for preventing their spread and ensuring healthy growth. and ensuring its healthy growth.

Potato leaf disease detection belongs to the problem of target detection in deep learning. Object detection algorithms are mainly divided into two-stage and one-stage object detection algorithms: two-stage object detection algorithms such as R-CNN, Fast RCNN (Xi *et al*., 2020), *etc*. Yang *et al.* (2020) proposed a potato leaf disease detection method combining Faster R-CNN network and composite feature dictionary. However, the image selected in the dataset are all single-target blade images, and the frame rate of model detection is only 1.6 frames /s and the recognition precision is only 81%. Single-stage target detection algorithms are such as SSD (Peng *et al*., 2018) (Single shot multiBox detector), YOLO (Xiu *et al*., 2022; Zhang *et al*., 2021; Cai *et al*., 2023), *etc*. Wang *et al*. (2021) perform channel pruning on YOLOv4 model to achieve the purpose of lightweight for potato soil block detection. With the continuous development of the one-stage detection algorithm, the current one-stage object detection algorithm has more advantages in detection speed and detection accuracy than the two-stage object detection algorithm.

Fan *et al*. (2019) designed a fast recognition method of disease area relying on ORB and SIFT key feature points, and constructed a disease recognition system based on SVM combining color and texture features. However, the number of late blight samples in the system is only 54 and the sample images are all single-leaf target images, which cannot meet the sample size required for practical experiments. Dang *et al.* (2020) proposed a recognition method combining color, texture and shape based on machine vision, However, the method used only 240 images as the total sample number, resulting in the identification precision of potato late blight disease only 77%, and the detection frame rate was only 10 frames per second. Xiao and colleagues (2017) proposed an image method using adaptive feature fusion and fast recognition. The recognition precision of this model reaches 95%, but because 124 features need to be extracted and sorted before 13 detection features are selected, the learning cost is too high and there are only 250 images in the experiment. Li *et al.* (2020) used the improved Faster-RCNN to detect bitter melon leaf diseases. This method has good robustness and high precision; however, this model can only achieve a detection frame rate of 0.5 frames per second, and has a large number of parameters and a long training time. In summary, the current research on the application of deep learning methods for potato disease detection is relatively limited. Most studies have adopted a two-stage detection algorithm. Regarding the construction of experimental data, most studies have utilized an insufficient amount of experimental data, typically around 200 samples, which fails to fully ensure the authenticity of the experimental results. Additionally, the model's detection speed is slow, reaching only up to 10 frames per second. Therefore, it is crucial to explore

Correspondence: Xia Liu, College of Plant Protection, Yunnan Agricultural University, Kunming 650201, Yunnan, China.
E-mail: liuxia_0213@163.com

Key words: Potato leaf disease; targeted detection; YOLOv5s; deformable convolution.

Contributions: all authors have contributed equally, read and approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

Conflict of interest: the authors declare no potential conflict of interest.

Funding: this work has been supported by the Major Science and Technology Special Project 202102AE090018 and 202402AE090017 of the Yunnan Provincial Department of Science and Technology.(Yunnan province Major Science and Technology Special Program: 202402AE090017, project name: Potato Major Diseases Green control system research).

Acknowledgments: personal data set of this work is collected and provided by Yunnan Agricultural University.

Received: 27 October 2023.
Accepted: 30 April 2024.

Journal of Agricultural Engineering 2024; LV:1587
doi:10.4081/jae.2024.1587

approaches that involve selecting a sufficient number of sample datasets and incorporating a one-stage algorithm in order to develop a rapid and accurate method for detecting potato diseases.

In response to the above problems, taking potato late blight, early blight and anthracnose as examples, this paper proposes an improved potato leaf disease recognition and detection model based on YOLOv5s model. By introducing deformable convolution module and attention mechanism, the feature extraction ability of the algorithm is enhanced, and the deblurring network is used to reduce the influence of data input caused by factors such as shooting jitter. The accuracy of the actual detection process of the model is further improved, and the accuracy and effectiveness of potato disease recognition are guaranteed by replacing the loss function optimization model, so as to provide an accurate and real-time disease detection model for potato disease recognition.

Materials and Methods

Dataset construction

A total of 1200 original images were collected in this experimental dataset. This includes 600 clear single-blade images, 200 blurry single-blade images, and 400 multi-blade images. These images are derived from the online open data set Plant Village (750 images) and the potato seed planting Experimental Base of Yunnan Agricultural University (450 images). Of these images, late blight images accounted for 870, early blight images accounted for 50, anthracnose images accounted for 85, healthy leaves accounted for 195.The details are shown in Table 1. As shown in Figure 1a, the image pixels in the Plant Village data set are low and all are single-leaf target images. On the contrary, as shown in Figure 1b, the UAV is used to capture high-quality images by close-range focusing, and then the acquired data is enlarged and clipped to obtain leaf images containing the disease conditions of single leaf and multi-leaf. Therefore, the potato leaf disease samples obtained by these two collection methods have certain randomness and meet the experimental requirements.

Test platform

The experimental training processing platform is equipped with an NVIDIA RTX 2060 GPU, an Intel i5 11400F CPU, and 16GB RAM. The experiments were conducted using the Pytorch deep learning framework in conjunction with the Python 3.8 programming language. The potato planting environment scene is shown in Figure 2.

Sample pretreatment

In order to improve the generalization ability of the trained model and avoid the phenomenon of overfitting in the network, the obtained images are enhanced. After the data set was expanded by Moscia, Mixup, image translation and rotation, and Gaussian noise, we obtained 3,600 rich and diverse image data. The enhanced data image is shown in Figure 3.

The YOLO data set format is utilized for model training, with an 8:1:1 ratio for the training data set, validation data set, and test data set respectively. LabelImg software was employed to annotate the final training set consisting of four target types. The image annotation results are displayed in Figure 4. During training, the resolution of acquired images was adjusted to 640 × 640 pixels as input images. The following training parameters were configured: each batch consisted of 32 samples; momentum was set at 0.937; and an initial learning rate of 0.005 was utilized.

Network analysis and improvement

YOLOv5s network model structure

The YOLOv5s model is a single-stage object detection method, comprising four main components: an image receiving module (Input), a Backbone module, Feature Pyramid Networks (Lin *et al*., 2017), and a prediction module (Head). Figure 5 illustrates the architecture of the YOLOv5s model. Initially, the network dynamically adjusts the input image to a size of 640×640 pixels .Subsequently, it extracts feature information from the input image using the backbone network module and merges these features through a feature merging module. Finally, the effective feature layer is simultaneously classified and position regression in the prediction module.

Deformable convolution

As the backbone network for YOLOv5s, CSPDarknet's advantage over the previous YOLO series is that it enhances the learning ability of convolutional neural networks while maintaining the accuracy and lightweight of the model. In this study, the traditional

Table 1. Number of images of potato diseases.

Disease name	Tag name	Number of images
Late blight	Late blight	870
Early blight	Early blight	50
Anthracnose	Anthracnose	85
Health	Health	195
Total number	1200	

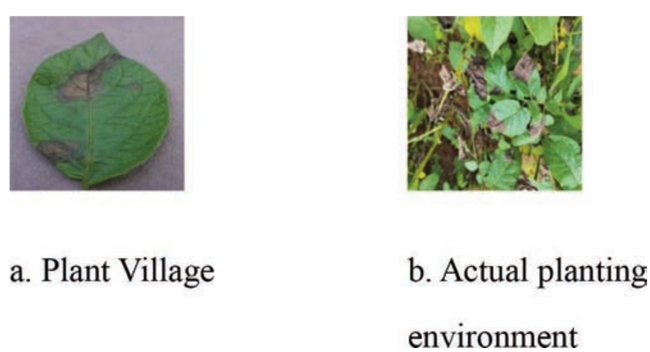

Figure 1. Dataset image.

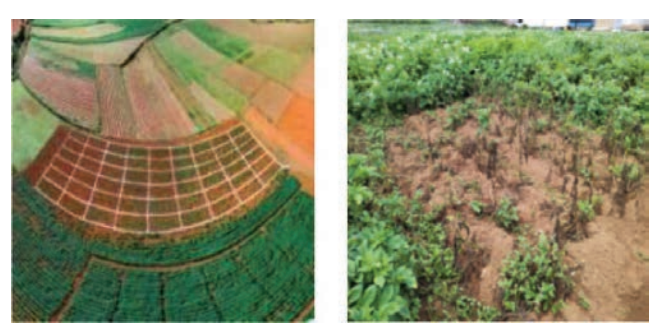

Figure 2. Potato planting environment.

convolution in CSPDarknet is replaced by deformable convolution (Dai *et al*., 2023). By introducing an offset in the receptive field, the receptive field is no longer rigid square, but closer to the actual shape of the disease spot. However, since the amount of parameters of deformable convolution is larger than that of traditional convolution, there is room to reduce the amount of parameters by model compression.

The definition of a traditional convolution structure can be formulated as follows:

$$y(p_0)=\sum_{p_k\in R} w(p_k)\cdot x(p_0+p_k) \tag{1}$$

The structures of deformable convolutions can be defined as follows:

$$y(p_0)=\sum_{g=1}^{G}\sum_{k=1}^{K} w_g m_{gk} x_g(p_0+p_k+\Delta p_{gk}) \tag{2}$$

Where: w_g, share projection weights within each group; p_k, position relative coordinates; m_{gk}, normalized modulation factor of the K sampling point in group g; Δp_{gk} learned offset of the K sampling point in group g.

As shown in Figure 6, the convolution region of the traditional convolution is a square structure (Figure 6a), but in part

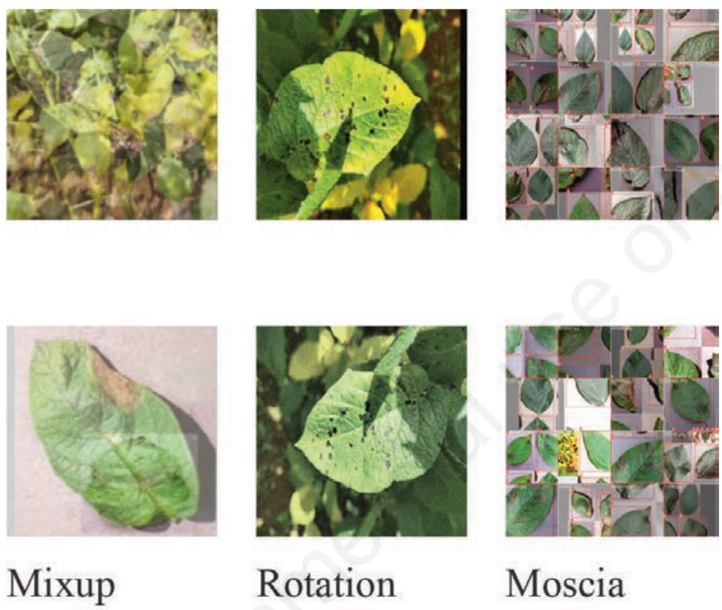

Figure 3. Data enhancement result.

Figure 4. LabelImg annotation diagram.

deformable convolution can simulate the actual shape of the object more realistically by using the self-learning offset (Figure 6b), and extract the irregular characteristic information of potato late blight, while ensuring that the subsequent convolution region always covers the shape edge of the disease spot.

CBAM module

The CBAM (Convolutional Block Attention Module) (Woo *et al.*, 2018) is introduced in this paper to enhance the disease target information and suppress background noise, thereby extracting key features of potato leaf disease from complex feature information. Figure 7 illustrates the structure of the CBAM module.

The intention of enhancing the channel's attention expression is reflected by incorporating a skip connection in this paper. The structure diagram is illustrated in Figure 8. The function expression can be formulated as follows:

$$F'' = Mc(F') \otimes F' + F' \tag{3}$$

Loss function optimization

The YOLOv5s model adopts CIoU_Loss (Zheng *et al.*, 2020) to compute the errors in center coordinates, width, and height between bounding box and ground truth. CIoU_Loss is defined as follows:

$$CIoU\_Loss = 1 - CIoU = 1 - IoU + \frac{\rho^2}{c^2} + \alpha v \tag{4}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2 \tag{5}$$

where ρ is the Euclidean distance between the centroid of bounding box and ground truth; c is the diagonal length of the outermost matrix of bounding box and ground truth; v measures the aspect ratio consistency parameter between bounding box and ground truth.

However, since the aspect ratio in CIoU_Loss describes the relative ratio, it can be seen that as long as the width and height

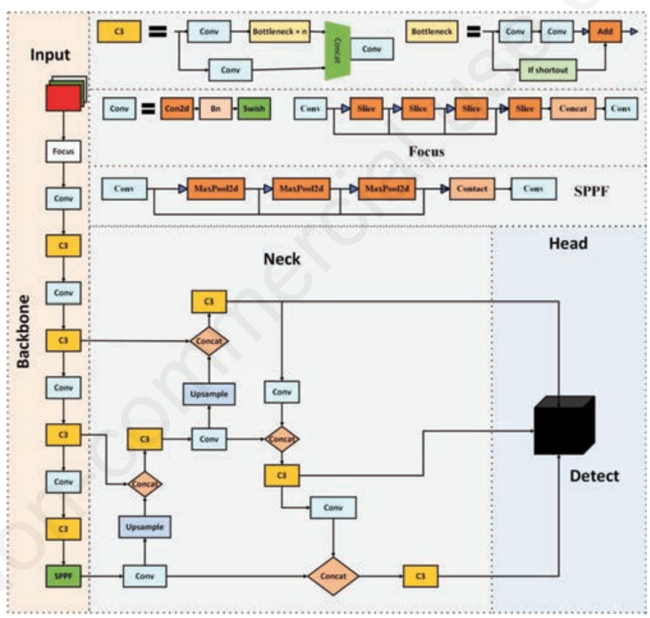

Figure 5. YOLOv5s network structure.

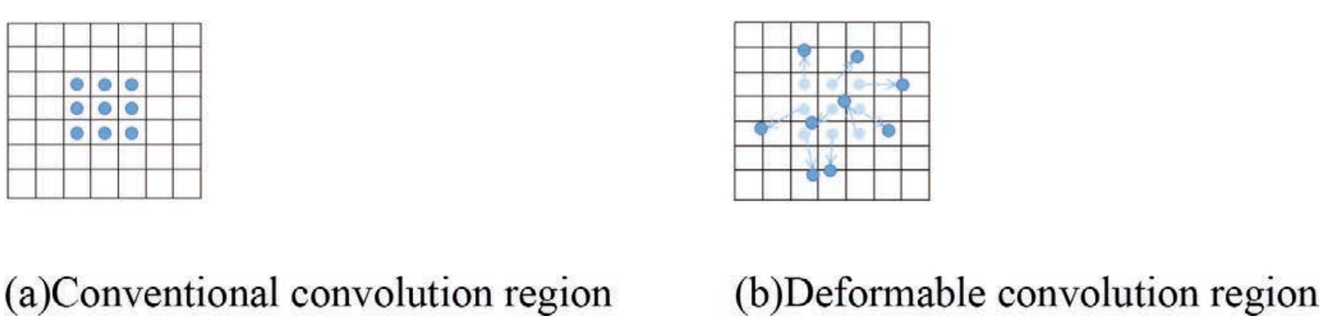

Figure 6. Standard convolution and deformable convolution region.

meet the condition that $w=kw^{gt}, h=kh^{gt}$, the penalty item added in CIoU will not work.

In this study, WIoU_Loss (Tong *et al*., 2023) is employed to replace CIoU_Loss. WIoU_Loss calculates the normalized distance between the center points of two bounding boxes and incorporates the width and height of the minimum surrounding box. The definition of WIoU_Loss is:

$$R_{WIoU} = \exp\left(\frac{(x-x_{gt})^2+(y-y_{gt})^2}{(W_g^2-H_g^2)^*}\right) \tag{6}$$

$$L_{IoU} = 1 - IoU = 1 - \frac{W_i H_i}{S_u} \tag{7}$$

$$L_{WIoU} = R_{WIoU} L_{IoU} \tag{8}$$

where Wg and Hg are width and height of the minimum surrounding box; x and y are the normalized distance between the centroids of two bounding boxes.

The utilization of WIoU_Loss as the loss function for boundary box regression significantly enhances both the speed and precision of model prediction.

SRN-DeblurNet

The SRN-DeblurNet (Tao *et al*., 2018) employs an encoder-decoder residual network architecture, consisting of an input module, two encoder modules, one ConvLSTM module, two decoder modules, and an output module. The model structure diagram is depicted in Figure 9.

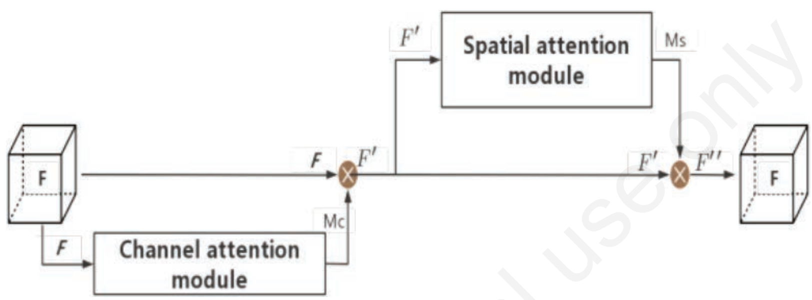

Figure 7. CBAM module.

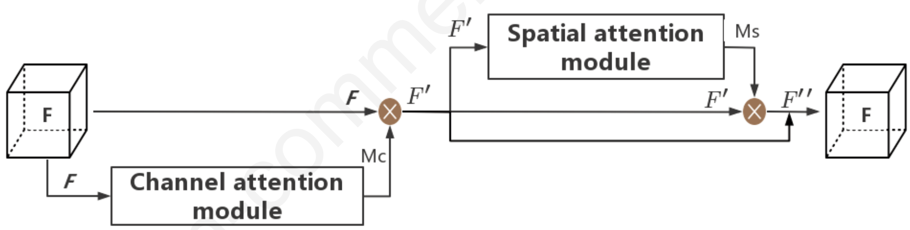

Figure 8. Add a connected CBAM module.

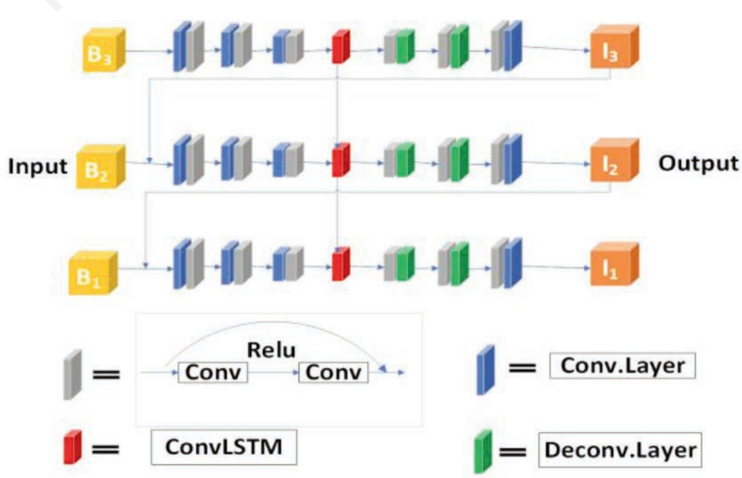

Figure 9. SRN-DeblurNet structure diagram.

The SRN-DeblurNet network has the problem of too many parameters and slow running speed. Therefore, under the condition of ensuring the accuracy of the model, the model has as few parameters, low calculation amount and fast speed as possible. The objects of parameter quantization mainly include weight, activation function output and gradient. The compressed SRN-DeblurNet network can improve the memory bandwidth and cache utilization more efficiently, and the calculation speed is usually improved by 2 to 3 times without great loss of accuracy.

During the training process, the 160×160 pixel image block B3 is randomly cropped, and the 320×320 pixel resolution subimage B2 and 640×640 pixel resolution subimage B1 are obtained by the nearest neighbor interpolation method. The output of the previous layer was used as the input of the next layer to restore the clear image from small to large, and the output1 of the last layer was used as the final restoration result.

$$L=\sum_{i=1}^{n}\frac{K_i}{N_i}\left\|I^i-I_*^i\right\|^2 \tag{9}$$

where I^i is the network output at the i-th scale; I_*^i is the truth value on the i-th scale; K_i is the weight for each scale, with initial value set to 1; N_i is the number of elements to be normalized in I^i.

Results

Experimentation

The results processed by the SRN-Deblur algorithm are compared and presented in Table 2 and Figure 10. The evaluation criteria for assessing the deblurring effect include Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM), which achieve values of 26.42 and 0.750 ,respectively. Table 2 shows that the deblurring time of SRN-Deblur network is slightly longer than that of DeblurGAN network, and the overall performance is better than DeblurGAN and deblurgan-v2 in the case of similar deblurring time. Figure 10 shows that SRN-Deblur restores the original natural information of the image to a certain extent. It can be seen from Figure 10 that SRN-deblur restores the original natural information of the image to a certain extent. In the comparison models, the deblurring results are better than DeblurGAN and Deblurgan-v2, and the model size is only 1/6 of DeblurGAN. The improved module in the dataset was subjected to ablation experiments to validate the efficacy of this paper in identifying potato leaf diseases, and the corresponding experimental results are presented in Table 3. The SRN network, CBAM module, DCN module, and WIOU loss function were consecutively incorporated. As presented in Table 3, YOLOv5s initially achieved a mean average precision (mAP) result of 0.835; however, the mAP increased to 0.84 after

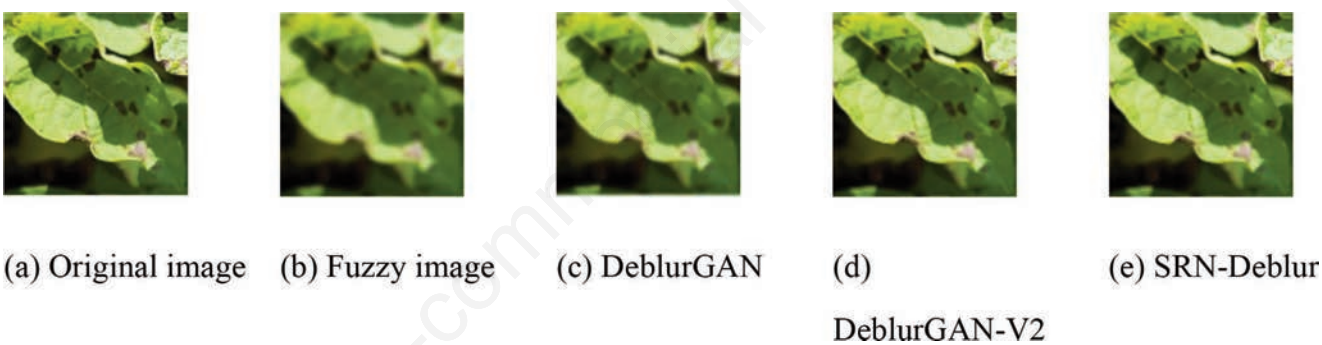

Figure 10. SRN-DeblurNet structure diagram.

Table 2. The performance comparison among various deblurring algorithms.

Model	Time/s	Model memory usage (MB)	PSNR	SSIM
DeblurGAN	0.198	44.6	21.96	0.733
DeblurGAN-V2	0.082	13.5	22.96	0.776
SRN-Deblur	0.270	8.0	26.42	0.750

Table 3. Ablation study (mAP).

Method	SRN	CBAM	WIOU	DCN	mAP
YOLOv5s					0.835
YOLOv5s	√				0.84
YOLOv5s	√	√			0.857
YOLOv5s	√		√		0.842
YOLOv5s	√			√	0.851
YOLOv5s	√	√	√	√	0.875

integrating the SRN deblurring network. Moreover, with the inclusion of three modules, the detection mAP improved to 0.857, 0.842, and 0.851, respectively, for each module - clearly indicating their positive impact on potato late blight leaf image target detection task. This validates our approach of enhancing local feature capture ability and improving the loss function as well. When all these enhanced modules are simultaneously combined, the target detection mAP of the SRN-YOLOv5s detection network on potato late blight reaches 0.875 - a significant improvement by four percentage points compared to that of the YOLOv5s model-thus demonstrating the effectiveness of our proposed enhancement method.

In order to verify the superiority of the proposed method, different models of early blight, late blight, anthracnose and healthy leaves were tested. As can be seen from Table 4, the enhancement method has a slight improvement on AP compared to other models.

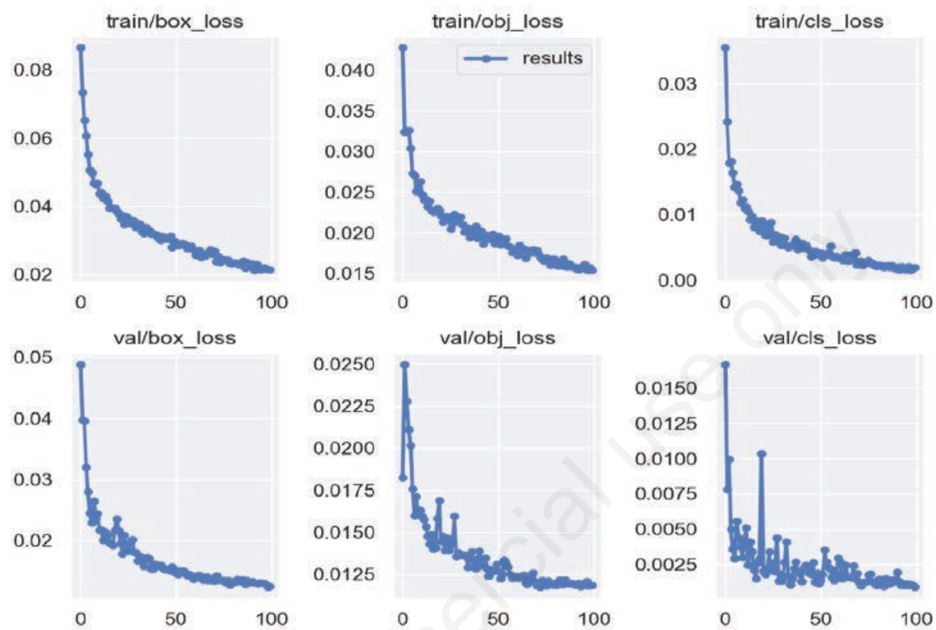

Figure 11. Loss curve of the YOLOv5s method.

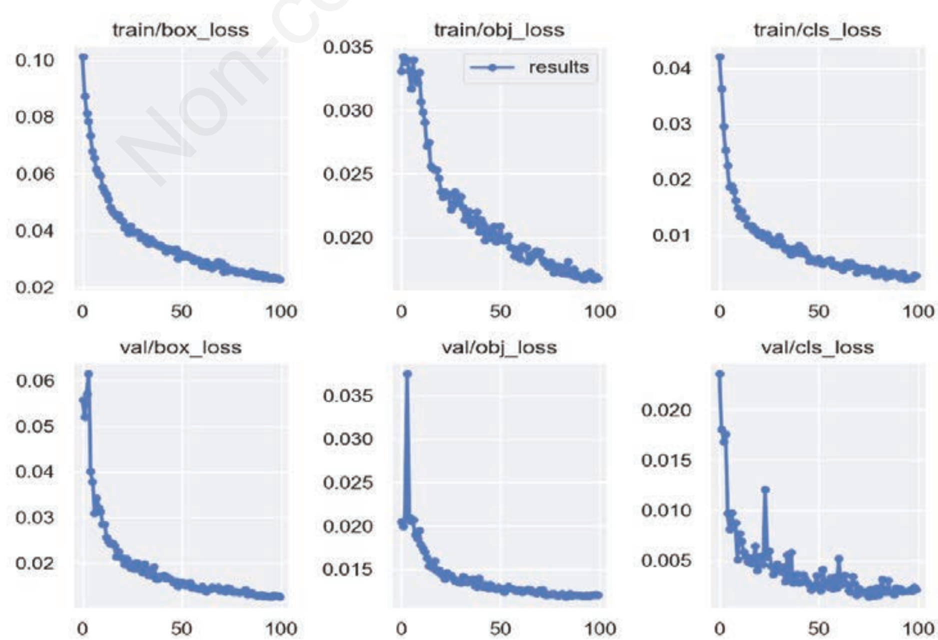

Figure 12. Loss curve of the proposed method.

The results in Table 4 demonstrate that, when compared to Faster R-CNN, SRN-Faster R-CNN, YOLOv5s, and DeblurGAN-YOLOv5s, the average precision of early blight detection improved by 10.9%, 9.1%, 15.3%, and 10.9% respectively. Moreover, the average precision of late blight detection increased by 7.3%, 4.7%, 4.0%, and 3.5%. Additionally, the average precision of anthracnose detection saw improvements of 8.5%, 4.8%, 6.2%, and 2.8%. Lastly, the average precision for detecting healthy leaves was enhanced by 15.2%, 12.9%, 6.4% , and 5.3%. These outcomes clearly illustrate the effectiveness of the proposed approach. As can be seen from Figures 11 and 12, for bounding box detection box_loss, the two are relatively close in value, and for obj_loss, the two are relatively close in value, but the curve of the method in this paper is smoother. On the obj_loss curve of the validation set val, YOLOv5s jumps between 0.175 and 0.125, and then fluctuates around 0.125. However, the method in this paper has been relatively stable below 0.125 after being reduced to 0.125. It can be seen from the curves and data that the loss curve of the proposed method on the test set val is smoother, which further verifies the effectiveness and advantages of the proposed method.

Test

After completing all the training, the test set images of the potato late blight leaf dataset were utilized to evaluate the performance of each model. The specific results are presented in Table 5. As depicted in Table 5, SRN-YOLOv5s outperforms Faster R-CNN, SRN-Faster R-CNN, YOLOv5s, and DeblurGAN-YOLOv5s with a precision rate improvement of 17.1%, 13.7%, 8.6%, and 5.8% respectively when detecting and recognizing potato late blight leaf images. Moreover, there is a recall rate increase of 17.8%, 12.8%, 8.9%, and 4.5% respectively for these models. In terms of detection speed, SRN-YOLOv5s exhibits faster performance compared to Faster R-CNN, SRN-Faster R-CNN, and DeblurGAN-YOLOv5s but slightly slower than YOLOv5s itself.

Table 4. The training outcomes of various models.

Model	AP comparison of different targets in the training dataset			
	Early blight	Late blight	Anthracnose	Health
Faster R-CNN	0.664	0.802	0.775	0.843
SRN-Faster R-CNN	0.682	0.828	0.792	0.866
YOLOv5s	0.62	0.835	0.798	0.931
DeblurGAN-YOLOv5s	0.664	0.84	0.832	0.942
SRN-YOLOv5s	0.773	0.875	0.86	0.995

Table 5. Detection results of potato late blight leaf images by different models.

Model	Precision	Recall	F1	Fps ($f \cdot s^{-1}$)	Model memory usage (MB)	Deblurring detection function
Faster R-CNN	0.732	0.702	0.717	11	110.2	No
SRN-Faster R-CNN	0.766	0.752	0.759	70	120.6	Yes
YOLOv5s	0.817	0.791	0.804	51	26.9	No
DeblurGAN-YOLOv5s	0.845	0.835	0.839	35	50.3	Yes
SRN-YOLOv5s	0.903	0.880	0.891	42	34.4	Yes

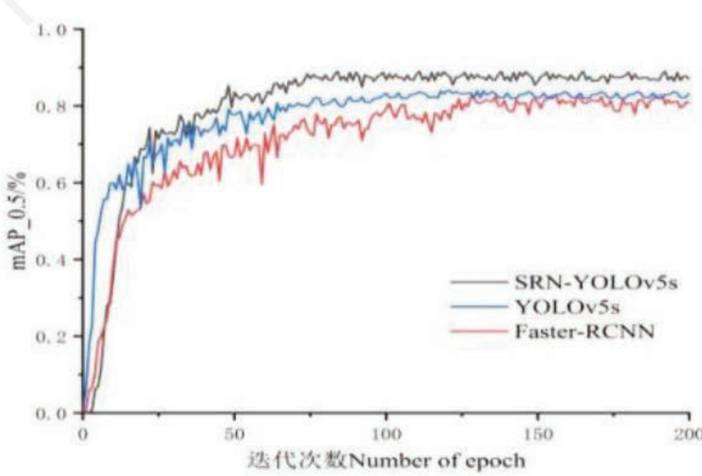

Figure 13. Changing curves of mean average precision.

Although SRN-YOLOv5s has fewer parameters and a lower detection speed compared to YOLOv5s, it significantly surpasses YOLOv5s in terms of precision and recall rate. This enhancement greatly improves both the precision (90.3%) as well as the detection speed (42 frames/s) for potato late blight leaf image recognition. Figure 13 illustrates the comparison of mAP change curves among different models.

For the optimization method, we selected five different field planting environments for testing. Figure 14 presents samples of the detection results of YOLOv5s and the optimized method. By observing Figure 14, it can be found that the late blip disease in the test images of Figure 14 a,c was correctly detected with a confidence of 0.8. In Figure 14b, the SRN-YOLOv5s detection confidence of 0.7 is higher than that of YOLOv5s detection confidence of 0.5; In Figure 14 d,e, the number of late blight leaves detected by SRN-YOLOv5s was higher than that detected by YOLOv5s. Therefore, compared with YOLOv5s, SRN-YOLOv5s can better solve the problems of missed detection, false detection and low accuracy.

According to the analysis of the above experimental results, the reasons for the difference in recognition effect are as follows:

i) The color, texture, and shape characteristics of diseased leaves are crucial for identifying them. However, in the early stages of the disease, there are no apparent signs of leaf lesions, and the differences in color information within diseased regions are minimal. Therefore, accurately extracting characteristic information from these regions remains challenging.
ii) Due to the similarity in appearance of dark-brown patches on early infected leaves caused by both early blight and late blight, accurate differentiation becomes challenging. Additionally, the recognition accuracy of diseased leaves is further reduced as different texture features may also be influenced by pixel variations in captured images.
iii) The leaves affected by anthracnose initially exhibit a pale yellow coloration, followed by the emergence of small brown spots that gradually curl inward as the affected area expands. In comparison to early blight and late blight, the distinct characteristics of anthracnose spots on leaves contribute to a higher recognition rate.

Discussion

Other researchers have studied automatic disease recognition in crops such as wheat, apples and tomatoes. According to the data results of literature (Li *et al*., 2022), the SSD-based method has only 81% recognition precision. However, in literature (Li *et al*., 2023), when YOLOv5 was used to conduct classification research on tomatoes, the model detection speed was 38 frames/s, slightly lower than 42 frames/s in this paper. In addition, the identification of tomato yellow leaf curl disease based on YOLOv5s method was carried out in literature (Zuo *et al*., 2023), and the combination of Plant Village data set and self-collected data set was adopted to enrich the image data, thus improving the generalization ability of the model. In addition, the number of original data sets used in literature (Sun *et al*., 2022) reached 1,400, which showed various diseases of apples to a large extent. These relevant studies prove that this paper has higher precision and faster detection speed than traditional detection methods, and has achieved good speed improvement compared with YOLOv5. At the same time, the data set containing single blade and multi-blade images can meet the requirement of target detection in the actual environment. In general, the detection accuracy and speed of the proposed method for potato leaf disease images meet the expected requirements, and the disease recognition accuracy is high, the speed is fast, and the robustness is good. In practical application research, crop disease images in natural environment are easily affected by illumination, occlusion, tilt, background noise and other factors. In-depth study of potato leaf disease recognition in the above complex scenes is the focus of the next research.

Conclusions

In this paper, we propose an improved YOLOv5 convolutional neural network-based algorithm, SRN-YOLOv5s, for detecting potato leaf diseases. The algorithm achieves a precision of 0.903 on the test set while maintaining a model weight size of 34.4 MB. Compared to the original model, our proposed algorithm demonstrates an improvement in precision by 8.6 percentage points with-

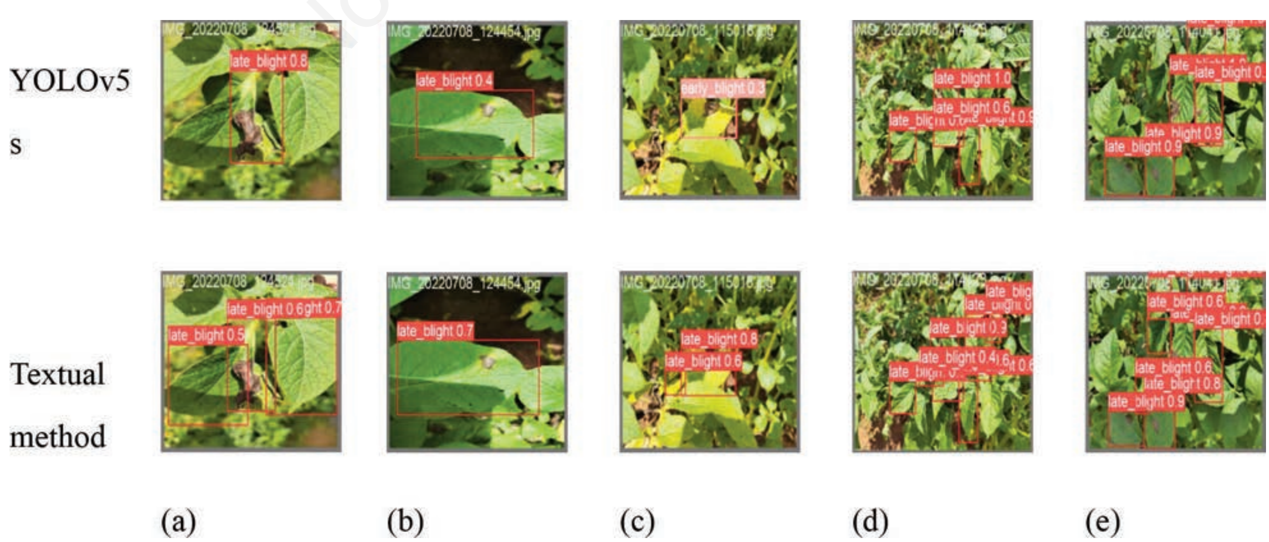

Figure 14. Visual result.

out significantly increasing the model's weight. Furthermore, when compared to YOLOv5 and Faster R-CNN methods that also utilize DeblurGAN technology, our algorithm exhibits distinct advantages. The deformable convolution was introduced as a replacement for the standard convolution in the backbone network of YOLOv5s, while the CBAM attention mechanism was employed to enhance the model's feature extraction capability. Additionally, WIoU_Loss was utilized as the loss function for bounding box regression. As a result, there were observed improvements in mAP by 1.6 percentage points, 2.2 percentage points, and 0.7 percentage points respectively. By combining the SRN network after model compression, the mAP value is increased by 0.5 percentage points, effectively mitigating the impact of low-quality images on the model. At the same time, the model achieved a detection frame rate of 42 frames per second, which exceeded other detection models in potato late blight detection, and performed well in meeting the requirements of edge equipment deployment.

References

Cai, S.P., Pan, W.H., Liu, H. 2023. Orchard obstacle detection based on D2-YOLO deblurring recognition network. Trans. Chin. Soc. Agr. Machin. 54:284-292.

Dai, J., Qi, H., Xiong, Y., Li, Y, Zhang, G., Hu, H., Wei, Y. 2017. Deformable convolutional networks. EEE International Conference on Computer Vision (ICCV), Venice, pp. 764-773

Dang, M.Y., Meng, Q.K., Gu, F., Gu, B., Hu, Y.H. 2020. Rapid recognition of potato late blight based on machine vision. Trans. Chin. Soc. Agr. Eng. 36:193-200.

Huang, C., Liu, W.C. 2016. Occurrence characteristice and monitoring advice of potato late blight in China in recent years. Plant Protec. 42:142-147.

Fan, Z.J., Li, X.X. 2019. Recognition of potato diseases based on fast detection and fusion features of ROI. Southw. China J. Agr. Sci. 32:544-550.

Li, H., Yan, K.H., Jing, H. 2022. Pathological detection and identification of apple leaves based on improved SSD. Sens. Microsyst. 41:134-137.

Li, J.H., Lin, L.J., Tian, K. 2020. Detection of leaf diseases of balsam pear in the field based on improved Faster R-CNN. Trans. Chin. Soc. Agr. Eng. 36:179-185.

Li, T., Ren, L., Hu, B. 2023. Grading detection of tomato hole-pan seedlings using improved YOLOv5s and transfer learning. Trans. Chin. Soc. Agr. Eng. 39:174-184.

Li, Y.H., Huo, C., Cao, J.F. 2022. Epidemics and infection characteristics of potato late blight in different seasons in Yunnan province. Southw. China J. Agr. Sci. 35:2046-2053.

Lin, T.Y., Dollar, P., Girshick, R., He, K., Hariharan, B., Belongie, S. 2017. Feature pyramid networks for object detection. Proc. IEEE Conf. on Computer Vision and Pattern Recognition. pp. 2117-2125.

Peng, H.X., Huang, B., Shao, Y.Y. 2018. General improved SSD model for picking object recognition of multiple fruits in natural environment. Trans. Chin. Soc. Agr. Eng. 34:155-162.

Sun, G.F., Wang, Y.L., Lan, P. 2022. Identification of apple fruit diseases using improved YOLOv5s and transfer learning. Trans. Chin. Soc. Agr. Eng. 38:171-179.

Tao, X., Gao, H., Wang, Y., Shen, X., Wang, J., Jia, J. 2018. Scale-recurrent network for deep image deblurring. arXiv: 1802.01770.

Tong, Z., Chen, Y., Xu, Z., Yu, R. 2023. Wise-IoU: bounding box regression loss with dynamic focusing mechanism. arXiv: 2301.10051.

Wang, X.Y., Li, Y.X., Yang, Z.Y. 2021. Detection method of clods and stones from impurified potatoes based on improved YOLO v4 algorithm. Trans. Chin. Soc. Agr. Machin. 52:241-247.

Woo, S., Park, J., Lee, J.Y., Kweon, I.S. 2018. Cbam: Convolutional block attention module. CBAM: Convolutional Block Attention Module. In: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (eds.), Computer Vision – ECCV 2018. Lecture Notes in Computer Science, vol 11211. Cham, Springer.

Xi, R., Jiang, K., Zhang ,W.Z. 2020. Recognition method for potato buds based on improved faster R-CNN. Trans. Chin. Soc. Agr. Machin. 51:216-223.

Xiao, Z.Y., Liu, H. 2017. Adaptive features fusion and fast recognition of potato typical disease images. Trans. Chin. Soc. Agr. Machin. 48: 26-32.

Xiu, C.B., Sun, L.L. 2022. Potato leaf bud detection method based on improved YOLO v4 network. Trans. Chin. Soc. Agr. Machin. 53:265-273.

Yang, S., Feng, Q., Zhang, J.H. 2020. Identification method for potato disease based on deep learning and composite dictionary. Trans. Chin. Soc. Agr. Machin. 51:22-29.

Zhang, Z.G., Zhang, Z.D., Li, J.N. 2021. Potato detection in complex environment based on improved YoloV4 model. Trans. Chin. Soc. Agr. Eng. 37:170-178.

Zheng, Z.H., Wang, P., Liu ,W., Li, J.Z., Ye, R.G., Ren, D.W. 2020. Distance-IoU loss: Faster and better learning for bounding box regression. Proc. AAAI Conf. on Artificial Intelligence 34:12993-13000.

Zuo, H.X., Huang, Q.C., Yang, J.H. 2023. Improved YOLOv5s-based detection method for tomato yellow leaf curl virus disease. Trans. Chin. Soc. Agr. Machin. 1-11.