Article

Two-Stage Detection Algorithm for Plum Leaf Disease and Severity Assessment Based on Deep Learning

Caihua Yao [1,2,†], Ziqi Yang [1,2,†], Peifeng Li [1,2], Yuxia Liang [1], Yamin Fan [1,2], Jinwen Luo [1,2], Chengmei Jiang [1,2] and Jiong Mu [1,2,3,*]

[1] College of Information Engineering, Sichuan Agricultural University, Ya'an 625000, China;
202105694@stu.sicau.edu.cn (C.Y.); 202206324@stu.sicau.edu.cn (Z.Y.); 202204858@stu.sicau.edu.cn (P.L.);
202205692@stu.sicau.edu.cn (Y.L.); 202205660@stu.sicau.edu.cn (Y.F.); 202105716@stu.sicau.edu.cn (J.L.);
202105718@stu.sicau.edu.cn (C.J.)
[2] Sichuan Key Laboratory of Agricultural Information Engineering, Ya'an 625000, China
[3] Ya'an Digital Agricultural Engineering Technology Research Center, Ya'an 625599, China
[*] Correspondence: jmu@sicau.edu.cn; Tel.: +86-133-4060-8699
[†] These authors contributed equally to this work.

Abstract: Crop diseases significantly impact crop yields, and promoting specialized control of crop diseases is crucial for ensuring agricultural production stability. Disease identification primarily relies on human visual inspection, which is inefficient, inaccurate, and subjective. This study focused on the plum red spot (*Polystigma rubrum*), proposing a two-stage detection algorithm based on deep learning and assessing the severity of the disease through lesion coverage rate. The specific contributions are as follows: We utilized the object detection model YOLOv8 to strip leaves to eliminate the influence of complex backgrounds. We used an improved U-Net network to segment leaves and lesions. We combined Dice Loss with Focal Loss to address the poor training performance due to the pixel ratio imbalance between leaves and disease spots. For inconsistencies in the size and shape of leaves and lesions, we utilized ODConv and MSCA so that the model could focus on features at different scales. After verification, the accuracy rate of leaf recognition is 95.3%, and the mIoU, mPA, mPrecision, and mRecall of the leaf disease segmentation model are 90.93%, 95.21%, 95.17%, and 95.21%, respectively. This research provides an effective solution for the detection and severity assessment of plum leaf red spot disease under complex backgrounds.

Keywords: deep learning; computer vision; plum disease detection; precision agriculture

1. Introduction

Leaf diseases are widespread and significantly impact crop quality and yield, often showing early symptoms on leaves [1], making leaf diseases crucial indicators for the early prevention and control of crop diseases. Plums, known for their rich nutritional content, are widely cultivated in China [2]. However, during the cultivation process, plum trees are highly susceptible to various diseases [3]. Mild infections can lead to a decline in fruit quality, while severe infections can cause tree weakening and inedible fruits [4,5], as with plum red spot [6]. At present, the diagnosis of most plum diseases mainly depends on a visual assessment by plant disease experts, which is time-consuming and labour-intensive and easily interfered with by the subjective judgment of the diagnosticians. Especially in large-scale planting, if the infected leaves are not identified and removed promptly, the disease can quickly spread to other parts and lead to large-scale diseases [7,8]. Therefore, the study of plum leaf disease detection is of great significance to agriculture, as it can guide growers in eradicating pathogens at an early stage, thus reducing the use of pesticides [9] and ensuring the safety of products.

Plum red spot [10,11] is caused by the infection of *Polystigma rubrum (Pers.)DC.*, occurs more seriously in the current stage of plum cultivation, and produces greater harm. Summer

is the peak season of plum red spot disease, and the high temperature and humidity of the climate make the pathogen multiply and spread rapidly, which is easy to cause a wide range of disease infections. Plum red spot often affects the leaves and fruits of plum trees. In the early stage of the infection, the leaves turn yellow and develop slightly raised spots. As the disease progresses, these spots expand and darken, causing leaves to fall off, affecting photosynthesis, which may ultimately result in the weakening of the plum tree. Fruits infected with the disease will produce orange-red circular spots on the surface, which will deepen in colour in later stages, making the fruits inedible. Diseased fruits are often deformed and easy to fall off before ripening, which leads to a reduction in the yield and quality of plums.

With the rise of precision agriculture, the advantages of computer vision technology in detecting plum leaf diseases have gradually become prominent. Compared to traditional artificial vision detection, computer vision technology offers fast speed, high precision, and multiple functions, and has been generally applied in agricultural domains, such as with the detection of corn phenotypes by computer vision technology [12] and remote sensing image segmentation of farmland [13], etc. Applying computer vision technology to the field of crop diseases, such as corn [14,15], potato [16,17], kiwifruit [18,19], and rice [20,21], can help to mitigate the adverse effects of crop diseases on agricultural production, promoting the development of agricultural production development towards high quality and high yield.

Current leaf disease detection models struggle with multiple leaves and complex backgrounds. Many types of research focus on detecting single-leaf diseases under simple backgrounds, mainly focusing on the colour characteristics of the disease. In contrast, research on the textural characteristics of the disease is insufficient and lacks a severity assessment of the disease. Xu [22] et al. used an improved YOLOv5 model to detect melon leaf diseases under complex backgrounds, only performing simple detection and location of diseases without further severity assessment. Shu [23] et al. proposed an improved DeepLabv3+ grape disease segmentation model using texture features of leaves and lesions. However, this research was limited to an experimental environment, reducing its effectiveness in complex backgrounds. Divyanth [24] et al. proposed a two-stage corn leaf disease segmentation model under a complex background, firstly using U-Net to segment the corn leaves from the complex background and then using DeepLabV3+ to segment the disease, which eliminated the effect of the complex background on the disease segmentation; but this study focused on the disease detection of single leaves under complex backgrounds.

Based on these issues, we collected a high-quality dataset of plum red spot leaves under natural conditions and proposed a two-stage plum leaf disease detection algorithm, and we also proposed a disease segmentation model MOC_UNet based on improved U-Net. We firstly used the advanced object detection model YOLOv8 [25] to strip out leaves from complex backgrounds in order to eliminate the complex background from the subsequent interference of disease spot segmentation. Then, we put the stripped diseased leaves into the disease segmentation model MOC_Unet to segment the leaves and disease spots accurately. Finally, the disease severity was preliminarily assessed by calculating the disease spot coverage rate. The MOC_UNet combined Dice Loss [26] with Focal Loss [27] to address poor training performance caused by the imbalance of pixel ratios between disease spots and leaves. To address the inconsistencies in the size and shape of leaves and lesions, we used ODConv [28]. Also, we introduced MSCA [29] so that the model could pay more attention to features of different scales and achieve better segmentation of target boundaries. The accuracy of leaf recognition reached 95.3%, and the mIoU, mPA, mPrecision, and mRecall of the leaf disease segmentation model reached 90.93%, 95.21%, 95.17%, and 95.21%, respectively. This indicates that the two-stage detection algorithm demonstrates high accuracy and strong robustness.

2. Materials and Methods

2.1. Data Acquisition and Processing

The experimental data in this study were collected from a plum plantation in Gulin County, Luzhou, Sichuan Province, using a Canon EOS60d camera to record in the summer of 2022. We shot from multiple angles to simulate the images of plum leaves collected by unmanned vehicles and drones under natural conditions in real life, as shown in Figure 1. The dataset included images of plum leaves of different ages, varieties, and weather environments. This study collected 447 images of plum leaves under natural conditions.

Figure 1. Plum leaf dataset under natural conditions.

To ensure the accurate detection of plum leaves and the data sufficiency of diseased leaves in the second stage, we used some stochastic data augmentation techniques, including rotation, translation, random cropping, adding noise, changing brightness, and flipping operations to expand the plum leaf dataset by ensuring that at least one operation worked, as shown in Figure 2. As much as possible, to simulate different light, complex and diverse environments were included to achieve the accurate detection of plum leaves in different environments.

Figure 2. Comparison of plum leaf dataset before and after data augmentation.

After detection of leaves by YOLOv8, we stripped the identified plum leaves and screened out 6321 images of leaves with red spot disease, as shown in Figure 3. We used the semantic segmentation annotation tool LabelMe to annotate red spot-diseased leaves and their lesions, and obtained the semantic segmentation dataset of plum red dot-diseased leaves. This dataset was then divided into training and validation sets with a 7:3 ratio.

Figure 3. Red spot-diseased leaves after detection.

2.2. Overall Algorithm Workflow

We propose a two-stage detection and severity assessment algorithm for plum leaf red spot disease, mainly consisting of three modules. Firstly, YOLOv8 is used to peel off a single diseased leaf from a complex background, and then it is sent to the disease segmentation model MOC_UNet to segment the leaves and disease spots. Finally, the coverage rate of disease spots is calculated by counting the number of disease spots and leaf pixels. Then the disease severity assessment results are output, as shown in Figure 4.
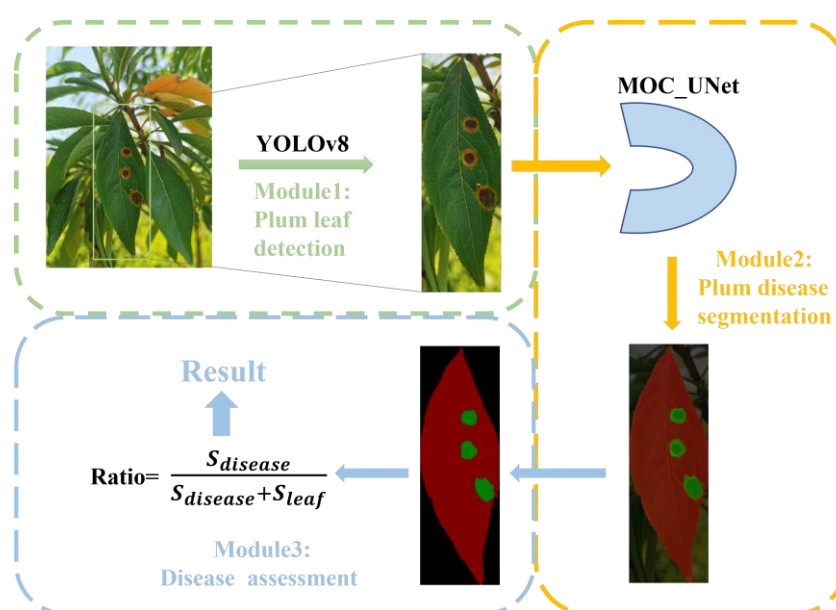

Figure 4. Flow chart of disease detection.

2.3. Plum Leaf Detection Model Based on YOLOv8

YOLO [30] is a one-stage object detection algorithm comprising three core components: neck, backbone, and head. The Yolo series algorithms can detect the position and category information of the target object simultaneously, allowing for end-to-end training and high-speed detection, making it suitable for detecting plum leaves.

YOLOv8 [25], a newer YOLO series algorithm, uses the C2f structure instead of the C3 structure of YOLOv5 [31] in Backbone and Neck, which ensures it is lightweight and at the same time obtains richer information about the gradient flow. Using the SPPF module reduces the computational amount to a certain extent and increases the Receptive Field. In terms of head components, the head part of YOLOv8 adopts a decoupled-head structure similar to that of YOLOX [32], which separates the classification and detection heads, and introduces an Anchor-Free Detection Head, which provides greater flexibility and can better adapt to targets with various shapes and sizes. As for the loss function, YOLOv8 uses VFL Loss as the classification loss and DFL Loss+CIOU Loss as the regression loss, which can somewhat improve the rate of convergence and performance of the model. Thus, YOLOv8 is suitable to be used as a plum leaf detection model in farmland scenarios. The structure of YOLOV8 is shown in Figure 5.
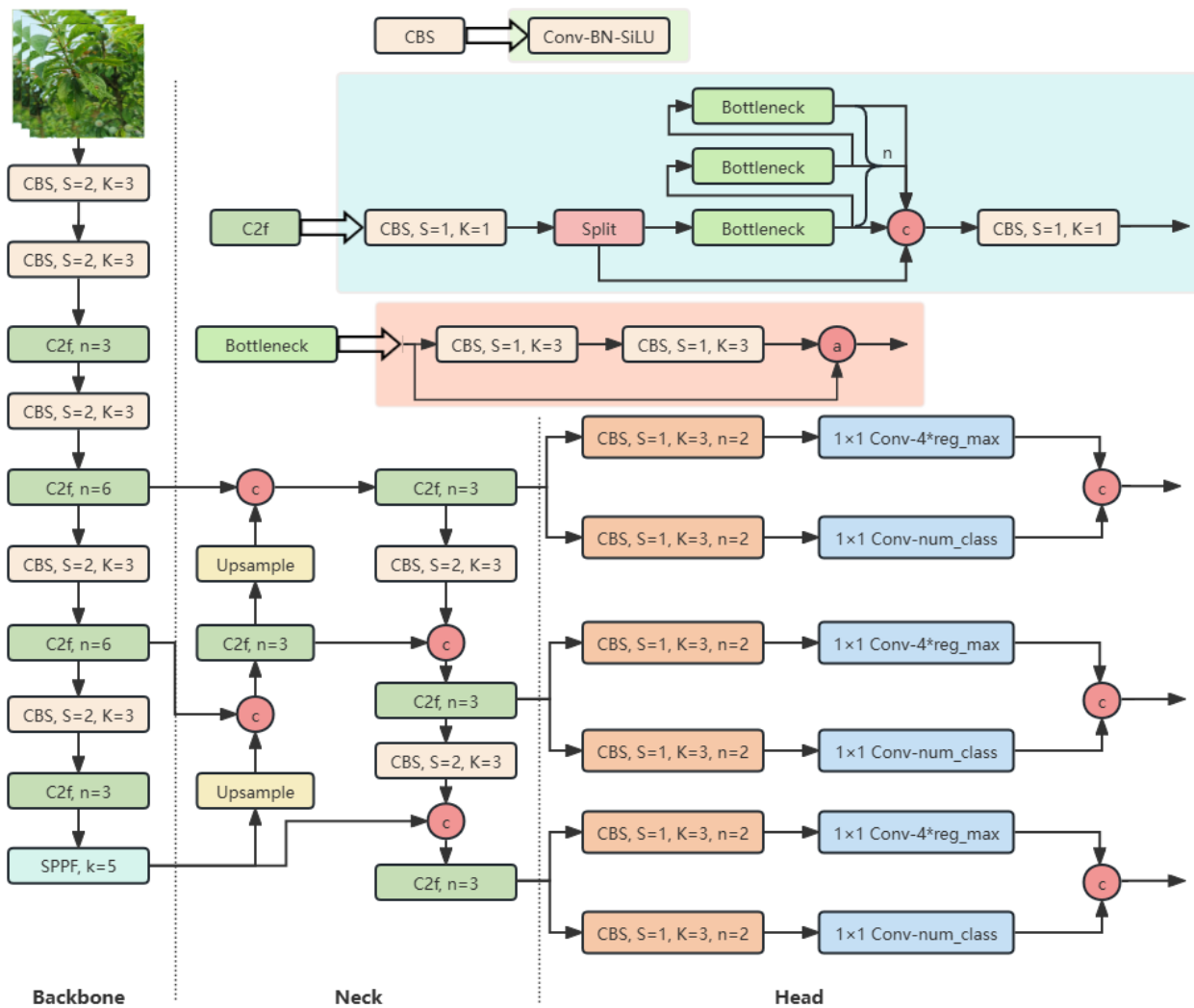

Figure 5. YOLOV8 structure diagram.

2.4. Leaf Disease Segmentation Model Based on MOC_UNet

2.4.1. U-Net Model

Image semantic segmentation is a crucial field in computer vision. It refers to pixel-level recognition of images, which means labelling the category of each pixel in the image. Semantic segmentation models like FCN [33], SegNet [34], PSPNet [35], DeepLab [36], and U-Net [37] are among the most representative.

U-Net is a semantic segmentation network for medical imaging applications inspired by FCN by Ciresan et al. in 2015. The U-Net algorithm is characterized by a "U-shaped" network structure composed of an Encoder and a Decoder. In U-Net, the skip connection mechanism is introduced to fuse the decoder's output features with the encoder's semantic features, which can effectively capture different levels of feature information and improve image segmentation accuracy and preserve details. Compared with other deep learning algorithms, U-Net can learn a highly robust network with less data. Therefore, it is especially suitable for few-shot, unbalanced data and tasks requiring detailed information retention.

The structure of the improved U-Net used in this paper, MOC_UNet, is shown in Figure 6.
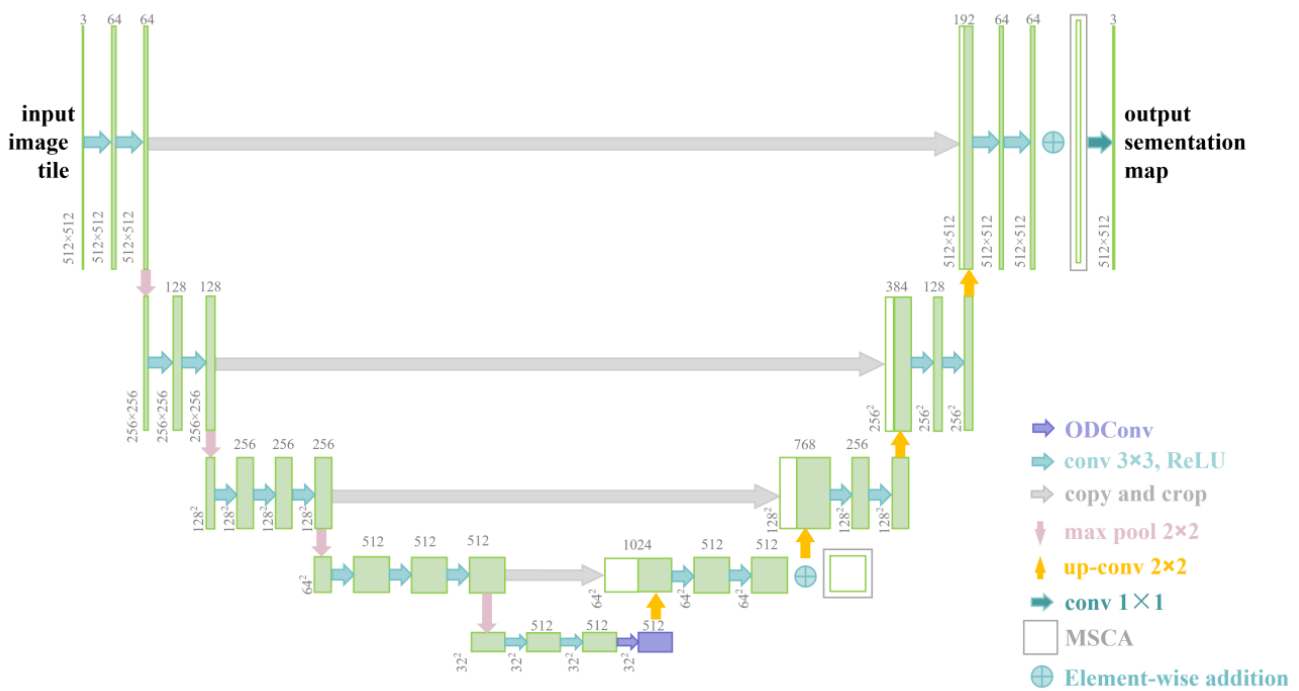

Figure 6. MOC_UNet Network Architecture.

2.4.2. Omni-Dimensional Dynamic Convolution Module

Most of our detection of diseased plum leaves is carried out in a complex back-ground, surrounded by many interfering factors affecting the feature extraction of the network, resulting in the scale features of the diseased leaves themselves not being fully acquired; in addition, the detection of diseased plum leaves is mostly carried out outdoors, which requires higher detection accuracy and detection speed. To address the above problems, we chose ODConv, whose deformable convolution with extra offset can successfully use the potential of spatial information to completely obtain the information within the effective area of the sampling point, thus obtaining a better performance [38], and compared with other dynamic convolution algorithms, ODConv has only one convolution kernel, and the number of parameters is much smaller, which ensures efficiency while ensuring accuracy. And its generalisation capability is sufficient for the outdoor detection of diseased plum leaves [39].

ODConv [28] can be regarded as a continuation of CondConv [40] (Conditionally Parameterized Convolution) and DyConv [41] (Dynamic Convolution). By leveraging a multi-dimensional attention mechanism to compute four types of attentions along all four dimensions of the kernel space in parallel, these attentions are multiplied by the convolution kernel W_i. Thus, a linear combined convolution weighted according to multi-dimensional attentions is generated, as shown in Figure 7. Different convolution combinations provide performance guarantees for capturing rich contextual cues, enhancing the feature extraction ability of the network for leaves and red spots with different shapes and sizes, to capture the edges and textures of different regions in the image more accurately and improve the segmentation accuracy and detail performance. ODConv is calculated by the following formula:

$$y = (\alpha_{w1} \odot \alpha_{f1} \odot \alpha_{c1} \odot \alpha_{s1} \odot W_1 + \ldots + \alpha_{wn} \odot \alpha_{fn} \odot \alpha_{cn} \odot \alpha_{sn} \odot W_n) * x \quad (1)$$
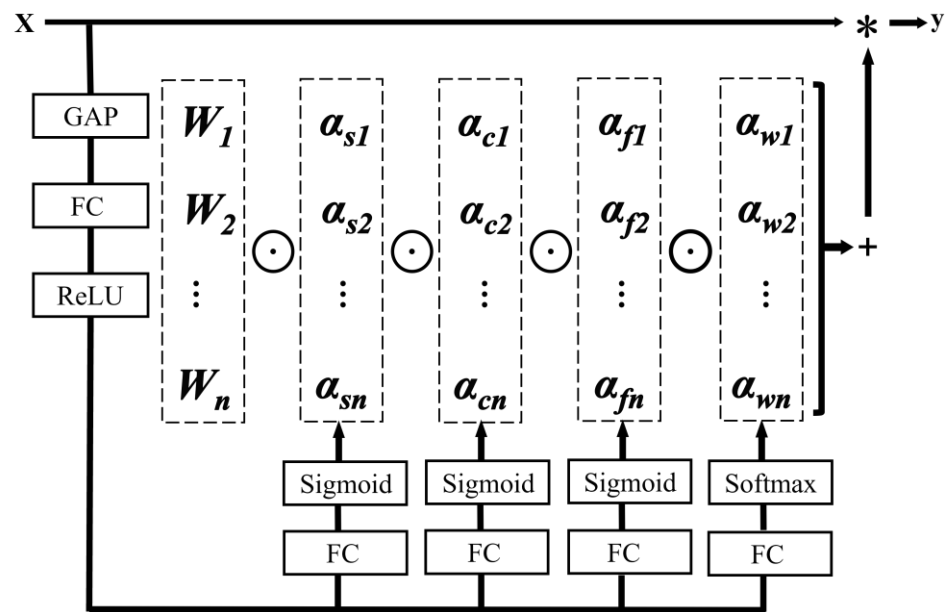

Figure 7. Schematic diagram of ODConv.

α_{wi}, α_{fi}, α_{ci}, and α_{si} denote the attention of four dimensions of the kernel, output channel, input channel, and space along the convolutional kernel W_i, respectively. $*$ denotes the convolution operation.

2.4.3. Multi-Scale Convolutional Attention Module

In the assessment of the severity of plum leaf disease by the ratio of the lesion area, the accurate segmentation of leaf and lesion is very important. Accurate boundary segmentation can ensure the accurate calculation of lesion coverage. Therefore, the model needs to capture and distinguish the edge features in detail. The MSCA attention mechanism can enhance the model's ability to perceive important information between image channels at multiple scales, especially in the processing of complex backgrounds and subtle spot edges, showing excellent performance. Qian et al. significantly improved the accuracy of density map estimation in cell-counting tasks through an innovative MSCA-UNet architecture, thereby enhancing the accuracy of cell counting [42].

As shown in Figure 8, the attention mechanism MSCA [29] consists of three parts: Firstly, local information is aggregated by depth-wise convolution to extract rich feature representations and expand the Receptive Field. Furthermore, multi-branch depth-wise strip convolution is used to capture multi-scale contextual information to comprehensively perceive features at different scales and accurately segment the leaves and lesions with different sizes and shapes. Lastly, the outputs of the 1×1 convolution are used as the atten-

tion weights to enhance the weights of leaves and disease spot features. Two depth-wise strip convolutions are used in each branch to simulate large kernel depth-wise convolutions with kernel sizes of 7, 11, and 21, respectively, to simulate different Receptive Fields. This structural design enables MSCA to process both local and global information, enhancing the robustness and accuracy of the model. Therefore, MSCA was applied in this study to effectively solve the shortcomings of tiny lesion capture challenges and inaccurate leaf edge segmentation in plum leaf disease segmentation. The expression for MSCA is as follows:

$$\text{Att} = \text{Conv}_{1\times1}\left(\sum_{i=0}^{3}\text{Scale}_i(\text{DWConv}(F))\right) \tag{2}$$

$$\text{Out} = \text{Att} \otimes F \tag{3}$$

where F and Out denote the input and output features, respectively; Att is the attention map; $\otimes$ is the multiplication operation of the element-wise matrix; and DWConv and Scale_i are depth-wise convolution and ith branch, respectively.
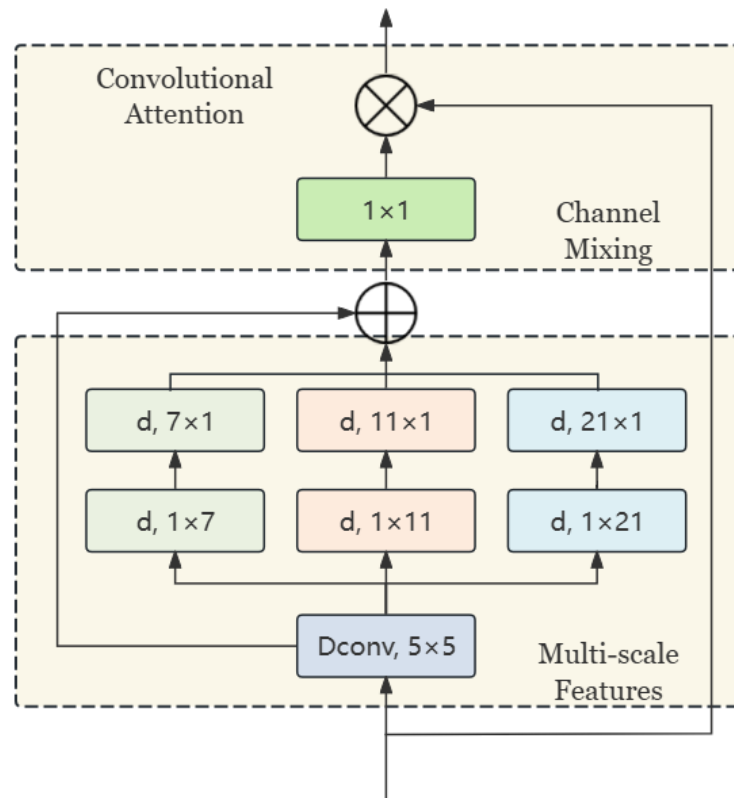

Figure 8. MSCA structure diagram.

2.4.4. Combined Loss

The number of pixels in the three categories of background, disease spots, and leaves in the plum disease leaf dataset is unbalanced, leading to an imbalance in model training. Therefore, this study combines Focal Loss and Dice Loss as a new loss function, which can pay more attention to foreground targets and inaccurately categorised samples while paying attention to the overall loss. It is defined in the following formula: L is the constructed loss function, L_F is Focal Loss, and L_D is Dice Loss.

$$L = L_D + L_F \tag{4}$$

(1) Focal Loss [27]: Focal Loss decreases the loss of accurately classified samples without changing the loss of inaccurately classified samples. It is helpful to improve

the accuracy of inaccurately classified samples by making the loss function tend to be inaccurately classified samples. The formula is as follows: where P_t represents the model's predicted value, α_t and γ are the two parameters that regulate the Focal Loss.

$$L_F = -\alpha_t(1 - P_t)^{\gamma}\log(P_t) \tag{5}$$

(2) Dice Loss [26]: Dice Loss can mitigate the negative effect of foreground and background imbalance in the sample, making the training pay more attention to the foreground region for mining. Its mathematical expression is as follows: $|X \cap Y|$ represents the number of intersecting elements between X and Y; $|X|$ and $|Y|$ represent the number of X and Y elements, respectively.

$$L_D = 1 - Dice = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{6}$$

2.5. Disease Severity Assessment

As there are no clear criteria for grading the severity of plum red spot disease, existing studies on grading the severity of red dot disease on plum leaves have mainly been judged by counting the number of spots on the leaves, which lacks accuracy. Combining existing methods for assessing leaf disease severity, disease spot coverage can provide an effective index for quantitative assessment and realize the precise grading of disease severity [43]. Considering the irregular shape of the spots and leaves, it is difficult to measure their areas manually. Therefore, this study used a method based on the number of pixels to calculate the areas of spots and leaves.

Using the trained MOC_UNet model to predict the diseased leaf images, we obtained the matrix after pixel point classification, where the pixel points took the values of 0, 1, and 2 to indicate the background pixels, healthy leaf pixels, and disease spot pixels, respectively. We used the sum of healthy leaf pixels and diseased spot pixels as the number of pixels of the intact leaf. By calculating the ratio of pixels of the diseased spot to pixels of the intact leaf, we can calculate the percentage of the area of the diseased spots in the leaf, and the calculation formula is as follows:

$$\text{Disease Ratio} = \frac{S_{disease}}{S_{leaf} + S_{disease}} \times 100\% \tag{7}$$

$S_{disease}$ denotes the number of lesion pixels; the S_{leaf} denotes the number of healthy leaf pixels.

2.6. Experimental and Evaluation Indicators

The experimental platform uses an Ubuntu (64-bit) operating system equipped with a 12-core Intel (R) Xeon (R) Platinum 8255C CPU @ 2.50 GHz processor. The GPU model is RTX3090, and the open-source deep learning framework Pytorch is used as the development environment. The Cuda version is 11.1, and the computer memory is 43 GB.

We used mean Average Precision (mAP), precision (Precision, P), and recall (Recall, R) as evaluation metrics for the plum leaf detection models, and mean Intersection over Union (mIoU), pixel accuracy (PA), precision (Precision, P), and recall (Recall, R) as evaluation metrics for plum leaf disease segmentation models. We also used the Matthews Correlation Coefficient (MCC) as a metric to evaluate the performance of different segmentation models. Table 1 shows the definitions of the parameters in the calculation formula of these evaluation indexes.

Table 1. Definitions of relevant calculation parameters for evaluation indicators.

Confusion Matrix		Predicted Results	
		Positive	Negative
Expected Results	Positive	TP	FN
	Negative	FP	TN

TP: positive samples predicted by the model to be in the positive category; TN: negative samples predicted by the model to be in the negative category; FP: negative samples predicted by the model to be in the positive category; FN: positive samples predicted by the model to be in the negative category; k denotes the number of sample categories.

mIoU is a standard metric for semantic segmentation, which is the mean ratio of the intersection and merge set of true labels and predicted values, and is calculated as follows:

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{\text{TP}}{\text{FN} + \text{FP} + \text{TP}} \tag{8}$$

PA indicates the ratio of the number of correctly categorised pixels to the total number of pixels, calculated as follows:

$$\text{PA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{9}$$

Precision indicates the probability that the prediction is an actual positive sample in a positive sample. The formula is as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

Recall indicates the proportion of pixels that the model correctly determines to be in the positive category out of all the pixels that are actually in the positive category. The formula is as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{11}$$

AP denotes the integral of the P-index to the R-index. mAP refers to the average of the AP values for all categories.

$$\text{AP} = \int_{0}^{1} P(R)dR \tag{12}$$

$$\text{mAP} = \frac{1}{K} \sum_{i=1}^{K} \text{AP}_i \tag{13}$$

The Matthews Correlation Coefficient is used to evaluate the quality of the model classification and is applicable to datasets with category imbalance, where a higher MCC indicates better performance and is calculated as follows:

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{14}$$

In order to compare the performance of different segmentation models in predicting disease spot coverage, we established a regression relationship between the disease spot coverage predicted by different models and the true measurements. We used the coefficient of determination R^2 and the Mean Absolute Percentage Error (MAPE) to evaluate the performance of the prediction effect. The larger the R^2, the better the model fit, and the smaller the value of MAPE, the better the accuracy of the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\overline{y}_i - y_i)^2} \tag{15}$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{16}$$

3. Results

3.1. Plum Leaf Detection Results

YOLOv8 is used to strip the leaves from the complex background, helping the subsequent model to accurately segment the disease. mAP, Precision, and Recall evaluate the performance of the plum leaf detection model. Higher values mean that the model is better at recognizing the plum leaves. The plum leaf detection results are shown in Figure 9.
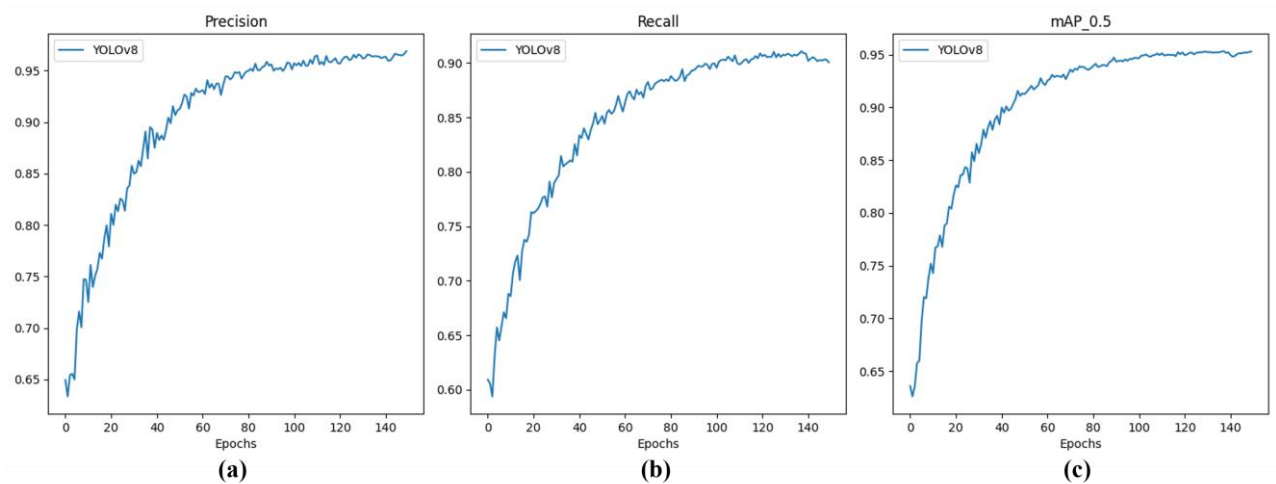

Figure 9. YOLOv8 results: (**a**) Precision; (**b**) Recall; (**c**) mAP@0.5.

As seen in Figure 9, the mAP@0.5 of the plum leaf detection model based on YOLOv8 reaches 95.26%, with Precision and Recall at 96.9% and 90.05%, respectively. These results suggest that YOLOv8 can recognize plum leaves accurately in complex backgrounds.

In order to better demonstrate the detection effect of plum leaves, three of the detection results are shown in Figure 10. It can be seen that YOLOv8 can accurately recognize the plum leaves in the foreground and detect the edge parts of the leaves accurately. Thus, YOLOv8 can meet the detection of plum leaves in real agricultural scenarios.

Figure 10. Effectiveness of YOLOv8 on the detection of plum leaves.

3.2. Leaf Disease Segmentation Results

In order to further verify the reliability of the disease leaf dataset obtained in the first stage and the performance of the disease segmentation model MOC_UNet proposed in this study, we compare it with the most representative semantic segmentation models, including PSPNet, DeepLabV3+, Segformer, and HRNetv2. To ensure a fair comparison, the comparison process uses the same training strategy and computing environment. As shown in Table 2, compared with PSPNet and DeepLabV3+, the U-Net network is more suitable for plum leaf and red spot segmentation, and our improved plum leaf disease segmentation model MOC_UNet has higher mIoU and mPA compared with other models, indicating a better segmentation effect is achieved by the improved U-Net model. And compared with other models, MOC_UNet has higher MCC, indicating that better classification results are achieved. So, in summary, MOC_UNet is more accurate at identifying and segmenting plum leaf red spot disease.

Table 2. Comparison of evaluation metrics for different network model segmentation.

Network Model	mIoU	mPA	mPrecision	mRecall	MCC
PSPNet	67.71%	76.40%	83.46%	76.40%	0.7325
DeepLabV3+	84.69%	91.29%	91.82%	91.29%	0.8790
Segformer	86.89%	92.61%	93.03%	92.61%	0.9001
HRNetv2	87.20%	92.41%	93.60%	92.41%	0.9033
U-Net	89.73%	94.35%	94.62%	94.35%	0.9234
MOC_UNet	90.93%	95.21%	95.17%	95.21%	0.9320

To verify the effectiveness of the improved disease segmentation model MOC_UNet, we carried out ablation experiments. The experimental results are shown in Table 3. The model's performance with the Combined Loss function is significantly improved, demonstrating that the combination of Focal Loss and Dice Loss can effectively overcome the drawbacks caused by the imbalance of different categories of pixels. The configurations using ODConv are effectively improved on mIoU and mPA, which shows that ODConv can effectively enhance the model's feature extraction capability for disease spots and leaves with different shapes and sizes. Especially, the model using Combined Loss, ODConv and MSCA simultaneously exhibits superior effectiveness compared to the original network, and the mIoU, mPA, mPrecision, and mRecall of MOC_UNet increase 1.2%, 0.86%, 0.55%, and 0.86%, respectively. Therefore, it indicates that all of our improvements can significantly enhance the segmentation performance of the leaf disease segmentation model.

Table 3. Results of ablation experiments.

Combined Loss	ODConv	MSCA	MIOU	mPA	mPrecision	mRecall
			89.73%	94.35%	94.62%	94.35%
√			90.44%	95.01%	94.83%	95.01%
	√		90.33%	94.61%	95.05%	94.61%
		√	89.99%	94.47%	94.81%	94.47%
√	√		90.76%	95.10%	95.08%	95.10%
√		√	90.54%	95.05%	94.90%	95.05%
	√	√	90.47%	94.68%	95.14%	94.68%
√	√	√	90.93%	95.21%	95.17%	95.21%

Figure 11 demonstrates the comparison of the effect of MOC_UNet with different segmentation models for plum disease leaf segmentation, and it can be found that the PSPNet, DeepLabV3+, Segformer, HRNetv2, and UNet models all have the problem of lesion misdetection (Figure 11c–g). Meanwhile, due to the complexity of the background, the other segmentation models are less effective for the boundary segmentation of leaves

and disease spots, while the leaf disease segmentation model MOC_UNet proposed in this study can well overcome the impact of the complex background on the image segmentation, it can segment leaves and disease spots accurately (Figure 11h), and the prediction results are similar to the labelled images.
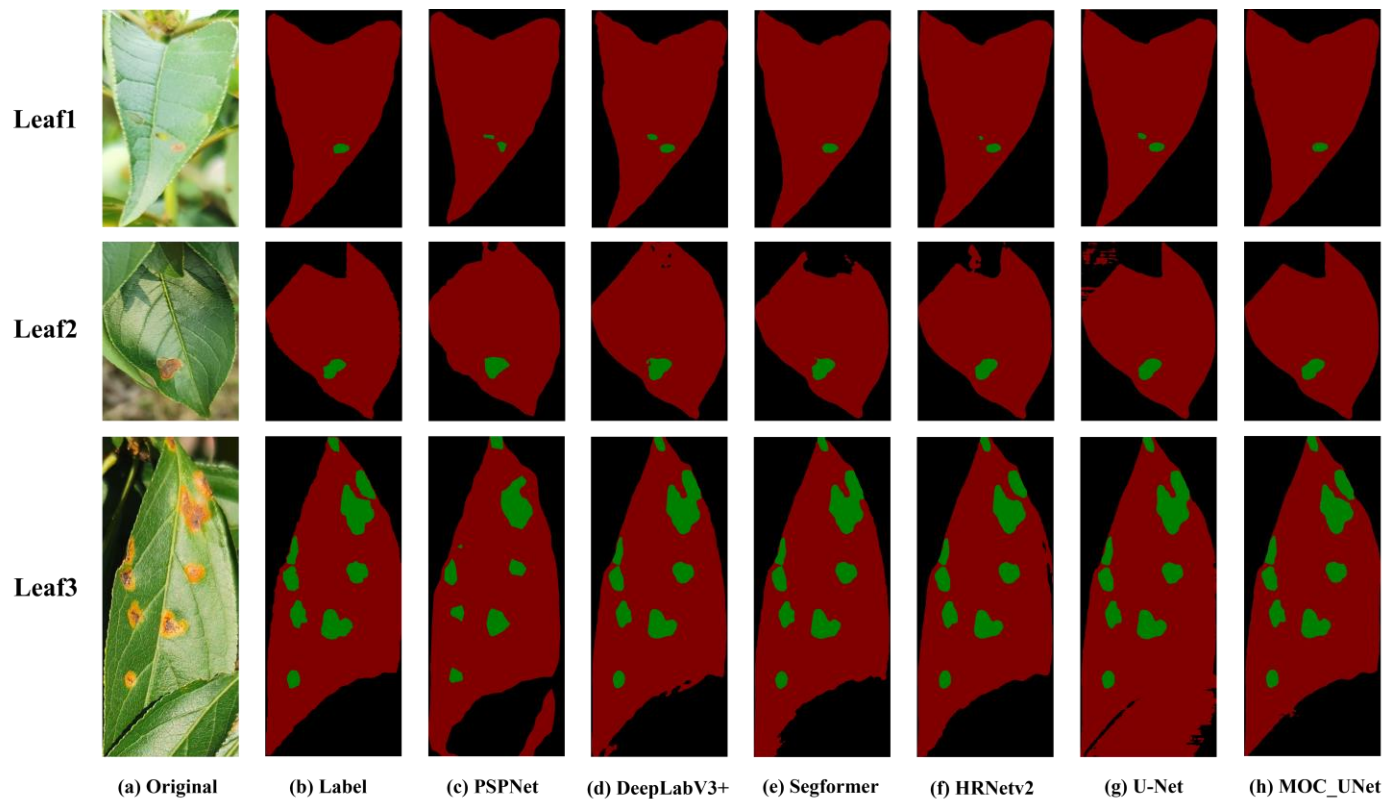

Figure 11. Comparison of the prediction effect of different segmentation models: (**a**) Original images; (**b**) Label images; (**c**) PSPNet; (**d**) DeepLabV3+; (**e**) Segformer; (**f**) HRNetv2; (**g**) U-Net; (**h**) MOC_UNet.

In order to test the effect of the leaf boundary prediction effect on the predicted lesion coverage, we predicted the lesion coverage of plum leaves in Figure 11 using the different models in Table 2, as shown in Table 4, and it can be observed that, due to the fact that there are misdetections in the boundary segmentation of leaves by PSPNet, DeepLabV3+, U-Net, Segformer, and HRNetv2, this results in a large degree of deviation between the predicted and true measurements of lesion coverage. In contrast, MOC_UNet was more effective in boundary segmentation, so the predicted value of lesion coverage was closer to the true value.

Table 4. Comparison of lesion coverage predictions from different models.

Measurement Methods	Leaf1	Leaf2	Leaf3
measured value	0.66%	2.19%	12.46%
PSPNet	0.63%	2.38%	8.71%
DeepLabV3+	0.97%	2.39%	12.51%
Segformer	0.73%	2.48%	12.51%
HRNetv2	0.71%	2.21%	12.03%
U-Net	0.93%	2.22%	10.18%
MOC_UNet	0.69%	2.21%	12.43%

To further validate the reliability of the disease segmentation model MOC_UNet in predicting the plum leaf red spot coverage, we employed linear regression analysis to assess

the relationship between the spot coverage predicted by the different models in Table 2 and the true measurements, which we tested using the test set. We used the coefficient of determination R^2 and Mean Absolute Percentage Error (MAPE) to evaluate the performance of the different models, as shown in Figure 12; the R2 for PSPNet, DeepLabV3+, HRNetv2, Segformer, U-Net, and MOC_UNet, respectively, were 0.70, 0.87, 0.92, 0.93, 0.93, and 0.96; and the MAPE were 30.97%, 17.27%, 14.31%, 13.38%, 12.05%, and 8.76%, respectively. The disease segmentation model MOC_UNet has the largest R2 and the smallest MAPE, so we believe that the plum leaf spot coverage predicted by MOC_UNet is closer to the true value, which can effectively help farmers to accurately detect plum leaf disease and initially determine the severity of the disease.
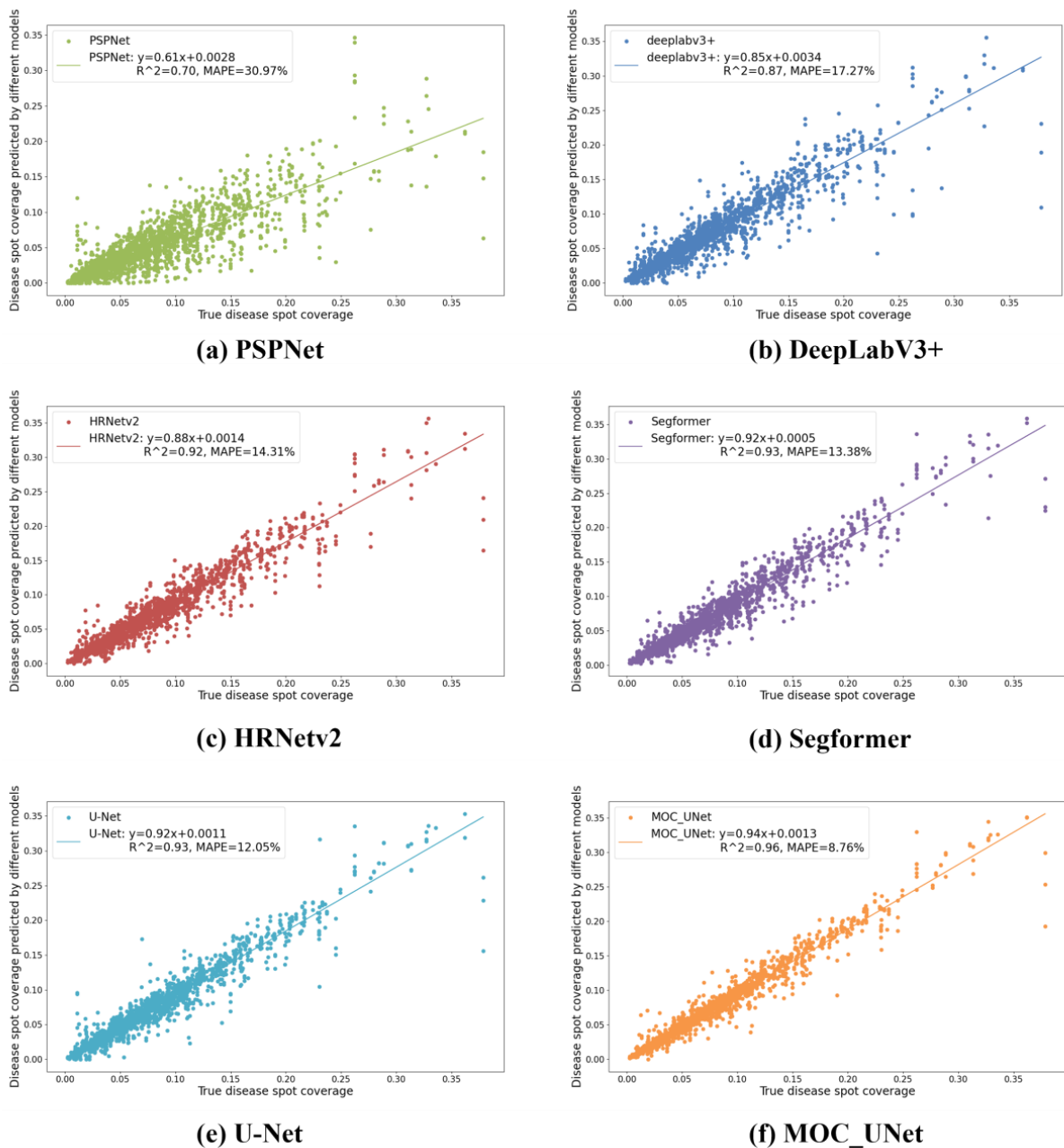

(a) PSPNet

(b) DeepLabV3+

(c) HRNetv2

(d) Segformer
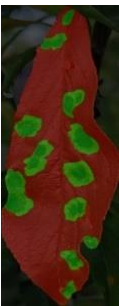
(e) U-Net

(f) MOC_UNet

Figure 12. Regression of predicted lesion coverage with true values for different models: (**a**) PSPNet; (**b**) DeepLabV3+; (**c**) HRNetv2; (**d**) Segformer; (**e**) U-Net; (**f**) MOC_UNet.

3.3. The Results of the Disease Severity Assessment

To better demonstrate the process of disease severity assessment, we present three plum leaves with varying severity of red spot disease as Table 5.

Table 5. Example of disease severity assessment.

Original Image	Segmentation Image	Labels	Value	Ratio	Disease Ratio
		background	418,965	41.14%	
		Plum Leaf	587,989	57.74%	1.89%
		Plum red spot	11,327	1.11%	
		background	476,665	38.78%	
		Plum Leaf	683,744	55.63%	9.14%
		Plum red spot	68,793	5.60%	
		background	322,539	36.93%	
		Plum Leaf	432,966	49.57%	21.40%
		Plum red spot	117,874	13.50%	

As Table 5 shows, disease coverage can clearly reflect the degree of disease proliferation and provide an intuitive indicator for assessing the severity of the disease, which can help growers realize precise disease control.

4. Discussion

Leaf disease detection, as one of the common methods for crop disease detection, can provide an important basis for early disease control. Traditional leaf disease detection mainly relies on the visual detection of relevant experts, with low accuracy and efficiency, and there is strong subjectivity, and it is difficult to rely on manual visual judgement to accurately and quantitatively assess the severity of the disease. Deep learning methods have been widely used for leaf disease detection, but the following shortcomings exist: (1) Most leaf disease detection algorithms only detect and localise the disease on the leaf [20] and cannot further assess the disease severity. (2) Most algorithms based on leaf disease segmentation are limited to simple backgrounds [21] or the case of a single leaf [22], ignoring the interference of complex backgrounds and multiple leaves in realistic scenes. The above limitations lead to the fact that most of the leaf disease detection algorithms have difficulty meeting the application of leaf disease detection in realistic scenes.

In this study, we explored a two-stage detection algorithm for red spot disease on plum leaves, addressing the shortcomings of research on leaf disease detection. We used the object detection model YOLOv8 to strip leaves, which can eliminate the interference of complex backgrounds. And we used the improved disease segmentation model MOC_UNet to

segment the leaf and the disease spots accurately. Finally, by calculating the spot coverage, we made a preliminary determination of the disease severity. Compared with the existing deep learning methods, the two-stage detection algorithm proposed in this study well overcomes the influence of a complex background and multiple leaves and is more suitable for real-life scenarios, and the algorithm for leaf spot segmentation is able to make a preliminary determination of disease severity, which can effectively help farmers to better and accurately grasp the severity of the disease and realise precise prevention and control of the disease.

The two-stage algorithm proposed in this study can be similarly extended to the detection and severity assessment of other leaf diseases. It also provides an effective method for related researchers to calculate the area of disease spots and diseased leaves. However, this study currently focuses only on the red spot disease of plum leaves, which still has limitations, such as less disease coverage and slower identification. We will continue to extend it to other plum diseases, such as plum black mold disease and bacterial leaf spot of plum, at a later stage. In addition, although the algorithm proposed in this study has high accuracy, it is difficult to meet the requirements for the real-time detection of plum leaf diseases due to its slow detection speed. We will also improve the detection speed and reduce the complexity of the network through network pruning in the later stage, so as to make it more suitable for practical agricultural production.

5. Conclusions

In order to detect red spot disease and assess the disease severity of plum leaves under a complex background, we proposed a two-stage recognization method. We used the YOLOv8 to strip plum leaves from complex backgrounds and then used the improved disease segmentation model MOC_UNet to segment leaves and red spots accurately. Finally, spot coverage was calculated to assess the severity of red spot disease. We combined Focal Loss with Dice Loss to eliminate the influence of the imbalance samples. ODConv was used to enhance the model's capability to extract leaf and spot features of various sizes and shapes, and MSCA was introduced to make the model better utilize the multi-scale feature information and enhance the model's ability to segment the boundary of the target area. The accuracy rate of leaf recognition is 95.3%, and the mIoU, mPA, mPrecision, and mRecall of the improved model MOC_UNet reached 90.93%, 95.21%, 95.17%, and 95.21%, respectively, which were improved by 1.2%, 0.86%, 0.55%, and 0.86%, respectively, compared to the original model. We also used regression analysis to compare the relationship between the lesion coverage predicted by different segmentation models and the true measurements. The coefficient of determination R^2 and the mean absolute percentage error MAPE of our proposed MOC_UNet were 0.96 and 8.76%, respectively, which had a larger R^2 and smaller MAPE compared to the other segmentation models, indicating that the predicted values of MOC_UNet for the disease spot coverage were more close to the real values, and that it could assess the severity of red spot disease on plum leaves more accurately. In summary, this paper proposed a two-stage method for high-precision detection and severity assessment of red spot disease on plum leaves. It can help growers to detect diseases early and help achieve precise prevention and control of diseases in actual agricultural production.

Author Contributions: Conceptualization, C.Y. and Z.Y.; methodology, C.Y. and P.L.; software, C.Y. and Y.L.; validation, Z.Y., Y.L. and P.L.; formal analysis, C.Y. and Z.Y.; investigation, C.Y.; resources, C.J., Y.F. and C.Y.; data curation, J.M., C.Y. and J.L.; writing—original draft preparation, C.Y., J.L. and Y.F.; writing—review and editing, P.L. and C.J.; visualization, C.Y.; supervision, J.L.; project administration, J.M.; funding acquisition, J.M. All authors have read and agreed to the published version of the manuscript.

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

References

1.   Li, L.; Zhang, S.; Wang, B. Plant disease detection and classification by deep learning—A Review. *IEEE Access* **2021**, *9*, 56683–56698. [CrossRef]
2.   Wangchu, L.; Angami, T.; Mandal, D. Plum. In *Temperate Fruits*; Apple Academic Press: Palm Bay, FL, USA, 2021; pp. 297–331.
3.   Seethapathy, P.; Gothandaraman, R.; Gurudevan, T.; Malik, I.A. Diseases, Pests, and Disorders in Plum: Diagnosis and Management. In *Handbook of Plum Fruit*; CRC Press: Boca Raton, FL, USA, 2022; pp. 133–176.
4.   Pennazio, S.; Roggero, P.; Conti, M. Yield losses in virus-infected crops. *Arch. Phytopathol. Plant Prot.* **1996**, *30*, 283–296. [CrossRef]
5.   Garcia, J.A.; Cambra, M. Plum pox virus and sharka disease. *Plant viruses* **2007**, *1*, 69–79.
6.   Neagu Frăsin, L. Integrated pestand disease management in sweet cherry and plum orchards. *Ann. Food Sci. Technol.* **2021**, *22*, 430.
7.   Jain, A.; Sarsaiya, S.; Wu, Q.; Lu, Y.; Shi, J. A review of plant leaf fungal diseases and its environment speciation. *Bioengineered* **2019**, *10*, 409–424. [CrossRef]
8.   Waggoner, P.E.; Green, J.S.A.; Smith, F.B. The aerial dispersal of the pathogens of plant disease. *Philos. Trans. R. Soc. London. B Biol. Sci.* **1983**, *302*, 451–462. [CrossRef]
9.   Applalanaidu, M.V.; Kumaravelan, G. A review of machine learning approaches in plant leaf disease detection and classification. In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6 February 2021; pp. 716–724.
10.   Labusca, A.V.; Manoliu, A.; Oprica, L. Influence of the attack of the fungus *Polystigma rubrum* (pers.) dc (red leaf spot) on nutritional value of fruits in different plum cultivars. *J. Exp. Mol. Biol.* **2011**, *12*, 139.
11.   Blackman, V.H.; Welsford, E.J. The Development of the Perithecium of *Polystigma rubrum*, DC. *Ann. Bot.* **1912**, *26*, 761–767. [CrossRef]
12.   Guan, H.; Deng, H.; Ma, X.; Zhang, T.; Zhang, Y.; Zhu, T.; Zhou, H.; Gu, Z.; Lu, Y. A corn canopy organs detection method based on improved DBi-YOLOv8 network. *Eur. J. Agron.* **2024**, *154*, 127076. [CrossRef]
13.   Sun, W.; Zhou, R.; Nie, C.; Wang, L.; Sun, J. Farmland segmentation from remote sensing images using deep learning methods. In Proceedings of the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXII, Online, 21–25 September 2020; p. 1152809.
14.   Verma, S.; Kumar, P.; Singh, J.P. A Unified lightweight CNN-based model for disease detection and identification in corn, rice, and wheat. *IETE J. Res.* **2023**, 1–12. [CrossRef]
15.   Rajeena PP, F.; Su, A.; Moustafa, M.A.; Ali, M. Detecting plant disease in corn leaf Using efficientNet architecture—An analytical approach. *Electronics* **2023**, *12*, 1938. [CrossRef]
16.   Tiwari, D.; Ashish, M.; Gangwar, N.; Sharma, A.; Patel, S.; Bhardwaj, S. Potato leaf diseases detection using deep learning. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020; pp. 461–466.
17.   Iqbal, M.A.; Talukder, K.H. Detection of potato disease using image segmentation and machine learning. In Proceedings of the 2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 4–6 August 2020; pp. 43–47.
18.   Yao, J.; Wang, Y.; Xiang, Y.; Yang, J.; Zhu, Y.; Li, X.; Li, S.; Zhang, J.; Gong, G. Two-stage detection algorithm for kiwifruit leaf diseases based on deep learning. *Plants* **2022**, *11*, 768. [CrossRef] [PubMed]
19.   Xiang, Y.; Yao, J.; Yang, Y.; Yao, K.; Wu, C.; Yue, X.; Li, Z.; Ma, M.; Zhang, J.; Gong, G. Real-Time Detection Algorithm for Kiwifruit Canker Based on a Lightweight and Efficient Generative Adversarial Network. *Plants* **2023**, *12*, 3053. [CrossRef] [PubMed]
20.   Ahmed, K.; Shahidi, T.R.; Alam, S.M.I.; Momen, S. Rice leaf disease detection using machine learning techniques. In Proceedings of the 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 24–25 December 2019; pp. 1–5.
21.   Wang, Y.; Wang, H.; Peng, Z. Rice diseases detection and classification using attention based neural network and bayesian optimization. *Expert Syst. Appl.* **2021**, *178*, 114770. [CrossRef]
22.   Xu, Y.; Chen, Q.; Kong, S.; Xing, L.; Wang, Q.; Cong, X.; Zhou, Y. Real-time object detection method of melon leaf diseases under complex background in greenhouse. *J. Real-Time Image Process.* **2022**, *19*, 985–995. [CrossRef]
23.   Shu, H.; Liu, J.; Hua, Y.; Chen, J.; Zhang, S.; Su, M.; Luo, Y.J.M.T. A grape disease identification and severity es-timation system. *Multimed. Tools Appl.* **2023**, *82*, 23655–23672. [CrossRef]
24.   Divyanth, L.; Ahmad, A.; Saraswat, D.J.S.A.T. A two-stage deep-learning based segmentation model for crop disease quantification based on corn field imagery. *Smart Agric. Technol.* **2023**, *3*, 100108. [CrossRef]
25.   Xiao, B.; Nguyen, M.; Yan, W.Q. Fruit ripeness identification using YOLOv8 model. *Multimed. Tools Appl.* **2023**, *83*, 28039–28056. [CrossRef]
26.   Li, X.; Sun, X.; Meng, Y.; Liang, J.; Wu, F.; Li, J. Dice loss for data-imbalanced NLP tasks. *arXiv* **2019**, arXiv:1911.02855.

27. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Cambridge, MA, USA, 20–23 June 1995; pp. 2980–2988.
28. Li, C.; Zhou, A.; Yao, A. Omni-dimensional dynamic convolution. *arXiv* **2022**, arXiv:2209.07947.
29. Guo, M.-H.; Lu, C.-Z.; Hou, Q.; Liu, Z.; Cheng, M.-M.; Hu, S.-M. Segnext: Rethinking convolutional attention design for semantic segmentation. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 1140–1156.
30. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of yolo algorithm developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [CrossRef]
31. Yao, J.; Qi, J.; Zhang, J.; Shao, H.; Yang, J.; Li, X. A real-time detection algorithm for kiwifruit defects based on YOLOv5. *Electronics* **2021**, *10*, 1711. [CrossRef]
32. Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. Yolox: Exceeding yolo series in 2021. *arXiv* **2021**, arXiv:2107.08430.
33. Lu, Y.; Chen, Y.; Zhao, D.; Chen, J. Graph-FCN for image semantic segmentation. In Proceedings of the International Symposium on Neural Networks, Moscow, Russia, 10–12 July 2019; pp. 97–105.
34. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
35. Zhu, X.; Cheng, Z.; Wang, S.; Chen, X.; Lu, G. Coronary angiography image segmentation based on PSPNet. *Comput. Methods Programs Biomed.* **2020**, *200*, 105897. [CrossRef] [PubMed]
36. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A. Deeplab: Semantic image seg-mentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [CrossRef]
37. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; pp. 234–241.
38. Zhang, Y.; Zhang, J.; Wang, Q.; Zhong, Z. Dynet: Dynamic convolution for accelerating convolutional neural networks. *arXiv* **2020**, arXiv:2004.10694.
39. Ma, J.; Zhang, Z.; Xiao, W.; Zhang, X.; Xiao, S.J.I.A. Flame and smoke detection algorithm based on ODConvBS-YOLOv5s. *IEEE Access* **2023**, *11*, 34005–34014. [CrossRef]
40. Yang, B.; Bender, G.; Le, Q.V.; Ngiam, J. Condconv: Conditionally parameterized convolutions for efficient inference. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. [CrossRef]
41. Chen, Y.; Dai, X.; Liu, M.; Chen, D.; Yuan, L.; Liu, Z. Dynamic convolution: Attention over convolution kernels. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 11030–11039.
42. Qian, L.; Qian, W.; Tian, D.; Zhu, Y.; Zhao, H.; Yao, Y. MSCA-UNet: Multi-Scale Convolutional Attention UNet for Automatic Cell Counting Using Density Regression. *IEEE Access* **2023**, *11*, 85990–86001. [CrossRef]
43. Wang, C.; Du, P.; Wu, H.; Li, J.; Zhao, C.; Zhu, H. A cucumber leaf disease severity classification method based on the fusion of DeepLabV3+ and U-Net. *Comput. Electron. Agric.* **2021**, *189*, 106373. [CrossRef]