

An improved DSCCA-UNet for apple leaf disease severity estimation and prescription map generation

Shixin Huo

Jiangsu Normal University

Na Duan

duanna08@163.com

Jiangsu Normal University

Zhizheng Xu

Jiangsu Normal University

Hui Feng

Jiangsu Normal University

Research Article

Keywords: Spots segmentation, Prescription map, Apple leaf diseases, DSCCA-UNet

Posted Date: August 7th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6911991/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

An improved DSCCA-UNet for apple leaf disease severity estimation and prescription map generation

Shixin Huo, Na Duan *, Zhizheng Xu, Hui Feng

School of Electrical Engineering & Automation, Jiangsu Normal University, Xuzhou, 221116, China.

Abstract

This paper investigates the problems of disease severity estimation and prescription map generation. By introducing dynamic snake convolution (DSC), a DSCC multi-scale feature extraction module is designed to build a new UNet architecture. The combination of DSCC module with VGG16 backbone can enhance the receptive field for segmented edge feature information which can achieve more detailed edge feature fusion. The channel attention (CA) and spatial attention (SA) in the CBAM attention module are disassembled and used in the skip connection part and the upsampling part, respectively. This new connection can obtain more location information of the spots and mitigate the effect of background on network learning. Moreover, an automatic pixel counting algorithm based on the improved DSCCA-UNet is designed to estimate the disease severity. Finally, the system of the apple leaf disease severity estimation and the variable prescription maps are obtained based on the PyQt5 tool and ArcGIS component. The experimental results show that the improved DSCCA-UNet model outperforms other mainstream semantic segmentation models. It can more effectively complete the tasks of disease severity estimation and prescription map generation in actual orchard scenarios.

Keywords: Spots segmentation, Prescription map, Apple leaf diseases, DSCCA-UNet

1 Introduction

Disease problems can affect apple yield and quality [1–3]. The accurate estimation of disease severity and the generation of prescription maps are essential for the reduction

of economic loss. The conventional method for estimating disease severity relies on expert experience. It employs the human eye to perceive the severity of disease symptoms within the visible range, by combining with a scale and a standard area map to judge disease severity [4, 5]. However, in a complex field environment, the reliability of visual estimation is low due to the heterogeneity of different diseases and the subjectivity of visual estimation.

Disease severity estimation based on traditional image processing methods is more accurate than visual estimation methods. The image quality is improved by pre-processing operations, such as grey scale conversion, noise removal and image enhancement. The shape, colour, texture and other features of the spots are then extracted manually and combined with threshold segmentation [6–8], area growth [9, 10] and edge detection [11] to obtain the final segmentation results. The ratio between the area of diseased spots and the area of healthy leaves are calculated to obtain the disease severity. However, background noise and lighting conditions are often present in real field environments. Traditional image processing methods rely on manually extracted features [12]. These low-level features are difficult to be adapted to the diverse and complex problem of disease severity estimation in complex field environments.

Recently, deep convolutional neural networks has been gradually applied to the task of image semantic segmentation for performing pixel-level classification [13, 14]. It is more efficient and accurate to use deep learning semantic segmentation method to get the pixel ratio of diseased spots and healthy leaves, then get the severity of disease. The mainstream semantic segmentation networks include UNet, HRNet, PSPNet, DeepLabv3+, Segformer and UNet++ [15–21]. These semantic segmentation networks have been widely applied in spots semantic segmentation tasks. In [22], for persimmon leaf spot and grey mould diseases, Jia et al. used deformable convolution in the UNet downsampling stage to accommodate the geometrical deformation of the diseases. The self-attention mechanism was used to obtain global spatial and contextual information to improve segmentation accuracy. For cucumber downy mildew and powdery mildew, Li et al. [23] proposed a semantic segmentation model based on the DeepLabv3+. In this model, hybrid attention and transfer learning were used to improve the segmentation accuracy. Meanwhile, Wang et al. proposed a two-stage model using DeepLabv3+ and UNet, in [24]. In the first stage of the model, the DeepLabv3+ model was used to segment and extract whole cucumber leaves. Then the results of previous stage were used to further segment and extract the diseased spots, thus achieving segmentation of diseased leaves and spots in a complex background. However, the two-stage semantic segmentation model consumed more computational resources and reduced efficient.

In [25], Lu et al. proposed the MixSeg model, which contains both encoder and decoder parts. In the encoder, the Mix Transformer, the position indexer and the local optimization module were designed to obtain global information and optimize the segmentation effect on local edges. In the decoder, shallow and deep features were integrated to facilitate the generation of accurate segmentation results. This model was tested in the complex environments. However, diseases are complex and diverse in real orchard environments. It is necessary to explore a more accurate and suitable segmentation method for more kinds of apple leaf disease in disease severity estimation.

Prescription map is a crucial tool for implementing variable-rate application (VRA) in precision agriculture [26–28]. It indicates the geographical coordinates and dosages needed for agrochemicals at specific locations throughout the field. This map is then transferred to the spraying unmanned aerial vehicles (UAV) or other machinery, guiding them to deliver precise treatments accordingly [29]. In the field of prescription map generation, the traditional generation techniques are achieved mainly through remote sensing techniques. In [30], Chen et al. generated prescription maps for rice false smut disease in complex environments based on near-earth remote sensing. Nowadays, it is possible to use deep neural networks to obtain spots classification results and thus generate prescription maps. In [31], Huang et al. captured high resolution UAV imagery over a rice field. For weed mapping and prescription map generation, fully convolutional networks (FCN) is applied for a pixel-level classification. The image-based prescription map generation method offers a cost-effective alternative by reducing dependence on remote sensing technologies. Therefore, it is significant to further explore how to generate prescription maps by leveraging image-related information.

This paper investigates an improved DSCCA-UNet model and an automatic pixel counting algorithm to increase the accuracy and efficiency of apple leaf disease severity estimation. Then a severity estimation system and a prescription map of apple leaf alternaria spot disease are generated by combining this method. The main contributions can be summarized as follows:

- (1) A DSCC multi-scale feature extraction module is designed and constructed by introducing dynamic snake convolution and convolutional block attention module (CBAM). The DSCC multi-scale feature extraction module is introduced in the middle of the last layer of encoder and decoder. The channel attention, spatial attention in the CBAM attention module are disassembled and used in the skip connection part and the upsampling part, respectively. It enhances the feature extraction and fusion capability of the model for the edge feature information of various scales.
- (2) An automatic pixel counting algorithm based on DSCCA-UNet is designed to estimate the disease severity. Subsequently, an interface for apple leaf disease severity estimation system is developed. The system can realise functions related to disease segmentation and disease severity estimation.
- (3) Combining the apple leaf disease severity estimation system and the ArcGIS tool, a vision-based method for generating prescription maps of apple leaf disease is proposed. For actual captured apple leaf images, the system is used for disease severity estimation to obtain prescription values. By introducing ArcGIS tools, the prescription maps are generated to visualise disease infection. It is more convenient for orchard management.

The remainder of this paper is as follows. Section 2 covers evaluation indicators and the disease severity grading criteria. Section 3 proposes the improved DSCCA-UNet model, the apple leaf disease severity estimation method and the prescription map generation method. In section 4, the apple leaf diseases dataset, the experimental setup and the results and analysis of the experiments are presented. The prescription map of alternaria spot disease is given in section 5. Finally, conclusions are provided.

2 Evaluation indicators and disease severity grade

This section introduces evaluation indicators and some preliminary knowledge. In the part of apple leaf spot segmentation, the evaluation metrics used for comparative experiments include mean intersection over union (MIoU), mean pixel accuracy (MPA), mean precision (Mprecision), F1 and parameters [32]. The formulas are

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{i=0}^k p_{ii} + \sum_{i=0}^k p_{ji} - p_{ii}} \quad (1)$$

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{i=0}^k p_{ij}}, MPrecision = \frac{\sum_1^k \frac{TP}{TP+FP}}{k} \quad (2)$$

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (3)$$

where P_{ij} represents the number of pixels that actually belong to class i but are predicted as class j , P_{ji} represents the number of pixels that actually belong to class j but are predicted as class i , and P_{ii} is the number of true positive pixels, the term $k+1$ refers to the number of segmentation classes, including one background class, TP denotes true positives, and FP denotes false positives.

For disease severity estimation, the evaluation metrics are root mean squared error (RMSE) and the R-Square (R^2) coefficient of determination. The formulas are

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (4)$$

where y_i represents the true value, $\hat{y}_i$ represents the estimated value, and $\bar{y}_i$ represents the mean of the predicted values. The R^2 value ranges from $[0, 1]$, where a higher R^2 generally indicates a better model fit.

Disease severity classification plays a crucial role in grading the severity of plant diseases and in plant disease management. Typically, disease grade is performed based on the proportion of the leaf area affected by lesions [33, 34]. The formula for calculating disease severity is

$$S = \frac{P_{lesion}}{P_{lesion} + P_{leaf}} \times 100 \quad (5)$$

where P_{lesion} refers to the number of pixels in the lesion area within the segmentation result, and P_{leaf} refers to the number of pixels in the healthy leaf area within the segmentation result.

The proportion of lesion area pixels is obtained by an automatic pixel counting algorithm. Using the Chinese Ministry of Agriculture's NY/T992-2006 standard, Quality of Work for Orchard Air-Blast Sprayers, the final disease severity is determined. The specific calculation criteria is shown in Table 1, the disease severity grade contains six classes. This grade can be derived by calculating the ratio of diseased area to total leaf area.

Table 1 The criteria of disease severity grade

Disease severity grade		Criteria
0		No disease
1		Lesion area occupies less than 10% of the total leaf area
3		Lesion area occupies 11%-25% of the total leaf area
5		Lesion area occupies 26%-40% of the total leaf area
7		Lesion area occupies 41%-65% of the total leaf area
9		Lesion area occupies more than 65% of the total leaf area

3 An improved DSCCA-UNet for disease severity estimation

In this section, the improved DSCCA-UNet model is proposed. The DSCC multi-scale feature extraction module is designed and introduced in the middle of the last layer of encoder and decoder. It can achieve more detailed edge feature fusion for apple leaf disease, which can further improve the accuracy of disease severity estimation. Meanwhile, the CBAM's channel attention and spatial attention components are separated, with channel attention applied to the UNet skip connections and spatial attention utilized in the upsampling stage. This new connection can obtain more location information of the spots and mitigate the effect of background on network learning. Finally, the apple leaf disease severity estimation method, and the prescription map generation method are proposed to construct the disease severity estimation system.

3.1 DSCC multi-scale feature extraction module

As is well known, edges occupy a relatively small proportion of pixels within an image, posing challenges for conventional convolution operations to effectively extract their features. To address this problem, dynamic snake convolution is incorporated to improve the network's ability to recognize the curved and elongated tubular segmentation edges of apple leaf spots. Based on the previous convolution position, dynamic snake convolution can adjust the direction of convolution motion and capture richer feature information than conventional convolution [35]. In contrast to deformable convolution, dynamic snake convolution pairs impose specific constraints on the process of adjusting kernel sizes to align with the shape of the object being detected. For multiple perspectives and facilitate efficient multi-view feature fusion, this approach can complement key features.

Figure 1 (a) illustrates a 3×3 conventional convolution. For 2D convolution, a 3×3 convolution kernel can be denoted as

$$K = \{(x-1, y-1), (x-1, y), \dots, (x+1, y+1)\} \quad (6)$$

where centre coordinates is (x, y) . For a dynamic snake convolution, define its centre coordinates as $K_i = (x_i, y_i)$. By introducing an offset Δ in the x-axis and y-axis directions, constraints are imposed on the convolution kernel, where $\Delta = \{\delta | \delta \in [-1, 1]\}$. Figure 1 (b) illustrates a 9×9 dynamic snake kernel, the specific position of each lattice in the convolution kernel K can be expressed as $K_{i \pm c} = (x_{i \pm c}, y_{i \pm c})$, with

$c = \{0, 1, 2, 3, 4\}$. The selection of each grid K_{i+c} is a cumulative process, starting from the central position K_i , with the distance from the central grid depending on the position of the previous grid. To ensure that the convolution kernel conforms to a linear morphological structure, the introduced offsets need to be accumulated. Using dynamic snake convolution leads to

$$K_{i\pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \sum_{i-c}^{i+c} \Delta y) \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \sum_{i-c}^i \Delta y) \end{cases} \quad (7)$$

$$K_{i\pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = (x_j + c, y_j + \sum_{j-c}^{j+c} \Delta y) \\ (x_{j-c}, y_{j-c}) = (x_j - c, y_j + \sum_{j-c}^j \Delta y) \end{cases} \quad (8)$$

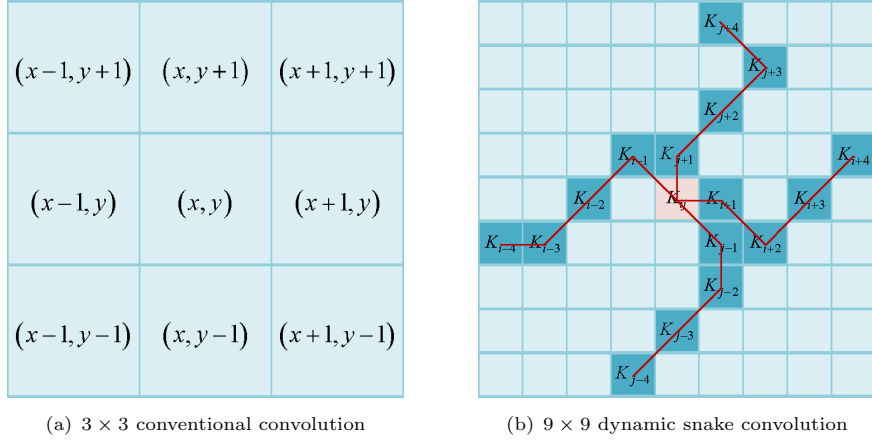


Fig. 1 Example diagram of two convolutions

The DSCC multi-scale feature extraction module is constructed to further enhance the receptive field for segmented edge feature information and achieve more detailed edge feature fusion. This module can enrich the representation of curved edge features across different scales. The module architecture draws inspiration from the atrous spatial pyramid pooling (ASPP) in [36, 37].

Figure 2 provides a detailed illustration of the specific components of the DSCC multi-scale feature extraction module. In the DSCC module, the input feature map is processed through three dynamic snake convolutions and a 1×1 standard convolution to extract multi-scale and multi-view feature information. The convolution kernel sizes for the three dynamic snake convolutions are 1, 3, and 5. The extracted features are then concatenated and passed through attention mechanism to further enhance the representation of spot feature information and suppress background interference. The feature map undergoes channel adjustment to produce the output feature map. The DSCC multi-scale feature extraction module is able to capture curved segmentation edge features at various scales by incorporating dynamic snake convolutions of different sizes. It can also effectively improve the segmentation accuracy of apple leaf spots.

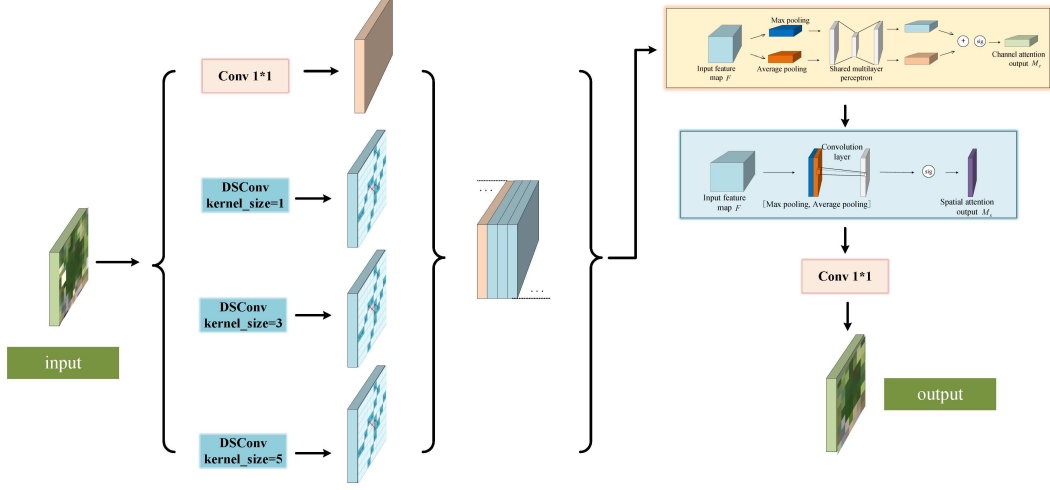


Fig. 2 DSCC multi-scale feature extraction module

3.2 The improved DSCCA-UNet model

The baseline UNet model uses a self-encoder structure, where the encoder extracts the image features and the decoder classifies each pixel and reconstructs the segmented output [38]. In addition, the skip connection between the encoder and decoder can effectively prevent the model from losing information during downsampling. Incorporating attention mechanism into the skip connection component enhances the network's ability to focus on relevant and beneficial information, thereby contributing to improved overall performance. The CBAM attention module is a lightweight mechanism composed of channel attention and spatial attention [39]. These two sub-modules perform attention operations on the channel and spatial dimensions, respectively. They can enhance the representation of information in each domain.

Figure 3 illustrates the details of the attention mechanism. The channel attention module first compresses the input feature map along the spatial dimension, generating two one-dimensional vectors, or feature descriptors, via max pooling and average pooling. These feature descriptors are then passed through a shared network consisting of a multilayer perceptron. The outputs of the shared network are summed and merged. Finally, after applying a sigmoid activation function, the channel attention feature map is obtained. This process can be expressed as

$$\begin{aligned}
 M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\
 &= \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c)))
 \end{aligned} \tag{9}$$

where F represents the input feature map for channel attention, F_{avg}^c and F_{max}^c denote the average pooling and max pooling features across the spatial dimensions, respectively. W_1 and W_0 are the weight coefficients, and σ denotes the sigmoid activation function.

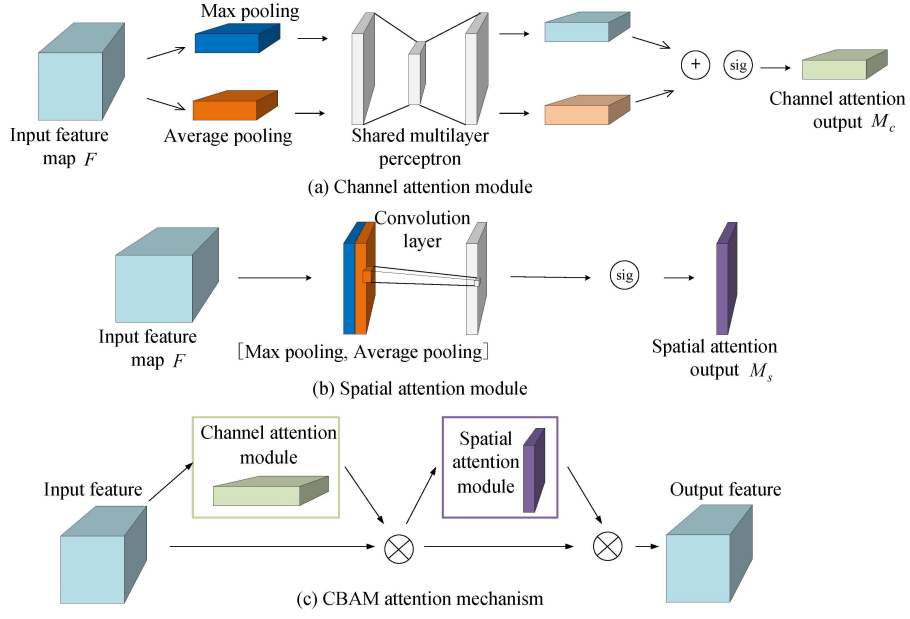


Fig. 3 Attention mechanism

For spatial attention, the input is first processed by max pooling and average pooling along the channel dimension. The resulting feature maps are then concatenated, followed by a convolution operation on the concatenated map. Finally, the spatial attention feature map M_s is obtained by applying the sigmoid activation function, which can be expressed as

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) = \sigma(f^{7 \times 7}(F_{avg}^s; F_{max}^s)) \quad (10)$$

where F represents the input feature map for spatial attention, while F_{avg}^s and F_{max}^s denote the average pooling and max pooling features along the channel dimension, respectively. The sigmoid activation function is denoted by σ .

To construct the improved DSCCA-UNet model, channel attention (CA) is applied in the skip connection stage, while the spatial attention (SA) is used in the upsampling stages of the model. This combination aims to restore more detailed feature information related to the spots, further improving segmentation accuracy. Figure 4 illustrates the specific details of the model. In addition to the attention mechanism, the VGG16 are utilized as the backbone network for feature extraction in the improved DSCCA-UNet model. The DSCC multi-scale feature extraction module is integrated between the final stages of the encoder and decoder to enrich multi-scale feature representation, mitigating the loss of fine-grained details.

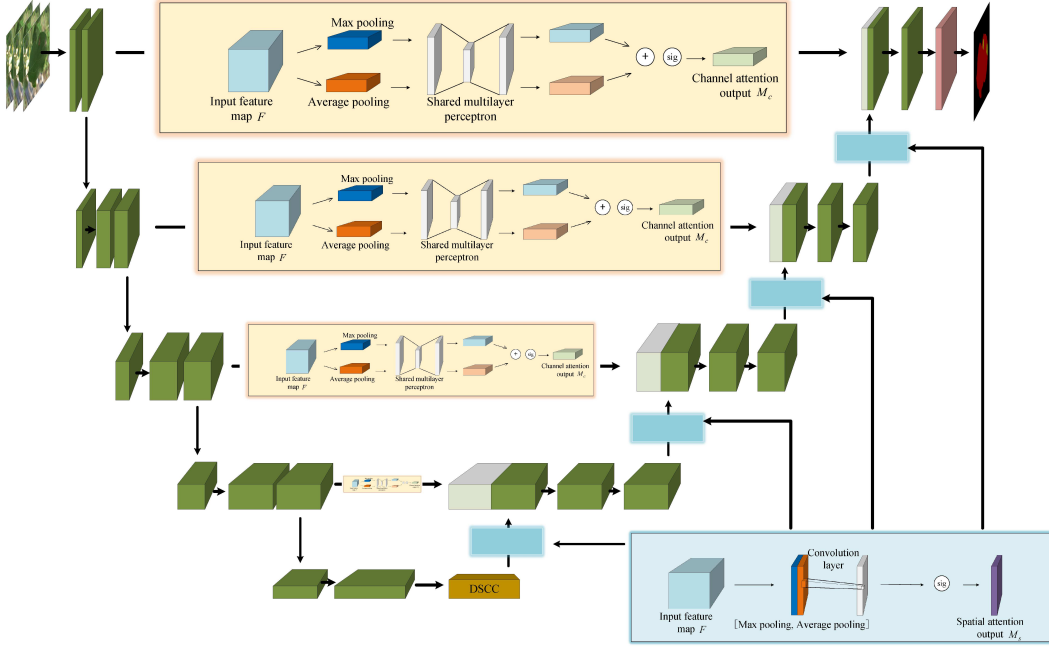


Fig. 4 DSCCA-UNet network model

3.3 Apple leaf disease severity estimation and prescription map generation

The process of apple leaf disease severity estimation is shown in Figure 5. This process has three parts. The required dataset is established, including the acquisition of images and data enhancement of the obtained images. The established apple leaf disease dataset is divided into training, validation and test sets in the ratio of 8:1:1. The training set is used for model training. The function of validation set is to evaluate the hyperparameters and preliminary estimation of performance. The test set serves to evaluate the model performance. Finally, the obtained model is used for apple leaf spot segmentation to obtain the final segmentation results. The pixel points of diseased spots and healthy leaves are counted separately to obtain the estimation results of disease severity.

For disease severity estimation, manual counting of pixel points in the predicted segmentation map can be time-consuming and inefficient. An automatic pixel counting algorithm is proposed to address this problem [40]. As described in Algorithm 1, the predicted images are first transformed into grayscale maps to facilitate the computation of the total number of pixels, wherein each pixel is subsequently categorized according to its corresponding class. The dataset includes six pixel classes, background, healthy leaf, and four disease types—alternaria leaf spot, grey spot, brown spot, and rust. During the counting process, a category flag variable i is set. When $i = 0, 1, 2, 3, 4, 5$, the pixels are counted for the background, healthy leaf, rust,

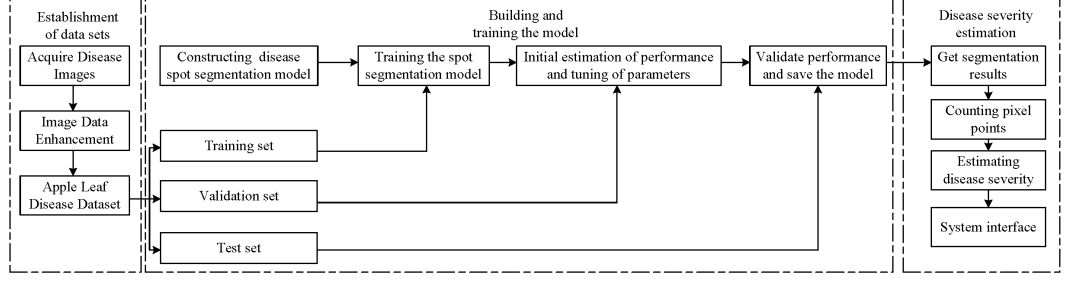


Fig. 5 The process of apple leaf disease severity estimation

Algorithm 1 Count Pixels for Leaf Pixels

Input:

- *image_path*: Path to the grayscale image to be processed
- *num_classes*: Number of disease classes, e.g., 6 (background, healthy leaf, rust, alternaria, grey, brown)
- *input_shape*: Target input size for the model
- *model*: Trained deep learning model for segmentation

Output:

- *class_counts*: Number of pixels for each class

Initialization:

- 1: *class_counts* $\leftarrow [0] \times \text{num_classes}$
- 2: *name_classes* $\leftarrow [\text{Background, Healthy leaf, Rust, Alternaria, Grey, Brown}]$
- 3: *image* $\leftarrow \text{read_image}(\text{image_path})$
- 4: *original_h, original_w* $\leftarrow$ dimensions of *image*

Preprocess Image:

- 1: *image* $\leftarrow \text{resize}(\text{image}, \text{input_shape})$
- 2: *image_data* $\leftarrow \text{preprocess}(\text{image})$
- 3: *image_data* $\leftarrow \text{expand_dims}(\text{image_data})$

Predict Pixel Classes:

- 1: *output* $\leftarrow \text{model}(\text{image_data})$
- 2: *output* $\leftarrow \text{softmax}(\text{output})$
- 3: *output* $\leftarrow \text{resize_to_original}(\text{output}, \text{original_h}, \text{original_w})$
- 4: *pixel_classes* $\leftarrow \text{argmax}(\text{output}, \text{axis} = -1)$

Count Pixels for Each Class:

- 1: *total_pixels* $\leftarrow \text{original_h} \times \text{original_w}$
- 2: **for** *i* $\in [0, \text{num_classes})$ **do**
- 3: *class_counts*[*i*] $\leftarrow \text{sum}(\text{pixel_classes} == i)$
- 4: **end for**

return *class_counts*

alternaria spot, grey spot, and brown spot categories, respectively. The percentage of diseased pixels relative to the total number of leaf pixels is then calculated for each category. By Table 1 and Algorithm 2, the final disease severity grades can be derived. These results can inform the development of the disease classification estimation system and the generation of prescription maps.

Algorithm 2 Disease Severity Estimation for Leaf Pixels

Input:

- *class_counts*: Number of pixels for each class (calculated in Algorithm 1)
- *num_classes*: Number of disease classes
- *name_classes*: Names of the disease classes

Output:

- *disease_labels*: Disease severity labels for each type

Calculate Disease Severity:

```

1: healthy_count  $\leftarrow$  class_counts[1]
2: for  $i \in [2, \text{num\_classes})$  do
3:   disease_count  $\leftarrow$  class_counts[ $i$ ]
4:   if disease_count == 0 then
5:     Append "No Disease" to disease_labels
6:   else
7:     severity_ratio  $\leftarrow \frac{\text{disease\_count}}{\text{healthy\_count} + \text{disease\_count}}$ 
8:     if  $0 < \text{severity\_ratio} \leq 0.1$  then
9:       Grade 1
10:    else if  $0.1 < \text{severity\_ratio} \leq 0.25$  then
11:      Grade 3
12:    else if  $0.25 < \text{severity\_ratio} \leq 0.4$  then
13:      Grade 5
14:    else if  $0.4 < \text{severity\_ratio} \leq 0.65$  then
15:      Grade 7
16:    else
17:      Grade 9
18:    end if
19:    Append corresponding severity label to disease_labels
20:  end if
21: end for
return disease_labels

```

Figure 6 illustrates the process of prescription map generation. The first step involves locating the operational plot and saving its latitude and longitude information. This information is then imported into ArcMap software to generate a grid of surface layers, which are rasterized to create the prescription map. The grid is subsequently converted into editable and readable map layers. The collected images are analyzed

for disease severity, and the estimation results are integrated into the surface map layer to complete the prescription map generation. This generation method reduces reliance on remote sensing technology. It also enhances the practical applicability of disease severity estimation results.

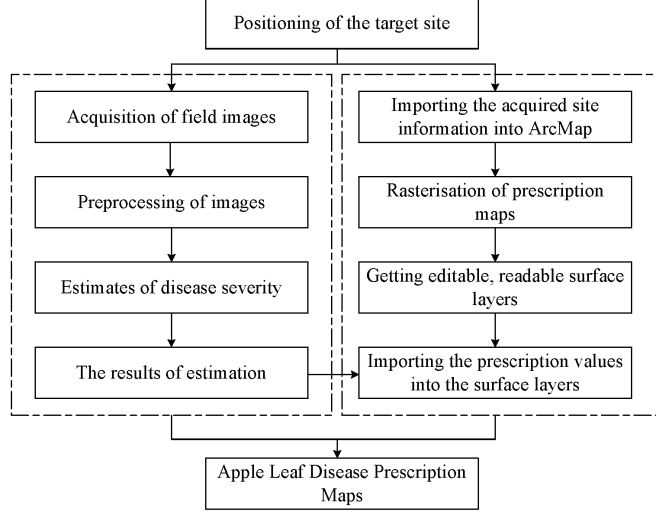


Fig. 6 The process of prescription map generation

4 Experimental results and analysis

In this section, dataset and data augmentation method are introduced. Ablation experiments and comparative experiments are conducted to demonstrate the advantages of the proposed method. The experimental results are also visualized and analyzed.

4.1 Apple leaf diseases dataset and experimental setup

The images for train are sourced from the publicly available ATLDS dataset, which contains images of apple leaves with varying degrees of disease. The data are acquired under varying weather conditions and at different time, encompassing both controlled laboratory environments and real-world cultivation field settings [41]. Approximately 48.1% of the images are taken in real cultivation fields. The dataset includes a total of 1,641 apple leaf images, consisting of 409 healthy leaves, 278 with spotted leaf spot, 395 with gray spot, 215 with brown spot, and 344 with rust. Figure 7 presents examples of images alongside their corresponding labels for each disease.

To enhance data diversity and prevent overfitting, the acquired dataset for each type of disease is augmented through random image transformations. The data augmentation strategy included image rotation (80% probability, $\pm 10^\circ$), horizontal flipping (50% probability), and scaling (30% probability) to 85% of the original size.

These operations increased the number of images to approximately 10 times the original amount, with corresponding label transformations applied to produce the final dataset.

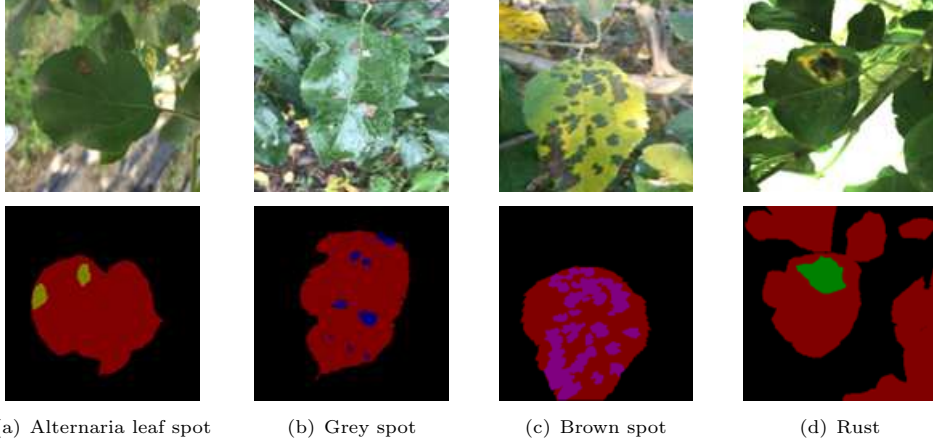


Fig. 7 Example of apple leaf disease images and labels

The apple leaf spot segmentation network model in this paper is developed using Pytorch 2.0.1 on a Windows 10 operating system. The hardware environment for network training included a 12th Gen Intel(R) Core(TM) i7-12700KF CPU with a 3.61GHz clock speed, 32GB of memory, and an NVIDIA GeForce RTX 3070 GPU, with GPU acceleration provided by CUDA 11.7. The network training parameters contain a batch size of 4, a learning rate of 0.0001, a momentum factor of 0.9, and the model is iterated 70 times. The final training weights are obtained and saved once the model’s loss function and mean intersection over union (MIoU) values stabilized. To validate the accuracy of the proposed network model for apple leaf spot segmentation, ablation experiments are conducted. The model is also compared with state-of-the-art semantic segmentation models.

4.2 Ablation experiments

Ablation experiments are conducted on the apple leaf disease dataset. It can verify the effectiveness of the proposed DSCC multi-scale feature extraction module and attention mechanism within the model. During model training, the UNet model is used as the baseline, and three optimized models are created by sequentially incorporating VGG16 as the backbone feature extraction network, the DSCC multi-scale feature extraction module, and the attention mechanism. Finally, the final DSCCA-UNet (UNet+DSCC+Attention) model is obtained by equipping the baseline UNet model with VGG16, DSCC, and the attention mechanism.

Table 2 shows the overall performance of the model after incorporating each module. Compared to the baseline UNet model, using VGG16 as the backbone feature extraction network improves MIoU, MPA, Mprecision and F1 by 1.29%, 0.66%, 0.95%

and 0.8% respectively. This indicates that the improvement in the backbone feature extraction network enhances the model’s ability to focus on useful information. By integrating the DSCC multi-scale feature extraction module and attention mechanism into the enhanced backbone feature extraction network, the model demonstrates significantly improved performance compared to the baseline UNet architecture. Notably, the combination of the DSCC module with the VGG16 backbone improves MIoU, MPA, Mprecision and F1 by 2.61%, 2.04%, 1.1% and 1.58%, over the baseline UNet model, respectively. It improves MIoU, MPA, Mprecision and F1 by 1.32%, 1.38%, 0.15% and 0.78% compared to using only the VGG16 backbone. Meanwhile, MIoU, Mprecision and F1 are improved by 0.38%, 0.67% and 0.23% respectively after using the DSCC multi-scale feature extraction module than the ASPP module. The DSCC module improves segmentation accuracy at class boundaries. It enables precise differentiation between lesions and leaves as well as between leaves and the background.

Table 2 The results of ablation experiment

Backbone	DSCC module	ASPP module	Attention mechanism	MIoU/%	MPA/%	Mprecision/%	F1/%
✗	✗	✗	✗	87.01	91.93	93.65	92.78
✓	✗	✗	✗	88.30	92.59	94.60	93.58
✓	✓	✗	✗	89.62	93.97	94.75	94.36
✓	✗	✗	✓	89.59	93.50	95.21	94.35
✓	✗	✓	✓	89.38	94.06	94.39	94.22
✓	✓	✗	✓	89.76	93.84	95.06	94.45

In conclusion, the proposed DSCCA-UNet model achieves MIoU, MPA, Mprecision and F1 values of 89.76%, 93.84%, 95.06%, 94.45% respectively, representing improvements of 2.75%, 1.91%, 1.41% and 1.67% over the baseline UNet model. The enhanced DSCCA-UNet model integrates the advantages of the DSCC module and attention mechanisms, leading to further improvements in segmentation accuracy and IoU performance. It contributes to more accurate pixel counting of lesions and leaves in the disease severity estimation task, reducing bias in severity assessments.

4.3 Comparative experiments among different models

In this section, the proposed DSCCA-UNet model is compared with mainstream semantic segmentation algorithms such as UNet, HRNet, PSPNet, DeepLabv3+, Segformer, and UNet++. It can further validate the effectiveness for lesion segmentation. To ensure fairness in the comparison, the same training strategies and environments are used across all models.

Table 3 shows the comparison of performance for different semantic segmentation models on apple leaf disease dataset. For segmentation accuracy, the improved DSCCA-UNet model outperforms other models. The DSCC module can enhance the representation of curved edge features at various scales, improving the model’s ability to classify edge pixels. Furthermore, the integration of the attention mechanism

strengthens the focus on both leaves and lesions, improving segmentation precision for these regions.

Table 3 Performance comparison of different models

Model	MIoU/%	MPA/%	Mprecision/%	F1/%	Parameters/MB
DeepLabv3+	84.20	89.40	92.70	90.94	54.71
HRNet	86.93	91.47	94.01	92.72	29.54
PSPNet	83.94	88.07	93.21	90.57	2.376
UNet	87.01	91.93	93.65	92.78	43.93
Segformer	86.19	94.35	90.40	92.33	3.720
UNet++	83.19	88.48	92.48	90.44	9.163
DSCCA-UNet	89.76	93.84	95.06	94.45	43.20

Compared to DeepLabv3+, HRNet, PSPNet, UNet and UNet++, the improved DSCCA-UNet model achieves a 5.56%, 2.83%, 5.82%, 2.75% and 6.57% increase in MIoU, respectively. For the MPA, it exhibits improvements of 4.44%, 2.37%, 5.77%, 1.91% and 5.36%. For Mprecision, the improvements are 2.36%, 1.05%, 1.85%, 1.41% and 2.58%, respectively. While for F1, it improves 3.51%, 1.73%, 3.88%, 1.67%, 4.01%, respectively. Although the MPA of the improved DSCCA-UNet is 0.51% lower than that of Segformer, it shows increases of 3.57%, 4.66% and 2.12% in MIoU, Mprecision and F1, respectively. It surpasses the baseline UNet model’s performance in these areas.

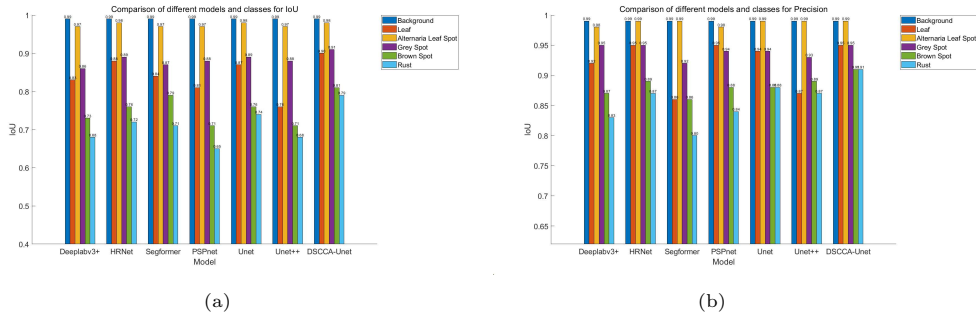


Fig. 8 The segmentation performance of different models across various classes

Figure 8 presents histograms of IoU, PA, and Precision for different semantic segmentation models on six categories, they are leaf, alternaria leaf spot, grey spot, brown spot, rust, and background. All models perform exceptionally well in the background and alternaria leaf spot disease categories. Therefore, the following discussion focuses on other categories. For the leaf category, the improved DSCCA-UNet model achieves an IoU of 90%, surpassing the best-performing HRNet by 2% and outperforming the worst-performing UNet++ by 14%. In the rust category, the improved DSCCA-UNet model reaches an IoU of 79%, 5% higher than the best-performing UNet and 14%

higher than the worst-performing PSPNet. For brown spot, the improved DSCCA-UNet model achieves an IoU of 81%, 2% higher than the best-performing Segformer and 10% higher than both PSPNet and UNet++. In the grey spot category, it achieves an IoU of 91%, at least 2% higher than other models. For Precision, the improved DSCCA-UNet demonstrates the most significant advantage in the rust category, outperforming the best original UNet by 3% and the worst-performing Segformer by 11%.

Figure 9 shows the segmentation results on several test images. It can visually compare the results of the semantic segmentation models. When the leaf exhibits extensive diseased areas with a high density of lesion pixels, the segmentation outputs of DeepLabv3+ and Segformer tend to display merged or confluent lesion regions, indicating a loss of boundary delineation between individual lesions. When diseased areas contain few lesion pixels, deep learning models often fail to distinguish lesions from leaves, noise, or other interferences. Compared to other semantic segmentation models in [13–18], the improved DSCCA-UNet exhibits fewer instances of missed or false detections of lesion pixels, and the issue of lesion region sticking is alleviated. The DSCC module enhances the model’s ability to capture curved and tubular edge features more effectively. The new type of attention mechanism connection can reduce background noise.

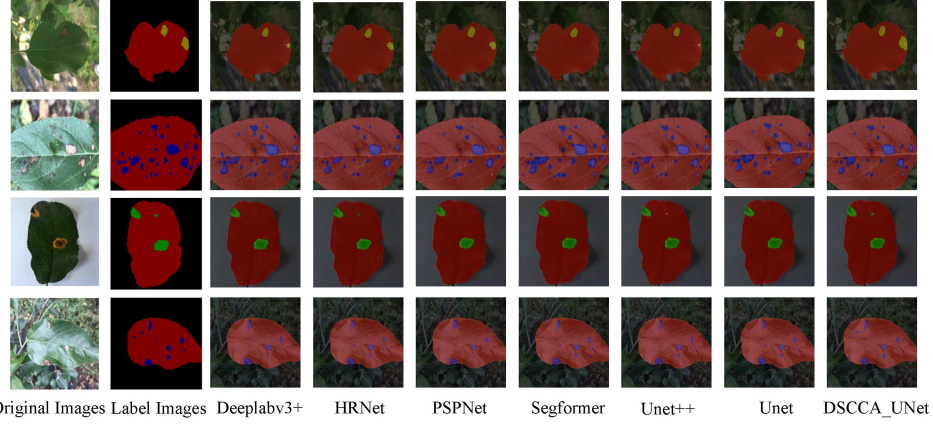


Fig. 9 Semantic segmentation results on the test set

5 The generation of severity estimation system

In this section, we select the apple orchard of Dashahe, Feng County, Jiangsu Province as the study area, the prescription maps for apple alternaria leaf spot diseases is generated. Furthermore, a disease severity estimation system is established based on the improved DSCCA-UNet model and the pixel counting algorithm. It can meet the needs for estimating disease severity and generating prescription maps.

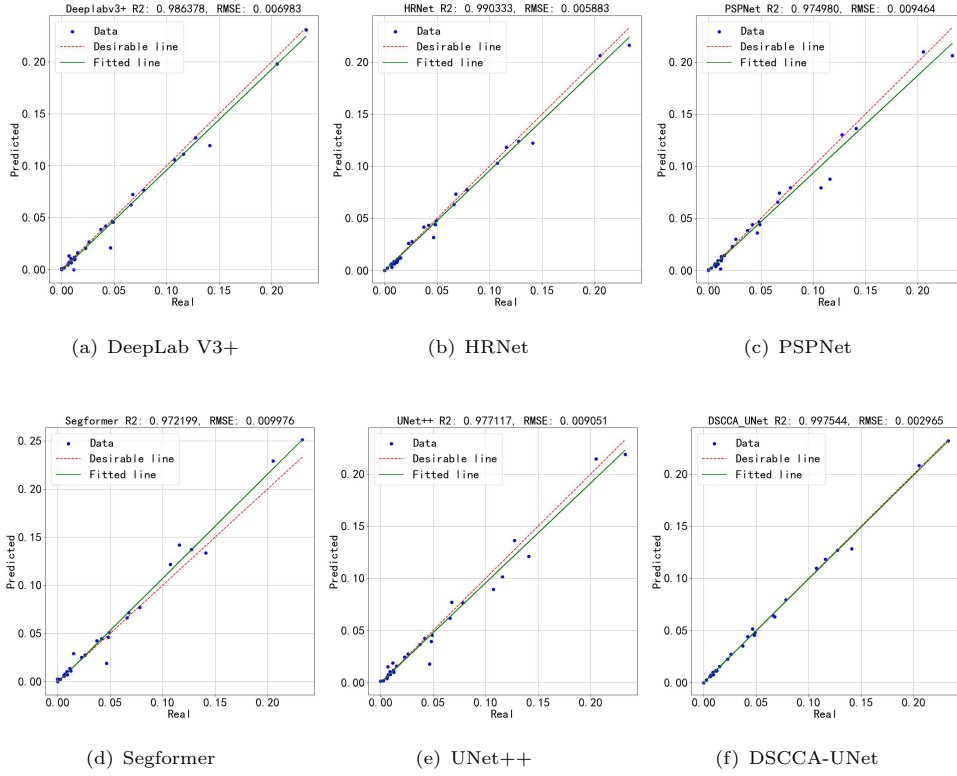


Fig. 10 Disease severity estimation results for different semantic segmentation models

In disease severity estimation, 30 images are randomly selected from the actual captured images to validate the effectiveness of the improved DSCCA-UNet model. The segmentation results are obtained by using the semantic segmentation models. The lesion and leaf pixel counts are calculated. The RMSE and R^2 for each model are computed, leading to the final disease severity estimation results. As shown in Figure 10, the fitted lines for Deeplabv3+, HRNet, PSPNet, and UNet++ fall below the desirable lines, indicating that these models tend to underestimate disease severity. This underestimation is related to missed detections and false positives in lesion pixel segmentation. On the other hand, Segformer tends to overestimate disease severity, implying the issue of lesion pixel merging in its segmentation results impacts its accuracy in severity estimation. Then, the improved DSCCA-UNet model achieves the best estimation performance, with an R^2 of 0.998 and an RMSE of 0.00297.

A disease severity estimation system for apple leaves is developed using PyQt5 framework. It can facilitate the estimation of apple leaf disease severity in real orchard scenarios. This system employs the improved DSCCA-UNet model, using the trained best.pt weights file to perform semantic segmentation on apple leaf images. Based on the segmentation results, the system counts the number of lesion pixels and healthy leaf pixels, and ultimately determines the severity grading by referencing the disease

severity grading criteria. As shown in Figure 11, the system allows users to upload images locally and receive both the semantic segmentation result and the disease severity grading.

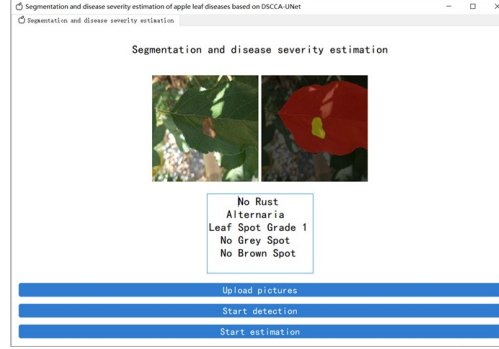


Fig. 11 The interface of apple leaf spot segmentation and disease severity estimation system

An experimental field in the apple orchard of Dashah, Feng County, Jiangsu Province, is selected as the study area of this paper to generate prescription maps. Firstly, the LocaSpaceViewer is used to locate the study field. The study area is delineated using the "Draw Surface" command, enabling the acquisition of the geographical location of the study field and the preservation of its corresponding latitude and longitude coordinates. The obtained field information is then imported into ArcMap software to generate a surface layer grid to rasterise the prescription map. Then the element class is added to generate the fishing net, so that the grid becomes an editable and readable surface map layer. The images collected in the orchard are uploaded to the apple leaf spot segmentation and disease severity estimation system to obtain the disease severity grade. Finally, the disease severity grades is imported into the corresponding surface layer as prescription values to generate the final apple leaf spot prescription map. In the case of apple alternaria leaf spot, Figure 12 shows a prescription map in the study field. Most of the disease severity grades are 1 and 3, with fewer occurrences of grades 5-9 due to the presence of anthropogenic treatments.

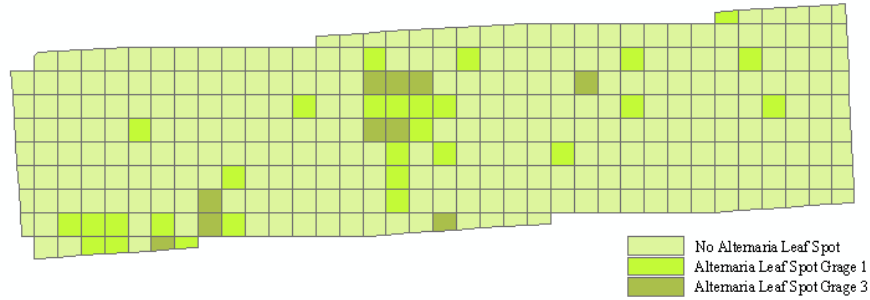


Fig. 12 Prescription map of apple alternaria leaf spot

6 Conclusion

This paper proposes an improved DSCCA-UNet model by incorporating the DSCC multi-scale feature extraction module and designing the new type connection of attention mechanism. This model can enhance the ability to capture features of curved spot edges and reduce background noise interference, effectively improving spot segmentation accuracy. A method is designed for disease severity estimation, which is combined the improved DSCCA-UNet and an automatic pixel counting algorithm. Experiments show that, the proposed model achieves MIOU, MPA, Mprecision and F1 of 89.76%, 93.84%, 95.06% and 94.45%, respectively. It shows improvements of 2.75%, 1.91%, 1.41% and 1.67% over the baseline UNet model. The improved DSCCA-UNet demonstrates significant advantages in segmentation accuracy, compared to semantic segmentation models such as DeepLabv3+, HRNet, PSPNet, and UNet++. It achieves an R^2 of 0.998 and an RMSE of 0.00297 in disease severity estimation, demonstrating the best estimation performance compared to other models. Based on the spot segmentation results and the automatic pixel counting algorithm, an apple leaf spot segmentation and disease severity estimation system is developed. A prescription map generation method is proposed for application in real-world orchard management scenarios, facilitating the practical implementation of prescription map generation tasks. In the future, we will investigate pest and disease early warning models.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

Data will be made available on request.

Funding statement

This work is supported by the National Natural Science Foundation of China (Grant No. 62173166), the Basic Research Program of Xuzhou (Grant No. KC22053).

References

- [1] Tang L., Yi J., Li X. Improved multi-scale inverse bottleneck residual network based on triplet parallel attention for apple leaf disease identification. *Journal of Integrative Agriculture* **2024**. 23(3), 901–922.
- [2] Liu B., Huang X., Sun L., Wei X., Ji Z., Zhang H. MCDNet: Multi-scale constrained deformable convolution network for apple leaf disease detection. *Computers and Electronics in Agriculture* **2024**. 222, 1–12.

- [3] Huo S., Duan N., Xu Z. An improved multi-scale YOLOv8 for apple leaf dense lesion detection and recognition. *IET Image Process* **2024**. 18, 4913–4927.
- [4] Bock C. H., Chiang K.S., Del P. E.M. Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy. *Tropical Plant Pathology* **2022**. 47(1), 25–42.
- [5] Del P. E. M., Nelson S. C., Pethybridge S. J. Evaluation of app-embedded disease scales for aiding visual severity estimation of cercospora leaf spot of table beet. *Plant Disease* **2019**. 103(6), 1347–1356.
- [6] Chen C., Wang X., Heidari A. A., Yu H., Chen H. Multi-threshold image segmentation of maize diseases based on elite comprehensive particle swarm optimization and otsu. *Frontiers in Plant Science* **2021**. 12, 1–20.
- [7] Song H., Wang J., Bei J., Wang Min. Modified snake optimizer based multi-level thresholding for color image segmentation of agricultural diseases. *Expert Systems with Applications* **2024**. 255, 1–30.
- [8] Ortega-Sánchez N., Rodríguez-Esparza E., Oliva D., Pérez-Cisneros M., Mohamed W.A., Dhiman G., Hernández-Montelongo R. Identification of apple diseases in digital images by using the Gaining-sharing knowledge-based algorithm for multilevel thresholding. *Soft Computing* **2022**. 26, 2587–2623.
- [9] Jothiaruna N., Joseph Abraham Sundar K., Karthikeyan B. A segmentation method for disease spot images incorporating chrominance in comprehensive color feature and region growing. *Computers and Electronics in Agriculture* **2019**. 165, 1–8.
- [10] Ma J., Du K., Zhang L., Zheng F., Chu J., Sun Z. A segmentation method for greenhouse vegetable foliar disease spots images using color information and region growing. *Computers and Electronics in Agriculture* **2017**. 142, 110–117.
- [11] Liu C., Zhao R., Pang M. Lung segmentation based on random forest and multi-scale edge detection. *IET Image Processing* **2019**. 13(10), 1745–1754.
- [12] Deng J., Huang W., Zhou G., Hu Y., Li L., Wang Y. Identification of banana leaf disease based on KVA and GR-ARNet. *Journal of Integrative Agriculture* **2024**. 23(10), 3554–3575.
- [13] Zuo X., Jiang J., Shen J., Yang W. Improving underwater semantic segmentation with underwater image quality attention and multi-scale aggregation attention. *Pattern Analysis and Applications* **2025**. 28(80), 1–12.
- [14] Apedo Y., Tao H. A weakly supervised pavement crack segmentation based on adversarial learning and transformers. *Multimedia Systems* **2025**. 31(266), 1–21.

- [15] Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* **2015**. Munich, Germany, 234–241.
- [16] Wang J., Sun K., Cheng T., Jiang B., Deng C., Zhao Y., Liu D., Mu Y., Tan M., Wang X., Liu W., Xiao B. Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2020**. 43(10), 3349–3364.
- [17] Zhao H., Shi J., Qi X., Wang X., Jia J. Pyramid scene parsing network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* **2017**. Honolulu, USA, 6230–6239.
- [18] Chen L.-C., Papandreou G., Kokkinos I., Murphy K., Yuille A. L. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2017**. 40(4), 834–848.
- [19] Xie E., Wang W., Yu Z., Anandkumar A., Álvarez J. M., Luo, P. SegFormer: Simple and efficient design for semantic segmentation with transformers. *ArXiv* **2021**. arXiv:2105.15203, 1–18.
- [20] Zhou Z., Rahman Siddiquee M. M., Tajbakhsh N., Liang J. UNet++: A nested U-Net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* **2018**. Granada, Spain, 3–11.
- [21] Zu Z., Lei H., Qu Z., Huang Z., Suo, W. Facilitating air transportation of agricultural systems via intelligent runway detection. *Computational Intelligence* **2025**. 41, 1–9.
- [22] Jia Z., Shi A., Xie G., Mu S. Image segmentation of persimmon leaf diseases based on UNet. *International Conference on Intelligent Computing and Signal Processing (ICSP)* **2022**. Xi'an, China, 2036–2039.
- [23] Li K., Zhang L., Li B., Li S., Ma J. Attention-optimized DeepLabV3+ for automatic estimation of cucumber disease severity. *Plant Methods* **2022**. 18(1), 93–109.
- [24] 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. *Computers and Electronics in Agriculture* **2021**. 189, 1–13.
- [25] Lu B., Lu J., Xu X., Jin Y. MixSeg: A lightweight and accurate mix structure network for semantic segmentation of apple leaf disease in complex environments. *Frontiers in Plant Science* **2023**. 14, 1–16.

- [26] Faıçal B. S., Freitas H., Gomes P. H., Mano L.Y., Pessin G., de Carvalho A., Krishnamachari B. Ueyama J. An adaptive approach for UAV-based pesticide spraying in dynamic environments. *Computers and Electronics in Agriculture* **2017**. 138, 210–223.
- [27] Wang L., Lan Y., Yue X., Ling K., Cen Z., Cheng Z., Liu Y., Wang J. Vision-based adaptive variable rate spraying approach for unmanned aerial vehicles. *International Journal of Agricultural and Biological Engineering* **2019**. 12(3), 18–26.
- [28] Wen S., Zhang Q., Deng J., Lan Y., Yin X., Shan J. Design and experiment of a variable spray system for unmanned aerial vehicles based on PID and PWM control. *Applied Sciences* **2018**. 8(12), 1–22.
- [29] Wang D., Zhao M., Li Z., Xu S., Wu X., Ma X., Liu X. A survey of unmanned aerial vehicles and deep learning in precision agriculture. *European Journal of Agronomy* **2025**. 164, 1–36.
- [30] Chen F., Zhang Y., Zhang J., Liu L., Wu K. Rice false smut detection and prescription map generation in a complex planting environment, with mixed methods, based on near earth remote sensing. *Remote Sensing* **2022**. 14(4), 1–13.
- [31] Huang H., Deng J., Lan Y., Yang A., Deng X., Wen S., Zhang H., Zhang Y. Accurate weed mapping and prescription map generation based on fully convolutional networks using UAV imagery. *Sensors* **2018**. 18(10), 1–12.
- [32] Guo W., Gong Z., Gao C., Yue J., Fu Y., Sun H., Zhang H., Zhou L. An accurate monitoring method of peanut southern blight using unmanned aerial vehicle remote sensing. *Precision Agriculture* **2024**. 25(4), 1857–1876.
- [33] Bao W., Zhao J., Hu G., Zhang D., Huang L., Liang D. Identification of wheat leaf diseases and their severity based on elliptical-maximum margin criterion metric learning. *Sustainable Computing: Informatics and Systems* **2021**. 30, 1–9.
- [34] Xing B., Wang D., Yin T. The evaluation of the grade of leaf disease in apple trees based on PCA-logistic regression analysis. *Forests* **2023**. 14(7), 1–22.
- [35] Qi Y., He Y., Qi X., Zhang Y., Yang G. Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. *IEEE/CVF International Conference on Computer Vision (ICCV)* **2023**. Paris, France, 6047–6056.
- [36] Qiu Y., Liu Y., Chen Y., Zhang J., Zhu J., Xu J. A2SPPNet: Attentive atrous spatial pyramid pooling network for salient object detection. *IEEE Transactions on Multimedia* **2023**. 25, 1991–2006.

- [37] Ou X., Wang H., Zhang G., Li W., Yu S. Semantic segmentation based on double pyramid network with improved global attention mechanism. *Applied Intelligence* **2023**. 53(15), 18898-18909.
- [38] Li Z., Zhang L., Yin S., Zhang G. MSCFF-Net: multi-scale context feature fusion network for polyp segmentation. *Multimedia Systems* **2025**. 31(189), 1–17.
- [39] Woo S., Park J., Lee JY., Kweon I.S. CBAM: Convolutional block attention module. *European Conference on Computer Vision (ECCV)* **2018**. Munich, Germany, 3–19.
- [40] Farjon G., Liu H., Edan Y. Deep-learning-based counting methods, datasets, and applications in agriculture: a review. *Precision Agriculture* **2023**. 24(5), 1683–1711.
- [41] Feng J., Chao X., Zhang Z., He D., Zhang J., Ye Z. A pooling module with multidirectional and multi-scale spatial information and its application on semantic segmentation of leaf lesions. *Precision Agriculture* **2023**. 24(6), 2416–2437.