Wheat Powdery Mildew Spore Images Segmentation Based on U-Net

Xinshen Liang and Botao Wang
Faculty of Information Technology, Beijing University of Technology, Beijing, China
Email: 634070097@qq.com, wangbt@bjut.edu.cn

Abstract. In recent years, there have been significant development in deep learning technology, which has also promoted the development of image segmentation. Wheat powdery mildew is a common crop disease, which is very harmful to crops. The segmentation task of wheat powdery mildew spore images is important for the target spores identification and spores counting. In this paper, an improved framework based on U-Net is proposed, comparing with original U-Net, we add the pyramid pooling module after the 1024-channel feature map in the down-sampling part to extract different sizes pooling feature and fuse into a global feature map, and adjust some skip connections in original U-Net to obtain more features that are effective for spore images segmentation, experiment shows that the improved U-Net structure has better segmentation performance than U-Net, and the segmentation miou (mean intersection over union) has reached 91.4% in wheat powdery mildew spore image dataset, which proves that our proposed architecture is effective and competitive in the wheat powdery mildew spore segmentation task.

1. Introduction
Wheat powdery mildew is a worldwide disease, and its spores often grow on the plant surface, affecting the growth of plants. It is common in the northern and southern provinces of China [1] and cause huge damage to crops and economy. Therefore, how to accurately segment the wheat powdery mildew spores images, realize the automatic monitoring of wheat powdery mildew is an important issue that need to be resolved.

Through investigation and literature review, we find that there are not many studies on wheat powdery mildew spore image segmentation and detection, so we also conducted an analogy study on some other spore and cell image segmentation methods. In Ref. [2], the K-means clustering algorithm is adopted, the B channel component of the RGB color image of the winter spores of wheat smut fungus is used as the clustering object, and the sum of the R, G and B component values is unchanged as the iteration termination condition Image segmentation. In Ref. [3], the fuzzy C-means algorithm is used to calculate the threshold of Canny edge detection, and then the watershed algorithm combining distance transformation and Gaussian filtering is used to segment the adhesion spores. In Ref. [4], author uses VGG-16 model which has been trained to initialize the parameters of SegNet network, and replace the classification pooling layer of VGG-16 with a fully connected layer for segmentation to segment blood cells. In Ref. [5], a convolutional neural network combined with a seed watershed method is used to segment the nuclei in the cell image. In Ref. [6], a CFL (Constrained Focal Loss) Loss function is proposed for DeepLabv3+ network training and anthracis spore segmentation. In Ref. [7], an improved architecture CSnet based on CNN (convolutional neural network) is proposed to segment low-resolution cell images.
The wheat powdery mildew spore image has many problems such as mass noise and blurred edges. According to the investigation and preliminary experiments, we found that the traditional methods like the K-means, canny edge detection and watershed method are sensitive to the noise and lack of robustness, so we tend to use the U-Net to segment our spore images.

U-Net is a segment net, which architecture proposed by Ronneberger in Ref. [8], which is widely used in image segmentation tasks in the field of computer vision. It has a very good effect on medical cell image segmentation. In Ref. [9], Huang et al. combine the U-Net neural network with improved level set method to finish the segment task of overlapping cervical smear cells. In Ref. [10], Mostayed et al. proposed an architecture which uses the content-adaptive convolution to replace the skip connections operation in the traditional U-Net for cell segmentation.

In order to investigate U-Net’s application in wheat powdery mildew spore images segmentation, in this paper, we propose an improved U-Net framework that adds the pyramid pooling module and adjust the skip connection number to achieve better segment result. This new architecture includes an encoder and decoder, a pooling pyramid module, and a skip connection between the 512-channel feature maps. The pooling pyramid module pools the feature map into four different scales, and fuses four different features from corresponding pyramid pooling layers. We use the skip connection to contact the feature map information in the down-sampling path with the same channel numbers feature map in the up-sampling path. We use the improved U-Net for image segmentation of wheat powdery mildew spores, and the experiment result shows that the improved U-Net architecture performs better in the spore segmentation task than the original U-Net.

2. Proposed Methodology

2.1. Dataset Processing
In this paper, 103 microscopic wheat powdery mildew spores images of different sizes collected by the equipment are carefully annotated as ground truth images, and we use various methods such as rotation, random color transformation of the RGB channel, random parameter Gaussian filtering, and random cropping of the image area to augment our image data, and the final 835 images can be use for the experiment in this article. The following figure 1 show the wheat powdery mildew spores and the ground truth pictures in our data set.

![Figure 1](image.png)

**Figure 1.** Wheat powdery mildew spores image and its ground truth.

2.2. Proposed Architecture
Figure 2 shows the U-Net architecture without modification. Our proposed U-Net framework is improved base on the original U-Net architecture. After the 1024-channel feature convolution layer, we add a pyramid pooling module to harvest different scale feature map of different sub-region. The pyramid pooling module was proposed in Ref. [11], it is a four-level pooling architecture with different sizes of 1×1, 2×2, 3×3 and 6×6 respectively, and average pooling operation is chosen, considering the weight of the global feature map, for balance weight of various feature map, if the pyramid has N levels, we use 1×1 convolution operation after four different levels to reduce the
number of corresponding feature channels to $1/N$, finally, the low-dimension feature map concatenated by different level sub-region is upsampled via bilinear interpolation to get the same size as original feature map. Figure 3 shows the structure of pyramid pooling module.

**Figure 2.** This is an original U-Net, consist of encoder-decoder architecture and four skip connections between the feature maps.

In order to avoid the convolution and pooling operation changing the original size information of the image, we add reflection padding operations in convolution and pooling operation, and add a batch normalization layer after convolution layer to avoid the gradient disappearing during training. L2 regularization is used to prevent overfitting. After the batch normalization layer, we choose rectified linear unit (ReLu) to increase network nonlinearity as the activation function, at the same time, skip connection is used to transfer and merge feature information between the same size feature maps. we remove first, second and third skip connection in the net, just keep the fourth connection between encoder’s 512-channel feature map with the same channel numbers feature map in the decoder, which can keep high level feature concatenate and reduce more low level feature concatenate compared with original U-Net.

In model training, we use binary cross entropy loss as the loss function of the training network. The representation of binary cross entropy loss is:

$$
loss = -\sum_{i=1}^{n} \hat{y}_i \log y_i + (1-\hat{y}_i)\log(1-\hat{y}_i)
$$

$$
\frac{\partial loss}{\partial y} = -\sum_{i=1}^{n} \hat{y}_i - 1 \hat{y}_i
$$

where $y_i$ is the ground truth of the $i$ pixel, and $\hat{y}_i$ is the model predict label, $n$ is the num of image pixels. In spore images, the number of spore pixels is much less than the number of background pixels, the imbalance can affect to our model learning result, we deal with this problem with reweighting the two classes, we calculate to determine that in the global ground truth data, the pixel ratio of the target to the background is close to 1:25 as figure 4 shows, so we set the class weight of the spore area in the ground truth 10 times in the experiment, to balance the class difference in number of pixels.

**Figure 3.** The pyramid pooling module for pooling feature map into different scales.

In the segmentation task of wheat powdery mildew spores, we need to consider establishing a quality standard for segmentation results to describe the quality of the segmentation results. We use miou (mean intersection over union) as the index to describe the segmentation target picture, which is expressed as follows:

$$
miou = \frac{1}{k+1} \sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} \frac{P_{ii}}{P_{ij} + \sum_{j=0}^{k} P_{ji} - P_{ii}}
$$
Figure 4. Histogram of the frequency of the two classes pixels.

where \(k + 1\) is the number of class including the background, \(P_j\) is the number of \(i\) class pixels that are misclassified to \(j\) class, \(P_{ii}\) is the \(i\) number of pixels that are classified correctly.

3. Experiments and Result Analysis

We divide 835 images into two set, training image set of 550 images and test image set of 285 images. Each epoch during training, 20% of the total number of training images were used as verification set. We design experiment to compare the segment result of different network, for pyramid pooling modules and the number of skip connections.

Table 1 shows the different network’s test miou comparison in dataset, figures 5-6 shows the different segmentation result of different method.

| Method | Pyramid Pooling Module | First Feature map Skip connection | Second Feature map Skip connection | Third Feature map Skip connection | Fourth Feature map Skip connection | Mean IoU % |
|--------|-------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|------------|
| 1      | ×                       | √                                 | √                                 | √                                 | √                                 | 84.097     |
| 2      | ×                       | ×                                 | ×                                 | ×                                 | ×                                 | 89.386     |
| 3      | √                       | ×                                 | √                                 | √                                 | ×                                 | 88.416     |
| 4      | √                       | ×                                 | ×                                 | ×                                 | ×                                 | 90.767     |
| 5      | √                       | ×                                 | ×                                 | √                                 | ×                                 | 88.638     |
| 6      | √                       | ×                                 | √                                 | ×                                 | ×                                 | 86.754     |
| 7      | √                       | ×                                 | ×                                 | √                                 | ×                                 | 87.461     |
| 8      | √                       | ×                                 | ×                                 | √                                 | ×                                 | 91.477     |

This research focuses on investigating the impact of pyramid pooling module and the skip connection. Figures 5-6 show the segmentation result via different method, comparing with figures 5b-5e, we can find that adding the pyramid pooling module is able to keep more effective segmentation result, and reduce incorrect segmentation of similar noise. Comparing figures 5d-5i and figures 6d-6i, we can find that keeping the fourth skip connection and removing first, second, third skip connection perform better, for which can reduce the harvest of low-level feature, like figures 5f and 6f, first skip connection between 64-channel feature maps take more low-level feature, the noise with similar color and shape is segmented mistakenly, the same phenomenon appear on figures 5g-5h and 6g-6h. As figures 5i and 6i show, just keep the fourth can reduce most incorrect segmentation, and achieve best segmentation effect.
Table 1 lists the mean intersection over union (miou) of different framework, comparing the method 1 and 3, method 2 and 4, we can also draw the conclusion of pooling pyramid module can improve the segmentation effect, and the comparison of method 3-8 shows that just keep fourth skip connection between 512-channel feature maps can improve the miou of segmentation. These experiments indicate that our proposed framework is able to effectively segment the wheat powdery mildew spores images.

4. Conclusion
In this paper, we propose an improved U-Net architecture that adds the pooling pyramid module, and removes some skip connections in the original U-Net framework, compared to original U-Net, new network architecture has achieved higher segmentation accuracy in the segmentation task of wheat powdery mildew spore image dataset, which proves the effectiveness of the framework. Next, we will investigate more strategies to train the improved U-Net, which is able to make it more suitable for our wheat powdery mildew spore segmentation task.
References

[1] Wang D P, Wang B T and Yan Y 2013 The identification of powdery mildew spores image based on the integration of intelligent spore image sequence capture device 2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Process pp 177-180.

[2] Deng J Z, Li M, Yuan Z B, Wang Z and Li J 2012 Image segmentation of winter spores of wheat smut fungus based on K-means clustering Journal of South China Agricultural University 33 (02) 266-269.

[3] Qi L, Jiang Y, Li Z H, Ma X, Zheng Z X and Wang W J 2015 Automatic detection and counting of rice blast fungus spores based on microscopic image processing Transactions of the Chinese Society of Agricultural Engineering 31 (12) 186-193.

[4] Tran T, Kwon O, Kwon K, Lee S and Kang K 2018 Blood cell images segmentation using deep learning semantic segmentation 2018 IEEE International Conference on Electronics and Communication Engineering (ICECE) pp 13-16.

[5] Kowal M, Żejmo M, Skobel M, Korbicz J and Monczak R 2020 Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm Journal of Digital Imaging 33 (1) 231-242.

[6] Zhao Y C, Lin F S, Liu S G, Zhu H H, Li H and Bai Y 2019 Constrained-focal-loss based deep learning for segmentation of spores IEEE Access 99 1-1.

[7] Liu Y, Yu N M, Fang Y and Wang D F 2018 Low resolution cell image edge segmentation based on convolutional neural network 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC) pp 321-325.

[8] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional networks for biomedical image segmentation Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th Int. Conf. (Munich) vol 9351 (Springer) pp 234-241.

[9] Huang Y M, Zhu H Q, Wang P Y and Dong D P 2019 Segmentation of overlapping cervical smear cells based on u-net and improved level set 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) pp 3031-3035.

[10] Mostayed A, Wee W and Zhou X F 2019 Content-adaptive U-Net architecture for medical image segmentation 2019 International Conference on Computational Science and Computational Intelligence (CSCI) pp 698-702.

[11] Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y 2017 Pyramid scene parsing network 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp 6230-6239.