

Segmentation of Corn Leaf Disease Based on Fully Convolution Neural Network

Zhen Wang^a, Shanwen Zhang^b

School of Information and Engineering, Xijing University, Xi'an, 710123, China
^awzwwk2313@163.com, ^bwscs1325@163.com

ABSTRACT. Aiming at the problem that traditional computer vision is susceptible to different illumination and complex background in the segmentation of corn leaf disease, a method of segmentation of corn leaf disease based on full convolution neural network is proposed. Firstly, pre-processing and data enhancement are performed on the acquired images to establish a training set and a test set for deep learning. Secondly, the training centralized image is input into the FCN, and the feature map is obtained through multiple convolution, pooling, and activation layers; The feature map is upsampled to obtain a feature map of the same size as the input image. Finally, the resolution of the segmented image is restored by the process of deconvolution, and the segmentation result is output. This method was used to segment the common corn leaf disease images. The results show that the method has better segmentation effect and the segmentation accuracy is over 96%.

KEYWORDS: Corn leaf disease image, Computer vision, Convolutional neural network, Image segmentation

1. Introduction

Corn disease is one of the main factors affecting corn yield. Common corn diseases include large spot disease, small spot disease, stripe rust and other diseases. These diseases seriously affect the yield and quality of corn. Plant leaves are the main part of the reaction diseases. According to the disease spots of the plant leaves, the degree of disease occurrence can be judged intuitively, and the disease can be prevented early. In recent years, with the continuous development of agricultural Internet of Things technology, different plant leaf growth information can be collected through a variety of sensors. Therefore, the automatic recognition of various leaf diseases can be realized by computer vision technology and pattern recognition technology. Disease image segmentation is one of the key factors affecting the recognition effect in the process of automatic recognition. Qi Zhihua et al. [1] improved the segmentation accuracy and reduced the segmentation time by using the improved fuzzy edge detection algorithm for wheat lesion segmentation. Hu Weizhen et al [2] first used the method of color moment and gray level co-

occurrence matrix to extract wheat disease characteristics, and then used median filtering combined with K-means clustering method to segment lesion image on wheat powdery mildew, stripe rust and leaf rust. Li Chao et al [4] used the adaptive learning local binary model to extract the feature differences between normal leaf and diseased leaf images to determine the lesion segmentation threshold. This method has fewer feature dimensions than the traditional feature extraction method. More precise segmentation. Qin Feng et al [6] used K-median clustering and linear discriminant analysis to segment the diseased leaves in the segmentation of the disease of the temporal lobe, and achieved a good segmentation effect. In order to improve the traditional threshold and the segmentation method of clustering on the segmentation of leaf lesions, Wang Xianfeng et al [5] first use SVM to segment the lesion image, then use image morphology to correct the segmentation image, and finally use this method in cucumber leaf spot. The segmentation test was carried out on the diseased leaves, and a good segmentation effect was obtained. Xia Yongquan et al [7] proposed an adaptive fuzzy C-means clustering algorithm, which uses the pixel gray value and its neighborhood mean as the input features of the initial cluster, and determines the optimal cluster number and fuzzy weighted index by cluster validity. , in the cotton lesion image segmentation to achieve better results. The above-mentioned traditional machine vision method is mainly used for the segmentation of plant leaf lesions. Although this method has achieved certain effects on the segmentation of plant leaf lesions, the operation of these traditional methods is more complicated and less robust. The feature extraction method is relatively simple, which makes the segmentation method less versatile.

At present, deep learning has achieved good application results in the fields of image segmentation and image recognition, and it has also been widely used in the segmentation and recognition of plant lesions. Yang Jindan et al [10] used convolutional neural network to segment strawberry leafy powdery mildew lesions, using hybrid pooled CNN-9 (convolution kernel scale 5x5, 3x3; convolution network depth 11) as the segmentation model, successful The leaf spot area of the strawberry leaf was segmented. Zhao Bing et al [11] replaced the last three fully connected layers of the traditional convolutional neural network with three convolutional layers on the segmentation of grape disease leaves, and extracted the characteristics of grape disease leaves by multi-layer convolution. The Matthew's coefficient is used to evaluate the segmentation effect. The results show that the method has high segmentation accuracy, which creates favorable conditions for accurate segmentation of diseased leaves and feature extraction of disease. Zhang et al [12] first used image median filtering and histogram threshold method to preprocess the image, and then used the full convolutional neural network to segment the image in the image of wheat sheath blight, stripe rust and leaf rust. Good results have been achieved in the segmentation. Xie Zeqi et al [13] constructed a basic model of the full convolutional neural network based on the traditional VGG16 convolutional neural network, and used the structural model to extract the lesion area on the cucumber leaf disease image. The experiment shows that the model has a high Segmentation accuracy. The above method of deep learning has achieved good results in plant disease leaf segmentation. Therefore, this paper proposes a method based on full convolutional neural network for corn leaf disease

segmentation, using convolutional layer instead of the last three of CNN. The connection layer performs upsampling on the feature map obtained by the convolution layer to preserve the spatial information of the original input image, and obtains the final segmentation result by classifying the upsampled feature map pixel by pixel.

2. Materials and Methods

2.1 Material

The image collection of corn leaf disease was collected from the Daejeon Experimental Base of the Agricultural Science Institute of Baoji City, Shanxi Province, and the image data was collected under normal lighting conditions using a Canon SX620 digital camera with a pixel value of 20.2 million. The collected images included six common corn diseases, such as corn leaf spot, small spot disease, leaf spot disease, brown spot disease, streak disease and round spot disease. 125 images were collected for each disease image, and a total of 750 images were collected. The collected images were uniformly compressed by the resolution of each image using software Python 3.6, and the compression size was 256×256 . Finally, the image was stored in JPEG format to construct a corn leaf disease image database. Part of the image is shown in Figure 1.



Figure. 1 Six kinds of corn diseased leaf images

In the 750 images collected, since the lesion area is small in size and the shape is irregular rectangle, three smaller areas and three similar lesion colors are selected when generating the candidate area, and 9 is generated. The boundary areas of the candidate regions are labeled with the training image to obtain 750 original training sets with sample labels.

Normalizing the scale of each blade image in the training set with the sample label: first, the minimum enclosing area included in the foreground of the leaf image is obtained, and the perimeter l of the area is compared with 200, if $l > 200$, Then, the original map and the foreground template map are scaled down according to the scale factor $s=200/l$; otherwise, they are unchanged. Then, centering on the enclosing area, the original color blade image and the 220×220 circular area in the target background binary image are cut out so that the foreground pixel is completely within the circle, and the out-of-circle pixel is completely the background. In order to solve the over-fitting problem of network model training,

the normalized image is subjected to four pre-processing operations of translation, rotation, scaling and color jitter. When each training sample is expanded to 20, the training set is expanded to 15000. An extended training sample set of images. The specific operations are as follows:

Translation: After each pre-processed training sample is shifted to 8 pixels in the upper left, upper right, lower left, and lower right directions, the center is centered, and the center 256×256 area is cropped, thereby making an image. Expanded into 5 different samples containing an image that was not translated.

Rotation: Each pre-processed training sample is randomly rotated 5 times in the range of rotation angle $[-5^\circ, 5^\circ]$, centered on the center of the circle, and the center 256×256 area is cropped, thereby expanding an image For 5 different samples.

Scaling: Each pre-processed training sample is randomly reduced by 5 times in the range factor $[0.85, 1]$, centered on the center of the circle, and the center 256×256 area is cropped, thereby expanding an image into 5 different samples.

Color jitter: The three color components R, G, and B of the RGB color of each training sample after pre-processing are uniformly multiplied by 5 times in the range of $[0.8, 1.2]$, and will exceed the value of 255. Set to 255 to avoid overflow distortion, then combine the three R, G, and B component images into a color RGB image centered on the center, cropping out the center 256×256 region, thereby expanding an image to five different sample.

All the training samples after pre-processing are averaged, that is, the pixel values corresponding to the corresponding positions of all the images are summed and divided by the total number of samples, and a three-channel mean image with a gray scale of 256×256 is obtained, and input from all the training samples is obtained. The mean of the corresponding pixel position is subtracted. Then all the obtained images are further scrambled to form an averaged training set.

2.2 Method

The Full Convolutional Neural Networks (FCN) used in this paper are based on the Alexnet network architecture, respectively. In order to reduce the number of parameters in the network model training process and solve the problem of insensitivity to the input image size, the fully connected layer in the original convolutional neural network is replaced by a convolutional layer. Since the image resolution is reduced by downsampling during the convolution process, a deconvolution operation is added to the last three convolutional layers of the network model to restore the image resolution.

The structure of the Alexnet full convolutional neural network based on Fig. 2 shows that the first few layers in the FCN are identical to the Alexnet structure, and the full connection layer is replaced by a convolutional layer. The basic structure mainly includes 8 convolution layers (Conv1~Conv8), 3 pooling layers (Pool1, Pool2, and Pool3), 1 upsampling (Deconv), and 1 crop layer. The crop layer is the result of upsampling. Crop to make the size equal to the original input image. The

parameters to be learned by the FCN come from the convolution kernel of each convolutional layer. The use of small channel numbers can reduce network parameters and reduce network complexity. In order to ensure the nonlinearity of the output of each layer of the network, each convolutional layer output results through a nonlinear activation correction linear unit (relu) function, relu can accelerate the convergence of the network.

The parameters of Conv8 and its subsequent upsampling layer in the FCN are training from the beginning, the learning rate of Conv1 to Conv7 is set to 0.001, the learning rate of the eighth convolutional layer weight is set to 0.1, and the upsampling after Conv8 The convolution kernel size in the layer is set to 63, step size 32, and the 9×9 core input is upsampled to 319×319, and then the crop layer is cropped to the same size as the original image according to the size and offset parameters of the original blade image. . Initially, the learning rate of the offset is set to 0.2, and the learning rate of the upsampling layer is set to 0.1, after which the learning rate is adjusted to 0.7 of the previous round per 1000 iterations. The final output of the FCN is the probability that each pixel belongs to the background and foreground, and each pixel is classified into a class with a higher probability, thereby obtaining a lesion image.

The image size of the training set is 256×256×3 as the input image. Convolution, maximum pooling, convolution, and maximum pooling operations are performed on the first four layers of Conv1, Pool1, Conv1 and Pool2, respectively. The obtained feature map sizes are respectively : 112 × 112 × 96, 56 × 56 × 96, 56 × 56 × 256, 28 × 28 × 256. The feature maps obtained by the previous convolutional layer are sequentially convoluted using three consecutive convolutional layers Conv3, Conv4, and Conv5, and the obtained feature map sizes are 28×28×384, 28×28×384, and 28×, respectively. 28 × 256. On the pooling layer Pool5, the maximum pooling operation is performed on the feature maps obtained by the Conv3, Conv4 and Conv5 convolutional layers, and the feature map size is 14×14×256. On three consecutive convolutional layers Conv6, Conv7 and Conv8, three different convolution operations are performed on the feature map obtained by Pool5, and the obtained feature map sizes are 9×9×4096, 9×9×4096, 9 respectively. ×9×2. The three consecutive convolutional layers Conv6, Conv7 and Conv8 are deconvolved by the upsampling layer Deconv, and the size of the feature map is 319×319×2. The result obtained by the cropping layer on the upsampling layer is cropped, and the lesion size is 256×256×2. The convolutional layer mathematical expression is as follows:

$$x^l = f(W^l x^{l-1} + b^l) \quad (1)$$

In the formula (1), x^{l-1} is output of the $l-1$ hidden layer, x^l is the output of the convolution layer in the l hidden layer, x^l is the input image of the input layer, W^l indicating the mapping weight matrix of the l hidden layer, b^l is bias of the l hidden layer $f()$ is the relu function, and its expression is $f(x) = \max(0, x)$.

In the training process of the network model, the loss function is used to evaluate the training effect and improve the training effect. The convolution layer operation is repeated in the network model training process until the model loss converges,

that is, the loss is reduced to a certain extent and then no longer reduced, and a network model capable of accurately detecting the disease is obtained. The mathematical expression of the loss function is as follows:

$$L(P) = \frac{1}{2N} \sum_{i=1}^N \|E_i(I_i, P) - D_i\|_2^2 \quad (2)$$

In equation (2), where P is the parameter that the network model needs to learn, I_i is the i training image on the training set, N is the number of images in the training set, D_i is the labeled disease image, E_i is the image of the lesion detected by the FCN. $L(P)$ is the loss obtained by calculating the Euclidean distance between the marked lesion image and the detected lesion image.

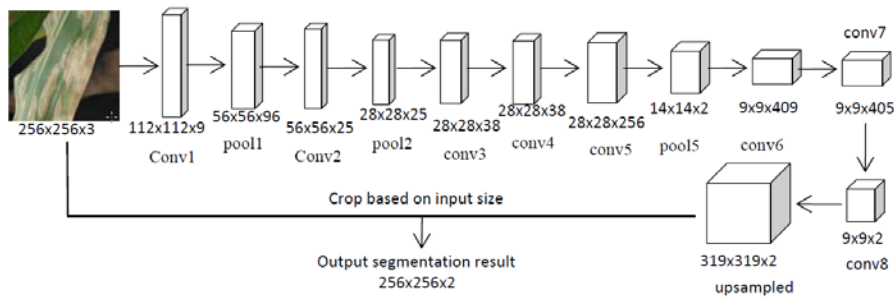


Figure. 2 FCN full convolution neural network structure

2.3 Evaluation criteria for segmentation results

In order to quantify the segmentation effect, the IOU (Intersection Over Union) function is used to evaluate the image segmentation effect. The IOU function is an important index for calculating the correlation between the real value and the predicted value of the segmented image. The value refers to the ratio of the number of pixels simultaneously appearing in two regions in the image to the total number of pixels after the region is merged. Its mathematical expression is as follows:

$$O[R, P] = \frac{|R \cap P|}{|R \cup P|} = \frac{TP}{TP + FP + FN} \quad (3)$$

In equation (3): R is the labeled ground truth, P is the image divided by the full convolutional neural network, TP is the number of identical pixels in the two regions, and FP and FN are different in the two regions. The number of pixels.

3. Results and Analysis

References In the corn leaf disease segmentation test based on the full convolutional neural network, the hardware environment used is mainly Ubuntu

operating system, the CPU is Intel Core i5-7200 2.8 GHz, the memory is 16GB, the GPU is NVIDIA GTX 1050ti, and the memory is 4GB. Implemented using the Tensorflow framework, using the GPU for acceleration during training and testing.

In the test process, the training set image marked in the database is input into the convolutional neural network to train the network model, and the convolution layer parameters are set as shown in Fig. 2. The main function of the pooling layer is to reduce the amount of calculation and prevent over-fitting. The methods of pooling layer mainly include average pooling and maximum pooling. It is found in experiments that selecting the maximum pooling method will result in segmented image resolution. The problem of descent, so this article chooses the average pooling method to downsample the input image. In the process of convolution calculation, the activation function is used to improve the nonlinearity of the network model. The commonly used activation functions are mainly sigmoid, tahn and relu. During the experiment, it is found that the relu activation function can greatly activate the neurons of the neural network. Increase the nonlinearity of the network model, so use relu as the activation function in the experiment. In order to improve the segmentation effect of the network model and reduce the training time, the network parameters are set using the fine-tune method. In the back propagation of training, the stochastic gradient descent method was chosen, and the learning rate was set to 1×10^{-12} , and the momentum factor was set to 0.99.

In order to show the effectiveness of the segmentation effect, N-cut based image segmentation algorithm, SVM-based segmentation method and segmentation method were used to segment the corn spot disease, small spot disease and leaf spot disease respectively. Table 1 lists the IOU index, network training time, and image segmentation time. Figure 3 shows the segmentation results for three different algorithms.

Table 1 The IOU index, training time and image segmentation time three image segmentation algorithms

Method	IOU index (%)	Training time (h)	Split time (s)
N-Cut	92.38	3.1	4.2
SVM	93.59	2.6	3.8
My Method	96.26	5.2	2.1

It can be seen from Table 1 that in the initial model training, the segmentation algorithm takes a long time to extract the input image due to the convolutional layer, but the trained model is compared with the traditional segmentation method in the later stage. Both the split time and the split effect have a good performance during use.

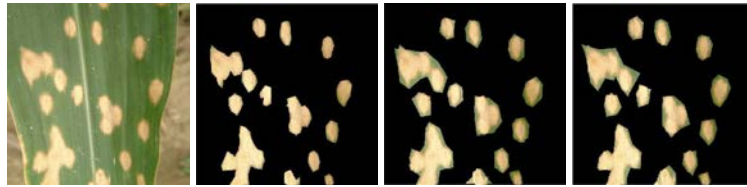


Figure. 3 Image Segmentation Effect of Corn Leaf Diseases by Different Segmentation Methods

It can be concluded from Fig. 3 that the segmentation method of this paper has a better segmentation effect, and the diseased area of the corn leaf is completely extracted. Based on the N-cut and SVM segmentation algorithms, only the contour regions of the disease can be extracted, but still contain a large number of normal parts of the corn leaves.

4. Conclusion

Aiming at the problem of image segmentation of corn leaf diseases, a segmentation method based on full convolutional neural network is adopted to solve the problem of low segmentation accuracy in image segmentation of corn pests and diseases in traditional segmentation methods and the parameters and convergence speed of existing convolutional neural networks during training. Slow and easy to fall into local optimum problems, this paper proposes a method based on full convolution neural network to provide a theoretical basis for crop leaf disease monitoring system.

The specific steps of the method mainly include: firstly preprocessing and data enhancement of the acquired image, establishing a training set and a test set for the network model; secondly, inputting the training centralized image into the FCN, after multiple convolution, pooling, and activation layers A feature map is obtained; then the feature map is upsampled to obtain a feature map having the same size as the input image; finally, the resolution of the segmented image is restored by a process of deconvolution, and the segmentation result is output.

The method is compared with N-cut and SVM, and the results show that the proposed method has better segmentation effect.

References

- [1] Li, H.Y., Chen, Z.Y., Zhou, J.C., Zhang, J.X.: Occurrence and Damage of the Main Diseases of Cucumber in Greenhouse During Growth Period and Shelf Life (in Chinese). *Modern Agricultural Sciences and Technology*, 20, 187--189, 191 (2010)

- [2] Sasaki, Y., Okamoto, T., Imou, K., Torii, T.: Automatic Diagnosis of Plant Disease-recognition Between Healthy and Diseased Leaf. *Journal of the Japanese Society of Agricultural Machinery*. 61, 119--126 (1999)
- [3] Zhao, Y.X., Wang K.R., Bai, Z.Y., Li, S.K., Xie, R.Z., Gao, S.J.: Research of Maize Leaf Disease Identifying System Based Image Recognition (in Chinese). *Scientia Agricultura Sinica*. 40, 698--703 (2007)
- [4] Sanyala, P., Patel, S.C.: Pattern Recognition Method to Detect Two Diseases in Rice Plants. *Imaging Science Journal*. 56, 319--325 (2008)
- [5] Cui, Y.L., Cheng, P.F., Dong, X.Z., Liu, Z.H., Wang, S.X.: Image Processing and Extracting Color Features of Greenhouse Diseased Leaf (in Chinese). *Transactions of the CSAE*. 21(supp.), 32--35 (2005)
- [6] Tian, Y.W., Zhang, C.H., Li C.H.: Study on Plant Disease Recognition Using Support Vector Machine and Chromaticity Moments (in Chinese). *Transactions of the Chinese Society of Agricultural Machinery*. 35, 95--98 (2004)
- [7] Tian, Y.W., Niu, Y.: Applied Research of Support Vector Machine on Recognition of Cucumber Disease (in Chinese). *Journal of Agricultural Mechanization Research*. 31, 36--39 (2009)
- [8] Cen, Z.X., Li, B.J., Shi, Y.X., Huang, H.Y., Liu, J., Liao, N.F., Feng, J.: Discrimination of Cucumber Anthracnose and Cucumber Brown Speck Based on Color Image Statistical Characteristics (in Chinese). *Acta Horticulturae Sinica*. 34, 1425--1430 (2007)
- [9] Pydipati, R., Burks, T.F., Lee, W.S.: Identification of Citrus Disease Using Color Texture Features and Discriminant Analysis. *Computers and Electronics in Agriculture*. 52, 49--59 (2006)
- [10] Pydipati, R., Burks, T.F., Lee, W.S.: Statistical and Neural Network Classifiers for Citrus Disease Detection Using Machine Vision. *Transactions of the ASAE*. 48, 2007--2014 (2005)
- [11] Camargo, A., Smith, J.S.: Image Pattern Classification for the Identification of Disease Causing Agents in Plants. *Computers and Electronics in Agriculture*. 66, 121--125 (2009)
- [12] Wang, N., Wang, K.R., Xie, R.Z., Lai, J.C., Ming, B., Li, S.K.: Maize Leaf Disease Identification Based on Fisher Discrimination Analysis. *Scientia Agricultura Sinica*. 42, 3836--3842 (2009)
- [13] Vapnik, V.N.: *The Nature of Statistical Learning Theory*. Springer-Verlag, New York (1995)
- [14] Song, K., Sun, X.Y., Ji J.W.: Corn Leaf Disease Recognition Based on Support Vector Machine Method (in Chinese). *Transactions of the CSAE*. 23, 155--157 (2007)
- [15] Ren, D., Yu, H.Y., Wang, J.H.: Research on Plant Disease Recognition Based on Linear Combination of the Kernel Function Support Vector Machine (in Chinese). *Journal of Agricultural Mechanization Research*. 29, 41--43 (2007)
- [16] Zhang, J., Wang, S.X.: A Study on the Segmentation Method in Image Processing for Plant Disease of Greenhouse (in Chinese). *Journal of Inner Mongolia Agricultural University*. 28, 19--22 (2007)

- [17] Mao, H.P., Zhang, Y.C., Hu, B.: Segmentation of Crop Disease Leaf Images Using Fuzzy C-means Clustering Algorithm (in Chinese). Transactions of the CSAE. 24, 136--140 (2008)
- [18] Tian, Y.W., Li C.H.: Color Image Segmentation Method Based on Statistical Pattern Recognition for Plant Disease Diagnose (in Chinese). Journal of Jilin University (Engineering and Technology Edition). 34, 291--293 (2003)
- [19] Selim, S.Z., Ismail, M.A. K-means-type Algorithm: a Generalized Convergence Theorem and Characterization of Local Optimality. IEEE Transactions on Pattern Analysis and Machine Intelligence. 6, 81--87 (1984)
- [20] Li, G.L., Ma, Z.H., Huang, C., Chi, Y.W., Wang, H.G.: Segmentation of Color Images of Cucumber Diseases Using K-means Clustering Algorithm. Transactions of the CSAE. 26(Supp.2), 32--37 (2010)
- [21] Li, Z.R., He, D.J.: Research on Identify Technologies of Apple's Disease Based on Mobile Photograph Image Analysis (in Chinese). Computer Engineering and Design. 31, 3051--3053, 3095 (2010)