Accurate recognition method of plant leaves based on multi-feature fusion

Lin, Ruikai, Ma, Junwei, Yu, Huiling, Zhang, Yizhuo

Ruikai Lin, Junwei Ma, Huiling Yu, Yizhuo Zhang, "Accurate recognition method of plant leaves based on multi-feature fusion," Proc. SPIE 11928, International Conference on Image Processing and Intelligent Control (IPIC 2021), 119280Q (22 October 2021); doi: 10.1117/12.2611757

Event: International Conference on Image Processing and Intelligent Control (IPIC 2021), 2021, Lanzhou, China

Accurate Recognition Method of Plant Leaves based on Multi-feature Fusion

Ruikai Lin[1], Junwei Ma[1], Huiling Yu*[1], Yizhuo Zhang*[1]
[1]Northeast Forestry University, Harbin, Heilongjiang Province, China 150000
*yhl@nefu.edu.cn; nefuzyz@163.com

ABSTRACT

During the use of a convolutional neural network to train a recognition model of plant leaves, the convolutional layers focus on the appearance of leaves in learning the features of them, while ignoring their internal texture features, thereby resulting in the misclassification of plant leaves with similar appearance. Aiming at this problem, this paper proposes an accurate identification method of plant leaves based on multi-feature fusion, which can be applied to extract the appearance and texture features of leaves simultaneously, and to conduct fusion and summation for these two types of features. The experimental results indicate that compared with the accuracy of the ordinary convolutional neural network recognition method and traditional machine learning method, the accuracy of this method has been improved substantially.

Keywords: convolutional neural network, leaf recognition, texture feature, appearance feature

1. INTRODUCTION

Composed of roots, stems, leaves and flowers, plants can be distinguished by the appearance and texture of these organs. Due to the higher universality and diversity than other five organs, leaves can be a reference basis for the classification of many kinds of plants. In addition, most leaves are flat, so they are prone to be stored as two-dimensional images. Thus, with the development of machine learning and computer vision technologies, an idea of identifying and classifying plants based on leaf images has gradually become a mainstream. Zhang et al. proposed a leaf recognition model based on a K-nearest neighbor algorithm [1], making the classification accuracy of 100 kinds of leaf images reach 91.37%. A support vector machine algorithm applied by Wang et al. achieved the accuracy of 91.41% in identifying plant leaves [2]. By adding 10 feature parameters to the leaf recognition model based on a support vector machine [3], Wei et al. raised classification accuracy to 95.8%. Receiving tremendous artificial intervention during feature selection, the above models presented poor generalization ability in recognizing other leaf images. With the rise of deep learning technologies, neural network-based image classification models bypasses the artificial feature engineering of traditional image recognition models. In addition to the ability of learning features independently, their training efficiency has been improved substantially. Wu et al. applied a probabilistic neural network [4] to recognize 32 kinds of plant leaves with the accuracy of about 93%. Zhang et al. combined a convolutional neural network and a support vector machine to classify plant leaves with the accuracy of 91.11% [5]. Huang et al. proposed a recognition model based on an upgraded AlexNet (a convolutional neural network) to make accuracy reach 94% [6].

It can be seen that the leaf recognition method based on the convolutional neural network can achieves higher classification accuracy due to its avoidance of human interference in applying traditional algorithms to select features. However, Lee et al. found that the convolutional neural network learns the features of leaf images based on their external morphological features [7], which causes the misclassification of many leaves with similar appearance. In response to this problem, a convolutional neural network recognition model based on feature fusion proposed in this paper is able to simultaneously learning the external morphological and internal texture features of plant leaves, which overcomes the shortcomings of traditional convolutional neural network classification models and avoids the misclassification of different plants' leaves with similar appearance into one category.

International Conference on Image Processing and Intelligent Control (IPIC 2021),
edited by Feng Wu, Fengjie Cen, Proc. of SPIE Vol. 11928, 119280Q

2. DEVELOPMENT ENVIRONMENT

2.1 Development Architecture

TensorFlow, a deep learning framework released by Google Brain, supports programming languages such as C++ and Python, and is widely applied in the field of computer vision. It describes the calculation process with data flow diagrams. In data flow diagrams, a node represents an independent operation, and an edge represents data used or generated in the process of calculation. This model can not only quickly build and deploy deep neural networks, but also facilitate the allocation of calculation amount among multiple computing nodes of a CPU or a GPU.

2.2 Experimental Environment

The computer was equipped with an Intel Core i7-8700K CPU with the benchmark speed of 3.20GHz, 16GB memory and an Nvidia GTX 1080Ti graphics card with the video memory of 11GB. The software environment of the experimental research embraced Ubuntu 16.04 (64-bit) operating system, Python 3.6, TensorFlow-GPU, CUDA and cuDNN.

3. MODEL DESIGN

3.1 Data Preparation and Preprocess

The leaf data used in the experiment were mainly from the UCI Folio Leaf Dataset published by the University of California, Irvine. In addition to the Folio Dataset, the leaf images of some common plants in Asia we collected were also added into the experimental data set.

The Folio Dataset consists of 640 high-resolution images of 32 different plants taken in full daylight. Each leaf is photographed against a white background. The leaf images of this data set, as shown in Figure 1a, can display the edges of the leaves better. Texture feature cropping, which is to sharpen a picture and cut the central image out, was performed on all the leaf images to highlight the texture features of the leaves. A total of 640 images based on original ones were created, as shown in Figure 1b. All 1280 images were scaled to 229×229 to facilitate subsequent process.

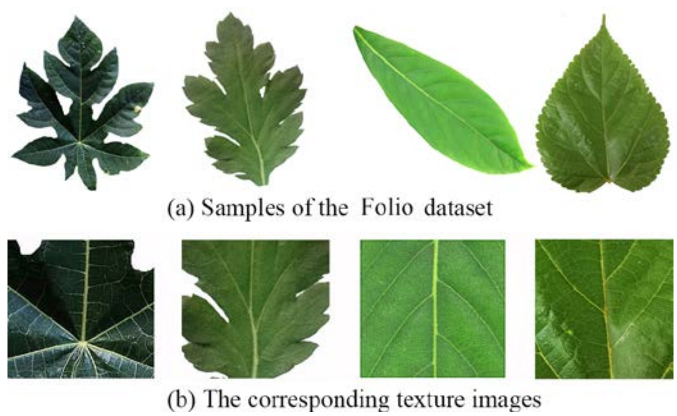


Figure 1. Folio - a data set of plant leaves

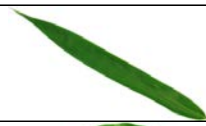
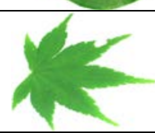
The expanded data set was collected from 44 kinds of plants, containing a total of 2816 leaf images with more complex backgrounds. Some samples are shown in Figure 2a. The preprocess steps of these images are as follows. Firstly, the images were segmented by using a Sobel edge detection algorithm and a watershed algorithm to obtain the foreground images of them as image data representing the appearance features of the leaves. Secondly, texture images were cropped from the foreground images to obtain the data of 2816 images representing leaf texture features. One sample is shown in Figure 2c. Finally, the resolution of the complete images and the texture images were all scaled to 229×229.

(a) The original image (b) Foreground image (c) Texture feature image

Figure 2. Supplementary leaf images

The appearance of 5 kinds of leaves was summed up, each of which represents some different plant species with similar shapes and large differences in internal texture. They are prone to be misclassified in case of inputting a single image into a convolutional neural network. The number of the plant species included in each common shape in the data set is shown in Table 1.

Table 1. Common leaf shapes included in the data set.

Type	Appearance images	Number of species
No.1		13
No.2		5
No.3		14
No.4		12
No.5		28

The 30% of the plant leaves included in each shape were randomly selected as a test data set, and the rest was regarded as a training set. There were 2419 training images and 1307 test images in total.

3.2 Overview of Convolutional Neural Networks

The Convolutional Neural Network (CNN) is a feedforward neural network specially for processing grid structure data, characterized by local perception and weight sharing [8]. Its neurons can interact with other neurons in a certain range and preserve spatial structure [9]. Directly taking original images as input [10], the network bypasses the feature engineering of image data in traditional image recognition technologies. So far, it has been widely used in the field of computer vision.

Its network structure is functioned as feature extraction and feature mapping, generally composed of a convolutional layer, a pooling layer and a fully connected layer. The convolutional layer is used to reduce parameter input, extract local features and form a feature image. Its mathematical expression is:

$$x_j^l = f(\sum_{i \in M_j} x_j^{l-1} \times \omega_{ij}^l + b_j^l) \tag{1}$$

Where, x_j^l was the jth image of the lth convolutional layer; M_j was the input feature image; ω_{ij} was a weight matrix; b_j^l was the bias term of the convolutional layer.

f() is a non-linear activation function. The commonly used are Sigmoid, tanh and ReLU functions. When the input value of tanh and Sigmoid functions is very large or very small, the output result varies in a small range. In this case, a phenomenon of gradient disappearance is liable to appear, which makes a deep network be locally optimal. Due to quick convergence, the ReLU function applied in this paper not only can reduce computational costs, but also retains the non-linear expression ability, so it is applicable for training deep networks. The mathematical expression of the ReLU function is:

$$f(x) = \max(0, x) \tag{2}$$

The pooling layer following the convolutional layer can compress input feature images, extract the main features, and reduce the computational complexity of the network. Its mathematical model can be expressed as:

$$x_j^l = f(\beta_j^l \times Down(x_j^{l-1}) + b_j^l) \tag{3}$$

Where, β was the weight. Down() represented down-sampling operation. In the paper, the MaxPooling layer was used to take the maximum value in a feature area as a value in a new abstract area.

Playing a role of classification, the fully connected layer integrates feature information learned by a network, and then maps it to a predicted target.

3.3 Establishment of a Dual Channel CNN

The appearance and texture images of a leaf were input into two feature extraction paths. The appearance feature vector $X(x_{1\times1\times1}, x_{1\times1\times2}, \cdots, x_{1\times1\times n})$ and texture feature vector $Y(y_{1\times1\times1}, y_{1\times1\times2}, \cdots, y_{1\times1\times n})$ were summarized and aggregated to obtain the fusion feature vector Z:

$$Z = \Sigma_1^n (x_{1\times1\times n} + y_{1\times1\times n}) \tag{4}$$

Finally, the learned fusion features were mapped to labeled space through the fully connected layer to complete the classification of the leaf images.

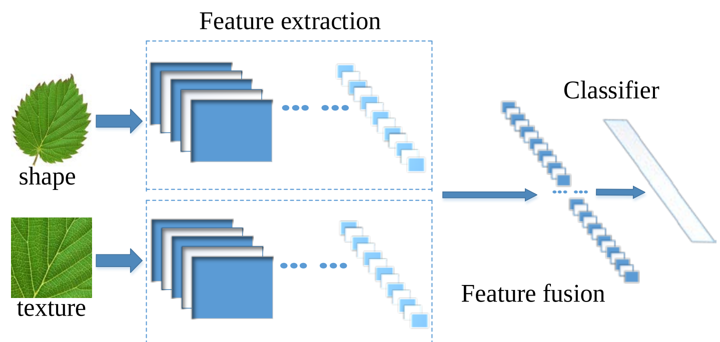


Figure 3. Model architecture

The appearance extraction path was used to learn the external appearance features of the leaves, and the texture extraction path was used to learn the internal texture features of them. The former, composed of 7 convolutional layers and 4 maximum pooling layers, used large convolution kernels of 11×11, 5×5 and 3×3 to facilitate the model to learn the overall appearance features of plant leaves. The model architecture with gradually shrinking convolution kernels was designed in consideration of the parameters and training time of the model. The latter, composed of 6 convolutional layers and 4 maximum pooling layers, used six 3×3 convolution kernels to learn the texture features of the leaves and constantly streamlined the extracted features through four MaxPooling layers.

4. EXPERIMENTAL PROCESS AND ANALYSIS

In order to solve the problem of fewer training set samples and improve the invariance of the recognition mode, we rotate, flip, scale and add Gaussian noise [11] to the leaf images in the experiment, and save the new images to the training data set. Moreover, we adopted a batch algorithm, using 32 training samples in each batch and 26 test samples for verification. The training set was traversed once in each of total 100 times iterations. A regularization factor was added and set to 0.005 in order to avoid over-fitting. A stochastic gradient descent method was adopted to reduce the time overhead. The initial learning rate was set to 0.01, and the learning rate gradually decayed to 0.001 as the training progressed.

An analysis on the Top-1 value (the probability that the recognition result with the highest probability after classification matches the label of the tested image) and Top-3 value (the probability that the recognition results with the top three highest probabilities after classification match the label of the tested image) of the experimental results was conducted to prove the strong generalization ability of the model. According to the test, the accuracy rates of Top-1 and Top-3 were 99.26% and 99.97% respectively when the model recognized a leaf image with single background, and they were 97.32% and 99.74% when the background was complex. A comparative analysis on the performance of an ordinary convolutional neural network recognition model, a curvature histogram + support vector machine recognition model and a multi-scale distance matrix + support vector machine recognition model on the data set of this paper was conducted in order to objectively reflect the superiority of the proposed model, and the comparison result is shown in Table 2.

Table 2. Performance comparison with other models.

Classification technique	Accuracy %		Standard deviation	
	Top-1	Top-3	Top-1	Top-3
The model proposed in this paper	99.26	99.97	0.18	0.10
Ordinary CNN model	96.22	99.46	0.26	0.21
Curvature histogram+SVM	82.84	94.62	1.63	0.21
Multi-scale distance matrix+SVM	82.55	95.12	1.61	1.10

It can be seen that that compared with the ordinary CNN recognition model, the dual channel CNN that learns the appearance and texture features of the leaves had a significantly improved recognition rate and standard deviation obtained on the basis of the tested data set, which indicated that both the appearance and texture of a leaf are important for leaf recognition and classification, and the dual channel CNN can heighten the model's ability in recognizing plant leaves.

5. CONCLUSION

Previous studies have shown that the convolutional neural network learns the features of leaf images mainly based on their edges, which causes the misrecognition of many leave with similar appearance. This paper proposes a leaf recognition model based on a dual channel convolutional neural network, which first extracts and fuses the appearance and texture features of a leaf, and then completes recognition and classification. According to the experimental results, the model can avoid the interference caused by the similar appearance of leaves. With outstanding robustness and much higher recognition accuracy than the ordinary neural network model and traditional image classification model, the model has high practical value.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding by the Natural Science Foundation of Heilongjiang Province (LH2020F002).

REFERENCES

[1] Zhang N, Liu W P. Plant leaf recognition method based on clonal selection algorithm and k-nearest neighbor. Computer Application (2013).

[2] Wang L J, Huai Y J, Peng Y C. Leaf-viewing plant species recognition based on multi-feature fusion of leaf images. Journal of Beijing Forestry University, 37.001(2015):55-61.

[3] Wei L, He D J, Qiao Y L. Plant leaves classification based on image processing and SVM. Journal of Agricultural Mechanization Research, 2013, 35(5):12--15.
[4] Wu S G, Bao F S, Xu E Y, et al. A leaf recognition algorithm for plant classification using PNN (probabilistic neural network) [C/OL] // Proceedings of 2007 IEEE International Symposium on Signal Processing and Information Technology. New York: IEEE, 2007 [2018--02--27]. https://ieeexplore.ieee.org/document/4458016.
[5] Zhang S, Huai Y J. Leaf image recognition based on layered convolutions neural network deep learning. Journal of Beijing Forestry University, 2016, 38(9):108--115.
[6] Huang F L, Shen T P, Jin L. Improved AlexNet convolutional neural network for leaf classification of Chinese herbal medicine. Journal of Anqing Normal University (2020).
[7] Lee S H, Chan C S, Wilkin P, et al. Deep-plant: plant identification with convolutional neural networks [C] // Proceedings of 2015 IEEE International Conference on Image Processing (ICIP) . Quebec: IEEE, 2015:452--456.
[8] Krizhevsky, Alex, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. Communications of the ACM 60.6(2017):84-90.
[9] Lu H T, Zhang Q C. Applications of deep convolutional neural network in computer vision. Journal of Data Acquisition and Processing, 2016, 31(1):1--17.
[10] Lecun, Y., L. Bottou. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86.11(1998):2278-2324.
[11] Wang Y, Sun W, Zhou X P. Research on Image Recognition Method of Chinese Herbal Plants Based on Deep Learning. Information on Chinese Medicine, Vol. 37, No. 6, 2020, 21-25, ISTIC (2020)