PURPOSE-LED
PUBLISHING™

PAPER • OPEN ACCESS

Leaf Classification Based on Convolutional Neural Network

View the article online for updates and enhancements.

Leaf Classification Based on Convolutional Neural Network

Peng Wu[1,a], Qian Zhou[2,b*]

[1,2]Shandong Institute of Commerce and Technology, Jinan, China

No. 4516, Tourist Road, Jinan City, Shandong Prov. China 250103

[a]email: 20050107@sict.edu.cn

[*]Qian Zhou: [b]email: zhouqian@sict.edu.cn

Abstract. Convolutional Neural Network (CNN), a very important neural network structure in deep learning, is a network model often used in image classification, target recognition and other fields. In botany, leaf classification and recognition is very important for identifying new or scarce tree species. In nature, plants are widely distributed, and the survival and development of all living things on the earth depend on plants. Identification of species by leaves and related research are of great help to the study the evolution law of plants, the protection of plant species and the development of agriculture. This paper uses convolutional neural network in artificial intelligence to identify the leaves of several kinds of trees collected by Kunming Institute of Botany, Yunnan Province, which can realize the automatic extraction of leaf image features, reduce tedious labor costs, and realize the use of artificial intelligence to classify leaves, thus providing an auxiliary means of artificial intelligence for botany research.

1. Introduction

CNN is a kind of feedforward neural network, which includes convolution layer with convolution calculation function and depth structure of multilayer neural network, and is one of the important algorithms of machine learning. With the rapid development of artificial intelligence and deep learning, the advantages of convolutional neural network in image classification become more and more prominent, which makes it widely used in industry, agriculture, military and other fields, including botany. Plants are widely distributed in nature. The classification of plant species is the first to proceed with the research of botany. Starting with the area and shape of the leaves to distinguish the tree species, this is of great significance for the study of the tree species itself, even the study of the regional ecological environment, and other research work of plant researchers. Accurate information on species types and spatial distribution is the basis of natural resources utilization and protection. Researchers at the Institute of Botany often go to the field to collect plant leaves. Traditional plant recognition mainly relies on manual recognition[1]. Due to the harsh field environment and inconvenient communication, the experience of collectors is very high[2]. Manual recognition limits the accuracy of leaf recognition to a great extent[3] [4]. In this paper, convolution neural network algorithm is used to identify the leaves of several typical trees in a large number of leaf samples from Kunming Institute of Botany, Yunnan Province. Convolutional neural network can realize automatic extraction of leaf image features, reduce tedious labor costs, realize artificial intelligence method to classify leaves, and provide artificial intelligence assistance for botany research.

The sample set used in this paper was collected by plant researchers in Kunming, Yunnan. Most leaves are collected from remote mountain forests at home and abroad. They are collected by tower

cranes and scanned into electronic images. After image acquisition, OpenCV will be used for image preprocessing, including denoising, changing image size and so on, to generate a standard and unified sample set. At the same time, the binarization operation of leaves will be carried out, which is convenient for calculating the leaf area, automatically calculating the leaf area index, transpiration rate, photosynthetic rate, etc., and improving the working efficiency. As shown in Figure 1, in the experiment, five representative leaves were selected for classification and recognition. See Table 1 for different folders and corresponding tree species for storing five leaves.

Fig. 1 Five kinds of leaves

Table. 1 Folders correspond to tree species

Folder	varieties of trees
Folder 1	Ficus langkokensis Drake
Folder 2	Diospyros xishuangbannaensis
Folder 3	Pometia pinnata
Folder 4	Canarium subulatum Guill.
Folder 5	Parashorea chinensis Wang Hsie

2. Introduction to Convolutional Neural Networks

Receptive field is the inspiration of CNN. In 1960s, Hubel et al. discovered the rule of the size of visual receptive field and summed up the concept of receptive field through experimental study of cat visual cortical cells. The input image tensor flows into each layer of neurons in the CNN, and after it flows out, it becomes the feature map tensor. Each pixel value on the feature map is mapped by a corresponding size area on the original image, which is called the receptive field. In other words, the pixel values within a certain area of the input tensor jointly determine the pixel value of a certain point of the output tensor, pixel value of a certain point of the output tensor is influenced by all pixels in the receptive field area of the input image.

CNN is a special deep neural network model with three characteristics. On the one hand, its neuronal connections are not completely connected. The pixel value of the output feature image is not calculated from all the pixel points of the input image (called full connection), but is influenced by the pixels in the receptive field area (called incomplete connection). However, the influence of pixel values within the receptive field of the input image on the corresponding pixels of the output feature map is not necessarily uniform. For example, it is possible that the middle influence is large and the surrounding influence is small, so when the influence range is expressed by the receptive field, it is necessary to add weight to the influence of each pixel position in the receptive field. The larger the value, the greater the influence. Adding influence weight to each pixel in the receptive field is the convolution kernel in CNN[5].

On the other hand, the second characteristic is that neurons in the same layer of CNN are sharing weights. In other words, different receptive fields of the input image use the same convolution kernel to generate each pixel in the feature map. Deep neural network structure is difficult to train and learn. The number of neuron weights is proportional to the complexity of the neural network model. Reducing the former can solve the problem of training difficulties to a certain extent.

The network structure of incomplete connection and weight sharing makes CNN imitate and approach the real biological neural network, which makes CNN more suitable for image processing and speech recognition than the fully connected network.

The third feature is that the convolution kernel parameters of CNN are obtained through training data learning, which avoids the complex artificial feature extraction. For example, traditional iris recognition and classification applications like this need to manually extract features such as length and width of different kinds of iris flowers, and classify iris flowers according to these manual features. There are differences in the quality of manually extracted features, such as whether the extracted features are closely related to the prediction results, and feature extraction will affect the prediction results. In the CNN network, the work of extracting features is no longer done manually, but the convolution kernel. The weights of convolution kernel are not assigned manually, but initialized as random values. With the increase of training times, the optimized weights are obtained through constant correction.

3. Neural model building

In this paper, Google TensorFlow machine learning framework and Keras are used to complete the construction, training and prediction of CNN neural network. After TensorFlow2.0 version 2.0, Keras was used as its advanced deep learning API. Keras is an advanced neural network API written in Python, which can run with TensorFlow as the back end. The development of Keras focuses on supporting rapid experiments. It can convert ideas into experimental results with minimal time delay.

Before using neural network, OpenCV is used to preprocess the image in advance, including removing noise and unifying size. At the same time, the binarization operation of leaves will be carried out, which is convenient for calculating the leaf area, automatically calculating the leaf area index, transpiration rate, photosynthetic rate, etc., and improving the working efficiency. Then build a convolutional neural network. It usually consists of several groups of convolution layers and maximum pooling layers, followed by flattening layers and multiple fully connected layers. Convolution layer realizes feature extraction of image by convolution operation. The number of channels of the image tensor varies depending on the number of convolution kernels, but the convolution layer does not change the image size. Convolution feature graph can be activated by relu function, thus adding non-linear elements to linear convolution neural network, which makes classification curve better distinguish different categories. The max-pooling layer only keeps the maximum value of pixels in the neighborhood, which reduces the size of the picture and keeps the salient features of the image. The saved computing power can enable the convolution layer to extract more features. Therefore, the later the convolution layer, the more convolution kernels there are.

After several rounds of convolution and max-pooling, usually three rounds, followed by flattening layer, the purpose is to flatten the image into a one-dimensional matrix, so as to prepare for multiplication with the next fully connected layer matrix. There can be two fully connected layers. It should be noted that the number of neuron matrix columns in the last fully connected layer must be the number of classifications to be identified. So far, the complete convolution neural network structure is basically completed, as shown in Figure 2. Generally, in order to reduce the over-fitting of neural network to the training sample set, shedding elements and regularization are added to the whole neural network.
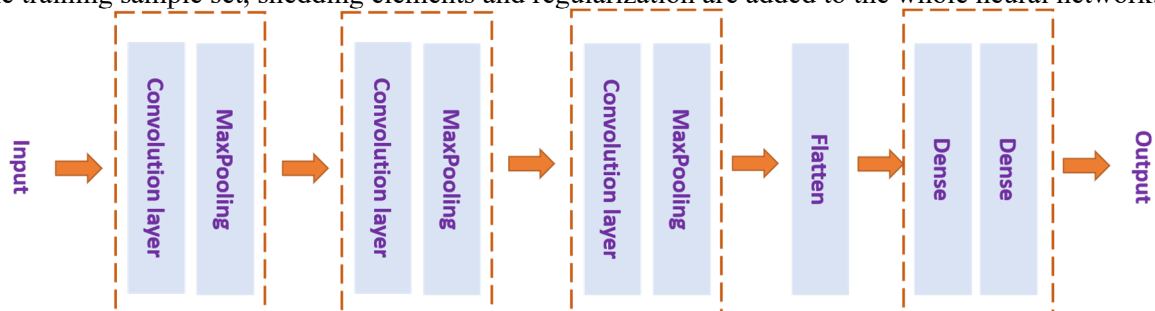

Fig. 2 Neural network model structure

The problem of over-fitting usually occurs when there are too many variables (features). In this case, the trained equation always fits the training data well, that is to say, the loss function may be very close to 0 or just 0. However, this curve tries every means to fit the training data, which leads to its inability to generalize to new data samples, so that new samples cannot be predicted. Too many variables (features) and very little training data will lead to over-fitting. Therefore, in order to solve over-fitting, dropout and regularization are introduced.

The principle of Dropout is that, in each training batch, the phenomenon of over-fitting can be obviously reduced by ignoring some feature detectors (letting some hidden layer nodes have values of 0). When propagating forward, let the activation value of a neuron stop working with a certain probability p, which can make the model more generalized, because it does not depend too much on some local features.

Regularization is an effective way to avoid model over-fitting and ensure generalization ability by controlling model complexity in machine learning. An index describing the complexity of the model is added to the loss function of the model. The commonly used function to describe the complexity of the model is L2 regularization. By limiting the weight, the model can not fit the random noise in the training data arbitrarily. After the weight is squared, the small weight becomes smaller and the large weight becomes larger, which also plays the function of feature selection, but does not make the weight become 0. Assuming that the network layer parameter to be regularized is W, the l2 regularization formula is as shown in Formula (1). Among them, Ein is the training sample error without regularization term, λ controls the size of regularization term, and a larger value of λ will constrain the complexity of the model to a greater extent; or vice versa, Dallas to the auditorium In practice, the regular term is usually added to the objective function (loss function), and the error of the whole objective function is propagated back, so that the regular term can influence and guide the network training. L2 regularization is commonly called "weight attenuation" in deep learning.

$$L = E_{in} + \lambda \sum_j W_j^2 \tag{1}$$

In addition, a Softmax activation unit is added behind the last full connection layer. Through it, the values of output neurons are mapped in the range of 0-1, and the sum is guaranteed to be 1 after normalization, so that the sum of probabilities of multiple classifications is exactly 1, that is, 100%.

4. Sample pretreatment and model training

All leaf images are scanned by scanner, and some simple image preprocessing needs to be done by OpenCV, such as reducing to uniform size, filtering to remove noise and so on. The processed image samples are randomly scrambled and divided into training set and test set according to the proportion of 80% and 20%. The model is trained by the model.fit () method in TensorFlow keras database, Adam algorithm is selected as the optimizer, the loss is calculated by cross entropy, and the performance is evaluated by classification accuracy.

As the basis of loss, cross entropy can correctly express the difference between predicted value and label value. The larger the difference, the greater the value of cross entropy, and vice versa. Secondly, it can be seen from the calculation process that the cross entropy is only related to the label category, and the more the predicted value belonging to this category approaches 1, the better. The mean square error, which is related not only to the real category but also to other items, requires the difference between the real values and the predicted values of all categories to be squared again. The more average the predicted values belonging to each category will make the mean square error get the minimum value. Different from classification problem, regression problem solves the problem of predicting specific values. For example, housing price forecast and sales volume forecast are all regression problems. What these problems need to predict is not a predefined category, but an arbitrary real number. Neural networks for solving regression problems generally have only one output node, and the output value of this node is the predicted value. For regression problems, the most commonly used loss function is the mean square error. As for cross entropy loss, since the complex similarity matrix between categories is difficult to quantify, we can only pay attention to the category to which the sample belongs, as long as the predicted value of the real category is closer to 1, which shows that it is more reasonable.

Training iteration times are also introduced in the model.fit () method. After the training, the loss curve and accuracy curve can be observed through the TensorBoard visualization panel, so that the super-parameters of the model can be easily adjusted. After the model training is completed, the weights of neural network are solidified, which can be used for image prediction and leaf species prediction.

In order to make the application of the program more portable and convenient, it is necessary to transplant the trained artificial neural network model to the mobile phone terminal. TensorFlow provides TensorFlow Lite for mobile terminal. The trained model can be converted into the model format required by TensorFlow Lite by freezing, and the TensorFlow Lite loading model can be imported and configured on Android mobile phone, so that the leaf shape prediction function can be added on Android mobile phone.

The specific method is to put the converted model file in the assets folder of Android Studio. Configure Android Studio to import TensorFlow Lite. Then call the model in AndroidStudio. TensorFlow Java API opens all required methods through TensorFlowInferenceInterface class. To use the model, first load the libtensorflow_inference.so library and initialize the TensorFlowInferenceInterface object. Then, by using the inferenceInterface object, including calling feed (), run (), fetch (), etc., the image is sent to the neural network model, and the prediction result is obtained. On the mobile phone side, it is also necessary to preprocess the pictures of leaves in the photo or picture library in advance.

TensorFlow Lite helps realize artificial intelligence classification and recognition on Android mobile terminal, and provides more convenient and flexible auxiliary means for plant researchers and enthusiasts.

5. Conclusion

Convolutional neural network in deep learning is widely used in image classification, which has important applications in various fields and can greatly improve classification efficiency. In botany, the classification and identification of leaves are very important for identifying new or rare tree species. Identification of species is of great help to plant evolution research, plant species protection and agricultural development, and accurate acquisition of species types and spatial distribution information is the basis of natural resources utilization and protection[6]. Researchers in the Institute of Botany often go to the field to collect plant leaves. Traditional plant recognition mainly relies on manual recognition, which is limited by environment and communication, and requires very high experience of collectors. Manual recognition limits the accuracy of leaf recognition to a great extent. In this paper, convolutional neural network algorithm is used to identify leaves. Convolutional neural network can realize automatic extraction of leaf image features, reduce tedious labor costs, realize artificial intelligence method to classify leaves, and provide artificial intelligence assistance for botany research. OpenCV is used to preprocess the leaf image, and TensorFlow and Keras are used to build, train and predict the leaf classification model. Combined with TensorFlow Lite, the model loading and prediction are deployed on Android mobile phone terminal, which improves the portability of application, thus providing a convenient and quick artificial intelligence auxiliary means for plant research.

References

[1]   Shuwei Wei,Shangyou Zeng,Yue Zhou,Xinjiao Wang. (2020) Tree Leaves Classification Based on Deep Convolution Neural Network. Modern electronic technology, 09 96-100.

[2]   Muhammad Azfar Firdaus Azlah,Lee Suan Chua,Fakhrul Razan Rahmad,Farah Izana Abdullah,Sharifah Rafidah Wan Alwi. (2019) Review on Techniques for Plant Leaf Classification and Recognition. Computers,Volume 8, Issue 4.

[3]   Febri Liantoni,Rifki Indra Perwira,Syahri Muharom,Riza Agung Firmansyah,Akhmad Fahruzi. (2019) Leaf classification with improved image feature based on the seven moment invariant. Journal of Physics: Conference Series, Volume 1175, Issue 1.

[4]   Xiumei Zhang,Yaqiu Liu. (2006) Recognition system of leaf images based on neuronal network.Journal of Forestry Research, Page 243-246+266.

[5]   Muammer Turkoglu,Davut Hanbay. (2019) Recognition of plant leaves: An approach with
      hybrid features produced by dividing leaf images into two and four parts.Applied Mathematics
      and Computation, Volume 352, 2019. PP 1-14.

[6]   K. Pankaja,V. Suma.(2020)Plant Leaf Recognition and Classification Based on the Whale
      Optimization Algorithm (WOA) and Random Forest (RF).Journal of The Institution of
      Engineers (India): Series B,Volume 101, 2020. PP 597-607.