Journal of Research in Science and Engineering (JRSE)
ISSN: 1656-1996 Volume-7, Issue-3, March 2025
DOI: 10.53469/jrse.2025.07(03).11

# Plant Leaf Recognition based on Mask R-CNN

**Linzhe Zou[1], Dong Wu[2], Liqiong Lu[2,*]**

[1]School of Computer Science and Engineering/School of Software, Guangxi Normal University, Guilin 541004, China
[2]School of Computer Science and Intelligence Education, Lingnan Normal University, Zhanjiang 524048, China
**Correspondence Author, hb_llq@126.com*

**Abstract:** ***Plants, as the main form of life and an important component of human life, are often as the preferred objects for automatic plant recognition research due to their distinct features, prominent internal textures, and significant differences in appearance. Firstly, a plant leaves recognition dataset containing 805 images was constructed. These leaves are classified as 8 different plant species, including blueberry, apple, cherry, grape, strawberry, capsicum, peach and potato. Subsequently, Mask R-CNN was used as a basic method to provide a baseline leaves recognition result. The plant leaves dataset has been made public on Baidu Network Disk and the link is https://pan.baidu.com/s/1YIE-QO25lQIlvQf5QFbKjQ?pwd=mn89.***

**Keywords:** Plant leaf recognition, CNN, Mask R-CNN**.**

## 1. Introduction

Plants, as brilliant treasures of the living world, not only play a fundamental role in the ecological balance of nature, but also play a pivotal role in various fields of human life, especially in the food chain. The in-depth study of plant leaves has immeasurable value in improving crop yield and quality, and meeting the growing food demand of humans. In the broad field of botany, the identification of plant species often relies on features such as leaf shape, color, and texture [1]. Although leaf color is usually green and susceptible to environmental influences, color features are not commonly used in identification due to their instability. In contrast, the shape characteristics of leaves vary significantly among different species, making them a key basis for recognizing plant species. In addition, leaf texture features, specifically referring to leaf veins and the information they contain, can reflect the internal structure of leaves. Due to the similarity of leaf vein structures within species and the large differences between species, texture features are also widely used in plant recognition.

Plant leaf recognition can be classified two kinds of methods. Traditional machine learning methods [2-4] and deep learning based methods. Traditional machine learning methods are often based on shape and texture features to design recognition algorithms. However, these methods often require high requirements for prepossessing, and the designed methods often only perform well on specific datasets. In contrast, deep learning methods can autonomously learn highly recognizable leaf features with sufficient samples, thereby achieving effective recognition of plant leaf images. Therefore, plant leaf image recognition methods based on Convolutional Neural Networks (CNN) are gradually emerging and showing great potential [5-9].

Traditional machine learning methods were proposed in earlier years. Yan Qing et al. proposed a supervised LLE method based on Fisher projection for plant leaf image recognition. This method uses Fisher projection distance to replace the geodesic distance of the sample, and calculates the weight of the sample based on this, which is added to the cost function of the LLE algorithm and the average recognition rate of this algorithm reaches 92.36% [2]. Aiming at the problem that low dimensional features cannot fully describe leaf information and have low recognition accuracy, a plant leaf recognition method based on multi feature dimensionality reduction is proposed. The experimental results show that using the trained support vector machine classifier to classify and recognize test leaf samples from Flavia database and ICL database, the average correct recognition rates are 92.52% [3]. A leaf recognition algorithm based on Complex Frequency Domain Texture Features (CFDTF) and KNN classifier is proposed to address the issue of incomplete and inaccurate description of leaves by spatial domain features and he average recognition accuracy of this algorithm is significantly improved, reaching 95.75% [4].

Deep learning plant leaf recognition methods are also researched widely. In 1991, Heymans et al. [5] chose to use neural network algorithms for the recognition of cactus leaf images based on the advantages of backpropagation paradigm, which was a highly prospective study. Mokhtarian et al. [6] innovatively improved the traditional curvature scale space algorithm for feature recognition of 12 different chrysanthemum leaf images in their study and the experimental results showed that the recognition rate of this method exceeds 98%, providing an efficient and accurate method for identifying chrysanthemum leaf images. To improve the accuracy of plant leaf recognition and reduce computational costs, a novel convolutional neural network leaf recognition method was proposed under the Pytorch framework, which combing deep convolutional generative adversarial network (DCGAN) and transfer learning (TL) and the results showed that this method can achieve a plant leaf recognition accuracy of 96.57% [1]. An improved YOLOv8 model was proposed based on a dataset of fresh leaves from five tea-plant varieties among Yunnan large-leaf tea trees. This method used dynamic Upsampling to replace the UpSample module in the original YOLOv8, reducing the data volume in the training process, the Efficient Pyramid Squeeze Attention Network was integrated into the backbone of the YOLOv8 network to boost the network’s capability to handle multi-scale spatial information and a Spatial and Channel Reconstruction Convolution module was introduced [7]. A Convolutional neural network (CNN) was employed with and without data augmentation, in addition to a DCNN Classifier model based on VGG16, to classify grape leaf diseases. The DCNN Classifier Model successfully utilized the strengths of the VGG16 model and modified it by incorporating supplementary layers to enhance its performance and ability to generalize. Systematic evaluation of metrics, such as accuracy and F1-score, was performed [8].

This paper firstly constructed a plant leaves recognition dataset where 805 plant leaf images and the corresponding label files (.JSON files) are displayed. These images are classified into eight species including blueberry, apple, cherry, grape, strawberry, capsicum, peach and potato. Then, we divided the dataset into a training set and a testing set. 80% randomly selected images are in the training set, the remaining 20% images are in the testing set. Lastly, Mask R-CNN is used to trained on the training set to get the plant leaf recognizing model and the model is tested on the testing set to get a baseline recognizing result.

## 2. Plant Leaves Recognition Dataset

A dataset of plant leaves containing 805 images was constructed. These plant leaf images come from two sources, some plant leaves are downloaded from the Internet, the other images are directly taken from the field manually. The dataset includes eight species of leaves images: blueberry, apple, cherry, grape, strawberry, capsicum, peach and potato. The plant leaves dataset has been made public on Baidu Network Disk and the link is https://pan.baidu.com/s/1YIE-QO25lQIlvQf5QFbKjQ?pwd=mn89.

Figure 1 lists the sample images of plant leaves in this paper. These leaf images have different textures, shapes, and leaf flatness, all of which can be used as features for recognition. Some plants have similar leaf shapes, such as blueberry Leaf, peach leaf and cherry leaf, apple leaf and potato leaf. In addition, the lighting of different parts of the leaves, the similarity of leaf textures, and the complex background deepen the difficulty of identifying leaf types and positions.

This paper calculates the total number of samples of various plant leaves in the dataset. The statistical results are shown in Figure 2. From Figure 2, it can be seen that there are 70 images of seaweed and soybeans each, while the images of other types are around 100. Overall, the images corresponding to each plant leaves are relatively balanced.

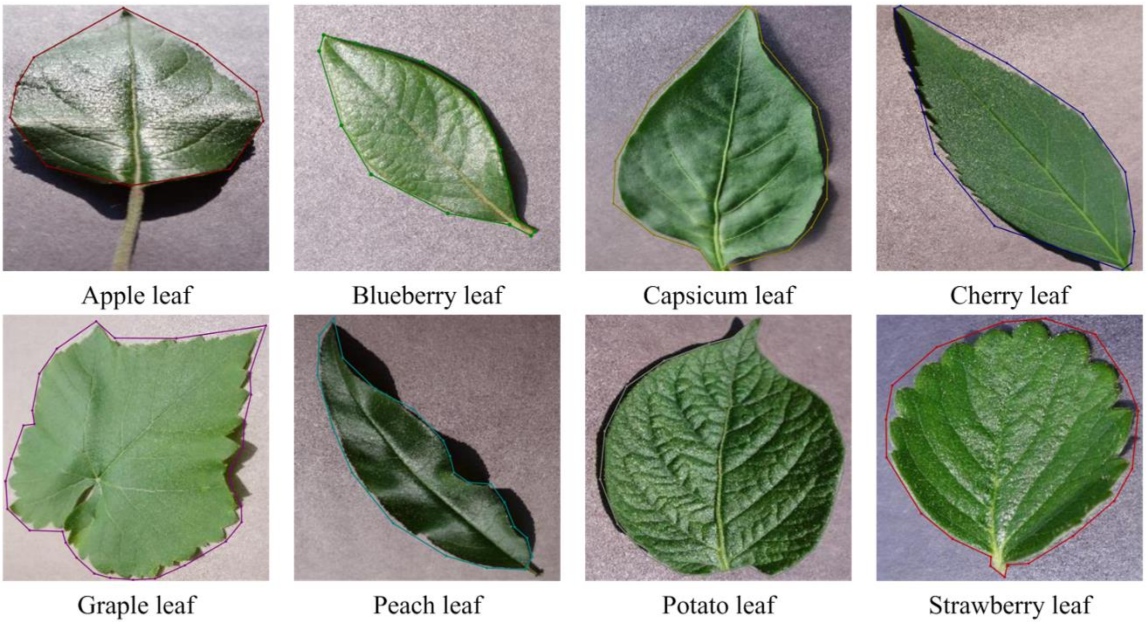


**Figure 1:** Sample images of plant leaves

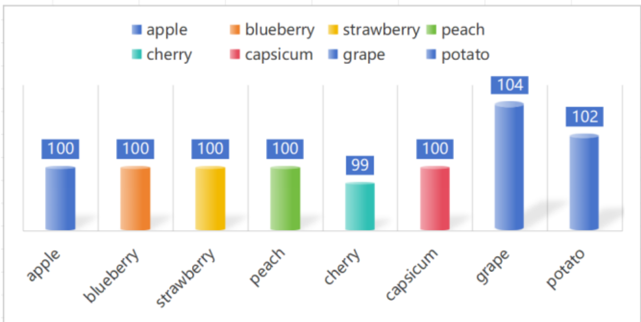


**Figure 2:** Statistics of plant leaves in the dataset

## 3. Plant Leaf Recognition Based on Faster Mask R-CNN

This paper chooses Mask R-CNN algorithm as the core technology for accurate recognition and segmentation of plant leaves. Mask R-CNN is an advanced deep learning algorithm that combines object detection and semantic segmentation tasks to generate high-quality segmentation masks for each object instance in an image.

All model training experiments in this paper were conducted on a local computer running Windows 10. The machine is equipped with an Intel Core i7-10875H CPU with a clock speed of 2.30GHz, ensuring efficient data processing capabilities. In terms of memory, a capacity of 16GB was adopted to meet the memory requirements during the model training process. In terms of graphics processing, the machine is equipped with NVIDIA GeForce GTX 1650 graphics card, which has 8GB of video memory and provides powerful computing power for graphics intensive tasks.

### 3.1 Experimental Set

In this paper, Mask R-CNN algorithm is implemented based on Tensorflow 2.4. Max epoch is set to 10, initial learning rate is set to 0.001, and momentum is set to 0.9 by default. The training takes about 2 hours to run, and the segmentation after training takes a few seconds. Figure 3 shows the changes of loss of training and testing data as the epoch number increases during the training process. It can be seen that the loss converges around 0.1.

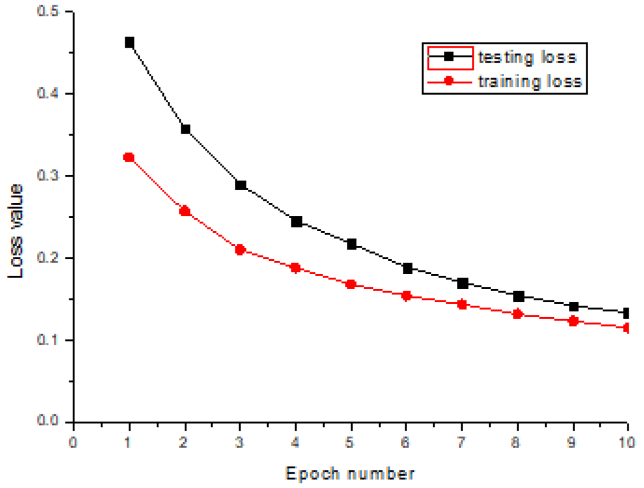


**Figure 3:** Changes of loss corresponding to epoch number increasing

### 3.2 Performance Evaluation Indicators for Plant Leaf Recognition

This paper uses accuracy (Precision, P), regression rate (Recall, R) and comprehensive indicators *mAP* (mean Average Precision) to evaluate the plant leaf recognition. The specific calculation methods for P and R are shown in the following formula.

$$P = \frac{FP}{FP+TP};$$

$$R = \frac{TP}{TP+FN};$$

In the above formula, TP is the number of correctly identified samples; FP is the number of error identified samples; FN is the number of undetected samples. Here, when the IOU between the recognition box and the ground truth box exceeds 90%, it is considered correct recognition. *mAP* is the average of the average Precision values for all categories, calculated based on the Precision Recall curve. It interpolates Precision at different recall rates and calculates the area under the interpolation curve.

### 3.3 Plant Leaf Recognition Results

Table 1 lists the plant leaf recognition results based on Mask R-CNN. It can be seen from Table 1 that the recognition results of all plant leaves are relatively good, with an mAP value of 0.894. Especially, grape has the highest AP value with 0.98. Strawberry, peach, cherry, capsicum all have a relatively high value in AP above 0.9. Apple, blueberry and potato has a lower value in AP below 0.85.

**Table 1:** plant leaf recognition performance

| *Type* | *P* | *R* | *AP (mAP)* |
|---|---|---|---|
| apple | 0.8 | 1 | 0.83 |
| blueberry | 0.6 | 1 | 0.64 |
| strawberry | 0.9 | 1 | 0.91 |
| peach | 0.95 | 1 | 0.95 |
| cherry | 0.9 | 1 | 0.93 |
| capsicum | 0.9 | 1 | 0.92 |
| grape | 0.95 | 1 | 0.98 |
| potato | 0.7 | 0.93 | 0.81 |
| all | 0.843 | 0.993 | 0.894 |

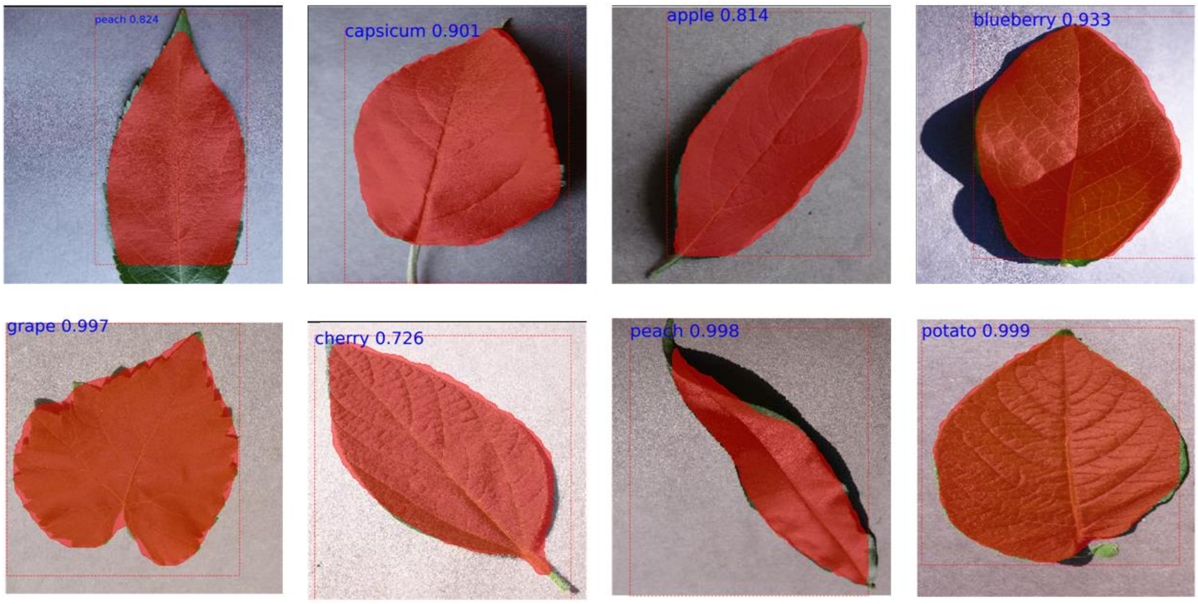


**Figure 4:** Successful plant leaf recognition samples.

Overall, after using the Mask R-CNN method for recognition, the accuracy and regression rate of all plants recognition are relatively high. Analyzing the regression rate, apple, blueberry, strawberry, peach, cherry, capsicum and grape all have the value of 1 in regression rate; potato has a relatively lower regression rate value. Analyzing the accuracy rate of all plants leaves recognition, the highest accuracy rate is achieved by peach and grape, reaching 0.95; strawberry, cherry and capsicum have a relatively high value above 0.9. While the accuracy rate of blueberry, apple and potato are not very ideal.

### 3.4 Sample Analysis of Plant Leaf Recognition

Successful leaf recognition samples and unsuccessful leaf recognition samples are displayed in Figure 4 and Figure 5.

Seen from Figure 4, the vast majority of plant leaves can be accurately identified including prediction of plant leaf types, positions, and leaf pixel levels. It can be seen from Figure 5, with a few plant leaves are recognized as multiple types, especially when the leaf area is relatively large, with shadows and unclear texture.

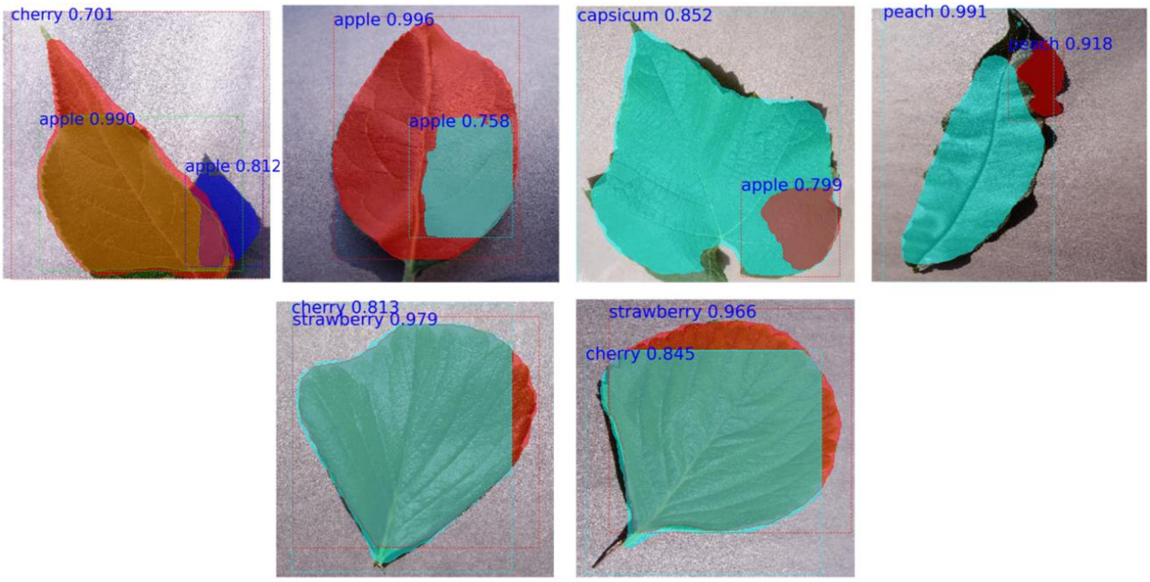


**Figure 5:** Unsuccessful plant leaf recognition samples.

## 4. Conclusion

This paper focuses on plant leaf recognition. Firstly, a dataset containing 805 images was constructed for plant leaf recognition with 80% images used for training and the remaining 20% used for testing. Then, Mask R-CNN was used to recognize these plant leaves and achieved an *mAP* of 0.894. From the results of plant leaf recognition, blueberry and potato has a relatively lower accuracy rate of recognition. Subsequent analysis can be conducted on the characteristics of the leaves of these two plants to seek methods to improve its recognition accuracy.

## Acknowledgments

The authors are grateful to the anonymous reviewers and the helpful suggestion given by the partners. The research was supported by the Technology Project of Zhanjiang (no. 2022A01005), the Guangdong province philosophy and social science planning project (no. GD24CJY21).

## References

[1] ZHU L K, YAN M, Huang J P. A novel convolutional neural network method for plant leaf recognition[J]. Journal of Northeast Forestry University, 2020, 48(04): 50-53.

[2] Yan Q, Liang D, Zhang J J. Supervised LLE algorithm for plant leaf image recognition based on Fisher transform[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(09): 179-183.

[3] Zheng Y L, Zhong G L, Wang Q. Plant leaf recognition method based on multi feature dimensionality reduction[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(03): 30-37.

[4] Mei X Y, Li X H, Bao W X. Plant leaf recognition algorithm based on complex frequency domain texture features[J]. Jiangsu Journal of Agricultural Sciences, 2019, 35(06): 1334-1339.

[5] Mokhtarian F., Abbasi S.. Matching shapes with self-intersections: application to leaf classification. [J]. IEEE Trans Image Process, 2004, 13(5): 653-661.

[6] Carranza-Rojas J., Mata-Montero E.. Combining leaf shape and texture for Costa Rican plant species identification[J]. CLEI Electronic journal, 2016, 19(1): 7-7.

[7] Wang C, Li H Y et al. Improved You Only Look Once v.8 Model Based on Deep Learning: Precision Detection and Recognition of Fresh Leaves from Yunnan Large-Leaf Tea Tree[J]. Agriculture, 2024, 14(12): 2324-2324.

[8] Kerehalli Vinayaka Prasad, Hanumesh Vaidya et al. Multiclass classification of diseased grape leaf identification using deep convolutional neural network (DCNN) classifier[J]. 2024, Scientific reports, 14(1): 9002-9002.