Blueberry Gray Mold Disease Recognition based on YOLOv5

1st Dianfeng Li
University of Electronic Science and Technology of China
Chengdu, China
ldfjiayou@163.com

2nd Wenchen Liu
University of Electronic Science and Technology of China
Chengdu, China
812288728@qq.com

***Abstract*—China is a large agricultural country, and agriculture has long been the mainstay of China's national economic progress and prosperity, and one of the most important industries in China. However, the problem of pests and diseases of agricultural products has continued to affect the continuous progress of agriculture for a long time, and every year will cause great losses due to the problems of pests and diseases of agricultural products. Deep learning technology can be used to identify plant diseases, which can not only improve the accuracy and speed of identification, but also contribute to the digital development of agriculture.**

This paper will focus on the blueberry stiff fruit disease, that is, the normal and diseased conditions of the leaves, flowers and fruits of the blueberry plant as the research object. The method of disease identification and the more diverse applications of convolutional neural networks in agriculture. The main research contents are as follows:

1)Through Gaussian blur, color enhancement, histogram normalization, geometric change, and gamma transformation, image preprocessing is performed on the existing original data set to expand the original data set.

2)Use the Make Sense image annotation tool to perform feature annotation on the expanded dataset to complete the dataset preprocessing.

3)According to the YOLOv5 algorithm, a convolutional neural network model was learned and designed to identify blueberry stiff fruit disease. The model was trained with 4,200 blueberry images, which were divided into 6 categories of normal and infected leaves, flowers and fruits. 1800 blueberry images are used as validation set. The lowest accuracy of the trained model is 92.6% for flower normal, and the highest accuracy is 99.5% for flower infection.

***Index Terms*—blueberry grad mold disease, convolutional neural networks, deep learning, image analysis, object classification**

I. Introduction

A. Background

Deep learning is an important technology that has brought us strong development. Convolutional neural networks (CNNs) as a typical deep learning algorithm have been widely used in high-tech industries [1]. CNNs have many applications, such as single-frame image recognition, video processing, and natural language understanding. Since the mid-1990s, CNNs have been continuously integrating theory and practice, but they have not caused an academic storm. In the early 21st century, CNNs had been quiet for more than ten years. However, in 2012, AlexNet's amazing performance made CNNs popular again [2].

China is a traditional agricultural country. Although China has developed unprecedentedly in various industries since the reform and opening up, agriculture has also developed unprecedentedly. However, as an agricultural country, China still has a big gap with other countries in the world.

Although China has accumulated a lot of agricultural production experience and achieved remarkable results, there is still a big gap between China's agricultural intelligence level and information level. Especially scientific and technological innovation is a weak link in China's agricultural development. How to apply deep learning technology to agriculture has become a research hotspot in the scientific community.

B. Summary of Research

Blueberry fruit rot is one of the most common and serious diseases in blueberry production [3]. Once blueberries are infected, the plants will suffer from a reduction in yield of about 40%, which will have a significant impact on the yield and quality of blueberries. After overwintering, if the fallen fruit is not cleaned up in time, its spores will germinate again and cause cyclic infection in the second year.

Fruit rot is a disease caused by fungi. In the early stage of the disease, mature spores of the pathogen will germinate on the surface of new leaves or flowers. The hyphae will develop on the surface cells of flowers and leaves, causing cell rupture and death, resulting in sudden wilting and browning of leaves, buds, flowers and other positions. After a period of time, fungal spores will cover leaves, stems and flower columns like powdery substances. At the same time, they will spread like open flowers and carry out secondary transmission. Finally, it will cause fruit wilting, dehydration and drying, showing a stiff state [4].The disease and health status of different parts of blueberry are shown in Fig 1.

II. Disease Data Image Processing

The original data often cannot fully meet the requirements of the test and can not be directly used in the test. The correct processing of the original data will promote the success of the test to a certain extent. This chapter will briefly introduce the original data of the test, and carry out statistical analysis on the original data, find out the problem, solve the problem, and achieve the data needed for the test.

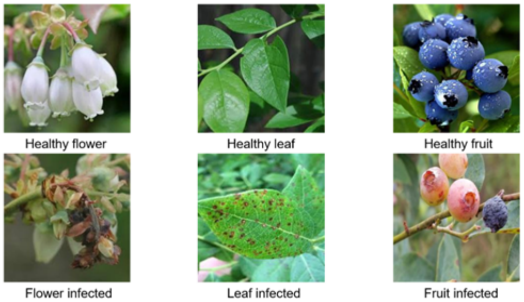

Fig. 1. Comparison chart of disease and health status in different parts.

TABLE I
THE NUMBER OF IMAGES IN THE ORIGINAL DATASET

Category	Blade	Flowers	Fruit	Total
number of normal images	45	83	56	184
number of disease images	123	76	41	240
total	168	159	97	424

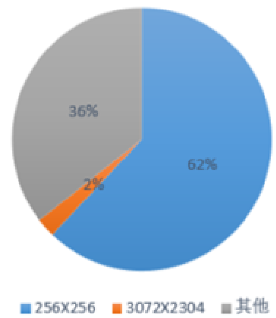

Fig. 2. Original image size distribution

A. Datasets introduction

This article analyzed six types of normal and diseased blueberry plant organs, including leaves, flowers, and fruits, with a total of 424 basic images. The number of images of the three organs of blueberry plants, leaves, flowers, and fruits, which are normal and invaded by diseases are shown in Table I.

B. Data Preprocessing

1) Image Data Analysis: The statistical analysis of the original data shows that:

1.The names of the image data are arbitrary, and there is no standard image naming format to determine the classification of images.

2.The total amount of information in the image data is very small, and there are significant differences between the image data of each category.

3.The size of the images is inconsistent. In the original dataset, 264 images have a size of 256×256, 10 images have a size of 3072×2304, and other images have different sizes. The proportions of each image size in the original dataset are shown in Fig 2. From Fig 2, it can be seen that mainly 256×256 and 3072×2304 sizes occupy 62% and 36% of the entire dataset respectively.

2) Image Normalization: The first part is the standardization of image naming [5]. The images are renamed in the form of “string + underscore + number”, where Hl, Hi, Hfl, Fli, Hfr, Fri are six categories of healthy leaf, leaf infected, healthy flower, flower infected, healthy fruit and fruit infected respectively. The underscore is just a separator, and the number is an image number. In addition to this, as the data is strengthened, the name of the image will also be extended according to this naming format.

The second part is the standardization of image size. Since the original images are all based on blueberry flowering and fruiting as the background, most of the effective images are located in the center of the original images. Around the original images are some invalid images. While removing some invalid data, the original images are center-cropped to obtain an image size of 224×224 and display the cropping effect shown in Fig 3.

3) Image Enhancement: A total of 424 color images were used in this experiment, including six types of normal or diseased organs of blueberry plants: flowers, leaves, and fruits. Based on this, when building a convolutional neural network model, the first thing to note is that convolutional neural networks use big data and calibration monitoring mechanisms to learn models. Sufficient training data is a prerequisite for achieving deep learning and data mining capabilities. However, it is not realistic to rely solely on collecting original data for data collection due to the time-consuming, laborious and expensive nature of collecting blueberry flower, leaf and fruit data. The amount of information currently available is not large enough to be applied to large-scale convolutional neural networks.

Secondly, in multi-classification problems, there is a significant causal relationship between the training results and the training results of convolutional neural networks. In the original data used in this experiment, there was an imbalance in the data. For example, there were only 45 normal leaf classifications compared to 123 abnormal leaf classifications. The sample ratio was imbalanced and could easily cause model training bias. Based on this, data enhancement was performed on different types of images, and modified data storage probabilities were used to solve the uneven distribution of sample data and achieve balanced sample numbers for each type.

Due to the lack of training samples, it is impossible to meet the training data of convolutional neural networks, and the number of samples in the data set is also imbalanced. It is urgent to expand the training samples of the model. However, it is difficult to collect samples manually, so existing data must be used for enhancement and expansion to ensure the promotion of the model [6]. Although there are many methods to improve data performance, in practical applications, the choice of data enhancement algorithm should be adapted to the actual situation. Choosing the appropriate method can

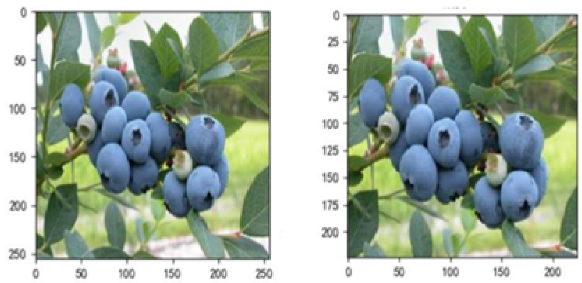

Fig. 3. Comparison chart before and after image size normalization

play a multiplier effect. After offline enhancement of different data enhancement methods, this article has improved in image frequency domain, color and geometric space.

4) Image Blur: Gaussian blur technology is an image blur technology. By using this technology to process images, noise in the image can be effectively reduced, and the detail level in the image can be effectively reduced [7]. The principle of the Gaussian blur algorithm is similar to that of image convolution operation. The value of each pixel is obtained by multiplying the value of the pixel itself and the pixel value in its area, just like a normal distribution convolution kernel. It is a convolution kernel centered on a pixel point, and its pixel value is determined by itself and the pixel value in its range.

Gaussian blur technology is a commonly used method in image denoising. It can suppress or eliminate the image noise caused by normal distribution. The size of the Gaussian blur convolution kernel, that is, its neighbor size, can control the degree of image blur. By continuously adjusting the size of the convolution kernel, the experimental data can be strengthened by Gaussian blur.

5) Histogram Equalization: Histogram normalization is a technique that adjusts the gray-level histogram of an image using a technical method to make its distribution more uniform and standardized [8]. In general, the gray-level histogram of an image has non-uniformity, and the histogram has non-linear characteristics. It reduces the gray level of high-frequency bands and increases the gray level of low-frequency bands. This can make the gray distribution more uniform. From the histogram point of view, the histogram on the histogram is flatter. From the data point of view, this is a nonlinear stretching transformation.

Since histogram normalization can increase the number of low-frequency gray levels, it can improve the contrast of the entire image and significantly improve the visual quality of the image. Histogram normalization can selectively enhance local and global characteristics of images, strengthen specific effective areas, suppress low-contrast areas, and use histogram normalization to improve detail expression in images.

6) Data Division: After preprocessing, the data set undergoes data enhancement. Each category of data contains about 1050 images, and finally 1000 images are selected as the training set and validation set data of the model, which are divided into 7:3. Fig 4 shows a specific segmentation process.

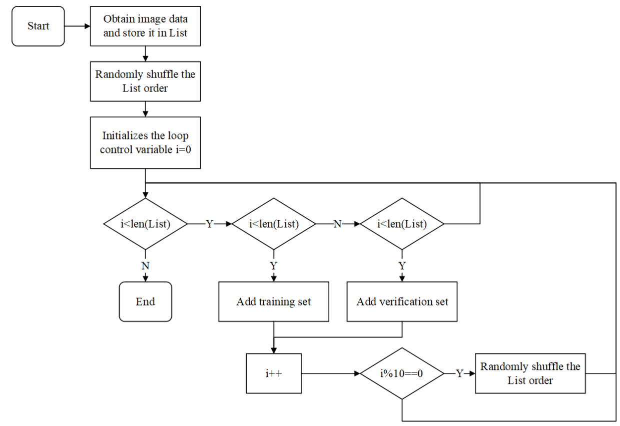

Fig. 4. Flow chart of model training set and test set data division

Because before segmenting the data, you first need to obtain the name of the image. Secondly, when segmenting, the names of the images are arranged in order. If you take the first 700 pictures in the list as training samples and the remaining 300 pictures as verification samples, it will cause insufficient differences between samples and low recognition and classification efficiency of training samples. Therefore, before segmentation, the collected image names are randomly sorted and stored in the training set or test set [9]. Every 10 rounds, the image name list is randomly shuffled again to effectively avoid problems caused by insufficient differences in training samples. The segmentation effect is shown in Fig 4.

III. Realization of Blueberry Gray Mold Recognition System Based on CNN

Due to the differences between different image data, the traditional CNN also has great differences in the ability to classify and recognize different images. In this article, the YOLOv5 algorithm was adopted, and partial pruning of the classical YOLOv5 network was carried out according to the characteristics and categories of the identified object blueberry plants, so as to make it more adaptable and the constructed model more lightweight, so as to realize the recognition system of blueberry gray mold disease.

A. Yolov5s neural network model

In traditional convolutional neural networks, it is generally through increasing the depth of the network to improve training efficiency. Although the method of adding network hierarchy can improve training efficiency, it will also cause a large number of parameters to be too many, gradient disappearance, overfitting, computational complexity and other negative effects. This article uses a shallow depth and lightweight YOLOv5s network model.

YOLOv5 is a single-stage object detection method based on the YOLOo series. It is an efficient implementation method based on YOLOv5. The YOLOv5 network model includes four types: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. YOLOv5s is the shortest one in YOLOv5, and the latter

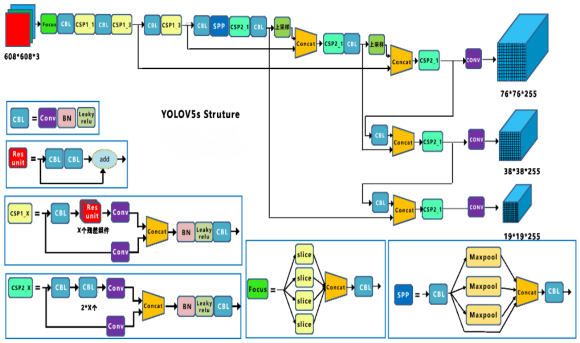

Fig. 5. YOLOv5s network structure diagram

three models are products of further deepening and expanding YOLOv5s. Due to the small size of the network model and low system performance requirements, it is easy to deploy. Therefore, this article prunes and optimizes its network model based on YOLOv5s. The structure of YOLOv5s can be divided into four modules: Input, Backbone, Neck (multi-scale feature fusion module), Prediction. Its network structure is shown in Fig 5.

In the input part, YOLOv5s first uses the IMlosaic data augmentation technology to achieve arbitrary cropping, scaling, and arrangement of images, thereby enriching the background of objects being detected. At the same time, when performing large-scale standardized operations, data processing of four images at a time can be used to obtain better results with a single GPU, improving the universality of the network. Secondly, adaptive anchor box structure is used to determine the initial anchor box parameters, and then the network automatically calculates the corresponding prior box parameters according to different data. Finally, adaptive scaling of images is used to achieve uniform scaling or magnification of images.

Compared with the YOLO series network model, YOLOv5s adds the Focus architecture backbone network to segment images as the center. Later, YOLOv5 adopted a CSP (cross partial network) architecture similar to YOLOv4. YOLOv4 only uses CSP structure in Backone, while YOLOv5 uses both CSPs in Backbone and Neck. In Backone, CSP1_X with residual structure is used. Due to the deep depth of the Backone network, adding residual structure improves the gradient between layers, which can effectively avoid gradient disappearance caused by network depth and obtain more fine-grained features. In Neck, module CSP2_X is used to divide the output of the main network into two parts and then merge them to enhance the feature fusion ability of the network while retaining more feature information.

Multi-scale feature fusion is an important step in target recognition. Early Neck used two sampling blocks up and down, which does not require feature layer aggregation operations like SSD, but directly tracks multiple layers of features. Currently, FPN, PAN, ASFF, BiFPN are widely used. The Neck of YOLOv5s adopts FPN (feature pyramid network) + PAN (pyramid attention network) storage structure.

B. Datasets feature labeling

Before model training, the collected data must be labeled. In this article, image labeling technology of Make Sence is used to label data sets. The key of feature marking is to carry out special labeling for data images, and the processing procedure is shown in Fig 6.

C. Experimental model training

The deep learning framework used in this experiment is Pytorch, the network model is YOLOv5s, and the training platform is Google Cloab. The size of the images used in training and testing is 224×224, with 700 pictures for each training set, totaling 4200. The check set consists of 300 images, 1800 in total.The whole experimental flow chart is shown in Fig 7.

In this article, torch_utils.prine() function is used to prune the YOLOv5s model in the val.py file to achieve 30% sparsity, that is, 30% of the model's weight parameters in the layer are equal to 0.

The default batchsize is 16. During the pre-training of the model, it is found that the model has better convergence, so the training cycle of the model is set to 100 minimum values. In terms of optimization algorithm, we use a relatively common stochastic gradient descent algorithm (SGD). The cross entropy loss function is used to calculate the model loss. In terms of excitation function, ReLU is used in all models, which reduces the dimension of data and the amount of calculation.

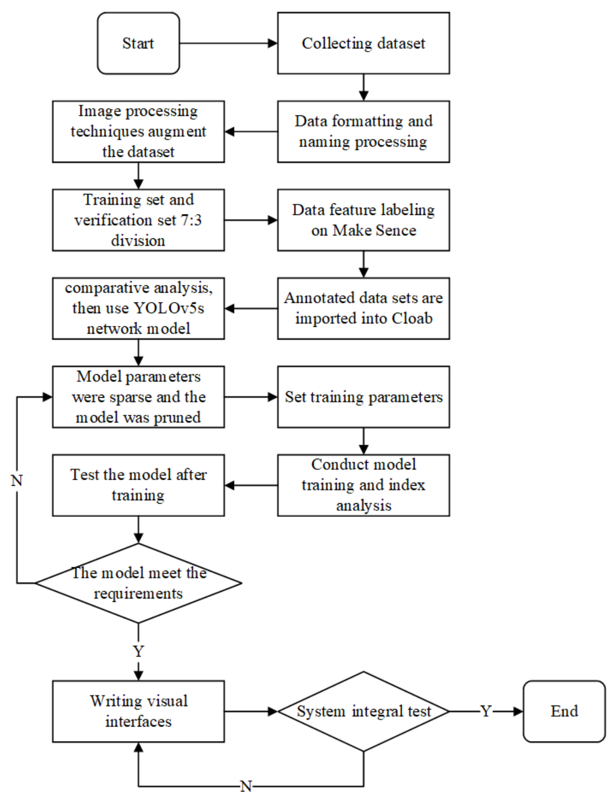

Fig. 6. Experimental flow chart

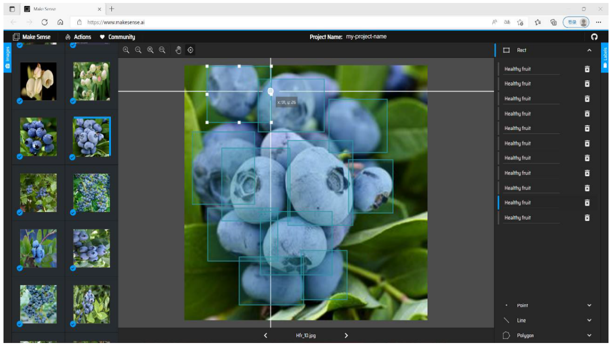

Fig. 7. Feature labeling on Make Sence

IV. Experimental results and model evaluation

In the training and training of the model, we are concerned about the loss of training and the accuracy of identification. However, evaluating the quality of a model cannot completely rely on two factors. In order to better test the effectiveness of the model, we also need to evaluate from multiple aspects. This chapter will show the results of the experiment and the relevant model evaluation indicators, and verify the indicators of the model according to the corresponding index calculation method.

A. Experimental result

Under the training of 100 epochs, the average accuracy of the model reaches 85.5%, meeting the requirements of the model. The model can be tested according to the prediction of the test set and related test results, and provides specific data according to the corresponding evaluation system. See the following for specific evaluation indicators.

The average accuracy of the model reached 85.5%, 88.4%, 80.1%, 91.5%, 73.1%, 93.3%, and 86.6% of the average accuracy of flower anomaly, fruit anomaly, fruit anomaly, and leaf anomaly.

During the training iteration, the model's performance during the training was analyzed by analyzing the numerical change of the model's loss function and checking the prediction accuracy of the test data. Relevant changes are shown in Fig 8.

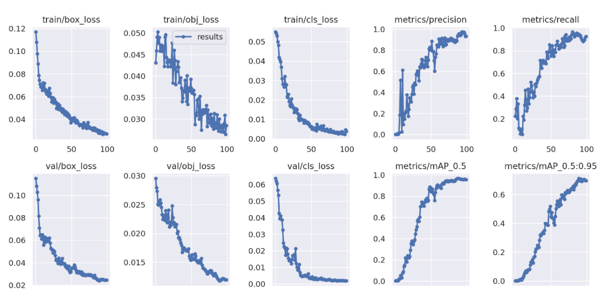

Fig. 8. Model training cycle loss function value and test average accuracy curve

As shown in Fig 8, during the training of the model, the value of the loss function shows a downward trend on the whole. After 70 rounds of training, the loss value is almost 0. At the same time, in the 20 rounds of training, the prediction accuracy of the model has been improved on the whole. After 80 rounds of training, the average accuracy is about 0.96. Both the reduction of training loss value and the accuracy of the test remain constant, which indicates that the performance of the model is improved through repeated training and parameter updating. Moreover, the two curves of the training cycling-loss function and the average accuracy of the training cycling-test tend to be horizontal at last, which indicates that the model has no over-fitting problems.

B. Model evaluation

Fig 9 shows the fuzzy matrix obtained from the convolutional neural network model.

In the training of the model, the precision, recall rate and PR chart index of the model were also analyzed and shown in Fig 10-12.

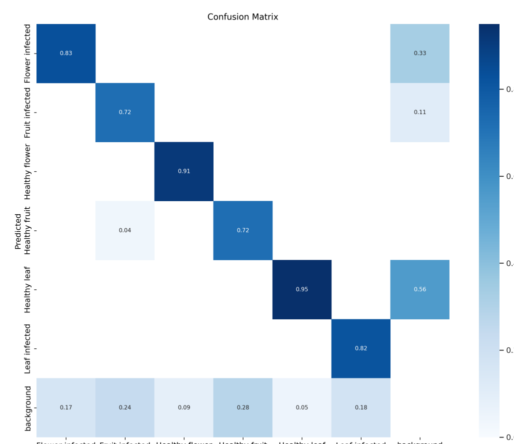

Fig. 9. Diagram of the confusion matrix generated by the model on the verification set

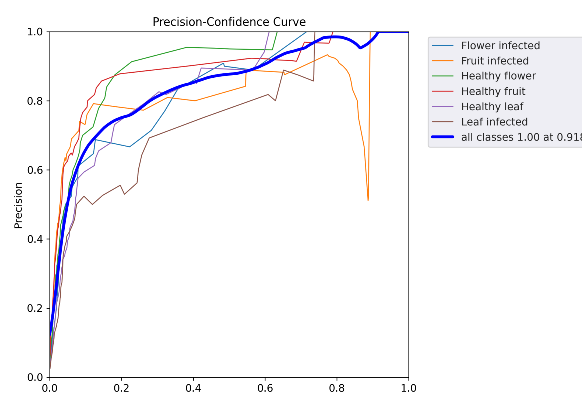

Fig. 10. Model training precision

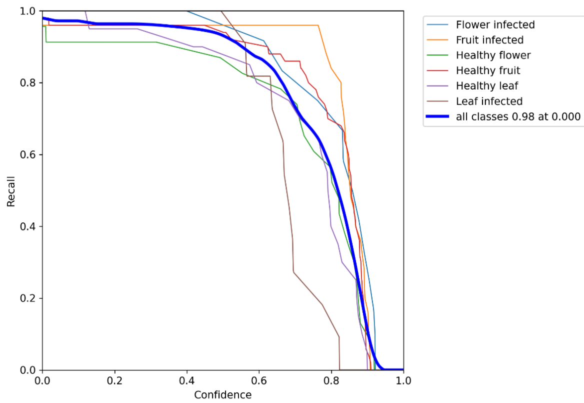

Fig. 11. Model training recall rate

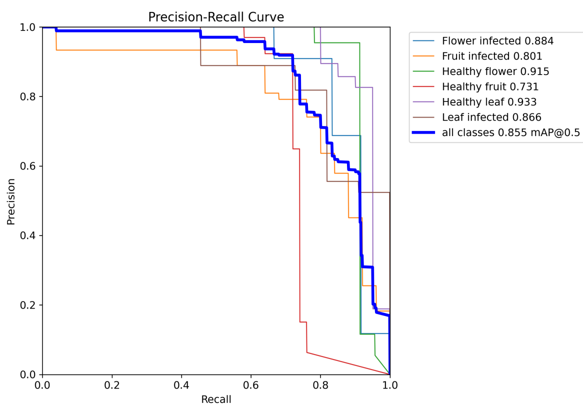

Fig. 12. Model training PR chart index

V. Summary and Prospect

A. Main work and innovation points

This article extends the image data from frequency domain, color and geometric space, and extends the inadequate and unbalanced image set to sufficient and balanced. Partial pruning of the YOLOv5s network model is carried out according to the characteristics and categories of blueberry plants of the identification object, so that it has better adaptability, and the constructed model is more lightweight, so as to realize the recognition of blueberry gray mold disease.

This model uses the YOLOv5s network structure. In the disease identification of flowers, leaves and fruits of blueberry plants, after 100 rounds of training, the recognition accuracy of this model reached 85.5%, which is good for all types of recognition accuracy and recall rate.

The above results indicated that the convolutional neural network YOLOv5 applied in this model could be used for the identification of flower, leaf and fruit diseases of blueberry plants.

B. Prospect of follow-up research work

In this article, the CNN was designed to identify the gray mold disease of blueberry.

First, the number of data sets used is insufficient. Even if the data set is expanded by image enhancement technology, there will still be the problem of similar data set features, resulting in slight overfitting phenomenon in model training.

Secondly, the neural network used is relatively complex, so it can be improved and tailored according to the actual situation to adapt to the actual application scenario.

Subsequently, more data sets can be collected, and the convolutional neural network model can be improved and tailored to optimize the model, so as to achieve better recognition effect.

References

[1] A.Krizhevsky, I.Sutskever and G.E.Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Communications of the ACM, vol.60, 2017, pp.84-90.

[2] Alex Krizhevsky,Ilya Sutskever,Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM,2017,60(6).

[3] Milholland R. D.. Blueberry Fruit Rot Caused byPhomopsis vaccinii[J]. Plant Disease,1983,67(3).

[4] Schilder A.M.C.,Gillett J.M.,Woodworth J.A.. THE KALEIDOSCOPIC NATURE OF BLUEBERRY FRUIT ROTS[J]. Acta Horticulturae,2002(574).

[5] J. L. Agraz et al., "Robust Image Population Based Stain Color Normalization: How Many Reference Slides Are Enough?," in IEEE Open Journal of Engineering in Medicine and Biology, vol. 3, pp. 218-226, 2022.

[6] M. S. Imtiaz and K. A. Wahid, "Image enhancement and space-variant color reproduction method for endoscopic images using adaptive sigmoid function," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 2014, pp. 3905-3908.

[7] K. Bahrami, A. C. Kot, L. Li and H. Li, "Blurred Image Splicing Localization by Exposing Blur Type Inconsistency," in IEEE Transactions on Information Forensics and Security, vol. 10, no. 5, pp. 999-1009, May 2015.

[8] Y. Chen, M. Yang and L. Zhang, "General Data Mining Model System Based on Sample Data Division," 2009 Second International Symposium on Knowledge Acquisition and Modeling, Wuhan, China, 2009, pp. 182-185.

[9] S. Patel and M. Goswami, "Comparative analysis of Histogram Equalization techniques," 2014 International Conference on Contemporary Computing and Informatics (IC3I), Mysore, India, 2014, pp. 167-168.