International Core Journal of Engineering Volume 7 Issue 2, 2021
ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

A Flower Classification Study based on SVM and VGG16

Jize Bi

Computer Science and Engineering college, Northeastern University, Shenyang, Liaoning, 100167, China.

Abstract

In modern society, AI technology has become more and more mature. Many intelligent algorithms, such as traditional machine learning and deep learning, have helped humans complete complex tasks in many ways. In this context, this article uses traditional machine learning and deep learning methods to implement a flower category recognition system. In this system, the user inputs a picture of a flower, the system extracts its features, and uses the trained model to complete the classification. The prediction result using the svm model can reach 65%, and the accuracy rate of the model using VGG16 can reach 85%. The implementation of this system represents an attempt to solve complex tasks using machine learning.

Keywords

SVM; Machine Learning; VGG16; Flower Classification.

1. Introduction

In modern society, the application of machine learning has become more and more widespread. In people's daily life, machine learning can already take on many tasks. For example, inferring users' preferences based on their operating information, unmanned driving, etc. In the work, it uses the traditional machine learning method-svm model, and the deep learning model-VGG16 to train a computer model that can recognize 21 kinds of flowers. The accuracy of the Svm model reached 65%, and the accuracy of the VGG16 model reached 85%. In the work, compared with the use of the svm model, the use of vgg16 deep learning for prediction has achieved a great improvement in speed and accuracy.

2. Data getting and Pretreatment

Data getting module: the main function is to get the flower images by writing python crawler programs, then divide these images into 21 types and put them into 21 respective folders. The main source of these images is google and baidu searching engine. Each type has 500 images. I divide the total over 10000 pictures into training set and test set. In other words, the ratio of training set to test set is 4 to 1. Because every picture has different size, which means that difference might affect the latter training process, I cut them into uniformly 224*224. That is because the standard image input size of the VGG16 is also 224*224. After cutting, I put the name and label information of these pictures in a .csv file. This file is going to be used in latter debugging and programming.

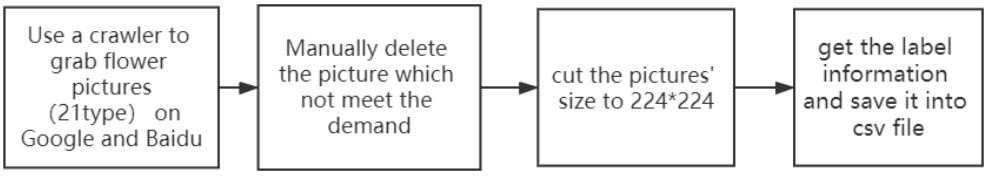


Figure 1. Data getting and Pretreatment process

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

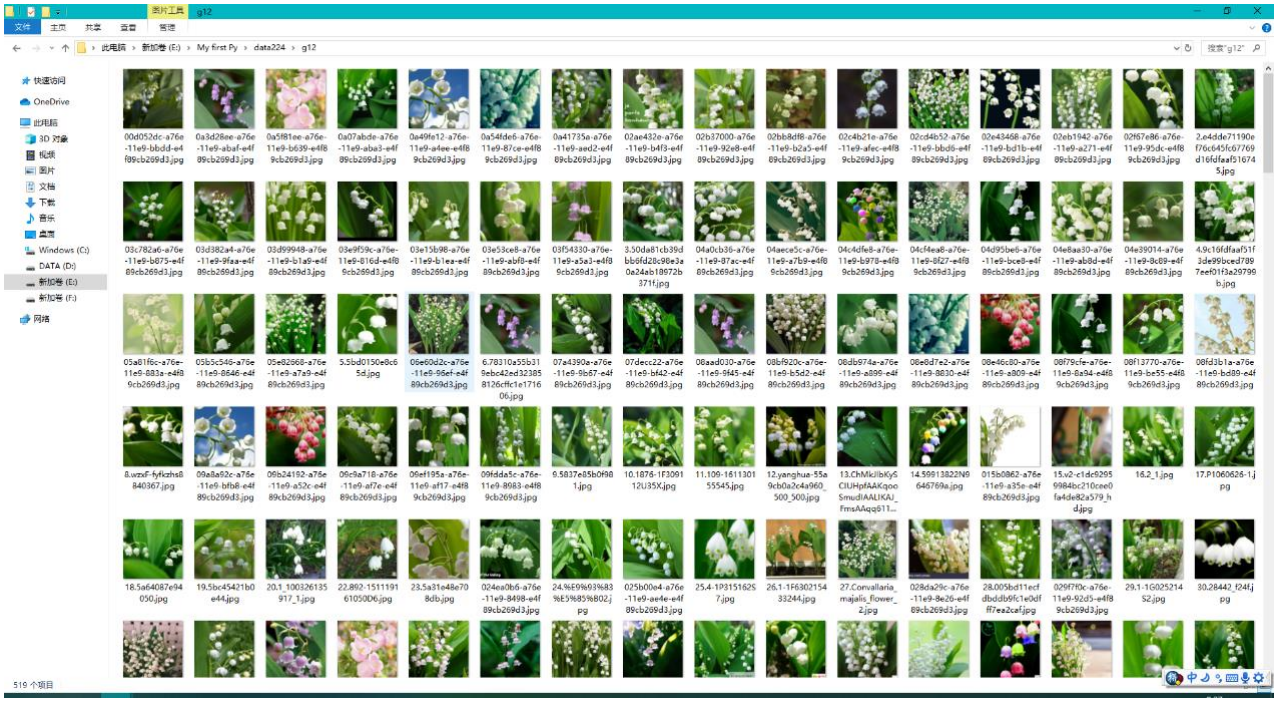


Figure 2. The data set after processing

3. Feature extraction, dimensionality reduction, feature transformation

3.1 Function summery

This module mainly extracts some distinguishing features in the pictures for subsequent processing.

I mainly extract the sift feature as the main feature to establish the Word of Bag model (Firstly, get sift points of all the pictures, then use kmeans to cluster them into 70 clusters, then build the wob model by detecting how many sift points of a picture in a certain cluster. Finally, every picture gets a feature which has a length of 70), as well as lbp, glcm, hog, hsv, and other feature. Additionally I reduce the dimensionality of the longer features to ensure the speed and accuracy of the calculation. After feature extraction, I continuously adjust the length of each feature to ensure the accuracy of the model.

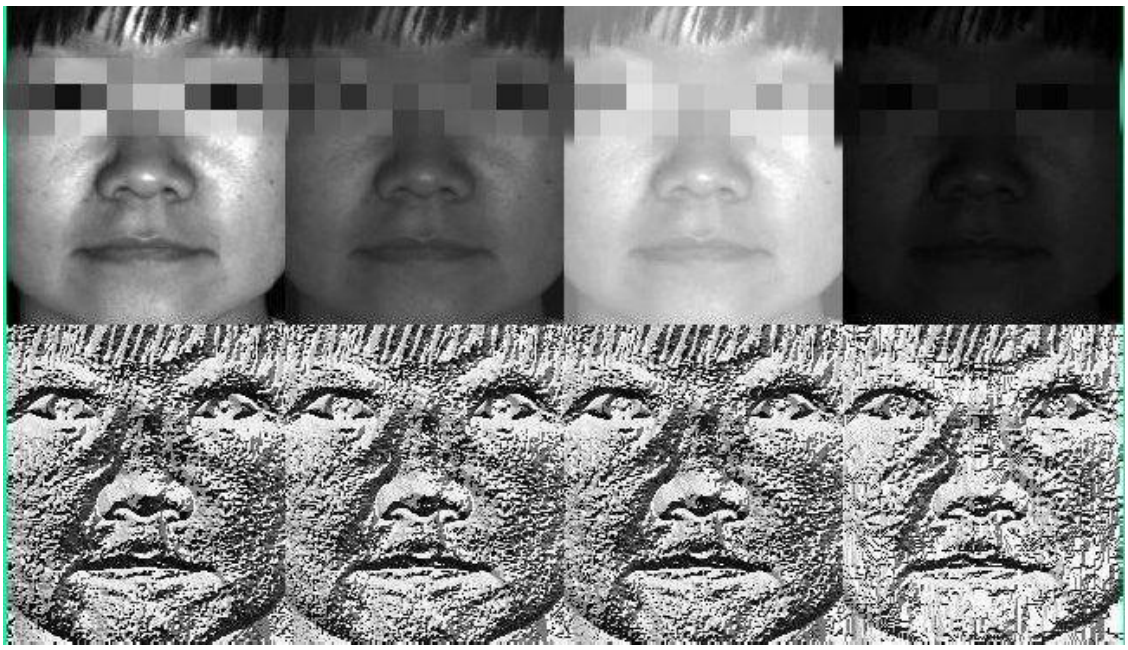


Figure 3. Robustness of lbp feature to light changes[1]

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

3.2 Implement

3.2.1 LBP

LBP refers to Local Binary Pattern, which is an operator used to describe the local features of an image. LBP features have significant advantages such as gray invariance and rotation invariance. The local_binary_pattern function in python's skimage. feature module that can be called directly to obtain the LBP feature of an image. I extracted 64-dimensional lbp features. What's more important is that lbp is very robust to changes in light intensity. Even if there is a big difference in the intensity of light in the pictures, lbp is relatively stable. Figure 3 shows an example of how lbp is robust to light changes.

3.2.2 GLCM

GLCM refers to the gray-level co-occurrence matrix. It can obtain its co-occurrence matrix by calculating the gray-scale image, and then obtain some eigenvalues of the matrix by calculating the co-occurrence matrix to represent some texture features of the image (the definition of texture is still difficult). The gray-level co-occurrence matrix can reflect the comprehensive information about the direction, the adjacent interval, and the change range of the image gray level. It is the basis for analyzing the local patterns of the image and their arrangement rules. Using the greycoprops and greycomatrix functions in the python texture module to obtain glcm features

3.2.3 HOG

Hog is a histogram of directional gradients, a feature descriptor used for object detection in computer vision and image processing. HOG features is to calculate and count the histogram of the gradient direction of the local area of the image. Since HOG operates on the local grid cells of the image, it maintains good invariance to the geometric and optical deformation of the image. The hog function in the python sklearn module can extract the hog feature. But because the hog feature is relatively longer, it is needed to use pca to reduce its dimensionality, and save the trained pca model for use in predicting pictures.

3.2.4 HSV

The HSV feature is a color space created by A. R. Smith in 1978 based on the intuitive characteristics of colors, also known as the Hexcone Model. The color parameters in this model are: Hue (H), Saturation (S), and Value (V). It describes the proportion of different colors in the entire image, but it cannot describe the positional relationship of colors. In python, we use cv2.calcHist to get the hsv feature of a picture. But in my project, the color difference of many flowers is particularly obvious. So this feature can play a very good effect. But for flowers of the same species with different colors, using this feature will reduce the accuracy of the species. Figures 4 and 5 show the relationship between a part of flower accuracy and flower color types.

Figure 4. The prediction accuracy of tulips when using hsv features

Figure 5. Color types of tulips

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

3.2.5 Color-related histogram

The color correlogram is another way of expressing the color distribution of an image. This feature not only describes the proportion of the number of pixels of a certain color in the entire image, but also reflects the spatial correlation between different color pairs. Experiments show that color correlation graphs have higher retrieval efficiency than color histograms and color aggregation vectors, especially for images with consistent spatial relationships. In the designing process, I used the autoCorrelogram function in the texture module to extract the color autocorrelation map features. This feature greatly improves the model accuracy in the flower name recognition system.

3.2.6 SIFT

sift refers to scale-invariant feature transform, it is a description for image processing. This description is scale-invariant and can detect key points in the image. It is a local feature description algorithm[2]. child. In python's cv2, there is the xfeatures2d.SIFT_create function that can be called directly to obtain the sift feature of an image. Figure 6 shows the result of feature point matching based on sift feature

Figure 6. Detecting and matching images based on sift features[3]

3.3 Dimensionality reduction

The method of dimensionality reduction which I use is the PCA (Principal Component Analysis) dimensionality reduction, that is, to retain several dimensions that best reflect the feature in multiple dimensions[4]. This dimensionality reduction method provides related functions in the python library. You only need to specify the size of the dimensionality yourself, and you can call it, which is relatively simple. The problem I encountered in this process is that the dimensionality reduction model after pca training cannot be saved and used repeatedly, so it cannot predict a single picture. After that, I used python's pickle module and used pickle. dump to serialize an object into a file to save the pca model. In the final system, use pickle. load to load the model to save the object.

3.4 Standardization of features

I used the linear normalization method in the development process. Because the dimension of each feature is different, it needs to be normalized to eliminate the difference. Figure 7 expresses the principle of linear normalization

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Figure 7. Principle of linear normalization

That is, the original data is scaled in proportion, where Xnorm is the normalized data, X is the original data, and Xmax and Xmin are the maximum and minimum values of the original data set, respectively.

3.5 Feature conversion and splicing

In the process of extracting features to extract features, I extracted all the features and saved them as a csv file after standardization, and tried to combine features with different permutations and combined them. When adjusting the parameters of the svm model, I can directly splice different files to improve efficiency. Figure 8 shows multiple files when extracting features.

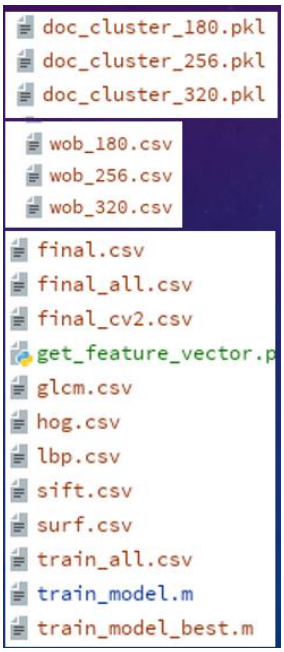


Figure 8. The file storing the features in the system

4. Classification model training and prediction

4.1 Function summery

Use the extracted image features to train the model, and use the trained model to complete the classification task.

4.2 Implementation

In this module, I used two models for predictive classification: svm (support vector machine) and adjusted VGG16 model[5].

4.2.1 svm model

Support Vector Machine (SVM)[6] was first proposed by Corinna Cortes and Vapnik in 1995. It has many unique advantages in solving small sample, nonlinear and high-dimensional pattern recognition, and can be extended to function simulations. In other machine learning problems. There is svm in python's sklearn that can be called directly.

In the process of tuning parameters, you can use the GridSearchCV function in the sklearn library to automatically adjust the parameters. You can also use handwritten loop parameters to adjust

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

parameters. The first method can often get accurate parameters, but the second method is more targeted because it is customized.

The feature vectors I used are: sift+glcm+lbp+color_hist+hsv+hog

4.2.2 VGG16 model

I adjusted the last few layers of the vgg model, as shown in Figure 9

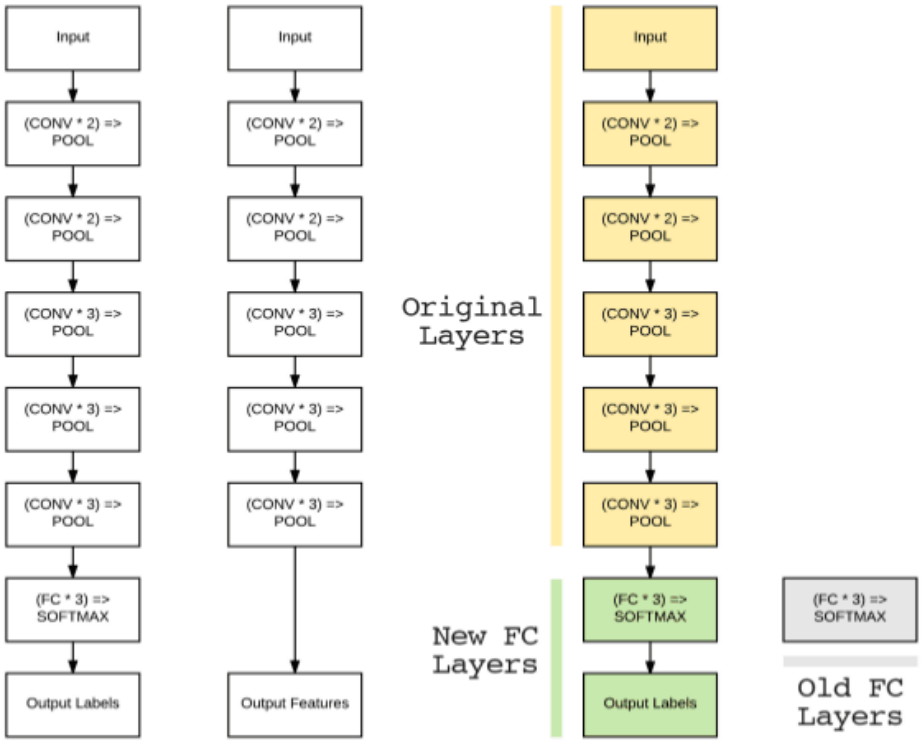


Figure 9. Structure of vgg16 model [7]

In Figure 9, each yellow square represents *convolution and one pooling. The green part on the right is used for network fine-tuning. Replace the original fully-connected layer with 3 fully-connected layers and one output layer. The yellow part means to freeze their parameters without fine-tuning, and the green part means that the parameters need to be fine-tuned to adapt to the new Model.

4.3 Model optimization

After building the model, I tried to optimize it on the basis of this model. Optimization includes two aspects: adjusting parameters and adding features, data normalization and adjusting algorithms.

In terms of adjusting parameters and adding features, I have tried various combinations of feature lengths, but within a certain range, it has little effect on the accuracy. I later added the feature of color-related histograms, which improved the accuracy a lot.

In terms of data normalization, during the training process, I found that the operation was time-consuming and the effect was not good. Finally, I found that the data was not normalized. I processed the data with maximum and minimum normalization, and finally successfully improved the accuracy.

Later, in terms of algorithms, when adjusting the svm parameters, I tried the Gaussian kernel function. Kernel functions such as polynomial kernel functions and exponential kernel functions. Finally, it is found that the Gaussian kernel function has the best performance in accuracy.

In addition, in the combination of module extraction features, different combinations have different effects. I enumerated different combinations of features and got a better feature extraction method on my model.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

5. User Interface

After realizing the function of predicting the input picture, the system realizes the user interface. In this interface, users can see the pictures they input and the pictures of the training set. Moreover, after the forecast is over, users will be able to view the forecast accuracy of each type of flower.

Figure 10 shows the main interface. The user will see this page when he first enters the webpage. The user uploads the data set from this webpage and goes to other pages.

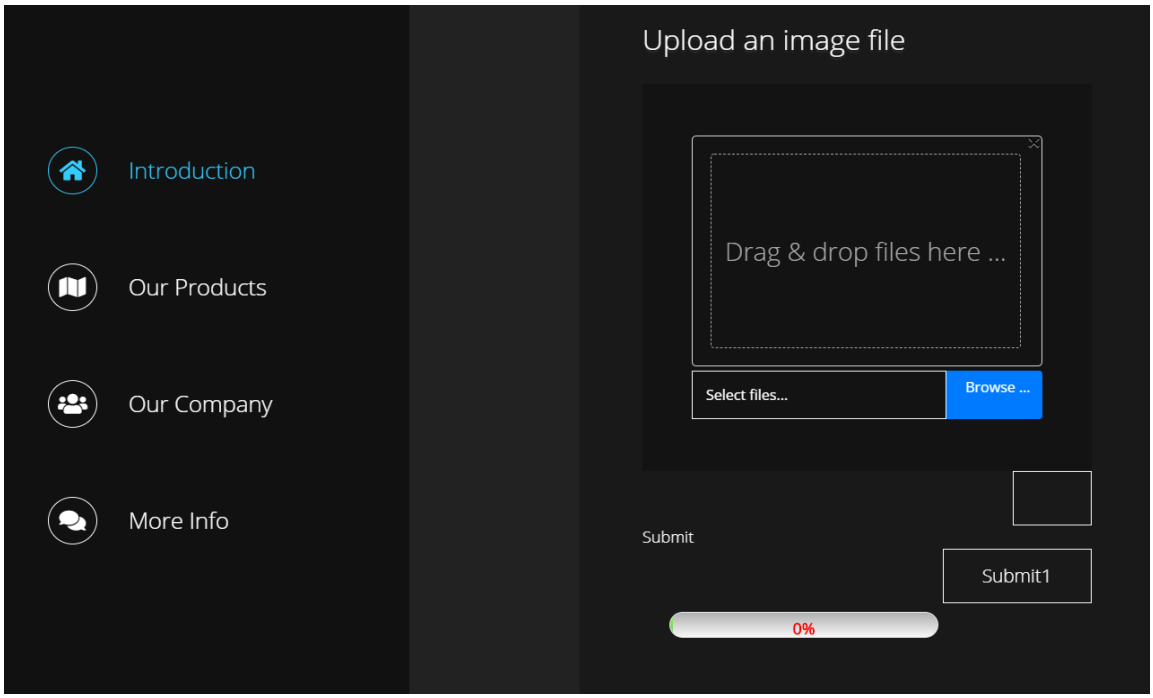


Figure 10. Main interface

Figure 11 shows the page displayed when the user enters to view the uploaded picture. The page uses the cls parameter to locate the path. When cls=0, the user picture is displayed, and the folder is located according to the cookie. When cls is between 1 and 21, the training set picture is displayed.

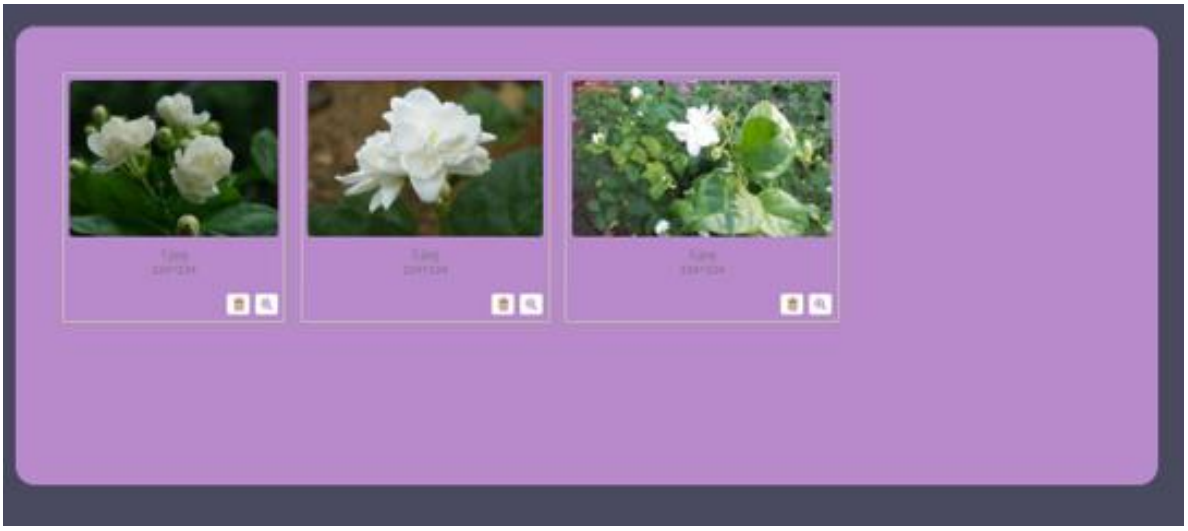


Figure 11. Showing picture interface

Figure 12 shows the prediction results without labels, and Figure 13 shows the accuracy of image prediction using traditional methods

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

Figure 12. Image prodiction

g0	NaN	NaN	0
g1	0.5151515151515151	0.9696969696969697	99
g2	0.7171717171717171	0.9696969696969697	99
g3	0.8686868686868687	0.98989898989899	99
g4	0.5252525252525253	0.9595959595959596	99
g5	0.8181818181818182	0.9797979797979798	99
g6	0.6565656565656566	0.9595959595959596	99
g7	0.7474747474747475	0.9393939393939394	99
g8	0.13131313131313133	0.5454545454545454	99
g9	0.5670103092783505	0.9278350515463918	97
g10	0.7272727272727273	0.9595959595959596	99
g11	0.7272727272727273	0.9696969696969697	99
g12	0.8080808080808081	0.9595959595959596	99

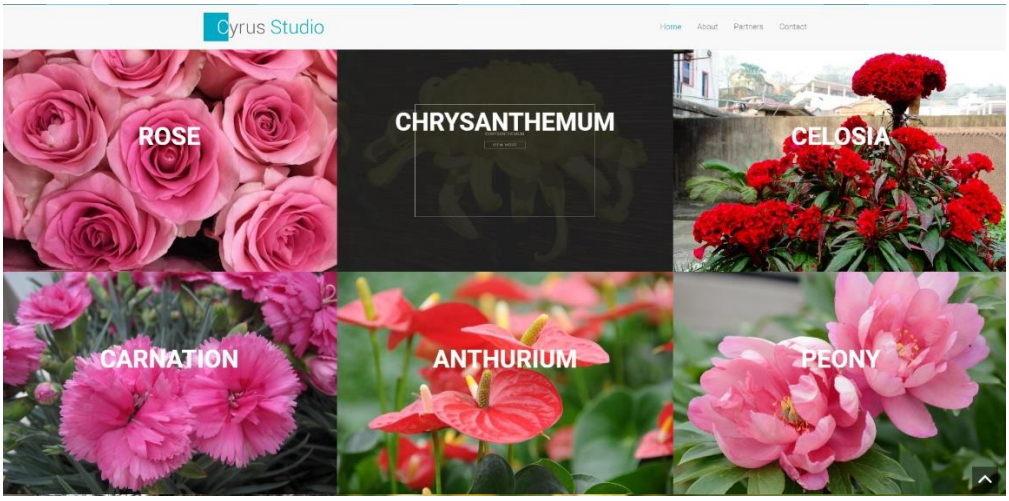
Figure 13. Accuracy of svm model

g1	0.9292929292929293	0	99
g2	0.9494949494949495	0	99
g3	0.9494949494949495	0	99
g4	0.8888888888888888	0	99
g5	0.98989898989899	0	99
g6	0.9393939393939394	0	99
g7	0.8383838383838383	0	99
g8	0.9595959595959596	0	99
g9	0.9489795918367347	0	98
g10	0.9494949494949495	0	99
g11	0.9595959595959596	0	99
g12	1	0	99
g13	0.9797979797979798	0	99
g14	1	0	99
g15	0.8888888888888888	0	99
g16	0.8080808080808081	0	99

Figure 14. Accuracy of VGG16 model

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

Figures 15 and 16 show that users can view the training set pictures and modify them (add or delete).



sss

Figure 15. Training Data Set-1

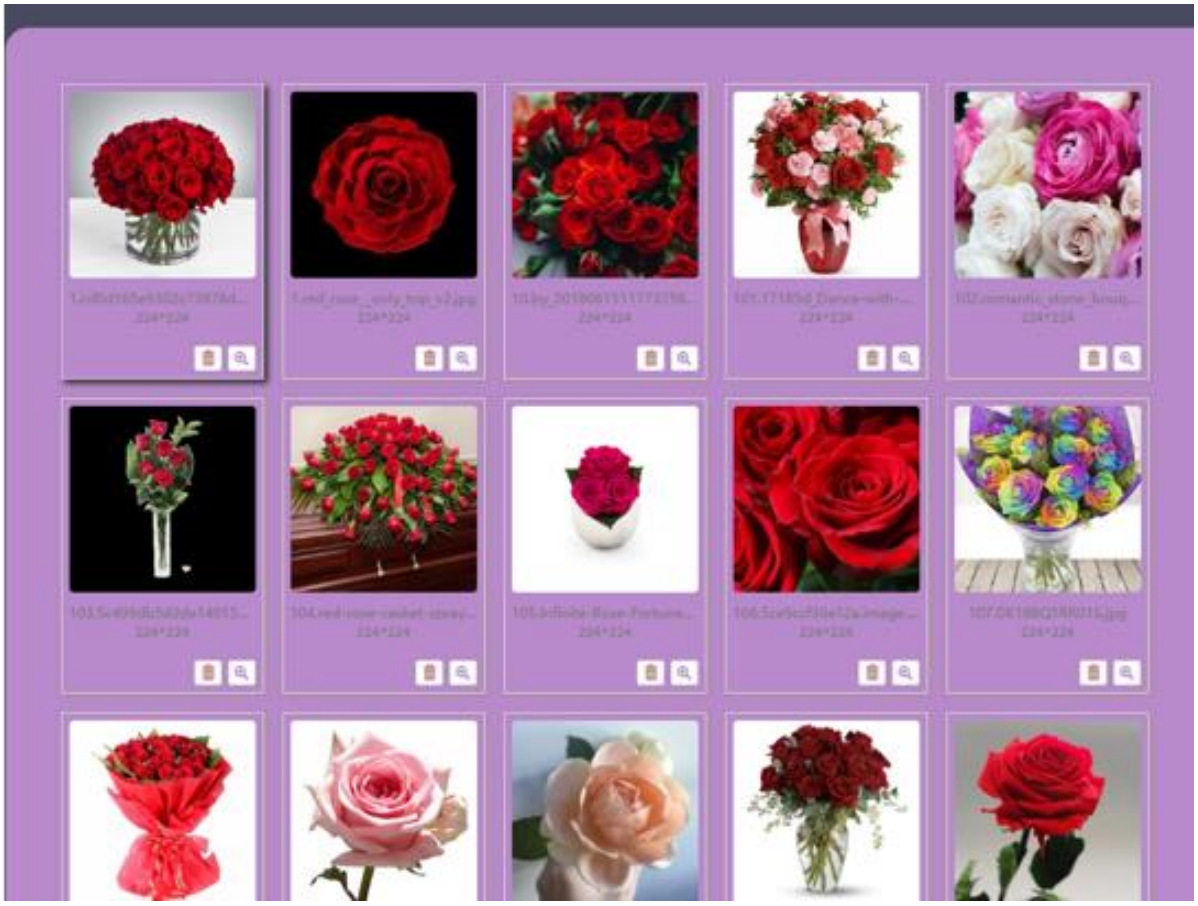


Figure 16. Training Data Set-2

6. Conclusion

This thesis uses two classification learning methods: svm and vgg16 models. By realizing the prediction of 21 kinds of flowers, the accuracy rates of 65% and 85% are reached respectively. Comparing the two methods at the same time, it can be seen that the accuracy of the neural network

ISSN: 2414-1895 DOI: 10.6919/ICJE.202102_7(2).0048

is higher and the prediction time is shorter. In the system of this article, a user-friendly interface is realized, which is more practical. Users can view their uploaded pictures, prediction results, training set pictures and other details in the system. Due to the limited time, the accuracy of prediction using the svm model is not very high. In future work, methods such as gradient descent can be further used to adjust parameters to further improve accuracy.

References

[1] https://www.cnblogs.com/thinkinpakho/p/10880797.html

[2] Wu ying. Product Image Similarity Algorithm Based on SIFT and Nearest Neighbor Matching[J]. Computer and Modernization, 2020(10):69-75.

[3] https://www.cnblogs.com/wenbozhu/p/10548794.html

[4] Zhang suzhi, Chen xiaoni,Yang rui, Li penghui, Caiqiang. Method of principal component analysis based on intra-class distance and inter-class distance[J].Computer Engineering and Design. 2020,41(08):2177-2183.

[5] He H S, Huang X X, Li H G, et al. Water body extraction of high resolution remote sensing image based on improved U-Net network[J]. Journal of Geo-information Science, 2020,22(10):2010-2022.

[6] Cortes, C. and Vapnik, V., 1995. Support-vector networks. Machine learning, 20(3), pp.273-297.

[7] https://www.pianshen.com/article/9718106687/.