Comp9517 20T3 Image Classification of Plants - Part1. Individual Component

Haojin Guo
The Univeristy of New SouthWales
z5216214@student.unsw.edu.au

Abstract— **The efficient classification of plant pictures has positive significance for biological research. This report mainly focuses on the classification of two plant pictures Arabidopsis and Tobacco. The traditional feature extraction method is used to extract image features, and then to encode the feature is as Bad of words(BoW), and finally to complete the image training, testing and, classifying. In addition, the classification effects of different models are compared.**

Keywords— ***classification, feature extraction, BoW, feature encoding***

I. INTRODUCTION

The technology of analyzing plant phenotype [1] with computer vision is widely used in the observation and research of comprehensive biological traits or characteristics in modern genetics, which including classification and identification of plant type, localization of their position in an image, and segmentation of the plant and its leaves.

The main task of the first part of this project: classification. In particular, implementing a Python solution to distinguish Arabidopsis plant images from Tobacco plant images. Here, the plant images are from the Plant Phenotyping Dataset [2]. But a significant prerequisite point is that deep learning method for features is not allowed, traditional feature extraction techniques from computer vision, hand-crafted or engineered features are required. Therefore, the main challenge may be the feature encoding, integrating the local features of an image into a global vector to represent the whole image.

There are two data sets in this project. This first set is called “Ara2013-Canon” including 166 files (165 “*_rgb.png” images and 1 “Metadata.csv” file). The size of the image Arabidopsis in this is much smaller than the image Tobacco in the second data set, which called “Tobacco” in the dataset including 62 “*_rgb.png” images and 1 “Metadata.csv”. This will lead to resize the sets of images in the preprocessing. Besides, the number and size of Arabidopsis leaves are larger than those of Tobacco leaves even these two types of pants are in the budding stage in the picture, which is the biggest difference between the two plant phenotypes.

In this report, the second section mainly shows the method of the training and learning process in the image classification task. The third section demonstrates the experimental setups and the method used to evaluate. The last section will be the results and discussion.

II. METHOD

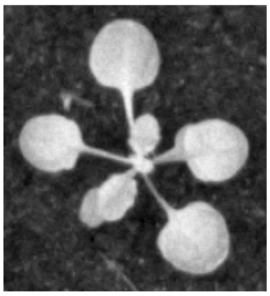
A. Preprocessing

The first step is reading the two image files separately, and meanwhile, resizing the two sets of images to the same size, such 250*250, which plays a vital role in the subsequent feature extraction and feature encoding. The images exist in the corresponding lists in matrix form. Besides, because the original image data does not contain the corresponding labels when reading two different images of folders, it is necessary to manually classify the two types of plants and then store the labels in two lists. The next step is to integrate the two image lists and two label lists into a dataset and a label set. In addition, the images are tried to be filtered, median filter, and morphological [3] processed in order to verify the effect of classification in its different preprocessing situations. Finally, the dataset and label set need to be divided into training and test sets by using the function of “train_test_split()” in scikilearn [4], which the training set accounts for 75% and the test set accounts for 25%.

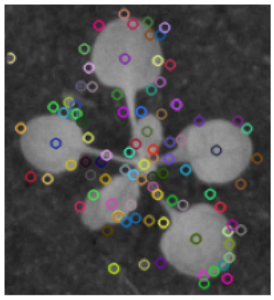
B. Features Extraction

In this project, SIFT in OpenCV is mainly used for image extraction. In 2004, D. Lowe of the University of British Columbia proposed a new algorithm called Scale Invariant Feature Transform (SIFT) in his paper "SIFT-the scale-invariant feature transform" [5], which can extract the key points of the image and calculate its descriptor. In the case of a grayscale image, using the descriptor “sift.detectAndComputer()” to extract the features of the image as a number array with a shape of *NumberOfKeypoints * 128*.

Furthermore, other techniques of extracting features, such Histogram and SURF, are also tested in this process, the purpose is to compare the differences in the evaluation of classification of different methods. The specific comparison results will be presented in the third section.



F.1 Gray image



F..2 Extract key points by SIFT

C. Feature Encoding

After using the feature extracting by SIFT to get the representative features of image, and put the features of all the images together to form an array list, the next important method is to encode them to form a bag of features(visual words). If there is no bag of features, it is impossible to establish a unified interpretable relationship between features and vectors.

The most popular method of feature encoding is Bag-of-Words (BoW). The word bag generation is based on the descriptor data to generate a series of vector data. And the main technique used to create the vocabulary is k-means clustering [6]. First, First, cluster analysis of descriptor data is realized through k-means, it will be divided into 100 clusters, and the central data of each cluster will be obtained, which will generate 100 words bags. According to the distance of each descriptor to these cluster centers, it is determined which cluster it belongs to. This generates its histogram to represent the data. Therefore, each image will form a meaningful feature vector k*1. In fact, bag-of-words clustering analysis essentially extracts higher-level features of the image.

D. *Classification and testing*

Using SVM for data classification training to obtain the output model. Here, the classification model training and export are realized through sciki-learn SVM (kernel = 'linear') training.

Furthermore, other techniques, such k-Nearest Neighbors (KNN) and RandomForestClassifier (RFC), are also tested in this process, the purpose is to evaluate the differences classifiers. The specific comparison results will be presented in the third section

III. Experiment Setup

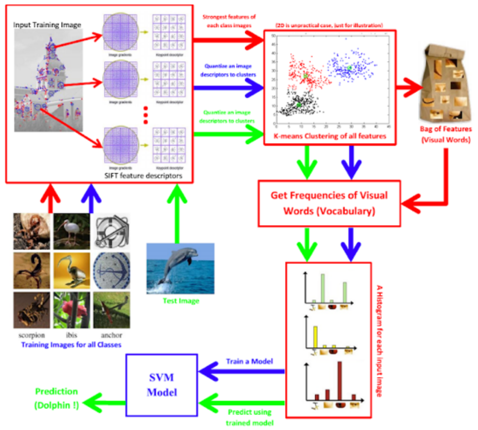
In the experiment, the two image data sets are " Ara2013-Canon" and "Tobacco", representing images of plants Arabidopsis and Tobacco. Besides, the size of each image is inconsistent, so it is necessary to read each image at the same size. In this experiment, the image size is set to 250*250 and a data set and a label set will eventually be generated.

Besides, each image is read as an RGB picture at the very beginning. But grayscale images are used in the entire classification learning process because the extractor initialization uses the SIFT. However, in the subsequent experimental process, the different classifiers accuracy results of the extractor SURF between the RGB image and the GRAY will be provided.

The processor of the computer we used in the working environment is 2.3 GHz Quad-Core Intel Core i5 with Intel i5 2.3GHz with 16 GB 2133 MHz LPDDR3 RAM. And Pycharm has been applied to compile and run our code, and the installed binaries include Python 3.7.7, scikit-learn 0.23.2, OpenCV 3.4.2, scipy 1.5.1.

The process of the entire project task follows the steps shown in Fi.3. [9] below.

Moreover, in this experiment, SURF and Histogram, and two extra classifiers KNN and RFC, are additionally used to extract features of images to verify and compare the different results of picture classification. On top of that to verify the effectiveness of the image classification results and tracking methods, the quantitative evaluation we applied is to calculate the parameters Precision, Recall, AUC. The comparison results and data statistics are shown in Table1.



Fi.3. System overview [9]s

Table.1. Evaluation of different classification models

Model	Precision	Accuracy	Recall	AUC
SIFT+SVM	1.000	0.977	0.907	0.953
SURF+SVM	0.972	0.947	0.813	0.914
HIST+SVM	0.976	0.977	0.930	0.961
SIFT+KNN	1.000	0.982	0.930	0.965
SURF+KNN	0.973	0.953	0.837	0.914
HIST+KNN	1.000	0.988	0.953	0.976
SIFT+ RFC	1.000	0.970	0.884	0.941
SURF+ RFC	1.000	0.988	0.953	0.976
HIST+ RFC	1.000	0.976	0.901	0.952

(Note: 1. All based on grayscale image. 2. The value of k in K-means clustering is 120. 4. Using clf = SVC(kernel="linear"), Not clf = LinearSVC())

Table.2. Evaluation of SIFT and SURF in RGB images

Model	Precision	Accuracy	Recall	AUC
SIFT+SVM	1.000	0.988	0.953	0.976
SURF+SVM	0.975	0.965	0.86	0.930

Table.3. Evaluation of SIFT and SURF after MorphologyEX Opening (Gray image)

	Precision	Accuracy	Recall	AUC
SIFT+SVM	1.000	0.988	0.953	0.977
SURF+SVM	0.948	0.953	0.860	0.922

IV. Results and Discussion

Evaluating different classification models to test the different effects of image classification learning under different preprocessing conditions, and recording the result in Table1,2,3.

A. Evaluation of different classification models

In the python code implementation process, there are 9 different classification models, which are the result of 3 extractors, SIFT, SURF, and Histogram, and 3 classifiers, SVM, KNN, and RFC, combinations.

From Table.1, it can be found that the final evaluation data results of the SIFT and Histogram-related models are quite similar firstly. The possible reason is that when SIFT and HIST extract different image features, they will get a fixed-length one-dimensional array, which is [128*1] and [256*1], respectively. Moreover, sift undergoes the process of histogram statistical analysis after feature encoding through k-means, and finally obtains a k*1 one-dimensional array. But the specific reasons should be analyzed accurately.

Secondly, the overall evaluation values of SIFT is better than SURF, when using SVM and KNN classifiers, however in the case of RFC, the latter is slightly better than the former.

Thirdly, calculating the results of SVM, KNN and RFC in three different feature extractors will find that the numerical results are not much different.

B. Evaluation of SIFT and SURF in RGB images

In the case of reading RGB images, the classification effect of SIFT is better than SURF.

C. Evaluation of under morphological opening

In this case, the classification effect of SIFT is better than SURF as well

D. Result

In summary, it can be concluded that the overall classification and evaluation effect of SIFT feature extraction is better than that of SURF. To be more specific, excluding the evaluation parameters Precision, Accuracy and Recall, AUC (sift) > AUC(surf) also. The area under the ROC curve is called AUC and can evaluate the performance of the model. Particularly, the prediction performance, this means that the closer the AUC value is to 1, the stronger the prediction ability of the corresponding model. Therefore, in this project, SIFT has a stronger effect on plant image classification than SURF.

However, there are indeed some inevitable errors or deviations in this project. This is because there is no opportunity to fully consider the influencing parameters in each task. For example, the random allocation of training set and test set will cause unstable results of running code, which will negatively affect experimental conclusions. As a result, the evaluation data obtained must not be very accurate.

References

[1] https://en.wikipedia.org/wiki/Phenotype

[2] Massimo Minervini, Andreas Fischbach, Hanno Scharr, Sotirios A. Tsaftaris, *Finely- grained annotated datasets for image-based plant phenotyping*, Pattern Recognition Letters, Volume 81, 2016, Pages 80-89, ISSN 0167-8655.

[3] https://docs.opencv.org/4.1.2/d4/d86/group__imgproc__filter.html#ga67493776e3ad1a3df63883829375201f

[4] https://scikit-learn.org/stable/index.html

[5] Lowe, G., 2004. *SIFT-the scale invariant feature transform*. *Int. J*, *2*, pp.91-110.

[6] https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.vq.vq.html#scipy.cluster.vq.vq

[7] https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html

[8] https://docs.scipy.org/doc/scipy/reference/cluster.vq.html

[9] http://heraqi.blogspot.com/2017/03/BoW.html

[10] https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc