《统计学习与数据科学导论课程设计》

课 程 论 文

(2023 -2024 学年第暑假学期)

Apple Quality Assessment and Recognition Using Geometric Fitting and SVM Classification

学生姓名：涂誉晋

提交日期：2024 年 8 月 28 日　　　　学生签名：涂誉晋

学　　号	202166321337	座位编号	
学　　院	数学学院	专业班级	21 信息与计算科学
课程名称	统计学习与数据科学导论课程设计	任课教师	夏立

教师评语：

本论文成绩评定：______分

Apple Quality Assessment and Recognition Using Geometric Fitting and SVM Classification

Yujin Tu

Abstract

In this paper, we use SVM classifier to build a model to classify images. We focus on an image dataset of fruit pictures. We try to use these data sets for training and get a complete apple recognition model. We follow the methods based on edge detection and connected domain analysis to construct an image feature extraction model, followed by using SVM as a classifier training and then feature classification to form a complete apple recognition model. Combined with the for the model evaluation, and get a high accuracy rate, and finally the images in Attachment 3 are feature classification and plot the histogram of the distribution of ID numbers of all apple images.

Keywords: Edge detection；Connected domain analysis；SVM classifier

1. Introduction

China is the largest producer and exporter of apples in the world. Due to the difference between the orchard environment and the laboratory environment, it is difficult for the existing robots to accurately recognize the obstacles such as leaf shading, branch shading, fruit shading and mixed shading. There is an urgent need to establish an image recognition model for apples with high recognition rate, fast speed and accuracy, and to analyze the images with data.

Support vector machine is a supervised learning algorithm that can be used for multiclassification problems. The idea is to find an optimal hyperplane that can divide different classes in the many feature spaces while maximizing the interval from the sample points to the hyperplane.

2. Method and model building

2.1 Edge detection

Since the image in Attachment 1 has strong light interference and mixed occlusion, this paper first performs some processing operations on the image to improve the quality and analyzability of the image, and also since the size of the images in Attachment 1 are all 270×180, there is no need to standardize the size of the images to uniform resolution again.

2.1.1 Color Space Conversion and Enhancement

HSV is a color space created based on the intuitive properties of colors, the parameters are expressed as hue (H), saturation (S) and luminance (V), which can intuitively express the hue, vibrancy, and lightness and darkness of the image, and HSV is more sensitive to colors compared to the RGB color space of the original image. The analysis found that there are only 200 images in Attachment 1, the dataset is small, and at the same time, the foreground color of the apple is isolated from the background color to a large extent, so in this paper, we first establish a color space conversion model to convert the image from RGB color space to HSV color space in order to better distinguish the color of the apple from the green of the background.

Figure 1 Original Image 27

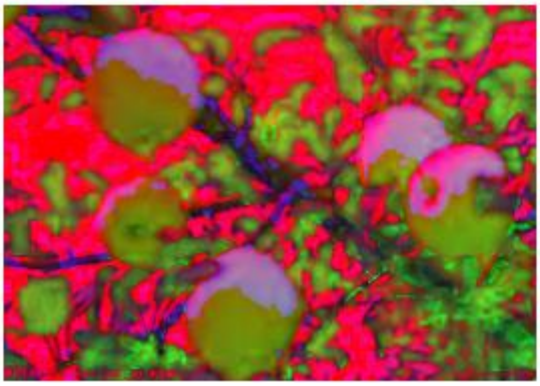

Figure 2 HSV Image

The comparison before and after the conversion shows that the separation property of HSV makes the color information easier to distinguish, making it easier to detect the presence of apples. At the same time, since the luminance (V) channel contains the brightness information of the image and is relatively insensitive to changes in lighting, it makes the effect of changes in lighting on apple color recognition much less in the case of foliage and ambient occlusion, enabling the presence of apples to be more prominently highlighted.

2.1.2 Median filter

The median filter is a nonlinear filter that could reduce the effect of pretzel noise by taking the median of the neighboring pixels, which preserves the edge information of the target apples and helps to smooth out the edge profile of the whole image in the case of more noise from leaf occlusion and ambient light, thus improving the accuracy of apple recognition.

For a matrix I of pixels on an image, consisting of individual pixel points (x, y) and a predefined window size W, the operation of median filtering can be expressed as:

$$I_{\text{median}}(x, y) = \text{median}\big(I(x-w, y-h), I(x-w, y), I(x-w, y+h)\big) \tag{1}$$

$$I(x, y-h), I(x, y), I(x, y+h), I\big(x+w, y-, I(x+w, y), I(x+w, y+h)\big) \tag{2}$$

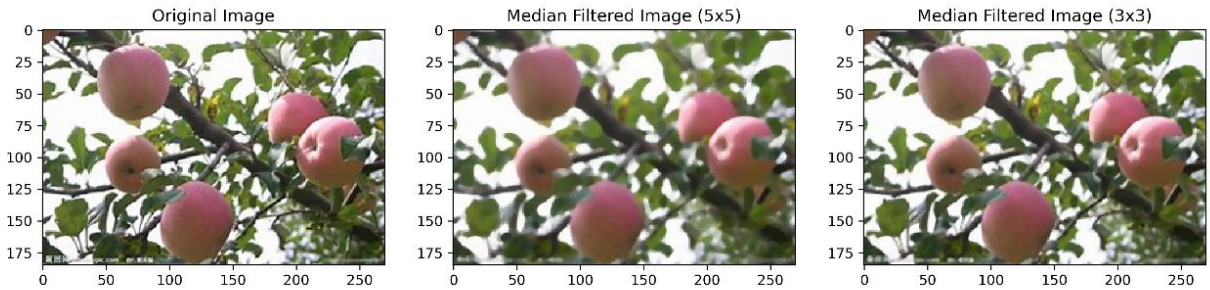

Figure 3 Before and after median filtering

Among them, the effect of median filtering is closely related to the choice of window size. Too small a window may result in the ineffective removal of large noise, while a larger window may result in the loss of too much detail in the image. Therefore, in this paper, the images in Annex 1, respectively, selected 3×3 and 5×5 window for median filtering comparison, the image before and after the filtering and the histogram for comparison.

Histograms are a common way to understand noise and show its effects. After comparing the three histogram curves below, it can be concluded that the histogram of the median filter with the 3×3 window is smoother, while that with the 5×5 window has less contrast with the original one, and at the same time the details of the image can be preserved more, probably because the noise present in the image in Attachment 1 is more small in size, and the effect of median filtering using a small window is more significant.

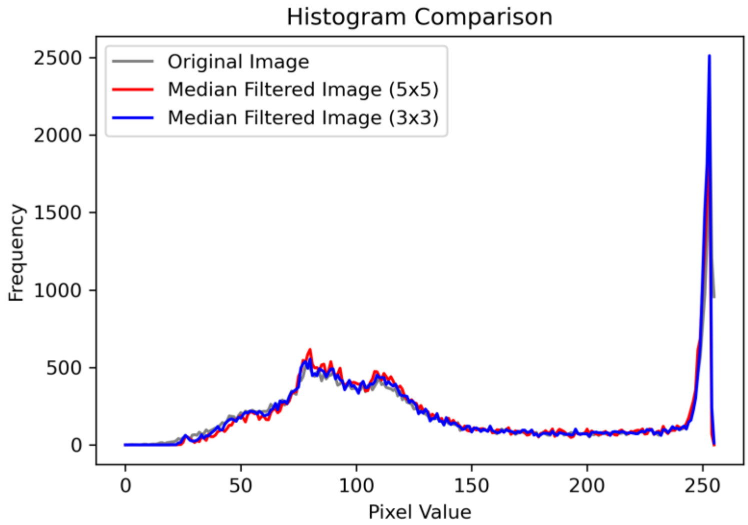

Figure 4 Histogram of median filter

2.1.3 Canny Contour Enhancement

In this paper, we found that even after the processing of the previous two methods, the recognition of apples will still be greatly affected by the occlusion of leaves, light, etc. Therefore, we will use Canny Edge Detection for contour enhancement of all the images in Attachment 1, which will help to capture the clear boundary between the apples and the background, and to form a clear and more continuous outline of the apples. For the threshold parameter of Canny edge detection, after many experiments, we set the threshold parameter as Double threshold detection $(50, 150)$, the gradient is calculated by the following formula, and the direction of the gradient is taken as.

$$\theta = \arctan\left(\frac{G_y}{G_x}\right) \tag{3}$$

$$G_x = \frac{\partial G}{\partial x}, G_y = \frac{\partial G}{\partial y} \tag{4}$$

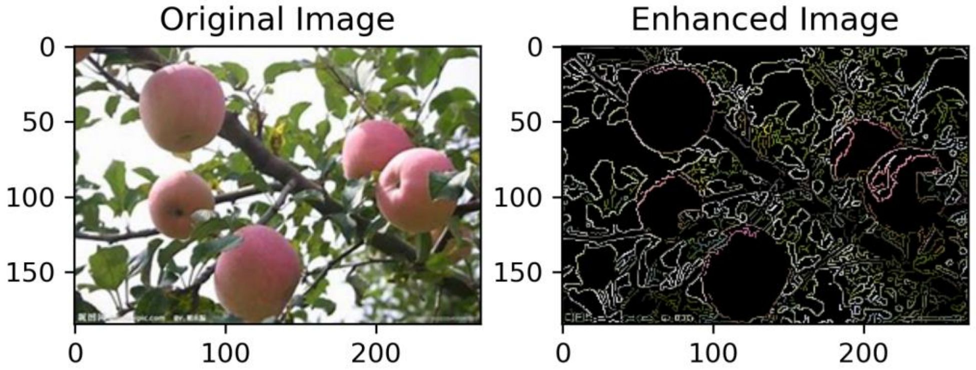

Figure 5 Canny Contour Enhancement

2.2 Connectivity domain analysis

Connectivity domain analysis can be used to identify and label the connectivity region in an image, i.e., in this paper, we use four-neighborhoods to define a connectivity domain, i.e., the masks of the top, bottom, left, right and right four neighboring elements are all 1,,which is identified as a strongly connected region, i.e., a single apple region, whereby the region where each apple is located is divided.

Set M as the binary mask matrix, where $M_{i,j}$ is the element of the ith row and jth column of the matrix, P is the number of columns of the matrix, N is the number of rows of the matrix, L is the labeling matrix,$L_{i,i}$ is the element of the ith row and jth column of the labeling matrix, C is the number of connected regions, i.e., the number of apple regions, and S_k is the size of the kth connected region. k size of the connected region, $k = 1,2, \dots ,c$.

Using the traversal algorithm, iterates through each \$M\$ value and if $M_{i,j} = 0$, then skips the element and proceeds to the next element. If $M_{i,j} = 1$, then check whether its four neighboring elements $M_{i-1,j}, M_{i+1,j}, M_{i,j-1}, M_{i,j+1}$ are 1. If they are 1, the region is labeled as a strongly connected region, i.e., an apple region.

$$\begin{bmatrix} M_{i-1,j} & M_{i-1,j-1} \\ M_{i+1,j} & M_{i+1,j+1} \end{bmatrix} \tag{5}$$

2.3 SVM classification model

For the features of many fruits such as apples, this paper considers that the features of apple images are linearly divisible in the original space, i.e., there exists an optimal hyperplane capable of completely separating different classes of data points, then for the training set Attachment 2 is linearly divisible, containing the feature vector set X and the label set Y, the number of samples m.

$$(X, Y) = \{(x_1, y_1), \ldots, (x_m, y_m)\} \tag{6}$$

Accordingly, the expression for the segmentation hyperplane, i.e., the decision function of the linear SVM, is obtained:

$$f(x) = w^T x + b \tag{7}$$

where w is the weight vector, x is the feature vector, and b is the bias, Therefore, during the training process using the data in Attachment 2, the goal of the SVM model is to maximize the interval from the sample points to the hyperplane while ensuring that the sample points are correctly classified, to obtain the optimal hyperplane:

$$f(x) = (\sum_{i=1}^{m} \alpha^* y^{(i)} x^{(i)})^T x + b^* = \sum_{i=1}^{m} \alpha^* y^{(i)} < x^{(i)}, x > +b^* \tag{8}$$

3. Experiments

After feature extraction for each fruit using the previous feature model, model training is started. In this paper, the model is trained to differentiate between apple and non-apple images by labelling apples with 1 and other fruit images with 0. Also, the data in Attachment 2 is divided into validation set and training set according to 3:7.

The common evaluation metrics for classification tasks are Confusion Matrix, Accuracy, Precision, Recall, where Accuracy is the most direct criterion for evaluating how good a model is, with the formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

where TP denotes the number of samples that are positive and predicted to be positive, FP denotes the number of samples that are negative but predicted to be positive, TN denotes the number of samples that are negative and predicted to be negative, and FN denotes the number of samples that are positive but predicted to be negative.

According to the confusion matrix, it can be seen that the model training is good and the accuracy is high.

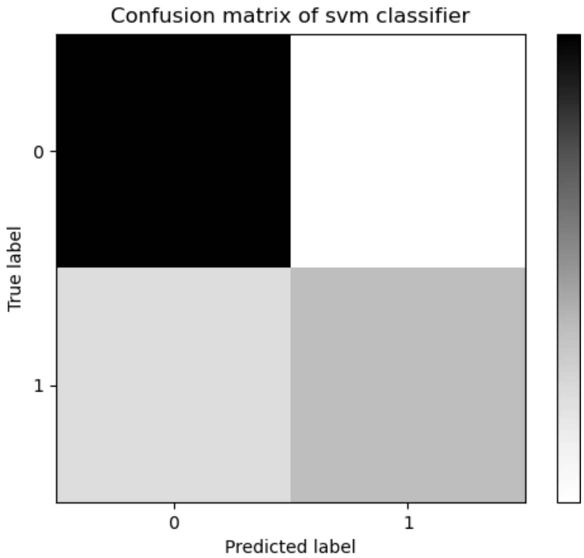

Figure 6 Confusion matrix of SVM classifier

The apples were finally identified by combining feature extraction and SVM classification models to classify the unlabeled fruits in Attachment 3, and the histogram of the distribution of the ID numbers of the apple images was also plotted:

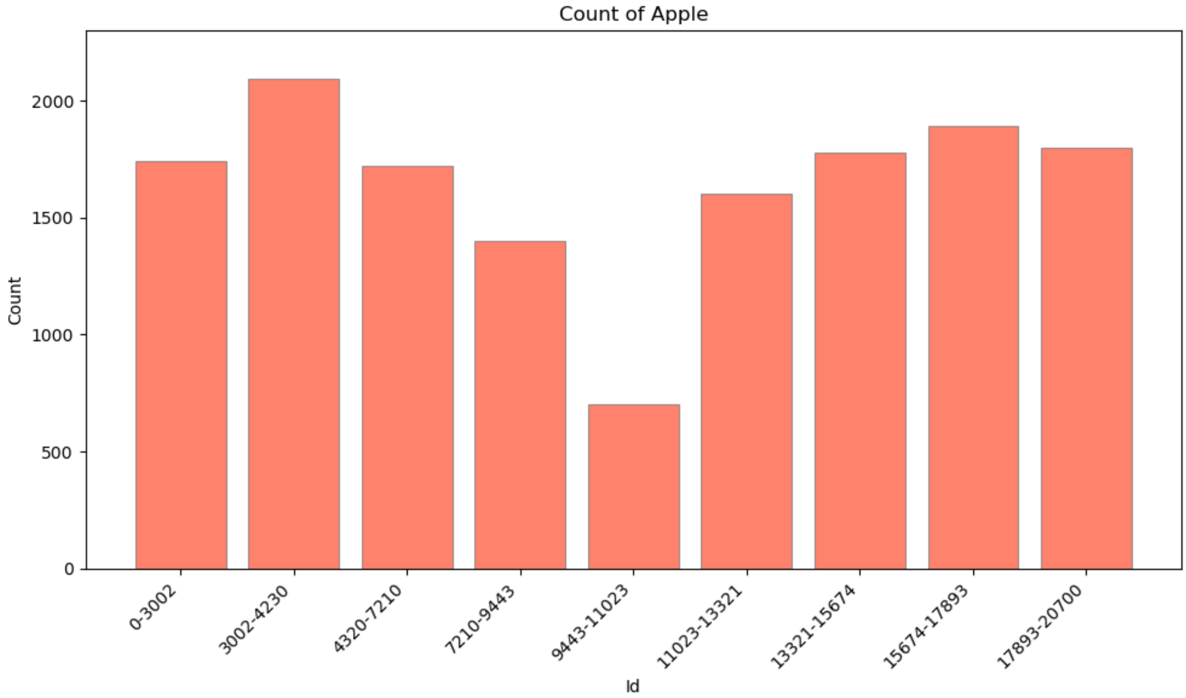

Figure 7 Count of Apple

4. Conclusion

In this paper, we successfully demonstrate a method for apple recognition using a combination of image preprocessing techniques and SVM classification. By addressing challenges such as occlusion and lighting variations, the proposed model achieves high accuracy in apple detection. The use of HSV color space, median filtering, and Canny edge detection significantly improves the quality of image feature extraction, leading to effective classification with the SVM model. The results and evaluation metrics confirm the model's robustness and its potential application in automated apple quality assessment and recognition.

In future work, we will expand the dataset to improve the generalization and robustness of the model, experiment with advanced image processing techniques such as deep learning for feature extraction and incorporate more diverse environmental conditions to enhance the model's adaptability.

Reference

[1] Argote I, Archila J, Becker M. Fruit Identification System in Sweet Orange Citrus (L.) Osbeck Using Thermal Imaging and Fuzzy[J]. 2016.

[2] Huirong X, Zunzhong Y, Yibin Y. Identification of citrus fruit in a tree canopy using color information[J]. Transactions of The Chinese Society of Agricultural Engineering, 2005.

[3] Dubey S R, Jalal A S. Adapted Approach for Fruit Disease Identification using Images[J]. Image Processing Concepts Methodologies Tools & Applications, 2012, 2(3): 51-65.DOI:10.4018/ijcvip.2012070104.