Vision system of apple picking robot based on twin support vector machine

Yuxiang Du
Guangzhou Nanyang Polytechnic Vocational College, Guangzhou 510925, China
yuxiangdu@npvc.edu.cn

Abstract. The key technology of picking robot is visual system, whose accuracy and stability directly and indirectly affect the efficiency of the system. Aiming at the intelligent sorting and grading problem of fruit tree picking robot, a vision system of apple picking robot based on twin support vector machine was proposed. First, the system divided the apples into mature and immature according to the color attributes of the collected samples, and segmented the lesion areas in the image using Gabor wavelet transform to identify the defects in the fruit. Secondly, according to the defect vector constructed by additional color and geometric features, the advanced twin support vector machine is used as a classifier to identify three types of defects. Finally, the performance of the system is evaluated with the indexes of accuracy, specificity, sensitivity and precision. Experimental results show that, compared with other classifier models, the proposed system has the best performance on all performance indexes, can greatly improve the accuracy of the picking robot, and promote the practical progress of the fruit picking robot.

Key words: Robot vision; Picking robots; Apple; Gabor wavelet transform; Twin support vector machine.

1. Introduction
With the rapid development of our country's economy and science and technology, the contradiction between the increasing demand for agricultural products and the slow development of modern agricultural productivity has become increasingly acute. The backward agricultural production has seriously hindered the economic development. At present, the domestic fruit harvesting is completed by human resources. With the aging of the population, gradually reducing human resources and increasing production costs, these basic national conditions greatly reduce the advantages of China's fruit market. In this context, agricultural robots came into being.

Because of the convenience and high efficiency of fruit picking robot, it has become an important research direction for scholars. It can not only liberate farmers' labor force, but also ensure the effectiveness of fruit picking. Of course, the advantages of fruit picking robot are not limited to these two aspects, and there are many undiscovered advantages to be developed and utilized. With the rapid development of domestic fruit industry, the market demand of intelligent fruit picking robots has greatly increased, which makes the demand of agricultural productivity improve the application value and practical significance of intelligent fruit picking robots.
The vision system of robot is the key of intelligent picking robot, which is closely related to the significance of eyes to human beings, and how to improve the recognition accuracy and positioning accuracy of picking robot. For example, Jhawar et al. [6] proposed a pattern recognition method based on nearest neighbor and linear regression to determine the maturity of oranges. Ramos et al. [7] proposed a non-destructive method based on linear estimation to classify the number and harvestable fruits of coffee trees. Sofu et al. [8] put forward an automatic system for apple variety sorting and quality evaluation. In this method, two industrial cameras fixed on the conveyor belt capture four photos of a single apple in real time, and classify them according to color, size and weight by classifier algorithm. However, most of the above methods can only classify the maturity or size, and cannot detect the defective apples affected by staining, scab and decay from the images. Therefore, this paper proposes an intelligent sorting and grading method for processing apples. Accurate classification of mature and immature apples with digital images is a completely nonlinear problem. Support Vector Machine (SVM) [9-10] has been successfully used to realize nonlinear systems characterized by complex decision boundaries, which proves its effectiveness in various pattern classification problems.

Twin support vector machines (TWSVM) is a new machine learning method based on statistical learning theory. As a variant algorithm of traditional SVM, TWSVM inherits its excellent learning ability, but its running efficiency is increased by 4 times. Therefore, this paper uses basic SVM classifier to classify mature apples and immature apples, and uses TWSVM to detect defects. Using the collected apple images and performance evaluation indicators, a specific system experiment was carried out, and the results verified the effectiveness of the system in maturity detection and disease classification.

2. System design

2.1. Apple image acquisition and preprocessing

In this paper, a hybrid contrast stretching method is used to improve the image contrast by using a Top Hat [11-12] filter and Gaussian function, and to capture the target on the image surface to classify apples. Top-hat transform is an advanced morphological operation, which has great potential in extracting micro-structures and fine features from images. The input images were respectively transformed by white top hat and filtered by Gaussian. The Gaussian removes noise and smoothes the image. Let the original RGB image with a size of 256×256×3 be represented, and the white-topped hat filtering image $T_{hat}(s,t)$ is given by Equation (1).

$$T_{hat}(s,t) = I(s,t) - I(s,t) * M$$  \hspace{1cm} (1)

Where $M$ is the structure element, which is initialized to the value 9.

Image $T_{hat}(s,t)$ contains the fine objects needed to classify the apple image. The Gaussian function given in Equation (2) is applied to smooth the image.

$$G(s,t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(I-\mu)^2}{2\sigma^2}}$$  \hspace{1cm} (2)

Where $I$ is the input RGB image, $\mu$ is the average value of the input image, and $\sigma$ is the scaling parameter.

A differential image $T_n(s,t)$ is obtained by equation (3).

$$T_n(s,t) = G(s,t) - I(s,t)$$  \hspace{1cm} (3)

Use image $T_n(s,t)$ to calculate the value of average factor.

$$\alpha = \frac{\max(T_n(s,t)) + \min(T_n(s,t))}{2}$$  \hspace{1cm} (4)
The average factor $\alpha$ is used to further smooth the differential images to eliminate false positives in the classification. This factor is added to the top-hat filtered image to generate $T_{\text{new}}(s,t)$ as shown in Equation (5), which is contrast enhanced without artifacts.

$$T_{\text{new}}(s,t) = T_{\text{hat}}(s,t) + \alpha$$

Finally, the contrast-improved image $T_{\text{new}}(s,t)$ is added to the Gaussian image to generate $I_\ell$ as shown in equation (6) where the visual contrast of the infected area is enhanced compared to the rest of the image.

$$I_\ell(s,t) = T_{\text{new}}(s,t) + G(s,t)$$

2.2. Feature extraction

The classification process of the system is divided into two stages, including maturity detection and disease classification, and adopts 24 features, including color, texture, shape and geometric attributes. Subsets of these features are used in different classification stages.

The red, green and blue components of the image are represented as R, G and B, respectively, and 18 color features from $f_1$ to $f_{18}$ are extracted from the R, G and B components. The RGB images were converted into gray images and the gray co-occurrence matrix texture features $f_{19}$ to $f_{22}$, i.e. contrast, correlation, energy and uniformity, were extracted. The feature $\chi$ is extracted from the first moment of the Hu invariant moment of the candidate image. The number of white pixels $f_{24}$ is the number of white pixels extracted from the original image by Gabor wavelet transform.

2.3. Gabor wavelet transform

Gabor wavelets have been widely applied to classification problems such as feature extraction and texture analysis of digital images [13–14]. The Gabor function is exponentially distributed near $y = 0$, as shown in Equation (7).

$$g_{u,v}(y) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha y^2} e^{-i\varepsilon y}$$

Where $\alpha^2$ is the variance and $\varepsilon$ is the frequency.

Gabor wavelet transform expands the input signal into a set of functions in time domain and frequency domain. The two-dimensional Gabor function $g(x,y)$ and its Fourier transform $g(U,V)$ are shown in Equation (8) and Equation (9).

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j\omega\right]$$

$$G(U,V) = \exp\left[-\frac{1}{2}\left(\frac{(U-\omega)^2}{\sigma_U^2} + \frac{V^2}{\sigma_V^2}\right)\right]$$

Where $\sigma_x$ and $\sigma_y$ are the standard deviations of $g(x,y)$. Here, $\sigma_U$ and $\sigma_V$ are denoted as $\sigma_U = 1/(2\pi\sigma_x)$ and $\sigma_V = 1/(2\pi\sigma_y)$.

In the spatial domain, the two-dimensional Gabor function is the product of the complex plane wave and the elliptic Gaussian function. In the frequency domain, the Gabor function is a two-dimensional elliptic gaussian function. In this system, the candidate images undergo multi-directional and multi-scale
Gabor wavelet transform, which is sufficient to extract features from the apple surface for disease detection.

2.4. Twin support vector machines

As an improved version of traditional machine learning, TWSVM is looking for a pair of non-parallel hyperplanes, so it has better classification ability and is very suitable to solve the problem of sample classification of approximate type. In addition, compared with the traditional SVM, TWSVM performs two SVM-type problem solving, so the computational efficiency is higher. The time complexity of the standard SVM is about $O(m^3)$, while that of the TWSVM is $O(2^*(m/2)^3)$, and $m$ represents the number of samples. It can be seen that the computational cost is about 1/4 of that of the standard SVM.

In actual application cases, most data samples are not simple binary classification. If simple linear TWSVM classification is used in the feature space of genomic data studied in this paper, satisfactory classification results can not be obtained. Therefore, for the nonlinear classification problem, that is, the linear non-time-sharing, we need to introduce a kernel function to solve the problem. The training sample set in the dimensional real number space is assumed to be $(x_i^j, y_i^j), \ i = 1, 2, \ j = 1, 2, ..., m$. The total number of samples was $m = m_1 + m_2$, where $m_1$ was the number of positive sample points and $m_2$ was the number of negative sample points. Then the way to find the nonlinear TWSVM hyperplane is:

$$K(x_i^T, C_T^T)u_1 + b_1 = 0, \ K(x_j^T, C_T^T)u_2 + b_2 = 0 \quad (10)$$

Where $u$ is the normal vector of the hyperplane, $b$ is the offset, and the subscript symbols of the two respectively represent positive class samples and negative class samples. $C^T = [A^T B^T]^T$, $K$ is the kernel function, and the Gaussian kernel radial basis function is used as the kernel function of the twin support vector machine. $A$ is a sample matrix consisting of $m_1 \times n$ positive class samples, and $B$ is a matrix consisting of $m_2 \times n$ negative class samples.

SVM is basically a binary classifier, which is a powerful tool for nonlinear classification, regression and multivariate function estimation. Its goal is to maximize the distance between the hyperplane and the boundary. The optimal boundary is defined as follows:

$$f(x) = w \cdot x + m \quad (11)$$

By solving the optimization problem represented, the values of the parameter $w$ and the parameter $m$ can be obtained.

$$\text{Minimize } \frac{1}{2} |w|^2 + C \sum \varepsilon_j \quad (12)$$

Where $C$ is the penalty parameter and $\varepsilon$ is the error, $\text{lab}_j(w \cdot x_j + b) \geq 1 - \varepsilon_j, \ j = 1, 2, 3, ... n$.

Implementations of TWSVM are available in the LIBSVM open source library for machine learning. In the proposed system, the apple maturity is distinguished by the basic SVM algorithm, and the disease classification is realized by TWSVM algorithm.

3. System implementation steps

After the image acquisition processing, the classification processing is performed. The implementation process of the proposed system is divided into two stages, as follows:

1) Maturity detection: this process is performed using a basic binary SVM classifier. A feature vector consisting of 14 features $\{f_1, f_2, ..., f_{14}\}$ was constructed from mature/immature apples and was specifically used to determine the maturity of the fruit.

2) Disease identification: this process was used to detect the black spot, ulcer and melanin infection in the defective apple sample. The candidate images are segmented in multi-scale and multi-direction by the Gabor filter to separate the infected area. The extent of the infected area is determined according to the number of white pixels in the segmented area. A feature vector, i.e., a defect vector, is constructed
from each candidate image, and the feature vector includes a set of 10 features \( \{ f_1, f_2, \ldots, f_{24} \} \) to clearly distinguish defects. In this stage, TWSVM is used to train the defect vector and classify the samples.

4. Experimental results and analysis

4.1. Experimental data set

The experimental system uses 1394 image acquisition card, as shown in Fig. 1.

The 1394 card interface occupies little resources of the host processor when transmitting data, can accurately control the transmission packet length, has no signal loss, is more suitable for multi-channel image acquisition and transmission at the same time, improves the real-time performance of experiments, and well meets the requirements of experiments on a visual system. The camera model is DH-HV1300FC with a measured maximum distance of 0.6m and a resolution of 12801024.

Images of 266 apple fruits, including ripe, immature, dark spots, canker and melanosis fruits, were collected. The training and test sets were constructed in a 1:1 ratio using these images. The parameter settings for the training and test sets are shown in table 1.

| Image type      | total quantity | Number of training | Number of test |
|-----------------|----------------|--------------------|----------------|
| Healthy mature  | 82             | 41                 | 41             |
| Health immaturity | 82           | 41                 | 41             |
| Black spot      | 26             | 13                 | 13             |
| Canker          | 48             | 24                 | 24             |
| Melanosis       | 28             | 14                 | 14             |

The candidate images were subjected to Gabor wavelet transform in four different scales and four directions to segment the infected area without overlapping with the healthy area. Taking healthy and mature images as reference, each test image is subtracted from the reference image to generate a difference image, and the difference image is subjected to Gabor wavelet transform in four scales and four directions.

4.2. Evaluation Indicators

For specific analysis, quantified performance metrics [15] were used: Accuracy, Specificity, Sensitivity, and Precision. These metrics are defined as follows:

\[
\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (13)
\]

\[
\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (14)
\]
\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{16}
\]

Where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

4.3. Analysis of results

To make a fair comparison, we evaluated the performance of the other three classifier models using MATLAB 2017. The overall performance index comparison of the classifiers is shown in Fig. 2.

![Fig. 2 Comparison of overall performance indicators of classifiers](image)

It can be seen that compared with other classifier models, the proposed system has the best performance in all performance indicators.

5. Conclusion

This paper presents a vision system of apple picking robot based on twin support vector machine. The classification process of the system is divided into two stages, including maturity detection and disease classification, and 24 features are used for image feature extraction. The basic SVM algorithm was used to distinguish the maturity of apple and TWSVM algorithm was used to achieve disease classification. The test results of four quantified performance indexes verify the effectiveness of the system. The system has the advantages of simple setup and high reliability, and is suitable for a variety of application scenarios in the field of apple picking.

Acknowledgments

Project: 2019 Curriculum teaching Reform project "Robot Technology" (NY-2019CQ2KC-45).

References

[1] Ht A, Tw A, Yl A, et al. Computer vision technology in agricultural automation —A review - ScienceDirect[J]. Information Processing in Agriculture, 2020,7(1):1-19.
[2] Zhao W, Wang X, Qi B, et al. Ground-Level Mapping and Navigating for Agriculture Based on IoT and Computer Vision [J]. IEEE Access, 2020(8):221975-221985.
[3] Terekhin E A. Recognition of abandoned agricultural lands using seasonal NDVI values[J]. Computer Optics, 2017, 41(5):719–725.
[4] Qizi A. Methods of Teaching Exact and Natural Sciences Using Computer Technologies to Agricultural Students [J]. International Journal of Psychosocial Rehabilitation, 2020, 24(5):7453-7460.
[5] Hashim N, Ad Ebayo S E, Abdan K, et al. Comparative study of transform-based image texture analysis for the evaluation of banana quality using an optical backscattering system [J]. Postharvest Biology and Technology, 2018(135):38-50.

[6] Jhawar, Jyoti. Orange Sorting by Applying Pattern Recognition on Colour Image [J]. Procedia Computer Science, 2016(78):691-697.

[7] Ramos P J, Prieto F A, Montoya E C. Automatic fruit count on coffee branches using computer vision [J], Computers and Electronics in Agriculture, 2017(137): 9–22.

[8] Sofu M M, Er O, Kayacan M C, et al. Design of an automatic apple sorting system using machine vision [J]. Computers and Electronics in Agriculture, 2016(127): 395-405.

[9] Jiang J, Zhang W, Guo D, et al. Prediction of real contact area of fractal rough surface based on FEM contact model results and support vector machine algorithm [J]. Journal of Computational Methods in Sciences and Engineering, 2020, 21(2):1-9.

[10] Martinez-Castillo C, Astray G, Mejuto J C, et al. Random Forest, Artificial Neural Network, and Support Vector Machine Models for Honey Classification [J]. eFood, 2020, 1(1):69-76.

[11] Zhu H, Liu S, Deng L, et al. Infrared Small Target Detection via Low-Rank Tensor Completion With Top-Hat Regularization [J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 2(58):1004-1016.

[12] Zhou W, Liu H, Zeng W G. Improved recognition algorithm of knuckles and palmprints based on top hat transformation [J]. Sensors and Microsystems, 2019, 38(03):133-136.

[13] Radivojevi D S, Mirkov N S, Maleti S. Human Activity Recognition based on Machine Learning Classification of Smartwatch Accelerometer Dataset [J]. FME Transactions, 2021, 49(1):225-232.

[14] Milosevic N, Dehghantanha A, Choo K. Machine learning aided Android malware classification [J]. Computers & Electrical Engineering, 2017(61):266-274.

[15] Soofi A, Awan A. Classification Techniques in Machine Learning: Applications and Issues [J]. Journal of Basic & Applied Sciences, 2017(13):459-465.