Apple Classification Based on Information Fusion of Internal and External Qualities

Xue Li, Liyao Ma, Shuhui Bi(✉), and Tao Shen

School of Electrical Engineering, University of Jinan, Jinan 250002, China
cse_bish@ujn.edu.cn

Abstract. Apple classification plays an important role in improving the sales of apples. Based on both the internal and external qualities of an apple, in this paper, we propose to classify apples by DS theory-based information fusion. Soluble solid content is selected for apple internal quality detection. Making near-infrared spectroscopy nondestructive testing, principal component analysis -Martensitic distance method and multiple Scattering correction are used to preprocess the spectral data collected. Partial least squares prediction model is established with genetic algorithm selecting the wavelength characteristics. The color, shape, diameter and defect of apple are taken as the important indexes of external quality detection, and the sample images are analyzed and studied. The RGB color model and HSI color model commonly used in image processing are introduced. Selecting the median filtering algorithm for image denoising, the prediction model of support vector machine is established. In order to effectively avoid the classification error caused by the traditional hard classification using threshold and to make the detection result more accurate, the analysis of uncertain factors was introduced in the aspect of apple classification, and DS evidence theory was used to fuse the prediction results of internal and external quality.

Keywords: Apple classification · Support vector machine · DS evidence theory · Partial least squares

1 Introduction

Apple is one of the most sold fruits in the world, as well as in China. It not only tastes sweet and sour, but also has a high nutritional value, with a "wisdom fruit" reputation. China has a long history of fruit cultivation and rich variety resources. It is the largest fruit producer in the world. The planting area and total output of orchards rank first in the world [1]. Data released by the Key Agricultural Product Market Information platform of the Ministry of Agriculture shows that the export of apples in China has been increasing gradually since 2010. According to statistics, the world's apple export in 2017 was about 76.21 million tons, of which China's export exceeded 43.43 million tons. As the world's largest apple producer, China accounts for more than 50% of the world's planting area, total output and consumption scale. However, according to the DATA of the United States Department of Agriculture (USDA), in 2017, the world's apple exporters were mainly

W. Fu et al. (Eds.): ICMTEL 2021, LNICST 387, pp. 388–397, 2021.
https://doi.org/10.1007/978-3-030-82562-1_36

European and American countries [2]. The main reason for the fact that the proportion of Apple exports in China is lower than that in Europe and The United States is that China's apples are not classified and packaged in strict accordance with the standards due to outdated post-harvest commercialization processing technology. The post-processing commercialization rate of apples in developed countries is over 90%, while that in China is less than 40% [3, 4]. Therefore, improving the grading level of apples after picking is the key to improving the competitiveness of Chinese apples in the international market. With the improvement of people's living standard, consumers' requirements on fruit are not limited to price, but pay more attention to quality, brand and other aspects. The diversified demands of consumers for apple's appearance, taste, nutrition, function and brand have determined the necessity and urgency of researching postpartum commodity processing technology. Classifying apples according to their appearance and taste not only meets the consumption needs of different groups of people, but also helps to improve the commodity value of apples.

Information fusion combines information from multiple sources, reducing its uncertainty and improving the accuracy of modern intelligent information systems in decision-making, planning and response. Since the 1970s, it has been widely applied in various aspects of military and national economy [5, 6]. DS evidence Theory [7, 8] introduces basic probability assignment and belief function, which can distinguish between "unknown" and "uncertain". The reasoning mechanism of DS theory is simple, and it is close to the thinking habit of human beings, showing its unique advantages [9].

In this paper, DS evidence theory is used to fuse the internal and external qualities of apples. For information fusion, both models for internal and external qualities are established. The final decision is then made with Demspter's rule of combination.

2 Preprocessing of Apple Data

2.1 Physicochemical Analysis of Apple's Internal Quality Data

439 red Fuji apples were selected as the research samples, and 32 sampling modules were scanned at room temperature to collect the near-infrared spectrum of apples. The spectra of different parts of the equator were collected 3 times for each sample, and the average spectrum was calculated as the sample spectrum. Due to the influence of sample background and other factors, noise interference and baseline drift often occur in the NIR spectrum, which greatly affect the accuracy of the model. Therefore, we need to preprocess the spectrum before modeling.

The accuracy of the sample data directly determines the validity of the model architecture. Abnormal parameters will reduce the stability and accuracy of the sample model [10]. Therefore, it is very important to screen out abnormal parameters and store reasonable parameters. Pa-md method is adopted to calculate and eliminate abnormal samples in the original spectrum.

Multiple scatter correction (MSC) is currently the world's one of the most common and most effective means of data processing, can be in the process of spectral measurement optical path difference to a certain extent, the correction, strengthen its detailed information. In this paper, MSC is used to deal with reflected light.

The Savitzky-Golay smoothing filter can improve the spectral smoothness and reduce the noise interference. According to the characteristics of the spectrum, combined with the modeling effect, the window width of 5 time leads to the best processing effect.

The spectral information was pretreated in the early stage, and the noise content was reduced. But the results are still not directly applicable. Further processing is needed to further reduce the redundant information. For this purpose, genetic algorithm (GA) is used to select the optimal characteristic wavelength and take it as the input of partial least squares (PLS) based prediction model.

2.2 Analysis of Apple External Quality Data

Taking Fuji apple as the research object, 439 apple samples were collected for grading processing. The features of each apple were extracted from four aspects including color, fruit shape, fruit diameter and defect, and a 7-dimensional feature vector was constructed for training and testing the grading model. During the processing of Apple's external quality, the images used are taken by CCD industrial camera, and all the collected images are color images.

It is very important to select a suitable color model for color image processing. The RGB color model is easy to understand and has a very good effect in the implementation of hardware devices. The H and S components of the HSI model are in line with the way people perceive colors, and the I component is not reflected in the color information of the image. Therefore, image processing is carried out based on the RGB and HIS color models.

Only when the apple area is completely segmented can the external quality features such as apple color and defects be accurately and efficiently extracted. However, the collected images contain background information as well as apple area [11]. Therefore, the background segmentation method with bimodal threshold of 15 can be used to completely segment the apple region, completely eliminating the interference of the background part and achieving a better effect. In the process of image acquisition and transmission, noise will be generated in the original image under the influence of light source and sensor, etc. Noise is usually random and unpredictable, which is difficult to conduct quantitative analysis [12]. Median filtering algorithm is selected to carry out filtering denoising process. Both color and gray images can retain the original image structure completely, and the filtering effect is the best.

When extracting the external features of apple such as color, shape, diameter and defect, the ratio of red and near-red H values in the apple image was taken as the color index, and the variance of R, G and B components was selected as the color distribution parameter by Fisher's coefficient method. Based on the obtained apple edge, the horizontal and vertical diameters of the apple image are obtained by calculating the minimum enclosing rectangle of the apple, and the shape index is calculated. The pixel diameter and pixel equivalent are obtained by calculating the minimum enclosing circle of an apple image, and the conversion between the actual diameter and the pixel diameter is realized. Since the gray level of the defect area is different from the normal apple area, the common defect area of apple is detected based on the Canny algorithm, and morphological operation and hole filling are introduced into the defect area to segment the defect area, and the ratio of the defect area to the apple area is taken as the defect feature [13, 14].

After extracting the characteristics of apple's external quality, the apple's external quality was graded based on support vector machine.

3 PLS Model for Internal Quality of Apples

PLS is one of the most common methods to construct regression model. It has a good design effect whether it is to analyze a single variable or to deal with multiple different variables at the same time.

The optimal characteristic wavelength selected by genetic algorithm is used as the input of PLS based prediction models. When using PLS to analyze SSC, the collected NIR data and soluble solid content were defined as the main components, and the absorbance and content matrix were used as independent variables. PLS synthesized the external relationship between the spectral matrix X and W. The internal relation between the two is obtained, and then X and W are decomposed into the following forms:

$$\mathrm{X} = \mathrm{TP} + \mathrm{E} \tag{1}$$

$$\mathrm{W} = \mathrm{VQ} + \mathrm{F} \tag{2}$$

$$\mathrm{V} = \mathrm{TB} \tag{3}$$

$$\mathrm{B} = (\mathrm{T^T T})^{-1}\mathrm{T^T W} \tag{4}$$

In which T and V are the score matrix of matrix X and W respectively, P and Q are the load matrix of the two matrices respectively, and E and F are the residual matrix of the two.

Finally, according to the formula above, the synthesis matrix B is obtained. B matrix is used to predict the result of the sample to be tested. Collect all kinds of data of test samples, and then decompose them according to formula 1 and 4 to find out the sample concentration W. As shown in Eq. (5).

$$W = tB \tag{5}$$

in which W is the concentration value of the sample to be tested, T is the spectral decomposition score of the samples to be tested.

4 SVM Model for External Quality of Apples

In the mid-1990s, Support Vector Machine (SVM) was proposed by Vapnik et al. [15]. SVM is a dichotomous question, and its main purpose is to maximize two types of sample data by looking for the optimal hyperplane interval. The white dots and the black dots

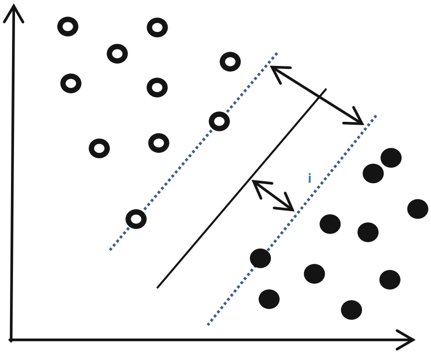

Fig. 1. The classification schematic diagram

in Fig. 1 represent the two types of sample data. The main purpose is to look for a line with the greatest distance between all the characteristic points to divide the two samples. Classification problems using SVM can be divided into linear separable, approximate linear separable and nonlinear separable.

Two linearly separable sample sets are assumed to be in the form

$$D = (Xi, Yi) \tag{6}$$

in which Xi ∈ R, Yi ∈ {−1, + 1}, i = 1,2,…,n.

Therefore, the linear discriminant function can be described as

$$h(x) = \alpha \cdot x + \beta \tag{7}$$

The corresponding segmentation surface equation is shown as

$$\alpha \cdot x + \beta = 0 \tag{8}$$

in which α is normal vector, β is the intercept.

Based on the segmentation surface, the feature space can be divided into positive and negative parts. To correctly classify all samples, the conditions as shown in Eq. (9) should be met:

$$y_i \cdot h(x) - 1 \geq 0, i = 1, 2, \ldots, n \tag{9}$$

The distance from x_i to the hyperplane can be calculated as

$$d = \frac{|h(x)|}{\alpha} = \frac{|\alpha \cdot x + \beta|}{\alpha} \tag{10}$$

Since all the points on the optimal hyperplane satisfy that the molecule is 1, so, the distance between the support vector and the hyperplane is expressed as

$$d = \frac{|\alpha \cdot x + \beta|}{\alpha} = \frac{1}{\alpha} \tag{11}$$

At this point, the classification interval is $\frac{2}{\|\alpha\|}$, If the interval is maximized, it is equivalent to making $\|\alpha\|^2$ the smallest. Therefore, the optimal classification surface can be transformed into

$$\min f(\alpha) = \frac{a^2}{2} \tag{12}$$

Lagrange function is defined as

$$L(\alpha, \beta, c_i) = \frac{\alpha^2}{2} - \sum_{i=1}^{n} c_i[y_i(\alpha \cdot x_i + \beta) - 1] \tag{13}$$

where c_i is Lagrangian multiplier.

Further, in order to calculate the minimum objective function, respectively for α, β partial derivatives and make them are equal to zero, with

$$L(\alpha, \beta, c_i) = \frac{\alpha^2}{2} - \sum_{i=1}^{n} c_i[y_i(\alpha \cdot x_i + \beta) - 1] \tag{14}$$

Therefore, the above problem is transformed into a dual problem of convex quadratic programming optimization, as shown in Eq. (15).

$$\begin{cases} \max \sum_{1}^{n} c_i - \frac{1}{2} \sum_{1}^{n} \sum_{1}^{n} c_i c_j y_i y_j (x_i x_j) \\ \sum_{1}^{n} c_i y_i = 0 \\ c_i \geq 0, i = 1, 2, \ldots, n \end{cases} \tag{15}$$

Assuming c_i^* is the optimal solution, the optimal hyperplane normal vector can be expressed as

$$\alpha^* = \sum_{1}^{n} c_i^* y_i x_i \tag{16}$$

The optimal classification function is

$$f(x) = \operatorname{sgn}\left\{ \sum_{1}^{n} c_i^* y_i (x_i \cdot x) + \beta^* \right\} \tag{17}$$

Nonlinear classification is when the sample data set is indivisible, the input feature vector is mapped to the high-dimensional space and the optimal hyperplane is obtained by selecting the appropriate kernel function.

Assume that A is the input space, B is the high-dimensional feature space, and there is A mapping function (x) from A space to B space. For all real numbers in A space, if there is A function that satisfies the relation

$$k(s, t) = \rho(s) \cdot \rho(t) \tag{18}$$

Then $k(s, t)$ is called the kernel function. Consider the objective function as

$$f(c) = \sum_{1}^{n} c_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j y_i y_j k(s, t) \tag{19}$$

Equation (15) is converted into the form

$$\begin{cases} \max \sum_{1}^{n} c_i - \frac{1}{2}\sum_{1}^{n}\sum_{1}^{n} c_i c_j y_i y_j k(s, t) \\ \sum_{1}^{n} c_i y_i = 0 \\ 0 \le c_i, i = 1, 2, \ldots, n \end{cases} \tag{20}$$

If classification or regression is performed directly in a high-dimensional space, there will be problems in determining the form and parameters of the nonlinear mapping function, the dimension of the feature space and etc. The kernel function can effectively solve these problems. Radial basis kernel function is one of the most widely used kernel functions. Compared with polynomial kernel function, it has strong locality and few parameters. As a result, this paper takes RBF as the kernel function of SVM.

The traditional support vector machine (SVM) is a typical two - class classifier. A one-to-many method to construct the multi-classifier is to sort out the samples of a certain category into one group, and the rest samples are classified into another group in the training, so that the samples in k categories can construct k SVMs. When doing this, the unknown samples will be classified into the categories with the maximum classification function value. The brief process of SVM is shown in Fig. 2.

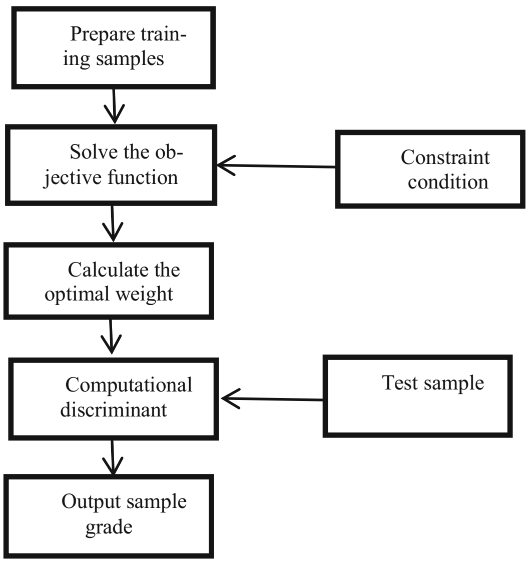

Fig. 2. Flow chart of support vector machine classification processing

5 Apple Classification Fusion Based on DS Evidence Theory

The previous apple classification models both make a hard division of the sample space, which easily leads to misclassification in areas near the decision boundary. In order to reduce the difference of results caused by hard classification, this paper introduces DS theory, which has a strong analysis advantage in dealing with uncertainty.

5.1 Method Description

The DS evidence theory was first proposed by Dempster in 1967 and then extended by Shafer in 1976 [7, 8]. When using belief functions to deal with problems, there is no need to generate accurate probability for each possible outcomes. The basic mass assignments provide a useful way to represent unknown and uncertainty.

Consider the previously discussed internal and external models, each model provide a basic mass assignment m of apple classification, which provides its belief of predicted result. Since each model has its own reliability, each basic mass assignment should be discounted by an evidence discount rate $\alpha \in [0, 1]$ with the following discounting formula

$$m^{\alpha}(U) = (1 - \alpha)m(U) + \alpha \tag{21}$$

$$m^{\alpha}(A) = (1 - \alpha)m(A), \forall A \subset U, A \neq \varphi \tag{22}$$

In which U is the frame of discernment, which contains all the possible outcomes.

Finally, the information fusion of two models is implemented by the combination of evidence. Let m_1 and m_2 be two independent basic probability assignments on 2^U, the fusion results in another basic probability assignment denoted as $m = m_1 \oplus m_2$, which is obtained with Dempster's rule of combination

$$K_1 = \sum\nolimits_{A_i \cap B_j = \phi} m_1(A_i)m_2(B_j) \tag{23}$$

$$m(C) = \begin{cases} \sum_{A_i \cap B_j = C} m_1(A_i)m_2(B_j), \ \phi \neq C \subset U \\ 0, C = \phi \end{cases} \tag{24}$$

5.2 Realization of Apple Classification Based on DS Evidence Theory

Suppose the Red fuji apples can be classified into three classes A, B and C. We have the frame of discernment $U = \{A, B, C\}$ and corresponding focal elements $F \in 2^U = \{A, B, C, AB, AC, BC, ABC\}$.

The PLS model provides m_1 considering the prediction based on internal quality. Similarly, the SVM model gives m_2 regarding to the external quality. Table 1 shows an example of the basic mass assignments of these two models.

The PLS soluble solid content prediction model and ELM external feature prediction model were combined with Demspter's rule, with the fusion result of

$$m_1 \oplus m_2(A) = 0.88, \quad m_1 \oplus m_2(AB) = 0.09, \quad m_1 \oplus m_2(ABC) = 0.03$$

Table 1. Basic mass assignment table

Prediction model	A	B	C	AB	AC	BC	ABC
PLS model m_1	0.6	0	0	0.3	0	0	0.1
ELM model m_2	0.7	0	0	0	0	0	0.3

From Table 1, it can be seen that these two prediction models provide relatively uncertain results (the basic mass assignment for A are 0.6 and 0.7). Yet when we combine the predictions together, the fusion result becomes much more certain that this apple belongs to class A.

6 Conclusions

In this paper, focused on red Fuji apples, the information fusion model of apple classification is discussed. The spectrum of apple was collected by near-infrared spectroscopy technology and then used for internal quality analysis. After spectral information preprocessing and spectral feature analysis, the prediction model of apple internal quality measurement was established based on PLS. The apple image was collected by machine vision technology and preprocessed with background segmentation and feature extraction. The prediction model of apple external quality was then established based on support vector machine. By using DS evidence theory method, the prediction model of apple internal quality measurement and the prediction model of external quality measurement were fused to improve the accuracy and certainty of Fuji Apple grading prediction model. In future work, more prediction models, as well as the generation approach of basic mass assignment will be considered.

Acknowledgements. This paper was supported by Shandong Provincial Key Research and Development Project (No. 2017GGX10116).

References

1. Zhixia, L., Jiyun, N., Jing, L., et al.: Analysis and suggestions on the development of Apple industry in China. Chin. Fruits **05**, 81–84 (2014)
2. Biao, Z.: Analysis on the annual production, processing and trade status of Apple industry in China in recent 7 years. China Fruit Tree **192**(04), 112–114 (2018)
3. Wu, M., et al.: Research on the status quo and development strategy of post-havest apple industry in China. **34**(10), 17 (2014)
4. Cerutti, A.K., Bruun, S., Donno, D., et al.: Environmental sustainability of traditional foods: the case of ancient apple cultivars in Northern Italy assessed by multifunctional LCA. J. Cleaner Prod. **52**, 245–252 (2013)
5. You, H., Guo-hong, W., Xin, G., et al.: Information Fusion Theory and Application. Electronic Industry Press, Beijing (2010)

6. Sun, B., Cheng, W., Ma, L., Goswami, P.: Anomaly-aware traffic prediction based on automated conditional information fusion. In: 21st International Conference on Information Fusion (FUSION), pp. 2283–2289 (2018)
7. Dempster, A.: Upper and lower probabilities induced by a multivalued mapping. Annals of Math. Stat. **38**(4), 325–339 (1967)
8. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton (1976)
9. Ma, L., Sun, B., Han, C.: Learning decision forest from evidential data: the random training set sampling approach. In: 4th International Conference on Systems and Informatics (ICSAI), pp. 1423–1428 (2017)
10. Lin, H., Zhang, H., Gao, Y., et al.: Hyperspectral identification of desert tree species based on Markov Distance method. Spectrosc. Spectral Anal. **34**(12), 3358–3362 (2014)
11. Wu, Y., Meng, T., Wu, S.: Research progress of image threshold segmentation method in 20 years (1994–2014). Data Acquisit. Process. **30**(1), 1–23 (2015)
12. Li, H., Suen, C.Y.: A novel non-local means image denoising method based on grey theory. Pattern Recogn. **49**, 237–248 (2016)
13. Qiang, L.: Research and Development of Apple Quality Grading Technology Based on Machine Vision. Heilongjiang University, Harbin (2019)
14. Li, S., Li, R., Du, G., Ding, S., Jiang, L., Liu, X.: Nondestructive identification analysis of oats of different brands based on near-infrared spectroscopy and optimized pretreatment. J. Food Saf. Qual. Inspection **10**(24), 8204–8210 (2019)
15. Zhang, H.: Research on multiple classification methods of support vector machine and its application in fund evaluation. Beijing Jiaotong University (2014)