J Bangladesh Agril Univ 18(3): 724–733, 2020 https://doi.org/10.5455/JBAU.108516

ISSN 1810-3030 (Print) 2408-8684 (Online)

Journal of Bangladesh Agricultural University

Journal home page: http://baures.bau.edu.bd/jbau

Apple Identity Recognition Based on SVM Model Parameter Optimization and Near Infrared Hyperspectral

Xiao Zhang[1], Chunlan Zhang[3], Huaping Luo[4], Xianchen Ren[1], Buxun Peng[1], Nannan Zhang[1,2]✉

[1]School of Information Engineering, Tarim University, Alar, Xinjiang 843300, China

[2]Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of education, China Agricultural University, Beijing 100083, China

[3]College of Life Sciences, Tarim University, Alar, Xinjiang 843300, China,

[4]School of Mechanical Electrification, Tarim University, Alar, Xinjiang 843300, China

ARTICLE INFO

Article history
Received: 23 Jul 2020
Accepted: 22 Sep 2020
Published online: 25 Sep 2020

Keywords
SVM; Apple; Hyperspectral; Identity recognition; Parameter optimization

Correspondence
Nannan Zhang
✉: zhangnannan@taru.edu.cn

ABSTRACT

Xinjiang Aksu apple with national geographical indication protection products is often counterfeited. In order to solve this problem, this paper uses hyperspectral imaging technology combined with SVM algorithm to identify apple samples from different producing areas and varieties. A total of 258 apple samples from different regions were collected by hyperspectral imager. The region of interest (ROI) of apple hyperspectral image was selected by ENVI software. Nine ROI were selected from the positive and nine from negative sides of each apple sample, and the average spectral value within the ROI was calculated. Then, SPA was used to reduce the dimension of the original spectral signal. Then, four different models were constructed by SVM, GS-SVM, GA-SVM and PSO-SVM. Different kernel functions were used to establish the hyperspectral apple classification prediction model. The results showed that the indicators of PSO-SVM with sigmoid as the kernel function were the best values. The accuracy was 91.6016%, precision was 96.1574%, recall was 88.6111%, F1 was 92.2269%. The model had high stability and prediction accuracy, which could meet the actual prediction needs. The results show that the near infrared hyperspectral based on SVM model parameters optimization can quickly identify apple species, which provides the basis for large-scale apple classification in the future, and also provides reference for standardizing apple trading market.

Introduction

Apple identification mainly refers to the classification and identification of varieties and producing areas. The varieties are influenced by genetic factors, while the producing areas are caused by environmental factors, soil characteristics and cultivation methods. The Red Fuji apple produced in Aksu area, due to the unique climate conditions such as large temperature difference between day and night, sugar gradually accumulates in the core and accumulates into a unique "ice sugar core" apple (Jiarui, 2019). The flesh is fine, crisp, juicy, and tastes excellent. It has become famous both at home and abroad as a brand characteristic, and has won the title of national geographical indication protection product. However, driven by interests, the phenomenon of counterfeiting Akesu "bingtangxin" apple has not been banned. For this reason, relevant departments in Aksu send a large number of professionals to fight against counterfeiting in the mainland every year. However, it is difficult to crack down on counterfeit goods and lacks technical support. Therefore, it is very urgent to study a fast identification method of apple identity features, which can not only make consumers understand the consumption, but also improve people's quality of life, and can better regulate the apple trading market, has important theoretical and practical significance.

At present, apple identity recognition mainly relies on sensory detection and chemical detection methods (Shunyu, 2015). There are many problems in sensory testing, such as strong subjectivity, inaccurate detection results, and chemical detection results are more accurate, but there are many problems such as time-consuming, laborious and expensive. Therefore, it is of great significance to find a fast and simple method to identify apple identity. Hyperspectral imaging technology has many advantages, such as continuous and numerous bands, high spectral resolution and

Cite This Article
Zhang, X., Zhang, C., Luo, H., Ren, X., Peng, B., Zhang, N. 2020. Apple Identity Recognition Based on SVM Model Parameter Optimization and Near Infrared Hyperspectral. *Journal of Bangladesh Agricultural University*, 18(3): 724–733. https://doi.org/10.5455/JBAU.108516

"combination of spectra". In recent years, it has become a research hotspot of agricultural product quality testing and has been widely used in the field of non-destructive testing of agricultural products. For example, in the detection of corn, millet, rice, bamboo leaves, tomatoes, walnuts, yuanzao, mango and other plant species in different regions and years for identification (JI Hai-yan, 2019, ZHAO Yi-kun, 2020, LIN Long, 2020, CHU Bing-quan, 2017, Mengjia et al., 2018, HE Yong, 2019, Wang, 2015, Pan, 2015, Yang et al., 2017, Williams and Kucheryavskiy, 2016, Zhang et al., 2012, Kong et al., 2013, Hao et al., 2010, Su et al., 2012, Liu et al., 2014, WANG Wei, 2019), but it is relatively rare to apply hyperspectral technology in the discrimination analysis of different kinds of apples. At the same time, because the full wave wavelength carries a lot of information, some of which are irrelevant to the establishment of the model, and even interfere with the modeling. Therefore, it is necessary to eliminate the useless information and select the relevant characteristic wave number for modeling.

Therefore, in this study, apple samples from different sources were taken as the research object. Firstly, SPA was selected to reduce the dimension of apple samples. Then, four different models, SVM(support vector machine),GS-SVM, GA-SVM and PSO-SVM, were used to establish the hyperspectral apple classification prediction model, and different kernel functions were used for debugging and comparison ,nondestructive identification of different kinds of apples provides a reference.

Materials and Methods

Test materials

Taking Xinjiang Aksu apple, Henan Lingbao apple, Gansu Jingning apple and Gansu Tianshui Huaniu apple as the experimental objects, all the apples used in the experiment were mailed in Xinjiang Aksu Hongqipo farm and other producing areas in October 2019. The selected apples had no surface defects, the diameter range was 65-85mm, and the size was uniform, a total of 258 apples, 144 of which were in Aksu area of Xinjiang, there were 82 samples of huaniu apple in Tianshui, 24 in Jingning and 8 in Lingbao, Henan. The purchased apples were stored in the freezer. Before the experiment, the apples were taken out in batches, and the experiment was started when the apples returned to room temperature. In the experiment, 1/2 samples of each type of apple were randomly selected as the modeling calibration set, and the remaining samples were used as the modeling prediction set.

Acquisition and correction of instrument and equipment and data

The hyperspectral system used in the experiment, as shown in Figure1, Zolix Gaia sorter (900~1700nm, 254 bands, spectral resolution 5nm, spectral sampling point 4nm).

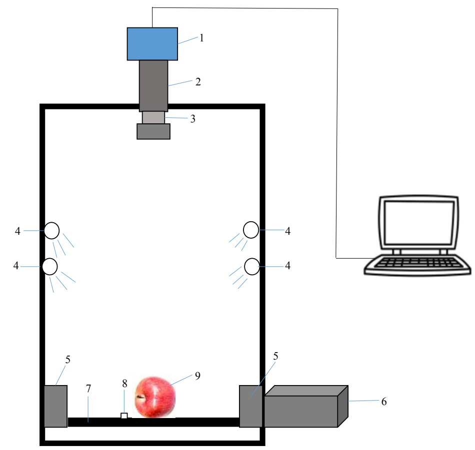


Figure 1. Schematic diagram of hyperspectral imaging system

1：CCD Camera, 2：Hyperspectral imager, 3：Imaging lens, 4：Light source, 5：Motor, 6：Motor controller, 7：Mobile platform, 8：Whiteboard, 9：Sample(Apple).

In order to reduce the influence of uneven illumination and dark current on the experiment, it is necessary to correct the collected hyperspectral data in black and white.

$$R=(I-B)/(W-B) \qquad (1)$$

In formula (1), I is the collected original hyperspectral data; B is the data collected by covering the camera lens (reflectance is close to 0); W is the data collected by aiming at the whiteboard (the reflectivity is close to 1); and R is the corrected hyperspectral data. The calibration tool is the SpewVIEW software which comes with the system, and the following softwares are Matlab 2016a, ENVI5.3 and Python 3.6.

SVM algorithm

The main idea of SVM algorithm (Sain and Stephan, 1997, Shankar et al., 2018, Gupta et al., 2018) is to establish a classification hyperplane as a decision surface, so that the isolation edge between positive and negative examples is maximized. It is an approximate realization of structural risk minimization, which can be used for pattern classification and non-linear regression, and is suitable for distinguishing different types of apple. There are four ways to select kernel function: linear kernel function, polynomial kernel function, radial basis function and sigmoid kernel function.

Model evaluation method

The evaluation criteria select the commonly used indicators in multi classification problems (Shanbai, 2019): accuracy, recall, precision and F1 evaluation model, and the calculation formula is as follows:

$$\text{Accuracy}=(TP+TN)/(TP+TN+FP+FN) \quad (2)$$
$$\text{recall}=TP/(TP+FN) \quad (3)$$
$$\text{precision}=TP/(TP+FP) \quad (4)$$
$$F1=2*\text{recall}*\text{precision}/(\text{recall}+\text{precision}) \quad (5)$$

In formula (2)-(5), TP indicates that positive samples are correctly predicted as positive samples; TN indicates negative samples are correctly predicted as negative samples; FN indicates positive samples are incorrectly predicted as negative samples; FP represents negative samples are wrongly predicted as positive samples; accuracy refers to the percentage of correct prediction results in total samples; recall refers to correct predictions in actual categories. The results show that the number of classification instances; precision represents the number of real and correct cases in the prediction category; F1 evaluation model represents that the precision and recall have the same importance, and the closer the four values are to 1, the better. In this study, the average values of the four indexes of the model were obtained by 20 repeated operations.

Results and Discussion

Sample original spectral curve and standard normal transformation (SNV)

The region of interest (ROI) of apple hyperspectral image is selected by ENVI software. As shown in Figure 2 (a), 9 ROI regions are taken from apple A (sunny side) and 9 ROI regions from apple B (shady side). The average value of 18 ROI regions is taken as a spectral record of the sample, as shown in Figure 2 (b). In the process of apple hyperspectral data collection, the irregularity of the sample surface will inevitably be caused the scattering phenomenon occurs, and then affects the real spectral information of apple. Therefore, SNV method is used to eliminate the influence of surface scattering, solid particle size and optical path change on near-infrared diffuse reflectance spectrum to achieve denoising effect (Chen et al., 2010), as shown in Figure 2 (c).

It can be seen from Figure 2(b) that the general trend and characteristic reflection peaks of different kinds of spectral curves are basically the same. There are two obvious reflection valleys near 1200 and 1430 nm, 1200 nm is the second harmonic absorption wavelength of C-H group, which represents the characteristic absorption peak of carbohydrate; 1400-1500nm is the first octave absorption wavelength of O-H and N-H groups, representing the characteristic reflection peaks of water and protein respectively. From Figure 2(c), it can be seen that the spectral curve of Gansu Jingning apple is not intersected with those of other 3 types of apples, which are easy to distinguish. After 1120nm, the spectral curves of the remaining 3 types of apples are significantly different, and the spectral curves processed by SNV can be classified into 4 types of apples.

PCA analysis

Principal component analysis (PCA) is an unsupervised analysis method. On the premise of ensuring the original information as much as possible, the dimension of multivariate data is reduced into a few new variables to reduce data redundancy, and then the differences of original variables can be understood and displayed. The PCA score map obtained from the first two principal components can directly show the sample state represented by the original data, and the aggregation and dispersion of sample points reflect the difference between sample points. As shown in Figure 3, the variance contribution rate of the first principal component is 79.49%, and that of the second principal component is 15.83%, totaling 95.32%. Therefore, the first two principal components can fully reflect the original data information.

The distribution of Xinjiang Aksu apple, Gansu Tianshui huaniu apple, Gansu Jingning apple and Henan Lingbao apple sample sites are relatively concentrated, indicating that the difference within each apple is small. At the same time, Xinjiang Aksu apple, Gansu Tianshui huaniu apple and Gansu Jingning Apple sample sites overlap, and the distribution of Henan Lingbao Apple sample points is relatively independent and the boundary is clear, indicating that Xinjiang Aksu apple, Gansu Tianshui huaniu apple and Gansu Jingning apple samples have small differences, while Henan Lingbao Apple samples have obvious differences. The above results showed that it was difficult to distinguish Xinjiang Aksu apple, Gansu Tianshui huaniu apple and Gansu Jingning apple in principal component analysis, but it was easier to distinguish Henan Lingbao Apple.

SPA dimension reduction and feature band extraction

Due to the high dimension of the original spectral data, if it is directly used as the input variable of SVM, the model will be too complex and the running time will be too long. At the same time, redundant spectral data will reduce the prediction accuracy of the model. The successive projections algorithm (SPA) selects the wavelength combination with the least linear relationship by projection, and finds the variable group with the lowest redundant information from the spectral information, so as to minimize the collinearity between variables, and retain most of the characteristics of the original data. The selected characteristic wavelength has clear physical meaning and strong explanatory power Force, therefore, can effectively improve the speed of modeling and the stability of the model, so as to achieve the purpose of simplifying the model (Fan et al., 2019). Therefore, before establishing SVM discriminant model, SPA algorithm is used for dimension reduction feature extraction. The optimal sample set of SPA model is selected by calculating the root mean square error (RMSE) of multiple linear regression models with different local subsets. When the RMSE value is the lowest, the subset represented is the optimal sample subset. Figure 4 (a) shows the RMSE values of different subset models, in which "□" represents the sample number of the optimal sample subset. From Figure 4 (a), it can be seen that when the number of variables is less than 19, the RMSE value as a whole shows a downward trend, and when the number of variables is greater than or equal to 19, the change trend tends to slow down. Therefore, this paper uses SPA to select 19 characteristic variables. Figure 4 (b) shows the selection of specific variables, and "□" represents the selected variables, and 19 wavebands are: 932.90002 nm, 1005.31 nm, 1037.1 nm, SP1169.35 nm, 907.95001 nm, 967.409997 nm, 1218.51 nm, 951.70001 nm, 1130.3199 nm, 1261.43999 nm, 1261.43999 nm, 1422.67 nm, 1324.76 nm, 1065.87 nm, 1104.45 nm, 1676.01 nm, 1707.92 nm, 1460.22 nm, 1392.12 nm, 942.28998 nm, its importance decreases in turn.

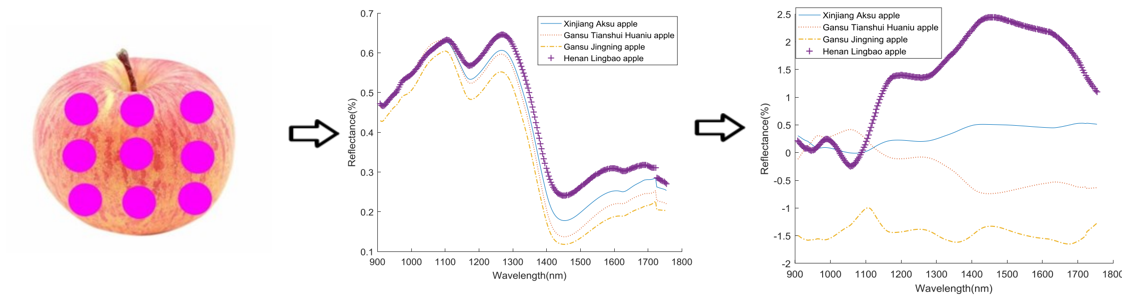


Figure 2. Main steps for image and spectra processing of apple fruit: (a) Identification of the Region of Interest (ROI), (b) Raw mean reflectance spectrum, and (c) Corrected spectral curve by SNV

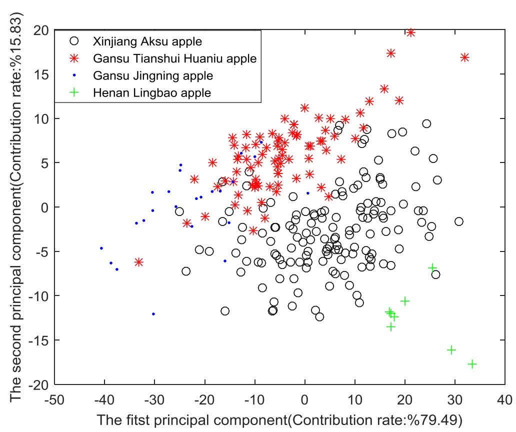
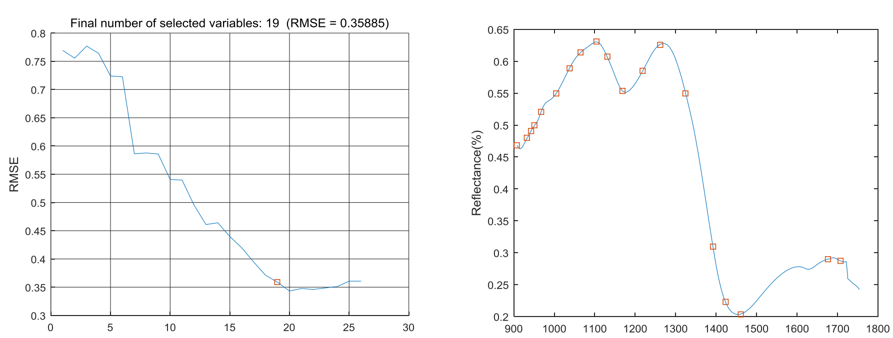


Figure 3. Principal component analysis



a. Number of screened variables

b. Distribution of characteristic wavelengths

Figure 4. Characteristic wavelengths extracted by SPA

SVM modeling and optimization

The data used in the latter four models are all SNV + SPA preprocessed data.

Modeling analysis of SVM

With 19 bands extracted by SPA algorithm as input and 4 kinds of apples as output, a recognition model based on SVM classifier is established. The libsvm 3.24 software package (Lei, 2011) developed and designed by Lin Zhiren of Taiwan University is used to predict four kinds of apples, and the parameters of SVM algorithm are set by default.

Four kernel functions of SVM, linear, polynomial, radial basis function and sigmoid, were used to construct hyperspectral apple classification prediction model. It can be seen from Table 1 that the prediction effect is better when polynomial is kernel function, accuracy is 89.8438%, precision is 95.4545%, recall is 87.5000% and F1 is 91.3043%. The prediction results of apple classification are shown in Figure 5.

GS-SVM

The grid search algorithm (Wang et al., 2020, Xiaoping Wang, 2020) is to search the spatial grid composed of the parameters to be solved according to a certain step size, and traverse all the points in the grid to find the optimal parameters. This method is a grid algorithm for comprehensive search. Generally, the parameters obtained by the grid algorithm are optimal and will not fall into local optimum. The second search interval should not be too large or too small. If it is too large, it will easily lead to long search time, and too small will easily lead to local optimization. In order to verify the performance of GS-SVM algorithm, the experimental parameters are set as follows: the population size L is 60, the maximum iteration G is 200, the crossover probability Pc is 0.8, and the mutation probability Pm is 0.2. In the fitness function, the accuracy weight w_p is set to 0.8, and the subspace reduced dimension w_s is set to 0.2. The termination condition is the maximum number of iterations or the fitness value of the optimal individual does not increase for 50 generations. In order to find the best parameters c and g for apple spectral classification data, the best parameters c and g are 36.7583, 0.43528 and 96.1832% respectively. The fine parameter selection is carried out on the basis of rough selection, and the results are shown in Figure 6, c is 11.3137, g is 1.4142. The four kernel functions of GS-SVM, linear, polynomial, radial basis function and sigmoid, were used to construct the hyperspectral apple classification prediction model. It can be seen from Table 2 that the prediction effects of the four kernel functions are the same. At this time, accuracy is 84.3750%, precision is 87.1429%, recall is 84.7222%, F1 is 85.9155%. The prediction results of apple classification are shown in Figure 7.

GA-SVM

According to the wavelength screening step of genetic algorithm(Shen et al., 2016) in **Figure 8**, 254 spectral data contained in the original spectral region of 1000-1700nm are divided into 20 sub regions, that is, the chromosome length is 20. The parameters of genetic algorithm are set as follows: population size is 20, maximum reproduction algebra is 200, crossover probability is 0.9, mutation probability is 0.05, c and g values are 0-100. When GA is used for parameter optimization, the relationship between individual fitness and evolutionary algebra is shown in Figure 9. After 200 iterations, the individual's optimal fitness reaches the maximum and remains stable. The optimal combination of parameters c is 24.7455 and g is 0.45862. The optimization results are substituted into the SVM algorithm model and trained, and then 129 samples of the test set are used as input to get the interference pattern matching results of GA-SVM model with polynomial as the kernel function, as shown in Figure 10. Using different kernel functions, the results are shown in Table 3. It can be seen that the accuracy is 89.7266%, precision is 94.7345%, recall is 86.5278%, F1 is 90.4392%.

PSO-SVM

Particle swarm optimization (PSO) algorithm is a group computing technology based on iterative optimization(YUAN Zi-ran, 2020). Particle swarm optimization algorithm is used to update the particle fitness until the global optimal solution is found. In this paper, PSO is used to optimize the parameters of SVM(Subasi, 2013), and the process flow is shown in Figure 11. Firstly, the PSO parameters are initialized. The particle dimension is 2, the number of particles in each dimension particle swarm is 20, the maximum optimization algebra of particle swarm optimization is 200, the local search ability c1 = 1.5 and the global search ability c2 = 1.7, the inertia weight factor ω = 1, the search range of penalty parameter c is 0.1 to 100, and the search range of kernel parameter g is 0.01 to 1000. The fitness curve of PSO training process is shown in Figure 12, and the optimal output parameters are c = 43.4486, g = 0.50312. With sigmoid as the kernel function, the optimized parameters are substituted into PSO-SVM prediction model, and the prediction results of training set and test set of four kinds of apples are obtained (Figure 13), and the accuracy rate is 90.7031%. Using different kernel functions, the results are shown in Table 4. It can be seen that when sigmoid is used as kernel function, the accuracy is 91.6016%, precision is 96.1574%, recall is 88.6111%, F1 is 92.2269%.

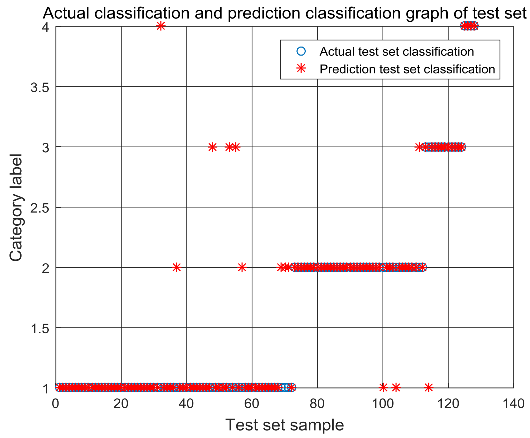


Figure 5. Classification result graph of test set

Table1. Comparison of four kernel functions of SVM

SVM kernel function	Accuracy（%）	Precision（%）	Recall（%）	F1（%）
linear	83.5938	87.3239	86.1111	86.7133
polynomial	89.8438	95.4545	87.5000	91.3043
radial basis function	78.9063	83.3333	83.3333	83.3333
sigmoid	56.2500	56.2500	100.0000	72.0000

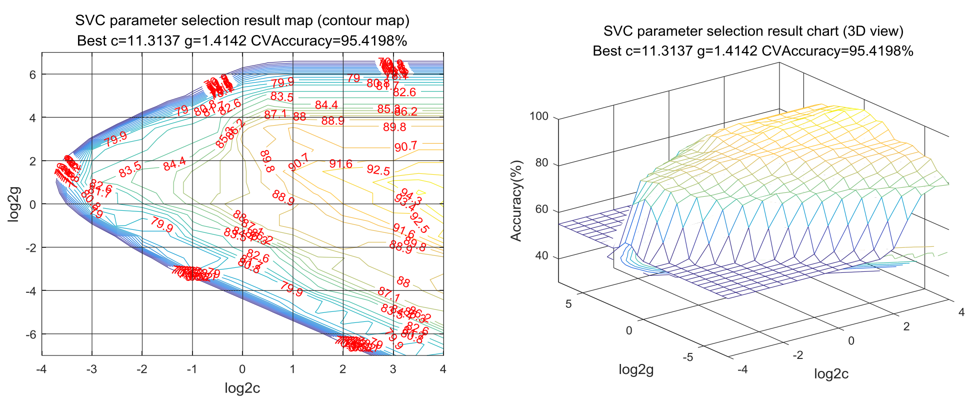


Figure 6. Result chart of fine parameter selection

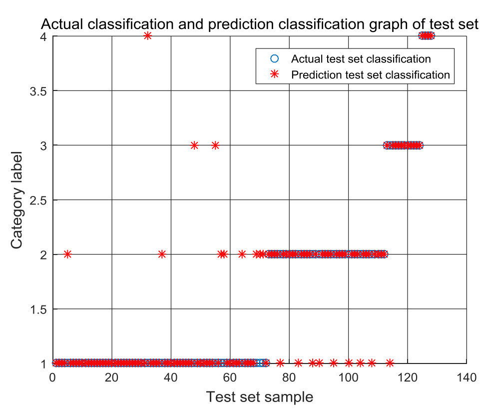


Figure 7. Classification result graph of test set

Table 2. Comparison of four kernel functions of GS-SVM

GS-SVM kernel function	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
Linear/polynomial /radial basis function /sigmoid	84.3750	87.1429	84.7222	85.9155

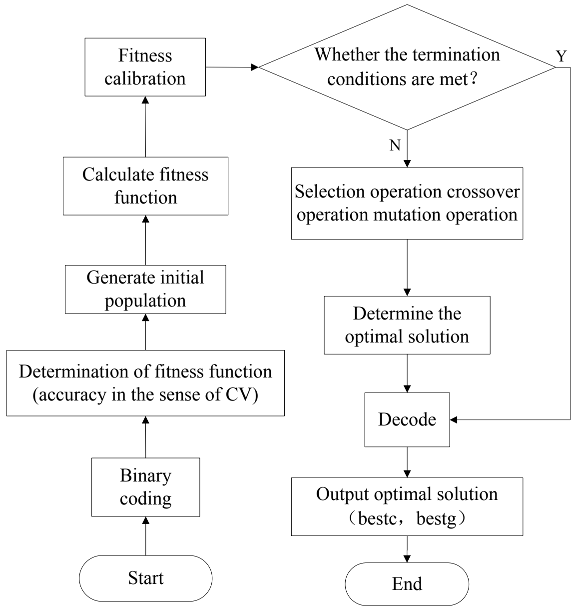


Figure 8. Algorithm flow chart of optimizing SVM parameters (c & g) by GA

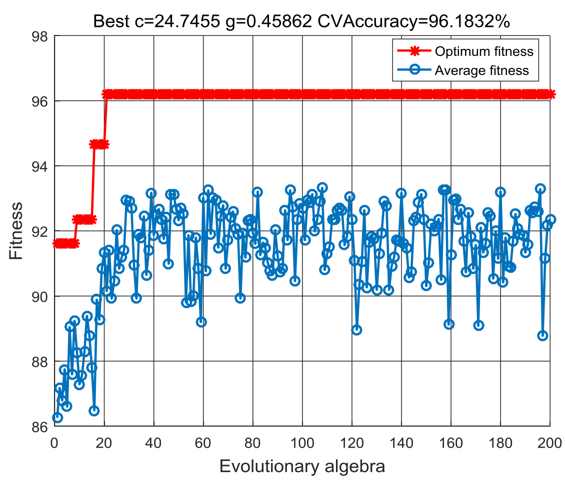


Figure 9. The Fitness curve of GA optimized parameters

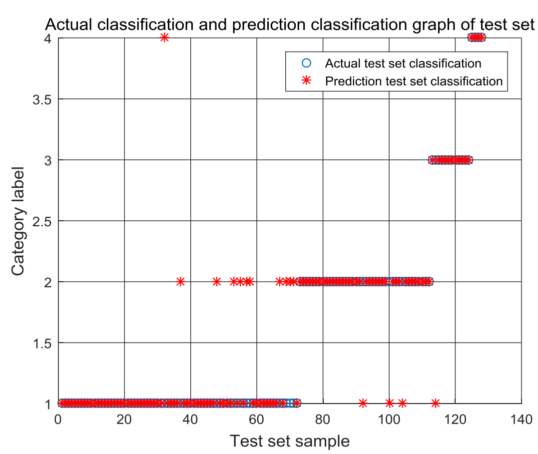


Figure 10. Classification result graph of test set

Table 3. Comparison of four kernel functions of GA-SVM

GA-SVM kernel function	Accuracy （%）	Precision （%）	Recall （%）	F1 （%）
linear	/	/	/	/
polynomial	89.7266%	94.7345%	86.5278%	90.4392%
radial basis function	84.5703%	88.4842%	83.4028%	85.8557%
sigmoid	89.6875	94.8622	86.3194	90.3854

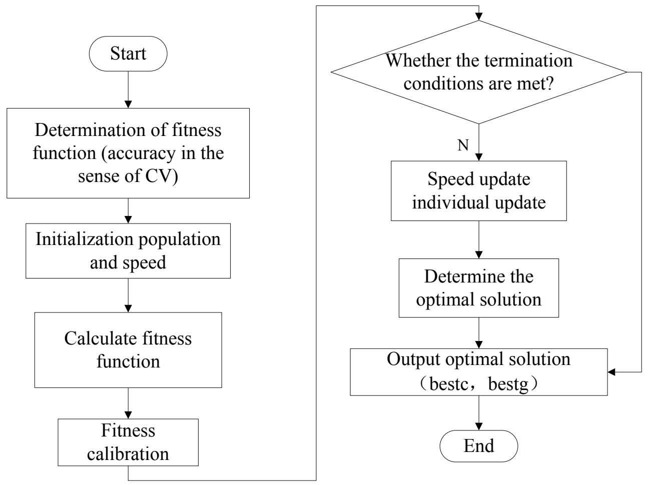


Figure 11. Algorithm flow chart of optimizing SVM parameters (c& g) by PSO

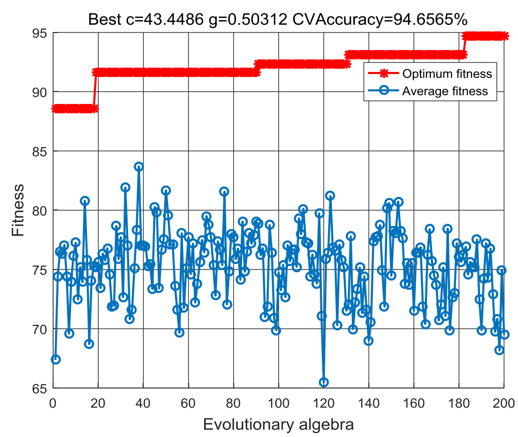


Figure 12. The fitness curve of PSO to find the best parameters

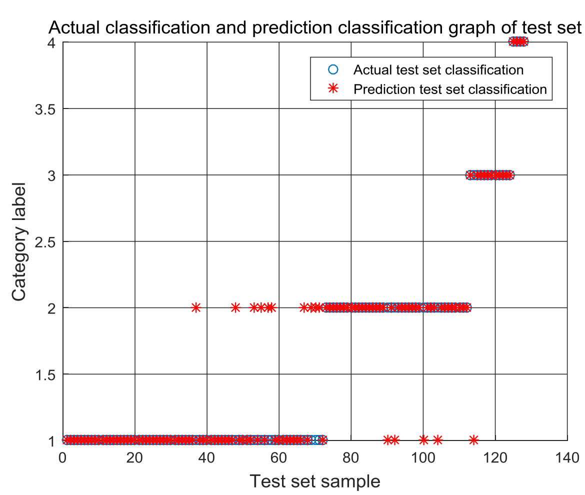


Figure 13. classification result graph of test set

Table 4. Comparison of four kernel functions of PSO-SVM

PSO-SVM kernel function	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
linear	/	/	/	/
polynomial	90.7031	95.0469	88.0556	91.4147
radial basis function	86.7578	90.0905	86.1111	88.0427
sigmoid	91.6016	96.1574	88.6111	92.2269

Conclusion

In this paper, Xinjiang Aksu apple, Henan Lingbao apple, Gansu Jingning apple and Gansu Tianshui,Huaniu apple were selected as the research objects. A total of 258 apple samples were collected. In the experiment, 1/2 samples of each type of apple were randomly selected as the modeling correction set, and the remaining samples were used as the modeling prediction set. Based on PCA, SVM, GS-SVM, GA-SVM and PSO-SVM, this paper focuses on the classification of apple using hyperspectral data. The main conclusions are as follows:

(1) The PCA algorithm was used to deal with the four kinds of apples. It was found that Xinjiang Aksu apple, Gansu Tianshui huaniu apple and Gansu Jingning apple samples overlapped and were difficult to distinguish. The distribution of Henan Lingbao Apple samples was relatively independent and the boundary was clear and easy to distinguish.

(2) SPA algorithm can effectively eliminate the irrelevant redundant information, greatly improve the correlation between spectra and apple category, and select fewer bands, the model is simple, and greatly improve the efficiency.

(3) After using GS, GA and PSO to optimize the SVM model, it is found that the PSO algorithm has the greatest improvement on the model. Compared with SVM model, F1 increases by 28.09% when kernel function is sigmoid, and 5.65% and 0.12% when kernel function is radial basis function and polynomial.

(4) From Table 1 to Table 4, the kernel functions of SVM, GS-SVM, GA-SVM and PSO-SVM are different when they have the best accuracy. When the kernel function of SVM and GA-SVM is polynomial, the value of F1 is the largest; the value of four kernel functions of GS-SVM is the same; when the kernel function of PSO-SVM is sigmoid, the value of F1 is the largest.

(5) As can be seen from Figure 5,7, 10 and 13, after optimization of SVM, GS-SVM and GA-SVM and PSO-SVM, PSO-SVM model can effectively distinguish Xinjiang Aksu apple and Gansu huaniu apple. Xinjiang Aksu apple, Gansu Jingning apple and Henan Lingbao Apple are the same Apple type, but their producing areas are different, and the result of differentiation is better. The results are the same as those of (Li Cai-hong, 2018), which use different varieties of apples and the same kind of apples from different producing areas, and have achieved good results.

This method provides technical support for the origin traceability of Red Fuji apple, and can effectively control the despicable means of adulterating and selling apples from different producing areas, and provides a new idea of detection and identification for apple products with national geographic indications.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

Chen, Q., Jiang, P. & Zhao, J. 2010. Measurement of total flavone content in snow lotus (*Saussurea involucrate*) using near infrared spectroscopy combined with interval PLS and genetic algorithm. *Spectrochimica Acta Part A Molecular & Biomolecular Spectroscopy,* 76, 50-55.

Chu Bing-Quan, Z. Y.-R., He Yong 2017. Discriminant Analysis of Bamboo Leaf Types with NIR Coupled with Characteristic Wavelengths. *Spectroscopy and Spectral Analysis,* 37, 1718-1722.

Fan, S., Li, J., Xia, Y., Tian, X., Guo, Z. & Huang, W. 2019. Long-term evaluation of soluble solids content of apples with biological variability by using near-infrared spectroscopy and calibration transfer method. *Postharvest Biology and Technology,* 151, 79-87.

Gupta, O., Das, A. J., Hellerstein, J. & Raskar, R. 2018. Machine learning approaches for large scale classification of produce. *Sci Rep,* 8, 5226.

Hao, Y., Sun, X., Gao, R., Pan, Y. & Liu, Y. 2010. Application of visible and near infrared spectroscopy to identification of navel orange varieties using SIMCA and PLS-DA methods. *Transactions of the Chinese Society of Agricultural Engineering,* 26, 373-377.

He Yong, Z. Q.-S., Zhang Chu, Cen Hai-Yan 2019. Identification of Walnut Origins and Varieties with Mid-Infrared Spectroscopy Analysis Technique. *Spectroscopy and Spectral Analysis,* 39, 2812-2817.

Ji Hai-Yan, R. Z.-Q., Rao Zhen-Hong 2019. Discriminant Analysis of Millet from Different Origins Based on Hyperspectral Imaging Technology. *Spectroscopy and Spectral Analysis,* 39, 2271-2277.

Jiarui, W. 2019. *Research on the Regional Brand Construction of Aksu Candy Heart Apple.* Wuhan Polytechnic University.

Kong, W., Zhang, C., Liu, F., Nie, P. & He, Y. 2013. Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis. *Sensors (Basel),* 13, 8916-27.

Lei, L. Y. W. X. Y. 2011. *Matlab Neural Network 43 Case Analysis,* BEIHANG UNIVERSITY PRESS, BEIHANG UNIVERSITY PRESS.

Li, C. , Li, L. , Wu, Y. , Lu, M. , Yang, Y. , & Li, L. 2018. Apple variety identification using near-infrared spectroscopy.*Journal of Spectroscopy*,2018, 1-7.

Lin Long, W. J.-Z., Liu Cui-Ling, Yu Chong-Chong, Liu Zhi, Yuan Yu-Wei 2020. Study on Hyperspectral Identification Method of Rice Origin in Northeast/Non-Northeast China Based on Conjunctive Model. *Spectroscopy and Spectral Analysis,* 40, 905-910.

Liu, W., Liu, C. & Zheng, L. 2014. Discrimination in varieties of rice seeds with multispectral imaging using support vector machine. *Transactions of the Chinese Society of Agricultural Engineering,* 30**,** 145-151.

Mengjia, D. U., Mao, B., Shen, F., Peng, L. I., Pei, F., Qiuhui, H. U. & Fang, Y. 2018. Identification of Geographical Origin of Rice Based on Fingerprint Information Fusion Model of Infrared Spectrum and Characteristic Volatile Compounds. *Food Ence*.

Pan, F. X. N. J. L. C. W. L. L. 2015. Clustering analysis of HPLC fingerprints of water extracts from different varieties of Mango Leaves. *Chinese Journal of Hospital Pharmacy,* 35**,** 710-713.

Sain & Stephan, R. 1997. The Nature of Statistical Learning Theory. *Technometrics,* 38**,** 409-409.

Shanbai, Z. J. X. W. X. J. Y. J. X. 2019. Identification of freshwater fish species based on near infrared spectroscopy and KPCA-SVM method. *Journal of Huazhong Agricultural University,* 38**,** 98-104.

Shankar, K., Lakshmanaprabu, S. K., Gupta, D., Maseleno, A. & De Albuquerque, V. H. C. 2018. Optimal feature-based multi-kernel SVM approach for thyroid disease classification. *The Journal of Supercomputing,* 76**,** 1128-1143.

Shen, L., Chen, H., Yu, Z., Kang, W., Zhang, B., Li, H., Yang, B. & Liu, D. 2016. Evolving support vector machines using fruit fly optimization for medical data classification. *Knowledge-Based Systems,* 96**,** 61-75.

Shunyu, J. 2015. The main factors and suggestions that affect the increase of farmers' grain (agricultural products) but not their income. *Jilin Financial Research***,** 59-61.

Su, X., Zhang, X., Jiao, B. & Cao, W. 2012. Determination of geographical origin of navel orange by near infrared spectroscopy. *Nongye Gongcheng Xuebao/transactions of the Chinese Society of Agricultural Engineering,* 28**,** 240-245.

Subasi, A. 2013. Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. *Computers in Biology and Medicine,* 43**,** 576-586.

Wang Wei, Z. X., Chu Xuan, Lu Yao, Jia Beibei 2019. Nondestructive Identification of Star Anise and Shikimmi by Visible/Near Infrared Hyperspectral Images. *Transactions of the Chinese Society for Agricultural Machinery,* 50**,** 373-379.

Wang, W. H., Xiaoguang; Yang, Xiaochen; Wang, Songlei; Wang, Jiayun 2015. Discrimination of origin about round jujube based on near-infrared spectroscopy. *Food Science and Technology,* 40**,** 344-347.

Wang, Z., Chen, W., Gu, S., Wang, Y. & Wang, J. 2020. Evaluation of trunk borer infestation duration using MOS E-nose combined with different feature extraction methods and GS-SVM. *Computers and Electronics in Agriculture,* 170.

Williams, P. J. & Kucheryavskiy, S. 2016. Classification of maize kernels using NIR hyperspectral imaging. *Food Chemistry,* 209**,** 131-138.

Xiaoping Wang, F. Z., Hsiang-Te Kung, Verner Carl Johnson & Aamir Latif 2020. Extracting soil salinization information with a fractional-order filtering algorithm and grid-search support vector machine (GS-SVM) model. *International Journal of Remote Sensing,* 41**,** 953-973.

Yang, S., Zhu, Q. B., Huang, M. & Qin, J. W. 2017. Hyperspectral Image-Based Variety Discrimination of Maize Seeds by Using a Multi-Model Strategy Coupled with Unsupervised Joint Skewness-Based Wavelength Selection Algorithm. *Food Analytical Methods,* 10**,** 424-433.

Yuan Zi-Ran, W. L.-F., Zhang Yang-Xi, Yu Ming, Yan Xin-Ru 2020. Hyperspectral Inversion and Analysis of Heavy Metal Arsenic Content in Farmland Soil Based on Optimizing CARS Combined with PSO-SVM Algorithm. *Spectroscopy and Spectral Analysis,* 40**,** 567-573.

Zhang, X., Liu, F., He, Y. & Li, X. 2012. Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. *Sensors (Basel),* 12**,** 17234-46.

Zhao Yi-Kun, Y. Y.-B., Shen Bing-Hui, Yang Yong-Qin, Ai Jun-Min, Yan Yan-Lu, Kang Ding-Ming 2020. Influence Factors in Near-Infrared Spectrum Analysis for the Authenticity Identification of Maize Single-Kernel Varieties. *Spectroscopy and Spectral Analysis,* 40**,** 2229-2234.