A Study on Classification of Fruit Type and Fruit Disease

Meng-Meng YANG[1,*], Rikuto Kichida[2],

[1] School of Computer Science and Technology, Nanjing Normal University, Jiangsu 210023, China

[2] Department of Energy Engineering and Science, Nagoya University, Nagoya, Aichi 464-8601, Japan

Email: 1320981274@qq.com

Keywords: fruit classification; fruit disease classification

Abstract: with the rise of biotechnology, the similarity between fruits is growing, which has some impact on our lives. Hence, it needs to design a system to classify the types of fruits. If the fruit is placed over a certain period of time, it will be infected. Meanwhile, it is also important to classify the disease of fruits. In this paper, we introduce several approaches to classify the type of fruits and fruit diseases automatically. Finally, we make an overall comparison based on these methods.

1. Introduction

Fruit nutrition is very rich and there are many effects [1]. For example, the fruit of the fiber for the pectin material beneficial defecation, and fiber ingredients can also promote the body's metabolic function. On the other hand, the fruit also contains natural pigments [2], which can effectively prevent cancer. At present, the post-processing of our fruit is by manual sorting, which inevitably arise from various problems, such as the workers' long-term monotonous repetition of work to make them fatigue, resulting in reduced accuracy of the test, the exception of workers on the classification of the standard accuracy of the test. Therefore, the automation of fruit classification [3] is an urgent need to improve worker productivity, improve classification accuracy.

If the fruit is placed at room temperature for more than the specified date, the color of the fruit will change, and then the fruit will be infected. It is harmful to human body if they eat this infected fruit. Therefore, it is important to design a system for automatic detection of fruit diseases [4].

2. Methods and Results

Wu [5] adopted a four-step to classify the type of fruits. Firstly, this paper used split-and-merge method to remove the background of the image. Secondly, extracting 79 features from a fruit image. Thirdly, in this paper, the authors used PCA to reduce the dimension to 14. Finally, three kinds of multi-class SVMs and kernels were chosen to classify. The results showed that the "Max-Wins-Voting SVM + Gaussian Radial Basis kernel" performed best.

Wang [6] employed a novel feature extraction method-wavelet entropy (WE). In this paper, 3-level decomposition was chosen to extract 10 features from each color channel for each fruit image, and a total of 30 features were extracted. The time spent on the training of 30 features is still relatively large, so it is necessary to reduce the dimension. After that, authors chose FSCABC-FNN and BBO-FNN as classifier. The accuracy of those two methods is the same, but the latter calculation time less than the former 5 seconds on average.

Ji [7] was based on the improvement of the paper [5], and the biggest difference between the two papers is the choice of the classifier. In this paper, authors proposed "FSCABC + FNN", afterwards, compared with four methods, it is found that the method has the highest accuracy, 89.1%.

Wu [8] proposed "PCA+BBO-FNN" method, and it is based on the paper [6] in order to improve the classification performance. By comparison, it found that the two methods were same in classification accuracy, but the average compute time of BBO-FNN was significantly less than FSCABC-FNN.

Wang [9] demonstrated that HPA's performance was better than other hybridizations, and SLFN has achieved good results in terms of classification. Therefore, Lu [10] proposed a novel "HPA + SLFN" method. At the same time, the method was compared with other five existing methods. The experimental results showed that our method has the highest classification accuracy for 20 kinds of fruits.

Table 1. Comparison of several classification algorithms

Author	Method used	Limitation	Accuracy
Wu [5]	PCA, multi-class SVMs +kernels, 5-fold stratified cross validation	For two similar fruit, the effect of classification is not good.	88.2%
Wang [6]	This paper proposed that four-step preprocessing can play an important role in the classification.	The camera's posture and position are different, resulting in unsatisfactory accuracy.	89.5%
Ji [7]	PCA + FSCABC + FNN	The classification result of the Hass Avocados was 76.2%.	89.1%
Wu [8]	PCA + BBO + FNN	The accuracy still need to be improved.	89.11%
Lu [10]	PCA+HPA-FNN	Passion fruit and green grapes were not recognized well.	89.5%
Kuang [11]	WSL-LW	Several areas contain very little information can be used to identify-purposes.	90.7%
Garcia [12]	HSV, Fourier descriptor, Hu moments and basic geometric characteristics + ACP + Bayesian	The appearance of similar fruit was prone to the wrong classification.	95.59%
Zawbaa [13]	RF+SIFT	The dataset and the number of fruit types were relatively small.	96.97% for apple
Rocha [14]	A combine many classifiers and features method	It is not advisable to combine weak and high classification errors and features with low classification errors.	Classification error was reduced by up to 15%.
Samajpati [15]	"LTP + Gabor + CLBP" with RF	only contains 80 images.	85%
Capizzi [16]	RBPNN+GLCM	N/A	97.25%

Kuang [11] used the proposed weighted score-level fusion to classify the type of fruits according to different single feature and complemental features. The result showed that "color + shape + edgeLBP + HOG + LBP" achieved the higher accuracy. Afterwards, this paper compared with other four state-of-the-art methods, and got a conclusion that WSL-LW (our method) achieved 90.7% which was higher than other methods.

Garcia [12] proposed a novel feature extraction method Firstly, the RGB space in the image of the fruit was captured, the image was mapped to the HSV space. The saturated background color of the fruit was divided from the background information. Secondly, the shape features were extracted by calculating the Fourier descriptor, Hu moments and basic geometric characteristics. Thirdly, the texture features were extracted by co-occurrence array of gray levels. Afterwards, ACP was applied to reduce size and then the species of the fruit was classified using the Bayesian classifier.

Zawbaa [13] proposed random forests (RF) as classifier. Firstly, this paper used color, shape features and SIFT methods to extract a number of features from a 90x90 pixel fruit image. After that, these features were used as inputs to RF and other classifiers such as SVM and K-NN. The result showed that RF achieved better classification accuracy on the whole.

Rocha [14] presented a method that can combine many features and classifiers. The dataset was the result of a local vegetable and fruit distribution center collected five months under different conditions in order to increase the accuracy of classification. Then extracted the structural, color and texture features from a fruit image. Meanwhile, this paper used four methods to subtract background. After that, this paper analyzed the extracted features and six different classifier.

3. Discussions

All of the above papers are classified on healthy fruits, but when the fruit is infected, the color of peel will change, which will have a significant impact on the fruit classification results. Samajpati [15] used random forest to detect the diseases of fruit. Firstly, extracting features from shape, color and texture. Secondly, the k-means clustering technique was used to segment the diseased part if the fruit was infected. Thirdly, RF classifier as used to detect the apple diseases.

Capizzi [16] converted the RGB image to a grayscale image and used the threshold to separate the image from its background. According to the threshold, the grayscale image was converted into a binary image, and the shape feature was extracted with the help of the binary image. Afterwards, in this paper, gray level co-occurrence matrix was used to extract the texture features. Finally, this paper was selected RBPNN as classifier to classify the five diseases of orange.

By comparison, we find that the methods are becoming more and more advanced than before, and the highest accuracy is 97.25%, which succeeds human eyes. However, these methods have shortcomings as listed in Table 1. We expect the readers can cooperate and solve these problems.

A problem is the unbalanced data in their private datasets. Some authors obtain much more images of one type of fruit than the other fruit type. This problem can be solved by data balance techniques, such as synthetic minority oversampling technique [17].

Another common problem is the classifier training. The training methods of several paper indeed trap into local minimal basins. Hence, we suggest to use bioinspired algorithms, such as particle swarm optimization (PSO) [18-20], biogeography-based optimization [21-23], artificial bee colony [24, 25], Jaya algorithm [26, 27], etc.

The classifiers can be replaced by advanced classifiers, such as extreme learning machine [28], fuzzy support vector machine [29-31], linear regression classifier [32, 33], twin support vector machine [34, 35], sparse autoencoder [36], logistic regression [37], etc.

4. Conclusion

In this paper, we conducted a comprehensive review and survey of the various classification methods. The future work is to select the best method and apply it into realistic world.

Acknowledgement

This paper is supported by Program of Natural Science Research of Jiangsu Higher Education Institutions (16KJB250025).

Reference

[1]. Safiri, S., M. Sani, et al., Fruit and Vegetable, Fat, and Sugar-Sweetened Beverage Intake Among Low-Income Mothers Living in Neighborhoods With Supplemental Nutrition Assistance Program-Education: Methodological Issues. Journal Of Nutrition Education And Behavior, 2017. 49(3): pp. 272-272

[2]. Estevez-Santiago, R., B. Olmedilla-Alonso, et al., Lutein and zeaxanthin supplied by red/orange foods and fruits are more closely associated with macular pigment optical density than those from green vegetables in Spanish subjects. Nutrition Research, 2016. 36(11): pp. 1210-1221

[3]. Seema, A. Kumar, et al. Automatic Fruit Grading and Classification System Using Computer Vision: A Review. in 2nd IEEE International Conference on Advances in Computing and Communication Engineering (ICACCE). 2015. Dehradun, INDIA: IEEE. pp. 598-603

[4]. Lin, Y.F., M.Y. Chen, et al., DNP and ATP induced alteration in disease development of Phomopsis longanae Chi-inoculated longan fruit by acting on energy status and reactive oxygen species production-scavenging system. Food Chemistry, 2017. 228: pp. 497-505

[5]. Wu, L., Classification of fruits using computer vision and a multiclass support vector machine. Sensors, 2012. 12(9): pp. 12489-12505

[6]. Wang, S., Y. Zhang, et al., Fruit Classification by Wavelet-Entropy and Feedforward Neural Network Trained by Fitness-Scaled Chaotic ABC and Biogeography-Based Optimization. Entropy, 2015. 17(8): pp. 5711-5728

[7]. Ji, G., Fruit classification using computer vision and feedforward neural network. Journal of Food Engineering, 2014. 143: pp. 167-177

[8]. Wu, J., Fruit classification by biogeography-based optimization and feedforward neural network. Expert Systems, 2016. 33(3): pp. 239-253

[9]. Wang, S., Y. Zhang, et al., Feed-forward neural network optimized by hybridization of PSO and ABC for abnormal brain detection. International Journal of Imaging Systems and Technology, 2015. 25(2): pp. 153-164

[10].Lu, S.Y., Z.H. Lu, et al. Fruit Classification by HPA-SLFN. in 8th International Conference on Wireless Communications & Signal Processing. 2016. Yangzhou, China: IEEE. pp. 11-17

[11].Kuang, H.L., L.L.H. Chan, et al., Fruit classification based on weighted score-level feature fusion. Journal of Electronic Imaging, 2016. 25(1), Article ID: 013009

[12].Garcia, F., J. Cervantes, et al., Fruit Classification by Extracting Color Chromaticity, Shape and Texture Features: Towards an Application for Supermarkets. IEEE Latin America Transactions, 2016. 14(7): pp. 3434-3443

[13].Zawbaa, H.M., M. Hazman, et al. Automatic fruit classification using random forest algorithm. in 14th International Conference on Hybrid Intelligent Systems (HIS). 2014. Kuwait, KUWAIT: IEEE. pp. 164-168

[14].Rocha, A., D.C. Hauagge, et al., Automatic fruit and vegetable classification from images. Computers and Electronics in Agriculture, 2010. 70(1): pp. 96-104

[15].Samajpati, B.J. and S.D. Degadwala. Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier. in International Conference on Communication and Signal Processing (ICCSP). 2016. Melmaruvathur, INDIA: IEEE. pp. 1015-1019

[16].Capizzi, G., G. Lo Sciuto, et al. Automatic Classification of Fruit Defects based on Co-Occurrence Matrix and Neural Networks. in Federated Conference on Computer Science and Information Systems. 2015. New York: IEEE. pp. 861-867

[17].ZHANG, Y.-D., Y. ZHANG, et al., Synthetic Minority Oversampling Technique and Fractal Dimension for Identifying Multiple Sclerosis. Fractals. 25(4), Article ID: 1740010

[18].Wu, L., Crop Classification by forward neural network with adaptive chaotic particle swarm optimization. Sensors, 2011. 11(5): pp. 4721-4743

[19].Yang, J.F. and P. Sun, Magnetic resonance brain classification by a novel binary particle swarm optimization with mutation and time-varying acceleration coefficients. Biomedical Engineering-Biomedizinische Technik, 2016. 61(4): pp. 431-441

[20].Ji, G., A comprehensive survey on particle swarm optimization algorithm and its applications. Mathematical Problems in Engineering, 2015. 2015, Article ID: 931256

[21].Chen, P. and S. Du, Pathological Brain Detection via Wavelet Packet Tsallis Entropy and Real-Coded Biogeography-based Optimization. Fundamenta Informaticae, 2017. 151(1-4): pp. 275-291

[22].Wang, S.-H. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimedia Tools and

Applications, 2016, DOI: 10.1007/s11042-016-4222-4 (Online).

[23].Wu, X., Smart detection on abnormal breasts in digital mammography based on contrast-limited adaptive histogram equalization and chaotic adaptive real-coded biogeography-based optimization. Simulation, 2016. 92(9): pp. 873-885

[24].Wu, L., Optimal multi-level Thresholding based on Maximum Tsallis Entropy via an Artificial Bee Colony Approach. Entropy, 2011. 13(4): pp. 841-859

[25].Wu, L., Magnetic Resonance Brain Image Classification by an Improved Artificial Bee Colony Algorithm. Progress in Electromagnetics Research, 2011. 116: pp. 65-79

[26].Rao, R.V., A. Liu, et al., Abnormal Breast Detection in Mammogram Images by Feed-forward Neural Network trained by Jaya Algorithm. Fundamenta Informaticae, 2017. 151(1-4): pp. 191-211

[27].Cattani, C. and R. Rao, Tea Category Identification Using a Novel Fractional Fourier Entropy and Jaya Algorithm. Entropy, 2016. 18(3), Article ID: 77

[28].Lu, S. and X. Qiu, A Pathological Brain Detection System based on Extreme Learning Machine Optimized by Bat Algorithm. CNS & Neurological Disorders - Drug Targets, 2017. 16(1): pp. 23-29

[29].Yang, J., Identification of green, Oolong and black teas in China via wavelet packet entropy and fuzzy support vector machine. Entropy, 2015. 17(10): pp. 6663-6682

[30].Li, Y., Detection of Dendritic Spines using Wavelet Packet Entropy and Fuzzy Support Vector Machine. CNS & Neurological Disorders - Drug Targets, 2017. 16(2): pp. 116-121

[31].Lu, H.M., Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation. IEEE Access, 2016. 4: pp. 8375-8385

[32].Du, S., Alzheimer's Disease Detection by Pseudo Zernike Moment and Linear Regression Classification. CNS & Neurological Disorders - Drug Targets, 2017. 16(1): pp. 11-15

[33].Chen, Y., Y. Zhang, et al. Wavelet energy entropy and linear regression classifier for detecting abnormal breasts. Multimedia Tools and Applications, 2016, DOI: 10.1007/s11042-016-4161-0 (Online).

[34].Yang, M., Dual-Tree Complex Wavelet Transform and Twin Support Vector Machine for Pathological Brain Detection. Applied Sciences, 2016. 6(6), Article ID: 169

[35].Chen, M., Morphological analysis of dendrites and spines by hybridization of ridge detection with twin support vector machine. PeerJ, 2016. 4, Article ID: e2207

[36].Chen, H. Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimedia Tools and Applications, 2017, DOI: 10.1007/s11042-017-4554-8 (Online).

[37].Zhan, T.M. and Y. Chen, Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression. IEEE Access, 2016. 4: pp. 7567-7576