Apple Defect Detection Method Based on Convolutional Neural Network

Zheng Xu, Tao Shen, Shuhui Bi, and Qinjun Zhao(✉)

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

Abstract. The appearance quality of apple is one of the important indicators for consumers to purchase. At present, the classification process of apple is still completed artificially, which not only wastes human resources, but also easily causes subjective misclassification. This paper proposes a convolutional neural network model to classify defective and defect-free apples. Apple images are collected by the smartphone camera, each type of apple has 312 images. The number of apple images is expanded through data enhancement technology, and randomly divided into training set, validation set, and test set according to the ratio of 6:2:2. The final classification accuracy is 99.2%.

Keywords: Deep learning · Convolutional neural network · Classification

1 Introduction

Apple is popular with consumers for its rich nutritive value and luscious taste. It is one of the most common fruits of people's daily life. China is the largest apple-producing country whose apple planting area and output account for more than 50% of the world. From 2003 to 2018, the apple output has maintained a steady growth trend, and reached a peak of 41.39 million tons in 2017 [1]. In addition, China is also the country with the largest export volume of apples in the world. From 1992 to 2017, China's total apple exports increased from 20 million dollars to 1.453 billion dollars, and the total exports increased from 38,300 tons to 1.3284 million tons (Un Comtrade Database, 1992- 2017), the proportion of global apple exports trade rose from 1.40% to 19.21% and 1.91% to 14.40% respectively [2]. However, China is not a powerful apple exporter, and China's apple exports account for 1.83%∼3.15% of the total volume merely. While the world's apple exports account for 8.30% of the total output approximately [3].

Supported by Key R&D projects of Shandong Province under grant 2019GNC106093, Shandong Agricultural machinery equipment R&D innovation plan project under grant 2018YF011, Key R&D projects of Shandong Province under grant 2019JZZY021005, Shandong Provincial Key R&D project 2017GGX10116.

W. Fu et al. (Eds.): ICMTEL 2021, LNICST 387, pp. 398–404, 2021.
https://doi.org/10.1007/978-3-030-82562-1_37

The inspection and classification of fruit quality are essential to improve the competitiveness of fruit products. Generally speaking, postpartum treatment of apples is mostly done by manual. So it may lead to some problems which can not meet the production needs, such as low classification accuracy and slow speed because of human subjectivity and visual fatigue. There needs to be a system that can identify whether it is worth based on the quality of two categories of fruit.

In recent years, a lot of researches on the detection and classification of apple have been done by Domestic and foreign scholars. In the work of [4], they extracted characteristic parameters like energy, entropy, and moment of inverse difference, then sent them to the neural network for classifying, the accuracy is 95.5%. The research was conducted by [5] combining with brightness correction technology, segmented the defect candidate areas (defects, fruit stalks, calyx) of apples, then randomly extracted the color, texture, and some other features, obtained 95.7% accuracy using AdaBoost. Subsequent research on the identification of apple defects [6] used Support Vector Machine (SVM), MLP, and K-Nearest Neighbor (KNN) classifier respectively, and SVM is the most accurate method with an accuracy of 92.5%. Most of the above studies include image acquisition, image processing, image features selection, and extraction. It does not only requires a lot of manpower, but the selection and extraction of features are also uncertain and complex. In response to this problem, the convolutional neural network uses the original image as the input, which can effectively learn the features from a large number of images. Thus the convolution neural network can avoid the complex features extraction process. In addition, the characteristic of weight sharing can also greatly decrease the complexity of the network and raise the efficiency of training. A recent study carried out by [7] used a convolutional neural network (CNN) on the dataset of apple images. It took the normal, calyx, fruit stem, and boundary image blocks as positive samples on the one hand, and on the other hand, rotting, scar, insect injury and other defects were made to negative samples, the accuracy is 97.3%. Compared with the above apple defects detection methods, it has a certain improvement in recognition efficiency and accuracy. In this paper, the dataset is expanded by translation, scaling, and rotation first, then sent to a 15-layer convolutional neural network model, and the results show that it can be implemented more accurately and quickly.

2 Image Pre-processing

Red Fuji apple was studied in this paper. The apple images are collected on the white background with the camera of the smartphone at 13MP resolution under the room lighting environment, as shown in Fig. 1. Image preprocessing consists mainly of the following two aspects. Firstly, preserving the part of the apple image that we are interested in through background removal. The second step is to expand the number of sample images through data enhancement to ensure that the network model is fully trained.

(a) Defect-free (b) Defective

Fig. 1. The original image captured by the camera

2.1 Background Removal

Choosing an appropriate color space model is critical in background segmentation, which is conducive to fast and accurate segmentation of the target area. Since the HSI color space is approximate to the way of the human eye perceives color, and minimally influenced by light intensity. So this paper converts RGB color space to HSI color space, and the conversion formula is as follows:

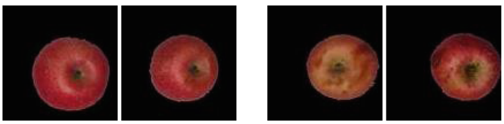
$$\begin{cases} I = \frac{1}{3}(R+G+B) \\ S = 1 - \frac{3}{R+G+B}[min(R,G,B)] \\ H = \begin{cases} \theta, & G \geq B \\ 2\pi - \theta, & G < B \end{cases} \end{cases}$$

where,

$$\theta = \arccos\left[\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(G-B)(R-B)}}\right].$$

It is found that the histogram distribution of the S component is double peaks obviously, so we can calculate the threshold value of the S component to segment the image. The specific steps are as follows:

(1) Median filtering of S component;
(2) Otsu algorithm [8] was used for threshold segmentation of S-component images;
(3) The binary image is obtained by the morphological operation;
(4) Multiply the binary image with the original image to remove the background;
(5) Set the image resolution to 100×100, as shown in Fig. 2.



(a) Defect-free (b) Defective

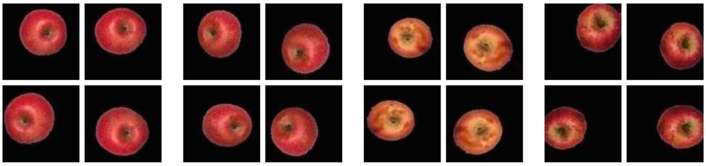
Fig. 2. Apple image with background removed

2.2 Data Enhancement

The deep learning model is complex and requires a sufficient number of samples for training. It's difficult to obtain enough samples due to the constraints of various factors, so the offline expansion of the sample images is carried out in the study. It mainly contains the following methods:

(1) Rotation: Rotate the image randomly to a certain angle;
(2) Shift: Shift the image to a certain range randomly in the horizontal or vertical direction;
(3) Flip: Reverse the image in the horizontal or vertical direction;
(4) Scale: Enlarge or reduce the image according to the specified scale factor;

The images in Fig. 2 are randomly enhanced by the above method, and the expanded images are shown in Fig. 3.



(a) (b) (c) (d)

Fig. 3. Data augmentation of the images in Fig. 2

3 The Convolution Neural Network

The convolutional neural network is a deep learning model or multilayer perceptron similar to the artificial neural network, which is often used to analyze visual images. It mainly consists of the convolution layer, pooling layer, and full connection layer. Also, it includes some auxiliary training modules such as activation function and classifier. Among them, the convolutional layer is the key part of the convolutional neural network. Using local connections and weight sharing to extract features of images can greatly decrease the number of parameters in the neural network. The activation function is a nonlinear mapping to the output of the convolutional layer. The pooling layer is used for decreasing the data dimension, simplifying the complexity of network calculation, and extracting the main features of the image. After then, the full connection layer connects all the features and sends the final value to the classifier.

The convolution neural network model used in this study is shown in Table 1. The input image is an RGB image with a resolution of 100×100, and the features of images are extracted by three convolution layers. The size of kernel/filter used for each convolution layer is 5×5, and 32 filters are used in the first two convolution layers, the third convolution layer uses 64 filters to extract deeper features. The relu activation function is used to nonlinear map the output of convolution layer, and the size of the pooling is 3 × 3. Finally, the extracted features are integrated through the fully connected layers, and then, are sent to the softmax classifier for classification.

Table 1. Summary model with 5 × 5 kernel size

Layers	Type	Activations	Learnables
1	Image Input (RGB)	100 × 100 × 3	–
2	32 Conv.Filter (5 × 5 × 3)	100 × 100 × 32	Weight 5 × 5 × 3 × 32 Bias 1 × 1 × 32
3	Relu	100 × 100 × 32	–
4	Max Pooling (3 × 3)	49 × 49 × 32	–
5	32 Conv.Filter (5 × 5 × 32)	49 × 49 × 32	Weight 5 × 5× 3 × 32 Bias 1 × 1 × 32
6	Relu	49 × 49 × 32	–
7	Max Pooling (3 × 3)	24 × 24 × 32	–
8	64 Conv.Filter (5 × 5 × 32)	24 × 24 × 64	Weight 5 × 5 × 3 × 32 × 64 Bias 1 × 1 × 64
9	Relu	24 × 24 × 64	–
10	Max Pooling (3 × 3)	11 × 11 × 64	–
11	Fully Connected (400)	1 × 1 × 400	Weight 400 × 7744 Bias 400 × 1
12	Relu	1 × 1 × 400	–
13	Fully Connected (2)	1 × 1 × 2	Weight 2 × 400 Bias 2 × 1
14	Softmax	1 × 1 × 2	–
15	Classification Output	–	–

4 Results

In this paper, we divided the expanded dataset into the training set, validation set and test set at the ratio of 6:2:2, which is used to train and test the neural network model respectively, we can see in Table 2.

After many experiments, the parameters were adjusted as follows: MaxEpoch is 40, BatchSize is 128, LearnRate is constant 0.001, the momentum stochastic gradient descent method was used for training. The loss and accuracy of training and validation sets during the experiment are shown in Fig. 4. When the accuracy of the validation set does not change for five consecutive times, we stop training, and the final iteration is 812 times. The losses and accuracies of this model are shown in Table 3, the losses of training and validation sets are 0.0122 and 0.0462 respectively.

Table 2. Dataset

Category	Training set	Validation set	Test set
Defect	1872	624	624
Defect-free	1872	624	624
Total	3744	1248	1248

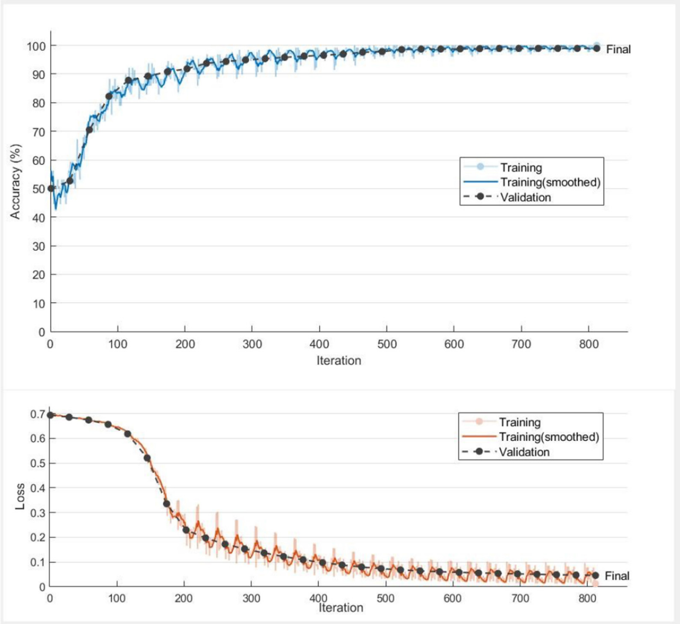


Fig. 4. Loss and accuracy plot from data training and validation

Finally, we test the model and get an accuracy of 99.2%. There are 10 apples were misclassified, including 9 defective apples and 1 defect-free apple, as shown in Table 4.

Table 3. The accuracy and loss of training and validation

	Training	Validation
Amount of data	3744	1248
Loss	0.0122	0.0462
Accuracy	100%	98.96%

Table 4. The result of testing

	Defective	Defect-free
Actual label	624	624
Predict labels	615	623

5 Conclusion

This research studies an apple classification method using the convolutional neural network. The classification process is completed through background removal and image expansion combined with CNN. In the training process, three convolution layers and pooling layers are used to extract the features of the sample images, and the appropriate parameters are adjusted according to the performance of the validation set. We terminate the training when the accuracy of the validation set no longer changes 5 consecutive times, and save the network model at this time. Finally, the reliability of the model is tested and very good results are obtained.

References

1. Yang, J., Meng, X. N., Xin, L.: Analytical report on China's apple market from 2017 to 2019. Deciduous Fruits **51**(5), 5–7 (2019)
2. Su, S.S., Huo, X.X.: Analysis on the changes of structural characteristics of global apple trade network and the transition of China's statue. Issues Agric. Econ. **6**, 99–109 (2020)
3. Zhang, B.: Analysis of China's apple industry's output, processing and trade status in the past 7 years. China Fruit Tree (4), 106–108 (2018)
4. Yu, M., Li, X., Yang, H.C.: Research on apple grading algorithm based on image recognition. Autom. Instrum. **7**, 39–43 (2019)
5. Zhang, B.H.: On-line identification of defect on apples using lightness correction and AdaBoost methods. Trans. Chinese Soc. Agric. Mach. **45**(6), 221–226 (2014)
6. Moallem, P., Serajoddin, A., Pourghassem, H.: Computer vision-based apple grading for golden delicious apples based on surface features. Inf. Process. Agric. **4**(1), 33–40 (2017)
7. Liu, Y., Yang, J.B., Wang, C.X.: Apple defect detection algorithm based on Convolution Neural Network. Electron. Meas. Technol. **40**(3), 108–112 (2017)
8. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. man Cybern. **9**(1), 62–66 (1979)