Apple Spoilage Classification

Haoran Zheng
College of Engineering
Boston University
Boston, MA 02215
haoranz9@bu.edu

Abstract

As a popular fruit, apples are rich in nutrients such as vitamins, proteins, and more. Apples are very nutritious and delicious when they are ripe and fresh, but they can also spoil and decay. In the food industry, it is necessary to use an automated method to classify food freshness. Deep convolutional neural networks are considered to be the most effective computer vision technique for classification tasks. This project uses a deep convolutional neural network to detect whether an apple is rotten or not. The CNN model for this project was designed with VGG16 as the architecture, and then a classifier block was constructed. The model was trained and tested on two types of image datasets, healthy and rotten apples. After testing, this CNN model can achieve 98% overall classification accuracy.

1 Introduction

Humans, as a carbon-based organism, need to constantly take in nutrients to obtain energy to sustain life. With the improvement of people's living standard, fruit, a food, is gradually becoming a favorite. Fruits are rich in vitamins, proteins and other nutrients. Apples, as a representative of this, are one of the most grown and consumed fruits in the world, and are also very popular. Apples are rich in vitamin A, vitamin B1, vitamin B2, vitamin B6, vitamin C and folic acid and other nutrients [1]. Although apples are easier to preserve compared to fruits such as bananas and plums, however, the presence of microorganisms such as bacteria causes them to spoil and become moldy [2]. For the food industry, spoiled and rotten apples can cause huge economic losses [3]. If the apples are not identified well, the food industry will also face subsequent troubles such as complaints from retailers and consumers after these spoiled apples are circulated to the market. For consumers, apple spoilage not only leads to a reduction in the nutritional content of apples, but also leads to health damage due to misuse of spoiled apples. According to the CDC, food-borne illnesses have been a cause for concern in recent years, with more than 3,000 Americans losing their lives each year as a result [4]. Therefore, the identification of rotten apples is very necessary for research on either side. Food spoilage detection methods based on digital image processing and machine learning as well as deep learning have been widely proposed [5, 6]. In this project, we aim to propose a CNN-based classifier for rotten apples.

2 Related methods for apple spoilage detection

2.1 Deep learning

Deep learning is an important research object in the field of machine learning, one of the latest trends in machine learning and artificial intelligence research [7]. Compared with the traditional shallow learning, deep learning has the characteristics of emphasizing the depth of model structure and the importance of explicit feature learning. Deep learning is widely used in various fields. In addition to

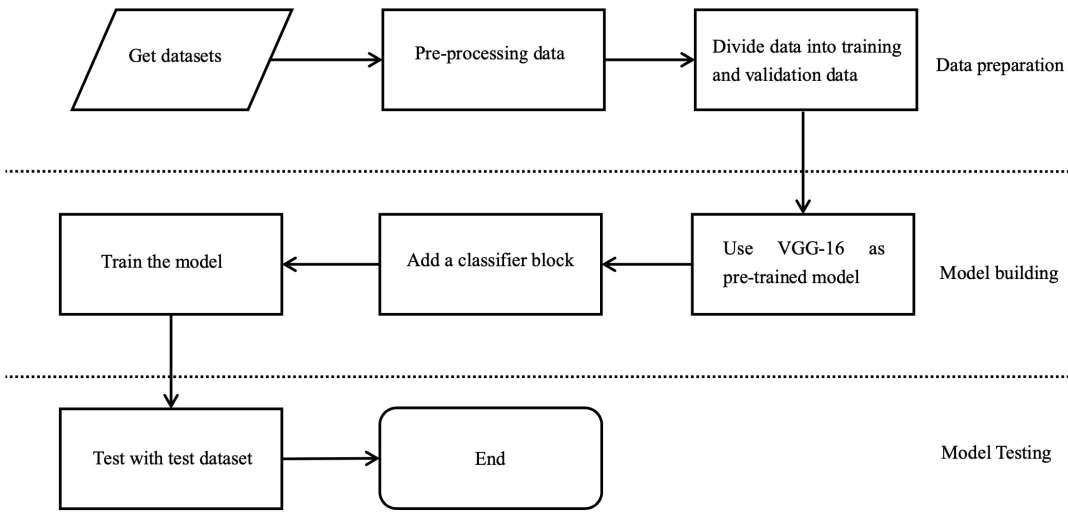


Figure 1: System flow Chart.

its application in food spoilage, deep learning has also been used to study the ripening problem of palm fruit [8].

2.2 Convolutional neural network

Convolutional neural network is one of the representative algorithms of deep learning, which is a class of neural networks containing convolutional computation and having depth structure. Convolutional neural networks can automatically learn features from large-scale data and generalize the results to unknown data of the same type. Therefore, convolutional neural networks are often used to classify images or identify pattern similarities between them. CNNs are composed of neurons whose weights and biases can be learned. To train the input image, CNNs can use convolutional layers as one of the main layers. These layers are tasked with extracting low-level features such as colors, spots and edges through filters. Other common layers are pooling layers, which serve to reduce the size of the image. The fully connected layer is also one of the important components. Training the CNN is divided into two steps: feedforward and backpropagation. In the feedforward step, the error of the network is calculated by comparing the difference between the output of the network based on a given image as input and the output of the actual label. In the backpropagation stage, the gradients of the parameters are calculated based on the network error, and then each weight matrix is updated based on the calculated gradients [9]. The essence of convolutional neural network includes shallow convolutional layer to extract the basic features, such as edges and contours. The deep convolutional layer extracts abstract features, and the fully connected layer performs scoring and classification based on feature combinations. In addition to its application in food spoilage, convolutional neural networks are also used to monitor kiwifruit under different timings as well as light conditions to obtain accurate harvest timing [10].

3 Flow Chart

The system flowchart of this project plan is drawn in Figure 1.

4 Dataset and data pre-processing

4.1 Dataset

The data used in this project is the Fruits Fresh and Rotten dataset used for classification, which is from Kaggle [11], by collecting, separating, and then labeling. For this project, only two parts of the data, freshapples and rottenapples, were used because the main object of study was apple. Thus, the dataset consisted of 4,035 images of fresh and rotten apples, including 1,693 images of fresh apples and 2,342 images of rotten apples. In this project, the images were separated into two categories by classes. In addition, some of the images in the dataset were rotated by certain angles, some were rotated horizontally/vertically, and some were added with manic points for the purpose of data augmentation. Data augmentation, in essence, is an image processing method that generates more data based on limited data, thereby increasing the number as well as the diversity of training samples, which in turn improves the generalization ability of the model, reduces the possibility of overfitting and increases robustness. Data augmentation is also used when processing the training data.

4.2 Data pre-processing

In the data preprocessing process, the same data augmentation methods are used, such as rotation, flip, and zoom. In the next step, the dataset is divided. The training data set is divided into two parts: training data and validation data. There are 2826 images in two categories in the training dataset and 1209 images in two categories in the validation dataset. Then the images are rized into (256, 256), while the data are divided into batches of batch size 32.

5 CNN model architecture

5.1 Fine-tuning pretrained network

A pre-trained model is a model that has already been trained with a dataset, commonly used pre-trained models are VGG16/19 and ResNet50, etc. When we encounter a deep learning task, our first intuition may be to train the network from scratch. But in reality, convolutional neural networks can have a large number of parameters, possibly in the range of millions. Training CNNs on small datasets, especially on small datasets with less than the number of parameters, can greatly affect the ability of CNNs to generalize, often leading to overfitting. Since the pre-trained model has already used a large number of datasets for training, it already has the ability to extract shallow underlying features and deep abstract features. Therefore, training from scratch without fine-tuning requires a large amount of data, computational time, and computational resources. Not doing fine-tuning also leads to non-convergence of the model, insufficient optimization of parameters, low accuracy, and low generalization ability of the model. If we use fine-tuning, we can avoid these problems. The fine-tuning step starts by cutting off the last set of fully connected layers from the pre-trained CNN model, and then, we replace the head with a new set of fully connected layers. Then, we freeze all the layers below the head and train the new fully connected layers.

5.2 VGG16

The depth of a CNN plays an important role in classification accuracy and proper detection performance, so the classification error decreases as the depth of the CNN increases. The commonly used CNN networks are AlexNet, ResNet, GoogLeNet, ResNet, etc. VGGNet with very deep is significantly better than AlexNet; although VGGNet is inferior to GoogLeNet in classification ability, its network topology is simpler than GoogLeNet [12]. Therefore, on balance, VGG is chosen for this project. VGG16 is one of the two VGG architectures introduced by Simonyan and Zisserman in 2014, which contains 13 convolutional layers with 3*3 kernels and 5 max-pooling layers with 2*2.

5.3 Dropout

The overfitting problem becomes a serious problem for convolutional neural networks. Standard backpropagation learning causes fragile co-adaptation, resulting in lack of generalization to data that has not been seen before [13] . Dropout can solve the overfitting problem. The key idea is to

Table 1: Model configuration

Layer	Output shape	Param number
input	(256, 256, 3)	0
vgg16	(8, 8, 512)	1471468
AvgPool2D	(3, 3, 512)	0
Flatten	(4608)	0
Dense1	(200)	921800
dropout1	(200)	0
Dense2	(100)	20100
dropout2	(100)	0
Dense3	(2)	202

randomly remove units from the neural network at any stage of training and then allow the remaining neurons to participate in the network training process, which prevents overfitting between units.

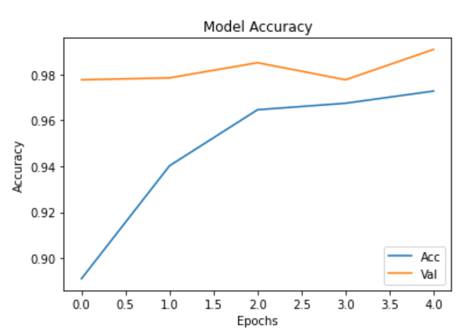
5.4 Modified model used in this project

For the application scenario of the apple spoilage classifier, we use a modified CNN model. We use a pre-trained model based on VGG16, set include_top to Fasle, and use the imagenet dataset. Then all layers below its head are frozen. The VGG16-based pre-trained model is followed by a classifier block. This classifier block uses an average-pooling layer with 3*3 and stride of 2. Then the flattening layer is applied. After the combination of two fully connected layers and the dropout layer. Finally the output of the dropout layer is connected to a fully-connected layer with two neurons, each corresponding to the probability of two labels. The configuration of the model used in this project is shown in Table 1.

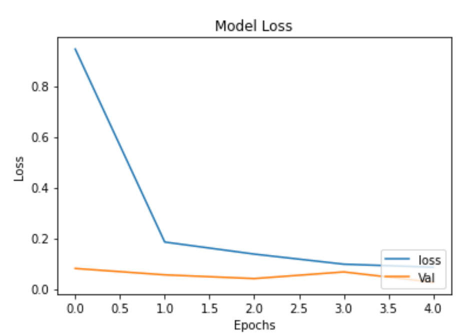
6 Model Training Results and Model Testing

6.1 Model training results

In this project, we used a modified CNN model of VGG16 with an additional classifier block, which includes avarage-pooling layer, dropout layer and dense layer. The model is trained after five epochs. Images of the model accuracy and loss with epoch change are shown in Figure 2. We can see that as epoch increases, the model accuracy increases and loss decreases.



(a) Model accuracy over epochs.



(b) Model loss over epochs.

Figure 2: Model accuracy and loss over epochs.

6.2 Testing on the full test data

We used test data from Kaggle [11], containing both healthy and rotten apples, containing a total of 996 healthy and rotten apples in two categories. After testing using these test data, our trained model has 98% accuracy.

Table 2: Test results

Test data	Accuracy
All test data	0.982
Healthy red apples	0.979
Healthy green apples	1.000

6.3 Testing after classifying the test data

The original training data was classified into red and green according to the color of the apples to test the model's effectiveness in testing different colors of apples. This idea was suggested by Prof. Chin during the presentation. Since it is difficult to determine the color of an apple when it is healthy after it rots, it is difficult to classify rotten apples by color. Therefore, I classified only healthy apples into red healthy apples and green healthy apples. After testing, our trained model has 98% accuracy in detecting red healthy apples and 100% accuracy in detecting green healthy apples. The results obtained in this section and Section 6.3 are placed in Table 2.

7 Conclusion and outlook

7.1 Conclusion

An automated vision-based system to differentiate fresh from rotten fruit would greatly reduce food waste, food-borne illness, and ecological losses. Classifying apples according to whether they are spoiled is an important need for the food industry. Traditional classification methods require many different steps. In this project, a deep CNN is used to overcome the complexity of traditional systems to examine whether apples are rotten or not. This project used the VGG16 architecture and an additional classifier block to build the CNN model. This classifier block uses avarage-pooling layers, dropout layers and dense layers. The dataset used in the project is from kaggle, and data augmentation is performed on the original training dataset when training the model, and then the training and validation data are divided according to a 7:3 ratio. When testing the model, the healthy apples in the test dataset were divided into red and green healthy apples. After testing, the accuracy of the trained model ranged from 98 to 100% after 5 epochs. These results show the high classification performance of the CNN model in this project for detecting healthy and rotten apples.

7.2 Areas of improvement

Although the trained model in this project achieved good classification results, there are many areas that can be improved. First, the number of datasets could be increased. When the dataset is too small, the training CNN model is likely to be overfitted, which affects the model's ability to generalize. Although data augmentation techniques were used, the number was still too small. Second, this project focuses on one fruit, apple, and the number of datasets can be increased while increasing the variety of fruit types in the data, such as adding bananas, peaches, and other common fruit images. Then the structure of the model need to be modified to achieve the purpose of spoilage detection for multiple fruit types. Finally, when fine-tuning the pre-training architecture VGG16, all layers below the head were frozen, and the other layers were not unfrozen. In the future, when the dataset is expanded or changed, it is also beneficial to modify the original convolutional layers during the fine-tuning process for new datasets. After the fully connected layers are learned, we can gradually unfreeze the rest of the network and continue training the model.

References

[1] Singh, S., & Singh, N. P. (2019). Machine learning-based classification of good and rotten apple. In Recent trends in communication, computing, and electronics (pp. 377-386). Springer, Singapore.

[2] Karakaya, D., Ulucan, O., & Turkan, M. (2019). A comparative analysis on fruit freshness classification. In 2019 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-4). IEEE.

[3] Valentino, F., Cenggoro, T. W., & Pardamean, B. (2021, July). A Design of Deep Learning Experimentation for Fruit Freshness Detection. In IOP Conference Series: Earth and Environmental Science (Vol. 794, No. 1, p. 012110). IOP Publishing.

[4] Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., ... & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. Emerging infectious diseases, 17(1), 7.

[5] Bhargava, A., & Bansal, A. (2021). Fruits and vegetables quality evaluation using computer vision: A review. Journal of King Saud University-Computer and Information Sciences, 33(3), 243-257.

[6] Teena, M., Manickavasagan, A., Mothershaw, A., El Hadi, S., & Jayas, D. S. (2013). Potential of machine vision techniques for detecting fecal and microbial contamination of food products: A review. Food and Bioprocess Technology, 6(7), 1621-1634.

[7] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

[8] Herman, H., Susanto, A., Cenggoro, T. W., Suharjito, S., & Pardamean, B. (2020). Oil Palm Fruit Image Ripeness Classification with Computer Vision using Deep Learning and Visual Attention. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 12(2), 21-27.

[9] Nasiri, A., Taheri-Garavand, A., & Zhang, Y. D. (2019). Image-based deep learning automated sorting of date fruit. Postharvest biology and technology, 153, 133-141.

[10] Song, Z., Fu, L., Wu, J., Liu, Z., Li, R., & Cui, Y. (2019). Kiwifruit detection in field images using Faster R-CNN with VGG16. IFAC-PapersOnLine, 52(30), 76-81.

[11] https://www.kaggle.com/sriramr/fruits-fresh-and-rotten-for-classification

[12] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[13] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), 1929-1958.