Fruit Image Recognition

Kiwi Yu
Data Analytics program
Georgetown University
Washington, D.C., U.S.A
ky240@georgetown.edu

Jiahui Chen
Data Analytics program
Georgetown University
Washington, D.C., U.S.A
jc3029@georgetown.edu

Yunxi Zhang
Data Analytics program
Georgetown University
Washington, D.C., U.S.A
yz721@georgetown.edu

Yuzheng Xie
Data Analytics program
Georgetown University
Washington, D.C., U.S.A
yx206@georgetown.edu

***Abstract*— Images could be classified in several ways. The project proposed to recognize five different classes of fruits which are the images from the fruit-360 dataset. The project applied several algorithms to the fruit images recognition. including CNN, VGG16, PCA, KNN, Decision Trees and Random Forest. The CNN model performs the best with accuracy of 99.99%. The applications of our work and future research are discussed.**

Keywords—Image Recognition, Fruit, CNN, VGG16, PCA, Decision Tree, Random Forest

I. Introduction

The appearances of fruits really play an important role in the fruits market, which impact their market value, the preferences and the choices of the consumers. Meanwhile, more and more machine learning algorithms could be applied to the fruits market, such as convolutional neural network (CNN) models. With the difference of the characteristics, like color, size, and shape, of various fruits, it is possible to assess the types and the outside quality of fruits. This technique could benefit the fruits market a lot: correcting the sales number and inventory; identifying the types of fruits, especially for those similar fruits like cilantro and parsley; improving customer satisfaction.

Given the perspective below, the project primarily focused on recognizing the fruits images. In other words, the project developed several models, including CNN model, VGG16 model, KNN model, Decision Tree model and Random Forest model to classify given fruits images correctly. Meanwhile, before constructing the KNN model, Decision Tree model, and Random Forest model, the project utilized the PCA method to reduce the dimensions of Fruits dataset.

Due to the variety of sizes, colors, and perspectives of the same objects, recognizing fruits might be challenging, and not all models are suitable for this problem. Thus, the project also compared the prediction accuracy of each model and found the best models with highest prediction accuracy.

II. Related Work

S.Arivazhagan et al. [1] discussed a new method to recognize by combining the different color, size, shape and texture together. They extracted the surface information and geometry information from the images to improve the performance. 'HSV representation' was used to represent luminance and chrominance and a co-occurrence matrix was constructed, from which the co-occurrence features could be calculated. Their classification was based on the Minimum Distance Criterion.

Jose Luis Rojas-Aranda et al. [2] conducted an image classification on lightweight Convolutional Neural Network Model, aimed at speeding up the checkout process for grocery stores. This paper classified fruits by an improved CNN architecture based on MobileNetV2, which proposed the addition of different input features, such as RGB color. The results show that the accuracy has been improved: the classification accuracy for fruits with no plastic bag is 95%; the accuracy for fruits in a plastic bag is 93%.

Zhenbo Li et al. [3] presented the improved VGG network model to train the vegetable image dataset. The research combined the output feature of the first two fully connected layers, and then added the Batch Normalization layers (VGG-M-BN) in order to improve the accuracy of the recognition of vegetables. Meanwhile, it verified that the ReLU activation function is better than the Sigmoid and Tanh in this model. Finally, the recognition accuracy of the improved VGG model is 96.5%, and the accuracy could be improved with the increment of the batch size.

Yudong Zhang et al. [4] used the principal component analysis (PCA) method to reduce the dimensions of features. With the dataset with less dimensions, they constructed SVM models to solve the automatic classification of fruits via computer vision.

Myint San et al. [5] processed the fruit360 dataset by four steps: pre- processing, boundary extraction, feature extractions, and classification. This study extracted the color features by the RGB color channel and got the morphological features by detecting the boundary of fruit. The authors conducted several classifiers, such as the Random Forest model, Decision Tree model, and KNN model. Finally, the Random Forest model performed better than others.

III. Dataset

The project utilized the "Fruits 360" dataset on Kaggle [6], includes images of 131 types of fruits and vegetables, and 90,483 images. Different varieties of the same fruit are stored as belonging to different categories.

The project chose five common categories of fruits to classify, considering the time it takes for the program to run. The list of five different fruits is: lychee, peach, kiwi, eggplant and banana. Among them, there are two files - test, train - in this dataset: the size of the training set is 2406, and the size of testing size is 808. Meanwhile, as for the training data, 20% of it was used for the validation set.

IV. Modeling

A. CNN

CNN refers to the convolutional neural network, which has multiple hidden layers for detecting the various patterns that exist in each image passed through the model. After each layer, an activation function is needed in order to guide the signal to the next layer, and in our case, ReLu is used[2]. A pooling layer would apply operations to each feature map.

Before building the model, we performed image data augmentation using *ImageDataGenerator* in order to expand the training dataset. Our CNN model is composed of four layers with 'rule' as activation, 2x2 kernel size and 'adam' as optimizer. We set the learning rate to 0.01. The dropout and early stop were added to avoid overfitting. We choose batch size = 128 and epchos = 15. The model stopped at epoch 13 and finished with an accuracy of 99.99%, the AUC curve is given in figure 1.

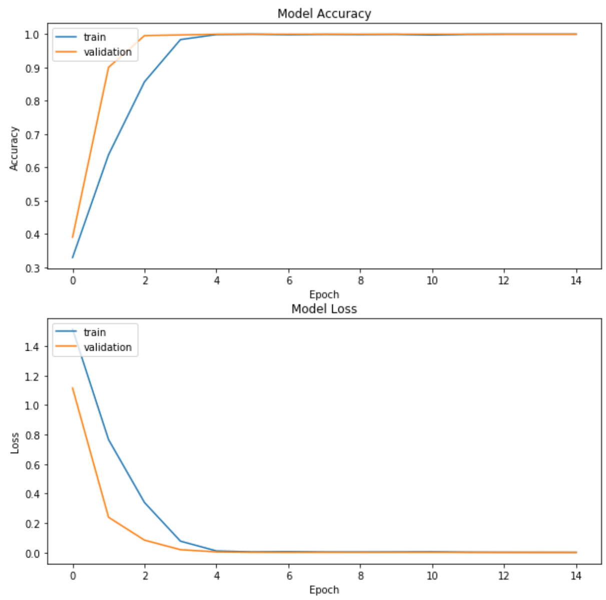

Fig. 1. Training and validation accuracy and loss of CNN

B. VGG16

In order to enrich our model selections, we further applied the VGG16. VGG16 is a transfer learning algorithm, namely a model which has already been developed and is reused as the starting point model for another task. The merits for the transfer learning model is its high accuracy and tangibility [3]. However, the drawback still exists, specifically the long time needed for applying the model. Just like CNN, we conducted data augmentation to expand datasets images. Our VGG16 is designed to be 100 x 100 x 3 with 'softmax' as activation function and 'adam' as optimizer. We used early-stop with a 0.01 learning rate. By exploring the accuracy for different steps, we could generate the best achievable one. The final accuracy for the VGG16 model is 98.7%.

Fig. 2. Training and validation loss of VGG16

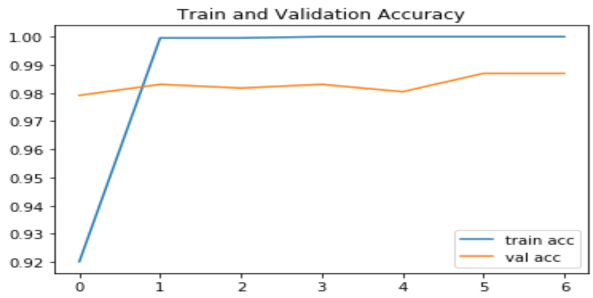

Fig. 3. Training and validation accuracy of VGG16

C. PCA

To figure out how the dataset would appear in lower dimension, the project utilizes Principal Components Analysis (PCA) method to reduce dimensionality [4]. In this way, the project could visualize and plot the dataset by representing the dataset similarly while greatly reducing the number of dimensions.

After scaling the training dataset, the project fits the PCA model. Then, the project visualized some sample images obtaining its principal components from the whole dataset (Figure 4 & Figure 5), whose plots show no color. Besides, it may be possible for the project to classify these fruits by considering a lower number of dimensions instead of all.

Fig. 4. Fruit images before PCA

Fig. 5. Fruit images after PCA

In order to check the validity of the PCA model, the project plots the number of components plotted against explained variance (Figure 6), which shows that it is feasible for the project to reduce the Fruit 360 dataset to 52 components with 99.87% variance retained.

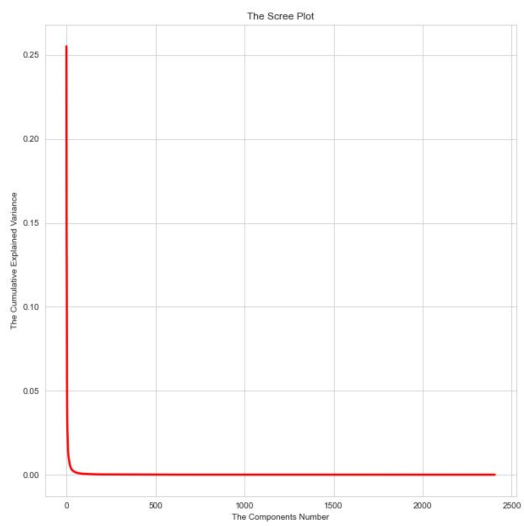

Fig. 6. The Scree Plot

Based on the PCA model result when the number of components equals to 5, the project also conducted T-SNE (T-distributed Stochastic Neighbor Embedding), a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions and set the number of components as 2. Furthermore, the project plots the two dimensions results of T-SNE (Figure 7) and marks different types of fruits by their images.

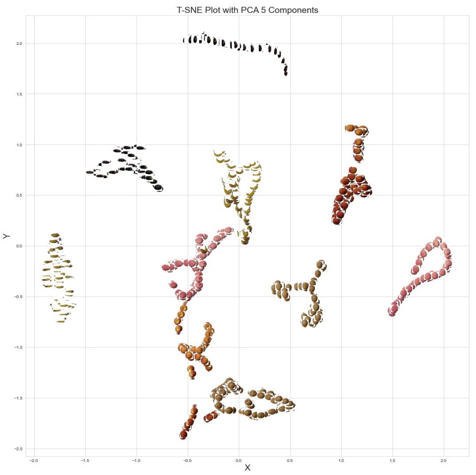

Fig. 7. T-SNE Plot (PCA 5 Components)

D. KNN

The k-nearest neighbors (KNN) algorithm is a supervised machine learning algorithm that could be used for classification problems, which assumes that similar things exist in close proximity [5]. Based on the results of the PCA model - a lower dimensional dataset, the project conducts the KNN model to classify and group images and predict where new fruits will be grouped.

After fitting the model, the project plots the accuracy from k equals 1 to 50 (Figure 8 & Table 1). From Figure 8 and Table 1, the project could discover that the highest accuracy on the test set appears when k equals to 1. Setting k as 1 means that the fruits are simply assigned to the class of that single nearest neighbor. Finally, it turns out that the accuracy of the KNN model is 97.65% with k equals to 1.

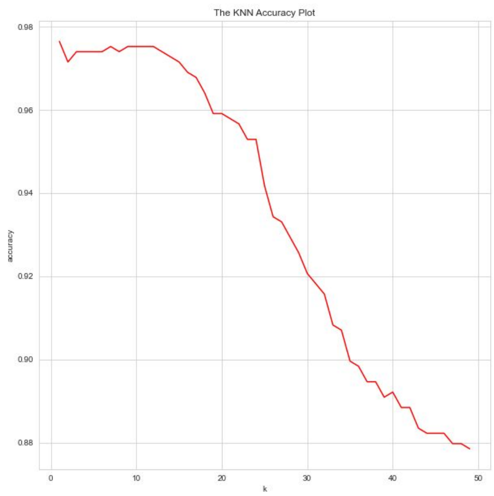

Fig. 8. The KNN Accuracy Plot

TABLE I. Accuracy for KNN under different k

	k	accuracy	accuracy_shift	difference
0	1	0.976456	0.971499	0.004957
1	2	0.971499	0.973978	-0.002478
2	3	0.973978	0.973978	0.000000
3	4	0.973978	0.973978	0.000000
4	5	0.973978	0.973978	0.000000
5	6	0.973978	0.975217	-0.001239
6	7	0.975217	0.973978	0.001239
7	8	0.973978	0.975217	-0.001239
8	9	0.975217	0.975217	0.000000
9	10	0.975217	0.975217	0.000000
10	11	0.975217	0.975217	0.000000
11	12	0.975217	0.973978	0.001239
12	13	0.973978	0.972739	0.001239
13	14	0.972739	0.971499	0.001239
14	15	0.971499	0.969021	0.002478
15	16	0.969021	0.967782	0.001239
16	17	0.967782	0.964064	0.003717
17	18	0.964064	0.959108	0.004957
18	19	0.959108	0.959108	0.000000
19	20	0.959108	0.957869	0.001239
20	21	0.957869	0.956629	0.001239

E. Decision Tree & Random Forest

Decision Trees (DTs) is a non-parametric supervised learning method that could be utilized in classification problems, whose goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features [5]. Meanwhile, Random forest (RF) [5] consists of a large number of individual decision trees, which could correct for decision trees' habit of overfitting to their training set. Based on the dataset from the PCA model, the project conducts the Decision Tree model and Random Forest model.

Since the Decision Tree model performs differently with several parameters, the project plots the accuracy of Decision Tree models which set the max depth of the tree from 1 to 50 (Figure 9 & Table 2). From Figure 9 and Table 2, the project could discover that the max depth of the tree with the highest accuracy on the test set is 9, and it turns out that the final accuracy of the Decision Tree model is 88.23% when max depth equals to 9.

TABLE II. Accuracy for Decision Tree under different depth

	max_depth	accuracy	accuracy_shift	difference
0	1	0.408922	0.629492	-0.220570
1	2	0.629492	0.827757	-0.198265
2	3	0.827757	0.864932	-0.037175
3	4	0.864932	0.874845	-0.009913
4	5	0.874845	0.853779	0.021066
5	6	0.853779	0.877323	-0.023544
6	7	0.877323	0.855019	0.022305
7	8	0.855019	0.882280	-0.027261
8	9	0.882280	0.881041	0.001239
9	10	0.881041	0.877323	0.003717
10	11	0.877323	0.861214	0.016109

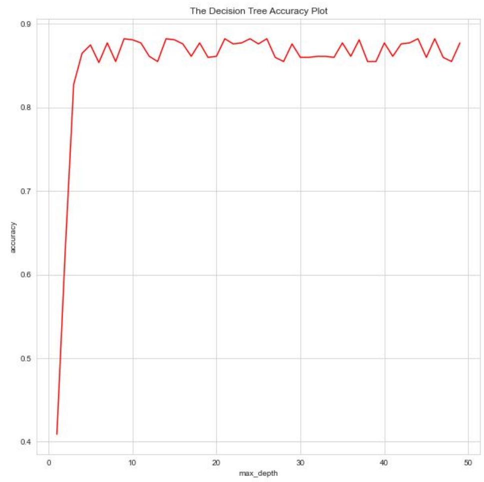

Fig. 9. The Decision Tree Accuracy Plot

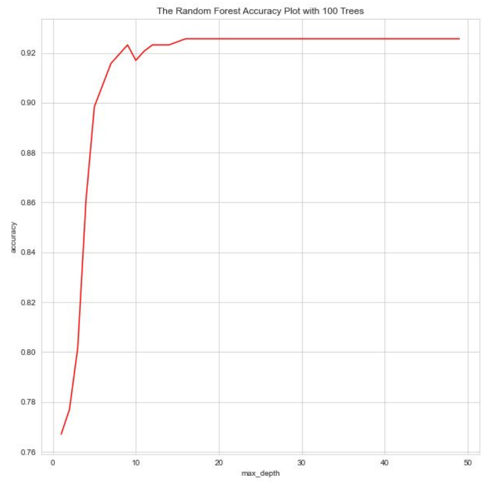

Fig. 10. The Random Forest Accuracy Plot with 100 Trees

Similarly, the project also plots the accuracy of Random Forest models which set the max depth of the tree from 1 to 50 with 100 trees(Figure 10 and Table 3). From Figure 10 and Table 3, the project could discover that the max depth of the tree with the highest accuracy on the test set is 16, and it turns out that the final accuracy of the Random Forest model is 92.57% when max depth equals to 16, which is a little higher than the accuracy of Decision Tree model.

TABLE III. Accuracy for Random Forest under different depth

	max_depth	accuracy	accuracy_shift	difference
0	1	0.767038	0.776952	-0.009913
1	2	0.776952	0.801735	-0.024783
2	3	0.801735	0.861214	-0.059480
3	4	0.861214	0.898389	-0.037175
4	5	0.898389	0.907063	-0.008674
5	6	0.907063	0.915737	-0.008674
6	7	0.915737	0.919455	-0.003717
7	8	0.919455	0.923172	-0.003717
8	9	0.923172	0.916976	0.006196
9	10	0.916976	0.920694	-0.003717
10	11	0.920694	0.923172	-0.002478
11	12	0.923172	0.923172	0.000000
12	13	0.923172	0.923172	0.000000
13	14	0.923172	0.924411	-0.001239
14	15	0.924411	0.925651	-0.001239
15	16	0.925651	0.925651	0.000000
16	17	0.925651	0.925651	0.000000
17	18	0.925651	0.925651	0.000000
18	19	0.925651	0.925651	0.000000
19	20	0.925651	0.925651	0.000000
20	21	0.925651	0.925651	0.000000

V. Conclusion

This project is designed to investigate the methodology on improving the accuracy of fruits' images recognition through various machine learning models. Before we apply the models, we implement the PCA method to reduce the dimensionality in order to achieve higher accuracy in the later stage. We built 5 machine learning models, namely decision tree, random forest, CNN, KNN, and VGG16, to train our dataset and get the testing accuracy. All the five models get great results, and CNN achieves the best among others with 99.99% accuracy. While CNN is currently the best model for our dataset, we still have reason to believe that VGG16 would even get better results in the future research when we enrich the dataset size or train more images in the larger ImageNet database. Although the VGG16 could have more significant uses in the future, we should also admit the significant amount of time cost under the VGG16 model.

Through this project, we have learned the methodology on classifying images using different models. As we learned more about the neural network, we realized the importance of the computing and algorithms to run these models. For the future research, we could use a more diverse dataset to test the machine learning models we have chosen. We are also going to find ways to solve the overfitting problem of the VGG16 model and improve its performance.

References

[1] Arivazhagan, S., Shebiah, R. N., Nidhyanandhan, S. S., & Ganesan, L. (2010). Fruit recognition using color and texture features. Journal of Emerging Trends in Computing and Information Sciences, 1(2), 90-94.

[2] Rojas-Aranda, J. L., Nunez-Varela, J. I., Cuevas-Tello, J. C., & Rangel-Ramirez, G. (2020, June). Fruit Classification for Retail Stores Using Deep Learning. In Mexican Conference on Pattern Recognition (pp. 3-13). Springer, Cham.

[3] Li, Z., Li, F., Zhu, L., & Yue, J. (2020). Vegetable Recognition and Classification Based on Improved VGG Deep Learning Network Model. International Journal of Computational Intelligence Systems, 13(1), 559-564.

[4] Zhang, Y., & Wu, L. (2012). Classification of fruits using computer vision and a multiclass support vector machine. sensors, 12(9), 12489-12505.

[5] San, M., Aung, M. M., & Khaing, P. P. (2019). Fruit Recognition Using Color and Morphological Features Fusion. International Journal of Image, Graphics and Signal Processing, 11(10), 8.

[6] Mureşan, H., & Oltean, M. (2018). Fruit recognition from images using deep learning. Acta Universitatis Sapientiae, Informatica, 10(1), 26-42