Fruits classification for retail store dataset using machine learning

Sengeni K, Rasathi K, Parkavi K
Department of Computer Science and Engineering, Agni college of technology
Affiliated to Anna University, Chennai.

Guided by:
Mrs. Abinaya P
abinaya.cse@act.edu.in
Assistant Professor
Department of Computer Science and Engineering, Agni college of technology
Affiliated to Anna University, Chennai.

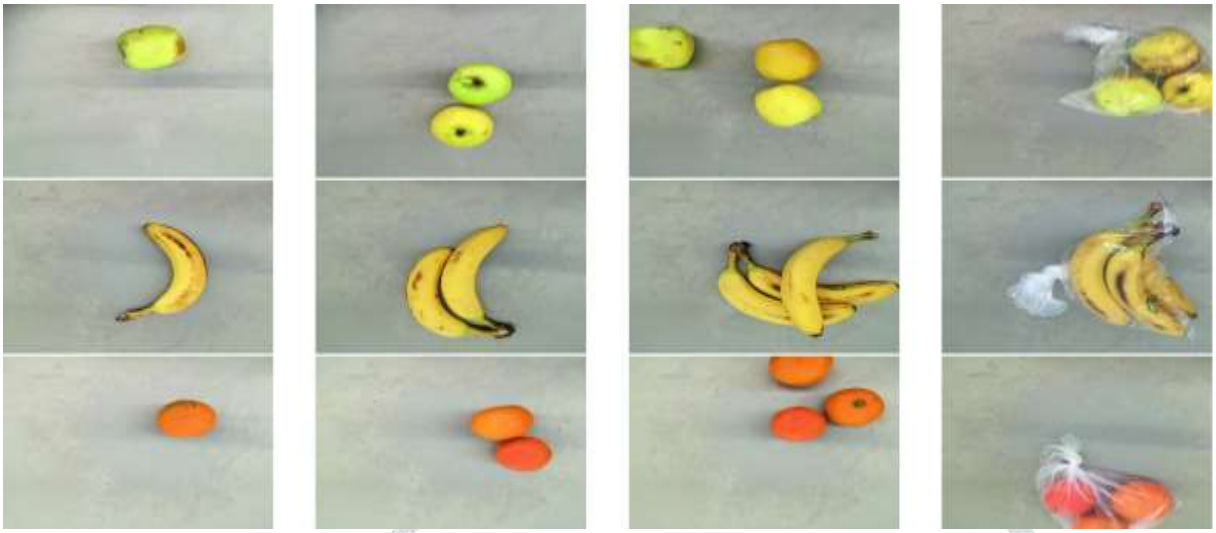
1.Abstract

Fruits are normally require them to be manually identified. This paper aims to find a best way of fruit classification method. This can be done by using supervised machine learning algorithms and image processing mechanism. It based on multi-feature extraction methods. Initially we preprocess the sample of fruit. It separating foreground and background using preprocessing method. The Scaling and cropping can be performed to reduce the image dimension. So that the processing is fast. Then we extract features from the fruit image and we also check the quality of fruits. Extracted features are fitted into the AdaBoost classifier machine learning algorithm. The obtained result will gives us the prediction accuracy. This result have been collected using a fruit image database . It consist of 5 different classes of fruits and 120 fruits image overall. Therefore average prediction accuracy of more than 85%.

2.Introduction

In retail stores the customers can rely on cashier or self-service system. In the stores many have products have barcodes that can be scanned and identify the product with cost. But some product does not have barcode for example fruits. The fruits are processed differently. The cashier or the customer need to manually identify the fruits. Fruit classification is a complex problem can be identified i)to check the classification of fruits with different types ii)to check the quality of the fruits difficult to achieve due to differences in shape and color. The problem is directly related to the purchase of fruits in stores, if that fruits can be inside a plastic bag it can be difficult to identify the fruits. This paper deals with the first type of classification(i.e., classification of fruits with different types), where fruits can be inside or without a plastic bag. The Recent advancements in Neural Network make them suitable for this problem. In addition , we created a new fruit dataset for five type of fruits also considering

fruits in transparent plastic bags. This results show an overall 85% classification accuracy for fruits without plastic bag, and 83% for fruits inside a plastic bag.



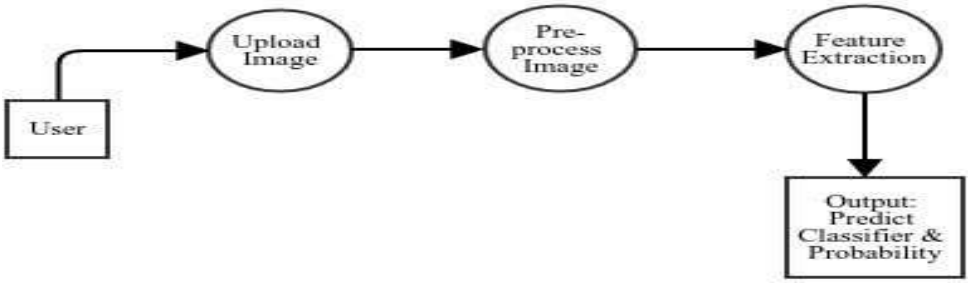
sample fruits of image

3.Algorithm

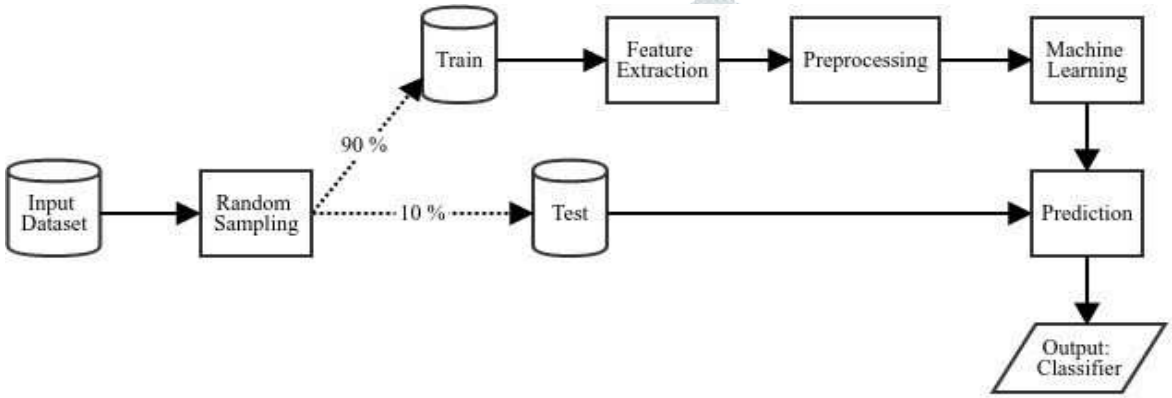
Our proposed system is to solve the classification of fruits in with or without plastic bag. Our main goal is to replicate a store environment, the fruits were placed over plastic bag and the photos were taken from the top . We choose to work with the five types of fruit. we introduce variation in the dataset, by taking images of the fruits at different position and side. we can also consider the checkout process of fruits are generally inside a transparent plastic bag, so that image fruits also taken with the plastic bag. There is no correct architecture for all problems

Fruits in retail stores normally require them to be manually identified. Our data set of images is introduced that considers three classes of fruits, inside or without plastic bags and also check the quality of the fruits in our project. It also increase the increase the accuracy of the speed.

There are different algorithms is used in this project. The Haar-Like Feature Extraction algorithm has been used to identify the fruits. Then Hue Histogram Feature Extraction algorithm which can find a number of pixels with particular hue then it produce results. It takes the color image and returns the normalized color histogram. The length and direction line of fruits can be found by using Edge Histogram Feature Extraction algorithm. This project mainly use the Ada Boosting algorithm to classify the fruits.



Event diagram



Abstract model of the system

4.Literature Review

Object recognition has been studied more than four decades. significant effort have been paid to develop representation schemes and algorithms aiming at recognizing generic objects in image taken under different imaging condition. object recognition along with machine learning algorithm and image processing has been implemented in fruit classification.

Yudong Zhang and Lenan wu have implemented fruits classification in their classification of fruits using computer vision and a multicalss support vector machine. In their research, first the fruit images were acquired by the digital camera. Second pre-processing of the image was carried out by the removing the background of each image by split-and -merge algorithm. The input was a database of 1,653 images with 18 categories of fruits, each image size was 256*256 . Third, the color histogram, texture and shape features of each pre-processed image were extracted to compose a feature space . Altogether 79 features were extracted from each 256*256 image. These 79 feature included 64 color features, seven textures features, and eight shape features. Fourthly, Principal Component Analysis(PCA) was used to reduce the dimension of features space. Achieved, object recognition is also related to content

based image retrieval and multimedia indexing as a number of generic object can be recognized.

5.Future Scope

In our data set of images is introduced that considers three classes of fruits, inside or without plastic bags and also check the quality of the fruits in our project. In future we also check the weight of the fruit and tells the cost of fruits based on their weight.

6.Conclusion

This project aims to classify the fruit image based on its Haar-Like, Hue and edge histogram features. The project is designed in such a way that can read image , extract future, preprocesses it and quality of fruits it implemented by machine learning algorithm and generate output based on input provider . The project has been able to classify the fruit image and quality of fruits. The cross validation score obtained is 84% with learning rate of and the prediction accuracy of system is above 80% . In some case system does not predict the fruit images even the provided input calls under the training category. With this result it can be concluded that the chosen ensemble machine learning algorithm is not suitable for fruit classification problem. So we can use Ada Boost algorithm it can provide better result.

7.References

1.M. Škrjanec. "Automatic fruit recognition using computer vision", Bsc Thesis, (Mentor: Matej Kristan), Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2013.

2."Welcome | Flask (A Python Microframework)", Flask.pocoo.org, 2016. [Online]. Available: http://flask.pocoo.org/. [Accessed: 19- Aug- 2016].

3."Scientific Computing Tools for Python", Scipy.org, 2016 [Online] Available: https://www.scipy.org/about.html [Accessed: 20- Aug- 2016].

4."scikit-learn: machine learning in Python — scikit-learn 0.17.1 documentation", Scikit-learn.org, 2016. [Online]. Available: http://scikit-learn.org/stable/. [Accessed: 20- Aug- 2016].

5."SimpleCV", Simplecv.org, 2016. [Online]. Available: http://simplecv.org/. [Accessed: 20- Aug- 2016].

6.Hameed, K., Chai, D., Rassau, A.: A comprehensive review of fruit and vegetable classification techniques. Image Vis. Comput. **80**, 24–44 (2018)CrossRefGoogle Scholar

7.Mureşan, H., Oltean, M.: Fruit recognition from images using deep learning. Acta Universitatis Sapientiae Informatica **10**(1), 26–42 (2018)CrossRefGoogle Scholar

8.Katarzyna, R., Paweł, M.: A vision-based method utilizing deep convolutional neural networks for fruit variety classification in uncertainty conditions of retail sales. Appl. Sci. **9**(19), 3971 (2019)CrossRefGoogle Scholar

9.Abadi, M., Agarwal, A., Barham, P., Goodfellow, I., et. al.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015). https://www.tensorflow.org/, software available from tensorflow.org

10.Bargoti, S., Underwood, J.: Deep fruit detection in orchards. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 3626–3633 (2017)Google Scholar

11.Femling, F., Olsson, A., Alonso-Fernandez, F.: Fruit and vegetable identification using machine learning for retail applications. In: 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), pp. 9–15. IEEE (2018)Google Scholar

12.He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)Google Scholar

13.Hossain, M.S., Al-Hammadi, M., Muhammad, G.: Automatic fruit classification using deep learning for industrial applications. IEEE Trans. Ind. Inf. **15**(2), 1027–1034 (2018)CrossRefGoogle Scholar

14.Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C.: Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of 'MangoYOLO'. Precis. Agric. **20**(6), 1107–1135 (2019)CrossRefGoogle Scholar

14.Rahnemoonfar, M., Sheppard, C.: Deep count: fruit counting based on deep simulated learning. Sensors **17**(4), 905 (2017)CrossRefGoogle Scholar

15.Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520 (2018)Google Scholar