# ABSTRACT

Fruit recognition and automatic classification of fruit is still a complicated task due to the various properties of numerous types of fruits. Fruits have different shapes, size, color texture and other properties. Some fruits share same characteristics like color, texture, size, etc. This paper proposes an algorithm to fruit classification method using machine learning algorithm based on multi-feature. Firstly, we pre-process the training sample of fruits image. Preprocessing includes separating foreground and background, cropping the image to reduce the dimension so that the processing is fast and then we divide the target area. Then, we extract fruit features, which includes color, texture and shape of the fruit image. Extracted features are fitted into the Tree Classifier machine learning algorithm. Finally, the results obtained from the machine learning network are cross validated with the input data (test sample). The output obtained will give us the prediction accuracy and class of the fruit it has recognized.

1.0 Introduction

Object Recognition is an important field in computer science. Object recognition is emerging technology to detect and classify objects based on their characteristics. It also employs pattern recognition on different objects. To achieve good object detection, classification and recognition, different machine learning algorithms and statistical methods are implemented. While using machine learning algorithms, it is not a guarantee that every algorithms give accurate result. The accuracy achievement can be different for different algorithms. Hence, we need to select the best algorithm which has maximum classification accuracy and prediction accuracy.

For fruit classification and detecting we implement a portion of object recognition with machine learning and neural network model. The rapid development of computer vision, image processing and recognition advancement in computer technology provide the fruit machine classification with the possibility. In recent years, fruit recognition using computer vision is gradually applied in agriculture sector, education sector and supermarkets [1]. Computer vision has been widely used in industries to aid in automatic and checking processes[2]. The important problem in computer vision and pattern recognition is the shape matching. Shape comparison and shape matching can be done using computer vision. Shape matching applications contain image registration, object detection and recognition, and images content based retrieval [3]. Many agricultural applications used image processing to automate its duty. Detecting crop diseases is one of these applications. The crop images are analyzed to discovered the affected diseases [4]. Also, image processing are used

1.1 Problem Statement

Automatic fruit classification is a challenging task. Algorithm being used, fruit database, fruit image shape, color plays vital role while classifying the fruit. The problems and characteristics may lead to lower accuracy and we can’t expect accurate result.

Many related works are been conducted in fruit classification using different classification algorithms but still there is an issue. Some research are done by taking only three fruits into consideration with 100% accuracy[5]. But considering only three fruit is not enough because it cant recognize other fruits out of sample.

Similarly, usage of proper algorithm should be taken into consideration while performing classification. Using Multiclass Support Vector Machine algorithm, gives success rate or accuracy in range of 70% to 75% [6]. This results in wrong interpretation of given input for out of sample fruits and similar fruits.

1.2 Objective of the project

The main objectives of project are:

a) To produce more accuracy in the result by using Tree Classifier Algorithm. Generating more accuracy means more accurate prediction for fruits.

b) To train the Algorithm with maximum number of training samples. More the training samples, higher will be the accuracy while predicting new class of fruit.

c) Develop a mechanism to train the samples with minimum amount of time by reducing the complexities.

d) Make code open source so that others can review, take reference and contribute.

e) Develop a platform where others can test their fruit image and get the accuracy of what the fruit is.

1.2 Scope of the project

Fruit classification is based on the fruit images that are available in the fruit database. So, the scope of the project is only limited to edible fruits that are available. Leafy vegetables like lettuce, cabbage, Spinach may not be in the scope of the project. So, when these are provided as input to the system, it can’t recognize and produce the desired result.

1.3 Project Features

Fruit Classifier project consists of certain key features. The features are as follows :

a) Visualization

b) Similar fruit recommendation ##MAY NOT NEED

a) Visualization

Visualization of the result is the key feature of the project. In visualization, user is prompted with the predicted fruit class and percentage of accuracy of that fruit. Visualization is done in web application by rendering the amount of accuracy percentage calculated by the Classifier Algorithm being used. If the prediction accuracy is below some threshold, then the nearest resemblance of that fruit is shown based on the prediction probability.

b) Similar fruit recommendation

It is another feature where user is shown fruits similar to the user has uploaded for the prediction. This is done on the basis of accuracy gained by the classes that are similar to the fruit provided by the user. Custom Google search API will be used to grab image from the Google image engine to render the image in the web application.

1.5 Requirement Analysis and Feasibility study

Visualization of the predicted accuracy and recommendation of similar class of fruit is the main requirement of this project and whether such visualization and recommendation can be performed or not is analyzed doing the feasibility study.

1.5.1 Requirement Gathering:

Visualization and similar recommendation is the main requirement of the project. For the purpose of these requirement, the crucial thing was to implement machine learning approach to classify the given input based on the training data. The training data was collected from Visual Cognitive Systems Laboratory University of Ljubljana Faculty of Computer and information Science Fruit Image Data Set [7]. The data consists of 30 fruits sample. While each class of fruit has more than 40 sample images. These images were used to train the machine using Tree Classifier Machine learning algotithm. The sets of data is definitely not enough to cover wider range and classes of fruit, but due to project deadline and the project is solely based on inplementation of the machine learning algorithm to automatically classify an image rather than as a product, these data sets were considered enough for the purpose of the project.

Some of the sample images are:



1.5.2 Feasibility Study

After gathering of the required resource, whether the completion of the project with the gathered resource is feasible or not is checked using the following feasibility analysis.

a) Technical Feasibility:

The project is technically feasible as it can be built using the existing available technologies. The recommendation of the similar fruits can be done using custom made Google image search. We can create our custom search engine which captures data from the web and displays in our web portal. This can be done using automation techniques.

b) Economic Feasibility:

The project is economically feasible as the cost of the project is involved only in the hosting of

the project. As the training samples increases, the training mechanism can consume more time and processing power. In that case better processor might be needed.

c) Operational Feasibility:

The project is operationally feasible as the user having basic knowledge about computer and

Internet can use while concept of Machine learning is a plus point. Furthermore the project can be easily used if the computer have Internet access and browser is installed in computer.

1.6 System Requirement – Hardware and Software Platforms

The software required by the user is:

a) Any Operating System ( Windows or Linux or Mac OS)

b) Any web browser supporting JavaScript.

The hardware required by the user is:

a) Computer : PC

b) Processor : Intel P2 and above.

c) RAM : 128 MB and above.

d) Input Device : Mouse, Keyboard.

The software required for the server is:

a) Operating System – Linux based OS

b) Python with supporting libraries (Scikit, SimpleCV, Flask)

c) Apache Webserver

The hardware required for the server is:

a) RAM 4 GB and above.

b) Minimum disk space 10 GB.

2.0 Literature Review

REFERENCES:

1. Patel H.N, Jain R.K and Joshi M.V, “Automatic Segmentation and Yield Measurement of Fruit using Shape Analysis”, International Journal of Computer Applications, Volume 45– No.7, May 2012

2. <http://www.ijera.com/papers/Vol3_issue6/GQ3611971205.pdf>

3. I. Oikonomidis and A. A. Argyros, ”Deformable 2D shape matching based on shape contexts and dynamic programming,” In ISVC 2009. LNCS , vol. 5876, R. Boyle et al. Eds., Springer-Verlag Berlin Heidelberg, 2009, pp. 460-469.

4. A. Camargo and J. S. Smith, ”An image-processing based algorithm to automatically identify plant disease visual symptoms,” Biosystems Engineering, vol. 102, no, 1, pp. 9-21, January 2009.

5. Cairo University, Faculty of Computers and Information, Cairo, Egypt Scientific Research Group in Egypt (SRGE) “Automatic fruit classification using random forest algorithm” 2014 IEEE

6. Classification of Fruits Using Computer Vision and Multiclass Support Vector Machine, Yudong Zhang and Lenan wu, 2012, SENSORS

7. <http://www.vicos.si/Downloads/FIDS30>