Quality Recognition of Banana Images using Classifiers

Dr. A.Anushya

Assistant Professor, Department of Computer Applications
e-mail: anushya.alpho@gmail.com

ABSTRACT: This research essence on banana images classification. Initially the researcher collects banana images (French Plantain is called as Nendran) and constructs the dataset. Gray Level Concurrence Matrix is used to features from image. Then classification algorithms such as Decision Tree and Neural Networks are used to group the images according to its quality. These algorithms are executed on MATLAB and Decision Tree and Neural Networks produced 71% and 75.67% of accuracy respectively.
Keywords: Image Classification, Gray Level Concurrence Matrix, Decision Tree, C 4.5 Algorithm and Neural Networks

INTRODUCTION

Banana is consumed by most of people habitually because of its health benefits and availability of nutrients. Also it is one of the cheapest fruit. Accordingly all and sundry can buying banana. French Plantain (Nendran, Ethakkai, Myndoli and Rajeli) is a prominent plantain variety of Kerala and Tamil Nadu. The fruits are large and plumpy with long pedicel. Fruits turn buff yellow on ripening [1]. In this article Banana and French Plantain are used interchangeably. It has different colors in its ripening stages. Thus the quality of banana classification based on images is problematic process. Image processing techniques are helping to stress-free the task. Gray Level Concurrence Matrix is working on images to extract the features from images. The classification algorithms are applied to classify the images in groups. Image classification is a task the group the images according to their categories. This research recognizes quality of banana images by the features such as color and texture.

Plentiful related research work was previously conducted classification and quality detection in banana images. The most relevant literature was explicated in this section. Anuja Bhargava et.al., piloted a comparative analysis on quality detection of fruits and vegetables using classification techniques like PCA, Multiclass SVM, BPNN, Neural Networks etc., through the features like size, shape, color and defect [ ]. Suresha M et.al., used decision tree to classify 269 vegetables in to 8 classes and reached the accuracy 95% []. Yan Zhang et al., proposed a novel CNN architecture to detect the 17312 banana images with or without defects [ ]. Artificial Neural Networks based frame work using color, brown spots development and Tamura features to classify the banana images by Fatima.M.A.Mazen et.al.,. To verify the result, comparative analysis was done by classification algorithms such as Support Vector Machine, Naïve Bayes, K-Nearest Neighbor and Decision Tree. Overall 100% class recognition accuracy is obtained [5]. Jonatha et.al., [6] applied Image processing methods by image filtering, segmentation and feature extraction for citrus fruits. Support Vector Machine used to classify the quality of citrus fruits and accuracy was gained 97.3%. D.Surya Prabha et.al., assessed 120 banana image in to 3 groups such as under matured, mature and over mature using mean color intensity algorithm and area algorithm. Mean color intensity algorithm produced 99.1% accuracy and area algorithm made 85% accuracy separately [7].

In this research, size is not a relevant feature for classification. Since French plantain is big in size normally than other banana variety. French Plantain is with different ripening stages such as unripe (green), mid ripe (yellow with green), ripe (golden yellow) and over ripe (yellow with black). For this research, only three categories such as mid ripe, ripe and over ripe are considering by the researcher. Thus the researcher collects images in the above mentioned sets and quality is recognized.

The rest of the paper is organized as follows. In section 2, Data source is described. Research methodology is depicted in Section 3, and Section 4 clearly shows the result analysis from the experiments and Section 5 concludes the paper while including argument on finding of the research and on probable future work.

DATA SOURCE

The dataset consists of 270 French Plantain images. These images are collected from retail shop in Pillaithoppu, Kaniyakumari District, Tamil Nadu, India. Because French Plantain is the most famous fruit in Kaniyakumari District, Tamil Nadu and Kerala in India. Images are collected under three categories. They are mid ripe, ripe and over ripe. 90 images are comprised in each category. These images are captured by Samsung J7 mobile camera. Appearance of the sample of image is in dataset with three classes exposed in Figure 1.

Figure 1: Sample dataset

BANANA IMAGES CLASSIFICATION

Classification is the process of learning a model that describes different classes of data. The classes are predetermined. The first step, of learning the model, is accomplished by using a training set of data that has already been classified. Each record in the training data contains an attribute, called the class label that indicates which class the record belongs to. Then, through the learning, it can be forecast the unknown labeled data. This section elucidates the research methodology followed in this research. GLCM features are extracted and Decision Tree and Neural Networks are used to group the real French Plantain images. Gray Level Concurrence Matrix is used for texture features. In the GLCM number of rows and columns are same as number of gray levels of that image. GLCM is used for number of applications in image processing. Images in database are resized and convert both images into Grayscale. Then the Converted Grayscale images are transformed to binary image. In binary image, count the number of pixels concealed by the images by using GLCM feature. Next, Calculate the GLCM feature of both images and find the difference in area and apply the Canny edge detection method on grayscale images. Afterward Extract the hue, saturation and intensity a from the uncropped test image. Now average of difference in GLCM feature, edge and color histogram is obtained. The above methods are repeated for all the images in the database [8]. As a final point, image classification is steered by Decision Tree and Neural Networks.

Decision tree is a classifier in the form of a tree structure and has decision node and leaf node. Decision node specifies a test on a single attribute, Leaf node indicates the value of the target attribute, Arc or edge is split of one attribute and Path is a disjunction of test to make the final decision. Decision trees classify instances or examples by starting at the root of the tree and moving through it until a leaf node. Uses a tree structure to model the training set and classifies a new record following the path in the tree and inner nodes represent attributes and leaves nodes represent the class. The most popular Decision tree algorithm is C4.5 algorithm, by first

introducing the basic methods of its predecessor, ID3 algorithm. The basic algorithm of C4.5 is discussed below.

C 4.5 Algorithm:

C4.5 builds decision trees from a set of training data using the concept of information entropy. The decision tree prescription for synthesizing an efficient decision tree can be stated as follows:

Step 1: Calculate initial value of entropy.

Step 2: Select that feature which results in the maximum decrease in entropy (gain in information), to serve as the root node of the decision tree.

Step 3: Build the next level of the decision tree providing the greatest decrease in entropy.

Step 4: Repeat Steps 1 through 3. Continue the procedure until all subpopulations are of a single class and the system entropy is zero.

Neural Networks is a field of Artificial Intelligence (AI) where we, by inspiration from the human brain, find data structures and algorithms for learning and classification of data. It is a set of connected input/output units where each connection has a weight associated with it. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct class label of the input tuples. It is also referred to as connectionist learning due to the connections between units. There are three layers: input, intermediate (called the hidden layer) and output. Several hidden layers can be placed between the input and output layers. Neural Networks (NN) as a Classifier, the learning algorithm of NN is a supervised learning method by training feed forward neural network using error back propagation technique to determine the parameters of neural network. Neural Networks is exploited because, prediction accuracy is generally high, robust, works when training examples contain errors, output may be discrete, real-valued, or a vector of several discrete or real-valued attributes, and fast evaluation of the learned target function, but the criticism is the long training time. Neural Networks learns by adjusting the weights so as to be able to correctly classify the training data and hence, after testing phase, to classify unknown data. The algorithm is as follows.

Neural Network Classifier

For this Neural Network algorithm, the input is given as N values $x_1, \ldots, x_N$ that form a feature vector x to obtain an output vector z that designates the class of the input vector x.

Step 1: A set $\{(x(q), t(q)) : q = 1,...,Q\}$ of exemplar vector pairs are set inputs.

Step 2: For each input values a weight is presented.

Step 3: Add all the updated input values with weights all together using a summation Function, $E = \sum q=1,Q \|z(q) - t(q)\|2$

Step 4: Here z(q) is the computed output for the input vector x(q), and t(q) is the target output (label) to which x(q) is supposed to map.

These two built-in algorithms are executed in MATLAB and results are explicitly exposed in next section sumptuously.

RESULT ANALYSIS

This section bounces the results produced by Decision Tree and Neural Networks on Banana image in MATLAB. A comparative analysis of Decision Tree and Neural Networks is also evaluated on Banana images. First GLCM features are extraction and then Decision Tree and Neural Networks have been executed in MATLAB on the real French Plantain images. Totally 270 images with 3 values such as mid ripe, ripe and over ripe are used for experiments. The steps involved using an image in image processing is exposed in Figure 2, Figure 3, Figure 4 and Figure 5.

Figure 2: Original image

Figure 3: After background removed

Figure 4: Gray scale image

Figure 5: Segmented image

The dataset is divided into training images and testing images. Here 240 images are used in training phase among these 80 images are mid ripe, 80 images are ripe and 80 images are over ripe. Also 30 images are used in testing phase whereas 10 images are mid ripe, 10 images are ripe and 10 images are overripe. Results are publicized in Table 1 noticeably. The classification accuracy obtained by Decision Tree is 71% and Neural Networks is 75.67% respectively.

Table 1: Classification on Banana Images

Algorithm	Dataset	Instances	Number of instances predicted		
			Mid ripe	Ripe	Over ripe
Decision Tree	Training	240	80	80	80
	Testing	30	73	69	71
Neural Networks	Training	240	80	80	80
	Testing	30	71	79	77

From the table, to rewording in a husk, Neural Networks produces enhanced accuracy than Decision Tree. This research triumph over its aim up to now. In next section conclusion of the research furthermore hopeful tracks for future augmentations are also pinched at the end of this article.

CONCLUSION

Banana is an epitome fruit for all age group people because of its high nutritional value. It is available in all markets. But the selection of banana is done very prudently. Because quality of banana affects the eating quality and market price of the fruit. Accordingly quality recognition is the essential process in food industries. In this research images of banana are captured and classified into dissimilar groups by Decision Tree and Neural Networks using GLCM features. Quality is recognized by the ripeness of banana. Decision Tree and Neural Networks are

provoked in MATLAB. Decision Tree is obtained 71% accuracy whereas Neural Networks twisted 75.67% accuracy. It is evidently explained that Neural Networks achieves higher than Decision Tree. In near future, other classification algorithms will be used on the same banana images to evaluate the better accuracy.

REFERENCES

[1]. Anushya , "Efficiency of feature selection on banana classification", International Journal of Scientific Research in Multidisciplinary Studies, Volume 5, Issue 12, 2019.

[2]. Anuja Bhargava et.al., "Fruits and vegetables quality evaluation using computer vision: A review", Journal of King Saud University – Computer and Information Sciences, Elsevier, 2018.

[3]. Suresha M et.al., " Texture Features and Decision Trees based Vegetables Classification", International Journal of Computer Applications (0975 – 8878) on National Conference on Advanced Computing and Communications - NCACC, April 2012.

[4]. Yan Zhang et al.," Deep indicator for fine grained classification of banana's ripening stages", EURASIP Journal on Image and Video Processing, 2018.

[5]. Fatima.M.A.Mazen et.al., "Ripeness classification of bananas using Artificial Neural Networks", Arabian Journal for science and Engineering, Jan, 2019.

[6]. Jonatha et.al., " Citrus Fruit quality classification using support vector machine", International Journal of Advanced Engineering Research and Science, Vol:6, Issue:7, July, 2019.

[7]. D.Surya Prabha et.al., "Assessment of banana fruit maturity by image processing technique", Journal of Food science and Technology, 52 (3), 2015.

[8]. Dr.A.Anushya, "Herbal Leaves Image Clustering Via K-Means", International Journal for Research in Applied Science and Engineering Technology, Volume 8, Issue 1, Pp:282-285 January , 2020.

[9]. Dr. A. Anushya, "Genetic Exploration for Feature Selection", International Journal of Computer Science and Engineering, Volume:7, Issue :2, 2019.