

Image Processing and Machine Learning for Automated Fruit Grading System: A Technical Review

Rashmi Pandey
Uka Tarsadia University,
Bardoli, Gujarat

Sapan Naik
Assistant Professor
Uka Tarsadia University,
Bardoli, Gujarat

Roma Marfatia
Uka Tarsadia University,
Bardoli, Gujarat

ABSTRACT

In India, demand for various fruits and vegetables are increasing as population grows. Automation in agriculture plays a vital role in increasing the productivity and economical growth of the Country, therefore there is a need for automated system for accurate, fast and quality fruits determination. Researchers have developed numerous algorithms for quality grading and sorting of fruit. Color is most striking feature for identifying disease and maturity of the fruit. In this paper; efficient algorithms for color feature extraction are reviewed. Then after, various classification techniques are compared based on their merits and demerits. The objective of the paper is to provide introduction to machine learning and color based grading algorithms, its components and current work reported on an automatic fruit grading system.

Keywords

Fruit grading, Machine learning, Color feature extraction, Classification

1. INTRODUCTION

In India, 70% of the agricultural labor and common man depends on the agriculture. As the major source of national income is from agriculture, it becomes the backbone of Indian Economy [1]. India ranks first among the other countries in the world, in the production of milk, pulses, jute and jute-like fibers; second in cereal crops, cotton, vegetables and fruits production; and is one of the leading producer of spices and plantation crops as well as fisheries and poultry [2]. If the overall production is good then it will directly increase the annual income of the cultivators and ultimately the national income of the country. Therefore currently researchers are trying to develop innovative and automated methods using science and technology to increase the production of agricultural industry [1].

1.1 Image processing and machine vision in agriculture

There are numerous applications in agriculture where image processing has been used as an analyzing tool. These applications includes land/aerial remote sensing of crops, detection and recognition of pathological stress conditions, detection of disease in leaf, stem, fruits and vegetables, shape and color characterization of fruits, identification of affected area by disease and effect on productivity. In fact, quantification of the visual properties of horticultural products and plants can play an important role to improve and automate agricultural management tasks [3]. Some of the specific applications are included:

- In Pest management, image processing is an interactive distance diagnostic tool for the detection of disease, insects and other organisms [4].
- In Crop assessment, image processing can be used to detect Weeds are unwanted crops that compete with growing crops for water, light, nutrients and other spaces. Remote sensing is most promising technique to detect weeds in crops [5].
- For quality inspection and grading for horticultural products, food products, grain classification and quality evaluation. Horticultural products are apples, mangoes, banana, pomegranates, grapes, oranges, nuts, peaches, pears, tomatoes and potatoes [6].
- For Harvesting of fruits and vegetables image processing can be used [6].
- Identification of different diseases such as fungus, foliar leaf spot, and Alternaria leaf spot on cotton leaves [7].
- Estimation of plant nitrogen content using multi spectral image sensors [8].
- Real time object tracking using color and texture segmentation [9].

Brief summary of different fruit grading system based on various features are described in section II. We demonstrate the novel algorithms for color based fruit grading with experimental results of different fruits in Section III. Different machine learning with its merit and demerits are described in section VI. We summarize our review and conclude this paper in Section VI.

2. FRUIT GRADING SYSTEM

Image processing has been widely used for grading of fruits into uniform categories (size, shape, color and texture, bruises, stem and calyx). Grading is being applied to many fruits and vegetables including apples, oranges, tomatoes, potatoes, carrots, green paper, peaches. In this section review is made on how different parameters can be used for automatic fruit grading system.

F.Pla and F.Juste presented thinning-based algorithm [11] to characterize fruit stems from profile images. In first iterations of thinning process, thinning algorithms makes skeletons of protrusions. A modified thinning method can detect and identify size and length of fruit stems or any other protrusion from binary patterns. Once the stem, or any protrusion from the fruit body, has been characterized and detected, it has to be located in order to supply this information to the de-stemming mechanism. The method has been applied to characterize a wide range of fruit stem shapes. 99% of stems

were correctly classified, therefore high accuracy was achieved.

To grade San-Fuji apples, neural network plays significant role to classify apples into five quality classes [12]. Pixels are obtained from image based on color and are given to the neural network as an input. Mean color of fruit is obtained from fruit image and grading is done based on fruit mean color, its variability and availability of diseased pixels and ratio of red color of fruit image. Accuracy is obtained up to 95% based on global classification rate (70%).

In [15], V.Leemans has graded apples into four classes according to European standards two varieties were tested: Golden Delicious and Jonagold. The image database included more than a 1000 images of fruits (528 Golden Delicious, 642 Jonagold) belonging to the three acceptable categories {Extra, I and II} and the reject (each class represents 60, 10 and 20% of the sample size). The image grading was achieved in six steps: image acquisition; ground colour classification; defect segmentation; calyx and stem recognition; defects characterization and finally the fruit classification into quality classes. The proposed method for apple external quality grading showed correct classification rates up to 72 to 78%, for Golden Delicious and Jonagold apples, respectively. The healthy fruit were better graded and an error rate decreases to 5 and 10%, respectively.

Automatic apple blemish detection system [25] has developed with Evolutionary Constructed (ECO) features for distinguishing bruises and blemishes from the stem end and calyx of apple images acquired in near-infrared spectrum. Iteratively transformations are applied to convert the input apple images into high quality features. 94% accuracy was achieved to detect true bruises and blemishes on apples.

In grading of Iyokan Orange fruit, a new method was proposed by Naoshi Kondo, Usman Ahmad, Mitsuji Monta, and Haruhiko Murase. From Iyokan Orange fruit image different parameters (roughness, color, and shape) were extracted. Other parameters like R.G color ratio, feret's diameter ratio and texture were also extracted. All parameters and weight of the fruit were given as a input to neural networks. Output of neural networks was evaluated based on reference value of sugar content or pH of the fruit.

A new method had proposed in [14] for recognition of fruit shape using concepts of attractor and fractal dimension. Neural network was used for quantitative measure and evaluation of the irregularity (or regularity) of the fruit shape. Proposed method is more reliable and more sophisticated for automated classification.

In [16], a new Fruit recognition system has been proposed, which combines color, shape and size in order to increase accuracy of recognition. Classification and recognition of fruit images are done based on features values obtained using nearest neighbors classification. Fruit recognition system classifies fruits accurately up to 90%.

Nur Badariah Ahmad Mustafa, Syed Khaleel Ahmed, Zaipatimah Ali, Wong Bing Yit, Aidil Azwin Zainul Abidin,

Zainul Abidin Md Sharif presented sorting and grading using Support Vector Machines and Fuzzy Logic [17]. The agricultural produce is classified based on fruit shape and size and grading. Accuracy obtained was very good.

Hongshe Dang, Jinguo Song, Qin Guo presented algorithm [18] for Fruit Size Detection and Grading System. Embedded grading system is used to implement the algorithm. Advantages of the algorithm are high accuracy of grading, high speed and low cost.

In [19], Quality grading of Coffee-Excelso Beans have been done using color, shape and size features that help us to find the green bean quality from a small sample.

In [20] paper, M. Khojastehnazhand, M. Omid, and A. Tabatabaefar, presented a novel approach for development of a sorting system for grading lemon based on color and size. The volume and color of fruit is determined and saved in database during calibration stage. During sorting stage, the calculated color and volume is compared with the saved information in the database. The final grade of fruit was determined. Accuracy achieved was 94.04%.

The system proposed in [21] can efficiently obtain the shape characteristic by drawing the lines and then class with K-means clustering method for the strawberry image. The proposed automated strawberry grading system uses $L^*a^*b^*$ color model. The size of fruit is decided based on the largest fruit diameter and color of strawberry is extracted using Dominant color method on a^* channel. The results show that the strawberry size detection error is not more than 5%, the colour grading accuracy is 88.8%, and the shape classification accuracy is above 90%. The average time to grade one strawberry is below 3s.

Yousef Al Ohali presented Computer vision based date fruit grading system [22]. In grading process of date fruit, Extracted features are flabbiness, size, shape, intensity and defects. Based on the extracted features it classifies dates into three quality categories (grades 1, 2 and 3) defined by experts. Date fruit having good shape, large size, high intensity, high flabbiness and no defects were branded as of the best quality. Author had used back propagation neural network (BPNN) as a classifier. Proposed system can sort 80% dates accurately.

Based on color and texture, a fruit recognition system was developed by S.Arivazhagan, R.Newlin Shebiah, S.Selva Nidhyanandhan, L.Ganesan [23]. Fusion of color and texture is proposed in this paper. Different parameters are intensity, color, shape and texture which are useful for grading of fruit. Minimum distance classifier is used for classification of fruit based on Wavelet transformed sub-bands.

In [24], Tajul Rosli B. Razak, Mahmud B. Othman, Mohd Nazari bin Abu Bakar, Khairul Adilah bt Ahmad, Ab Razak Mansor had presented automated mango grading system using fuzzy image analysis. In proposed system, the process of mango grading had done based on size, color and skin feature.

Table I gives brief summary of different fruits with different parameters and their accuracy.

Table 1. SUMMARY OF DIFFERENT FRUITS AND THEIR PARAMETERS

| Types of fruits | Parameters | Accuracy | Research Reference |
|-----------------|--|---------------------------|----------------------------------|
| Apples | Color | 95% for San-Fuji | Kazuhiro Nakano,1997 |
| | Color, stem and calyx, defect | 78 % for Golden Delicious | V. Leemans, 2002 |
| | | 72 % for Jonagold | |
| | Color, shape, texture | 72 % for Jonagold | V. Leemans, 2004 |
| | Bruises , Stem end and calyx | 89% | Xu Qiabao, 2009 |
| 94% | | Dong Zhang, 2013 | |
| Iyokan Orange | Color, shape, roughness | - | Naoshi Kondo, 2000 |
| Orange | Color | 70% | P. Levi,1988 |
| | Shape | 75% | D.Slaughter,1987 |
| | Intensity and color | 80% | F. Juste,1991 |
| Fruit* | Stem | 99% | F.Pla and F.Juste,1995 |
| | Shape | - | T. Morimoto, 2000 |
| | Color, shape, size | 90% | Woo Chaw Seng,2009 |
| | Shape, size | 90% | Nur Badariah Ahmad Mustafa, 2009 |
| | Size | High | Hongshe Dang, 2010 |
| | Color, texture | - | S.Arivazhagan,2010 |
| Coffee-Excelso | Color, shape, size | 90% | Eduardo Carrillo, 2009 |
| Lemon | Color, size | 94.04% | M. Khojastehnazhand, 2010 |
| Strawberry | Size, shape and color | 88.8% | Xu Liming,2010 |
| Date | Flabbiness, size, shape, intensity and defects | 80% | Yousef Al Ohali,2011 |
| | Color | More than 87% | Dah-Jye Lee,2011 |
| Tomatoes | Shape | 68% | Whitaker,1987 |
| | | 84%,87.5% | Md.Rokunuzzaman,2013 |
| | Color | 90% | F. Buemi ,1995 |
| | | 95% | Dah-Jye Lee,2011 |
| Mango | Size, color and skin feature | More than 80% | Tajul Rosli B. Razak,2012 |
| | Color and FD | 85.19%,88.89% | Hong Zheng,2012 |

*: Generalized algorithm for all fruits FD: Fractal dimension

3. COLOR FEATURE EXTRACTION AND ALGORITHMS

Most of the grading system depends on the color extracted from the image. Color is most visually striking feature of any image. Therefore, color features extraction plays an important role in developing grading system and also to identify defective fruits from normal fruits. In this section color feature extraction techniques of fruit grading system are explained.

Guo Feng and Cao Qixin presented in [36], an intelligent fruit sorting system which was based on color image processing. Fruit area was first segmented from an image using OHTA color space and blob extraction was applied to detect fruit contour then color ratio was calculated using HSI color space which act as classification feature. At last bayes classifier is applied for fruit sorting. Performance of proposed system is stable and satisfying. Average accuracy for classification is 90%.

The algorithm successfully estimate lemon based on color and size in [20]. RGB value of image was determined and HSI value was calculated from RGB. Color of fruit is determined by average value of hue component and stored in database. By comparing the information during sorting phase with available information stored inside database, grading of lemon was determined with 94.04 % accuracy.

Intelligent Color Vision System for Ripeness Classification of Oil Palm Fresh Fruit Bunch (FFB) was developed in [42]. Images of oil palm FFBs were collected and analyzed. Color feature was extracted from those images and used as the inputs for Artificial Neural Network (ANN) learning. The images were segmented into two parts, which were fruits area and spikes area. After the fruits area was obtained for each image, color features were extracted. Hue for each fruit pixel was calculated and a hue histogram representing the feature vector for each image was obtained. This feature vector represents the parameter for ripeness classifier. Multi layer Perceptron (MLP) managed to classify the ripeness of oil palm FFBs with 91.67% correct classification. In the second method, PCA was used to obtain a number of principal components that represented the new features to be fed into MLP. By using only 6 features, MLP managed to correctly classify the ripeness of oil palm FFB with 93.33%..

P. Sudhakara Rao and S. Renganathan presented algorithm [39] for automatic sorting and grading of apple. The Apple fruit images in different dimension were captured to cover the entire fruit. By means of selective hue component elimination, the background was totally removed and three apples image were extracted. The retained image data is used for extracting information like size, shape, color etc. Image is then converted into binary by threshold segmentation and then edge detection was applied. Finally, contour was analyzed for sorting of apples.

Comparative analysis of five different color spaces RGB, XYZ, HSI, $L^*a^*b^*$ and Lu^*v^* were examined by J. Blasco, N. Aleixos, J. Gómez, and E. Moltó in [40] for sorting of citrus fruits. Linear discriminant analysis (LDA) was used to obtain classification functions in each colour space, using the color coordinates. The highest classification accuracy achieved was 87% by using HSI color space.

An automated machine vision system was developed by M.Z. Abdullah, J.Mohamad-Saleh, A.S.FathinulSyahir, and B.M.N.Mohd Azemi in [41]. In proposed system, color and shape was considered. Artificial classifiers such as linear

discrimination analysis and multi-layer perceptron neural network were used to detect star fruit maturities. Images of star fruits were collected and background was removed from image to extract star fruit. RGB component of star fruit image was transformed in HSI and hue component was used for classification. Wilks lamda method was used for feature reduction. Average classification rate achieved was more than 95%.

A new fruit recognition system has been proposed in [16]. Area of the fruit region was excluded from background region of the image in order to extract the correct color feature values of the fruit. Mean of colour was calculated from RGB component. After calculating color of fruit, shape is measured by threshold segmentation. From geometrical properties (Area, perimeter, diameter, etc) of fruit, size is calculated and KNN was applied for classification. Recognition rate achieved was near about 90%.

In [24], Tajul Rosli B. Razak, Mahmud B. Othman, Mohd Nazari bin Abu Bakar, Khairul Adilah bt Ahmad, Ab Razak Mansor had presented automated mango grading system using fuzzy image analysis. In proposed system, the process of mango grading had done based on size, color and skin feature. Firstly size of mango is determined by calculating area of image object. Secondly RGB component was extracted from image and mean of three colors was determined. Then edge detection algorithm was applied for shape analysis. Fuzzy inference rule is applied for grading mango into different classes. An advantage of this method is the use of fuzzy inference engine without depending on human experts.

The system is proposed in [21] which can efficiently obtain the shape characteristic by drawing the lines and then class with K-means clustering method for the strawberry image. The colour of the strawberry is determined using the Dominant Color method on a^* channel. Gray value of window is calculated from left to right in order to obtain major color window. Average color value of dominant color window is calculated. Size is described by the largest fruit diameter. The results show that the strawberry size detection error is not more than 5%, the color grading accuracy is 88.8%, and the shape classification accuracy is above 90%. The average time to grade one strawberry is below 3s.

Hong Zheng, Hongfei Lu presented algorithm in [35] for detection of browning degree of mango based on least-square support vector machine (LS-SVM) as a function of fractal analysis and CIELab parameters. Images of mango were collected and background was removed. Feature parameter including $L^*a^*b^*$ value and fractal dimension (FD) was extracted from images. The average $L^*a^*b^*$ values from image were obtained. LS-SVM classifier based on $FD + L^*a^*b^*$ gives 100% accuracy. Based on FD and $L^*a^*b^*$ values, classification accuracy obtained was 85.19 up to 88.89%

Changyong Li, Qixin Cao, and Feng Guo presented new dominant color histogram matching method for fruits classification [43]. Images of fruit were acquisitioned and image segmentation algorithm is used to extract fruit's image from background in OHTA color space. Color image was transformed from RGB to HSV color space. The dominant color histogram matching method was used for color grading. Grade judgment result was given by calculating and comparing the similarity between the inspected sample histogram and standard template histogram for each grade. Fruit samples will be assigned grades based on the similarity between inspected sample histogram and standard histogram

template. Experiment result showed histogram matching method has high accuracy up to 97.2% and is suitable for real-time application.

Computer vision based automatic classification system for apple fruit is introduced in [44]. The fruit area that is foreground of the image is separated from its background by threshold method, by putting some threshold to the value plane image from HSV plane images. Then the v-plane image is divided into several equal parts which are called as windows. The statistical features like mean and standard deviation are calculated. The window which contains only black part that is only background and very small part of foreground is removed from the window list. Then these features are fed to the nearest neighbor (NN) supervised classifier for fruit classification. The resultant will be binary classification (defected apple or non-defected apple). The apples which contain stem end / calyx are classified as defects in initial classification. These defected apples are again classified as defected apple or stem end / calyx apple which is a natural part of the apple. Proposed method is effective and efficient.

In [45], a new system was developed to acquire RGB images from ruby red grapefruits with normal and five common diseased peel conditions (i.e., canker, copper burn, greasy spot, melanose, and wind scar). Total 39 images samples were used. Texture features were determined from the transformed hue (H), saturation (S), and intensity (I) or using region-of-interest images using the color co-occurrence method for each fruit sample. Algorithms for selecting useful texture features were developed based on a stepwise discriminant analysis, and 14, 9, and 11 texture features were selected for three color combinations of HSI, HS, and I, respectively. Classification models were constructed using the reduced texture feature sets through a discriminant function based on a measure of the generalized squared distance. The model using 14 selected HSI texture features achieved the best classification accuracy (96.7%). Average classification accuracy and standard deviation were 96.0% and 2.3%, respectively. Proposed Model is robust for classifying new fruit samples.

A new approach for grading pineapples based on color is presented in [46]. Images of pineapples were collected and background was removed. RGB component was extracted from pineapple image. During the sorting stage, the color values obtained are saved in a database for training the Neural Network (NN). The test results show the highest accuracy achieved for grading pineapples is 75%.

Direct color mapping technique is used in [33] for automatic color grading of fruit. Proposed method uses preselected

colors to calculate a unique set of coefficients for color space conversion. The three-dimensional RGB color space is converted into a small set of color indices unique to the application. Tomato and date maturity evaluation and date surface defect detection are evaluated by proposed method. Proposed method is simple but effective. This new direct color mapping concept can be applied to a variety of color grading applications that requires easy setting and adjustment of color preferences.

The whole image data collecting system was developed by Kazuhiro Nakano in [12] for colour grading of apple. Apple image was extracted by removing background and average colour gradients ($\bar{R}, \bar{G}, \bar{B}$), variances (V_R, V_G, V_B) and chromatic coordinates (r, g, b) were calculated from three primary colors. Nine colour characteristic data was obtained from entire apple. Neural network was used for color classification. Average accuracy for grading achieved is 70.76%.

In [38], a method called organization feature parameters (OFP) was used in grading of apple. Colour RGB signals of apple were transformed to HSI for colour discrimination. The hue component were divided in to eight equal intervals C_i ($i=1..8$). Finally, 17 colour feature parameters ($\bar{R}, \bar{G}, \bar{B}, V_R, V_G, V_B, r, g, b, c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$) were obtained for the apple.

V. Leemans, H. Mageinb, M.F. Destain presented in [15] apple grading system based on European standards. Mean of colour was calculated from 1000 images of apple and estimated ground color area of golden delicious and Jonagold. Correct classification rates for golden delicious and Jonagold are 78% and 72%.

In paper [37], ANN was used to classify Empire and Golden Delicious apples based on surface characteristics of the apple images. Firstly textural and histogram features were extracted from the images at selected wavelengths. Then images of apples with surface characteristics (good, defect and stem/calyx) in the view were used in classification applications. Back propagation neural network (BPNN) using textual feature performed 93.8% success rate in recognizing empire apples and 89.7% for golden delicious apples. BPNN using histogram features were significantly lower than textual feature.

Table II summarizes color feature extraction technique for different fruits and vegetables.

Table 2. Color Feature Extraction Methods and their Accuracy

| Types of fruits | Method | Merits | Demerits | Accuracy | Research Reference |
|-----------------|---|---|--|--|----------------------------|
| Fruit* | Mean of colors from RGB color space | 1. System is robust enough and is also able to recognize any input of fruit images that are being captured in any conditions. | 1. Types of fruits considered in recognition process are limited. 2. K-nearest neighbor classifier is sensitive to noise. 3. Time consuming process to calculate distance from each pixel. | 90% | Woo Chaw Seng,2009 |
| Lemon | HSI color space | 1.Good capability of representing the colors of fruit 2.Useful for separating objects with different Colors 3. Algorithm is easily adaptable for grading/sorting different agricultural products such as cucumber and eggplant. | 1. Numerically unstable at low saturation. 2.Misclassification rate is 5.96% | 94.04% | M. Khojastehnazhand,2010 |
| Apple | RGB color space to Nine color characteristic data | 1. Judgment ratio for classification is up to 95%. | 1. Ability to recognize fruit in A-grade is poor. | 70.76% | Kazuhiro Nakano,1997 |
| | HSI color space to 17 color characteristic data | 1.OFPs method was more accurate than BP-ANN 2.High grade judgment ratios were achieved in the classification of two of four apple color grades, i.e., 'Extra', and 'Reject' | 1. OFPs method was poor than SVM for identification. 2.Very slow process to grade apple in class I and class II | - | Zou Xiaobo,2007 |
| | Mean of colors from RGB color space | 1.For healthy fruit, correct classification rate is 90%-95% 2.Time saving process based on individual defect recognition | 1. Need better segmentation technique for bruise detection 2. Misclassification rate is high as compared to other techniques. | 78% for golden delicious apples 72% for Jonagold apples | V. Leemans,2002 |
| | HSV color space | 1. Effective approach to classify apples in defected and non defected class. | 1. In Hue plane image and saturation plane image, the contrast between healthy skin and defected skin is very low. We get poor result or false result with apple fruit images. | - | Shivleela R Arlimatti,2012 |
| Citrus | RGB, XYZ, HSI, La*b* and Lu*v* | 1 .HSI has good capability for grading citrus fruit 2. No false detection was obtained using these technique in the conditions under which the tests were conducted. | 1. Method is unstable at low saturation | HSI : 87% | J. Blasco,2007 |
| Star fruit | HSI color space, Wilks lamda method | 1. Proposed method produce 100 % success rate in detecting star fruits in different categories | 1. Numerically unstable at low saturation 2. Classification is affected if lighting levels are varied | More than 95% | M.Z. Abdullah,2006 |
| Mango | Mean of colors from RGB color space | 1. Use of if-then rules allow humans to make a best choice in reasoning | 1. Selection of membership function is difficult. | More than 80% | Tajul Rosli B. Razak, 2012 |

| | | | | | |
|----------------------|--|---|---|---|-------------------|
| | | process to ensure the consistency of decision | | | |
| | Fractal analysis and CIELab parameters | 1. LS-SVM classifier based on FD + L*a*b* with 100% accuracy rate is better than two classifiers FD and L*a*b*. | - | L*a*b* :85.19% FD :88.89% L*a*b* + FD :100% | Hong Zheng, 2012 |
| Strawberry | Dominant color method | 1. Memory efficient and faster processing for strawberry grading. The average grading time for one strawberry is less than 3 s. 2. Proposed method have higher sorting precise rate. | 1. Algorithm is not generalized for different type of fruits. | 88.8% | Xu Liming,2010 |
| | Dominant color histogram matching method | 1. High accuracy and is suitable for real-time application. | 1. Algorithm is less sensitive to lighting variations | 97.2% | Changyong Li,2009 |
| Tomato | Direct color mapping | 1. The overall grading results based on both color index and consistency averages are consistent with human grading. 2. Simple ,Effective and high accuracy | 1. Algorithm is not generalized for different type of fruits. | 95% | Dah-Jye Lee,2010 |
| Ruby red grapefruits | RGB color space to HSI color space | 1. Proposed method is robust for classifying new fruit according to their peel conditions. | 1. Hue and saturation can be somewhat influenced by low lighting variations. | 96.0% | Dae Gwan Kim,2009 |
| Pine apple | RGB color space | 1. Proposed method is better for color display | 1. Method is not suitable for color image, due to the high correlation between R, G, and B component. 2. Need better preprocessing algorithm to increase accuracy of grading pineapple | 75% | J. I. Asnor,2013 |
| Date | Direct color mapping | 1. The overall grading results based on both color index and consistency averages are consistent with human grading. 2. Simple ,Effective and high accuracy | Algorithm is not generalized for different type of fruits. | More than 87% | Dah-Jye Lee,2010 |

*: Generalized algorithm for all fruits
MLP: Multi Layer Perceptron
FD: Fractal Dimension
PCA: Principal Component Analysis
OFP: Organization feature parameter

4. MACHINE LEARNING TECHNIQUES

For classification of fruits into quality categories, statistical and syntactical classifiers are used. Some of the classifiers are:

4.1 Linear Discriminant Classifier (LDC)

LDC is a statistical classifier, which assumes that data is linearly separable and tries to find a linear decision boundary separating the data into different classes [47].

Date fruit classification was performed with Linear Discriminant Analysis [48]. First, new feature are constructed from linear combination of original features. Then after, this new features are used to separate the set of features into their respective categories. A test for apple fruit classification using

LDC was performed by Bernard Gosselin and Devrim Unay [49]. It searches for a linear decision boundary that separates the feature space into two half-spaces by minimizing the criterion function.

$$g(x) = wx + w_0$$

4.2 Nearest neighbour classifier (k-NN)

K-NN is a statistical classifier that focuses on similarity (proximity) of samples measured by a distance metric. It assigns data to the most represented category within its closest k neighbors' [47]. The Fruit Recognition System using the KNN algorithm as a classifier to classify fruit based on mean color values, shape roundness value, area and perimeter values of the fruit is presented by Woo Chaw Seng and Seyed

Hadi Mirisaee [16]. In current work, fruit area and perimeter are being chosen to represent the fruit size. This fruit size feature is used to distinguish one type of fruit from another. Then after inserting training data, the system is ready for use. The Euclidean distance determines the distances between the features values of the test input fruit with stored training fruits; whereas KNN algorithm will find out the 'K' shortest distance or closest examples to the input fruit then assign the input fruit to the class where majoring of the 'K' closest examples is from. For Fruits Recognition System, the value chosen for 'K' is '1', which means classification of input fruit sample is based on where is the class of the closest fruit example is from.

4.3 Support Vector Machines (SVM)

SVM is one of the powerful classification algorithms that have shown state-of-the-art performance in different varieties of classification tasks [50]. SVM is a new method that is used for classification of both linear and nonlinear data [51]. SVM first nonlinearly maps data to a high-dimensional space by using kernel functions. Then after, in that high-dimensional space it tries to find the linear optimal hyper plane that separates data with maximum margin. Originally SVM was proposed for only 2-class problems, but for multi-class problem we can extend SVM using near-against-one or one-against-all strategies [47].

SVM was used as decision making process for weeds identification in [52]. SVM approach is used for making the decision whether particular area needs to be sprayed or not. The proposed system works in two stages. First is the off-line process, where training is performed with the set of cells requiring to be sprayed and not to be sprayed and also decision function is computed. Second is the on-line process, where decision making is performed for each new incoming cell, based on the decision function computed in off-line process.

The LS-SVM, which solves a set of linear equations instead of solving a quadratic programming problem, is used for the automatic detection of browning degree on mango fruits in [35]. Proper kernel function and optimum kernel parameters are of importance in LS-SVM classifier and hence, RBF kernel was used as the kernel function due to its effectiveness and speed in training process.

Support Vector Machine-Multiclass Forward Feature Selection (SVM-MFFS), Successive Projection Analysis (SPA) and Uninformative Variable Elimination (UVE) algorithms were used in [50] to select representative wavelengths. Selected wavelengths are passed as an input to classifier to identify different brands of sesame oil.

4.4 Artificial Neural Network (ANN)

Artificial neural networks can be regarded as an extension of many classification techniques. These networks are inspired by the concept of the biological nervous system, and have proved to be robust in dealing with the ambiguous data and the kind of problems that require the interpolation of large amounts of data [53].

Neural Network (NN) was used for color grading of Apple fruit by Kazuhiro Nakano [12]. The first neural network (A) is

used to classify whether the color of the surface is normal red or abnormal red. The second Neural Network (B) is used to grade the color of the apples into the final five quality categories (AA, A, B, C, D).

I.Kavdir and D.E. Guyer proposed a new method in which ANN was used to classify Empire and Golden Delicious apples based on surface characteristics of the apple images [37]. Firstly textural and histogram features were extracted from the images at selected wavelengths. Then images of apples with surface characteristics (good, defect and stem/calyx) in the view were used in classification applications with two cases; two-class classification, and five-class classification. Two-class classification includes a good apple group and a defective (or stem/calyx) apple group. Five-class classification includes all the defective (leaf roller, bruise and puncture on Empire, and bruise bitter pit and russet on Golden Delicious) and good apple groups (good tissue and stem/calyx views).

In [13], a neural network was used in order to calculate the sugar content and acid content of iyokan oranges. A neural network with four inputs nodes, one output node and one hidden layer was constructed. The input data were grouped into four categories and output data were grouped into five categories. From total 30 iyokan oranges, 25 oranges were used in training the neural network and remaining 5 were used for testing. Kalman filter learning model was used for training the neural network.

T. Morimoto, T. Takeuchi, H. Miyata, Y. Hashimoto proposed an algorithm, for evaluation of fruit shape using neural network [14]. To characterize fruit shape, one-dimensional profile data consisting of radii between the centroid of the fruit and sampling points on the fruit profile was used. The error back propagation learning algorithm was used for training the neural network.

4.5 Rule Based System (Fuzzy system)

Rule-based systems are also known as knowledge-based systems, where the knowledge bases are represented in the form of a set of rules. The rules are teased out of domain experts or based on sample learning [54]. Rule-based inspection procedure proposed is four step process. First step is Blob extraction from original object images. Second step is feature extraction. Third step is rule base construction. Fourth step is recognition, where rule base constructed can be used for recognizing and separating good and defective apples.

4.6 C4.5

Syntactical classifiers use a hierarchical perspective where it is assumed that patterns are formed using simpler sub-patterns. Decision trees are typical examples of syntactical classifiers [47]. Decision tree is a flow chart like tree structure, where each internal node represents a test on the attribute, branches of tree represents the outcome of the test performed, and each leaf node of tree holds the associated class label [51].

Table III summarizes merits and demerits of machine learning techniques.

Table 3. Merits and Demerits of Machine Learning Techniques

| Machine learning techniques | Merits | Demerits |
|-----------------------------|--|---|
| LDA | LDA gives better results in the case when the normality assumptions are fulfilled and are appropriate for the development of linear classification models. | 1. May over fit the data 2. LDA implicitly assumes Gaussian distribution of data. |
| KNN | 1. Simple implementation. 2. Classes don't have to be linearly separable. 3. Robust, few parameters to tune (distance metric and k). | 1. Sensitiveness to noisy or irrelevant data. 2. Testing procedure is time consuming because of calculation of distance to all known instances. |
| SVM | 1. SVM is well suited to work with high dimensional data. 2. Classification accuracy is more as compared to other conventional classification techniques. 3. SVM is robust enough, even when training samples have some distortion | 1. Selection of kernel function and kernel parameters for mapping original data into higher dimensional data is difficult. 2. Learning process can be time consuming |
| ANN | 1. Robust and user friendliness and can handle noisy data. 2. Well suited to analyze complex problem. | 1. Scalability problem. 2. Require large number training samples. 3. Requires more processing time |
| RBS | 1. Robust. 2. Not sensitive to the changing environment. Memory requirement is little. | Determination of membership function is a difficult task. |
| C4.5 | 1. Relatively faster learning speed. 2. Convertible to simple and easy to understand classification rules. 3. Comparable classification accuracy with other methods | 1. An induced tree may over fit the training data. 2. Too many branches may reflect anomalies due to noise or outliers. 3. Poor accuracy for unseen samples. |

RBS: Rule Based System (fuzzy system)
 SVM: Support vector machines
 KNN: K- Nearest neighbour
 LDA: Linear discriminant analysis
 ANN: Artificial Neural Network

5. CONCLUSION AND FUTURE WORK

The literature review revealed that lot of research has been done on fruit grading system using image processing and machine learning applications. Image processing systems are capable of replacing labour work for inspection of fruit grading. The major problem for tackling with complex task is inclusion of knowledge in automatic grading system. Knowledge may be included implicitly in the form of training such as neural networks or in the form of rules such as fuzzy rule based system. It may be in the form of 2-class or multi-class problem such as support vector machine. Classification accuracy varies between 75%-96%. Further accuracy can be increased by combining multiple classifiers, but it results in heavy increase of computation time. Applications of machine learning have been reviewed in this paper and merits and demerits are described based on application.

Several methods for grading and sorting of fruits based on feature parameters have been reviewed and experimental results are summarized. Some methods are at more advanced stage than others because each method is based on estimation of feature parameters. One of the color feature extraction technique fractal analysis and CIELAB parameters had proved its best with 100% accuracy. Other techniques like dominant color method, dominant histogram matching method and direct color mapping techniques has achieved accuracy between 85 to 97%, but scope of these methods are limited. Further improvements can be done with different types of fruit with different parameters in direction to achieve

high speed and high accuracy for sorting and grading of different types of fruits.

Effectiveness of method depends on correlation between measured feature parameter and quality factor. Earlier grading method based on color did not require any high speed electronic technologies. Their sorting speeds were also not too high, and the accuracy was highly dependent on how well the measured parameters correlate with the quality factors. The combination of new imaging acquisition and high-speed image processing techniques has provided new direction for researchers to develop many new and improved techniques for grading and sorting of fruits.

6. REFERENCES

- [1] Agriculture Economics and Importance of Agriculture in National Economy website [Online] <http://agriinfo.in/?page=topic&superid=9&topicid=185>.
- [2] The Economic Survey 2012-13, Agricultural and Processed Food Products Export Development Authority (APEDA), The Union Budget 2013-14, Press Releases, Media Reports website [Online] <http://www.ibef.org/industry/agriculture-india.aspx>
- [3] Cunha, J.B., "Application of image processing techniques in the characterization of plant leaves," *Industrial Electronics, 2003. ISIE'03. IEEE International Symposium on*, vol.1, no., pp.612, 616 vol. 1, 9-11 June 2003

- [4] B.D. Mahaman, M. Maliappis, H.C. Passam, A.B. Sideridis, V. Zorkadis, Y. Koumpouros, "Image processing for distance diagnosis in pest management," *Computers and Electronics in Agriculture*, pp. 121-131, April 2004.
- [5] D.W. Lamb and R.B. Brown, "Remote-Sensing and Mapping of Weeds in Crops", *J. agric. Engng Res.*, pp. 117-125, 27 September 2001.
- [6] Tadhg Brosnan and Da-Wen. Sun, "Inspection and grading of agricultural and food products by computer vision systems-a review", *Computers and Electronics in Agriculture*, pp. 193-213, 2002.
- [7] Mr. Viraj A. Gulhane, Dr. Ajay A. Gurjar, "Detection of Diseases on Cotton Leaves and Its Possible Diagnosis", *International Journal of Image Processing*, vol. 5, no. 5, pp. 590-598, 2011.
- [8] V. K. Tewari, Ashok Kumar Arudra, Satya Prakash Kumar, Vishal Pandey, Narendra Singh Chandel, "Estimation of plant nitrogen content using digital image processing," *International Commission of Agricultural and Biosystems Engineering*, vol. 15, no. 2, pp. 78-86, July 2013.
- [9] Ercan Ozyildiz, Nils Krahnst-over, Rajeev Sharma, "Adaptive texture and color segmentation for tracking moving objects", *Pattern recognition*, pp. 2013-2029, 2002.
- [10] Kanali, C., Murase, H., Honami, "Three-dimensional shape recognition using a chargesimulation method to process image features", *Journal of Agricultural Engineering Research*, pp. 195-208, 1998
- [11] F. Pla, F. Juste, "Thinning-based algorithm to characterize fruit stems from profile images", *Computers and Electronics in Agriculture*, vol. 13, pp. 301-314, 1995.
- [12] Kazuhiro Nakano, "Application of neural networks to the color grading of apples", *Computers and Electronics in Agriculture*, Elsevier, pp. 105-116, 1997.
- [13] Naoshi Kondo, Usman Ahmad, Mitsuji Monta, Haruhiko Murase, "Machine vision based quality evaluation of Iyokan orange fruit using neural networks", *Computers and Electronics in Agriculture*, vol. 29, pp. 135-147, 2000.
- [14] T. Morimoto, T. Takeuchi, H. Miyata, Y. Hashimoto, "Pattern recognition of fruit shape based on the concept of chaos and neural networks", *Computers and Electronics in Agriculture*, vol. 26, pp. 171-186, 2000.
- [15] V. Leemans, H. Mageinb, M.-F. Destain, "On-line Fruit Grading according to their External Quality using Machine Vision", *Journal of Automation and Emerging Technologies*, Belgium. Biosystems Engineering, pp. 397-404, 2002
- [16] Woo Chaw Seng and Seyed Hadi Mirisae, "A New Method for Fruits Recognition System," *Electrical Engineering and Informatics*, vol. 01, pp. 130-134, August 2009.
- [17] Mustafa, N.B.A.; Ahmed, S.K.; Ali, Z.; Yit, W.B.; Abidin, A.A.Z.; Sharrif, Z.A.M., "Agricultural produce Sorting and Grading using Support Vector Machines and Fuzzy Logic," *Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on*, vol., no., pp. 391, 396, 18-19 Nov. 2009
- [18] Hongshe Dang, Jinguo Song, Qin Guo, "A Fruit Size Detecting and Grading System Based on Image Processing", 2010 Second International Conference on Intelligent Human-Machine Systems and Cybernetics, vol. 2, pp. 83-86, August 2010.
- [19] Eduardo Carrillo and Alexander Aristizabal Penaloza, "Artificial vision to assure coffee-Excelso beans quality," in EATIS, Czech Republic, pp. 35, 2009
- [20] M. Khojastehnazhand, M. Omid, and A. Tabatabaeefar, "Development of a lemon sorting system based on color and size," *African Journal of Plant Science*, vol. 4(4), pp. 122-127, April 2010.
- [21] Xu Liming and Zhao Yanchao, "Automated strawberry grading system based on image processing," *Computers and Electronics in Agriculture*, vol. 71, no. Supplement 1, pp. S32-S39, April 2010.
- [22] Yousef Al Ohali, "Computer vision based date fruit grading system: Design and implementation," *Journal of King Saud University - Computer and Information Sciences*, vol. 23, no. 1, pp. 29-39, January 2011
- [23] S. Arivazhagan, R. Newlin Shebiah, S. Selva Nidhyandhan, L. Ganesan, "Fruit Recognition using Color and Texture Features", *Journal of Emerging Trends in Computing and Information Sciences*, VOL. 1, NO. 2, pp. 90-94, Oct 2010.
- [24] Tajul Rosli Bin Razak, Mahmod Bin Othman (DR), Mohd Nazari Bin Abu Bakar (DR), Khairul Adilah BT Ahmad, and AB. Razak Bin Mansor, "Mango Grading By Using Fuzzy Image Analysis," *In proceedings of International Conference on Agricultural, Environment and Biological Sciences*, Phuket, 2012.
- [25] Dong Zhang, Kirt D. Lillywhite, Dah-Jye Lee, Beau J. Tippetts, "Automated apple stem end and calyx detection using evolution-constructed features," *Journal of Food Engineering*, Volume 119, Issue 3, December 2013, Pages 411-418, ISSN 0260-8774.
- [26] V. Leemans and M.F. Destain, "A real-time grading method of apples based on features extracted from defects," *Journal of Food Engineering*, vol. 61, no. 1, pp. 83-89, January 2004.
- [27] Xu Qiabao, Zou Xiaobo, and Zhao Jiewen, "On-Line Detection of Defects on Fruit by Machinevision Systems Based on Three-Color-Cameras Systems," *Computer and Computing Technologies in Agriculture II*, vol. 3, pp. 2231-2238, 2009.
- [28] P. Levi, R. Falla, R. Pappalardo, "Image controlled robotics applied to citrus fruit harvesting. Procedures", ROVISEC-VII, Zurich, 1988.
- [29] D. Slaughter, R.C. Harel, "Color vision in robotic fruit harvesting", *Trans. ASAE* 30 (4) (1987) 1144, 1148.
- [30] F. Juste, F. Sevilla, "Citrus: A European project to study the robotic harvesting of oranges," in: Proc. 3rd Int. Symp. Fruit, Nut and Vegetable Harvesting Mechanization, Denmark, Sweden, Norway, 1991, pp. 331, 338.
- [31] Whitaker, Miles, Mitchell and Gaultney, "Fruit location in a partially occluded image," *Trans. ASAE* 30 (3) (1987) 591, 597.

- [32] F. Buemi, M. Massa, G. Sandini, AGROBOT: a robotic system for greenhouse operations, in: Proc. 4th Workshop on Robotics in Agriculture & the Food Industry, IARP, Toulouse, 1995, pp. 172,184.
- [33] Dah-Jye Lee, James K. Archibald, and Guangming Xiong, "Rapid Color Grading for Fruit Quality Evaluation Using Direct Color Mapping," IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, vol. 8, no. 2, pp. 292-302, November 2011
- [34] Md. Rokunuzzaman, H. P.W. Jayasuriya," Development of a low cost machine vision system for sorting of tomatoes", CIGR Journal, Vol. 15, No.1,pp. 173-180,2013.
- [35] Hong Zheng and Hongfei Lu, "A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.)," Computers and Electronics in Agriculture, vol. 83, pp. 47-51, January 2012.
- [36] Guo Feng and Cao Qixin," Study on Color Image Processing Based Intelligent Fruit Sorting System", Proceedings of the 5th World Congress on Intelligent Control and Automation, pp. 4802-4805, June 15-19, 2004.
- [37] I.Kavdir and D.E. Guyer, "Comparison of Artificial Neural Networks and Statistical Classifiers in Apple Sorting using Textural Features," Journal of Biosystems Engineering, pp. 331–344, November 2004.
- [38] Zou Xiaobo, Zhao Jiewen, and Li Yanxiao, "Apple color grading based on organization feature parameters," Pattern Recognition Letters, vol. 28, pp. 2046-2053, June 2007.
- [39] P. Sudhakara Rao and S. Renganathan," New Approaches for Size Determination of Apple Fruits for Automatic Sorting and Grading", Iranian journal of electrical and computer engineering, Vol. 1, No. 2, November, 2002.
- [40] J. Blasco, N. Aleixos, J. Gómez, and E. Moltó, "Citrus sorting by identification of the most common defects using multispectral computer vision," Journal of Food Engineering, vol. 83, no. 3, pp. 384-393, December 2007.
- [41] M.Z. Abdullah, J. Mohamad-Saleh, A.S. Fathinul-Syahir, B.M.N. Mohd Azemi,"Discrimination and classification of fresh-cut starfruits (*Averrhoa carambola* L.) using automated machine vision system", Journal of Food Engineering ,pp. 506–523,2006.
- [42] Norasyikin Fadilah, Junita Mohamad-Saleh, Zaini Abdul Halim, Haidi Ibrahim and Syed Salim Syed Ali ," Intelligent Color Vision System for Ripeness Classification of Oil Palm Fresh Fruit Bunch", pp. 14179-14195,2012.
- [43] Changyong Li, Qixin Cao, and Feng Guo," A method for color classification of fruits based on machine vision", WTOS 8, vol. 2, pp. 312-321, February, 2009.
- [44] Shivleela R Arlimatti," Window Based Method for Automatic Classification of Apple Fruit", International Journal of Engineering Research and Applications ,Vol. 2, Issue 4, pp.1010-1013, July-August 2012.
- [45] Dae Gwan Kim, Thomas F. Burks, Jianwei Qin, Duke M. Bulanon," Classification of grapefruit peel diseases using color texture feature analysis", International Journal of Agricultural and Biological Engineering ,Vol. 2, No.3,pp. 41-50,September,2009.
- [46] J. I. Asnor, S. Rosnah, Z. W. H. Wan, and H. A. B. Badrul," Pineapple Maturity Recognition Using RGB Extraction",World Academy of Science,Engineering and Technology,vol.78 .pp. 147-150,2013.
- [47] Bernard Gosselin, Olivier Kleynen, Vincent Leemans, Marie-France Destain and Olivier Debeir Devrim Unay, "Automatic grading of Bi-colored apples by multispectral machine vision," Computers and Electronics in Agriculture, November 2010.
- [48] Haiwei Dong and Nikolaos Mavridis Abdulhamid Haidar, "Image-Based Date Fruit Classification," International Congress on Ultra Modern Telecommunications and Control Systems, vol. IV, 2012.
- [49] [49] Bernard Gosselin and Devrim Unay, "Thresholding based segmentation and apple grading by machine vision".
- [50] Yifei Xu, Li Li , Xiaoli Li, Yong He Shuiguang Deng, "A feature-selection algorithm based on Support Vector Machine-Multiclass for hyperspectral visible spectral analysis," Journal of Food Engineering, May 2013.
- [51] Micheline Kamber Jiawei Han, Data Mining: Concepts and Techniques, 2nd ed.: Morgan Kaufmann.
- [52] Gonzalo Pajares, Xavier P. Burgos-Artizzu, Angela Ribeiro Alberto Tellaache, "A computer vision approach for weeds identification through Support Vector Machine," Applied Soft Computing, pp. 908–915, 2011.
- [53] D. S. Jayas, J. Paliwal, and N. S. Visen, "Multi-layer Neural Networks for Image Analysis of Agricultural Products," Silsoe Research Institute, 2000.
- [54] Zhiqing Wen and Yang Tao, "Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines," Expert Systems with Applications, pp. 307–313, 1999.