Machine Learning-Based Detection and Grading of Varieties of Apples and Mangoes

Anuja Bhargava and Atul Bansal

Abstract Computer vision is a consistent and advanced technique for image processing with the propitious outcome an enormous potential. A computer vision has been strongly adopted in the heterogeneous domain. It is also applied to the various domains of agriculture that improve the quality of automation, growth of the economy, and the productivity of the nation. Fruits and vegetable quality highly affects the evaluation of the quality and export market. Recently, automatic visual inspection becomes very important for grading of fruits applications. In this paper, multiple features with support vector machine classifier-based automatic detection and grading of apple and mango are done. Firstly preprocessing is done using histogram equalization to smooth the image. Then, fuzzy c-means clustering is used for segmenting the defected region. Secondly, the combination of statistical, textural, and geometrical features is used to extract the information. Finally, the detection and grading are done using the SVM classifier and achieve accuracy with 98.48 and 95.72%. The agriculture industry achieves the direction of research and support technology for the detection and grading of fruits using multiple features.

Keywords Apple · Mango · Fresh · Rotten · SVM · Features

1 Introduction

Fruit detection and quality grading always endure a hot topic in the agriculture research field. Traditionally, detection and grading are done by labor manually [1]. This will lead to scarcity of consistency and due to shortage of labor results in research for solutions automatically. One of the most difficult processes is to detect and grade fruits visually [2]. Computer vision and image processing are some of the most important techniques used for the identification of features in many agricultural

A. Bhargava (✉) · A. Bansal
Department of Electronics and Communication, GLA, University, Mathura 281406, India
e-mail: anuja1012@gmail.com

A. Bansal
e-mail: atul.bansal@gla.ac.in

V. Goyal et al. (eds.), *Proceedings of International Conference on Communication and Artificial Intelligence*, Lecture Notes in Networks and Systems 192,
https://doi.org/10.1007/978-981-33-6546-9_43

products [3, 4]. Different algorithms have been tested and developed to detect and grade fruits automatically [5]. Moradi et al. [6] present the detection of defects in apples using a fuzzy c-means algorithm and achieves 91.00% accuracy. Razak et al. [7] proposed digital fuzzy image processing, content predicted analysis, and statistical analysis to grade mango and achieves 80.00% accuracy. Ashok and Vinod [8] presented quality grading of apples with 83.33% accuracy for probabilistic neural networks. Nandi et al. [9] propose mango grading using a multi-attribute decision theory with a prediction accuracy of 96.00% accuracy. Nandi et al. [10] presented the maturity and quality detection of mangoes using fuzzy and achieve an 87.00% recognition rate. Sahu and Potdar [11] identify defect a maturity of mango fruit using Image Processing Toolbox. Jawale and Deshmukh [12] proposed real-time rotten apple detection using ANN. Naik and Patel [13] presented a fuzzy classifier to grade mango ($L * a * b *$ color space) that achieves 89.00% accuracy. Khan et al. [14] classify apple disease using a genetic algorithm and achieve a 98.10% recognition rate. Nosseir [15] presented the classification and identification of multiple rotten fruits using linear SVM and achieves 96.00% accuracy. Singh and Singh [16] presented a classification of good and rotten apples using texture features with a 98.90% recognition rate. Bhargava and Bansal [17, 18] proposed a quality grading and detection of multiple fruits by machine learning.

This study aims to analyze an algorithm for the detection and grading of fruit and also to identify the performance of fruit algorithms. Furthermore, this study aims to ease meaningful and reliable conclusions. To that end, the proposed method aims to develop an algorithm that detects fruit variety among apple and mangoes and grades under two categories: fresh and rotten. This algorithm is based on a combination of features to attain a successful recognition rate. This work is the first to present a method to detect a variety of fruits in the vision system of agriculture.

2 Proposed Methodology

The purpose of this presented approach is to build a system that detects the fruit then classifies the variety of apples and mangoes among multiple varieties and assesses the sorting into two groups, i.e., fresh and rotten. Figure 1 presents an outline of the proposed method for the detection and categorization of fruits. These steps are explained in the following section.

2.1 Image Acquisition

The algorithm proposed uses two distinct varieties of fruits: apple and mango with five varieties of each fruit. Each set consists of fresh and rotten fruits. The characteristics of the dataset used are shown in Table 1. Figure 2 shows different varieties of fruits used.

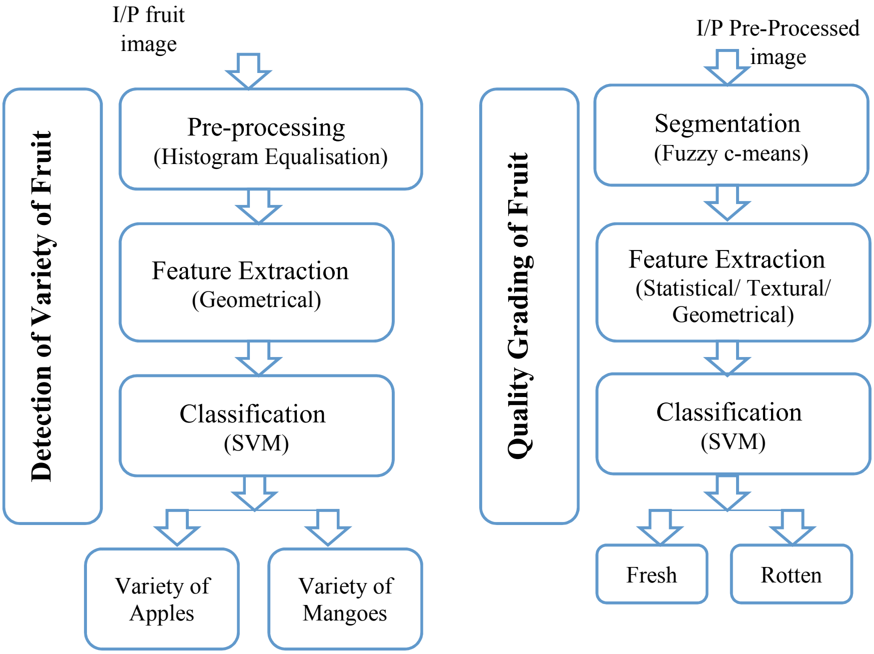


Fig. 1 Basic process for detection and grading of a variety of apples

Table 1 Dataset attributes

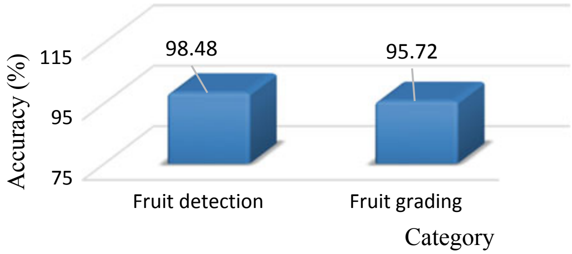
S. no.	Authors	Fruit variety	Quantity fresh fruit	Quantity rotten fruit
1	Blasco et al. [19]	Golden delicious	74	26
2	Unay and Gosselian [20]	Jonagold	1120	984
3	Blasco et al. [19]	Multigolden	97	80
4	Purdue Univ. [21]	Fuji	100	60
5	Purdue Univ. [21]	York	100	60
6	Naik [13]	Kesar	40	20
7	Naik [13]	Langdo	40	10
8	Naik [13]	Rajapuri	40	20
9	Naik [13]	Totapuri	40	25
10	Naik [13]	Madrashi aafush	30	30

2.2 Preprocessing

Image preprocessing is carried out before the actual analysis [22] to extract particular information. Preprocessing refers to data enhancement for the reduction of distortion

Fig. 2 Sample of a fresh and rotten database used **a** Golden delicious apple, **b** Kesar mango

Fig. 3 Fruit detection and grading accuracy



and noise illumination for the correction of degraded data. It also includes binarization, grayscale conversion, and filtering, smoothing, detection of edges, etc., used for the enhancement of the image.

In this paper, histogram equalization is done using grayscale imaging which results in the whitening of the image. Mathematically, the histogram equation is expressed as:

$$h(v) = \left(\frac{\text{CDF}(v) - \text{CDF}_{\min}}{(A \times B) - \text{CDF}_{\min}} X(L-1)\right) \quad (1)$$

where $\text{CDF}_{\min}$ is cumulative distribution function minimum value, $A \times B$ is some pixels in an image, L is the gray level of the image.

2.3 Segmentation

Image segmentation separates a particular set of pixels into multiple segments from the digital image. The object evaluation is done by separating the background area from the foreground image. Various segmentation methods are utilized such as Otsu segmentation, k-means clustering, fuzzy c-means clustering, and color segmentation. Among all segmentation methods, the important segmentation technique is color segmentation due to which the spot of every disease has a different color.

In this paper, fuzzy clustering is done because it is effective for segmenting images in a controlled environment. In this technique, data points are partitioned into a specific number of clusters. It minimizes the objective function fork cluster centroids and given partitioned fuzzy data, n [23].

Mathematically, the membership function,

$$F = \mu_{cd} = 1; \text{c belongs to d}$$

$$F = \mu_{cd} = 0; \text{c does not belong to d}$$

The condition to insure that set is exclusive and exhaustive:

$$\sum_{d=1}^{k} \mu_{cd} = 1, \quad 1 \le c \le n \tag{2}$$

$$\sum_{c=1}^{n} \mu_{cd} > 0, \quad 1 \le c \le k \tag{3}$$

$$\mu_{cd} \epsilon \{0,1\}, \quad 1 \le c \le n; \quad 1 \le c \le k \tag{4}$$

The generalize objective function is given as

$$J(F, Z) = \sum_{c=1}^{n} \sum_{d=1}^{k} \mu_{cd}^{\varphi} d_{cd}^{2} \tag{5}$$

where

k is number of clusters

n is data point number.

2.4 Feature Extraction

Segmentation develops separated pixels shapes with distinct sizes. The determination of the fruit category depends upon the pixels taken together or independently. Our experiment shows that 13 statistical (mean, RMS, variance, standard deviation, smoothness, skewness, inverse difference moment, kurtosis) and textural (contrast, correlation, energy, homogeneity, and entropy), 14 geometric features (area, eccentricity, major axis length, minor axis length, centroid, bounding box, eccentricity, orientation, convex hull, convex area, solidity, extrema, diameter, extent) [24] are used for best performance in the grading system.

Table 2 Detection and grading accuracy using SVM

S. no.	Category	Performance (%)			Execution time (s)
		Accuracy	Sensitivity	Specificity	
1	Fruit detection	98.48	97.41	99.38	75.34
2	Fruit grading	95.72	93.87	97.02	112.56

2.5 Classification

The extracted features from training images are the input to the classifier. The different characteristics of multiple varieties of apple and mango fruit are learned by the classifier. In this experiment, we use a statistical classifier known as SVM. It is a type of learning system which uses hypothesis space in higher dimensional space of linear function which implements statistical learning theory. It consists of two parts: linearly separable and nonlinear separable. “SVM is a supervised learning method that is based on the minimization procedure of structural risk” [25].

3 Results

The standards of commission European [26] for fruits describes one dismiss and three acceptable conditions. However, ample literature abides of fresh/rotten categorization because of the adversity of the compilation of database and sorting processes. A number of training and testing images used for detection and grading are 1780 and 1216, respectively, in SVM classifier. In the pursuit of detection and grading of apple and mango two category sorting, we have inspected 27 features while SVM is used for classification. The ten datasets of apples as mentioned in Table 1 have been trained with all features. Finally, the grading is done by SVM as shown in Table 2 and Fig. 3.

4 Conclusion and Future Work

In this research, a computer vision-based fruit detection and grading are introduced. Firstly preprocessing is done using histogram equalization to smooth the image. Then fuzzy c-means clustering is used for segmenting the defected region. Secondly, the combination of 27 features is extracted. Finally, the detection and grading are done using SVM classifiers and achieve accuracy with 98.48 and 95.72%. Furtherly, the more concluded and powerful system may be generated with enriched performance in the future. The solution proposed in this work must be tested for multiple fruit images to access its generality, robustness, and accuracy and must be established in a real-life detection and sorting domain.

References

1. Ashraf MA, Kondo N, Shiigi T (2011) Use of machine vision to sort tomato seedlings for grafting robot. Eng Agric Environ Food 4(4):119–125
2. Iraji MS, Tosinia A (2011) Classification tomatoes on machine vision with fuzzy the Mamdani inference, adaptive neuro fuzzy inference system based (Anfis-Sugeno). Aust J Basic Appl Sci 5(11):846–853
3. Zhang B, Huang Z, Li J, Zhao C, Fan S, Wu J, Liu C (2014) Principle, developments and applications of computer vision for external quality inspection of fruits and vegetables: a review. Food Res Int 326–343
4. Dubey SR, Jalal AS (2015) Application of image processing in fruits and vegetables analysis: a review. J Intell Syst 24(4):405–424
5. Bhargava A, Bansal A (2018) Fruits and vegetables quality evaluation using computer vision: a review. J King Saud Univ Comput Inform Sci. https://doi.org/10.1016/j.jksuci.2018.06.002
6. Moradi G, Shamsi M, Sedaaghi MH, Moradi S (2011) Apple defect detection using statistical histogram based fuzzy c-means algorithm. Iran Conf Mach Vision Image Process
7. Razak TRB, Othman MB, Bakar MNBA, Ahmad KAB, Mansor AB (2012) Mango grading by using fuzzy image analysis. In: International conference on agricultural, environment and biological sciences, pp 18–22
8. Ashok V, Vinod DS (2014) Automatic quality evaluation of fruits using probabilistic neural network approach. In: International conference on contemporary computing and informatics (IC3I). IEEE, pp 308–311
9. Nandi SC, Tudu B, Koley C (2014) Computer vision based mango fruit grading system. In: International conference on innovative engineering technologies
10. Nandi CS, Tudu B, Koley C (2016) A machine vision technique for grading of harvested mangoes based on maturity and quality. IEEE Sens J 16:6387–6396
11. Sahu D, Potdar RM (2017) Defect identification and maturity detection of mango fruits using image analysis. Am J Artif Intell 5–14
12. Jawale D, Deshmukh M (2017) Real time automatic bruise detection in (apple) fruits using thermal camera. In: International conference on communication and signal processing, pp 1080–1085
13. Naik S, Patel B (2017) Thermal imaging with fuzzy classifier for maturity and size based non-destructive Mango (*Mangifera Indica L.*) grading. In: International conference on emerging trends & innovation in ICT, pp 15–20
14. Khan MA, Lali MIU, Sharif M, Javed K, Aurangzeb K, Haider SI, Altamrah AS, Akram AT (2019) An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection. IEEE Access 7:46261–46277
15. Nosseir A (2019) Automatic classification for fruit types and identification of rotten ones using k-NN and SVM. Int J Online Biomed Eng 15(03)
16. Singh S, Singh NP (2019) Machine learning based classification of good and rotten apple. Recent Trends Commun Comput Electron 377–386
17. Bhargava A, Bansal A (2019) Quality evaluation of mono & bi-colored apples with computer vision and multispectral imaging. Multimed Tools Appl. https://doi.org/10.1007/s11042-019-08564-3
18. Bhargava A, Bansal A (2019) Automatic detection and grading of multiple fruits by machine learning. Food Anal Methods. https://doi.org/10.1007/s12161-019-01690-6
19. Blasco J, Aleixos N, Molto E (2003) Machine vision system for automatic quality grading of fruit. Biosys Eng 85(4):415–423
20. Unay, Gosselin (2005) Artificial neural network-based segmentation and apple grading by machine vision. In: International conference on image processing
21. Purdue University. https://engineering.purdue.edu/RVL/Database/IFW/database/index.html
22. Cubero S, Aleixos N, Molto E, Gomez-Sanchis J, Blasco J (2011) Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food Bioprocess Technol 487–504

23. Ashok V, Vinod DS (2014) Using K-means cluster and fuzzy C means for defect segmentation in fruits. Int J Comput Eng Technol 11–19
24. Ou X, Pan W, Xiao P (2014) Vivo skin capacitive imaging analysis by using grey level co-occurrence matrix (GLCM). Int J Pharm 460(2):28–32
25. Burges (1998) A tutorial on support vector machines for pattern recognition. Data Mining Knowl Discov 2:121–127
26. Anonymous (2004) Commission Regulation (EC) No 85/2004 of 15 January 2004 on marketing standards for apples. Off J Eur Union L 13:3–18