Multi-Parameter Based Mango Grading Using Image Processing and Machine Learning Techniques

AMRUTA SUPEKAR[1]
MADHURI WAKODE [2]

Department of Computer Engineering,
SCTR'S Pune Institute of Computer Technology,
Savitribai Phule Pune University Pune,
Maharashtra, India, 411043
[1]amruta.off@gmail.com
[2]mswakode@pict.edu

Abstract. Mangoes, the king of fruits is globally exported and locally consumed on a large scale. During exports and local marketing, delivering good quality fruits and satisfying certain pre-defined standards is important. This post-harvest operation of quality checking, known as mango grading is usually performed manually. But manual grading can be in-consistent, erroneous and labor-intensive. A computer vision based grading solution will result in consistent and accurate sorting. Such a mango grading system based on external parameters namely ripeness, size, shape, defects was developed in this research work. Image processing techniques were applied to extract the color, geometric and shape related features. These features were further utilized by pre-trained random forest classifiers to determine the mango ripeness (unripe/mid-ripe/ripe), size (small/medium/large) and shape (well-formed/deformed) category. K-means clustering was applied for defect segmentation to determine the mango defect category as (non-defective/mid-defective/completely-defective). Final grading was performed using a grading formula that combines the parameter specific quality scores assigned, according to predicted categories. Ripeness, size and shape classification performed on a created dataset of Dashehari mangoes achieved a test accuracy of 100%, 98.19% and 99.20% respectively. Formula based integrated grading could grade mangoes with 88.88% accuracy.

Keywords: computer vision, image processing, machine learning, mango grading, post-harvest technology

(Received April 16th, 2020 / Accepted December 1st, 2020)

1 Introduction

India is prominent producer and exporter of many fruits. The king of fruits, mango is one of the most important products exported by India to countries like U Arab Emts, U.K, Oman, Qatar and USA. There are about 1000 mango varieties produced in different states of India. Alphonso, Totapuri, Dashehari, Kesar, Pairi, Neelum are some of them [2]. Along with high export demands, mangoes are also consumed locally on a large scale. Such exported or local marketed mangoes are required to satisfy some preset quality expectations. A failure to meet such quality requirements can lead to huge economic losses, thereby reducing vendor economic gains. Hence it is required to sort mangoes according to grading parameters like size, shape, maturity, defects, firmness, nutrients etc. This post-harvest process is called mango grading where mangoes are classified into different pre-defined grades [1]. Proper grading can help to increase vendors/farmers confidence about his product and open up new export opportunities. Parameter specific sorting can help vendors to take appropriate transport and marketing related decisions. Eg Unripe mangoes can be exported to far-away places thereby reducing food wastage, slightly defective mangoes can be sold at reduced prices etc. Thus precise and accurate sorting/classification of mangoes is essential. However, this operation is usually performed manually which can be in-consistent, in-accurate, time consuming and labor intensive. In recent years, computer vision based solutions have been used for many agricultural practices like yield estimation, weed management, fruit/leaf disease identification [10, 29, 34] etc. Similar computer vision techniques can be applied to the crucial post-harvest grading process, leading to accurate, consistent and reliable grading. Many attempts have been made to perform image based mango grading, however, most of the researchers utilize single mango view [22, 24, 19, 7, 23, 17]. Analyzing multiple views of mango and then combining results for final grading would be a more reliable solution. Also very few research works have considered all four external grading parameters. [24]. In this research work, an attempt to grade Dashehari mangoes has been made. Four external grading parameters namely ripeness, size, shape and defects are considered. Two views (front and back) for each mango have been examined. The overall system consists of seven modules- image acquisition, image pre-processing and background removal, ripeness determination, size determination, shape determination, defect determination and grade determination. The ripeness, size and shape determination modules utilize image extracted features and perform classification using random forest classifiers. Defect determination module applies k-means clustering image segmentation approach for identifying mango defects. As far as the survey done, our attempt is the first one to utilize K-means clustering for defect segmentation. The final grading module utilises a formula based approach. Experimentation was carried out on a dataset of Dashehari mangoes. The current system applies to Dashehari mangoes, however, if relavant dataset is available, it can be applied for other varieties too. The developed grading solution can be incorporated in a grading machine with proper mechanical setup and be utilized by mango vendors/exporters. As far as survey done, this is first attempt to perform multi-parameter grading for Dashehari mangoes.

2 Related Work

Many researchers have made an attempt to perform mango classification using image processing and machine learning techniques. Many research works have performed single parameter classification while few have considered multiple parameters. Presented survey is divided into ripeness, size, shape, defect and multi-parameter classification.

Ripeness Classification: Mango ripeness can be closely associated with its peel color properties. As mango ripens its color changes from green to yellowish green to yellow. Such peel color variations can be analyzed using different color spaces like RGB, HSV, Lab, YIQ etc. Many researchers have explored the RGB color features [24, 16, 19, 23]. However, RGB color space being device dependent, sometimes does not provide good color based analysis [24, 16]. An attempt to classify Kent mangoes into two ripening stages achieved a recognition rate of 90% [16]. Accuracy of 84% was gained for shelf life prediction using support vector regression for 5 mango varieties [24]. Better classification results are obtained when other color spaces have been considered [22, 19, 35, 23, 17, 25, 7]. A five stage ripeness classification based on RGB color space achieved an error rate of 3.5% using Quadratic Discriminant Analysis. 114 mangoes were used for analysis [23]. RGB features combined with HSI color features gained an accuracy of 96% for six stage ripeness classification, performed on a test set of 25 mangoes [19]. A 100% accuracy was gained while performing a detailed analysis of ripeness using CIEL*a*b*, HSB color and physicochemical features [35]. The correlation of near infrared spectral data with ripeness was studied. Spectral data was used for four stage ripeness classification which gained an accuracy of more than 80% [30].An attempt to correlate specific gravity with ripeness was made by [13]. Specific gravity was determined using image estimated volume and its comparison with actual gravity determined by Archimedes's principle was made. [13].

Size based classification: Mango size can be closely correlated with its area. Many size based classification works performed by researchers have utilized area as a primary feature [22, 4, 26, 20, 28]. Few researchers perform size classification based on image extracted mango length and width called major and minor axis [24, 6, 25]. Classification accuracy of 87.5% and 89.5% was obtained when area was used as the only feature in [4] and [28] respectively. Major and minor axis when employed for size based analysis have gained good recognition rate [24, 25]. An attempt to predict actual length and width of mango has been made in [24] using image extracted major-minor axis. Only 3% error between actual and predicted length was obtained [24]. A size based fuzzy classification of 517 mangoes into 4 classes achieved an accuracy of 96.58% [25]. An attempt to estimate mango volume and weight from image, using disk method was performed. Size classification into 3 classes based on estimated mass achieved an accuracy of 94% [12]. An attempt to perform size based classification using deep learning approach was made in [21]. Features extracted from CNN when given as input to SVM classifier, achieved best accuracy of 72.46% by MobileNet [21].

Defect based classification: Mango defects can be identified from its image using image segmentation techniques. Most of the research works have utilized fixed thresholding approach for defect segmentation [11, 26, 24, 27, 16, 7, 32]. Various channels from different color spaces like YCbCr, Lab, RGB, HSL, HSI etc have been explored for thresholding. Based on defect area, mangoes are classified as defective and non-defective. The b channel of Lab color space was used for defect segmentation in [26] and [25]. It gained an accuracy of 93.33% and 96.95% respectively. Defect segmentation based on S (saturation) channel thresholding of HSI color space achieved classification accuracy of 90.66% [7]. It was observed that defects were evident in blue component of RGB color space. Hence R-B and G-B values were utilized for defect identification which achieved an accuracy of 90% [24]. The H (hue) component of HSL color space was utilized for defect identification. A four class defect classification of 180 mangoes gained an accuracy of 88.6% [27]. An attempt to perform neural network based defect classification for oranges, using color and texture features is done in [9] and [33]. Five type defect classification of 400 oranges achieved an error rate of 2.75%. Shape features were also utilized [9]. Two category defect classification achieved an accuracy of 94.5% [33].

Shape based classification: Shape related features like eccentricity, cross-ratio, extent and Fourier descriptors have been explored for shape based classification [31, 22, 24, 12]. However, Fourier descriptors have resulted in best shape recognition. Centroid based Fourier descriptors were extracted and SVM was used as machine learning classifier in [24] and [31]. Shape based classification of 300 mangoes achieved an accuracy of 100% [31]. Shape classification of 200 mangoes belonging to 5 varieties achieved an accuracy of 91% [24].

Multi-parameter grading: Few research works have attempted grading/sorting based on multiple parameters. A size, shape and ripeness based classification into 4 classes using deep learning approach was performed in [21]. Best accuracy of 83.97% was obtained by MobileNet for dataset of 2432 images [21]. SVM classification on Kent dataset [8], based on geometric, textural and histogram features into three classes, gained an accuracy of 97% [15]. Grading of 200 mangoes (5 varieties) based on maturity, size, shape and defects has been performed in [24]. Machine learning techniques like support vector regression, SVM, MADM (multi attribute decision making and fuzzy classification were utilized. Four grade classification gained an accuracy of 88% [24]. Shape, size and maturity based three category grading of 900 mango images using fuzzy and decision making techniques has achieved an accuracy of 90% [22]. Similar 3 class grading was also performed in [25]. Size and color properties were utilized for classifying mangoes into three grades namely export, class1 and class2 in [5]. Maturity and size based classification into 4 grades using artificial neural network has gained an accuracy of 94% [36].

3 Dashehari Mango Dataset

In this research, dataset for Dashehari mangoes was prepared. Dashehari is one of the Indian mango variety grown in states like Bihar, Gujrat, Haryana, Himachal Pradesh etc. [2]. Mangoes at unripe stage were purchased from a local vendor and images were captured at different ripening stages using Iphone 6s plus, 12 megapixel camera. Thus in total, images of 85 mangoes at different ripening stages were acquired. Two views, front and back for each mango were captured which resulted in 170 images. Out of the 85 mangoes, 16 test mangoes were selected for grading. Images of remaining 69 mangoes were utilised for training purpose. Data augmentation with three rotations namely 90, 180. and 270. was applied to 69 mango images(front and back) to create 552 images. The final training dataset therefore consists of 552 mango images corresponding to 276 mangoes. Some of the images from created dataset are shown in Figure 1. The front and back view of a defective mango is shown in Figure 2. It can be observed that one view is completely defective as opposed to other, hence inspecting both views for grading is important. Mango weight for all 85 mangoes was also measured using a weighing machine.

4 Methodology

The proposed mango grading methodology is depicted in Figure. 3. The overall grading system consists of seven modules:

4.1 Image Acquisition

Images were acquired using Iphone 6s plus, 12 megapixel camera. Images were acquired during day-light and two views(front and back) of each mango were captured. In order to facilitate the background removal, a

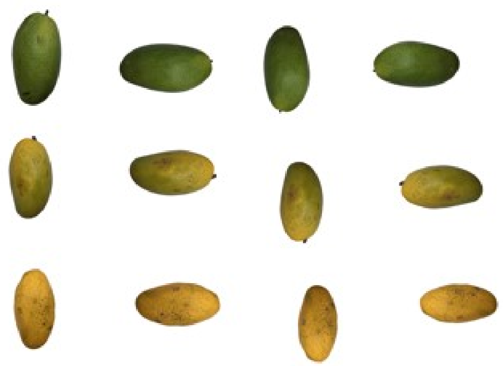


Figure 1: Dashehari Dataset Sample Images

Figure 2: Front and Back View of Mango

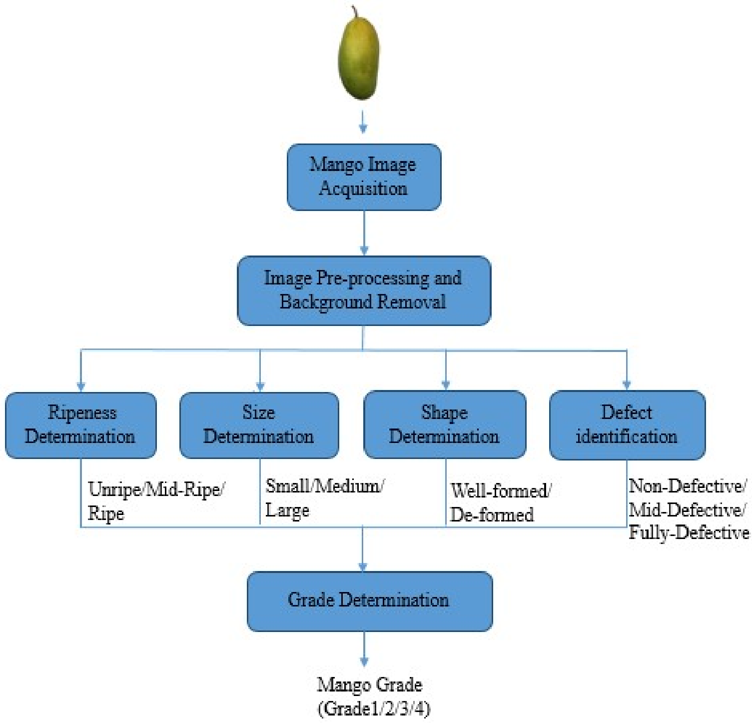


Figure 3: Proposed Mango Grading Methodology

white background was used during image acquisition. Each image is a digital colored image, 4032 * 3024 in size. Distance of each mango from camera was kept constant.

4.2 Image Preprocessing and Background Removal

Preprocessing operations like image resizing, median filtering and RGB to gray scale conversion were applied on acquired image. The acquired image was resized to half its original size. This helps in speeding up further image processing operations. Median filtering helps in removing the salt and pepper noise present in acquired image. The image background being white in color, mango segmentation was achieved using simple thresholding applied on the gray-scale image. Output of thresholding operation is a binary image where mango area is white and background is black in color. This binary image was anded with original colored image to obtain a mango segmented image. Background removal process is depicted in Figure. 4

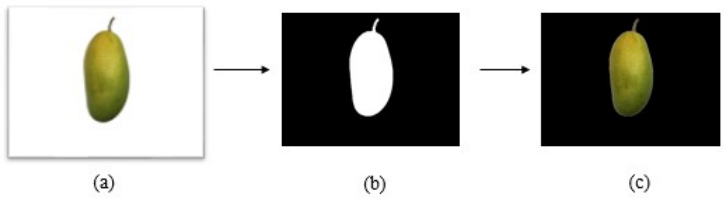


Figure 4: Mango Segmentation Steps (a-original image, b-binary image after thresholding, c-mango segmented image)

4.3 Ripeness Determination

As mentioned earlier, ripeness can be closely associated with mango peel color. Initially all color components related to RGB, HSV and CIEL*a*b* color spaces were extracted. Recursive feature elimination was applied for feature selection, which selected mean H(Hue), R(Red) and a* as important features. The features were first tried individually for ripeness classification, however it was observed that red and a* together could achieve a perfect ripeness classification. Thus mean R and mean a* were selected as final color properties. Color features from both the mango views were combined to determine mango ripeness category as unripe, midripe or ripe. A two length feature vector [mean R, mean a*] was formed. During training phase such feature vectors were combined to create a color feature CSV dataset, which was then used for training a random forest classifier. During testing phase, the extracted feature vector was utilized by pre-trained random forest model to predict mango ripeness category. The ripeness determination training and testing phases are depicted in Figure. 5.

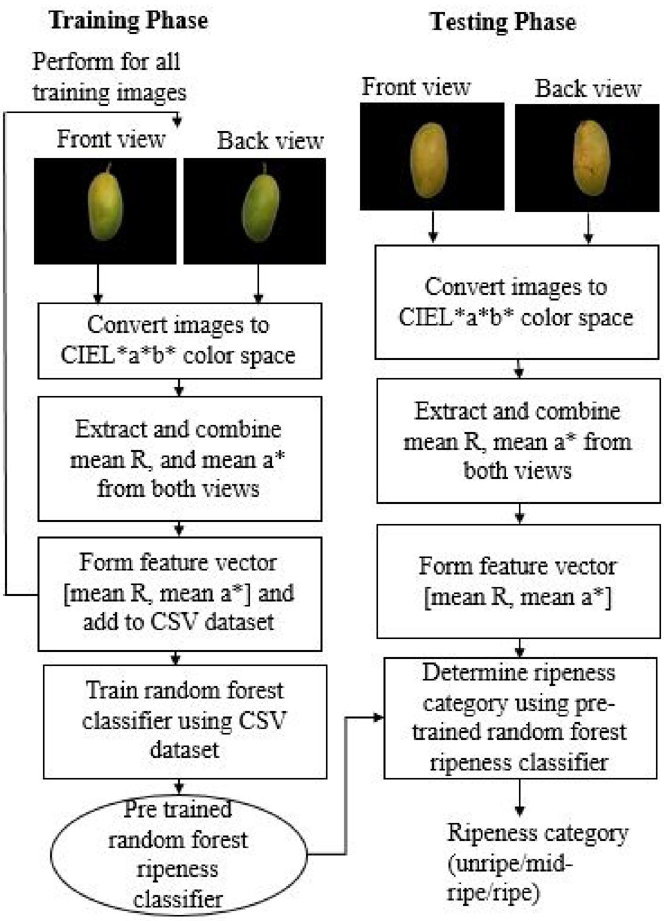


Figure 5: Ripeness Determination: Training and Testing Phases

4.3.1 Random Forest Classifier

Random forest is a parallel ensemble technique which uses decision trees as base classifier. It consists of many decision trees which differ from each other due to bagging and feature randomness. The final prediction is determined using majority voting. Random forests are known to perform very well when compared to other machine learning classifiers[14, 18, 3]. In this work, when training the classifiers, performance of four machine learning supervised algorithms namely support vector machines, neural networks, k-nearest neighbour and random forests was compared and random forest performed better. Thus it was utilized.

4.4 Size Determination

With respect to size, mango was classified as small, medium or large. According to inputs taken from mango exporters, mangoes are usually graded by their weight. Hence in this research work, initially, we have labelled mangoes according to weight. Then an attempt to classify mangoes from image extracted, geometric features was made. Mango shape is similar to an ellipse. Thus contour detection and elliptical fitting operations were performed to determine the mango major and minor axis. Mango area was calculated from the binary image obtained during segmentation. Image extracted major axis, minor axis and area formed the feature vector for size classification. Similar to ripeness determination module, a random forest classifier was trained and utilized to determine mango size category. Size determination module steps are depicted in Figure. 6. Ellipse fitting is depicted in Figure. 7.

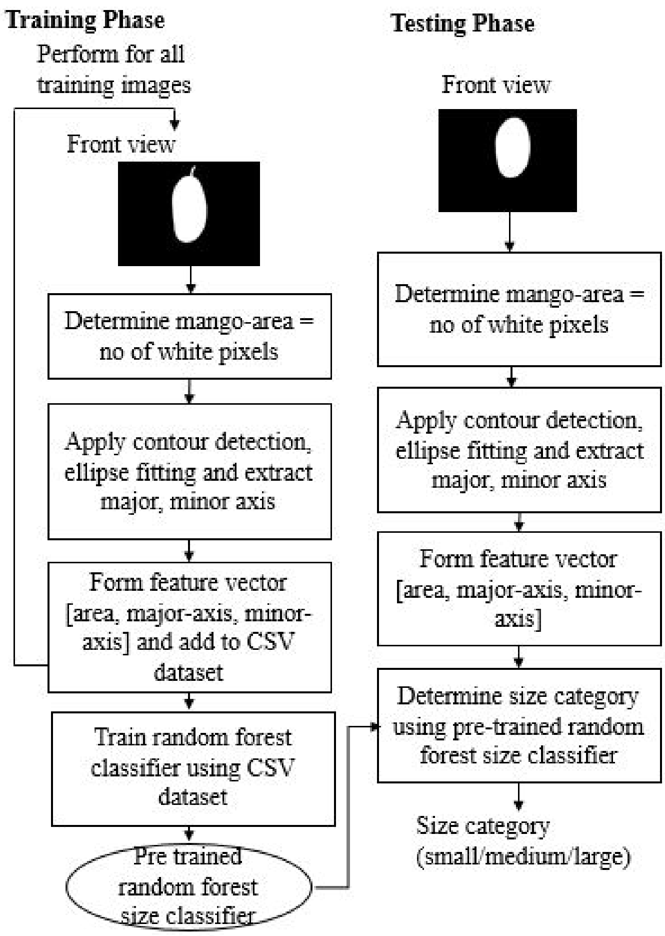


Figure 6: Size Determination: Training and Testing Phases

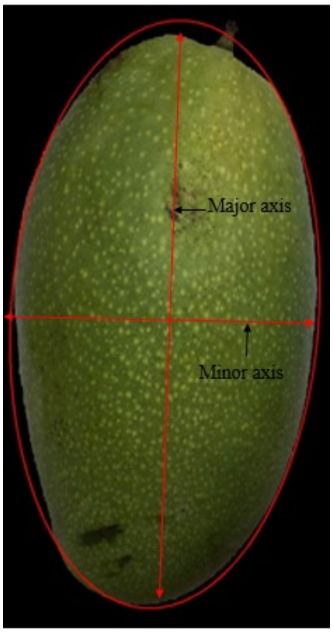


Figure 7: Ellipse Fitting with Major and Minor axis

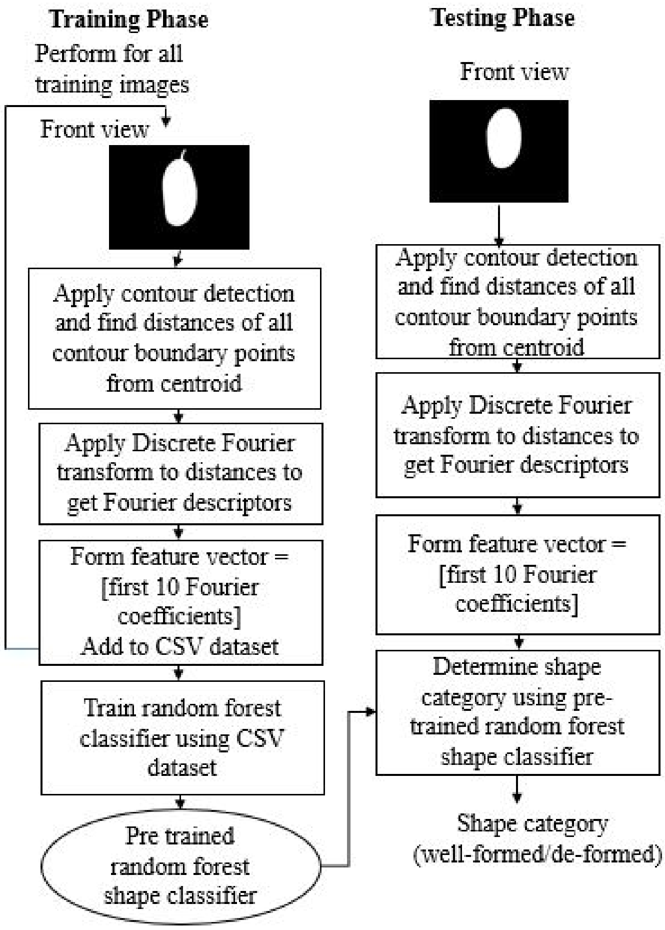


Figure 8: Shape Determination: Training and Testing Phases

4.5 Shape Determination

According to shape, mango was classified as well-formed and deformed. In this research work, Fourier descriptors were utilized for shape classification. Fourier descriptors are features that can represent boundary shape of an object. They are scale, translation and rotation invariant. Contour detection was applied on binary image to find the biggest contour (mango) and centroid was obtained from contour image moments. Distance of centroid from contour boundary points were computed and discrete Fourier Transform was calculated. Magnitude of derived DFT give the Fourier coefficients which were utilized as shape features. First ten Fourier coefficients were taken as shape feature vector. Similar to ripeness determination module, a random forest classifier was initially trained and then utilized to find the mango shape category. Shape determination module steps are depicted in Figure 8.

4.6 Defects Determination

According to defects, mango was classified into one of the three categories: non-defective, mid-defective or completely-defective. Mango defect category was determined from the defect percentage. Defective area from both the mango views were combined to find total defects. Defects can be identified using image segmentation operation. In this work, defect segmentation was performed using an unsupervised machine learning method called K-means clustering. In K-means clustering, the N data points or observations are grouped into M clusters. The data point is assigned to cluster whose, mean or centroid is closest to it. For image segmentation, data points are pixel values and clusters are different image colors. Initially, defect segmentation was tried using fixed global thresholding on green channel. However, since mangoes could be at any ripening stage, determining fixed threshold value was difficult. K-means clustering with no of clusters = 2 was also tried, however due to illumination effects in mango image, non-defective part was also getting segmented as defective. Hence K-means clustering with no of clusters(N) = 3 was first applied on the green component of RGB image. Further the pixels which fall into cluster with centroid value <=55 were assigned pixel value 0 and pixels belonging to other clusters were assigned value 255. This resulted in defect segmented image with defective area marked black and non-defective area white. From the defect area, defect percent was calculated using equation 1. Defect percent of front and back views were added to get total defect percentage.

$$\%df = dfa/ma * 100 \quad (1)$$

dfa = defective-area
ma = mango-area

According to total defect percent, mango defect category is determined as shown in Table 1:

Table 1: Defect Category Determination Thresholds

Total defect-percent	Defect category
<=5	Non-defective
> 5 and <= 20	Mid-defective
> 20	Completely-defective

4.7 Grade Determination

Once the ripeness, size, shape and defect category of mango is found, final step is to determine its grade. In this work, grade determination was done based on a grading formula. The formula was devised according to inputs obtained from mango exporters and mango grading standards [1]. Initially according to mango category, a quality score for each parameter was assigned to mango for eg. small mangoes are assigned lower size score than large mangoes. Scores are assigned as 1,2,3 for small, medium, large respectively and 1,2,3 for ripe, mid-ripe and unripe mangoes. Thus, according to size, preference was given to large mangoes and according to maturity, unripe mangoes are preferred. Shape and defects are given highest importance while grading, hence shape scores are 0/1 for deformed/well-formed mangoes. According to defects, preference was given to less defective mango, hence a completely defective mango was given score of 0. The grading formula used in this research work is as given in equation 2. Highest weight-age is given to shape and defect parameters, while ripeness and size are given lower but equal weight-age. Grade is determined according to computed composite score of mango.

$$\begin{aligned} total\_score &= (shape\_score \times defect\_score) \\ &\times (0.5 \times ripeness\_score + 0.5 \times size\_score) \end{aligned} \quad (2)$$

Another generalized formula which can be used is given in equation 3.

$$total\_score = \sum parameter_{weight} * parameter_{score} \quad (3)$$

Where parameter can be ripeness, size, shape or defect. The advantage of using a grading formula is that change in parameters and their weights can lead to a different grading all together. For eg, If ripeness is not relevant for a grading standard, then it can be skipped. Similarly, variable weights can be assigned to parameters according to their importance. For eg, If a standard prefers only large mangoes, highest weightage can be given to size parameter. The grading standards according to formula, used in this work (equation 2), are depicted in Table 2.

Table 2: Mango Grading Specifications

Mango Grade	Grading Specifications
Grade 1	Only best quality mangoes well-formed in shape and free of defects belong to this grade.They should be unripe and large in size.
Grade 2	Mangoes in this grade should be best quality w.r.t shape and defects.They can be unripe (small/medium), mid-ripe (large/medium) or ripe (large).
Grade 3	Slightly defective, but well-formed mangoes, of any ripening stage and any size belong to this grade. Also best quality mangoes but mid-ripe (small) or ripe (small/medium) also fall into this category.
Grade 4	Mangoes bad in quality, i.e either deformed or completely defective belong to this grade. These mangoes are usually rejected by vendors.

5 Results and Discussion

Experimentation was carried out on a dataset of Dashehari mangoes. The results obtained for ripeness, size, shape classification along with defect identification and final grading are discussed below:

5.1 Ripeness Classification

According to ripeness, mango was classified into 3 stages. For ripeness, 2 views of each mango were analyzed. The 552 mango images obtained after augmentation (69 mangoes, 2 views, 4-way augmentation) were utilized for ripeness classification. 552 mango images correspond to 276 mangoes. Out of these 276 mangoes, 68 belonged to unripe, 120 to mid-ripe, and 88 to ripe stage. RGB and Lab related color features were extracted and classification was performed using random forest classifier. The dataset of 276 mangoes was divided into 80% training and 20% test set. A 10-fold cross-validation performed on mangoes with mean R and mean a* as features could achieve a training an testing accuracy of 100%. The confusion matrix for te set is as shown in Table 3. The ripeness classificatio results obtained for different combination of color fea tures are also given in Table 4. Thus it can be see that best classification performance was achieved usin RGB (R) and Lab (a*) color components. The varia tion of selected color features (a* and R) with ripenin stages is depicted in box plots shown in Figure.9 and 1(It can be observed that values of R and a* componen go on increasing with ripening stage.

Table 3: Confusion Matrix For Ripeness Classification

		Estimated			
		Unripe	Mid-ripe	Ripe	Total
Actual	Unripe	17	0	0	17
	Mid-ripe	0	24	0	24
	Ripe	0	0	15	15
	Total	17	24	15	56

Table 4: Classification performance for Different Color Feature Combinations

Color Features	Training Accuracy	Test Accuracy
H	97.72%	98.21%
R	95%	92.85%
a*	96.81%	100%
H and R	98.63%	100%
R and a*	100%	100%
H and a*	98%	98.21%
H, R and a*	99.54%	100%

In the above table, H, R and a* correspond to color components of HSV, RGB and CIELa*b* color spaces respectively

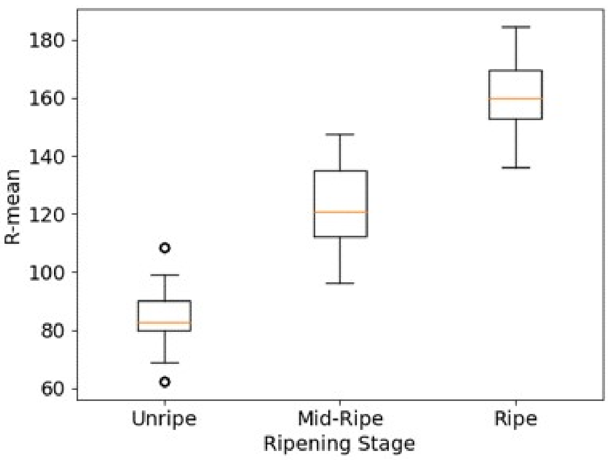


Figure 9: R-mean Color Variations

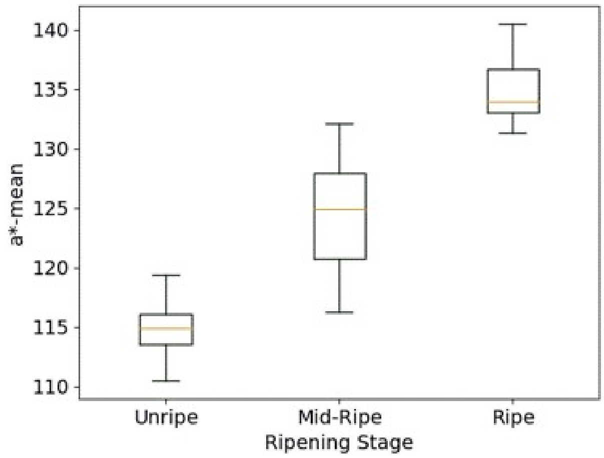


Figure 10: a*-mean Color Variations

5.2 Size Classification

According to size, mangoes were classified into three categories. For determining mango size category, only one of the mango view was utilized. Geometric features namely mango major axis, minor axis and mango area were extracted and fed to a random forest classifier to determine mango size. Out of the 552 mango images, 176 belonged to large, 240 medium and 136 to small category. Labelling was initially done on based on weights. Thus the developed size classification model, is trying to predict mango size according to weight, but from its image without taking actual weight. The dataset was divided into 80% training and 20% test sets. Ten-fold cross validation was carried out using 441 images which obtained a training accuracy of 95.24%. The trained model when tested on test set achieved an accuracy of 98.19%. The classification results on test set are depicted in Table 5. It can be observed that large mangoes are mostly correctly classified. Misclassification is present in between medium and small sized mangoes.

Table 5: Confusion Matrix For Size Classification

		Estimated			
		Large	Medium	Small	Total
Actual	Large	26	0	0	26
	Medium	0	49	2	51
	Small	0	0	34	34
	Total	26	49	36	111

5.3 Defect Identification

As mentioned earlier, K-means clustering using three clusters was utilized for defect segmentation. Two views of each mango were analyzed. The process of defect segmentation is depicted in Figure 11. It was observed that that proposed defect identification approach, was successful in segmenting defects of all the 170 mango images. Simple global thresholding approach and K-means clustering with N=2, resulted in incorrect defect identification as depicted in Figure 12. Hence K-means clustering with N=3 was utilized. Some of the mangoes and segmented defects are shown in Figure 13.

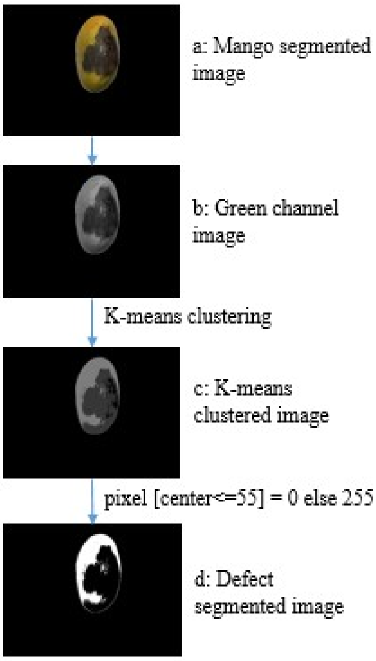


Figure 11: Defect Segmentation Steps

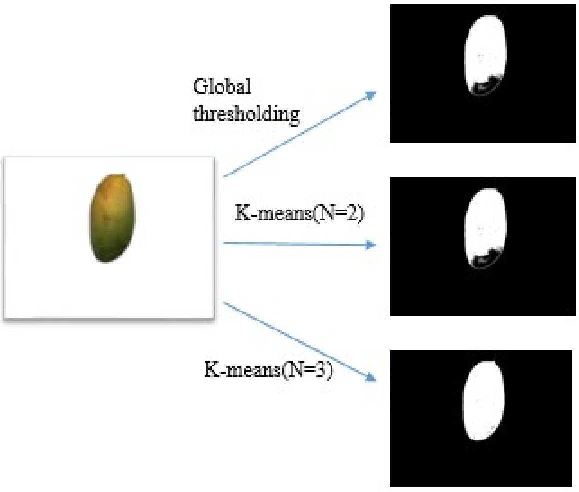


Figure 12: Defect Segmentation Approaches

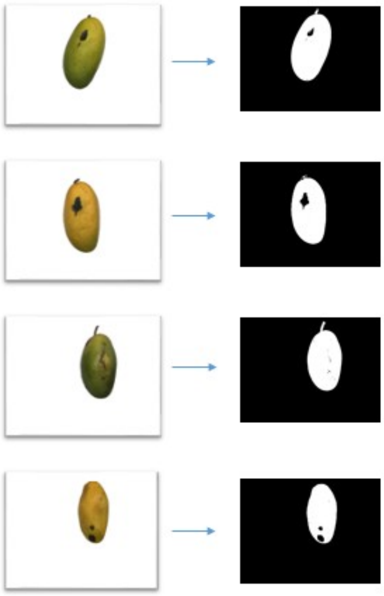


Figure 13: Mangoes and Segmented Defects

5.4 Shape Classification

According to shape, mangoes were classified as well-formed and deformed. In the created Dashehari dataset, all mangoes were well-formed in shape. Hence mango images from Kesar dataset [22] were taken and considered as deformed. A total of 210 well-formed (Dashehari) and 210 deformed (Kesar) mango images were used for shape based classification. Mangoes were classified based on Fourier descriptors using a random forest classifier. First ten Fourier coefficients were successful in obtaining very good classification results. The available dataset of 420 images was split into 70% training and 30% test sets. Ten-fold cross-validation using 294 training images achieved an accuracy of 99.32%. The trained shape classifier when tested on test set of 126 images obtained a recognition rate of 99.20%. The classification results for test set are as depicted in Table 6.

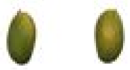
Table 6: Confusion Matrix For Shape Classification

		Estimated		
		Deformed	Well-formed	Total
Actual	Deformed	57	1	58
	Well-formed	0	68	68
	Total	57	69	126

5.5 Grade Determination

Mangoes were classified into one of the four grades (Grade1, Grade2, Grade3 and Grade4) according to standards described in Table 2. Grading was performed based on a formula, once ripeness, size, shape and defect category was found. 18 mangoes (16 Dashehari mangoes and 2 Kesar mangoes) were used as test set. The proposed system graded 16 mangoes correctly. Thus a grading accuracy of 88.88% was achieved. Error in grading was due to the incorrect size category prediction. Grading results for few mangoes along with their predicted categories according to different parameters and final grade is shown in Table 7. Incorrect predictions have been depicted in red color.

Table 7: Mango Grading Results

Mango Views	Shape	Defect	Ripeness	Size	Grade
	Well-formed	Non-defective	Unripe	Large	Grade 1
	Well-formed	Non-defective	Unripe	Medium	Grade 2
	Well-formed	Non-defective	Mid-ripe	Small	Grade 3
	Well-formed	Non-defective	Ripe	Large (Medium)	Grade 2 (Grade 3)
	Well-formed	Complete defective	Mid-ripe	Medium	Grade 4
	De-formed	Non-defective	Unripe	Medium	Grade 4

6 Conclusion

In this research work, an attempt to grade Dashehari mangoes based on ripeness, size, shape and defects has been made. A dataset of 85 Dashehari mangoes was prepared. Two views of mangoes are considered, hence developed system is more reliable. Initially mango category according to each grading parameter was determined which were then utilized to find mango grade (Grade1, Grade2, Grade3, Grade4). The random forest classifier was successful in achieving perfect ripeness and shape based classification. However according to size, there was some misclassification between medium and small mangoes was observed. The ripeness, size and shape based classification on test set achieved an accuracy of 100%, 98.19% and 99.20% respectively. The K-means clustering based segmentation approach also performed well in identifying mango defects. Finally, a mango grading formula was utilized to determine mango grade. Out of the 18 test mangoes, 16 were accurately graded, hence 88.88% grading accuracy was achieved. Thus image processing and machine learning techniques were successfully applied for mango grading. As future work such a system can be developed for other mango varieties. Further grading parameters like firmness, sweetness can be added. A multi-variety grading system which first identifies mango variety and then performs appropriate grading could also be developed.

References

[1] Agmark mango grading standards. `https://dmi.gov.in/Documents/FuitsVegGrd.pdf`. Accessed: 2020-04-20.

[2] Apeda export statistics. `https://agriexchange.apeda.gov.in/product_profile/prodintro/Mango.aspx`. Accessed: 2020-04-20.

[3] Al-Quraishi, T., Abawajy, J. H., Chowdhury, M. U., Rajasegarar, S., and Abdalrada, A. S. Breast cancer recurrence prediction using random forest model. *Recent Advances on Soft Computing and Data Mining*, pages 318–329, January 2018.

[4] Alejandro, A., Gonzales, J., Yap, J., and Linsangan, N. Grading and sorting of carabao mangoes using probabilistic neural network. *AIP Conference Proceedings*, 2045, 2018.

[5] Balbin, J. R., Fausto, J. C., Janabajab, J. M. M., Malicdem, D. J. L., Marcelo, R. N., and Santos, J. J. Z. Profiling and sorting mangifera indica morphology for quality attributes and grade standards using integrated image processing algorithms. *Second International Workshop on Pattern Recognition*, June 2017.

[6] Behera, S. K., Sangita, S., Sethy, P. K., and Rath, A. K. Image processing based detection size estimation of fruit on mango tree canopies. *International Journal of Applied Engineering Research*, 13:6–13, 2018.

[7] BermÃºdez, A. M., Padilla, D. B., and Torres, G. S. Image analysis for automatic feature estimation of the mangifera indica fruit. *Research Article, IngenierÃa y Desarrollo*, 31, 2013.

[8] Blasco, J. and Cubero, S. Cofilab kent dataset. `http://www.cofilab.com/portfolio/mangoesdb/`. Accessed: 2020-04-20.

[9] Capizzi, G., Sciuto, G. L., Napoli, C., Tramontana, E., and Wozniak, M. Automatic classification of fruit defects based on co-occurrence matrix and neural networks. *Proceedings of the Federated Conference on Computer Science and Information Systems*, 5:861–867, 2016.

[10] Chouhan, S. S., Singh, U. P., and Jain, S. Applications of computer vision in plant pathology: A survey. *Archives of Computational Methods in Engineering*, 27:611–632, 2020.

[11] Ganiron, T. U. Size properties of mangoes using image analysis. *International Journal of Bio-Science and Bio-Technology*, 6:31–42, 2014.

[12] Ibrahim, M. F., Saâad, F. S. A., Zakaria, A., and Shakaff, A. Y. M. In-line sorting of harumanis mango based on external quality using visible imaging. *Sensors (Basel)*, 16, 2016.

[13] Khalid, N., Abdullah, A., Shukor, S., A.S, F. S., Mansor, H., and Dalila, N. Non-destructive technique based on specific gravity for post-harvest mangifera indica l. cultivar maturity. *Asia Modelling Symposium*, 2017.

[14] Kumar, M., Jindal, M. K., Sharma, R. K., and Jindal, S. R. Performance evaluation of classifiers for the recognition of offline handwritten gurmukhi characters and numerals: a study. *Artificial Intelligence Review volume*, 53:2075â2097, June 2019.

[15] L.Agilandeeswari, M.Prabukumar, and Goel, S. Automatic grading system for mangoes using multiclass svm classifier. *International Journal of Pure and Applied Mathematics*, 116:515–523, 2017.

[16] Limsripraphan, P., Kumpan, P., Sathongpan, N., and Phengtaeng, C. Algorithm for mango classification using image processing and naive bayes classifier. *Industrial Technology Lampang Rajabhat University Journal*, 12:112–125, 2019.

[17] Marcus Nagle, K. I., Romano, G., Mahayothee, B., Sardsud, V., and MÃ$\frac{1}{4}$ller, J. Determination of surface color of âall yellowâ mango cultivars using computer vision. *International Journal of Agriculture Biological Engineering*, 9:42–50, 2016.

[18] Masetic, Z. and Subasi, A. Congestive heart failure detection using random forest classifier. *Computer Methods and Programs in Biomedicine*, 130:54–64, July 2016.

[19] Mim, F. S., Galib, S. M., Hasan, M. F., and Jerin, S. A. Automatic detection of mango ripening stages â an application of information technology to botany. *Scientia Horticulturae*, 237:156–163, 2018.

[20] Momin, M., Rahman, M., Sultana, M., Igathinathane, C., Ziauddin, A., and Grift, T. Geometry-based mass grading of mango fruits using image processing. *Information Processing in Agriculture*, 4:150–160, 2017.

[21] Naik, S. Non-destructive mango (mangifera indica l., cv. kesar) grading using convolutional neural network and support vector machineâ. ssrn electronic journal. *SSRN Electronic Journal)*, February 2019.

[22] Naik, S., Patel, B., and Pandey, R. Shape, size and maturity features extraction with fuzzy classifier for non-destructive mango (mangifera indica l., cv. kesar) grading. *IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development*, 2015.

[23] Nambi, V. E., Thangavel, K., Shahir, S., and Thirupathi, V. Comparison of various rgb image features for nondestructive prediction of ripening quality of alphonso mangoes for easy adoptability in machine vision applications: A multivariate approach. *Journal of Food Quality*, 39:816â825, 2016.

[24] Nandi, C. S., Tudu, B., and Koley, C. Machine vision technique for grading of harvested mangoes based on maturity and quality. *IEEE Sensors Journal*, 16:6387–6396, 2016.

[25] Pandey, R., Gamit, N., and Naik, P. S. A novel non-destructive grading method for mango (mangifera indica l.) using fuzzy expert system.

International Conference on Advances in Computing, Communications and Informatics, 2014.

[26] Pandey, R., Gamit, N., and Naik, S. Non-destructive quality grading of mango (mangifera indica l) based on cielab colour model and size. *IEEE International Conference on Advanced Communication Control and Computing Technologies*, 2014.

[27] Patel, K. K., Kar, A., and Khan, M. A. Common external defect detection of mangoes using color computer vision. *Journal of The Institution of Engineers (India): Series A*, page 559â568, 2019.

[28] Rojas-Cid, J., Perez-Bailon, W., Rosas-Arias, L., Roman-Ocampo, D., and Lopez-Tello, J. Design of a size sorting machine based on machine vision for mexican exportation mangoes. *IEEE International Autumn Meeting on Power, Electronics and Computing*, 2018.

[29] Roya, P., Kislaya, A., Plonskia, P. A., Luby, J., and Islera, V. Vision-based preharvest yield mapping for apple orchards. *Computers and Electronics in Agriculture*, 164:611–632, 2019.

[30] Rungpichayapichet, P., Mahayothee, B., Nagle, M., Khuwijitjaru, P., and MÃ$\frac{1}{4}$ller, J. Robust nirs models for non-destructive prediction of postharvest fruit ripeness and quality in mango. *Postharvest Biology and Technology*, 111:31–40, 2016.

[31] Saâad, F., Ibrahim, M., Shakaff, A., Zakaria, A., and Abdullah, M. Shape and weight grading of mangoes using visible imaging. *Computers and Electronics in Agriculture*, 115:51–56, 2015.

[32] Sahu, D. and Potdar, R. M. Defect identification and maturity detection of mango fruits using image analysis. *American Journal of Artificial Intelligence*, 1:5–14, 2017.

[33] Thendral, R. and Suhasini, A. Automated skin defect identification system for orange fruit grading based on genetic algorithm. *Current science*, 112:1704–1711, 2017.

[34] U.Rehmana, T., Zaman, Q. U., Chang, Y. K., Schumann, A. W., and Corscadden, K. W. Development and field evaluation of a machine vision based in-season weed detection system for wild blueberry. *Computers and Electronics in Agriculture*, 162:1–13, July 2019.

[35] Velez-Rivera, N., Blasco, J., Chanona-Perez, J., Calderon-Dominguez, G., de Jesus Perea-Flores, M., Arzate-Vazquez, I., Cubero, S., and Farrera-Rebollo, R. Computer vision system applied to classification of âmanilaâ mangoes during ripening process. *Food and Bioprocess Technology*, 8:1183â1194, 2013.

[36] Yossya, E. H., Pranataa, J., Wijayaa, T., Hermawana, H., and Budihartoa, W. Mango fruit sortation system using neural network and computer vision. *Information Processing in Agriculture*, 116:596–603, 2017.