



## Hierarchical approach for ripeness grading of mangoes

Anitha Raghavendra<sup>a,\*</sup>, D.S. Guru<sup>b</sup>, Mahesh K. Rao<sup>a</sup>, R. Sumithra<sup>b</sup>

<sup>a</sup> Maharaja Institute of Technology Mysore, Belawadi, S.R.Patna Taluk, Mandya 571477, India

<sup>b</sup> Department of Studies in Computer Science, Manasagangothri, University of Mysore, Mysore 570006, India

### ARTICLE INFO

#### Article history:

Received 12 February 2020

Received in revised form 6 October 2020

Accepted 10 October 2020

Available online 14 October 2020

#### Keywords:

Alphonso mango

L\*a\*b color space

Threshold based classifier

Support vector machine

### ABSTRACT

Grading of fruits based on their ripeness has been a topic of research for the last two decades. Identifying the ripened mangoes has become more of an art than science and is a challenging task. This study aims at introducing a system to grade mangoes with four classes based on their ripeness. The study was demonstrated through an extensive experimentation on a newly created dataset consisting of 981 images of Alphonso mango variety belonging to four classes viz., under-ripen, perfectly ripen, over-ripen with internal defects and over-ripen without internal defects. In this study, a hierarchical approach was adopted to classify the mangoes into the four classes. At each stage of classification, L\*a\*b color space features were extracted. For the purpose of classification at each stage, a number of classifiers and their possible combinations were tried out. The study revealed that, the Support Vector Machine (SVM) classifier works better for classifying mangoes into under-ripen, perfectly ripen and over-ripen while the thresholding classifier has a superior classification performance on over-ripen with internal defects and over-ripen without internal defects. Further, to bring out the superiority of the hierarchical approach, a conventional single shot multi-class classification approach with SVM was also studied. The results of the experimentation demonstrated that the hierarchical method with an accuracy of 88% outperforms the counterpart conventional single shot multi-class classification approach in addition to several existing contemporary models.

© 2020 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Mango is one of the most popular delicious fruits produced and consumed throughout the world. India is one amongst the most prolific mango growers across globe. India constitutes about 57% of the worldwide production of mangoes with a demand in local and international markets (Khoje and Shrikant, 2012). According to Agricultural and Processed Food Products Export Development Authority (APEDA), gradation and inspection processes would fetch better market price and hence they are necessary steps in post-harvest handling of mangoes. Automatic grading of mangoes has thus become inevitable, not only for export purpose, but also for consumption including food processing industries, as it brings down human intervention while dealing with large quantities. In recent years, machine vision based technology has become more popular in the fields of agriculture and food industry. Machine vision based grading of fruits and vegetables is considered efficient as it avoids inconsistency, inaccuracy, time consuming and labor intensive. There has been an extensive research done on computer vision based grading of fruits and vegetables. Apple stem-end identification, calyx areas and surface defect detection (Li and Weikang, 2002),

segmentation of mango region from mango tree image (Guru and Shivamurthy, 2013), automatic grading of Jonagold apples using hierarchical method based on their external defects (Leemans and Destain, 2004), grading of tomatoes based on defect discrimination (David et al., 2019), automatic segregation of good olives from defective ones (Barreiro et al., 2008), fusion based approach for categorization of fruits and vegetables (Rocha et al., 2010), ripeness based classification of Golden delicious apples (Stefany et al., 2017) and a hybrid method of classifying citrus fruits into immature and diseased on tree canopy images through their detection (Lu and Won, 2018; Sharif et al., 2018) are a few interesting works to mention in literature.

A considerable work could be traced on automatic grading and sorting of mangoes specifically based on attributes related to external appearances such as shape, size, color, weight and defects (Roomi et al., 2012; Sa'ad et al., 2015). A work on grading of mangoes using shape features (region-based and contour based) and wavelet descriptor is reported in (Khoje and Shrikant, 2012). Further, Khoje and Shrikant (2013) have used feed forward neural network and support vector machine for grading of mangoes based on their size features. Identification of external defects on mangoes using texture features is also addressed (Musale and Pradeep, 2014; Pauly and Deepa, 2015).

In addition to the above mentioned attributes, the ripeness parameter is also taken into consideration for grading of mangoes (Vyas et al., 2014; Nayeli et al., 2012; Salunkhe and Aniket, 2015; Chhabra et al., 2011; Mansor and Othman, 2014; Nandi et al., 2014). However, most

\* Corresponding author.

E-mail addresses: [anithbg@gmail.com](mailto:anithbg@gmail.com), [anitharaghavendra\\_ece@mitmysore.in](mailto:anitharaghavendra_ece@mitmysore.in) (A. Raghavendra), [dsg@compsci.uni-mysore.ac.in](mailto:dsg@compsci.uni-mysore.ac.in) (D.S. Guru), [maheshkrao\\_ece@mitmysore.in](mailto:maheshkrao_ece@mitmysore.in) (M.K. Rao).

of the works reported are on only classification of unripen and ripen fruits. Indeed, ripeness evaluation of mangoes plays a very important role at the consumer selection, export level and food industries. Mangoes in different stages of ripening are used to make products like sweets, candy, juice, pulps, ice-creams, etc. Prediction on ripeness stages of mangoes would also help in making up decision during transportation depending on time requirement. Hence, automatic grading of mangoes based on their ripeness is a necessary requirement in real time applications. In general, classification of unripen and ripen mangoes would be easier due to existence of high inter class variations. On the other hand classification of under-ripen, perfect- ripen and over-ripen mangoes is very challenging because of low inter class variation. Therefore, this study is on building up classification model which could stratify mangoes under different stages of ripeness. Over-ripen mangoes generally have some black spots on their external appearances and hence get rejected with an assumption that there may be some internal defects. But, most of the over-ripen mangoes with black spots being not internally defected are safe for consumption. To the best of our knowledge, no work could be traced in literature on classification of over-ripen mangoes with black spots into internally defected or not internally defected and thus has been taken up as a part of this study.

In order to address effectively the above mentioned issues, investigating on a suitable feature extraction technique is required. As color is the common parameter to classify different stages of ripening mangoes, different color features were studied extensively in literature (Stefany et al., 2017; Leon et al., 2006; Khoje and Bodhe, 2016; Khoje and Bodhe, 2015). The  $L^*a^*b$  color space coordinates have been used for the classification of fruits. As CIE  $L^*a^*b$  color space coordinates suit human perception better and its features also show objective color with fidelity (Riquelme and Ruiz-Altisent, 2008) and assess color differences, individual values of  $L^*$ ,  $a^*$  and  $b^*$  can be used to distinguish between any two samples based on their color characteristics. It shall be noted that 'L' indicates lighter/darker, 'a' represents more red/green while 'b' refers more yellow/blue in general. Hence, in this study,  $L^*a^*b$  color space features are recommended to be adopted.

Generally, many important real world classification problems have been resolved using a hierarchical method, where classes are to be organized into levels of hierarchy (Silla and Freitas, 2011). The main objective of this study was to develop a computer vision system that can be used to grade mangoes effectively. The specific objectives of this study were to choose efficient features and to develop suitable classifiers to

grade mangoes based on their ripeness with a hierarchical procedure. This proposed system can be applied to offline mango post harvest management, as precision agriculture aims to have efficient utilization of resources for achieving targeted production of fruits. Overall, the following are the major contributions of this work:

- Extensive study on ripeness evaluation of mangoes for grading.
- Proposal of a hierarchical multi-stage classification approach.
- Exploitation of  $L^*a^*b$  color space features for effective representation.
- Study on suitability of different six classifiers and their possible combinations for optimized performance.
- Creation of dataset of 981 mango images of four different classes.

## 2. Materials and methods

### 2.1. Experiment setup and dataset collection

Experiments were conducted on Alphonso variety of mangoes. This variety has four classes such as under-ripen, perfect-ripen, over-ripen black spots without internal defects and over-ripen black spots with internal defects. A total of 230 mangoes were collected from local farm, Mysore, Karnataka, India. Later the mangoes were subjected to a ripening process guided by different ripening procedures. An overview of the proposed system is shown in Fig. 1. Images were acquired from day 5 after they were kept for ripening, irrespective of the ripeness treatments which were applied on mangoes. Due to the type of ripeness treatment, on day 5 some mangoes had ripened and some had not ripened. Therefore, mangoes were first classified by an empirical method to generate the ground truth. Detailed description on the created dataset is given in Table.1. Some sample images of mangoes of all four classes are shown in Fig. 2 as examples.

The image acquisition system was composed through android mobile camera of resolution 16MP mounted perpendicularly in front of an imaging table at a distance of 15 cm, as shown in Fig. 3. The camera was connected via a data cable to an Intel core with i5-4500U CPU, 4GHz, and 16GB physical memory. Two focusing lights (Built-In Air TTL, Recycling: 0.05 to 1.2 Sec, Li-Ion Battery: 350 Full Power Flashes, LED Modeling Light, 9 Stop Power Range, 76 Ws Output) were placed in diagonal direction to the table at two different corners. Images were captured for each mango sample in different orientations chosen

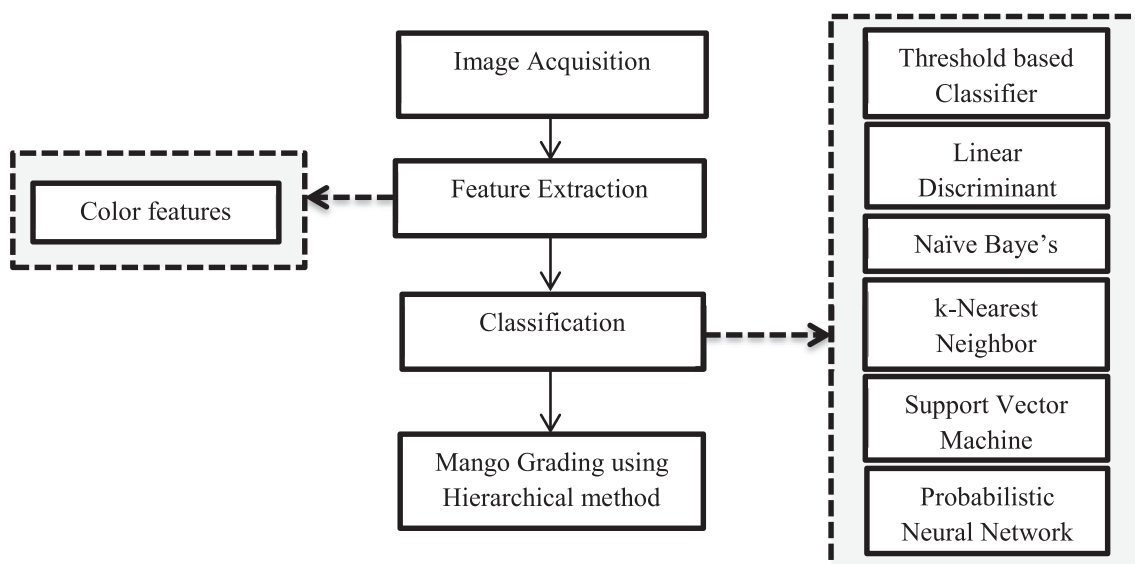


Fig. 1. An overview of the proposed system.

**Table 1**  
The description on the dataset created.

Category	Number of Samples on different days of ripening process					Total number of images
	Day1	Day 5	Day 7	Day 9	Day 11	
Raw	230	–	–	–	–	–
Under-ripen	0	85 340 images (85 × 4 views)	0	0	0	340 images (under-ripe)
Ripen (only 121 samples were used)	0	70 280 images (70 × 4 views)	38 152 images (38 × 4 views)	13 52 images (13 × 4 views)	–	484 images (perfect ripe)
Over-ripen with black spots	0	30 30 images (30 × 1 view)	80 80 images (80 × 1 view)	30 30 images (30 × 1 view)	17 17 images (17 × 1 view)	107 images (without internal defects) 50 images (with internal defects)
Decayed samples	0	45	67	75	26	–
Total images	–	–	–	–	–	981 images

randomly. To make the task of region segmentation simple and easier, the background is made up of black cardboard. The size of each image was set to 4608 × 3456 at the time of acquiring images.

$$\mu_i = \frac{1}{NM} \sum_{a=1}^N \sum_{b=1}^M f_i(a, b) \quad (1)$$

## 2.2. Proposed method

### 2.2.1. Feature extraction

During ripening process of mangoes, changes in skin color and texture are the predominant features irrespective of the method being used for ripening. In our study, initially color features were adopted to represent each stage of ripeness in mangoes. To quantify the color present in a segmented region of a mango, color features from different color spaces such as RGB, L\*a\*b, HSV, CMY and YCbCr were extracted (Gonzalez and Richard, 2002). Four color features were extracted from each individual color space. For example, in case of RGB space, features representing each individual channel (R, G, B) and their fusion as a separate feature were extracted. As mean value provides the average value of all the pixels in an image, means of individual channels of a color space were considered as features.

Let  $f(a, b)$  be an image of size  $N \times M$  expressed by three channels say  $f_i(a, b): i = 1, 2, 3$  in a particular color space. Here  $a$  and  $b$  denote the spatial coordinates. For the image  $f$ , three mean values ( $\mu$ 's) are computed corresponding to each channel using Eq. (1).

In addition to these three features, fusion of these features ( $u$ ) is also taken as a feature in the respective color space and is given by

$$u = \sum_{i=1}^3 \frac{\mu_i}{3} \quad (2)$$

In general, if  $K$  is the number of color spaces each having three different channels then each color space has four different features where three correspond to each individual channel and the other one correspond to their fusion resulting with totally  $4K$  number of features. In this study  $K$  is set to five as mentioned above, there are five color spaces.

As texture features (Haralick and Shanmugam, 1973; Abdur et al., 2013; Dengsheng et al., 2000) were also found to be equally important in literature, in this study the effectiveness of texture features such as Local Binary Pattern (LBP), Grey Level Co-Matrix (GLCM-Haralick) and Gabor were also studied against color features.



**Fig. 2.** Sample images of mangoes of all four classes (a) under-ripen (b) perfect-ripen (c) over-ripen, black spot without internal defects (d) over-ripen, black spot with internal defects.

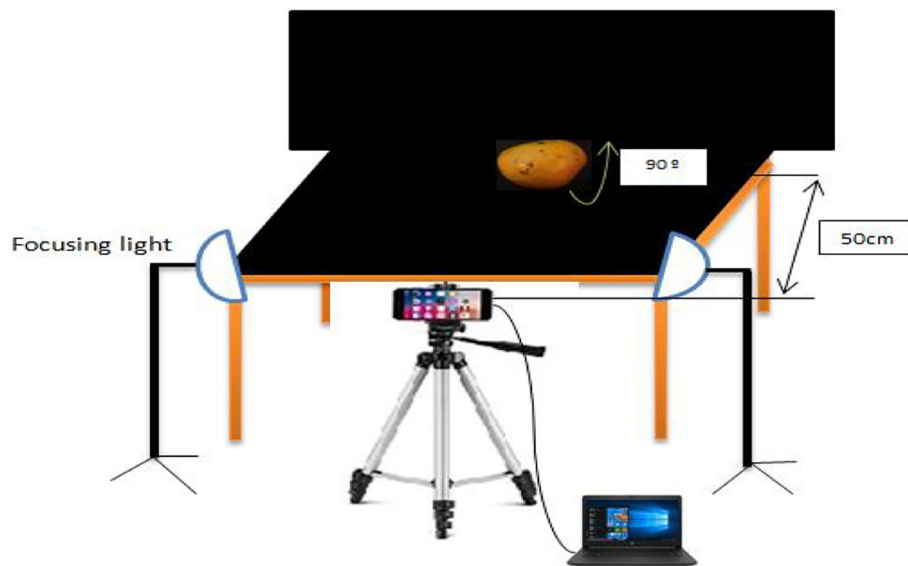


Fig. 3. Image acquisition system.

### 2.2.2. Classification

After extraction of features, mango images were classified using six different classifiers. Naïve Bayes's (NB), k-Nearest Neighbor ( $k$ -NN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Probabilistic Neural Network (PNN) and threshold based classifiers were considered (Islam et al., 2007; Alan and Izenman, 2013; Kim and Choi, 2014; Musavi et al., 1994). The NB classifier, a linear simple probabilistic classifier, predicts a probability distribution over a set of classes for a given observation. The LDA is a linear classifier which separates two or more classes of samples based on a linear combination of the features. The SVM is a discriminative classifier formally defined by a class separating hyperplane. The K-NN is a simple classifier which takes a decision based on the nearest neighbors. In PNN, the parent Probability Distribution Function (PDF) of each class is approximated by a Parzen window and a non-parametric function. The PDF of each class and the class probability of a new input data are estimated by Bayes' rule. Along with these classifiers, a new classifier was designed based on the threshold value of individual color spaces. Threshold based classifier was designed based on the differences between the samples of two different classes for each color space. The threshold values were empirically fixed using training samples. Subsequently based on the threshold values, testing samples were categorized. To validate the model, the dataset was divided into two subsets, a training set and a testing set at each level of classification.

### 2.2.3. Hierarchical method of classification

In order to have a high degree of discrimination, hierarchical structure was adopted for classification assuming a binary classification model at every stage of ripeness as shown in Fig. 4.

Let  $[S_1, S_2, \dots, S_n]$  be a set of  $n$  sample images of mangoes belonging to 4 classes say  $C_j$ :  $j = 1, 2, 3, 4$ , representing different ripening stages such as under-ripen, perfect ripen, black spotted over-ripen without internal defects and black spotted over-ripen with internal defects respectively. The first level of classification, involves binary classification between  $C_1$  and the remaining classes ( $C_2UC_3UC_4$ ). Once the first level classification was done, the second level of classification was considered between  $C_2$  and  $C_3UC_4$ . Finally, the last stage of classification was considered between  $C_3$  and  $C_4$ .

### 2.2.4. Model optimization

In order to optimize a model, it is recommended to study on the performances of each classifier with respect to each individual four color

features of all K number of color spaces at each level of the proposed hierarchical method of classification. Based on the performances of classifiers at each level of classification, the best performing classifier and the corresponding color feature of the respective color space is fixed up. To optimize the system, the dataset was divided into T1% for training and T2% for testing at every level of classification. Let P be the number of classifiers used at every level of the hierarchical method of classification. The study was recommended to be carried out for L number of trials by choosing the training and testing samples randomly.

During optimization, the average accuracy of each classifier across all L trails was calculated and preserved with respect to each color feature of all color spaces. Hence this optimization involves 12PKL trails of empirical study. As the interest is in finding the best classifier and the feature combination for each level, the highest accuracy amongst all trails at a particular level was noted down. The classifier and feature combination corresponding to that highest accuracy was identified to be the method of classification for that particular level. Overall, the proposed hierarchical approach for ripeness grading of mangoes is the sequence of the identified best performing combination of classifier and color feature.

## 3. Experimentation

### 3.1. Performance analysis and model optimization

In this subsection, the recommended hierarchical method of classification has been set up to function at its maximum performance through an extensive experimentation. For this purpose,

the dataset has been divided into a training subset and a testing subset with 60:40 ratio at every level of hierarchical method of classification. The average accuracies over 100 trails for each of the six classifiers at each level of classification (three levels) considering four possible features at a time in each of the five color spaces were computed as explained in Section 2.2.4. Fig. 5 presents the computed accuracies.

Based on the computed accuracies, the best performing classifiers and the corresponding color feature of the respective color space were identified at each level. Fig. 6 gives the plot of highest accuracies obtained in each color space for all the three levels.

In this study, based on the computed accuracies, it is observed that the feature fusion of L,\*a and \*b had attained highest accuracy over rest of the color space features both at level I and level II, while the



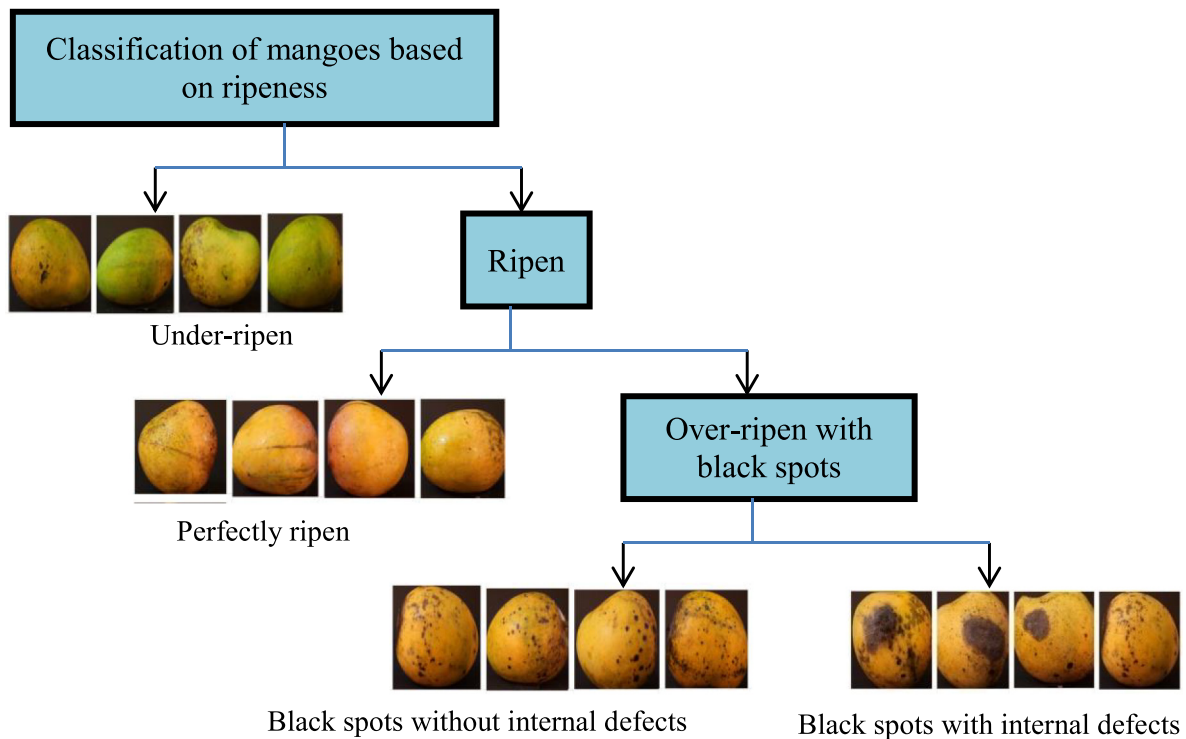


Fig. 4. Hierarchical method for classification of mangoes based on ripeness.

individual color feature  $*b$  of  $L^*a^*b$  color space had attained highest accuracy over the rest of the color space features at level III. It is quite interesting observation that at all three levels irrespective of a classifier, the feature responsible for better performance is from  $L^*a^*b$  color space and hence in the subsequent discussion, only the  $L^*a^*b$  color space is presented. Therefore, the accuracies obtained for all six classifiers using  $L^*a^*b$  fusion feature at level I and level II, and using the individual color feature  $*b$  at level III are tabulated in Table 2.

From Table 2, it is learnt that the SVM classifier performs better than the others at level I and level II, and the threshold based classifier performs better than the others at level III. Hence, the combination of SVM and  $L^*a^*b$  fusion feature is identified to be the method of classification adopted at both level I and level II. On the other hand, for level III, the combination of threshold based classifier and  $*b$  feature is identified to be the best performing one. Thus the optimized model recommended in this study is a hierarchical method of classification with three levels where in, the first level classifies the class of under-ripen mangoes from rest of other classes using SVM with  $L^*a^*b$  fusion feature, second level classifies the class of perfectly ripen mangoes from the class of over-ripen mangoes using SVM with  $L^*a^*b$  fusion feature and third level uses a threshold based classifier with  $*b$  feature to further classify over-ripen mangoes into internally defected and not internally defected mangoes.

As stated in Section 2.2.1, the performance of the recommended system was also analyzed by considering texture features in order to bring out the effectiveness of the color features. The results of  $L^*a^*b$  color features were compared with that of the texture features such as LBP (52 features), GLCM (14 Haralick features) and Gabor (500 features). In Fig. 7 the highest accuracies obtained for all the three types of texture features out of all six classifiers are shown against each of the three levels of classification along with the accuracies obtained by the recommended hierarchical method of classification using  $L^*a^*b$  color space.

From Fig. 7, the texture features were found to be the discriminating features at only level II. In fact it is not surprising as it shall be noticed that in Fig. 2(b and c), there is a considerable difference in terms of textures between perfect-ripen and over-ripen mangoes due to presence of

wrinkles as mangoes get ripened. However, there is no consistency in performance at all levels using texture feature. Hence,  $L^*a^*b$  color space features appear to be the best features even when compared to all other color spaces in addition to the texture features. Since the four classes of mango samples are found to be more distinguishable in  $L^*a^*b$  color space, a 3-dimensional plot of the three feature values of all samples of the dataset is shown in Fig. 8 for visualization.

In Fig. 8, a perceptible separation of under-ripen class from rest of the classes shall be noted. Therefore, a better classification result was obtained at first level. However, in second and third levels, the results were comparatively less due to less inter class variations as perfectly ripen samples appear to be more similar to that of over-ripen as for as color is concerned. Further, the performance at level III is slightly less than that of level II because of the fact that color is not uniform across all samples of level III due to presence of irregular black spots making class separability a difficult one. Hence, well-known statistical classifiers such as SVM were shown to be performing poor when compared to a threshold based classifier at level III.

### 3.2. Discussion and comparative analysis

Conventionally, the problem being addressed in this study is assumed to be a multiclass classification problem, where a single classifier is trained using a particular type of feature to categorize all four classes of mangoes. Therefore, to promote the recommended hierarchical method of classification with three levels, the study has been extended towards performance analysis of a single shot multi-class classification approach using all six different classifiers based on each individual four features (including fusion) of  $L^*a^*b$  color space. Each of the six classifiers was trained on each of the four features of the color space by considering 60% samples of the dataset and confusion matrices were built for the remaining 40% samples. Based on the computed confusion matrices, the accuracy of each classifier was computed for each of the individual feature type and the maximum amongst them was noted down. It was as expected, observed that all classifiers have had maximum accuracy when they were trained on fusion of  $L^*a$  and  $*b$  features. Fig. 9 shows

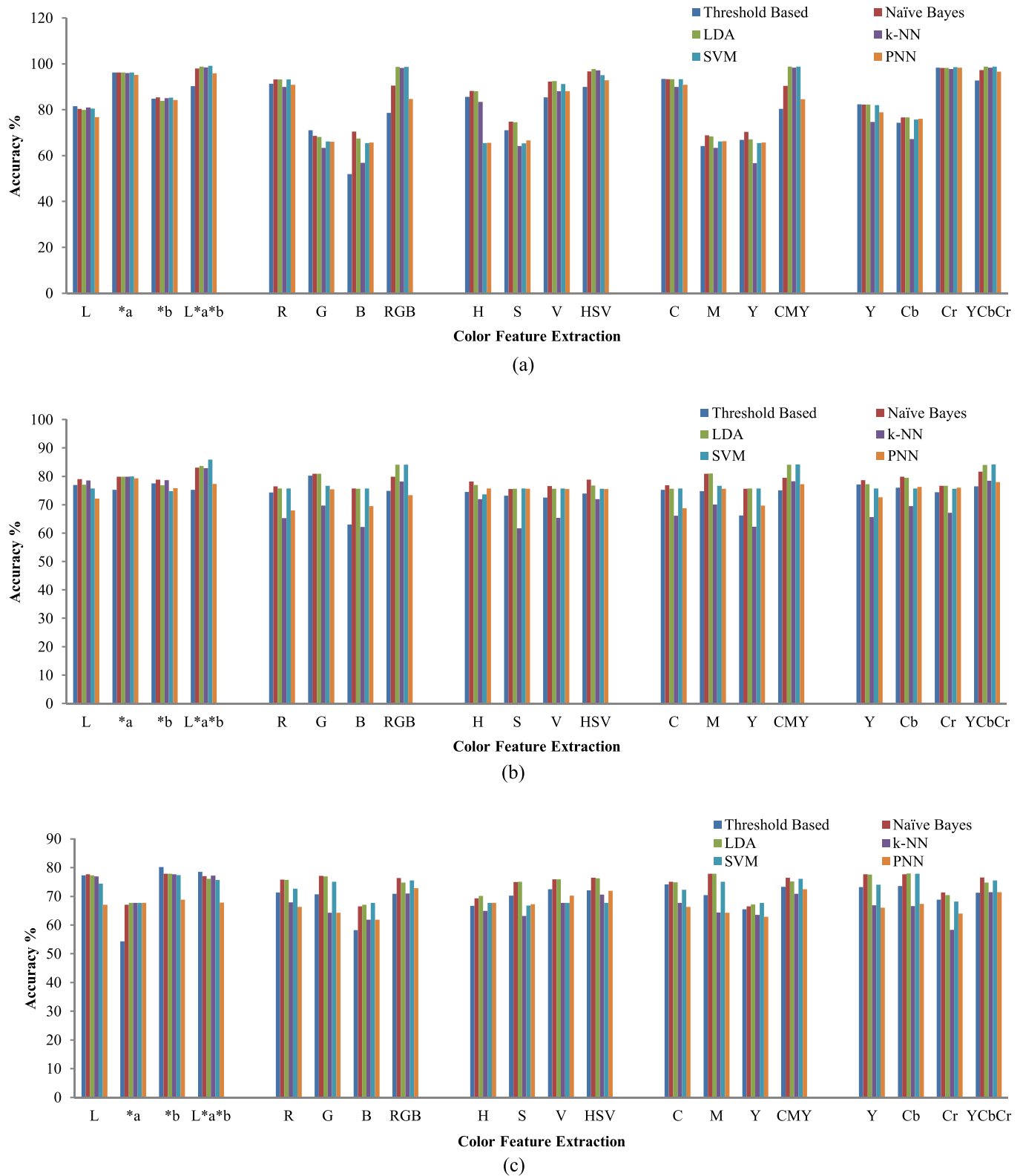


Fig. 5. Accuracies computed for various classifiers considering possible color spaces against each feature at a) I Level (b) II Level and (c) III Level.

the accuracies obtained for each of the six classifiers when they were used for single shot multi class classification with L\*a\*b feature fusion.

In Fig. 9, the SVM classifier attains 86.5% accuracy which is slightly higher than that of all other classifiers. Table 3(a) depicts the confusion matrix obtained for SVM classifier when it is used as a single shot

multi-class classification approach. The 40% of the samples of the dataset considered for computing this confusion matrix involved 136, 193, 43 and 20 number of samples respectively belonging to under-ripen, perfectly ripen, over-ripen back spots without internal defects and over-ripen back spots with internal defects.

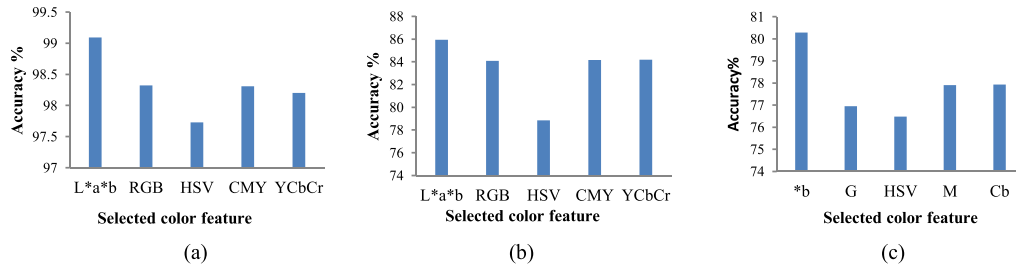


Fig. 6. The maximum accuracy obtained in each color space irrespective of classifier and feature combination at (a) Level I (b) Level II and (c) Level III.

Table 2

Accuracies obtained for different classifiers in  $L^*a^*b$  color space.

Classifiers	Levels		
	I	II	III
Threshold Based	96.20	77.51	80.28
Naïve Bayes	97.95	83.09	77.85
LDA	98.77	83.60	77.88
SVM	99.09	85.94	77.43
k-NN	98.44	82.92	77.72
PNN	95.81	79.32	68.80

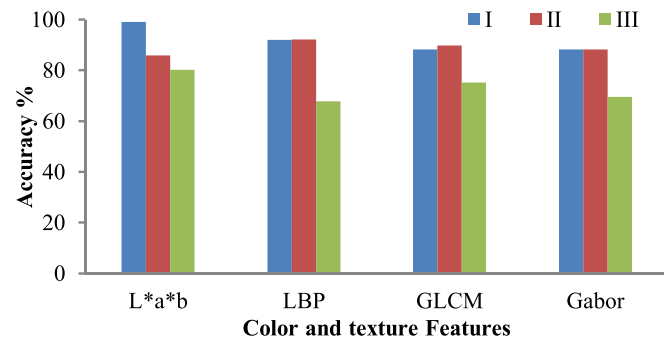


Fig. 7. Comparison of  $L^*a^*b$  color feature with texture features at three levels of hierarchy.

In a similar line, the same 40% of the samples were classified by the use of the recommended hierarchical method of classification and the confusion matrix was generated as given in Table 3(b).

Based on the confusion matrices (Table 3), measures such as accuracy, precision, sensitivity and specificity were computed for both the recommended hierarchical method of classification and the SVM based single shot multi-class classification method and are shown in Fig. 10.

single shot multi-class classification method.

From Fig. 10, it is clear that the recommended hierarchical method of classification achieves considerably better results when compared to that of SVM based single shot multi-class classification method which is shown to be performing better than rest of the classifiers.

Moreover, the recommended hierarchical method of classification has also been compared against existing contemporary works reported on classification of mangoes using ripeness. The works which have been considered for comparative analysis are the only works that could be traced on ripeness based classification/grading of mangoes. The comparative analysis made in this study has considered different qualitative attributes such as mango variety, feature type and classification method in addition to quantitative attributes such as number of samples and accuracy. Table 4 gives details on each of the considered work including the proposed method on all qualitative and quantitative attributes.

As observed in Table 4, the works reported in literatures have considered majorly two classes of mangoes, unripe and ripe which exhibit high inter class variation and only few have considered an additional

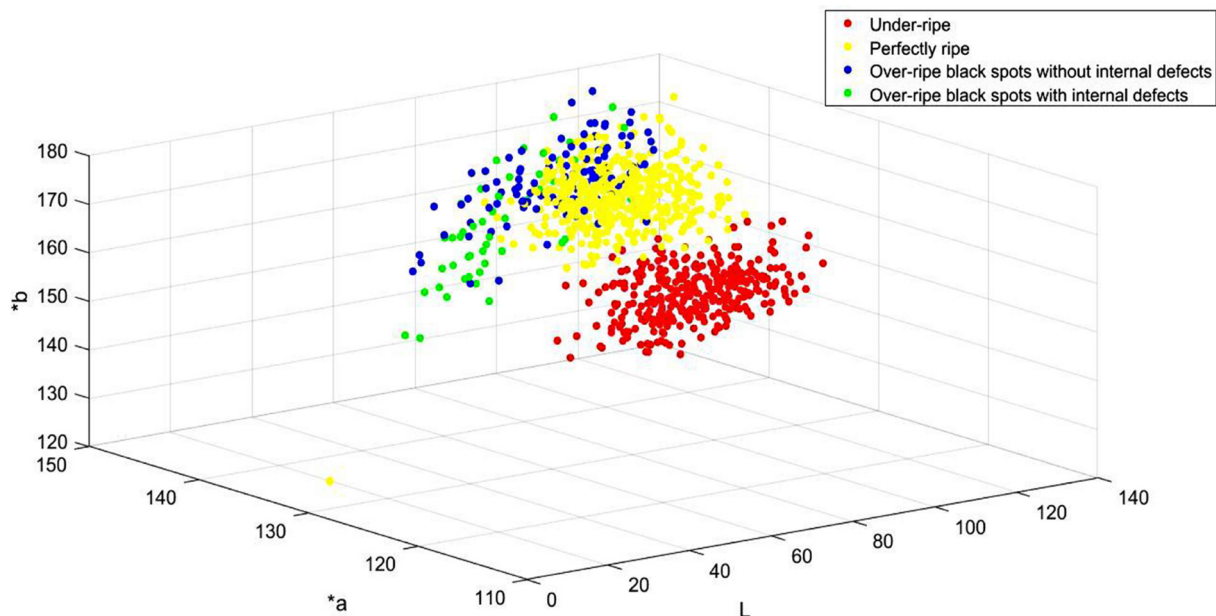


Fig. 8. Three Dimensional plot of  $L^*a^*b$  feature values of all samples of four classes of mangoes.

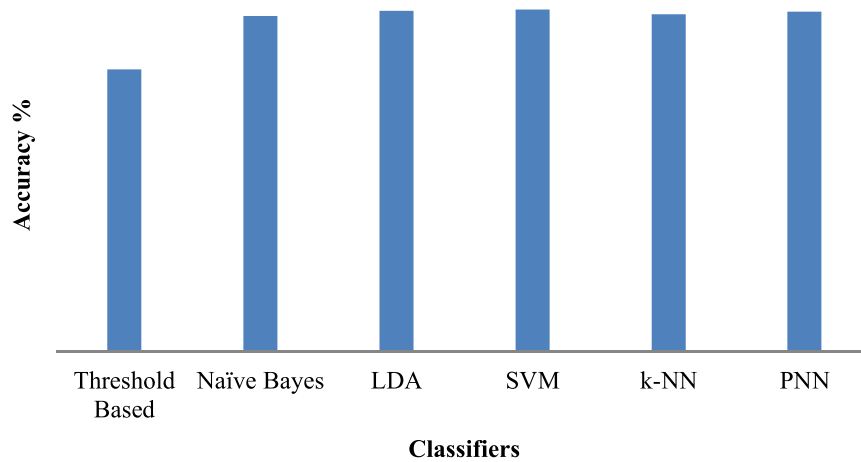


Fig. 9. Accuracies of various classifiers when used for single shot multi class classification with L\*a\*b feature fusion.

Table 3

Confusion matrices obtained for (a) SVM classifier with L\*a\*b feature fusion when used as a single shot classification approach (b) the recommended hierarchical method of classification.

	C1	C2	C3	C4
C1	133	3	0	0
C2	2	183	8	0
C3	0	27	11	5
C4	0	5	5	10

	C1	C2	C3	C4
C1	134	2	0	0
C2	2	174	12	5
C3	0	17	24	2
C4	0	0	5	15

class of mangoes called over-ripen. The classification methodologies adopted in all these works are like a single shot multi- class classification approach which has already been demonstrated to be not superior to the recommended hierarchical method of classification.

On contrary, in this study, in addition to two classes viz., under-ripen and perfectly ripen, two more classes such as over-ripen black spots without internal defects and over-ripen black spots with internal defects were taken at the third level of classification. This added ability of our model is a good contribution of this study because black spotted over-ripen mangoes without any internal defects can be identified so that they can still be put to use in food product manufacturing industries instead of being rejected with an assumption of having internal defects. At each level of classification, suitable classifiers and feature

combination were adopted achieve better results. It is interesting to infer through comparative analysis that the dataset considered in this study is reasonably sized when compared to that of the most of the existing works. It could also be observed that the accuracy, 99.09% obtained by the hierarchical method of classification at level I is relatively higher than that of the other works while classifying under-ripen and ripen classes. At level II and level III, the proposed method exhibits respectively 85.94% and 80.28% of accuracies. As our system is a hierarchical method with three levels of classification where in different classifiers and feature combination were adopted at each level having different accuracies of classification, it is introduced to compute overall system accuracy as follows (Ashok, 1978). Let there be in general four different classes say C1, C2, C3 and C4 respectively with say w, x, y and z number of samples. Let  $\alpha$ ,  $\beta$  and  $\gamma$  were the accuracies of the hierarchical method respectively at level I separating C1 from the remaining classes (C2UC4UC3), at level II separating C2 from rest of the classes (C3UC4) and at level III separating C3 and C4. The system accuracy A is computed as

$$A = \frac{w\alpha + x\alpha\beta + y\alpha\beta\gamma + z\alpha\beta\gamma}{w + x + y + z} \quad (3)$$

For our test dataset with  $w = 136$ ,  $x = 193$ ,  $y = 43$  and  $z = 20$ , and the obtained accuracies  $\alpha = 99.09\%$ ,  $\beta = 85.94\%$  and  $\gamma = 80.68\%$ , the overall system accuracy, A is computed to be 87.34%, which is also tabulated in Table 4.

The entire work presented in this paper has been realized on a i5 core computer with 4GHz having 16GB memory with a support of image processing toolbox in Matlab.

#### 4. Conclusion

In this study, a successful attempt was made to explore the applicability of a hierarchical method of grading of mangoes based on their ripeness. The significance of the L\*a\*b color space features in discriminating four different classes of Alphonso mango variety was brought out. The empirical analysis made in this study argues that the hierarchical approach has a superior performance than a single-shot multiclass approach. Further, the study on six classifiers and their possible combinations with different color features revealed that, the usage of SVM and simple threshold classifier was more effective. The success of this study on post-harvest management of Alphonso mangoes motivated for further studies of other varieties of mangoes in addition to consideration of additional parameters such as store time, temperature and moisture. The recommended method of classification can be deployed for offline application in post-harvesting management of mangoes.

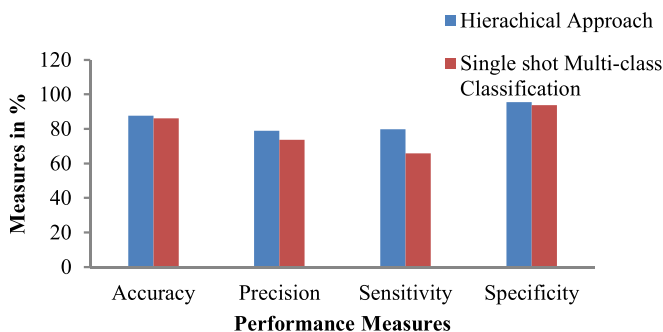


Fig. 10. Different measures computed for hierarchical method of classification and SVM based.



**Table 4**

Qualitative comparison with other well-known methods of mango ripeness classification.

	Number of classes (Mango variety)	Number of samples	Feature type	Classification method adopted	Accuracy
Vyas et al., 2014	3 unripe semi ripe ripe (Totapuri, Badami*)	150 mangoes	L*a*b	Rule based classification (Histogram of L*a and *b)	94.97%
Salunkhe and Aniket, 2015	4 Green stage unripen Ripen Fully ripen (Alphonso)	80 mangoes	RGB and HSV	Rule based classification (Relative changes in the components i.e. R/G, R/B, S/H)	90.4% (RGB) 84.2% (HSV)
Chhabra et al., 2011	2 unripe ripe	200 mangoes	HSV	Neural Network	95%
Mansor and Othman, 2014	3 under-ripe ripe over-ripe (Chokanan)	60 mangoes	RGB	Fuzzy logic	87%
Nandi et al., 2014	4 Raw Semi-matured Matured Over-matured (Kumrapali, Amrapali, Sori, Langra, Himsagar)	750 mangoes	RGB	Gaussian Mixture model and Fuzzy logic	less than 90% in all varieties
Proposed Study	4 1) Under-ripen, 2) Perfect- ripen, 3) Over-ripen black spots without internal defects and 4) Over-ripen black spots with internal defects. Alphonso	230 mangoes	L*a*b	Hierarchical method of classification (sequence of classifier and feature combination)	Level I-99.09% Level II 85.94% Level III- 80.68% Mean Accuracy- 87.34%

\* Badami is also called Alphonso.

## CRediT authorship contribution statement

**Anitha Raghavendra:** Conceptualization, Methodology, Software, Writing - original draft. **D.S. Guru:** Visualization, Investigation. **Mahesh K. Rao:** Supervision. **R. Sumithra:** Software, Writing - review & editing.

## Acknowledgments

We thank all the reviewers for their critical cum encouraging comments which helped us a lot in significant improvement on quality and readability of the content of the manuscript.

## References

- Abdur, Md. Rahim, Shafiul, Md. Azam, Nazmul, Hossain, Rashedul, Md. Islam, 2013. Face recognition using local binary patterns. *Glob. J. Comput. Sci. Technol.* 13 (4) ISSN: 0975-4172.
- Alan, Julian, Izenman, 2013. Linear discriminant analysis-modern multivariate statistical techniques. Springer in Statistics. 8, pp. 237–280. <https://doi.org/10.1007/978-0-387-78189-1>.
- Ashok, V. Kulkarni, 1978. On the mean accuracy of hierarchical classifiers. *IEEE Trans. Comput.* 27 (8), 771–776.
- Barreiro, Pilar, Constantino, Valero, Margarita, Ruiz, Altisent, 2008. Olive classification according to external damage using image analysis. *J. Food Eng.* 87 (3), 371–379.
- Chhabra, Manish, Abhishek, Gupta, Prateek, Mehrotra, Smarti, Reel, 2011. Automated detection of fully and partially ripened mango by machine vision. *Proceedings of the International Conference on Soft Computing for Problem Solving-Springer*. 131, pp. 153–164.
- David, Ileri, Eisa, Belal, Cedric, Okinda, Nelson, Makange, Changying, Ji, 2019. A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. *Artif. Intelligence Agric.* 2, 28–37.
- Dengsheng, Zhang, Aylwin, Wong, Maria, Indrawan, Guojun, Lu, 2000. Content-based image retrieval using Gabor texture features. *IEEE Transactions on Pattern Analysis and Machine Learning*, pp. 392–395.
- Gonzalez, C. Rafael, Richard, E. Woods, 2002. *Digital Image processing*. 2nd edition. Prentice Hall.

- Guru, D.S., Shivamurthy, H.G., 2013. Segmentation of mango region from mango tree image. *Proc. First International Conference, MIKE*. Springer, pp. 201–211 LNAI. 8284.
- Haralick, R.M., Shanmugam, K., 1973. Textural features for image classification. *IEEE Trans. Syst. Man. Cybernetics* 3 (6), 610–621.
- Islam, Md J., Wu, Q.M. Jonathan, Ahmadi, Majid, Maher, A., Sid-Ahmed, 2007. Investigating the Performance of Naive- Bayes Classifiers and k-Nearest Neighbor Classifiers. 5 (2) pp. 133–137.
- Khoje, Suchitra, Bodhe, S.K., 2015. Comparative performance evaluation of fast discrete curvelet transform and color texture moments as texture features for fruit skin damage detection. *Springer J. Food Sci. Technol.* 52, 6914–6926.
- Khoje, Suchitra, Bodhe, S.K., 2016. Application of color texture moments to detect external skin damages in guavas. *World Appl. Sci. J.* 27 (5), 590–596.
- Khoje, Suchitra, Shrikant, Bodhe, 2012. Performance comparison of Fourier transform and its derivatives as shape descriptors for mango grading. *Int. J. Comput. Appl.* 53 (3), 17–22.
- Khoje, Suchitra, Shrikant, Bodhe, 2013. Comparative performance evaluation of size metrics and classifiers in computer vision based automatic mango grading. *Int. J. Comput. Appl.* 61 (9), 1–7.
- Kim, S., Choi, J., 2014. An SVM-based high-quality article classifier for systematic-reviews. *J. Biomed. Inform.* 47, 153–159.
- Leemans, V., Destain, M.F., 2004. A real-time grading method of apples based on features extracted from defects. *J. Food Eng.* 61, 83–89.
- Leon, K., Mery, D., Pedreschi, F., Leon, J., 2006. Color measurement in L\*a\*b units from RGB digital images. *Food Res. Int.* 39, 1084–1091.
- Li, Qingzhong, Maohua, Wang, Weikang, Gu, 2002. Computer vision based system for apple surface defect detection. *Comput. Electron. Agric.* 36, 215–223.
- Lu, Jun, Won, Suklee, Haogan, Xiuwen, Hu, 2018. Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis. *Biosyst. Eng.* 171, 78–90.
- Mansor, Ab Razak, Othman, Md., 2014. Fuzzy ripening mango index using RGB color sensor model. *Res. World J. Arts Sci. Commerce* 5 (2) E-ISSN2229-4686 ISSN2231-4172.
- Musale, Sandeep, Pradeep, Patil, 2014. Identification of defective mangoes using Gabor wavelets: a non-destructive technique based on texture analysis. *Int. J. Agric. Innov. Res.* 2 (6), 992–996.
- Musavi, M.T., Chan, K.H., Hummels, D.M., Kalantri, K., 1994. On the generalization ability of neural network classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (6), 659–663.
- Nandi, C., Tudu, B., Koley, C., 2014. Machine vision based techniques for automatic mango fruit sorting and grading based on maturity level and size. *Springer* 8, 27–46.
- Nayeli, V. Rivera, José, J. Chanona Pérez, Reynold, F. Rebollo, José, Blasco, Georgina, C. Domínguez, de María, J.P. Flores, Israel, A. Vázquez, 2012. Description of maturity stages of mango 'Manila' by image analysis and ripening index. *CIGR-Ageng Conference* [www2.atb-potsdam.de](http://www2.atb-potsdam.de).

- Pauly, Leo, Deepa, Sankar, 2015. A new method for sorting and grading of mangoes based on computer vision system. Intl. Adv. Comput. Conf. <https://doi.org/10.1109/IADCC.2015.715489>.
- Riquelme, M.T., Barreiro, P., Ruiz-Altisent, M., Valero, C., 2008. Olive classification according to external damage using image analysis. *J. Food Eng.* 87 (3), 371–379.
- Rocha, Anderson, Daniel, C. Hauagge, Jacques, Wainer, Siome, Goldenstein, 2010. Automatic fruit and vegetable classification from images. *Compt Electron. Agric.* 70, 96–104.
- Roomi, S., Md, Mansoor, Jyothi, R. Priya, 2012. Classification of mangoes by object features and contour modeling. *Int. J. Pure Appl. Math.* <https://doi.org/10.1109/MVIP.2012.6428786>.
- Sa'ad, F.S.A., Ibrahim, M.F., Shakaff, Md A.Y., Zakaria, A., Abdullah, M.Z., 2015. Shape and weight grading of mangoes using visible imaging. *Compt Electron. Agric.* 115, 51–56.
- Salunkhe, R. Pralhad, Aniket, A. Patil, 2015. Image processing for mango ripening stage detection: RGB and HSV method. *IEEE Conf.* <https://doi.org/10.1109/ICIIP.2015.7414796>.
- Sharif, Md, Attique, Md Khan, Zahid, Md Iqbal, Faisal, Md Azam, Ikram, M. Ullah, Lali, Younus, Javed, Md, 2018. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Compt. Electr. Agric.* 150, 220–234.
- Silla, N. Carlos, Freitas, A. Alex, 2011. A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery.* 22. Springer, pp. 31–72.
- Stefany, Cardenas-Perez, Jorge, Chanona-Perez, Juan, V. Mendez-Mendez, Georgina, Calderon-Dominguez, Ruben, Lopez-Santiago, Maria, J. Perea-Flores, Israel, Arzate-Vazquez, 2017. Evaluation of the ripening stages of apple by means of computer vision system. *Bio-system Engineering.* 159, pp. 46–58.
- Vyas, M. Ankur, Bijal, Talati, Sapan, Naik, 2014. Quality inspection and classification of mangoes using color and size features. *Intern. J. Compt Appl.* 98 (1), 1–5.