I.J. Image, Graphics and Signal Processing, **2023, 4, 83-95**
Published Online on August 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2023.04.07

Building a Medium Scale Dataset for Non-destructive Disease Classification in Mango Fruits Using Machine Learning and Deep Learning Models

Vani Ashok*
Department of Computer Science and Engineering, JSS S&TU, Mysuru, Karnataka, India
E-mail: vanisj@sjce.ac.in
ORCID iD: https://orcid.org/0000-0002-2064-7464
*Corresponding Author

Bharathi R K
Department of Computer Applications, JSS S&TU, Mysuru, Karnataka, India
E-mail: rkbharathi@sjce. ac.in
ORCID iD: https://orcid.org/0000-0002-7616-2415

Palaiahnakote Shivakumara
Department of Computer System and Technology, University of Malaya (UM), Kuala Lumpur, Malaysia
E-mail: inshiva@um.edu.my
ORCID iD: https://orcid.org/0000-0001-9026-4613

Received: 13 February, 2023; Revised: 16 March, 2023; Accepted: 25 April, 2023; Published: 08 August, 2023

Abstract: The growing quality and safety concern about fresh agricultural produce among consumers have led to the development of non-destructive quality assessment and testing techniques of fruits and vegetables. Humans judge the quality of fruits based on sensory attributes like taste, aroma etc. The shape, size, color, presence of defects which are external to fruits also influence the degree of consumer acceptability of produce. The traditional time consuming, manual fruit quality inspection is replaced by automated, fast, consistent, non-destructive techniques using computer vision in combination with learning algorithms. But the lack of benchmark datasets for agricultural produce has made an objective comparison of the proposed methods difficult. Hence, the proposed work aims to build a medium scale dataset for mango fruits of “Alphonso” cultivar with three classes: chilling injury, defective and non-defective. The reliability of the proposed dataset consisting of 2279 color images of mango fruits with 736 samples in chilling injury class, 632 samples in defective class and 911 samples in non-defective class, was established using a novel approach of developing a predictive model based on discriminant function analysis (DFA) which assigns group membership to each sample of the dataset. Extensive benchmarking analysis is established on the validated dataset using statistical and deep learning algorithms like support vector machine (SVM) and convolutional neural network (CNN), respectively. SVM achieved significant disease classification accuracy of 95% and 91.52% accuracy was achieved by custom CNN. The results of the proposed work indicate that the proposed color image dataset of mango fruits can be used as a benchmark dataset by other researchers for objective comparison in quality evaluation of mango fruits.

Index Terms: Dataset, Non-destructive, Discriminant Function Analysis, Support Vector Machine, Convolutional Neural Network.

1. Introduction

The acceptability of fresh horticultural produce by consumers or global market depends on the quality parameters which are broadly classified into external and internal components. Some of the external quality attributes that influence the appearance of fruits are size, shape, color, defects and the internal quality parameters constitute flavor, firmness, crispness, nutrition constituents and defects. The traditional method of quality assessment involves manual inspection of each fruit individually or random sampling from large batches by graders or trained personnel and classifying it into its appropriate category based on external quality parameters. These methods are found to be subjective and inconsistent in nature [1]. Also, the older method of assessing the internal quality parameters of fruits is carried out by either puncturing or destroying the fruit.

Recent reviews on advanced sensing technologies for non-destructive fruit quality evaluation highlight the use of computer vision, near infrared and infrared spectroscopy, mechanical contact, chemical sensors, acoustic sensors and various imaging techniques operating at different bands of electromagnetic spectrum like X-ray, computed tomography, magnetic resonance imaging, nuclear magnetic resonance, near infrared, ultrasound and hyper-spectral imaging [2,3]. Computer vision system has been used by the food industry for automated external quality inspection and was found to be a faster, non-destructive, powerful and reliable quality evaluation technique for agricultural produces by research community over past few decades [4].

2. Literature Survey

From past two decades, numerous research works have been published in the arena of external quality analysis of fruits and vegetables by computer vision system based on huge hypothesis space of feature extraction methods and classifiers [5].

Arakeri M P [6] has developed a complete computer vision based quality grading system for tomatoes consisting of hardware and software. The system captures the images of fruits in real time, a trained multilayer neural network classifies the fruits into ripe/unripe and defective/no-defective based on statistical and textural color features obtained from RGB color space and the hardware moves the tomato to respective bins depending on the classification results. The maturity level was determined with 100% accuracy and defect detection was achieved with 96.47% accuracy. Deulkar Shweta S et al. [7] have also proposed a quality grading system for tomatoes using Otsu thresholding and k-means clustering for segmentation and extraction of color features. The SVM classifier trained on these color features classified the test samples into three grading levels with maximum accuracy of 95%. Ireri D et al. [8] have explored the use of computer vision to grade the tomatoes into four quality classes based on color and textural features in RGB color space and to classify fruits into defective and non-defective groups in L*, a*, b* model. SVM classifier with RBF kernel function performed classification and grading with 98.9% and 97.09% accuracy, respectively.

Pise et al. [9] proposed an automated grading tool for harvested mangoes based on quality and maturity. The size, shape and color values in RGB color space were obtained and classification was performed by Naïve Bayes algorithm based on calculated posterior probability of each class. A non-destructive machine vision system was presented by Castro W et al. [10] to classify cape gooseberry fruits into seven different classes based on ripeness level. The color features extracted from RGB, HSV and L*a*b* color spaces, combined with different machine learning models like SVM, ANN, k-NN, decision tree, were exercised to classify fruits. The results revealed that SVM based on L*a*b* features outclassed others with an average classification accuracy of 92.65%. Narendra V G [11] employed soft computing techniques, BPNN and PNN, for detection of defects in orange, lemon, sweet lime and tomato fruits with classification accuracy ranging between 85%-95%. Kaur R et al. [12] have recommended grading of guava fruits into four categories (maturity levels and defect) based on geometrical and local tetra patterns textural features. Many classifiers, SVM, k-NN, decision tree, and ANN, were trained on these features and among all ANN resulted in maximum accuracy of 100%.

Recently the popularity of deep learning in computer vision field has given way for the design of systems that provide end-to-end solutions for image classification problems with huge feature space and considerable number of classes [13]. Pande et al. [14] have explored deep learning neural network approach for classifying three types of fruits; apples, oranges and lemons. The same method was proposed for grading apples into four quality classes: three acceptable grades and one defective. The knowledge gained by deep convolutional neural network (CNN) Inception V3 model on ImageNet dataset was transferred to features of fruits (360 samples) to achieve fruit classification accuracy of 90% and grading success rate of 85%. Computer vision with deep learning approach has also been explored for grading bananas by Ucat R C et al. [15]. 1116 sample images were pre-processed, defects were extracted and finger size value features were used to train and test CNN model. The defect classification accuracy for four classes was found to be more than 90% for proposed four classes. A computer vision system for mango fruit defect detection using CNN with a classification accuracy of 98% has been proposed by Nithya R et al. [16]. The experimentation was conducted on publicly available mango dataset containing 50 samples in each of “good” and “defective” classes. The data augmentation was performed on original dataset to increase the number of samples from 100 to 800.

The major observation from the related works is that researchers have employed various imaging techniques to develop their own datasets for defect detection in horticulture products with limited samples in each class. Also, the size of publicly available mango datasets which have been used in few studies after heavy data augmentation is found to be of limited size with just two classes (defective and non-defective). And, it is evident from the previous works that the use of small sized dataset results in overfitting of the learning models yielding implausible classification accuracy whereas use of heavy data augmentation results in biases in dataset and also the quality assurance of augmented data are expensive. The dearth of benchmark datasets for agricultural produce has made an objective comparison difficult in the present scenario. Thus, a sincere effort is made to build a medium scale mango fruits dataset for investigating the possibility of solving the challenging non-invasive quality analysis problem. An extensive benchmarking analysis is also performed on this dataset using statistical machine learning and deep learning algorithms like SVM and convolutional neural network, respectively. To that end, the following contributions are proposed in this work:

- Building a manually labeled color image dataset of mango fruits belonging to “Alphonso” cultivar and establishing the reliability and validity of the dataset by employing predictive model based on discriminant function analysis (DFA).
- Performing an extensive benchmarking analysis on this dataset by evaluating the performance of SVM and deep CNN. The results suggest that both manually crafted visual feature based classifier and deep CNN model that takes raw input and performs end-to-end learning with direct mapping from input to output, achieved significant performance on defect classification in the proposed dataset of mango fruits.

3. Dataset Preparation and Validation

The mangoes are susceptible to injuries and diseases during its ripening stage and subjected to numerous post-harvest disorders affecting fruit quality. Stem-end-rot, scab, sap burn, spongy stem-end, sooty mold, soft nose, spongy tissue, internal flesh breakdown are some inherent physiological disorders commonly found in mangoes [17]. The post-harvest handling of fruits, if not done under proper conditions, also results in physical and physiological disorders like spongy tissue caused due to mechanical injury, chilling injury that occurs when mangoes are stored below 10^{o}C, internal and external heat injuries due to improper heat treatment. In the proposed work, a dataset consisting of images of “Alphonso” cultivar mango fruits belonging to “Non-defective”, “Defective” and “Chilling injury” class was developed. The proposed model for mango fruits dataset preparation includes the following steps: Sample collection and preparation; Image acquisition and pre-processing; Feature extraction and validation.

3.1. Sample Collection and Preparation

Alphonso variety mango fruits were acquired from an orchard near Mysuru, Karnataka, the southern state of Republic India from two batches of 2016 and 2017 harvests. The fruits were cleaned from dirt and dust. Before the experimentation, the farmers, who are the natural graders and fruit experts from Department of Horticulture, Mysuru, Karnataka were consulted to visually inspect these fruits and to grade them manually. Based on appearance, surface defects and Codex Standard [18] for mangoes, fruits were graded into two quality classes: non-defective and defective. According to Codex Standard, “Extra class”, which contains all the clean, superior quality mango fruits free from latex stains, bruises, blemishes, insect damage, defects, uniform size and maturity level, are considered as non-defective class. And, the remaining fruits with visible defects due to insects, skin blemishes due to sunburn, sunscald, skin breaks, scab, cracks, latex stains, fungal infection, stem end rot, harvest wounds, overripe, bruises or punctures are considered as defective class. A set of healthy mangoes were carefully chosen and inflicted with chilling injury by storing the fruits at 4℃ -5℃ for 3 to 4 weeks and then exposing the fruits to normal temperature for one week, thereby generating third quality class, chilling injury. Depending on the cultivar, temperature and extent of exposure, the symptoms of chilling injury include skin discoloration, uneven ripening, pitting, poor taste, reduced aroma and flavor, increased vulnerability to decay and pulp browning in severe cases [19].

3.2. Image Acquisition and Pre-processing

Among several approaches to be considered for identifying significant features of images, better image acquisition method which reveals more information with less background noise, no illumination irregularities and reduced blurring, aid in generating superior experimental data. Better sample preparation with reduced data corruption also assists in finding better image features. In the proposed work, the images are acquired using Sony Cyber-shot DSC-WX7 digital camera. A total of 2279 color images are acquired with a resolution of 640x480 pixels and 60.5KB size each. The image acquisition is done between 12 P.M to 3 P.M under natural light. Camera is positioned so as to avoid the formation of fruit shadow and a uniform white background is used to ease the segmentation process.

In the pre-processing stage, the acquired images are cropped and resized to a dimension of 200x150 pixels with 5.58KB size to lessen computational complexity. After further visual inspection, the samples are labeled and the class of each sample is determined. The quality of fruits is categorized into one of three classes: “chilling injury” (Class 1) with 736 samples, “defective” (Class 2) with 632 images and “non-defective” (Class 3) with 911 samples and in total 2279 images in the dataset. The overall sample distribution in the proposed dataset is shown in Fig. 1.

The sample images belonging to three pre-defined classes are shown in Fig. 2. Chilling injured mangoes exhibit greyish-scald like discoloration and pitting. The defective samples includes mangoes with all other physical defects such as sap/latex burn, anthracnose, fungal decay, insect damage, soft nose, bruising, scab, scars etc. High quality fruits are considered as non-defective samples.

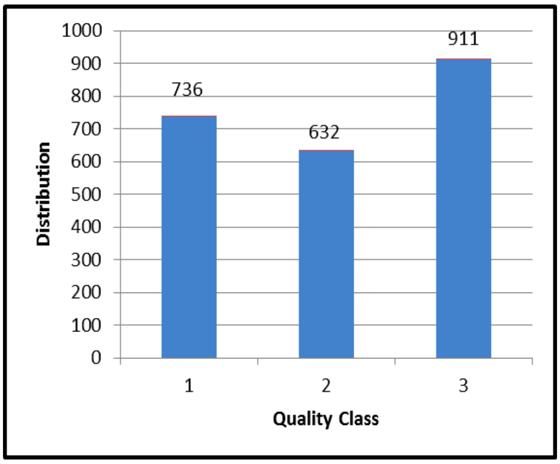


Fig. 1. Overall sample distribution in the proposed dataset.

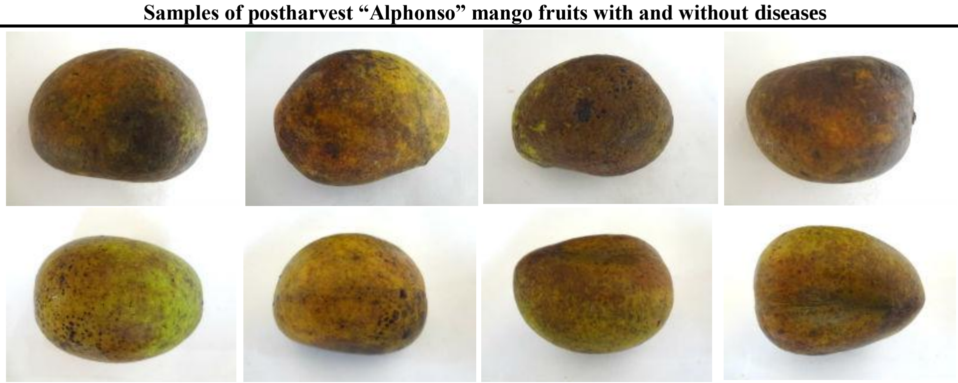
Classes	Samples of postharvest "Alphonso" mango fruits with and without diseases
Chilling injury	
Defective	
Non-defective	

Fig. 2. Sample images of Mango fruits dataset.

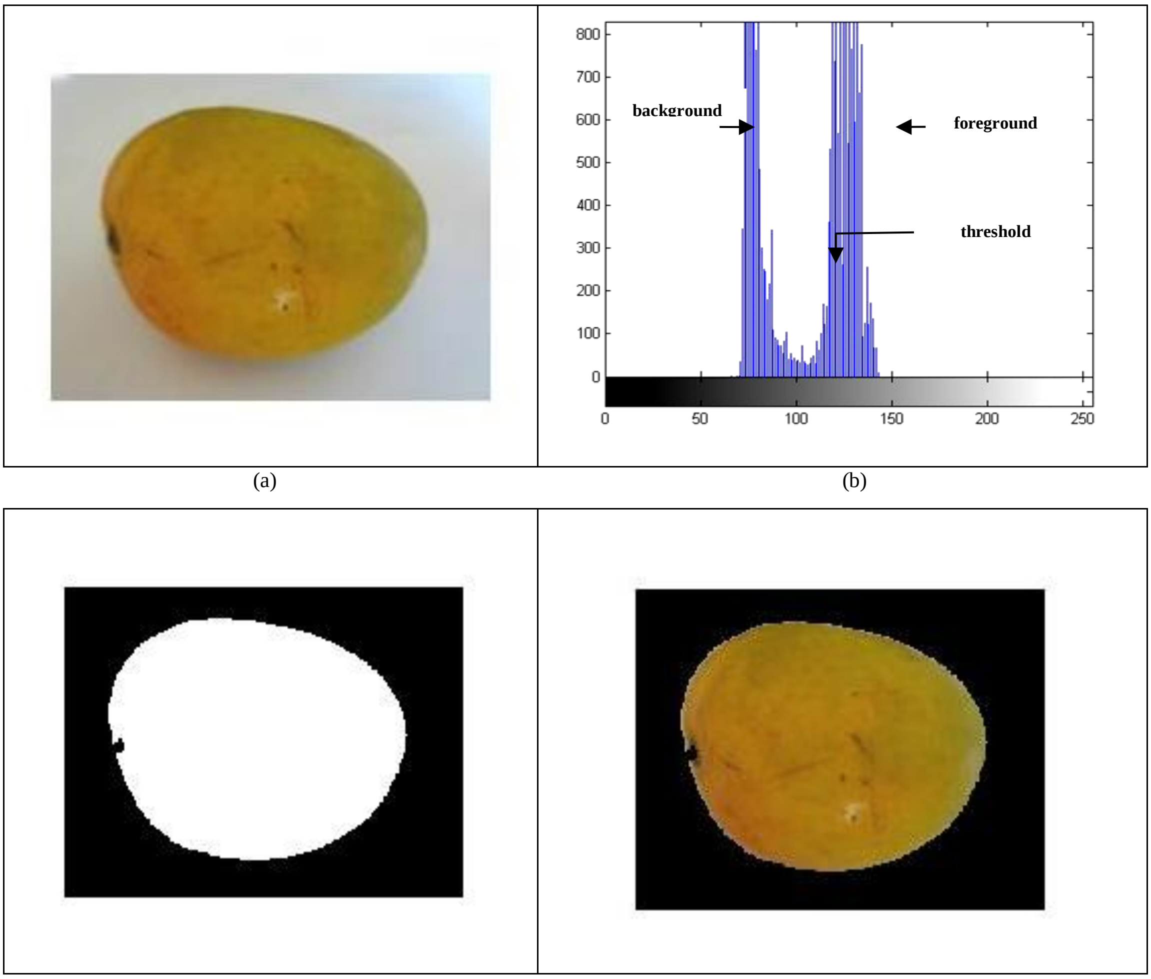
As next step in pre-processing techniques, a better image segmentation method which outputs more accurate region of interest with very less information loss was performed. The histogram based global thresholding was performed in YCbCr color space, as it facilitates more effective segmentation by clearly differentiating fruit pixels from the background color pixels [20]. The histogram dependent segmentation is suggested for images containing large homogenous regions as it separates regions where all area of the object and background are homogenous with an exception to the separation between the object and background [21]. In our work, since the images were acquired with uniform background, simple and easy to implement global thresholding is used where a region is recognized based on the pixels with similar intensity values.

The process of segmentation on a sample image is shown in Fig. 3 and the method includes the following steps:

1. Image in RGB color space (Fig. 3(a)) is converted to *YCbCr*.
2. An initial image is computed by thresholding *Cb* channel information of the image with the estimation of a global threshold using statistical approach. The obtained high contrast gray value image, *I*, has a bimodal histogram as shown in Fig. 3(b), where the left distribution corresponds to the background pixels and the right to the pixels of fruit image. From this image, a binary image *J* is obtained by creating a separation between foreground and background estimating global threshold T. It is represented by Eq. 1.

$$J(i,j) = \begin{cases} 1 & \text{if } I(i,j) > T \\ 0 & \text{else} \end{cases} \tag{1}$$

   where '1' (value 255) is assigned for all pixels in foreground i.e., fruit part and '0' (value 0) is assigned to all pixels of background.
3. A final binary image, shown in Fig. 3(c), is computed using morphological operations on the segmented image to remove isolated pixels in the background and to fill holes in the foreground.
4. Finally, the binary image is mapped to RGB color space to obtain the segmented image as shown in Fig. 3(d).



(a) (b) (c) (d)

Fig. 3. Segmentation based on global threshold. (a) Original image (b) Histogram (c) Binary image (d) Segmented image.

3.3. Feature Extraction and Dataset Validation

The reliability and validity of the proposed dataset was established by building a novel predictive model which determines the group membership of each sample based on discriminant function analysis [22]. The model is composed of discriminant function based on 603 predictor/independent variables which include the following intensity features extracted from R, G and B color channels of each sample: 162 standard intensity features and fourier descriptors, 21 Hu invariant moments, 420 statistical texture information extracted using Haralick approach. The descriptions of these features are given Table 1. The standard intensity features like mean, kurtosis, standard deviation, skewness, mean Laplacian and mean boundary gradient give the overall image intensity information [23]. Hu moments offer features invariant to magnification, translation and rotation and Haralick approach based statistical features provide information about intensity values distribution based on the computation of co-occurrence matrix [24,25]. Sixteen texture features from Discrete fourier transform (DFT) image transformation and 32 from Discrete cosine transform (DCT) were also extracted [26]. These features are considered as independent variables and are assessed to provide the discrimination between dependent variable i.e., groups: chilling injury, defective, non-defective.

Table 1. Summary of extracted features for the predictive model.

Type of Intensity features (Number of features extracted from each color channel of a sample image)	**Description of features**
Standard Intensity (6)	Intensity Mean, Standard deviation, Kurtosis, Skewness, Mean Laplacian, Mean boundary gradient
Filter Banks: DCT (16), DFT (32)	Int-DCT (1…4:1...4), DFT Abs (1,…,4:1,...,4), DFT Ang (1…4:1...4)
Hu Invariant Moments (7)	Hu with intensity (1….7)
Haralick Statistical Textures (140)	14 basic T_x features both in mean and range for 1…5 pixels

The discriminant score (D_i), of discriminant function is predicted as the weighted combination of the predictor variables (x_i) as given in Eq. 2.

$$D_i = a + b_1x_1 + b_2x_2 + \cdots + b_nx_n \tag{2}$$

Here, 'a' is constant of discriminant function and 'b' is discriminant weight for predictor.

As the number of independent variables is large, instead of entering independent variables together, stepwise method was employed to select the best predictor variables to be used in the model. It was guided by F value that indicates the discriminating ability of the variable between groups. Initially the stepwise method starts with step 0, a model without any predictor variable. And, at each step a variable is added into the model if its F value exceeds the 'Entry criteria', 3.84 and is removed from the model if it falls below the 'Removal criteria', 2.71. In total, 65 predictor variables entered the model in 79^{th} step. In the present study, there is higher order discriminant analysis where the number of discriminant functions is one less than the number of groups [22]. Since the proposed model is a 3-way classification problem with three dependent groups, higher order discriminant analysis was employed with two discriminant functions F1 and F2. The first discriminant function maximizes the interclass variance and the second function also maximizes the variance between the groups but controlling the first function.

The chosen two discriminant functions are inferred through standardized and unstandardized discriminant coefficients of 65 variables that entered the model. The predictor variables that entered the model, their standardized and unstandardized discriminant function coefficients are given in Table 2. The correlation coefficients of 65 variables associated with each discriminant function indicate that that among 65 predictor variables which entered the model, 45 are Haralick texture features and 20 are basic intensity and fourier features. The Hu invariant moment features did not contribute to any of the discriminant functions as they did not enter the model and hence are considered to be insignificant in predicting the group membership.

Table 2. Canonical discriminant function coefficients.

SI. No.	**Predictor Variables**	**Standardized Discriminant Coefficients**		**Unstandardized Discriminant Coefficients**	
		F1	**F2**	**F1**	**F2**
1	R-Intensity Skewness	.088	.078	.138	.122
2	R-Mean Laplacian	-.043	-.367	-.141	-1.208
3	R-DCT(3,2)	.002	.166	.000	.034
4	R-DCT(4,3)	-.131	.011	-.048	.004
5	G-Intensity Mean	-.709	-.171	-.036	-.009
6	G-Intensity StdDev	.168	.000	.025	.000
7	G-Intensity Kurtosis	-.026	-.208	-.017	-.135
8	G-Intensity Skewness	-.072	.063	-.141	.125
9	G-Fourier Abs (1,1)	-.873	-.203	.000	.000
10	G-Fourier Abs (2,1)	.063	-.120	.000	-.001

11	B-Intensity Mean	-4.460	-.219	-.355	-.017
12	B-Intensity StdDev	-1.360	-.312	-.360	-.083
13	B-Mean Boundary Gradient	-.099	.237	-.036	.086
14	B-Fourier Abs (1,1)	2.243	.334	.001	.000
15	B-Fourier Abs (1,2)	-.239	.050	-.001	.000
16	B-Fourier Abs (1,4)	-.076	-.087	-.001	-.001
17	B-Fourier Abs (2,1)	-.373	.017	-.002	.000
18	B-Fourier Abs (2,2)	-.177	-.153	-.001	-.001
19	B-DCT(1,1)	1.085	-.015	.030	.000
20	B-DCT(2,1)	-.279	.085	-.035	.010
21	R-Tx 5,d1(range)	-.404	-2.41	-29.19	-174
--	---	--	--	--	--
--	---	--	--	--	--
65	B-Tx12,d5(range)	.614	-.407	9.176	-6.092
	(Constant)	-----	-----	-159.21	90.857

Tx c, d (mean, range): Haralick textural features, where c is the number of neighbor pixels and d is the texture type.

The standardized discriminant function coefficients indicate the discriminating ability: the larger absolute values better discriminating ability of corresponding variables. The unstandardized discriminant coefficients are used to build the prediction equations as the linear weighted sum of predictor variables. The actual prediction equations built from unstandardized coefficients to classify new cases are given in Eq. 3 and Eq. 4.

$$\begin{gathered} \mathit{Function}\ 1\!: \ -159.212\ +.138\ \mathrm{R-Intensity\ Skewness}\ -. \\ 141\ \mathrm{R-Mean\ Laplacian}\ +\ .000\ \mathrm{R-DCT(3,2)}\ -\ .048\ \mathrm{R-DCT(4,3)} \\ -\ 0.036\ \mathrm{G-Intensity\ Mean}\ +\ \ldots\ldots -\ 23.069\ \mathrm{B-Tx12,d5(mean)}\ +\ 9.176\ \mathrm{B-Tx12,d5(range)} \end{gathered} \tag{3}$$

$$\begin{gathered} \mathit{Function}\ 2\!: 90.857 + .122\ \mathrm{R-Intensity\ Skewness}\ -1.208\ \mathrm{R-Mean\ Laplacian} \\ + .034\ \mathrm{R-DCT(3,2)}\ -\ .004\ \mathrm{R-DCT(4,3)}\ +\ .009\mathrm{G-Intensity\ Mean}\ +\ \ldots\ldots\ -2.086\ \mathrm{B}- \\ \mathrm{Tx12,d5(mean)} - .6.092\ \mathrm{B-Tx12,d5(range)} \end{gathered} \tag{4}$$

The discrimination between the groups by these two functions can be visualized by plotting the individual scores for the two discriminate functions, F1 and F2. The all-groups scatter plot which is the combined plot of canonical discriminant functions for three groups, chilling injury, defective and non-defective, with respect to two discriminate functions F1 and F2 is shown in Fig. 4. Within each group more the samples are nearer to the group centroid, less discrimination among samples and more probability of falling into the same group.

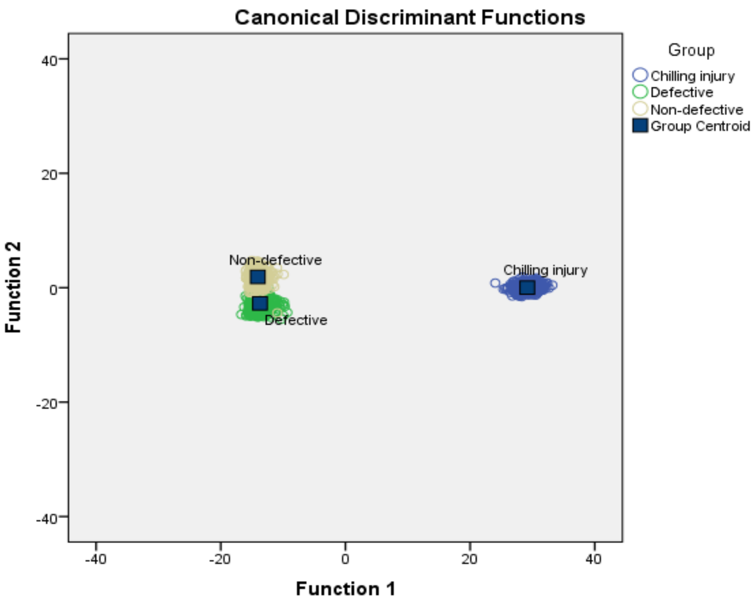


Fig. 4. Combined canonical discriminant functions plot.

So, the classification accuracy based on prior probabilities was calculated, the result of which is tabulated in Table 3. The predicted group membership of each sample in all three groups is specified in the table. The hit ratio, ratio of the correctly classified observation samples to the total number of samples in each group, gives the group membership accuracy and it is 98% for the proposed model i.e., 98% of original grouped cases are correctly classified. The model excels in identifying chilling injury and defective samples with accuracy of 100% and 98.8%, respectively, and performs moderately for defective class with an accuracy of 94.6%.

Table 3. Classification results[a].

Group		Predicted Group Membership			Total
		Chilling injury	Defective	Non-defective	
Original Count	Chilling injury	736	0	0	736
	Defective	0	598	34	632
	Non-defective	0	11	900	911
%	Chilling injury	100.0	0.0	0.0	100.0
	Defective	0.0	94.6	5.4	100.0
	Non-defective	0.0	1.2	98.8	100.0
a. 98.0% of original grouped cases correctly classified.					

4. Mangoes Disease Classification

The present study is extended to develop an automatic defect classification model for mangoes based on the proposed dataset using traditional statistical machine learning algorithm like SVM whose performance depends on the hand-engineered features. The reliability of the proposed medium-sized mangoes dataset for disease classification is also investigated by exploring deep learning approach i.e., custom built deep convolution neural network (CNN) model which provides end-to-end solutions for image classification problems based on raw input.

4.1. Disease Classification using SVM

SVM is basically a 2-way classifier which converts initial feature space of overlapped classes into high dimensional feature space where the classes are linearly separable. Then it implements hyperplane that has the largest distance to the nearest training data point of any class to build a good separation between the classes [27]. However, the classification ability of the SVM model lies in the kernel function that computes the mapping function in the transformed space. In the present work, linear kernel function is used for the classification of mangoes images into three classes. A huge feature space consisting of 2324 intensity features which include standard intensity features, fourier descriptors, Hu invariants, Haralick textural features, local binary patterns (LBP), contrast features and features from Gabor and DCT filter banks was extracted from gray, R, G, B, H, S, V color channels of each sample. The large feature space of these specific intensity features was considered for the proposed model as the validated experiments have established that such feature space provides a hypothesis space rich enough to solve common problems related to food quality evaluation.

Sequential forward selection algorithm with Fisher score objective function (SFS-Fisher) was used for selecting optimal features from the huge feature space [28]. As specified in Table 4, common 29 best features are selected by SFS-Fisher algorithm for SVM classifier with linear kernel function. All these features are intensity features majority being the basic, filter banks and LBP textural features. However, none of contrast features entered the classification model.

Table 4. Details of selected features for SVM classifier with linear activation functions.

Intensity Features				
Basic	Filter Banks	Haralick	LBP	Hu with intensity
g-Mean Boundary Gradient, S-Intensity StdDev, G-Intensity Skewness, B-Intensity StdDev, R-Mean Boundary Gradient, G-Intensity Mean, R-Intensity StdDev, g-Intensity StdDev, G-Intensity StdDev, S-Mean Laplacian, S-Intensity Kurtosis, S-Intensity Skewness	S-Gabor-J, H-Gabor(6,8), S-Fourier Abs (1,2), G-DCT(1,1), G-Gabor(4,3), g-Gabor(1,8), B-DCT(3,2)	G-Tx13,d1 (mean), G-Tx13,d2 (mean)	R-LBP(1,5) [8,u2], B-LBP(1,10) [8,u2], H-LBP(1,10) [8,u2], H-LBP(1,8) [8,u2], H-LBP(1,16) [8,u2], H-LBP(1,15) [8,u2], H-LBP(1,44) [8,u2]	B-Hu-moment-4

The SVM model with linear kernel achieved 95% accuracy for disease classification. This indicates that given a lot of features, the proposed data is already linearly separable and SVM has found the best separating hyperplane. Original dataset is divided into train and test datasets in the ratio of 80:20 and then 10-fold cross-validation with 95% confidence interval is used. The accuracy plot of SVM classifier with linear kernel function for selected 29 features is shown in Fig. 5. It shows that the performance accuracy of the classifier is 95% in confidence interval range of (93.95%, 96.05%) with CI=95% for 29 selected features.

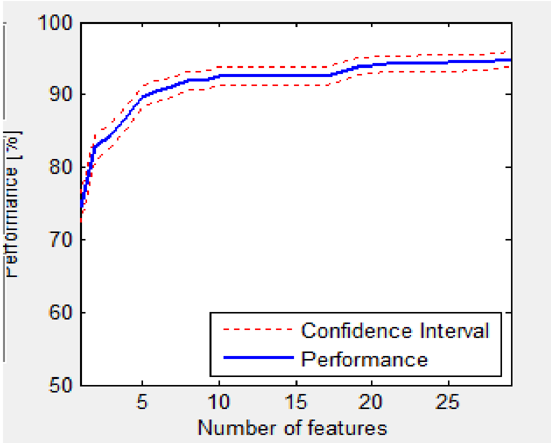


Fig. 5. Performance of SVM-linear with 29 best features and 95% accuracy.

4.2. *Disease Classification using Deep CNN Model*

In the proposed work, the CNN model which uses representation learning with deep learning to extract the features for classification thereby eliminating the need for traditional approaches to hand-engineer features was explored for defect detection and classification in mangoes.

A custom CNN model with the architecture as given in Table 5 is proposed for defect classification in mango fruits. It contains three sets of layers, each having two CONV layers followed by max pool layer. The number of filters is 64, 128 and 256 in first, second and third set of CONV layers, respectively. The filter size is 3x3, padding is 1 and stride is 1 in all CONV layers. CONV layer is followed by pooling layer where no parameter learning occurs. Max-pooling which decreases the input size by a factor of 2, is performed over 2x2 pixel window in both x and y direction with a stride of 1. Finally a FC layer with 128 units and another FC layer with softmax classifier for three object classes are added to the model. For every epoch, the model learns 1,670,211 parameters.

Table 5. Architecture model of custom CNN.

Layer (Type)	Description	Output shape	# of parameters
Input layer	64x64	(64, 64, 3)	0
Conv	64;3x3;p=1,st=1	(62, 62, 64)	1792
ReLU		(62, 62, 64)	0
Conv	64;3x3;p=1,st=1	(60, 60, 64)	36928
ReLU		(60, 60, 64)	0
Maxpool	2x2,st=1	(30, 30, 64)	0
Dropout	0.2	(30, 30, 64)	0
Conv	128;3x3;p=1,st=1	(28, 28, 128)	73856
ReLU	0.2	(28, 28, 128)	0
Conv	128;3x3;p=1,st=1	(26, 26, 128)	147584
ReLU		(26, 26, 128)	0
Maxpool	2x2,st=1	(13, 13, 128)	0
Dropout	0.3	(13, 13, 128)	0
Conv	256;3x3;p=1,st=1	(11, 11, 256)	295168
ReLU		(11, 11, 256)	0
Conv	256;3x3;p=1,st=1	(9, 9, 256)	590080
ReLU		(9, 9, 256)	0
Maxpool	2x2,st=1	(4, 4, 256)	0
Dropout	0.4	(4, 4, 256)	0
Flatten		4096	0
FC	(4096+1) x 128	128	524416
ReLU		128	0
Dropout	0.4	128	0
FC	(128+1) x 3	3	387
Softmax classifier		3	0
Total trainable parameters			1,670,211

The outputs of all convolutional layers are fed to Rectified linear unit (ReLU) rectifier function which adds non-linearity. Dropout regularization method which sets the weights of randomly selected neurons to zero is added after every convolutional layer to avoid overfitting and to speed up the convergence of the model. The convolution stride (st) is fixed to 1 pixel. To preserve the spatial resolution of input, the spatial padding (p) of convolutional layer is set to 1.

The input image size is fine tuned to 64x64 pixels. Total of 2279 mango fruits input images are normalized so that their pixel intensities are within [0, 1] range and the input dataset is split into 70:15:15 for training, testing and validation datasets. The proposed model is then trained on train dataset for 50 or 100 epochs, found with little trial and error, and then test dataset is used as validation dataset to get an idea of the generalization error of the model during training. Finally, the model is evaluated on unseen images of validation dataset only once. The batch size that specifies the number of samples in train dataset that must be processed before estimating the error gradient and updating the weights of the model was set to 64. AdaBound optimization algorithm, which covers both faster convergence and better generalization, with initial and final learning rate initialized to 0.0001 and 0.1, respectively, was used to minimize the cross entropy error loss [29]. The performance of the CNN model was analyzed in terms of classification accuracy and model loss. The model was trained on Intel Xeon® CPU with NVIDIA Quadro K600 GPU. The computational complexity of training CNN model was estimated by measuring training time which is the time required to learn the class.

5. Results and Discussion

The experimentation setup for SVM classifier comprises of dividing the original dataset into train and test datasets in the ratio of 80:20 and then using 10-fold cross-validation with 95% confidence interval is used. The SVM classifier with linear kernel function achieved 95% accuracy for disease classification at a computational time of 0.61 seconds. Since the proposed dataset is an unbalanced dataset with unequal number of image samples in each class, along with accuracy, the performance of the classifier was also evaluated using confusion matrix on the validation set which forms 20% of original dataset and is given in Table 6. The primary diagonal elements are True Positives and the non-diagonal elements are False Positives. From the confusion matrix it is observed that the mispredictions are very less and hence significant classification accuracy.

Table 6. Validation confusion matrix.

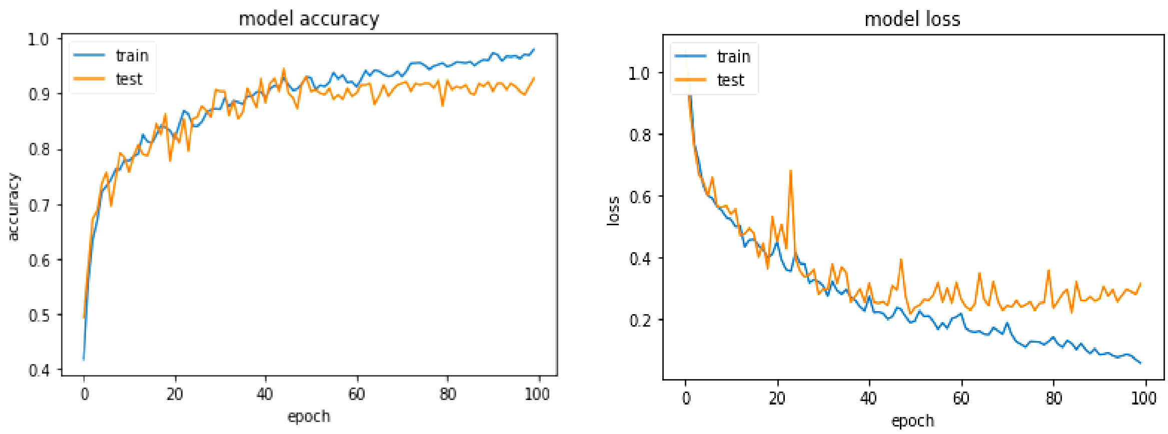
Class		**Prediction**			**Total**
		Chilling injury	**Defective**	**Non-defective**	
Ground Truth	Chilling injury	146	4	0	150
	Defective	3	104	9	116
	Non-defective	1	7	182	190

The experimentation setup for CNN model includes normalizing the images of the proposed dataset and splitting the dataset into 70:15:15 for training, testing and validation purpose. The performance of the CNN model is analysed in terms of classification accuracy and model loss. Initially, the performance of the proposed custom CNN was explored without adding dropout layers for 50 epochs. The train accuracy obtained was 96.3% and the validation accuracy was 90.12%. The test loss increased at the end of training (between 40-50 epochs) and there was clear deviation in the plots of train and test accuracy and the model loss. So, dropout layers were added after every set of layers and also after first fully connected (FC) layer as given in Table 5 and the model was trained. Dropout layers with dropout rate of 0.2, 0.3, 0.4 and 0.4 are added respectively and then the model was trained for 50 and 100 epochs. The performance of the custom CNN model for different number of epochs with and without dropout regularization is summarized in Table 7. Maximum validation accuracy of 91.52% was obtained for CNN model trained for 50 epochs with dropouts.

Table 7. Performance summary of custom CNN model.

Accuracy and Loss	**Without Dropouts**	**With Dropouts**	
	50 Epochs	50 Epochs	100 Epochs
Train Accuracy in %	96.3	93.98	97.99
Train Loss	0.128	0.1474	0.0575
Validation accuracy in %	90.12	91.52	90.35
Validation Loss	0.398	0.280	0.260
Train time per epoch	3s4ms	3s2ms	3s2ms

The model accuracy graph and loss graph, the plots of train and test accuracy and loss over 100 epochs for the proposed CNN model with dropouts are shown in Fig. 6(a) and Fig. 6(b), respectively. It is observed that after 50 epochs the model exhibits slight overfitting as the gap between the train and test plot increases in both accuracy and loss. So ideal number of training steps is 50 and if the number of epochs is increased beyond 50 the dataset size should also be increased.



(a) (b)

Fig. 6. Learning curves of CNN model. (a) Model accuracy for 100 epochs with dropouts (b) Model loss for 100 epochs with dropouts.

6. Conclusion

In the present study, a medium sized benchmark dataset of color images was proposed for non-destructive quality evaluation of mango fruits. The reliability of the dataset comprising of 2279 mango fruits color images falling into 3 classes, chilling injury, defective and non-defective, was established by using a novel predictive model based on DFA. It is also observed that among 65 predictor variables that entered the model, 13 are standard intensity features, 7 features are fourier descriptors and the remaining 45 are Haralick approach based statistical texture features. Hence, it can be concluded that these features have significant discriminating ability and their contribution in disease classification in mango fruits is more. The Hu invariant moments which could not make it to the predictive model are insignificant and may not be considered for disease classification of mangoes. The plot the individual scores for the two discriminate functions, F1 and F2 and also the classification accuracy obtained using prior probabilities indicate that there is distinct discrimination among the samples belonging to three samples, thus establishing the reliability of the dataset.

Further, the validity of the proposed dataset is established by exploring a classic machine learning classifier SVM and a deep CNN model for mango disease classification. SVM with linear kernel function achieved maximum accuracy of 95% with highest precision of 0.97 for chilling injury, 0.95 for non-defective and 0.90 for defective class as derived from confusion matrix. The performance results obtained from custom CNN model is reasonably good for both 50 and 100 epochs. However, adding dropout regularization method and training for fewer epochs i.e., 50 epochs, have further improved the performance of the model with 91.52% classification accuracy. The learning curves, train and test accuracy plots and loss plots indicate that the custom deep CNN model is a suitable fit for the proposed dataset. However, the performance of the CNN model can be improved further by increasing the size of the dataset through data augmentation. Hence, from the results it is concluded that the proposed color image dataset of mango fruits can be used as a benchmark dataset by other researchers for objective comparison in quality evaluation of mango fruits.

References

[1] Pace B M, Cefola D P, Cavallo and Attolico G, Automatic procedure to contactless and non-destructive quality evaluation of fruits and vegetables through a computer vision system, in VI International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain Model-IT 1311, pp. 99-106, 2019.
[2] Bhargava, Anuja, and Atul Bansal, Fruits and vegetables quality evaluation using computer vision: A review, Journal of King Saud University-Computer and Information Sciences 33, pp. 243-257, no. 3 2021.
[3] Ali M M, Hashim N, Abd Aziz S, Lasekan O, An overview of non-destructive approaches for quality determination in pineapples, J. of Agricultural and Food Engineer, pp. 1-7, 2020.
[4] Naik S, Patel B, Machine vision based fruit classification and grading-a review, International Journal of Computer Applications, 170(9), pp. 22-34, 2017.
[5] Nturambirwe J F, Opara U L, Machine learning applications to non-destructive defect detection in horticultural products, Biosystems Engineering, pp. 60-83, 2020.
[6] Arakeri M P, Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry, Procedia Computer Science, 79, pp. 426-33, 2016.
[7] Deulkar Shweta S, Sunita S. Barve, An automated tomato quality grading using clustering based support vector machine, 3rd International Conference on Communication and Electronics Systems (ICCES), IEEE, 2018.
[8] Ireri D, Belal E, Okinda C, Makange N, Ji CA, Computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing, Artificial Intelligence in Agriculture, 2, pp. 28-37, 2019.

[9] Pise, Dnyaneshwari, G. D. Upadhye, Grading of harvested mangoes quality and maturity based on machine learning techniques, IEEE International Conference on Smart City and Emerging Technology (ICSCET), pp. 1-6, 2018.
[10] Castro W, Oblitas J, De-La-Torre M, Cotrina C, Bazán K, Avila-George H, Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces, IEEE Access, 7(27), pp. 389-400, 2019.
[11] Narendra V G, An intelligent computer vision system for vegetables and fruits quality inspection using soft computing techniques, Agricultural Engineering International, CIGR Journal, 21(3), pp. 171-8, 2019.
[12] Kaur R, Kumar M, Juneja M, Data Generation and Fruit Grading Using Different Classifiers, 5, pp. 189-96, 2018.
[13] Zhang X, Yang J, Lin T, Ying Y, Food and agro-product quality evaluation based on spectroscopy and deep learning: A review. Trends in Food Science & Technology, pp. 431-41, 112, 2021.
[14] Pande, Aditi, Mousami Munot, R. Sreeemathy, and R. V. Bakare, An Efficient Approach to Fruit Classification and Grading using Deep Convolutional Neural Network, IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1-7, 2019.
[15] Ucat R C, Cruz J C, Postharvest Grading Classification of Cavendish Banana Using Deep Learning and Tensorflow, International Symposium on Multimedia and Communication Technology (ISMAC), pp. 1-6, 2019.
[16] Nithya R, Santhi B, Manikandan R, Rahimi M, Gandomi A H, Computer vision system for mango fruit defect detection using deep convolutional neural network, Foods, 11(21):3483, 2022.
[17] Supekar A D, Wakode M. Multi-parameter based mango grading using image processing and machine learning techniques. INFOCOMP Journal of Computer Science, 19, pp. 175-187, no. 2, 2020.
[18] Preface-mango, Post-Harvest Profile of Mango, Government of India, Ministry of Agriculture, (Department of Agriculture & Cooperation), Directorate of Marketing & Inspection, Branch head office, Nagpur, 2013.
[19] Patel K K, Khan M A, Kumar Y, Yadav A K, Novel techniques in post-harvest management of mango—an overview, South Asian J Food Technol Environ 5, pp. 821-835, no. 2, 2019.
[20] Ashok V, Vinod DS, Combining Discriminant Analysis and Neural Networks for Detection of Internal Defects in Mangoes using X-Ray Imaging Technique, International Journal of Innovative Technology and Exploring Engineering, 9(2S), pp. 188-194, 2019.
[21] Bhargavi K, Jyothi S, A survey on threshold based segmentation technique in image processing, International Journal of Innovative Research and Development, 3(12), pp. 234-9, 2014.
[22] Denis D J, SPSS data analysis for univariate, bivariate, and multivariate statistics, John Wiley & Sons, 2018.
[23] Mery D, Lillo I, Loebel H, Riffo V, Soto A, Cipriano A, Aguilera JM, "Automated fish bone detection using X-ray imaging", Journal of Food Engineering, 105(3), pp. 485-92, 2011.
[24] Humeau-Heurtier A. Texture feature extraction methods: A survey. IEEE access 7, pp. 8975-9000, 2019.
[25] Haralick RM, Shanmugam K, Dinstein IH, Textural features for image classification, IEEE Transactions on systems, man, and cybernetics, pp. 610-21, (6) 1973.
[26] Chellappa R, Bagdazian R, Fourier coding of image boundaries, IEEE Transactions on Pattern Analysis and Machine Intelligence, (1), pp. 102-5, 1984.
[27] Vapnik V, The nature of statistical learning theory, Springer science & business media, 2013.
[28] Mery D, Soto A, Features: the more the better, Proceedings of the 8th conference on signal processing, computational geometry and artificial vision, pp. 46-51, 2008.
[29] Kingma DP, Ba J, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980, 2014.

Authors' Profiles

Vani Ashok is an Associate Professor in the Department of Computer Science and Engineering. She graduated from University of Mysore, Mysore. She obtained her Masters Degree and Ph.D from Visvesvaraya Technological University, Belagavi. Her research areas of interests are artificial intelligence, machine learning, computer vision, image processing and food quality analysis.

Bharathi R K is a Professor in the Department of Computer Applications. She obtained M.Sc, Technology (by Research) degree from University of Mysore, and Ph.D from from Mangalore University, Mangalore. She has presented her research findings in Moscow, Singapore, Thailand, Romania and America.Her research areas of interests are Biometrics, Pattern Recognition, Computer Vision and Data Analytics.

Palaiahnakote Shivakumara is an Associate Professor in the Department of Computer System and Technology, University of Malaya, Malaysia. He received B.Sc., M.Sc Technology by research and Ph.D. degrees from the University of Mysore, Mysore, Karnataka, India. He worked as a Research Fellow at the National University of Singapore, Singapore from 2005-2007 and 2008-2013. Based on his work, he has published more than 300 research papers in national, international conferences and journals.

How to cite this paper: Vani Ashok, Bharathi R K, Palaiahnakote Shivakumara, "Building a Medium Scale Dataset for Non-destructive Disease Classification in Mango Fruits Using Machine Learning and Deep Learning Models", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.15, No.4, pp. 83-95, 2023. DOI:10.5815/ijigsp.2023.04.07