Classification and Grading of Harvested Mangoes Using Convolutional Neural Network

Hafiz Muhammad Rizwan Iqbal & Ayesha Hakim

Published online: 11 Jan 2022.

Submit your article to this journal ☑

View related articles ☑

View Crossmark data ☑

Taylor & Francis
Taylor & Francis Group

∂ OPEN ACCESS ● Check for updates

Classification and Grading of Harvested Mangoes Using Convolutional Neural Network

Hafiz Muhammad Rizwan Iqbal and Ayesha Hakim

Department of Computer Science, Muhammad Nawaz Shareef University of Agriculture, Multan, Pakistan

ABSTRACT

Mango (*Mangifera Indica L. Family Anacardiaceae*) is a climatic fruit with a short shelf life. A significant percentage of fruit is wasted each year due to the time-consuming manual grading and classification process. There is a need to replace the traditional methods by adopting automation technologies in the agriculture sector. This paper presents a deep learning-based approach for automated classification and grading of eight cultivars of harvested mangoes based on quality features such as color, size, shape, and texture. Five types of data augmentation methods were used: images rotation, translation, zooming, shearing, and horizontal flip. We compared three architectures of 3-layer Convolutional Neural Network (CNN): VGG16, ResNet152, and Inception v3 on augmented data. The proposed approach achieved up to 99.2% classification accuracy and 96.7% grading accuracy respectively using the Inception v3 architecture of CNN.

Introduction

Mango (*Mangifera Indica L.* Family *Anacardiaceae)* is the second major fruit crop in Pakistan. South Punjab region in Pakistan is famous for mango production throughout the world. Pakistan produces 8.5% of the world's mangoes that ranks it the fifth largest mango producer in the world. Punjab and Sindh are the major mango growing provinces of the country that accounts for 94% of total mango production in Pakistan (Mangan and Ruthbah, 2018; Planning Commission of Pakistan, 2020; SMEDA, 2020). Mangoes from the Multan region are well known all over the world for their quality traits that could be attributed to the warm, frost-free climatic conditions that are ideal for mango production. More than 450 cultivars of mango are cultivated in Pakistan (SMEDA, 2020). The most famous and export-quality cultivars grown in Punjab are Chaunsa (Summer Behsht), Chaunsa (Black), Chaunsa (White), Dosehri, Langra, Sindhri, Anwar Ratool, and Fajri. Chaunsa and Sindhri are the major export cultivars of Pakistan. The global consumption of mangoes has increased due to increase in world population, and also increased usage of mango in cooked food, jams, drinks, milkshakes, and salad. The demand for mangoes is expected to increase in the future.

Some of the major hurdles in exporting mangoes include a lack of standardized grading protocol, certification, and improper packaging that reduces its demand in the international market. During the COVID-19 pandemic, the mango industry faced challenges in picking, sorting, and grading due to the unavailability of workers, as lockdown restricted the movement of people. As a result, workers feared coming to the farms for harvesting, grading, and packaging. Another major issue is the lack of automation technology for harvesting, classification, and grading of mangoes. Traditional ways of mango classification are manual where they are sorted based on their shape (ovate or oblong), color, and size. A few photos were taken from various areas of Pakistan as shown in Figure 1.
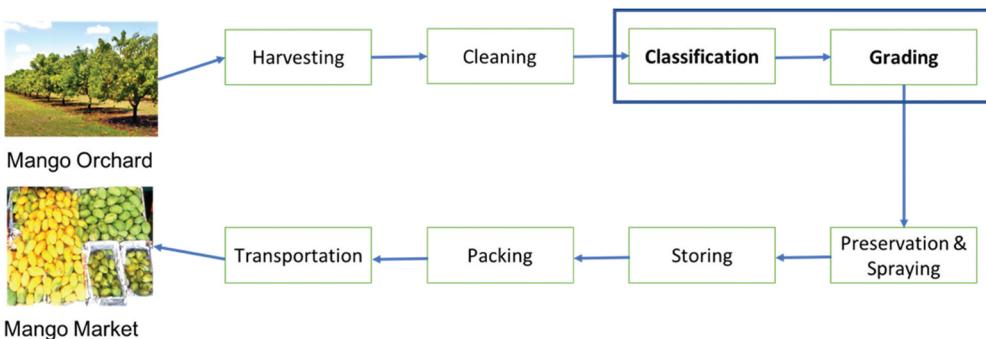
Figure 1. Manual classification and grading of Mango in different areas of Pakistan. **Photo courtesy**: The Dawn News June 04 (2017); Arab News June 16 (2019), and Pakistan Today taken on June 18 (2020).

Due to the lack of adequate postharvest management and supply chain mechanisms in the mango market in Pakistan, 30% to 40% of fruit gets wasted during post-harvest handling (Mangan and Ruthbah, 2018). To overcome all the above challenges and increase the export of Pakistani fresh mangoes to both local and international markets, there is a need to adopt automation technologies in Pakistan by developing an efficient and intelligent automated system for mango's cultivar classification and grading.

In this work, we are proposing an automated approach for the mango classification and grading according to a standardized protocol of the international market by using the deep learning technique convolutional neural network (CNN). CNN is a deep learning technique that imitates the human brain's working for creating patterns, and processing data for decision-making.

Figure 2 presents a flow diagram of postharvest activities of mango fruit. To meet export quality standards of Pakistani mangoes, we are focusing on automating the process of classification and grading only. Classification refers to the identification of mango into eight exportable cultivars of Pakistan (Chaunsa Summer Bahisht (S.B), Chaunsa (Black), Chaunsa (White), Dosehri, Langra, Sindhri, Anwar Ratool, and Fajri) and grading refers to further classification of a specific mango cultivar into quality standards (Extra Class, Class-I, and Class-II).

Figure 2. Flow diagram of Mango post-harvest activities. This paper focuses on the automation of 'classification' and 'grading' activities only.

The proposed automated classification and grading system will enhance the overall process by enabling good quality mangoes to reach the market on time that would be useful to strengthen the export and agriculture economy of Pakistan.

Background

Recently, researchers are focusing on developing methods for the automated classification of fruits into various categories with respect to quality and cultivar. Various image processing techniques and machine learning classifiers such as Support Vector Machine, K-nearest neighbor classifier, and Artificial Neural Network are used for automatic fruit classification systems. This section includes the literature review for mango classification and grading, respectively.

Classification Techniques

There have been several techniques discussed in the literature for classifying fruits such as mangoes, date fruits, olives, and grapes. A Genetic Adaptive Neuro-Fuzzy Inference System (GANFIS) has been presented by (Anurekha and Sankaran, 2020) to improve the classification and grading of mango fruit. The proposed system interpreted the 2D mango images and extracted various features from the image. The covariance matrix was generated by the extracted features and the adaptive neuro-fuzzy inference technique is applied to perform classification and grading. The authors claimed the GANFIS algorithm has produced results up to 99.2% accuracy.

Zhang et al. (2019) presented a 13-layer CNN for the classification of 18 fruit types with an average accuracy of 94.9%. However, the used dataset was clean that limits its use for practical applications. For mango cultivar classification, Abbas et al. (2018) used the B11 model that is usually used for medical image processing. They achieved an average accuracy of 88% on a limited dataset of 28 images only.

Image processing and machine learning techniques have been used by (Naik and Patel, 2013; Naik et al., 2015; Pandey et al., 2013) to improve the quality of the fruit grading and efficient mango classification. The parameters including the shape of the fruit, size, and color were extracted for nondestructive fruit grading and classification. The authors suggested that an automatic sorting system can perform fast, reduce manual labor, and save time to meet the growing demand for high-quality fruit on time. In Lihuan et al. (2017) an electronic nose (odors sensing device) has been used to detect mango fruit quality and develop a quality predictive model by linear fitted eigenvalues.

Grading Techniques

A study has been proposed by Thong et al. (2019) to evaluate the quality of mango fruit before export to the market. A control system was manufactured for mango grading based on image processing technology incorporated with artificial intelligence to classify mangoes in terms of color, shape, size, and volume. Mangoes were tested in a laboratory test bed that was comprised of CCD (charge-coupled device) cameras and conveyors. The first conveyer kept the mangoes moving, and the installed cameras found defects on the mango surface such as black spots and damaged broken. The second conveyor performed the mass calculation by scanning the mango fruit length, width, height. Finally, the mangoes that meet the set criteria were graded into the concerned class.

Ali and Thai (2017) worked on the inspection of two kinds of fruits: apple and mango. They developed an automated fruit grading system to detect the defects on the surface by capturing the fruit's image using a camera placed onto a rotating desk. CCD and complementary metal-oxide-semiconductor (CMOS) sensors are used to estimate the size and shape of fruits to grade and inspect the quality of fruits.

In Naik and Patel (2017) work has been done on mango grading based on maturity and size. As normally, mango's maturity is predicted by its skin color, while in some exceptional cultivars of mango such as in Langdo, skin color remains the same throughout its lifetime. For such cases, reflective

imaging (normal imaging) failed to predict mango maturity. As a solution to this problem, thermal imaging or infrared imaging and an x-ray is used for maturity prediction. In this work, the authors proposed a Fuzzy inference system to detect the maturity of the fruit. FLIR ONE thermal camera was used for image acquisition; the size of the mango was predicted through weight, eccentricity, and area parameters. A fuzzy classifier was used for predicting size features. The accuracy of results received through the system was claimed to be 89%.

Grading of harvested mangoes was done by Pise and Upadhye (2018) where mangoes were graded into three types: yellow mango, green mango, and red mango. The authors claimed that in an experiment conducted on two machine learning methods Posterior analysis gives more accuracy as compared to Naive Bayes and Support Vector Machines (SVM). Another automated grading of mango was suggested by Raghavendra and Rao (2016), where grading was done with the help of maturity detection. The proposed system uses the k-means algorithm for the grading of harvested mango after the Red, Green, Blue (RGB) to Hue Saturation Value (HSV) color scale conversion for better object detection. Mangoes were graded into five classes and the overall accuracy of the system was claimed to be 84%.

A non-intrusive method was applied by Semary et al. (2015) for internal defect detection in mango fruit. The proposed method was a very difficult process. This method identified the physiological disorders by using Near-Infrared Imaging.

Human inspection is a time-consuming process while chemical detection involves the destruction of fruits. Rivera et al. (2014) applied a system based on hyperspectral images to mechanically evaluate induced damage at different ripening stages of mango. To the best of our knowledge, there is no automated system for mango cultivar classification and grading based on the defined export standards in Pakistan. In this manuscript, we performed comparison of three models of CNN for mango cultivar classification and grading into three classes: Extra Class, Class I, and Class II, according to the standardized protocol of the international market defined by Codex Standards (Francis, 1998). These standards are currently being followed by Pakistan for mango exports.

Material and Methods

The graphical description of the methodology is presented in Figure 3. Section 3.1 presents image acquisition and pre-processing, Section 3.2 describes Convolutional Neural Network (CNN) architectures and classification of an image into one of the eight mango cultivars: Chaunsa Summer Bahisht (S.B.), Chaunsa (Black), Chaunsa (White), Dosehri, Langra, Sindhri, Anwar Ratool, and Fajri. Section 3.3 presents the grading of the classified image into three grades: Extra Class (superior quality mango), Class I (slight tolerance of defected skin or color), and Class II (defected mangoes).

Image Acquisition and Pre-processing

The images of eight common cultivars of mangoes were collected from Haji Ghulam Muhammad Mangana mango orchard located in Multan, Pakistan for database collection. Based on horticulture expert opinion, the best time for image acquisition was decided to be around twenty days after starting the stone-hardening stage of mango fruit to one week

Before the start of mango harvesting. The development cycle of each cultivar slightly differs from the other. Therefore, the image acquisition has been carried out between June to August 2020 to get images of the selected cultivars. Images were captured with a white background using a Nikon 7000 camera between 9:00 AM to 12:00 PM in the natural light intensity. The camera was fixed on a tripod to control its movement to avoid blurred and fuzziness of images. The dataset was divided into two parts: classification dataset and grading dataset. For classification purposes, 200 images of each mango cultivar were captured totaling a dataset of 2400 images. A sample of 'Dosehri' mango images from the classification dataset is shown in Figure 4. For grading, another 200 images of each grade of mangoes
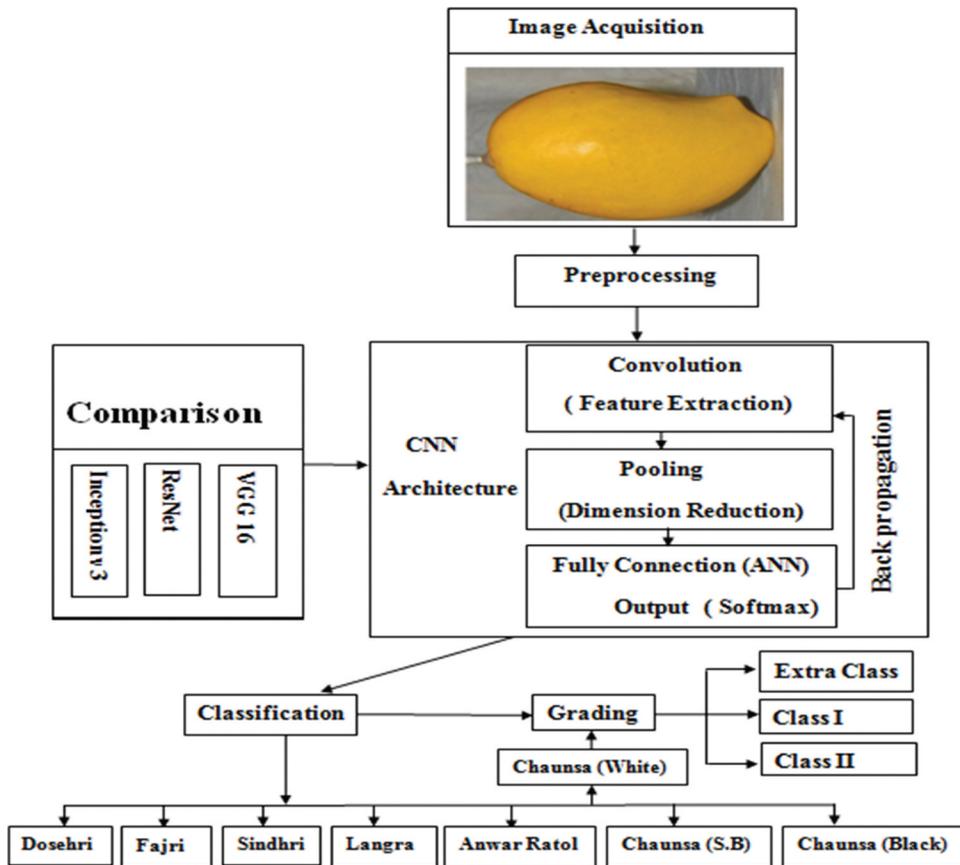
Figure 3. Graphical description of methodology. Major steps involve image acquisition, pre-processing, classification into one of the eight Mango cultivars and then grading according to the international Mango quality standards. CNN refers to convolutional neural networks, ANN refers to artificial neural network, Chaunsa (S.B) refers to the cultivar Chaunsa Summer Bahisht.

Figure 4. Images of 'Dosehri' Mango at different angles to train system for classification.

were captured totaling 600 images per cultivar. A sample of 'Chaunsa' (White) images from the grading dataset are shown in Figure 5. The complete dataset of classification and grading is available online.

Figure 5. Images of 'Chaunsa' (White) captured for Mango grading into 'Extra Class,' 'Class I,' and 'Class II.'

Pre-Processing

All images were resized to 280 × 260 that was the optimal size to input into the classifier. The images greater or smaller than this size adversely affected the results. The images have been captured using a white background to make the segmentation process easy. Edge-based segmentation was used to find the edges of mango. Sobel filter (horizontal) was used to detect horizontal edges, Sobel Filter (vertical) was used to detect vertical edges. Figure 6 shows sample segmented images of 'Dosehr*i*' mango.

Data Augmentation

It is one of the techniques that can be used to fix the overfitting problem. This technique improves the generalization ability of models by enhancing the size of the dataset artificially. In deep learning, the size of dataset directly influences the performance of model; training of deep learning models on more dataset make the models more efficient. As a large dataset is not always available for tainting the model, therefore, existing data was augmented to obtain a generalized model.

The specific data augmentation techniques are chosen within the context of the training dataset which is provided by *Keras* deep learning library *ImageDataGenerator*. The data augmentation techniques used for image pre-processing are images rotation, translation, zooming, shearing, and horizontal flip (as shown in Figure 7).
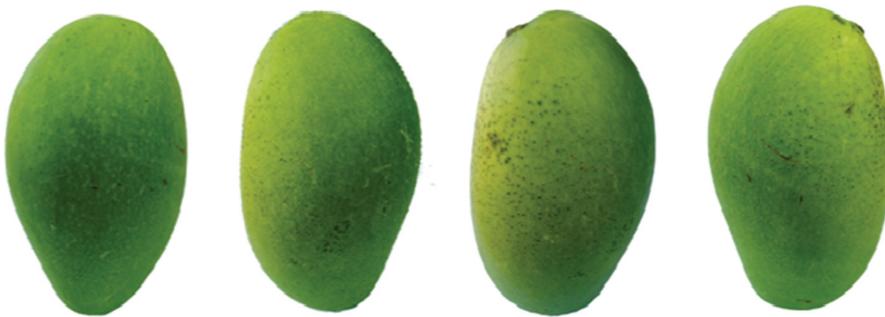
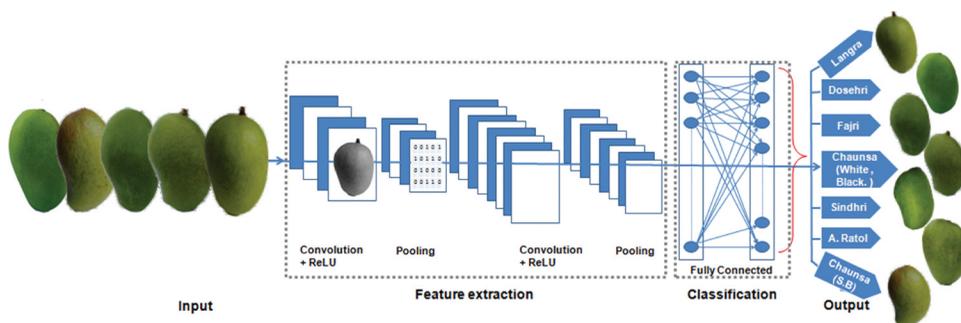**Figure 6.** Segmented images *of Dosehri* Mango cultivar.

Figure 7. Augmented images created by ImageDataGenerator.

Rotation: Images are rotated randomly between 0 and 360 degrees. Images with different positions of the object are produced by rotating a base image clockwise or anticlockwise at a specified angle. Each resulting image is unique datum that is used for training the model. We have set the rotation angle to 40°.

Translation

This method shifts the image into areas either along the x-axis or y-axis (or both). We have set a magnitude rate of 0.3 for translation.

Zooming

A zoom augmentation method added new pixel values around the base image randomly to expand the training dataset by creating various sizes of images. The image can be zoomed-in or zoomed-out by the **zoom_range** argument in the **ImageDataGenerator**. We had set zooming augmentation in the range of [0.8, 1.2], i.e., between 80% (zoom in) and 120% (zoom out).

Shearing: In shearing, the images were stretched along with the x-axis or y-axis. Image shearing refers to shifting one part or selection of image to one direction and the other part to the opposite direction. We had used a shear value of 0.2 to ensure class preservation.

Flipping

Flipping augmentation rearranged an image while preserving the features of the image. An image could be flipped horizontally or vertically depending upon the type of image. To use these augmentations, we set horizontal_flip = True and vertical_flip = True in ImageDataGenerator constructor.
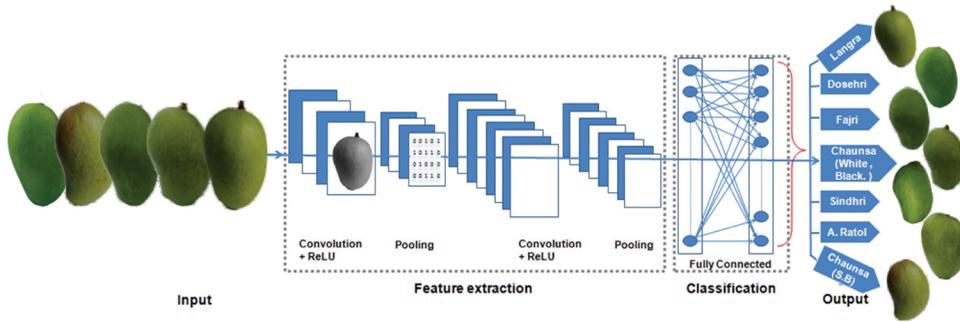
Classification

CNN was applied for mango's cultivar classification and grading that applies a particular geometric knowledge according to the constructed architecture to learn the features based upon the local region. The shared weights, sub-sampling, and local receptive fields enable CNN for backward propagation (supervised learning) to train deep architectures (LeCun et al., 1999). A CNN is a highly layered structural neural network, generally consisting of

The following three layers. Figure 8 illustrates an overview of the recognition and classification process of mango cultivars using convolutional neural networks.

Convolutional Layer

In the first layer of CNN, convolution is the core operation that performs image processing to detect unique features while keeping the relation between pixels; convolution is an integral transformation operation. Corresponding convolution kernels were used to detect certain features. Convolution kernels are small matrices that are applied to the original image to perform certain tasks as blurring, sharpening,

Figure 8. Three-layer CNN architecture consisting of convolutional + rectified linear unit (ReLU), pooling, and classification layers. For details, see Section 3.2.

or detecting specific features. To perform one convolution operation, a kernel convolved with the input matrix horizontally (row-wise) and vertically (column-wise). The convolution kernels are smaller in size as compared to the input image, and kernel parameters remain constant in the complete process of recognition. This guides the weight-sharing scheme for each receptive neuron (Liu, 2018).

In this layer, an image in the form of a matrix was treated as input and convolved with the corresponding filters. As the result of convolution operation, a feature map was obtained, a bias value was added to each element of the feature map (resulting matrix) and the sum of all feature maps was computed. The obtained output is used as input for the next layer.

Pooling Layer

After the convolution operations, pooling was performed in the next layer of CNN to reduce the dimension of the inputs. Pooling also speeds up the model by reducing the computation. In the pooling layer, the kernel was slid over the input (output of the previous layer) to record maximum value. The commonly used pooling type is maximum pooling which takes maximum value from the pooling window. In inception v3 architectures, convolution was performed with a 3×3 sized kernel, using stride 1 with the same padding (Nivrito et al., 2016).

Fully connected layer

In this layer, images were classified into categories. The operations of previous layers (convolution and pooling) reduced the size of the input by reserving the meaningful features in matrix form. A fully connected layer requires data in a one-dimension vector form to classify objects in labeled classes. Therefore, flattening was performed before this layer to produce output in a one-dimension vector form. Some operations like drop outing neurons, condensing, and softmax were performed in this layer to produce output in the desired classes.

The inception network is the state-of-the-art of deep learning architecture to solve the problems related to image recognition and detection. In this study, the Inception v3 module had been implemented using *Keras* with *TensorFlow* libraries as backend to perform classification and grading of mango cultivars. The classifier was trained using the *FitGenerator* method and ImageNet weight for transfer learning. The input image format for this model is 299×299. Inception v3 architecture is based on the asymmetric building blocks and symmetric building blocks, which are comprised of convolutions, activation function rectified linear unit (ReLU), average pooling, maximum pooling, concatenation, dropouts, and fully connected layers (Shlens, 2016).

In this architecture, small (effective) number of neurons and bottleneck layer (1×1 convolutions) were used to reduce the computation requirement. Inception architecture's main purpose is to extract the feature at multi-levels. This architecture convolutions of *1×1, 3×3*, and *5×5* are computed together in the same module of the network.

Dropout is used to avoid overfitting and proved helpful to increase the validation accuracy, inhere dropout was set as 0.2. Data augmentation was performed to give support to the model in terms of memorizing images with different aspects (Wu et al., 2015) which helped to reduce over-fitting. *Keras* contain the class *ImagedDataGenerator* that is used for images rotations, translations, zoom, shearing, and flipping.

Hyper-parameter tuning

These are variables related to the training process and optimization that determines the CNN structure. As the best numbers of hyper-parameters depend on the nature of each task and each dataset, there is a need to set them manually to get the best-required result. There is no magic number for hyper-parameter tuning that works everywhere. We achieved the best results at a 0.001 learning rate with a batch size of 32. Learning rate ranges between 0.0 and 1.0 that refers to the hyper-parameter that controls how much to change the model in response to the test loss of classification with changing weights. Test loss of classification refers to the loss calculated by taking the mean squared difference between the actual and predicted values. In our model, high classification accuracy and low test-loss was achieved at learning rate of 0.001 that means the model requires relatively more training epochs given smaller changes to weight to get optimal results. Training time is measured in minute:second.

Batch size is a set of the sample used in one iteration; batch size of 32 was selected by keeping in mind every effect of large batch size. The best epoch number found was until 50, where epoch refers to the number of times the neural network is trained with complete dataset, and momentum was set as 0.5 for optimal training speed and accuracy of the proposed model. Complete results are listed in Table 1.

There are several models of CNN algorithm which have different number of layers and structure; each has its specialty and performs better based on nature of dataset and their application. In this paper, the performance and accuracy of three of the CNN models namely Inception v3, VGG16, and ResNet152 was compared on the proposed dataset. Inception v3 model was widely used for image classification, in this architecture small number of neurons, and bottleneck layer (1×1 convolutions) are used to reduce the computation requirement. The advantage of this architecture is that multi-level feature is extracted in such way that the performance of the network is improved. For example, filter of (5x5) extracts general features and local features by (1x1) at the same time. In VGG16 convolution and max pooling, layers were arranged consistently throughout the whole architecture and two fully connected layers followed by a *softmax* were used for output at the end. Convolution layers used kernels of 3×3 sizes with a stride 1, always used the same padding and the max-pooling layer has 2×2 kernels with stride 2. The VGG16 is a deeper architecture as a large number of hyper-parameter are used here which slows down the training of the model and requires more computational resources (Wahed and Nivrito, 2016).

Table 1. Results achieved by hyper-parameter tuning at 85% and 15% training dataset and testing dataset respectively, at 50 Epochs with 32 batch size. Learning rate is measured in the range of 0.0 to 1.0. In our model, high accuracy and less test loss was achieved at learning rate of 0.001 that means the model requires relatively more training epochs given smaller changes to weight to get optimal results. Training time is measured in minute:second.

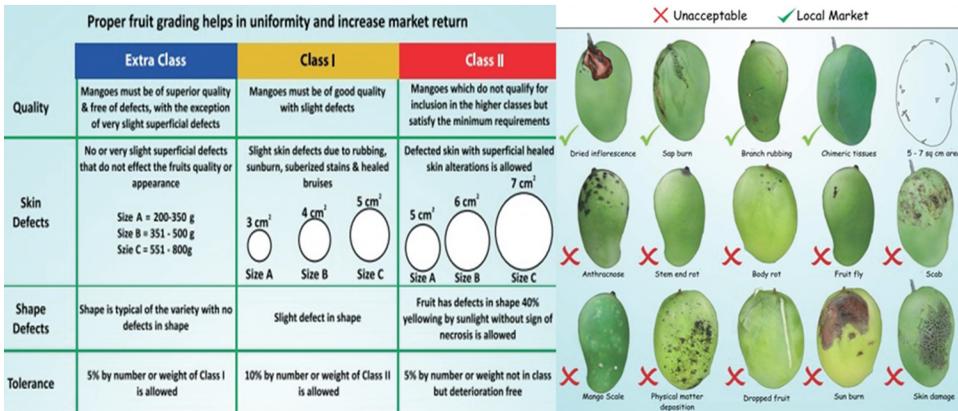| Cultivar	Per class accuracy of classification	Test loss of classification	Learning rate	Training time(minute:second)
Dosehri	98.5%	0.024	0.001	3:52
Sindhri	100%	0.001	0.001	3:52
Anwar Ratol	100%	0.001	0.001	3:52
Fajri	100%	0.001	0.001	3:52
Chaunsa White	98.2%	0.024	0.001	3:52
Chaunsa Black	100%	0.001	0.001	3:52
Chaunsa S.B	98.5%	0.024	0.001	3:52
Langra	98.1%	0.024	0.001	3:52

Figure 9. Mango grading standard. Source: Trade Related Technical Assistance program in Pakistan (TRTA-II, 2015).

ResNet152 used global average pooling instead of fully connected layers which reduced the model size up to 102MB. Another thing in this architecture is that all 150 layers stack periodically and double the number of filters with stride two, and finally fully connected layer is for output the classes (Khan et al., 2020).

ResNet covered the drawbacks of the VGG16 module by using a shortcut connection to skip the unnecessary layer, which reduced the computation cost and improved the performance. The accuracy achieved by each architecture is compared in Tables 2 and 3.

Mango Grading

To improve the value of mangoes in the international market, there is a need to adopt standardization norms in Pakistan (Mustafa et al., 2006; Ghafoor et al., 2018). Grading of mangoes is performed by external inspection of mango size, color, and texture. According to UN/ECE Standards mangoes are divided into three grades (Figure 9); *Extra Class*: Mangoes of this grade are of superior quality and defects free or very slight defects, that do not affect the mango quality and appearance. 5% tolerance by number or weight is allowed in the extra class. *Class-I*: This class holds good quality mangoes while slight skin defects due to rubbing, sunburn, suberized stains are acceptable. 10% tolerance ($3cm^2$-$5cm^2$) by number or weight of such defects is allowed in class I. *Class-II*: Mangoes of this grade do not qualify for inclusion in the higher classes but satisfy the minimum requirements. Defected skin with superficial healed skin is allowed ($6cm^2$-$8cm^2$), and mango with defects in shape 40% yellowing by sunlight without a sign of necrosis is allowed (Francis, 1998). In this paper, we are proposing an automated system using CNN for grading the images of classified mango cultivars as per grading standards. The results of classification and grading have been discussed in the next section.

Results and Discussion

The proposed work is an attempt to make an efficient model for mango classification and grading. A comparative analysis was conducted between various architectures (VGG16, ResNet152, and Inception v3) of CNN, the inception v3 module performed better results on the used dataset.

The model was fine-tuned at various hyper-parameter values to attain maximum accuracy. Dataset was split into 85% and 15% training data set and testing dataset respectively, at 50 epochs and 16 batch sizes model achieved 96% test accuracy and 0.2 test loss of mango cultivar classification. The test accuracy was increased from 96% to 98% by increasing the epoch from 50 to 70 and the batch size from

16 to 32. It was observed that epoch and batch size have a great effect on decreasing loss from 0.2 to 0.16 and a significant increase in accuracy. Furthermore, accuracy started to decline by increasing the epoch from 70 to 100 at batch size from 32, test accuracy declined to 97.12% and test loss increased to 0.19.

It was observed that the model was overfitting by increasing the epoch, therefore epochs were not increased beyond the 100 epochs to overcome the overfitting problem. Finally, by testing the model at various hyper-parameter values the best results of 99.16% test accuracy and 0.1 test loss were achieved at 50 epochs with 32 batch sizes as listed in Table 1. The highest test accuracy and minimum loss per class of mango cultivar at 50 Epoch with 32 Batch sizes that model achieved for mango cultivars classification is illustrated below in Figure 10.

The mango grading dataset was also split into 85% and 15% training dataset and testing dataset respectively. The proposed model was tuned at various hyper-parameter values for best results, the model achieved 96.66% accuracy and 0.2 test loss at 50 epochs with 32 batch size, as shown in Figure 11. An increase in the number of epochs from 50 to 70 and batch size from 16 to 32 caused the increase in test loss to 0.31 and decreased test accuracy to 94.32%.

In a few cases, the model was confused in two grades (Class I) and (Class II) due to the similarity of appearance of defected mangoes in both grades that caused the overfitting problem. To avoid the overfitting problem, the number of epochs was not increased further, and the model is re-tuned at 50 epochs with 32 batch size. This technique helped in improving the accuracy from 94.3% to 96.7% and declining the test from 0.31 to 0.19.

Confusion Matrix for Classification Results

The confusion matrix evaluated the recall, precision, and F-1 Score of the model prediction. Recall is the ratio of correctly predicted images to all available images in the dataset, precision is the ratio of correctly predicted images to the total predicted images, and F1 score is the weighted average of
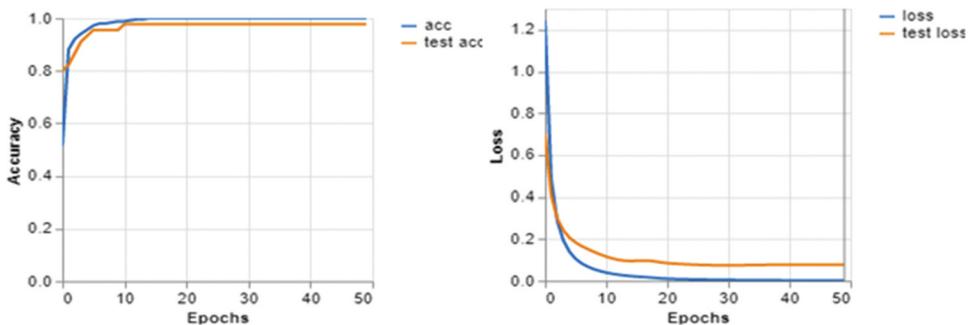
Figure 10. Test accuracy and test loss achieved at 50 epoch with 32 batch **size** for Mango cultivar classification.
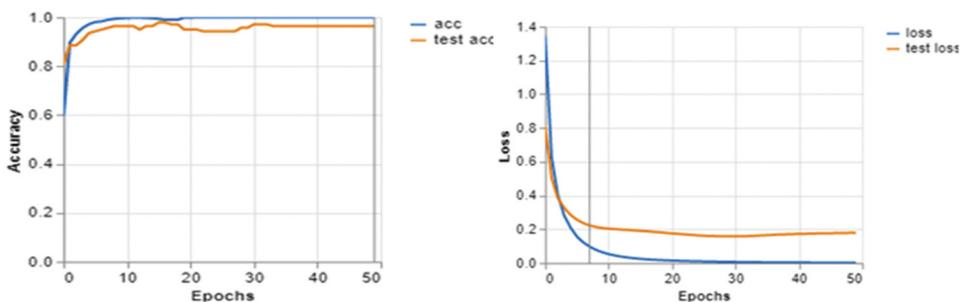
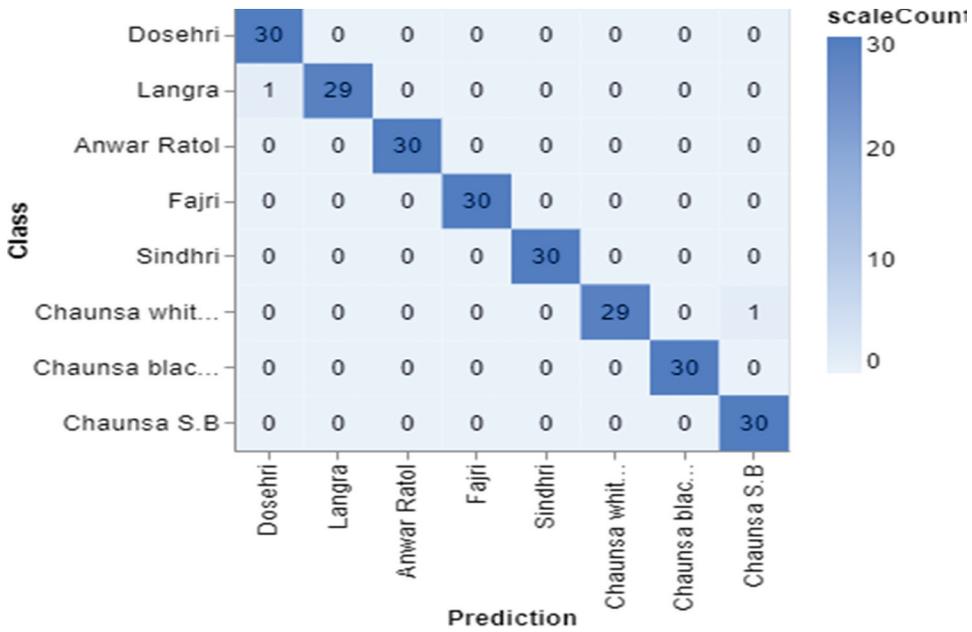**Figure 11.** Test accuracy and test loss achieved for Mango grading.

Figure 12. Confusion matrix for Mango cultivar classification.

precision and recall. All of these ratios are measured in the range of 0.0 to 1.0 where 1.0 refers to the maximum score. In Figure 12, Y-axis shows the actual class whereas X-axis shows the model's prediction against the respective class. The model outputs one false negative and one false positive, such as it incorrectly predicts 'Dosehr*i*' instead of the actual mango cultivar 'Langra' due to oval-shaped appearance of both cultivars. The model was also confused in *Chaunsa S.B* and *Chaunsa White* due to the apparent similarity of both mango cultivars. Table 2 shows mango classification results obtained by using Inception V3, ResNet 152, and VGG 16.

Confusion Matrix for Grading Results

In Figure 13, it has been observed that models are mostly confused in two grades, Class-I and Class-II, due to similarity in appearance of defected mangoes in both grades. The model predicted the wrong Class-I instead of actual mangoes of Class-II and vice versa. The grading model achieved 96.7% average accuracy. Table 3 shows mango grading results obtained by using Inception V3, ResNet 152, and VGG 16.

Conclusion and Future Work

This paper presented a comparison of three architectures of Convolutional Neural Network *for mango cultivar classification and grading*. Through experiments, we found that Inception V3 performed better than the other two architectures on a dataset of eight cultivars of mango. The reason for choosing CNN is that there is a minor difference in the appearance of some mango cultivars such as Chaunsa (white), Chaunsa (Black), Chaunsa (S.B.) that makes it harder to extract

Table 2. Mango classification results.

Model	Recall	Precision	F-1 Score
InceptionV3	**0.99**	**0.99**	**0.99**
ResNet 152	0.98	0.97	0.98
VGG 16	0.97	0.95	0.96
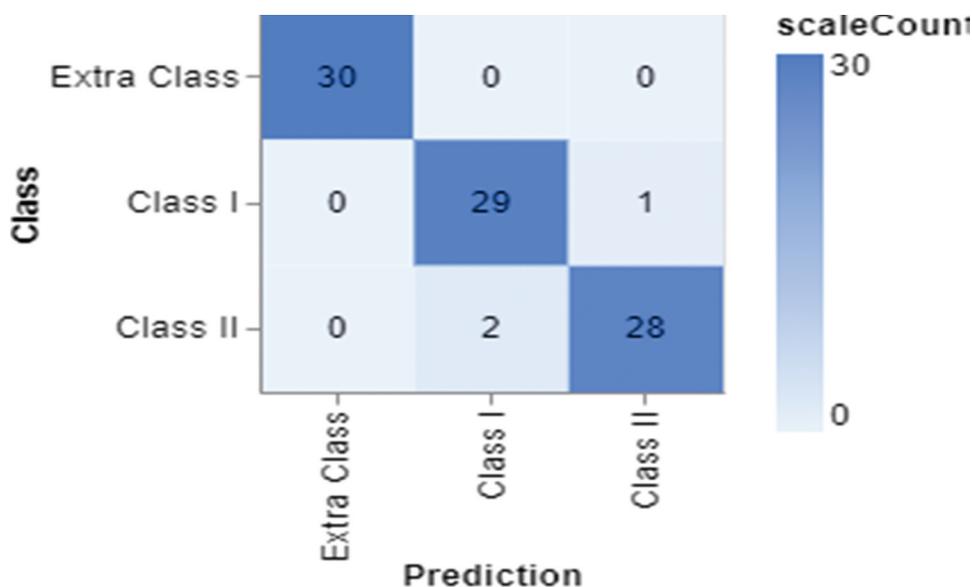
Figure 13. Confusion matrix for Mango grading.

Table 3. Mango grading results.

Model	Recall	Precision	F-1 Score
InceptionV3	**0.97**	**0.97**	**0.96**
ResNet 152	0.95	0.93	0.94
VGG 16	0.95	0.91	0.93

features to be used in traditional machine learning classifiers. The CNN is used to learn high-level features from data itself in an incremental manner, as the nested layers in the neural networks put data through hierarchies of different concepts, which eventually learn through their errors. This eliminates the need for hard-core feature extraction as needed in traditional machine learning techniques (O'Mahony et al., 2019).

Deep learning has provided an optimal solution for mango classification into various cultivars and grading. CNN learns important features of images like shape, texture, color, size, and defects automatically which reduces the error caused by manual classification and grading as in traditional machine learning techniques. In the future, we will update the system to classify mango images on trees before harvesting. Feature extraction will be improved using the method presented by (Zhang et al., 2021) and relevant techniques.

The advancement in computer vision technologies and availability of low-cost hardware can replace labor work for mango classification and grading in Pakistan as automated classification and grading system not only speeds up the time of the processing but also minimizes error and reduces the cost. The proposed model can be implemented at a large scale for automated grading of mangoes according to the International standards to enhance mango export in Pakistan.

Acknowledgments

Disclosure Statement

Funding

Data Availability

The dataset is available online at Mendeley Data: Rizwan Iqbal, Hafiz Muhammad; Hakim, Ayesha (2021), "Mango Cultivar and Grading Dataset," Mendeley Data, v1. http://dx.doi.org/10.17632/5mc3s86982.1

References

TRTA II: Trade Related Technical Assistance Program. 2015. http://trtapakistan.org

The Dawn News. 2017. June 04. https://www.dawn.com

Abbas, Q., M.M. Iqbal, S. Niazi, M. Noureen, M.S. Ahmad, M. Nisa, and M.K. Malik. 2018. Mango classification using texture & shape features. Int. J. Comput. Sci. Netw. Secur. 18:132–138.

Ali, M.A., and K.W. Thai. 2017. Automated fruit grading system. 2017 IEEE 3rd International Symposium in Robotics and Manufacturing Automation (ROMA), 19-21 September 2017. Kuala Lumpur. IEEE 1–6.

Anurekha, D., and R. Sankaran. 2020. Efficient classification and grading of MANGOES with GANFIS for improved performance. Multimed. Tools Appl. 79(5–6):4169–4184. doi: 10.1007/s11042-019-07784-x.

Arab News. 2019. June 16. https://www.arabnews.pk/

Francis, J. 1998. Product profile: Mango (Mangifera indica L.). Tropical Fruits Newsletter. (Mar). 13–15.

Ghafoor, A., A. Adeel, and A. Maqbool 2018. Report of mango farm survey in Punjab.

Khan, A., A. Sohail, U. Zahoora, and A.S. Qureshi. 2020. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 53 (8):5455–5516.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1999. Shape, Contour and Grouping in Computer Vision Lecture Notes in Computer Science. 1681. Berlin, Heidelberg: Springer. doi:10.1007/3-540-46805-6_1.

Lihuan, S., W. Liu, Z. Xiaohong, H. Guohua, and Z. Zhidong. 2017. Fabrication of electronic nose system and exploration on its applications in Mango fruit (M. indica cv. Datainong) quality rapid determination. J. Food Meas. Charact. 11(4):1969–1977. doi: 10.1007/s11694-017-9579-1.

Liu, Y.H. 2018. Feature extraction and image recognition with convolutional neural networks. First International Conference on Advanced Algorithms and Control Engineering, 10-12 August, 2018, 1087. IOP Publishing Ltd. Taiwan, 062032. doi:10.1088/1742-6596/1087/6/062032.

Mangan, T., and U. Ruthbah. 2018. Draft report on mango farm survey in Sindh, Pakistan.

Mustafa, K., A. Mahmood, and B. Ahmad. 2006. Barriers against export of mango from Pakistan: Role of sanitary and phytosanitary agreement of WTO. Research Report. Pak. Sci. Found.

Naik, S., and B. Patel. 2013. Usage of image processing and machine learning techniques in agriculture-fruit sorting. Mumbai, India: CSI Communications.

Naik, S., and B. Patel. 2017. Thermal imaging with fuzzy classifier for maturity and size based non-destructive Mango (Mangifera Indica L.) grading. 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), Feb 3-5, 2017. Pune, India. IEEE, 15–20.

Naik, S., B. Patel, and R. Pandey. 2015. Shape, size, and maturity feature extraction with fuzzy classifier for non-destructive Mango (Mangifera Indica L., cv. Kesar) grading. 2015 IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR) 10-12 July 2015, Chennai, India. IEEE, 1–7.

Nivrito, A., M. Wahed, and R. Bin. 2016. Comparative analysis between Inception-v3 and other learning systems using facial expressions detection. Dhaka, Bangladesh: BRAC University, 48.

O'Mahony, N., S. Campbell, A. Carvalho, S. Harapanahalli, G.V. Hernandez, L. Krpalkova, D. Riordan, and J. Walsh. 2019. Deep learning vs. traditional computer vision. Science and Information Conference March 16-19, 2019 Tokyo, Japan. Springer, 128–144.

Pakistan Today. 2020, June 18. https://www.pakistantoday.com.pk

Pandey, R., S. Naik, and R. Marfatia. 2013. Image processing and machine learning for automated fruit grading system: A technical review. Int. J. Comput. Appl. 81(16):29–39. doi: 10.5120/14209-2455.

Pise, D., and G. Upadhye. 2018. Grading of harvested mangoes quality and maturity based on machine learning techniques. 2018 International Conference on Smart City and Emerging Technology (ICSCET), 5 January 2018. Mumbai, India. IEEE, 1–6.

Planning Commission of Pakistan. 2020. https://www.pc.gov.pk/

Raghavendra, A., and M. Rao. 2016. A survey on internal defect detection in fruits by non-intrusive methods. Int. J. Latest Trends Eng. Technol. 6:343–348.

Rivera, N.V., J. Gómez-Sanchis, J. Chanona-Pérez, J.J. Carrasco, M. Millán-Giraldo, D. Lorente, S. Cubero, and J. Blasco. 2014. Early detection of mechanical damage in Mango using NIR hyperspectral images and machine learning. Biosyst. Eng. 122:91–98. doi: 10.1016/j.biosystemseng.2014.03.009.

Semary, N.A., A. Tharwat, E. Elhariri, and A.E. Hassanien. 2015. Fruit-based tomato grading system using features fusion and support vector machine. Intelligent Systems Conference November 10-11, 2015 London, UK. Springer, 401–410.

Shlens, J. 2016. Train your image classifier with inception in TensorFlow. electronic resource], https://research.google blog.com/2016/03/train-your-image-classifier-with.html.

SMEDA. 2020. https://smeda.org/

Thong, N.D., N.T. Thinh, and H.T. Cong. 2019. Mango classification system uses image processing technology and artificial intelligence. 2019 International Conference on System Science and Engineering (ICSSE), 20–21 July 2019. Dong Hoi City, Quang Binh Province, Viet Nam. IEEE, 45–52.

Wahed, R.B., and A. Nivrito. 2016. Comparative Analysis between Inception-V3 and Other Learning Systems using Facial Expressions Detection. Thesis.

Wu, Q., Y. Qu, P. Gong, T. Wang, T. Gong, and Y. Man. 2015. Evaluation of the efficacy of keratinized mucosa augmentation techniques around dental implants: A systematic review. J. Prosthet. Dent. 113:383–390.

Zhang, C., S. Bengio, M. Hardt, B. Recht, and O. Vinyals. 2021. Understanding deep learning (still) requires rethinking generalization. Commun. ACM 64(3):107–115. doi: 10.1145/3446776.

Zhang, Y.-D., Z. Dong, X. Chen, W. Jia, S. Du, K. Muhammad, and S.-H. Wang. 2019. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. Multimed. Tools Appl. 78 (3):3613–3632. doi: 10.1007/s11042-017-5243-3.