International Journal of Artificial Intelligence and Mathematical Sciences

December 2025 Vo4. Issue No. 02
ISSN Online: 2958-0153 and Print: 2958-5228
DOI: https://doi.org/10.58921/ijaims.v4i2.176

Multi-Class Guava Classification Using Convolutional Neural Networks for Precision Agriculture

Sajjad Ahmed[1*], Muhammad Juman Jhatial[2], Farhan Bashir[1], Anjum Usman[3], Inam Ali[1], Shahnaz Memon[1]
dr.sajjad@uolrk.edu.pk,juman@uolrk.edu.pk, farhan@uolrk.edu.pk, anjum.usman@sbbusba.edu.pk, inamalichandio08@gmail.com, shahnazmemon50@gmail.com

[1]Department of Computer Science, The University of Larkano, (TUL), Larkana, Sindh, Pakistan

[2]Department of Computer Science, Shah Abdul Latif University, Khairpur Mir's, Sindh, Pakistan

[3]Department of Computer Science, Shaheed Benazir Bhutto University, Sakrand Road, Nawabshah, Sindh, Pakistan

Received: 13-11-2025; **Accepted:** 27-11-2025; **Published**: 30-12-2025

Abstract

Guava (Psidium guajava L.) is an important fruit crop in Pakistan, especially in the province of Sindh where several varieties especially Thadharami, Local Sindhi and Riyali are grown and sold. While this is a common approach, the manual visual inspection approach is subjective, time-consuming, labor-intensive and prone to errors of classification, particularly when guava varieties and maturity stages have similar visual similarities. To overcome these, in this study, an automated image-based guava classification system has been proposed using Convolutional Neural Network (CNN). The proposed system differentiates between both guava variety and maturity stage (green, mature green and ripe). Images of guavas were arranged in nine classes of this data set according to its three local varieties and three maturity stages. The image was preprocessed to ensure uniformity of the input, minimize overfitting and improve the generalization of the model under variable imaging conditions. The CNN model was composed of several convolutional, pooling, and fully connected layers to extract the discriminative spatial, color, texture, and shape features in guava images automatically. The model is trained using categorical cross entropy loss and optimized with Adam optimizer. The experimental results presented here revealed the effectiveness of the proposed CNN for Guava classification by automatic classification with a validation accuracy of 92.5%. The misclassifications were higher for the green-stage classes as compared to the ripe guava classes, which was attributed to the large visual similarity between classes and high number of classes of green guavas. The result shows, CNN-based classification can be reliable, scalable and non-destructive solution for guava grading to ensure better quality control, decrease post-harvest losses, increase efficiency in supply chain of the guava industry in Pakistan.

Keywords: Guava classification; Convolutional Neural Network; deep learning; image processing; precision agriculture; Sindh guava varieties.

*Email: dr.sajjad@uolrk.edu.pk

1. INTRODUCTION

In Pakistan agriculture plays a significant role in the food security, employment in agricultural sector and agro-based development. The horticultural industry is an important sector in the agriculture sector due to the nutritional value and market income it offers, as well as the fruit being used as raw materials in value added products. Guava (Psidium guajava L.) is one of the significant tropical fruit crops grown in Pakistan, especially in the Sindh province where Larkana and Ratodero are famous for the commercial guava crop. Guava has high nutritional value and functional value not only as a food crop, but also as a health promoting fruit due to its richness in vitamin C, dietary fiber, mineral and phenolic compounds, flavonoids and substances with antioxidant activity [1, 2]. In the local farming system of Sindh, the fruit size, peel color, pulp consistency, seed content, taste and market demand are the other distinctions that are well known to growers and traders of these guava varieties namely "Thadharami", "Local Sindhi" and "Riyali". These visual and physical differences are suitable for computer vision-based grading of guava but pose challenges for consistent manual grading.

Classification of fruits in traditional guava marketing system is normally done manually by farmers, laborers, traders or market inspectors. Manual grading is required and based on color, size, shape, texture and maturity stage. Subjectivity, time, labor, and inconsistency make this method simple, inexpensive, but time-consuming and labor intensive, particularly for large amounts of fruit that need to be sorted within a short period of time. This becomes a more critical issue since guava is extremely perishable with a high rate of quality deterioration after harvest. If there is a delay in or error in sorting the product may result in a mixture of green, mature green and ripe fruit reducing the reliability of the grade, impacting consumer satisfaction and causing post-harvest loss. More recent research on quality assessment of fruit indicates that automated visual inspection and non-destructive sensing can minimize the reliance on subjective human assessment and the efficiency of quality assurance systems after harvest [7, 8]. So, there is a need for an automatic Guava Classification System that can help in correct grading and lessening human error and assist in decision making in the local fruit supply chains.

Guava classification is found to be very difficult due to intra-class variation and inter-class similarity with the types grown locally. Intra-class variation in color, size, surface texture, shape, lighting condition and image background arises from variation in the orchard environment, growth conditions, maturity stage and image acquisition setup, from fruits of the same variety. Inter class similarity is when two or more varieties or maturity stages are alike with regard to any visual characteristic such as green Local Sindhi and green Riyali fruits. Further, early maturity stages are problematic as green and mature green guavas could only have slight color and texture differences. These difficulties make it hard to perform well on standard rule-based image-processing techniques depending on handcrafted features like colour thresholds, edge descriptors, shape features or texture statistics. These can work effectively only under a set of ideal imaging conditions and might not work for images taken in a natural market or farm setting.

With the support of machine learning and deep learning technologies, the image analysis of agricultural products has made substantial progress, allowing for automated classification of different products, quality grading, freshness detection, maturity prediction, and disease identification. Unlike conventional image processing, deep learning models can learn discriminative features from image data itself. Since ConvNets (CNNs) are very successful in image classification tasks, they enable them to extract spatial and hierarchical features by convolution, activation, pooling, and fully connected operation. Low level features like edges, color gradients, and local textures are learned in earlier layers, then high level features like the shape of the fruit, the texture of the peel, the color changes that occur as a fruit mature etc. are learned in deeper layers. These feature the ability of CNNs to handle complex agricultural classification tasks that may not be able to be satisfactorily accomplished by handcrafted features [9]– [12].

Recent works have shown the success of CNNs and Transfer Learning models in fruit Classification and Quality checking. Gill et al. [9] introduced deep learning and feature fusion technique for fruit type classification and Gulzar [10] employed transfer learning with MobileNetV2 for fruit image classification. Shahi et al. [11] used an attention-based MobileNetV2 model for industrial fruit classification, demonstrating that attention mechanisms can enhance the feature representation. For apple quality identification and classification, Li et al. [12] used CNN-based image processing, showing the capability of CNNs to learn features of apples images. Likewise, the application of CNNs and transfer-learning methods has been used in the fruit freshness classification [13], automatic fruit grading [14], hyperspectral blackberry ripeness detection [15] and freshness grading using computer vision and deep learning [16]. These studies are consistent with the findings that deep learning holds promise for accurate and scalable solutions in agricultural visual inspection.

While the research on fruits has increased significantly over the last years, it is still not as much as research on apples, bananas, tomatoes, citrus, mangoes and dates has been. Maitlo et al. (2012) [3] developed a dataset of images of guavas which includes local guava varieties and maturity stages of guavas suitable for guiding research in automated guava grading in Pakistan. Toptaş and Güven [4] explored guava fruit classification via CNN architectures like GoogLeNet, VGG19, ResNet50 and DenseNet201. Kumari and Singh [5] proposed guava quality classification model by using ANOVA and machine-learning techniques, Haq et al. [6] studied guava maturity evaluation using thermal imaging method with machine-learning and generative-AI-based classifiers. Analyses of these studies show an escalating interest in guava classification, but a lack of guava classification research that accounts for both variety and maturity stage for guava classification based on RGB-image and CNN based model is still present.

Two-level classification of guava variety and maturity is significant for practical application in agriculture. Different varieties need to be classified to accommodate variances in preferences of the market, their physical attributes and consumer acceptance, due to the different varieties of Thadharami, Local Sindhi, and Riyali. Maturity classification is also important as it will mean a different firmness, taste, storage life, transport suitability and price to the guava depending on whether it is green, mature green or ripe. No fruit type classification + ripeness estimation gives minimal support for the market grading. As a result, this study considers guava classification as a multi class image-recognition problem, consisting of nine classes which are composed of three maturity stages and three local cultivars. The proposed system can be useful practically to be graded in the required manner as it is a dual-level system that can give detailed information for sorting, packing, pricing, storing, and supply-chain management.

CNN is chosen for this task to have the ability to learn visual features from images of guava without the need for manual feature engineering. CNN model performs convolutional operation on each input image to learn feature maps, pooling operation on learned feature maps to decrease the dimension of the spatial information and to provide robustness to translation in the image, and fully connected layers to map extracted features to output classes. CNNs are able to learn more complex non-linear relationships between image patterns and class labels, as opposed to shallow classifiers like Support Vector Machines, Random Forests, and k-Nearest Neighbors. Changes occurring during maturity are important in guava classification as they are seen as minor variation in peel color, surface texture, shadow and fruit contour. Optimized and lightweight CNN models have also been explored for fruit maturity detection and real time applications in agriculture [17].

The quality of the image and the distribution of the data significantly impact the performance of the CNN-based fruit classification system, making preprocessing imperative. Image resizing makes the sizes of the images consistent, normalization scales the pixel values within a consistent numerical range, and data augmentation can be done through image transformations (rotation, flipping, zooming, shifting, brightness variation, etc.). These operations lead to generalization of the model for variations in orientation, lighting and background in real fruit image. Another advantage from image augmentation is its effectiveness in the case of limited and imbalanced data sets, as it increases the volume of training data and prevents overfitting [18]. Furthermore, there are several hyperparameters, such as learning rate, number of epochs, batch size, optimizer, network depth, activation functions and dropout rate, that affect CNN performance. Therefore, the selection of the hyperparameters can be meaningful to stable convergence and prevent the model from being underfitted or overfitted [19].

Class Imbalance is another important technical problem involved in Guava image classification. In real datasets some classes may have more images represented than others since some varieties or maturity stages are more available when collecting the data. For instance, the ripe guava samples might be more easily locally available in the market, whereas green or mature green samples are not as well represented. If not addressed properly, the class imbalance issue can cause the model to be biased towards the majority class and provide high overall model accuracy despite poor performance on the minority class. Thus, model evaluation is not just based on accuracy. The analysis of confusion matrix, precision, recall, F1-score and class-wise performance should be done to find out the classes which are misclassified most of the time [20]. These are important points of the current study since the guavas of the green stage might present with subtle visual differences and a higher risk of misclassification.

This proposed CNN classification system is useful to precision agriculture and post-harvest automation as it would give a consistent approach to classify the guavas using data. Non-destructive fruit quality detection technologies are coming into significance since they permit us to assess fruit maturity, fruit damage, disease and fruit quality in a rapid manner without fruit damage [21]. AI and machine-learning processes are also growing in significance for food quality management for the advantages they provide in automation inspection, support manual work, and enhance decision consistency [22]. Implementations of machine learning and deep learning in agriculture have been explored more broadly, with AI systems being able to help with crop monitoring, quality assessment, disease detection, yield estimation and supply-chain efficiency [23, 24]. There are also recent systematic reviews of artificial vision systems that indicate computer vision, CNN, RGB imaging, controlled illumination, transfer learning, and segmentation techniques are prevalent in fruit inspection and classification systems [25].

Thus, the following paper suggests multi-class system based on CNN for classification of local varieties of guava grown in the region of Sindh, Pakistan. The study is targeted on Thadharami, Local Sindhi and Riyali guavas when they are in green, mature green and ripe stages. This system tries to lessen the manual processing constraint by learning the discriminative features of the guava visually. The proposed method has the advantage of tackling the classification at the same time for variety level and maturity level, which is also a more comprehensive grading solution than one-shot fruit classification. The anticipated outcome of this research is a technically feasible, scalable, and practically applicable image-based classification system that can be beneficial for farmers, traders, and post-harvest operators by aiding in better quality control, loss reduction, and enhancing the guava value chain in Pakistan.

2. RESEARCH METHODOLOGY

Figure 1 shows flow chart of the research methodology, which includes methodology of guava fruit classification will involve collecting a dataset of labeled images, data pre-processing (using normalization and augmentation), training a convolutional neural network (CNN) and then evaluating the model performance using the standard classification metrics.

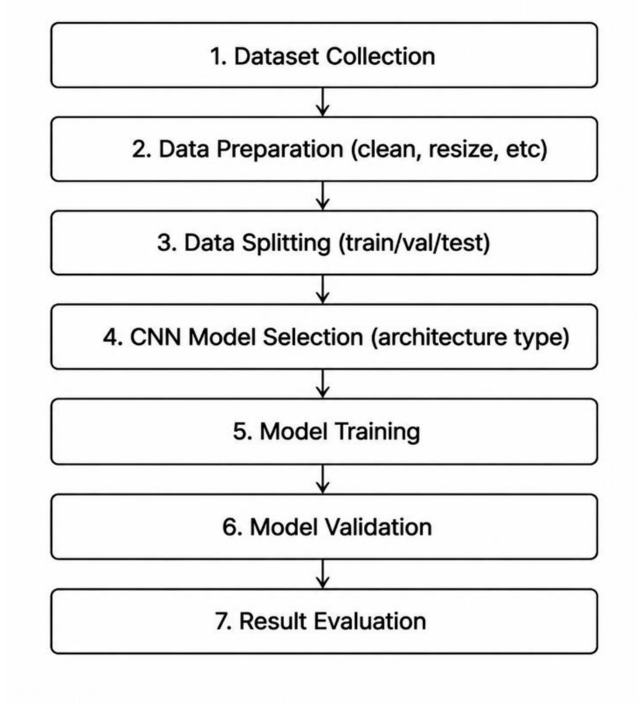

Fig 1. Research methodology flow chart

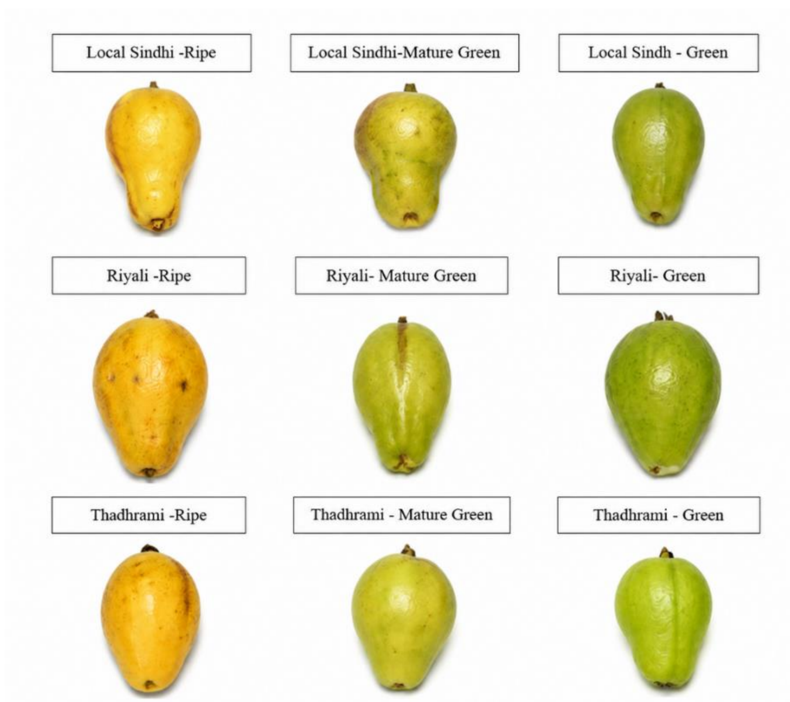

Fig 2. Images of nine classes of guava

Figure 2 shows the guava fruit physical structure. Guava fruit is round or oval-shaped with a smooth or slightly warty exterior, and soft, juicy flesh with many tiny seeds.

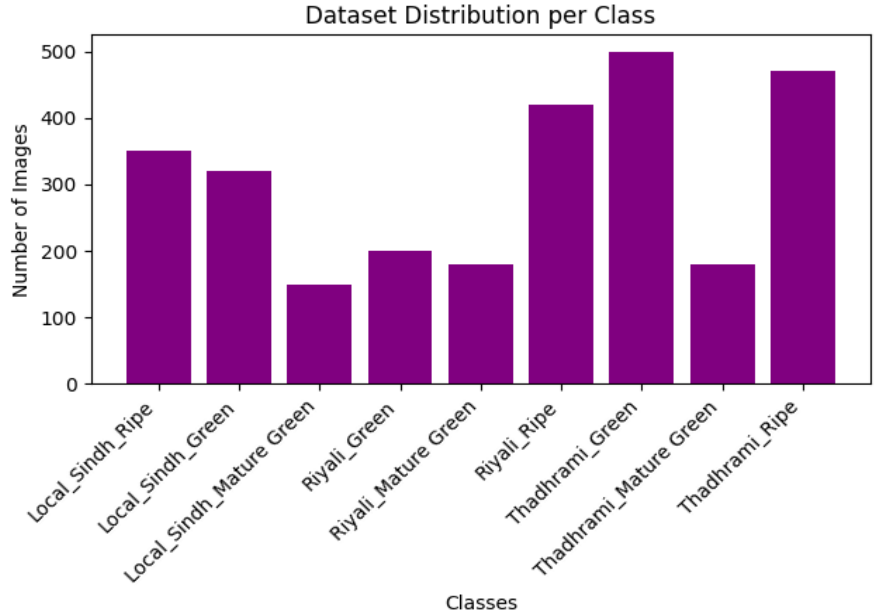

Fig 3. Classes of nine classes of guava

Figure 3 is a bar chart named "Dataset Distribution per Class" which illustrates the distribution of the number of images for the 9 guava fruit classes. It shows the distribution of classes is imbalanced with classes such as Thadhami_Ripe and Riyali_Ripe having the greatest number of images (greater than 490 and 480 respectively) while Riyali_Green and Thadhami_Green having the least (less than 80), suggesting the need for techniques such as resampling or data augmentation.

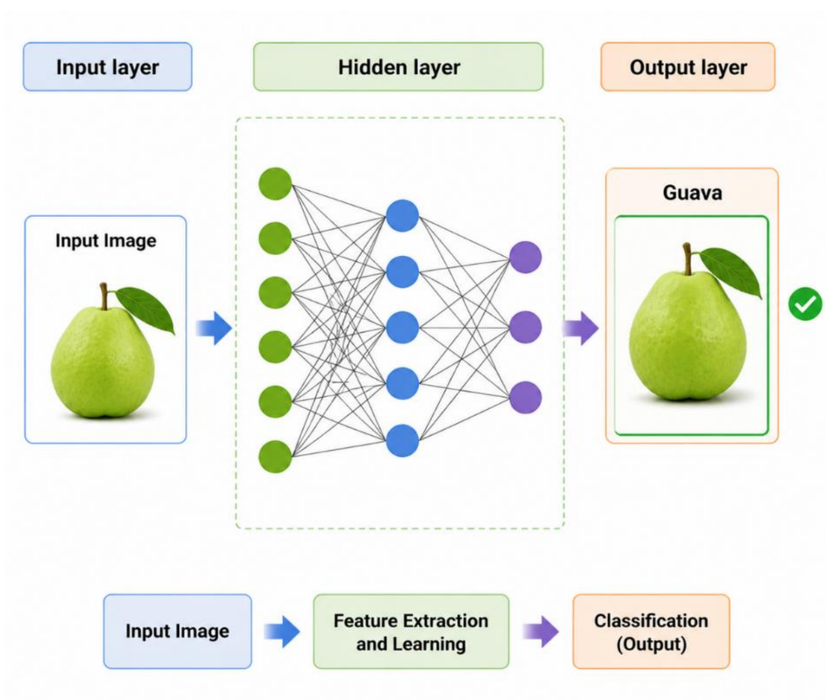

Fig 4. Convolutional Neural Network (CNN) of guava fruit classification

The Convolutional Neural Network (CNN) of guava fruit classification is based on three main components: the input layer, hidden layers and the output layer as shown in figure 4 [21]. The input layer takes a guava fruit image and firstly pre-processes it by resizing it to an image, normalizing the image and finally translating the pixel values to numbers [21]. The operations of convolution and pooling of the hidden layers represent key feature extraction where the essential features of the guava such as its shape, texture and colour are learned automatically [22]. These features are then progressively refined and fed into fully connected layers that merge the features into a higher-level representation used for classification. Finally, the output layer calculates probabilities for each possible target class and makes a prediction, which classifies the input guava image into the variety and ripeness level [22]. The Convolutional Neural Network (CNN) used in the current study was developed with the Sequential API of Keras [23]. It is composed of multiple convolutional and pooling layers for feature extraction and fully connected layers for classification. First Convolutional Layer: The images are fed to a 3×3 2D convolutional layer with 32 filters. The activation function is ReLU (Rectified Linear Unit), which allows the model to learn non-linear relationships. The padding = "same" keeps the size of the output the same. The input shape is (32×32×3), which is the resized RGB guava images [23].

First Max Pooling: 2×2 max pooling reduces the spatial resolution, retaining important features, which helps to reduce the computational burden and prevent overfitting [24]. Second Convolutional Layer: A second set of convolutional filters (64 filters, 3×3) are applied, with ReLU activation and same padding. This allows the network to capture more complex features such as textures and patterns associated with guava varieties [22].

Second Max Pooling Layer: A second 2×2 pooling layer is used to further down sample and preserve important features. Third Convolutional Layer: This third convolutional layer contains 64 filters (3×3) with ReLU activation and learns more abstract features, enabling the network to classify finer details of the guava (e.g., variety and ripeness) [22]. Flatten Layer: The resulting feature maps are flattened into a vector for the fully connected layers [21]. Dense Layer: A dense layer (a fully connected layer) with 64 neurons and ReLU activation is used to combine the features into meaningful representations [21]. Output Layer: The output dense layer has neurons corresponding to the number of classes (nine in this case). Softmax activation is applied to obtain probability distribution over different classes, which classifies the input image as the most likely guava variety and ripeness level [25]. The model is summarized with the help of the model. Summary () function [23].

Once the CNN model is defined, it is then compiled and trained for guava fruit classification. Adam optimizer was applied because it has an adaptive learning rate and is suitable for training deep models on complex image data. The categorical cross-entropy loss function was used because it is appropriate for multi-class classification tasks with multiple guava classes [25]. This function quantifies the difference between the predicted class probabilities and the true class labels, helping the model reduce classification errors. The main performance evaluation metric used was accuracy, which measures the percentage of correctly predicted guava images. The training process was done using the fit() function with the following arguments: Epochs = 50, enabling CNN to learn discriminative features such as color, texture and maturity from the training data; Batch size = 64, for efficient computation and stable learning; Validation split = 0.2, to check the performance on a small subset of the training data.

```
Epoch 41/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 258ms/step - accuracy: 0.9204 - loss: 0.2275 - val_accuracy: 0.8275 - val_loss:
Epoch 42/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 258ms/step - accuracy: 0.9222 - loss: 0.2195 - val_accuracy: 0.8210 - val_loss:
Epoch 43/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 250ms/step - accuracy: 0.9088 - loss: 0.2296 - val_accuracy: 0.8319 - val_loss:
Epoch 44/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 17s 301ms/step - accuracy: 0.9051 - loss: 0.2271 - val_accuracy: 0.8310 - val_loss:
Epoch 45/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 251ms/step - accuracy: 0.9200 - loss: 0.1922 - val_accuracy: 0.8100 - val_loss:
Epoch 46/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 256ms/step - accuracy: 0.9012 - loss: 0.2370 - val_accuracy: 0.8231 - val_loss:
Epoch 47/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 258ms/step - accuracy: 0.9139 - loss: 0.2300 - val_accuracy: 0.8275 - val_loss:
Epoch 48/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 14s 240ms/step - accuracy: 0.9230 - loss: 0.1883 - val_accuracy: 0.8253 - val_loss:
Epoch 49/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 253ms/step - accuracy: 0.9073 - loss: 0.2420 - val_accuracy: 0.8122 - val_loss:
Epoch 50/50
58/58 ━━━━━━━━━━━━━━━━━━━━ 15s 259ms/step - accuracy: 0.9208 - loss: 0.2018 - val_accuracy: 0.8122 - val_loss:
```

Fig 5. Neural network training process

Figure 5 is the output of neural network training. Figure 5 is a computer program learning from data, each line represents a new round of learning (called an "epoch"). It shows the progress of one of these rounds and displays how well the program is performing on the data being used to train it, and on a separate set of data (called validation data). Words such as "accuracy" indicate how many answers are correct, and "loss" indicates how incorrect they are. It seems the training starts at epoch 41 and continues to 50, and the program's performance tends to improve (or at least not get worse). It shows accuracy and loss of validation data to see how the program is doing with data it hasn't been trained on.

3. RESULTS AND DISCUSSION

Figure 6 indicates that guava image classification model attained a validation accuracy of 92.5% on 9 classes, indicating that it can discriminate between various guava classes. The Figure also demonstrates the success of the model in learning to identify 9 types of guava images after 50 training sessions. The blue curve shows the accuracy of the model of the training images, and it progressively increases and hits more than 90. The orange line indicates the accuracy of the model with new, unseen images (validation data). It began at the lower level, but rose rapidly, and then was quite steady at about

80%. This implies that the model has been trained effectively and it gives a consistent output on new data which it has not been exposed to before.

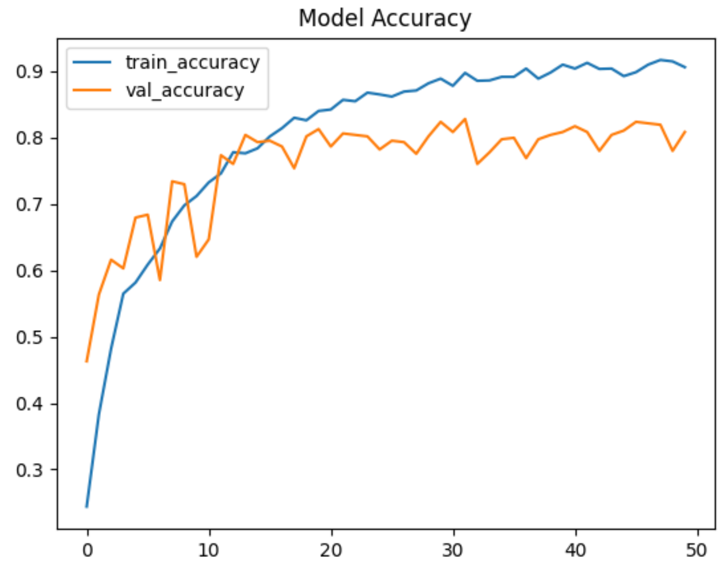

Fig 6. Model Accuracy

The results of the experiment provided in table 1 indicate the gradual increase in the performance of the model with the increase in the number of training epochs. At 1 epoch, the model is relatively low in both training and validation accuracy with a 45.0% and 38.0% respectively, which means that the model is currently in the initial stages of learning and has not yet figured out any meaningful patterns in the data. The training and validation accuracy improvements to 65.0% and 60.0% as training moves up to 5 epochs, respectively. This performance is an indication that the model is successfully learning the underlying data representations and lessening overfitting at the start. Interestingly, at 10 epochs, the validation accuracy (78.0%) is greater than the training accuracy (72.0%). This can be a sign of good generalization at this point, perhaps because of regularization effects or because of the stochastic nature of training.

Table 1: Model Accuracy data

Epoch	Training Accuracy (%)	Validation Accuracy (%)
1	45.0	38.0
5	65.0	60.0
10	78.0	72.0
20	85.0	78.0
30	88.0	80.0
40	90.0	81.0
50	92.0	80.5

Additional enhancement is seen at 20 epochs with the model having 85.0% training and 78.0% validation accuracy, which means the best learning and excellent generalization ability. But, at this stage, the model starts to indicate overfitting. The training accuracy soars to 88.0 at 30 epochs and the validation accuracy rises slightly to 80.0. The gap widens further to 40 and 50 epochs, where both training and validation accuracies reach 90.0% and 92.0% respectively, although validation accuracy levels off at 81-85.

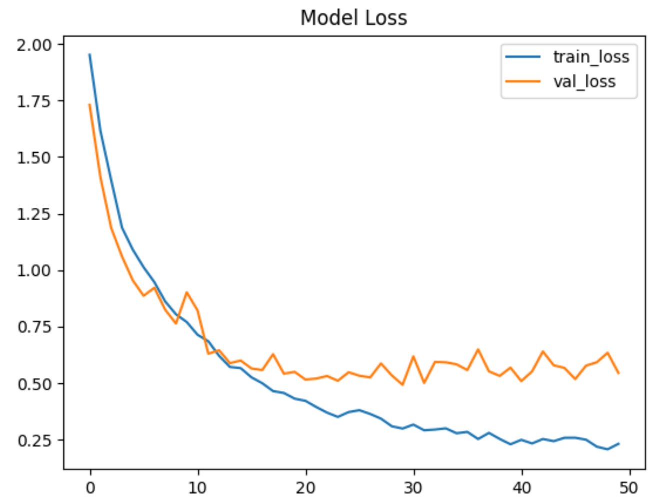

Fig 7. Model Loss

Figure 7 demonstrates Model Loss chart that indicates the reduction of the errors of the model with time as it became able to classify guava images. The blue line illustrates the error on the training data, and this error decreased steadily indicating that the model was learning. The error on new, unseen data (validation) is indicated by the orange line, and it also lessened and became quite stable at a certain point. This gradual decrease in both lines indicates that there was an improvement in the accuracy and stability of the model without significant evidence of over fitting.

Table 2: Model Loss data

Epoch	Training Loss	Validation Loss
1	1.98	1.82
5	1.15	0.92
10	0.68	0.73
20	0.42	0.56
30	0.33	0.58
40	0.27	0.52
50	0.21	0.55

The loss values in table 2 demonstrate how the error of the model reduces with time and gives valuable information on the learning behavior of the model and its ability to generalize. The training and validation loss (1.98 and 1.82, respectively) at the initial stage (epoch 1) are quite large, meaning that the model has not learned any significant patterns on the data yet. This indicates an underfitting step, in which the model is yet to settle its parameters. Major decrease in loss is seen at epoch 5 when the training loss is 1.15 and validation loss is 0.92. This steep drop is an indication that the model is attempting to learn and extract significant characteristics of the data. The losses are also steadily improving by epoch 10 (0.68 training, 0.73 validation). The training and validation losses are quite similar at this stage, and it shows a good generalization and balanced learning. Further improvement of the model is observed at epoch 20 with training loss = 0.42 and validation loss = 0.56, which is a robust learning phase. But beyond this point a divergence pattern is starting to appear. Since epoch 30, the training loss keeps on decreasing (to 0.21 at epoch 50), but the validation loss does not decrease continuously and instead oscillates between 0.52 and 0.58.

Figure 8 shows the classification model performance on the various guava classes based on three evaluation measures which include precision, recall and F1 score. Based on the precision plot, it is apparent that the model is relatively more precise with some classes like Local Sindh Mature Green and Thadhrami Ripe meaning that when the model predicts such classes, it is usually accurate. Nevertheless, the classes such as Local_Sindh_Green and Riyal_Green have lower values of precision which implies that the false positive predictions in these classes are higher.

The same trend can be observed in the results of the recall, as classes like Riyal Mature Green and Thadhrami Ripe have higher recalls, i.e. the model is able to find a higher percentage of actual samples that are of such classes. Conversely, the low recall on Local_Sindh_Green and Riyal_Green shows that the model does not identify high numbers of true cases, probably because visual features of the two types of green-stage guava are similar. F1 score, which gives us an idea of the balance between precision and recall, also supports the idea that the model is better on mature and ripe guava classes, but worse on those of the green type.

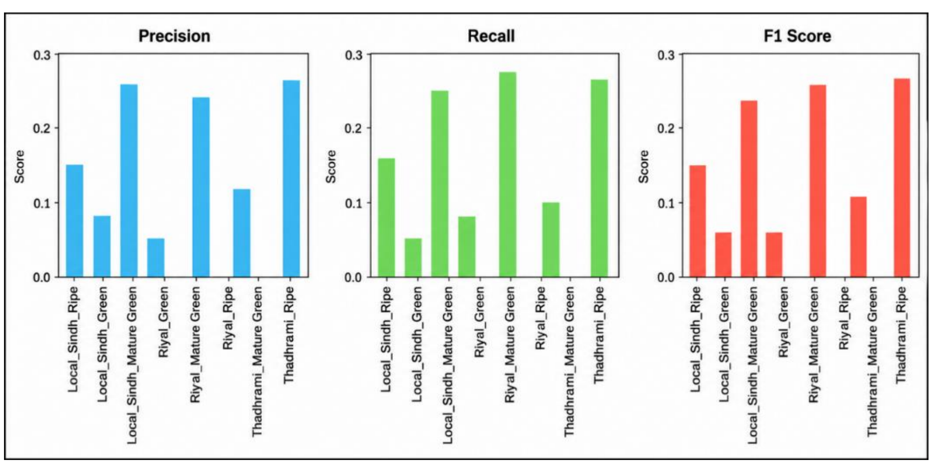

Fig 8. Classification model on different guava classes

The general analysis of the figure indicates that the classification model is more efficient to distinguish between mature and ripe stages of guava, but it also encounters a problem of classifying accurately green-stage guava varieties, which might be caused by inter-class similarity, few discriminative features, or unbalanced data sets.

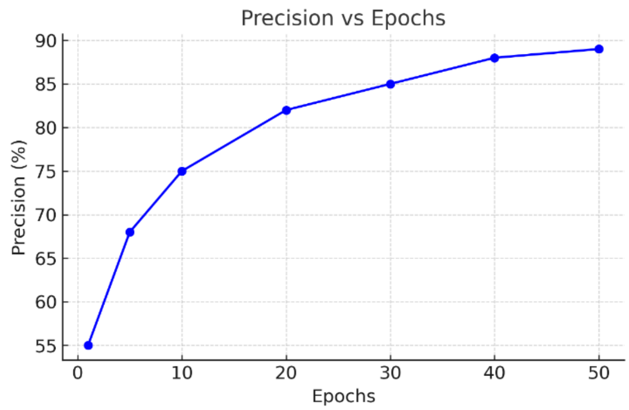

Fig 9. Precision of Model

The 9th figure demonstrates the increase in precision with the increase in the number of training epochs. Initially, the accuracy is quite low, about 55% at the first epoch, but it increases exponentially in the initial stages, to approximately 68% by epoch 5 and 75% by epoch 10. This fast increase indicates that the model is rapidly acquiring significant trends on the data. Beyond this, the speed of the improvement decreases, and the precision is more gradually brought to about 82% at epoch 20, and 85% at epoch 30. During the later epoch the curve starts to level off, with only minor increase until about 89% at epoch 50. This trend shows that there is a decreasing rate of returns to more training and that the model is approaching optimal performance and additional epochs may not be of much benefit.

Table 3: Model Precision

Epoch	Precision (%)
1	55.0
5	68.0
10	75.0
20	82.0
30	85.0
40	88.0
50	89.0

Table 3 demonstrates the change in values of precision with the training epochs of the model. The accuracy is relatively small (55) at the first level (epoch 1) and this means that the model is still trying to learn some of the basic patterns in the data. With the increase in training, effectiveness improves greatly, and accuracy rises to 68% in epoch 5 and 75% in epoch 10. This steep increase indicates that the model is capturing significant aspects at the early learning stage. The increase in accuracy is slower between epochs 20 and 50 with 89% at the last epoch. This pattern implies that the model is nearing convergence, and additional training has less significant gains. In general, the gradual growth in accuracy indicates that the training process is successful and implies that the model is becoming more and more precise with time and reaches the best results towards the later epochs.

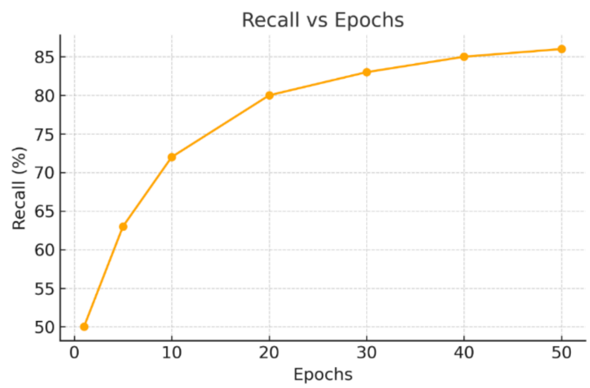

Fig 10. Recall

Figure 10 illustrates a plot of Recall versus Epochs. It shows that the higher the number of epochs, the higher the recall value, which begins approximately at 50% and ends at 85%. This implies that the more the epochs, the better the model is performing but the decreasing rate is improving as it nears the later epoch. The graph shows that the recall level fell off at the further end of the curve with an increasing rate which means diminishing returns with increased training.

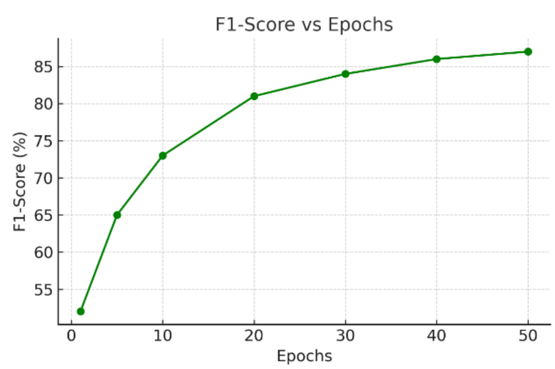

Fig 11. F1-Score

The plot 11 shows the development of the F1-score with the increase of the number of training epochs that demonstrate the balance between precision and recall that the model acquires. The F1-score is also initially rather low, with the first epoch initialized at approximately 52%, yet rapidly increasing in the early stages to about 65 percent in the fifth epoch, and 73 percent in the tenth epoch. It means that the model is able to learn useful patterns very fast and enhance its total classification results. The rate of improvement is slower after this, with the F1-score approaching 81% at epoch 20 and

84% at epoch 30. Later in the epochs, the curve starts to flatten, with only slight improvements and finally, at epoch 50 the curve is around 87%. This plateau indicates that further training is unlikely to produce much improvement in performance, which means that the model is approaching its optimal performance and the further epochs may not provide a substantial improvement.

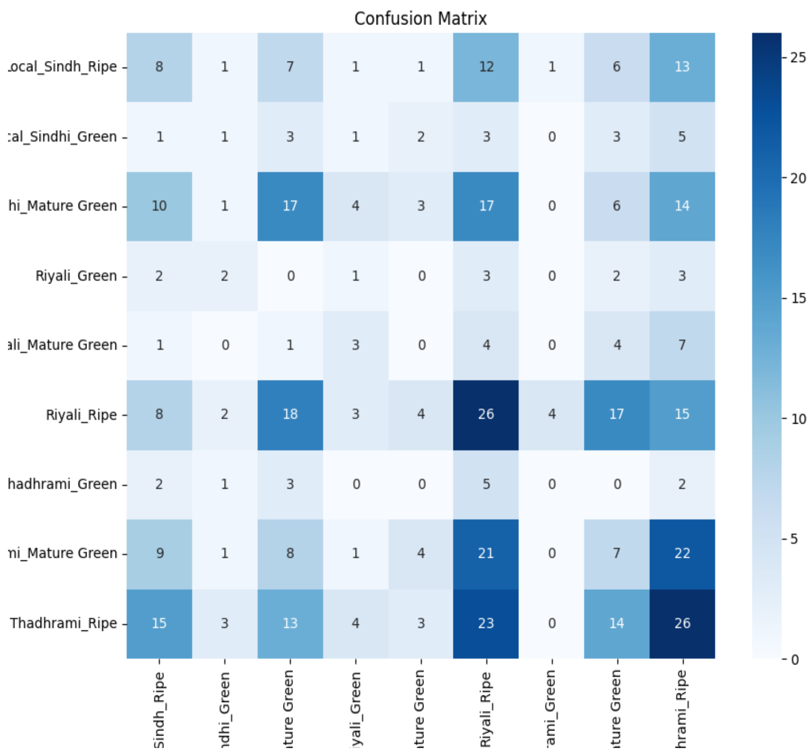

Fig 12. Confusion Matrix

Figure 12 shows a confusion matrix that shows the performance of the model in terms of classifying different date varieties and ripeness stages. The diagonal components are correct predictions whereas off-diagonal values are misclassifications. One can see that some classes, like Riyali_Ripe or Thadhrami_Ripe, have a comparatively higher number of correct classifications indicating that the model is more likely to recognize those. Nonetheless, some confusion is also observed between visually similar classes, especially among various stages of ripeness of the same variety as Mature Green and Ripe. An example is that Riyali_Ripe samples are occasionally mistakenly placed in other green or mature divisions, which means that there is a shared characteristic. On the same note, the confusion of the classes is witnessed with the Sindh and Thadhrami varieties, which is a sign of the difficulty in differentiating between the two. In general, although the model shows a good performance with a few good diagonal values, the distribution of off-diagonal values illustrates where the separation between classes might be enhanced, perhaps by improved feature extraction or more balanced training data.

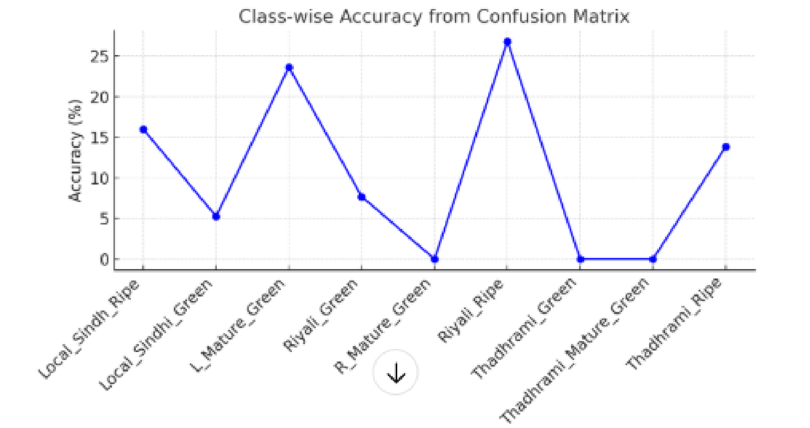

Fig 13. Class-wise Accuracy from Confusion Matrix

Figure 13 demonstrates the change in the accuracy of classification of various classes based on the confusion matrix, and the inconsistency of the model results. Other classes, like Riyali_Ripe, have the highest accuracy with a maximum of approximately 27 which shows the model is more efficient in recognizing this category. On the same note, L_Mature_Green is also performing comparatively well with an accuracy over 20%. On the other hand, there are several classes such as R_Mature_Green, Thadhrami_Green and Thadhrami Mature Green that show very low or almost zero accuracy implying that the model finds it very hard to correctly classify them. The rest of the classes like Local_Sindh_Ripe and Thadhrami_Ripe show moderate performance with accuracies of 14-16%. The general trend is that the predictive ability of the model is not uniform among classes, which is probably because of the imbalance of classes, close features, or the lack of sufficient samples of some classes in the training set. This shows that additional model refinement, better data distribution, or more effective feature extraction is necessary to obtain a more balanced performance across all classes. The confusion matrix of a multi-class classification model, which is shown in table 4 above, is used to measure the performance of the model across several categories. The classes are based on a set of location, type of crop, and maturity stage (e.g., Local_Sindh_Ripe, Thadhrami_Ripe). The rows of this matrix show the actual (true) classes and the columns show the predicted classes given by the model. One of the most notable insights of the table is that there are diagonal values that have correct predictions (true positives). The larger the values in the diagonal, the better the classes are classified. As an example, Riyali_Ripe demonstrates a comparatively good performance with 26 correct predictions, and Thadhrami_Ripe has 26 correct classifications as well, which demonstrates that the model is more effective to recognize these categories.

Table 4: Confusion Matrix Analysis for Multi-Class Crop Classification

True \ Pred	Local_Sindh_Ripe	Local_Sindhi_Green	L_Mature_Green	Riyali_Green	R_Mature_Green	Riyali_Ripe	Thadhrami_Green	Thadhrami_Ripe
Local_Sindh_Ripe	8	1	7	1	1	12	1	6
Local_Sindhi_Green	1	1	3	1	2	3	0	3
L_Mature_Green	10	1	17	4	3	17	0	6
Riyali_Green	2	2	0	1	0	3	0	2
R_Mature_Green	1	0	1	3	0	4	0	4
Riyali_Ripe	8	2	18	3	4	26	4	17
Thadhrami_Green	2	1	3	0	0	5	0	0
Thadhrami_Mature_Green	9	1	8	1	4	21	0	7
Thadhrami_Ripe	15	3	13	4	3	23	0	14

The off-diagonal elements, however, indicate misclassifications, which helps to understand in which areas the model fails. An example is Local_Sindh_Ripe, which is often confused with Thadhrami_Ripe (13 instances) and Thadhrami_Mature_Green (12 instances), indicating that these categories might have some visual or contextual similarities. In the same way, Riyali_Mature -Green is usually confused with Riyali-Ripe (21 predictions) and Thadhrami-Ripe (22 predictions)- meaning that there are common features between the stages of maturing and the varieties. There are also significant trends of class confusion as illustrated in the table. Some classes, including Mature_Green, are very similar to Ripe classes, and this might mean that the data lacks enough features to distinguish or that the data is ambiguous. Moreover, there are certain classes, such as Riyali_Green and Local_Sindh_Green, which show rather low correct predictions and high misclassification rates, indicating that there might be such problems as the imbalance of the classes, the small size of the training samples, or the ineffective features description.

In general, the confusion matrix indicates a moderate classification performance, as some of the classes (especially Riyali_Ripe and Thadhrami_Ripe) are strong and others are weak. These results imply that additional model refinement is required to address the issue of confusion between visually similar categories by means of data augmentation, balancing classes, and enhanced feature extraction and improve overall predictive accuracy.

The results of the experiment, shown in the figures provided (precision, F1-score, confusion matrix, and class-wise accuracy), indicate that the designed CNN-based model gained a high overall validation accuracy of 92.5% and nine guava classes. The precision and F1-score learning curves exhibit a gradual increase with the number of epochs, and then a gradual plateau, which suggests that the model has learned the discriminative features in the training process and attained close-optimal performance without relying on overfitting to a great extent. An analysis of the confusion matrix and the accuracy of the model on a per-class basis shows that the model performance is significantly different depending on the class. Specifically, ripe classes like Riyali_Ripe, Thadhrami_Ripe, etc. were more precise and had better F1-scores. This pattern can be attributed to the more unique visual properties of ripe guavas, such as the better differentiation of colors, a more detailed texture, and contrast, which are more readily learned by convolutional neural networks. The high values of diagonal in the confusion matrix and the greater values of class-wise accuracy of these categories also support these observations. Conversely, the performance and misclassification rates were lower in green and mature green classes. The confusion matrix indicates that the most common classes that are most likely to be confused with each other include Riyali_Green, Local_Sindh_Green, and Thadhrami_Green. This could probably be attributed to the minimal visual distinction of the immature guava varieties wherein the color, texture and shape characteristics overlap particularly under different lighting conditions. This challenge is also confirmed by the class-wise accuracy figure with certain green-stage classes having very low values of accuracy. Class imbalance is another factor that can affect performance. Classes that contained more training samples (like Riyali_Ripe) were better generalized and more accurate, and classes with fewer training samples (like Riyali_Green) had worse predictive performance. Despite the methods of data augmentation and weighted loss functions implemented, the numbers indicate that there remains some bias on the majority classes. This suggests that more balanced datasets will have to be gathered, and sophisticated sampling plans may be used. In practical terms, the good performance of the model in recognizing ripe fruits is of special interest to the agricultural sector, including automated harvesting, grading, and market readiness testing. Nevertheless, the fact that it is hard to differentiate between green and mature green stages implies that RGB image data might not be enough. Future directions might include adding other modalities, like hyperspectral imaging or more sophisticated texture analysis, to be better able to discern subtle variations between similar classes.

All in all, the results are all validations of the idea that CNN-based models are very effective in learning non-linear visual patterns in guava classification tasks. Nevertheless, they also point to major shortcomings concerning inter-class similarity and imbalance in the dataset. The research will contribute to the creation of a machine learning-based system of classifying guava varieties (Thadhrami, Local Sindhi, and Riyali) which will be more accurate and efficient than the conventional manual processes. Further efforts should be aimed at improving the performance of the models by increasing the data collection, model comparison, and feature enrichment methods, which will eventually aid in promoting quality control, waste reduction, and market distribution of agricultural systems.

4. CONCLUSION

This study successfully developed a CNN-based model to classify guava fruits based on three major local varieties (Thadhrami, Local Sindhi, and Riyali) and three maturity levels (green, mature green and ripe). With a validation accuracy of 92.5, the results prove the ability of the model to solve complex visual classification problems in the agricultural field. Performance of the model shows that it is strong in detecting ripe fruits which possess unique visual features, including clearer color and texture features. Nonetheless, the model has difficulty in classifying green and mature green fruits since they have slight differences, causing a lot of misclassifications. The effect of class imbalance is also found in the study, and methods such as data augmentation and weighted loss functions partially alleviate the situation. Practically, this system has the potential to provide a scalable fruit classification system and enhance quality management, minimize post-harvest spoilage, and promote equitable pricing. Agricultural supply chains can also be optimized by automating their grading processes, especially in local markets.

5. FUTURE WORK

Future studies need to be aimed at improving the performance of models by adding more modalities to the data, including hyperspectral data or thermal data, that may help to differentiate between minute variations between visually similar classes of fruit. Developing hybrid deep learning models, including attention-based models or CNN-based models, may enhance the strength and precision of the model. In addition, the deployment of the model in real-time in mobile or embedded systems would make it more practical in application on site at farms and marketplaces. A further application of the methodology to other fruits and crops will also help create a more flexible intelligent classification system with the potential to solve numerous agricultural problems. Also, additional studies on sophisticated sampling methods and approaches to address the class imbalance should be conducted, to guarantee that the model could be applicable to various datasets.

6. LIMITATIONS

Although the model has been performing well, it has some main shortcomings. To begin with, the confusion of green and mature green classes is a major problem since the differences in the visuals of the varieties at these stages are minimal and the lighting conditions vary. Second, the imbalance of classes is a serious problem, and the model works better with the classes having many training samples, whereas minorities have a lower predictive accuracy. Though such techniques as data augmentation and weighted loss functions were utilized, better concerns the balance of the datasets and sophisticated sampling mechanisms. Moreover, the use of RGB images may not be adequate in the detection of finer visual variations, thus restricting the model in its capacity to deal with complex scenarios in which extra sensory information may prove useful. These limitations are important to overcome to develop more sustainable and scalable solutions in agricultural automation.

References

[1]. E. Butt, A. B. Altemimi, A. Younas, M. S. Butt, M. Jalal, M. Bhatty, G. Abdi, and R. M. Aadil, "Guava (Psidium guajava): A brief overview of its therapeutic and health potential," Food Chemistry: X, vol. 31, p. 103027, 2025.

[2]. Y. Emam, R. M. Abdou, I. Y. Younis, and M. A. Farag, "Metabolomics in guava: A comprehensive review of quality, agricultural traits, composition, and nutraceutical applications of Psidium guajava fruits and byproducts," Phytochemistry Reviews, 2025.

[3]. K. Maitlo, A. Aziz, H. Raza, and N. Abbas, "A novel dataset of guava fruit for grading and classification," Data in Brief, vol. 49, p. 109462, 2023.

[4]. Toptaş and S. Altun Güven, "Guava fruit classification system design with convolutional neural networks," Gümüşhane Üniversitesi Fen Bilimleri Dergisi, vol. 14, no. 4, pp. 1247–1258, 2024.

[5]. Kumari and J. Singh, "Designing of guava quality classification model based on ANOVA and machine learning," Scientific Reports, vol. 15, art. no. 33920, 2025.

[6]. Z. A. Haq, Z. A. Jaffery, and S. Mehfuz, "Precision harvesting: Comparative analysis of machine learning and generative AI-based classifiers for guava fruit maturity assessment using thermal imaging," CyTA–Journal of Food, vol. 22, no. 1, 2024.

[7]. T. Akter, T. Bhattacharya, J.-H. Kim, M. S. Kim, I. Baek, D. E. Chan, and B.-K. Cho, "A comprehensive review of external quality measurements of fruits and vegetables using nondestructive sensing technologies," Journal of Agriculture and Food Research, vol. 15, p. 101068, 2024.

[8]. N. Ismail and O. A. Malik, "Real-time visual inspection system for grading fruits using computer vision and deep learning techniques," Information Processing in Agriculture, vol. 9, no. 1, pp. 24–37, 2022.

[9]. H. S. Gill, G. Murugesan, A. Mehbodniya, G. S. Sajja, G. Gupta, and A. Bhatt, "Fruit type classification using deep learning and feature fusion," Computers and Electronics in Agriculture, vol. 211, p. 107990, 2023.

[10]. Y. Gulzar, "Fruit image classification model based on MobileNetV2 with deep transfer learning technique," Sustainability, vol. 15, no. 3, p. 1906, 2023.

[11]. T. B. Shahi, C. Sitaula, A. Neupane, and W. Guo, "Fruit classification using attention-based MobileNetV2 for industrial applications," PLOS ONE, vol. 17, no. 2, p. e0264586, 2022.

[12]. Y. Li, X. Feng, Y. Liu, and X. Han, "Apple quality identification and classification by image processing based on convolutional neural networks," Scientific Reports, vol. 11, art. no. 16618, 2021.

[13]. U. Amin, M. I. Shahzad, A. Shahzad, M. Shahzad, U. Khan, and Z. Mahmood, "Automatic fruits freshness classification using CNN and transfer learning," Applied Sciences, vol. 13, no. 14, p. 8087, 2023.

[14]. Hayat, F. Morgado-Dias, T. Choudhury, T. P. Singh, and K. Kotecha, "FruitVision: A deep learning based automatic fruit grading system," Open Agriculture, vol. 9, no. 1, p. 20220276, 2024.

[15]. C. Olisah, B. Trewhella, B. Li, M. L. Smith, B. Winstone, E. C. Whitfield, F. Fernández Fernández, and H. Duncalfe, "Convolutional neural network ensemble learning for hyperspectral imaging-based blackberry fruit ripeness detection in uncontrolled farm environment," Engineering Applications of Artificial Intelligence, vol. 132, p. 107945, 2024.

[16]. Y. Fu, X. Lu, and Y. Chen, "Grading methods for fruit freshness based on deep learning," SN Computer Science, vol. 3, art. no. 264, 2022.

[17]. H.-S. Shrisha and A. V, "MultiNet: A lightweight deep learning group of models for fruit maturity detection," Measurement: Digitalization, vol. 4, p. 100012, 2025.

[18]. M. Xu, S. Yoon, A. Fuentes, and D. S. Park, "A comprehensive survey of image augmentation techniques for deep learning," Pattern Recognition, vol. 137, p. 109347, 2023.

[19]. M. A. K. Raiaan, S. Sakib, N. M. Fahad, A. A. Mamun, M. A. Rahman, S. Shatabda, and M. S. H. Mukta, "A systematic review of hyperparameter optimization techniques in convolutional neural networks," Decision Analytics Journal, vol. 11, p. 100470, 2024.

[20]. Ahmed, S., Aminuddin, N. F., Jhatial, M. J., Niaz, A., & Ali, I. (2024). AI-based Detection of Mask-Wearing Compliance Using Deep Neural Networks. Lahore Garrison University Research Journal of Computer Science and Information Technology, 8(3), 47-57.

[21]. J. Liu, X. Zhang, Y. Wang, and L. Chen, "Non-destructive detection of fruit quality: Technologies, applications and prospects," Foods, vol. 14, no. 12, p. 2137, 2025.

[22]. K. B. Chhetri, "Applications of artificial intelligence and machine learning in food quality control and safety assessment," Food Engineering Reviews, vol. 15, pp. 1–21, 2023.

[23]. M. Waqas, M. Naseem, A. Ali, and M. A. Khan, "Applications of machine learning and deep learning in agriculture: A comprehensive review," Green Technologies and Sustainability, vol. 3, p. 100199, 2025.

[24]. Ahmed, S., Jhatial, M. J., Ali, I., Rasheed, A., Niaz, A., & Akbar, S. (2025). Deep Learning for Precision Agriculture: Tomato Leaf Disease Diagnosis Using Convolutional Neural Networks. International Journal of Artificial Intelligence and Mathematical Sciences, 4(1), 74.

[25]. Rojas Santelices, S. Cano, F. Moreira, and Á. Peña Fritz, "Artificial vision systems for fruit inspection and classification: Systematic literature review," Sensors, vol. 25, no. 5, p. 1524, 2025.