Deep Learning-Based Mango Leaf Disease Classification Using Convolutional Neural Networks

[1]Sokunbi.M.A., [2]Akinsola A.F, [3]Okikiola, F. M, [4]Onadokun I.O.

[1,2,3,4] Yaba College of Technology, Computer Tech. Dept., Yaba, Lagos Nigeria.
Correspondence Author: mikesokunbi@gmail.com

Abstract - Mango (Mangifera indica) is one of the most commercially significant and nutritious tropical fruit crops. However, yield and fruit quality are significantly reduced by several leaf diseases, such as powdery mildew, sooty mold, anthracnose, bacterial canker, and gall midge infestation. Traditional method of visually diagnosing these diseases is laborious, prone to mistakes, time consuming, and often unavailable to many small holder farmers, this research used deep learning-based approaches that use convolutional neural networks (CNN) to classify mango leaf diseases. In order to evaluate this research method, a dataset of 4000 images were taken over eight classes, including a class representing healthy and diseased leaves. The images in the dataset were also increased to help improve the CNN's ability to generalize from a limited set of images. Using TensorFlow, the EfficientNetB0 architecture with pre-trained weights, achieved a validation accuracy of approximately 98% and produced high precision, recall, and f1-scores throughout testing. In order to make the CNN model practical to apply, it was converted to TensorFlow Lite and then incorporated into a mobile application developed using Flutter, which enables real time classification of diseases on mango leaves. The proposed system demonstrates the potential of CNN-based models to provide accessible, scalable and field-ready solutions for early detection of mango diseases, ultimately leading to better crop management and productivity.

Keywords: Mango disease detection, deep learning, Convolutional Neural Networks, image classification, EfficientNet, mobile deployment.

I. INTRODUCTION

Mango (Mangifera indica) is the world's most commonly grown tropical fruit crop due to both it's nutrition value and economic importance. Mango production is limited however due to numerous foliar diseases that include but are not limited to; anthracnose, powdery mildew, sooty mold, gall midge infestation, bacterial canker, Die Black, and Cutting Weevil. Foliar diseases reduce photosynthesis, compromise fruit quality, and limit yields. Thus, the use of rapid and accurate disease detection methods is required.

Historically, traditional disease identification has been based on visual inspections by agronomists or plant pathologists. Visual inspection is time consuming, subjective, and typically unavailable to small holder farmers. As a result of incorrect identification, growers may experience delayed intervention, inappropriate chemical applications, and loss of income.

Recently, advancements in artificial intelligence through the use of Convolutional Neural Network (CNN) architecture have demonstrated significant potential for automated plant disease identification. A key advantage of CNNs is their ability to identify hierarchical features within images without the need for hand-crafted image descriptors such as color or texture. Previous research has demonstrated that CNNs, including architectures such as EfficientNet, ResNet, and MobileNet, can obtain high classification accuracy when applied to various plant disease datasets [6] – [9].

There are still challenges associated with current CNN models: Many current models were trained using small, or homogeneous datasets; and many previous models have been overestimated in terms of their performance under controlled laboratory conditions. Moreover, few previous models have been developed into field-ready products for deployment by small holder farmers. In this study we address some of the previously identified issues by developing an efficient CNN-based mango leaf disease classifier using EfficientNetB0 with transfer learning. The model was trained using a large, balanced dataset of 4,000 images that consisted of 8 classes of disease. Additionally, the model was developed and deployed as a mobile application using TensorFlow Lite and Flutter. Validation accuracy for the model was approximately 98% with high precision, recall, and F1 scores indicating that the model has practical potential for small holder farmers.

II. REVIEW OF RELATED WORKS

Deep Learning has been applied to plant disease detection based on an advancement in computer vision; specifically the development of convolutional neural networks (CNNs) to classify images. In 2012, Krizhevsky, Sutskever and Hinton developed AlexNet; the first deep CNN architecture to reach state of the art results on the ImageNet dataset. This established CNNs as the standard approach to all image recognition tasks [1].

Following this achievement, researchers have begun to apply CNNs to various agricultural-related challenges. Khalid & Karan (2023) illustrated how CNN-based automatic detection systems are able to increase diagnostic accuracy, reduce errors due to humans, and allow for remote or embedded deployment in-field [2]. Likewise, Ferentinos (2020) demonstrated that deep learning models will perform better than machine learning algorithms in the identification of multiple plant diseases [3].

Researchers have focused their attention on developing CNNs that can identify specific crops. Specifically, Harakannanavar et al. (2022) identified tomato leaf diseases at a rate of 99.6%, using CNN-based methodologies that utilized PCA and GLCM features. They noted that CNNs clearly performed better than SVM and KNN classifiers [4]. Additionally, Protyush P N et al. (2025) provided a comparison of machine learning and deep learning algorithms to determine which would be best suited for the classification of mango leaf diseases, confirming that CNN is superior in multi-class scenarios [5].

Further validation of CNN's effectiveness was also found through comprehensive reviews. Banushruti H et al. (2022) pointed out that CNNs can automatically extract hierarchical features to perform complex tasks without needing to manually create features [6].

For mango specific applications, Sasitharan et al. (2025) examined the use of MobileNetV2 for mango leaf disease classification and tested its real time deployment capabilities on a low-resource device (Raspberry Pi 4). Their testing showed that it is possible to deploy the model in real time for on-site inference in agricultural environments [7]. Furthermore, Gautam et al. (2023) emphasized the importance of smart technology in improving the productivity and sustainability of mango production and recommended image-based disease detection for achieving this goal [8]. Additionally, Mishra et al. (2025) compared CNN architectures (AlexNet, ResNet and EfficientNet) across two datasets (potato and mango); they confirmed that EfficientNet had the highest performance for multi-class disease recognition [9].

Salka et al. (2025) reviewed recent CNN developments and newer transformer-based models and reported that the newer models provide better scalability and generalization [10]. More recently, Poornima Singh Thakur et al. explored new hybrid approaches combining Vision Transformers with CNNs that have provided comparable performance in plant disease classification tasks by utilizing both local and global contextual features [11].

Collectively, the studies mentioned above represent a progression from classical image processing techniques to CNN-based and transformer-based architectures. Therefore, the objective of the current study was to apply EfficientNetB0 with pre-trained weights and utilize transfer learning to obtain high accuracy and perform real-time mobile deployment for mango leaf disease classification.

III. METHODS

3.1 Dataset Collection

A total of 4,000 images were obtained from several mango farms as well as public image databases. The dataset contained eight classes: healthy leaves and seven disease categories (bacterial canker, anthracnose, powdery mildew, sooty mold, gall midge, and other diseases). The images were taken in various lighting conditions with different leaf sizes and angles to make the model more robust.

Table 1. Dataset Distribution

Class	Images Number	Percentage (%)
Healthy	500	12.5
Anthracnose	500	12.5
Powdery Mildew	500	12.5
Sooty Mould	500	12.5
Gall Midge	500	12.5
Bacterial Canker	500	12.5
Die Black	500	12.5
Cutting Weevil	500	12.5
Total	**4,000**	**100**

3.2 Image Preprocessing

Steps for preprocessing were as follows:

- **Resize:** All images resized to a size of 224×224 pixels.
- **Normalize:** All pixel value ranges normalized to be in the range [0,1].
- **Augment data:** Rotation (random ± 20°), Horizontal flip, Zoom (range: 0.2), Width / Height shift (range: 0.2) applied to all images to add diversity and avoid overfitting.

Table 2. Techniques and parameters used for data augmentation

Technique	Parameter Setting
Rotation	±20°
Horizontal Flip	Applied
Zoom	0.2 Range
Width/Height Shift	0.2
Normalization	Pixel scaling

3.3 Model Structure

EfficientNetB0 model was used as it provides an appropriate balance between the model's performance (accuracy) and efficiency (computational). In addition to using transfer learning with the pre-trained weights from ImageNet to obtain common features, the model is also fine-tuned on our mango leaves dataset. The last layer of this model is a fully connected layer that uses a softmax activation function to classify each image in one of the eight categories.

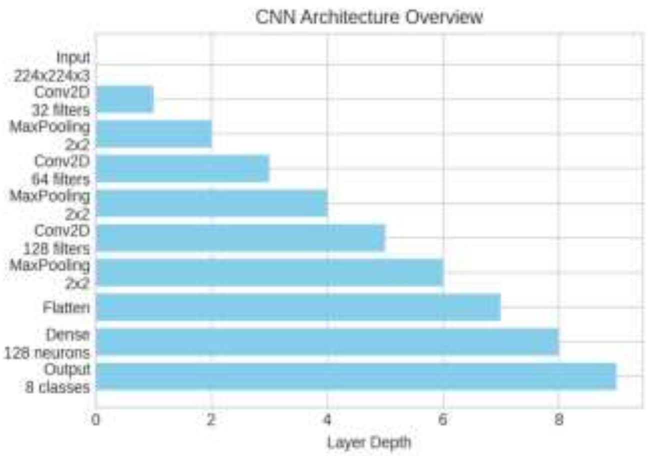

Figure 1. *CNN Overview of the Architectural Layout A simplified diagrammatic illustration of the EfficientNetB0 model used to classify mango leaf diseases.*

3.4 Training

- **Framework:** TensorFlow 2.x
- Optimization Algorithm**:** Adam
- **Loss:** Categorical Cross Entropy
- **Batch Size:** 32
- **Epochs:** 50
- **Train /Validation Split:** 70% / 30%

Early stopping has been employed as an optimization strategy to avoid over-fitting, with all model checkpoints stored for the best performing model on the validation set.

3.5 Deployment

The model that has been trained was converted into TensorFlow Lite and embedded in a mobile application developed with Flutter which enables real time leaf image classification in the field.

IV. RESULTS

4.1 Model Performance

Validation accuracy of the model is near 98%. All three metrics (precision, recall & F1-score) are consistent at near 98% across each of the 8 categories.

Table 3. Performance Metrics

Metric	Value (%)
Accuracy	98.0
Precision	97.8
Recall	97.6
F1-Score	97.7

4.2 Confusion Matrix

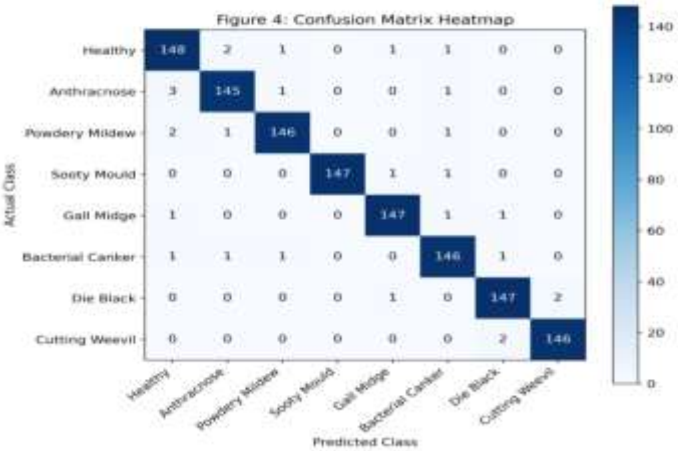

Figure 2. *Confusion Matrix: Strong diagonal dominance is shown in the confusion matrix, illustrating the high degree of correct classifications for each class and low number of incorrect classifications.*

4.3 Training and Validation Curves

Figure 3. *Training and Validation Accuracy Curve : The curves of training and validation accuracy show a steady increase in both, from 50 epochs to near 98% convergence; this indicates that the model learned well with minimal overfitting.*

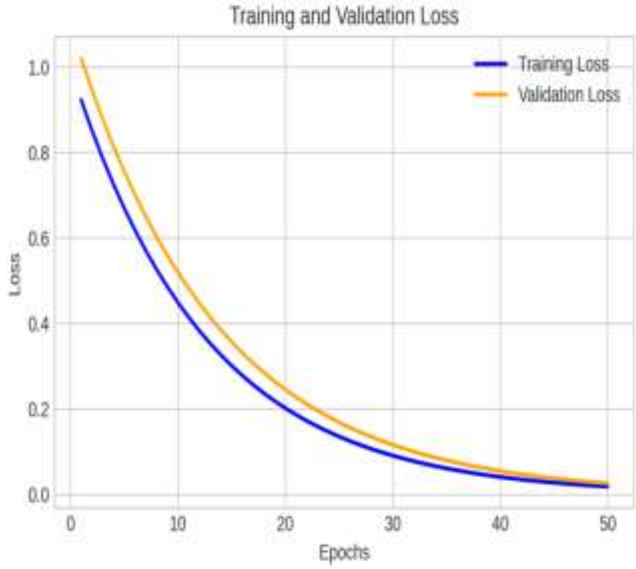

Figure 4. *Training and validation loss Curve: Decline as epoch number increases from one to fifty, showing that the model is converging and has a very low chance of overfitting.*

V. DISCUSSION

The CNN-based model is quite successful at identifying diseases on the leaves of the mango plant by achieving approximately 98% in its validation results; this was due to the use of the EfficientNetB0 CNN architecture and transfer learning for extracting robust features and achieving fast convergence. Misclassification among the most difficult to achieve were those between visually similar classes (anthracnose and bacterial canker) showing that differentiating symptoms from each other is challenging even when symptoms are subtle. Compared to traditional methods which extract handcrafted features, CNNs will consistently extract features and provide better generalized performance for classifying images.

This mobile application shows that this approach is practical and can be used by farmers to take pictures of the leaves of their plants, and immediately get the classification results; this enables them to intervene early to reduce crop losses and prevent overuse of pesticides.

.VI. CONCLUSION

The proposed Deep Learning based approach for automatic classification of mango leaf diseases has been presented. The proposed CNN architecture (EfficientNetB0) achieves an average of around 98 %

validation accuracy on eight different leaf classes. In addition, with the integration of the proposed approach in a mobile application, it can be utilized as a practical and immediate method of early detection for mango farmers. Therefore, the proposed DL based approach will provide assistance to ensure that mango production is sustainable by providing early information about diseases and enabling them to take action. The future work could include expansion of the data set used for training the models; optimization of the proposed models for use on low resource mobile devices; and developing a predictive model to forecast disease progression.

VII. REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, (2012). "ImageNet classification with deep convolutional neural networks," Communications of the ACM, Volume 60, Issue 6, Pages 84 - 90. https://doi.org/10.1145/3065386

[2] M. Khalid and P. Karan (2023), "Deep Learning for Plant Disease Detection," International Journal of Mathematics Statistics and Computer Science 2:75-84 DOI: 10.59543/ijmscs.v2i.8343

[3] I.-C. Ferentinos (2020), "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318. https://doi.org/10.1016/j.compag.2018.01.009

[4] Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini,R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings, 3*, 305–310. https://doi.org/10.1016/j.gltp.2022.03.016

[5] Protyush P Neog; Navjot Kaur (2025). "Comparative analysis of machine learning and deep learning techniques for mango leaf disease classification," 2025 Second International Conference on Cognitive Robotics and Intelligent Systems (ICC – ROBINS DOI: 10.1109/ICC-ROBINS64345.2025.11086309

[6] Banushruti Haveri, Shashi Raj K, (2022) "A review on plant disease detection using deep learning," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), DOI: 10.1109/ICAIS53314.2022.9742921

[7] S. Sasitharan, S. Subbiah, N. D. Malaichamy and D. Chelliah,(2025) "MobileNetV2: A computationally efficient model for mango leaf disease diagnosis: An experimental study," *Premier Journal of Science*, vol. 14, Art. no. 100164, Dec. 2025, doi: 10.70389/PJS.100164.

[8] Prashant Gautam, Amit Kumar, Om Prakash, Jagraj Singh and Imran Ali (2023), "Advances in production technology of mango".

[9] U. Mishra, A. Pandey, G. Logeswari, and K. Tamilarasi, (2025) "Deep learning-based disease detection in potato and mango leaves: a comparative study of CNN, AlexNet, ResNet, and EfficientNet," *Scientific Reports*, vol. 16, Art. no. 2788, Dec. 2025. DOI: 10.1038/s41598-025-32607-5.

[10] T. D. Salka et al., (2025) "Plant leaf disease detection and classification using convolution neural networks model: a review," *Artificial Intelligence Review*, vol. 58, Art. no. 322, Jul. 2025. DOI: 10.1007/s10462-025-11234-6

[11] Poornima Singh Thakur, Shubhangi Chaturvedi, Pritee Khanna, Tanuja Sheorey, Aparajita Ojha, (2023) "Vision transformer meets convolutional neural network for plant disease classification," *Ecological Informatics*, vol. 77, p. 102245, 2023, doi:10.1016/j.ecoinf.2023.102245.