International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-15 Issue-1, May 2026

Resource-Efficient MobileNetV2 Model for Multiclass Plant Disease Prediction Using Real-Time Data in Smart Farming

Iita Anatolia, Srinu Sesham, Mateus Abisai, Kenneth Gideon

Abstract: ***Agriculture remains a cornerstone of Namibia's economy, yet small-scale crop farmers continue to face significant productivity losses due to late or inaccurate diagnosis of plant diseases. Tomato, a major crop in the country's semi-arid regions, is highly susceptible to fungal and bacterial infections that spread rapidly under local climatic conditions. Manual inspection is labour-intensive, subjective, and ineffective for large-scale monitoring. In the literature, many studies have used high-quality datasets to train deep learning models. However, these datasets are not real-time and rarely reflect Namibia's specific atmospheric and climatic conditions. To address this challenge, this study uses a blended dataset combining the Plant Village Tomato Leaf Dataset from Kaggle with real-time images collected from small farms in Namibia. The study further investigates a resource-efficient and reliable deep learning model, namely MobileNetV2, for multiclass classification of plant diseases. The proposed framework using the MobileNetV2 model is benchmarked against the VGG16 and ResNet50 models, both trained and fine-tuned on the blended dataset. The models are compared in terms of the overall prediction accuracy from the multiclass confusion matrix and their computational cost. The results indicate that the proposed multiclass classification model based on the MobileNetV2 architecture has achieved the best performance near to 90% accuracy, compared to VGG16 (88.33%) and ResNet50 (58.02%), while incurring minimal computational cost. The model achieved fast predictions with reasonable accuracy, enabling mobile deployment to monitor crop health in the field. The results show that MobileNetV2 offers a low-cost way to assess tomato crop health and support farmers in Namibia using digital technologies.***

Keywords: ***Lightweight CNNs, MobileNetV2 Model, Multiclass Classification, Accuracy, Digital Technologies.***

Nomenclature:
DL: Deep Learning
CNNs: Convolutional Neural Networks
ML: Machine Learning
FC: Fully Connected
GAP: Global Average Pooling
AUC: Area Under the Curve
AP: Average Precision
PR: Precision-Recall
ROC: Receiver Operating Characteristic

I. INTRODUCTION

The agricultural sector is confronted with many new challenges in the 21st century, with climate change, pest invasions, and plant diseases together endangering security and economic stability in Namibia [1, 2]. Of the many crops farmed in the country, tomato (Solanum lycopersicum) is the most popular vegetable because of its massive consumer base in local markets. But several plant diseases plague tomato cultivation in Namibia, and late blight is a major threat to production if left untreated. Disease diagnosis in practice is largely based on visual inspection, which is time-consuming and requires expert knowledge [3]. This can result in late or incorrect diagnosis, leading to overuse of pesticides, compromised crop quality, and monetary losses for smallholder farmers. In recent years, digital technologies and Artificial Intelligence (AI) models have transformed plant disease diagnosis [4, 5]. In particular, computer vision with deep learning (DL), such as pre-trained convolutional neural networks (CNNs), is effective for the automatic detection of leaf diseases, compared to machine learning (ML) approaches using hand-crafted features [6]. By employing transfer learning with a pretrained CNN, models can be trained even with limited data, leveraging knowledge learned from large datasets such as ImageNet.

While significant advancements have been made in plant disease recognition globally, little work has been done to tailor these approaches to Namibia's local environmental and agricultural practices, where leaf texture and disease patterns may not align with datasets from other regions. This is largely due to differences in climate, especially high temperature variability and limited irrigation water [2]. Although CNN-based techniques have been successful in detecting diseases with high accuracy worldwide, there are some research gaps in developing lightweight CNN models, particularly for Namibian plant datasets. Published research mainly relies on data from non-African countries, which may not translate well to the tomato leaf textures and diseases found in the Namibian region.

Different approaches and their classification accuracies, along with the datasets used in the study, are listed in **Table I**.

The present study investigates the application of the MobileNet model to a blended dataset [7].

Manuscript received on 28 April 2026 | Revised Manuscript received on 04 May 2026 | Manuscript Accepted on 15 May 2026 | Manuscript published on 30 May 2026.
*Correspondence Author(s)

Iita Anatolia, Student, Department of Electrical and Computer Engineering, University of Namibia, Ongwediva, Namibia. Email ID: iitaapn@gmail.com

Srinu Sesham*, Senior Lecturer, Department of Electrical and Computer Engineering, University of Namibia, Ongwediva, Namibia. Email ID: ssesham@unam.na, ORCID ID: 0000-0002-2996-6023.

Mateus Abisai, Lecturer, Department of Electrical and Computer Engineering, University of Namibia, Ongwediva, Namibia. Email ID: mabisai@unam.na

Kenneth Gideon, Lecturer, Department of Electrical and Computer Engineering, University of Namibia, Ongwediva, Namibia. Email ID: kgideon@unam.na

Table I: Literature Review on the Recent Performance of a Multi-Class Disease Model on the Tomato Crop

Authors (Year)	Experimental Setup	Dataset Used / Source	Major Findings / Contribution
Wajid et al. (2023) [6]	CapsNet for 10 disease classes	Standard tomato dataset (PlantVillage)	Accuracy 96.4%; demonstrates CapsNet effectiveness
Sun et al. (2025) [8]	CSWinT + multi-kernel local/global fusion	Custom dataset (field-collected tomato leaves)	mAP50 97.2%; effective under complex backgrounds
Dhakyan aikabdel et al. (2023) [9]	Transformer and CNN model comparison	PlantVillage dataset	DenseNet121 accuracy 99%; ViT shows competitive results
Sowmya et al (2025) [10]	Inception v4 + YOLOv8 for tomato disease classification/ localization	Custom-curated tomato dataset	Accuracy 96%, mAP@0.5 86%; supports real-time detection
Howard et al. (2025) [11]	MobileNetV2 + residual CNN	Public PlantVillage dataset	Lightweight, explainable CNN; suitable for mobile deployment
Zhong et al (2023) [13]	Lightweight CNN for embedded/mobile deployment	Custom field dataset (9,000 annotated images)	High accuracy with minimal parameters; edge-device friendly
Prama et al. (2023) [14]	Comparative study on tomato & corn datasets	PlantVillage + custom corn dataset	Tomato accuracy 95.1%;
Gulzar (2023) [15]	Customized MobileNetV2 is developed and applied	40 Fruit data for multi-class classification	Achieved 99% accuracy. However, it is not applied to Tomato leaf disease classification accuracy.
Adebisi et al [16]	Five CNN models are trained and implemented in an FPGA for Potato disease detection.	Potato Leaves Dataset taken from the Public PlantVillage dataset	Achieved 98.8% accuracy. But the classification problem has only three classes, and the models are not resource-efficient.
Zhang et al (2024) [17]	Dual attention (channel + spatial) CNN	Public PlantVillage dataset (imbalanced classes)	Addresses class imbalance; strong recognition in complex backgrounds
Wang et al. (2024) [18]	Attention-mechanism + multi-scale fusion (YOLOv6 backbone)	Roboflow dataset (PlantVillage + field images)	Precision 92.9%, Recall 95.2%, mAP 93.8%; effective for small lesions
This Work	VGG16, ResNet50, and MobileNetV2 were trained and fine-tuned for tomato leaf disease detection	Blended dataset: PlantVillage (Kaggle) + real-time Namibian farm images	MobileNetV2 reached about 90% accuracy with very low computational cost, making it suitable for real-time, mobile-based crop disease detection in Namibia's smallholder farming systems.

Here, the Plant Village tomato leaf dataset from Kaggle is combined with real-time images captured from small farms in Namibia, and its performance is compared with highly used DL models in the literature, such as VGG1 and ResNet50 [8]. This research focuses on developing a resource-efficient and reliable tool for early disease detection to help smallholder farmers improve crop management practices and reduce crop losses. Hence, we studied and applied MobileNetV2 and compared its performance with that of VGG16 and ResNet50. The MobileNetV2 model is lightweight because it uses a special type of convolution called depth-wise separable convolution, which reduces the number of computations compared to standard convolutions. This operation decouples filtering and combining, making the network faster and lighter. In addition, MobileNetV2 uses a linear bottleneck, which retains relevant features while reducing memory usage. These features enable MobileNetV2 for real-time applications and make it easy to implement on edge and mobile devices [12].

The following are the main aims of the study:

(1) to verify the effects of the pre-trained MobileNetV2 CNN model by blending the dataset for tomato disease types in real-time.

(2) to evaluate the accuracy vs model size of the MobileNetV2 model, compared to the VGG16 and ResNet50 models, to achieve an economical deep learning model.

The findings of this research will support the creation of a resource-efficient, automated and precise multiclass plant disease prediction for Namibian tomato crop farmers.

The rest of this paper is structured as follows. In Section 2, we describe the problem, including the use of a multiclass classification model to assess tomato plant health using a merged dataset. Section 3 details the proposed approach, data preprocessing, and the application of resource-efficient CNN models. Section 4 outlines the experimental setup and results from using common multi-class metrics. Finally, Section 5 concludes the paper, discussing the results, limitations and future work.

II. PROBLEM STATEMENT

The tomato crop is susceptible to bacterial and fungal diseases that can reduce yield if not detected early. Manual inspection of tomato leaves is labour-intensive, subjective, and impractical for large-scale monitoring [19]. This motivates the development of automated disease-detection frameworks that accurately classify leaf images into multiple disease categories. The problem is inherently a multi-class classification problem, where each leaf image belongs to one of K disease classes, including a healthy class.

Let the tomato leaf dataset be represented as [20]:

$$D = \{(X_i, y_i) \mid X_i \in \mathbb{R}^{H \times W \times 3},\ y_i \in \{1,2,\dots,K\}\}_{i=1}^{N},$$

where X_i is the $\mid$ ^{Th} input image of size $H \times W \times 3$, y_i is the corresponding ground truth label representing one of K disease classes, and N is the total number of images [21]. The objective is to learn a function

$$f_\theta : \mathbb{R}^{H \times W \times 3} \rightarrow [0,1]^K,$$

Retrieval Number: 100.1/ijrte.A836315010526
DOI: 10.35940/ijrte.A8363.15010526
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

parameterized by θ, which outputs a probability distribution over all classes for a given input image:

$$\hat{y}_i = f_\theta(X_i), \hat{y}_i \in \mathbb{R}^K, \sum_{k=1}^{K} \hat{y}_{i,k} = 1,$$

where $\hat{y}_{i,k}$is the predicted probability that image $X_i$belongs to the class k. The model parameters are optimized by minimizing the categorical cross-entropy loss:

$$\mathcal{L}(\theta) = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \mathbf{1}\{y_i = k\} \log(\hat{y}_{i,k}),$$

where $\mathbf{1}\{\cdot\}$is the indicator function. The optimal parameters $\theta^*$are obtained as:

$$\theta^* = \arg\min_\theta \mathcal{L}(\theta).$$

This problem can be represented through the following null and alternative hypotheses:

H_0 : $f_\theta(X)$ cannot correctly classify the specific disease present in a given leaf image.

H_1: $f_\theta(X)$ can accurately classify the disease class of a given input leaf image.

To discriminate above hypothesis, pre-trained CNNs (VGG16, ResNet50, and MobileNetV2) are used to obtain the hypothesis function f_θ , leveraging transfer learning for efficient feature extraction [22].

III. METHODOLOGY

A. Data Preprocessing

The preprocessing steps for the tomato leaf dataset are defined mathematically as follows [23].

All images were resized to the same dimensions to keep the dataset uniform:

$$\tilde{X}_i = R(X_i), \tilde{X}_i \in \mathbb{R}^{224\times224\times3} \tag{1}$$

where $R(\cdot)$is the resize operator and X_i *This* is the original image.

Then, the pixel intensities of the resized images are scaled within the range [0,1] to accelerate model convergence during training:

$$\tilde{X}_i^{norm} = \frac{\tilde{X}_i}{255}, \tilde{X}_i^{norm} \in [0,1]^{224\times224\times3} \tag{2}$$

Data augmentation is applied to the normalised images to improve model generalisation and reduce overfitting. This process involves a combination of geometric and photometric transformations [24]:

$$\tilde{X}_i^{aug} = A(\tilde{X}_i^{norm}), \tilde{X}_i^{aug} \in [0,1]^{224\times224\times3} \tag{3}$$

where the augmentation operator $A(\cdot)$is expressed as:

$$A(\tilde{X}_i^{norm}) = B \circ F \circ Z \circ R_{rot}(\tilde{X}_i^{norm}) \tag{4}$$

Here, R_{rot} represents rotation, Z denotes zoom, F corresponds to flipping, and B adjusts brightness.

The augmented dataset is subsequently partitioned into training and validation sets using a stratified approach, which preserves the class proportions across both sets:

$$D_{aug} = \{(\tilde{X}_i^{aug}, y_i)\}_{i=1}^{N} \tag{5}$$

$$(D_{train}, D_{val}) = \text{StratSplit}(D_{aug}, p), p = 0.8 \tag{6}$$

where

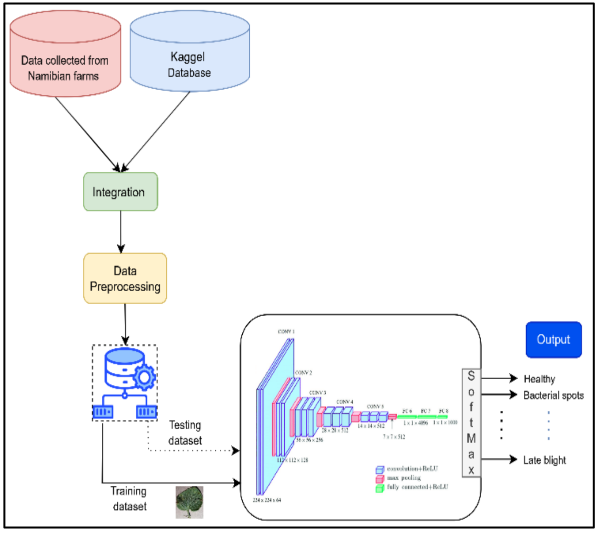
$$| D_{train} |= \lfloor p \cdot N \rfloor, | D_{val} |= N - | D_{train} | \tag{7}$$

and the stratified split ensures that for each class k, the proportion in the training and validation sets remains approximately the same:

$$\frac{\text{count}(y=k \in D_{train})}{|D_{train}|} \approx \frac{\text{count}(y=k \in D_{val})}{|D_{val}|}, \forall k \in \{1,2,\ldots,K\} \ (8) \qquad (8)$$

B. Multiclass Classification using VGG16

Fig. 1 shows the proposed framework for tomato leaf multiclass prediction using the VGG16 model. The architecture processes 224×224×3 leaf images through an initial 7×7 convolution and max-pooling layer. The backbone comprises four stages with [3, 4, 6, 3] bottleneck blocks, each using [1×1→3×3 →1×1] convolution patterns and residual connections, progressively increasing feature dimensions from 64 to 2048 channels. Global average pooling produces a 2048-dimensional vector, which is fed to a classification head with fully connected (FC) layers (2048 → 1024 → 512 → 256 → 10), using batch normalisation, dropout regularisation, and SoftMax activation for multi-class disease probability prediction across 10 tomato disease categories.



[Fig.1: Proposed Framework for Tomato Leaf Multiclass Prediction using VGG16 Architecture]

The algorithm begins with the preprocessed and augmented training day_hat_iy_hat_itaset $D_{train} = \{(\tilde{X}_i^{aug}, y_i)\}_{i=1}^{N_{train}}$ and the validation dataset $D_{val} = \{(\tilde{X}_i^{aug}, y_i)\}_{i=1}^{N_{val}}$. The goal is to predict the class labels $\hat{y}_i$ for the validation (test) images.

Let f_{VGG16} denote the VGG16 network pre-trained on ImageNet, configured without the top (fully connected) layers, and designed to process RGB images of size 224×224.

This can be expressed as [24]:

$$f_{VGG16} \leftarrow \text{VGG16}_{\text{ImageNet}}^{(\text{-top})}(224,224,3)$$

The convolutional base is frozen to perform feature extraction. For each augmented input image $\tilde{X}_i^{aug}$, the corresponding feature representation is obtained as:

$$F_i = f_{VGG16}(\tilde{X}_i^{aug}), \qquad F_i \in \mathbb{R}^{7\times7\times512}$$

To transform the feature maps into a one-dimensional (1D) vector, a global average pooling (GAP) layer is applied:

$$F_i^{flat} = \text{GAP}(F_i), \qquad F_i^{flat} \in \mathbb{R}^{512}$$

The flat features are processed by fully connected (FC) layers, followed by a Softmax function to obtain the classification output. This process can be expressed as:

$$Z_i = W_1 F_i^{flat} + b_1,$$

$$A_i = \text{ReLU}(Z_i),$$
$$A_i^{drop} = \text{Dropout}(A_i, p = 0.5),$$

$$\hat{y}_i = \text{Softmax}(W_2 A_i^{drop} + b_2), \hat{y}_i \in \mathbb{R}^K,$$

where, K represents the number of output/target classes, considered as 10.

The weights (model parameters) are optimized using cross-entropy loss L_{CE}, which can be expressed as:

$$\mathcal{L}_{CE} = -\frac{1}{N_T} \sum_{i=1}^{N_T} \sum_{k=1}^{10} y_i^{(k)} \log(\hat{y}_i^{(k)}), \tag{9}$$

where, N_T is the number of training samples.

The weights of the model are updated using the Adam optimizer, given by:

$$\theta_v \leftarrow \theta_v(t-1) - \eta \nabla_\theta \mathcal{L}, \tag{10}$$

where η denotes the learning rate in the Adam optimizer.

After initial training, the top convolutional blocks of VGG16 are unfrozen and fine-tuned with a smaller learning rate to adapt the pre-trained weights to the tomato disease dataset. The final prediction on the validation dataset is obtained as:

$$\hat{y}_{val} = \arg\max \hat{y}_i, \qquad i \in \mathcal{D}_{val}, \tag{11}$$

The overall accuracy using the VGG16 model is expressed as:

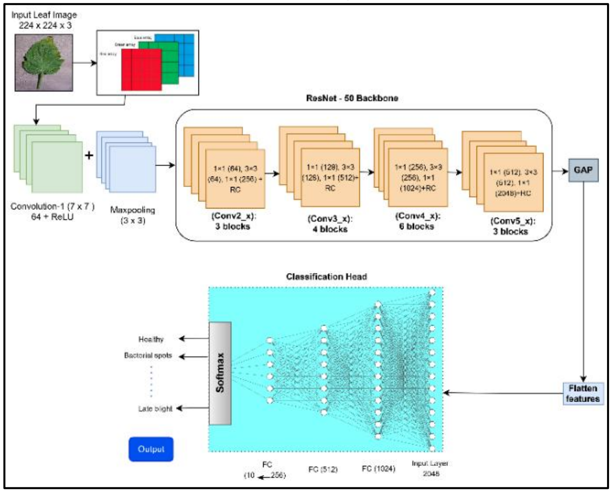
$$\text{Accuracy} = \frac{1}{N_{val}} \sum_{i=1}^{N_{val}} \mathbf{1}(\hat{y}_i = y_i). \tag{12}$$

C. Multiclass Classification using ResNet50

Fig. 2 shows the architecture of ResNet50 considered for multi-class tomato plant disease prediction [25]. The ResNet50 architecture processes 224 × 224 × 3 leaf images through an initial 7 × 7 convolution and max pooling layers. The backbone employs four stages with [3, 4, 6, 3] bottleneck blocks using $[1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1]$ convolution patterns and residual connections (RC), progressively increasing feature dimensions from 64 to 2048 channels. Global average pooling produces a 2048-dimensional vector, which is fed to a classification head with fully connected layers. $(2048 \rightarrow 1024 \rightarrow 512 \rightarrow 256 \rightarrow 10)$, using batch normalization, dropout regularization, and softmax activation for multi-class disease probability prediction across ten tomato disease categories.

A pre-trained ResNet50 model is loaded as:

$$f_{\text{ResNet50}} \leftarrow \text{ResNet50}_{\text{ImageNet}}^{(-\text{top})}(224,224,3)$$



[Fig.2: Architecture of ResNet50 Considered for Multiclass Disease Prediction]

Each residual block in the network computes its output as:

$$F_i^{l+1} = F_i^l + \mathcal{F}(F_i^l, W^l), \qquad l = 1, \ldots, L,$$

where $\mathcal{F}(\cdot)$ represents the residual mapping composed of convolution, batch normalization, and ReLU activation, and L denotes the total number of residual blocks [25].

After passing through all residual blocks, the feature representation of the $i -$ {Th} augmented image is obtained as:

$$F_i = f_{\text{ResNet50}}(\tilde{X}_i^{aug}), \qquad F_i \in \mathbb{R}^{7\times7\times2048}.$$

GAP operation is applied to get the 1D flat features:

$$F_i^{flat} = \text{GAP}(F_i), \qquad F_i^{flat} \in \mathbb{R}^{2048}.$$

Retrieval Number: 100.1/ijrte.A836315010526
DOI: 10.35940/ijrte.A8363.15010526
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The flattened feature vector is then fed into fully connected layers, followed by a Softmax
output layer:

$$Z_i = W_1 F_i^{flat} + b_1,$$
$$A_i = \mathrm{ReLU}(Z_i),$$

$$A_i^{drop} = \mathrm{Dropout}(A_i, p = 0.5),$$

$$\hat{y}_i = \mathrm{Softmax}\left(W_2 A_i^{drop} + b_2\right), \hat{y}_i \in \mathbb{R}^K,$$

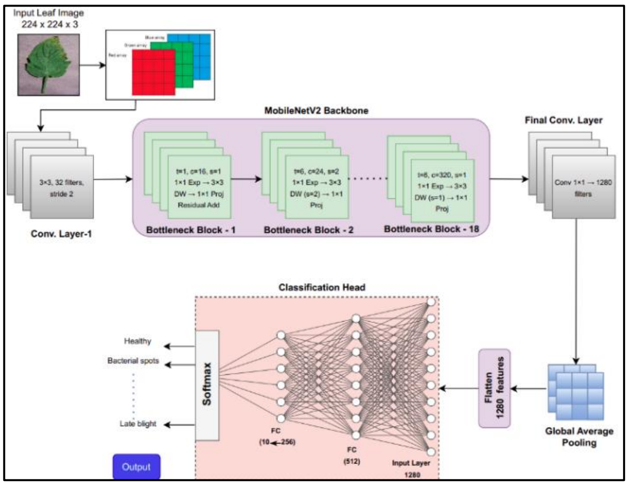
where $K = 10$ corresponds to the number of tomato disease classes.

The loss function is minimised during training according to Eq. (9), and the model parameters are updated using Eq. (10). After the initial training, the top residual blocks of ResNet50 are unfrozen and fine-tuned with a lower learning rate to adapt the pre-trained features to the tomato leaf dataset. The final predictions on the validation set are computed using the overall accuracy obtained with the ResNet50.

D. Proposed Multiclass Classification using MobileNetV2

The architecture of MobileNetV2 used for multiclass prediction is shown in Fig. 3. The proposed tomato disease classification system uses a MobileNetV2 model for leaf images [26]. The architecture takes images of size 224 × 224 × 3 using a first convolutional layer with 32 filters of size 3 × 3 and stride 2 to capture initial spatial features. The main feature extraction is based on the bottleneck residual blocks with inverted residuals and linear bottlenecks, where each block is a sequence of pointwise convolution for channel expansion, depth-wise convolution for spatial feature processing, and pointwise convolution for channel expansion. This efficient design allows the network to learn hierarchical representations while being computationally efficient. After 18 such bottleneck blocks, a final 1 × 1 convolution layer increases the channels to 1280, which are then aggregated via global average pooling.

The classification head consists of three fully connected layers with ReLU6 activations and dropout regularization, progressively reducing dimensionality from 1280 to 512, then to 256, and finally to K output neurons corresponding to disease classes, with softmax activation producing final probability distributions for multi-class tomato disease prediction [20].



[Fig.3: Architecture of MobileNetV2 Considered for

Multiclass Prediction [26]]

The input leaf image is represented as:

$$\boldsymbol{X} \in \mathbb{R}^{224\times224\times3}$$

where each pixel has RGB channels normalized to [0,1].

A pre-trained MobileNetV2 model is initialized as:

$$f_{\text{MobileNetV2}} \leftarrow \text{MobileNetV2}_{\text{ImageNet}}^{(-\text{top})}(224,224,3)$$

The initial convolution layer operates as [20]:

$$Z_0 = Conv2D(X; W_0, b_0), Y_0 = BatchNorm(Z_0), H_0 = ReLU6(Y_0)$$

where $W_0 \in \mathbb{R}^{3\times3\times3\times32}$ are convolutional filters, $b_0 \in \mathbb{R}^{32}$ are bias terms, and output $H_0 \in \mathbb{R}^{112\times112\times32}$. The ReLU6 activation function is defined as:

$$ReLU6(x) = min(max(0,x),6)$$

Providing bounded activation to improve quantisation performance.

For each bottleneck block $i = 1,2,\ldots,17$:

Expansion phase with pointwise convolution:

$$Z_i^{exp} = W_i^{exp} * H_{i-1} + b_i^{exp}, H_i^{exp} = ReLU6(BatchNorm(Z_i^{exp}))$$

Depthwise convolution:

$$Z_i^{dw} = W_i^{dw} \odot H_i^{exp} + b_i^{dw}, H_i^{dw} = ReLU6(BatchNorm(Z_i^{dw}))$$

where $\odot$ denotes depthwise convolution.

Projection phase with pointwise convolution:

$$Z_i^{proj} = W_i^{proj} * H_i^{dw} + b_i^{proj}, H_i^{proj} = BatchNorm(Z_i^{proj})$$

Residual connection:

$$H_i = \begin{cases} H_i^{proj} + H_{i-1} \text{if stride} = 1 \text{ and same dimensions} \\ H_i^{proj} \text{otherwise} \end{cases}$$

Final convolution and global pooling:

$$Z_{final} = W_{final} * H_{18} + b_{final}, H_{final} = ReLU6(BatchNorm(Z_{final}))$$

$$g = GAP(H_{final}) = \frac{1}{7\times7}\sum_{i=1}^{7}\sum_{j=1}^{7} H_{final}\,(i,j,:)$$

$$g \in R^{1280}$$

Custom classification head for $K = 10$classes:
First dense layer:

$$z_1 = W_1 g + b_1, h_1 = ReLU6(z_1), h_1^{drop} = Dropout(h_1; p = 0.2)$$

where $W_1 \in \mathbb{R}^{512\times1280}$, $b_1 \in \mathbb{R}^{512}$.

Second dense layer:

$$z_2 = W_2 h_1^{drop} + b_2, h_2 = ReLU6(z_2), h_2^{drop} = Dropout(h_2; p = 0.2)$$

where $W_2 \in \mathbb{R}^{256\times512}$, $b_2 \in \mathbb{R}^{256}$.
Final output layer:

$$z_3 = W_3 h_2^{drop} + b_3$$

where $W_3 \in \mathbb{R}^{10\times256}$, $b_3 \in \mathbb{R}^{10}$.
SoftMax activation for multi-class classification:

$$P(y = k \mid X) = \frac{\exp(z_{3,k})}{\sum_{j=1}^{10} \exp(z_{3,j})}$$

for $k = 1,2,\ldots,10$

The loss function uses categorical cross-entropy:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{10} y_{i,k}\log(P(y_i = k \mid X_i))$$

where N is the batch size, $y_{i,k}$ is the ground truth indicator, and θ represents all trainable parameters.

IV. SIMULATION RESULTS

The experimental simulations are carried out using a blended dataset, comprising the PlantVillage tomato leaf dataset and images captured from Namibian farms in the Oshana region. The dataset contains 14,529 labelled images spanning 10 disease classes, including tomato mosaic virus, leaf mould, healthy leaves, early blight, spider mite infestation, septoria leaf spot, late blight, tomato yellow leaf curl virus, target spot, and bacterial spot. The dataset was divided into training and validation sets in a 4:1 ratio, with 11,623 samples for training (80%) and 2,906 for validation (20%), using stratified sampling to preserve the class distribution. Image preprocessing was performed using an image data generator that applied normalisation, rotation, flipping, and zooming augmentations to improve model generalisation. Three transfer learning architectures, such as VGG16, ResNet50, and MobileNetV2, which are pre-trained on ImageNet, are evaluated. Each model undergoes a two-phase training process: first, the convolutional base is frozen to extract features from the input images, and subsequently, the top layers are unfrozen and fine-tuned to adapt the pre-trained weights to the target tomato leaf dataset. The models are optimised using the widely used Adam optimiser (Eq. (10), and class weights are applied to mitigate class imbalance across disease categories. The following standard performance metrics are used to evaluate the trained models.

The overall Accuracy (Acc) measures the proportion of correctly predicted samples and is defined as [27],

$$Acc = \frac{\sum_{i=1}^{K} TP_i}{\sum_{i=1}^{K}(TP_i + FP_i + FN_i + TN_i)}, \quad (13)$$

where TP_i, FP_i, FN_i, and TN_i represent the true positives, false positives, false negatives, and true negatives for class i, and $K = 10$ is the total number of classes.

Precision (P_i) quantifies the proportion of correctly identified instances among all predictions for class i:

$$P_i = \frac{TP_i}{TP_i + FP_i}. \quad (14)$$

Recall (R_i), also known as sensitivity, measures the fraction of correctly predicted positive samples for each class:

$$R_i = \frac{TP_i}{TP_i + FN_i}. \quad (15)$$

The F1-score, a harmonic mean of precision and recall, provides a balanced measure of performance:

$$F1_i = \frac{2P_iR_i}{P_i + R_i}. \quad (16)$$

For the multiclass case, the macro-averaged F1-score, precision, and recall are calculated as [25],

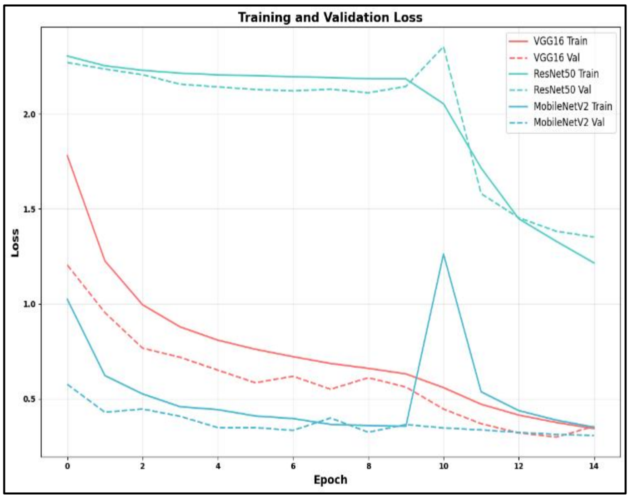
$$F1_{macro} = \frac{1}{K}\sum_{i=1}^{K} F1_i, \; P_{macro} = \frac{1}{K}\sum_{i=1}^{K} P_i,$$

$$R_{macro} = \frac{1}{K}\sum_{i=1}^{K} R_i \quad (17)$$

Additionally, the Area Under the Curve (AUC) and Average Precision (AP) are macro-averaged across all classes to measure overall discriminative performance and resilience to class imbalance. The performance metrics are used in the experimental evaluation (see Results) and in the interpretation of the results and comparative analyses in this study.

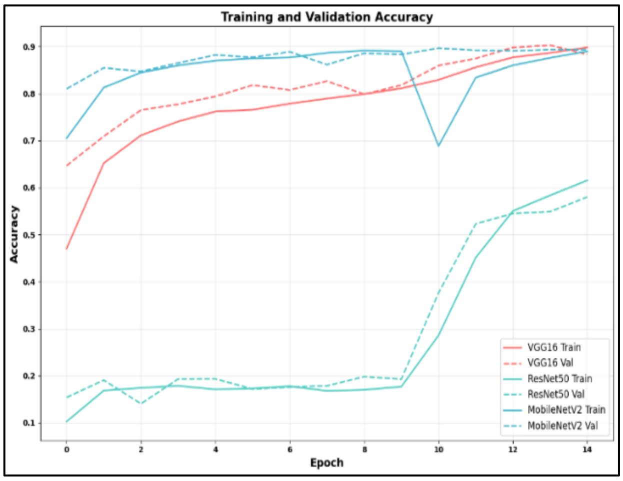
Fig. 4 shows the training and validation Loss vs epochs for VGG16, MobileNetV2 and ResNet50. A steady decline in loss (training and validation) is observed in the VGG16 model (indicating good convergence) and in the MobileNetV2 model (moderate convergence). ResNet50, on the other hand, exhibited slower convergence and a higher loss value, suggesting that the model may have under-fitted the tomato leaf dataset due to its deeper architecture and larger parameter count. This behaviour is in line with the loss term, where a best-performing model should minimise the number of misclassifications across all classes, thereby reducing the denominator term for false alarms.

Fig. 5 illustrates the progression of accuracy (training and validation). VGG16 achieved the highest validation accuracy of approximately 90%, followed by MobileNetV2, whereas ResNet50 underperformed with an accuracy below 30% after fine-tuning. The accuracy behaviour matches Equation (1), where a higher ratio of true positives to total predictions yields improved performance. This trend indicates that VGG16 effectively captured discriminative spatial features for disease classification, whereas deeper models struggled to adapt due to the limited dataset size relative to their depth.



[Fig.4: Training and Validation Error curves for VGG16, ResNet50 and MobileNetV2 Over Epochs]

The evolution of accuracy (training and validation) is shown in Fig. 5. VGG16, followed by MobileNetV2, achieved the highest validation accuracy (around 90%), while ResNet50 performed poorly (accuracy below 30%) after fine-tuning. The accuracy trend is consistent with Equation (13): a higher true positive rate leads to higher accuracy. This suggests that VGG16 learned discriminative spatial features for disease classification, whereas deeper models struggled to adapt because the dataset was too small relative to their complexity. The VGG16 accuracy curve is more linear than the others, indicating a stable learning process with high P_macro and R_macro values. **Figs.** 6, 7, and 8 show the Confusion Matrices for each model, showing class-wise prediction accuracy. The VGG16 matrix, with strong diagonal elements, indicates it can identify most disease classes, especially the yellow leaf curl virus and healthy plants. ResNet50 was more likely to misclassify visually similar classes, such as Early Blight and Target Spot, confirming its poorer generalisation. The lower F1-scores in certain classes are due to high false positives or false negatives, leading to lower precision and recall. As a result, the macro-average F1 scores of VGG16 reveal its strong classification performance across all classes.

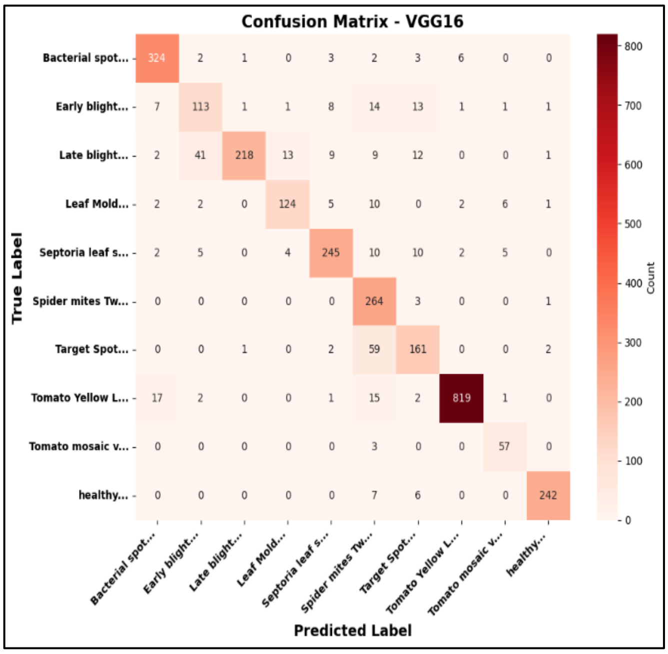


[Fig.5: Training and Validation Accuracy Curves of VGG16, ResNet50, and MobileNetV2]

Table II: Performance Comparison of CNN Models for Tomato Leaf Disease Classification

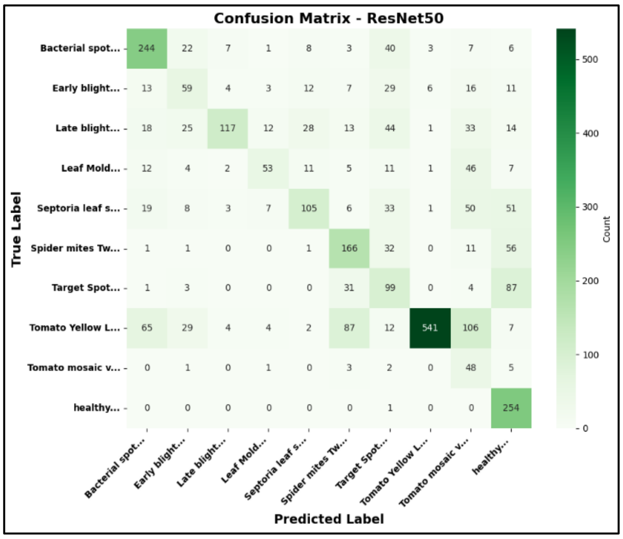
Model	**Accuracy (%)**	**Precision**	**Recall**	**F1-Score**	**Parameters (M)**	**Training Time (s)**
VGG16	88.33	0.8965	0.8833	0.8845	14.782	2639.94
ResNet50	58.02	0.6926	0.5802	0.5976	23.851	2374.35
MobileNetV2	89.83	0.9005	0.8933	0.8937	2.423	2262.20

Table II compares VGG16, ResNet50, and MobileNetV2, which were used to classify diseases on tomato leaves. The table provides various measures of model performance in terms of accuracy, precision, recall, and F1 score, in addition to model complexity expressed as the number of parameters (millions) and training time in seconds. Among the models, MobileNetV2 achieved the highest accuracy of 89.83%, followed by VGG16 (88.33%) and ResNet50 (58.02%). As for precision and recall, MobileNetV2 also performed better, with values of 0.9005 and 0.8933, respectively. VGG16 ranked second on both performance measures, while ResNet50 lagged on both. As regards computational resources, MobileNetV2 used fewer parameters (2.423 M) than both VGG16 (14.782 M) and ResNet50 (23.851 M). Moreover, the latter model took significantly less time to train than VGG16 and ResNet50, slightly faster than the other two models. Furthermore, MobileNetV2 trained much faster than the other two models, making it a lightweight yet accurate model for real-time detection of tomato leaf diseases in the Namibian field.



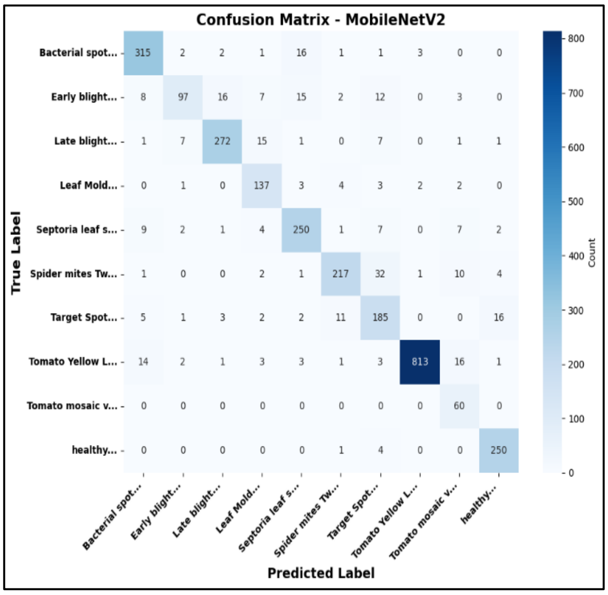
	Bacterial spot...	Early blight...	Late blight...	Leaf Mold...	Septoria leaf s...	Spider mites Tw...	Target Spot...	Tomato Yellow L...	Tomato mosaic v...	healthy...
Bacterial spot...	324	2	1	0	3	2	3	6	0	0
Early blight...	7	113	1	1	8	14	13	1	1	1
Late blight...	2	41	218	13	9	9	12	0	0	1
Leaf Mold...	2	2	0	124	5	10	0	2	6	1
Septoria leaf s...	2	5	0	4	245	10	10	2	5	0
Spider mites Tw...	0	0	0	0	0	264	3	0	0	1
Target Spot...	0	0	1	0	2	59	161	0	0	2
Tomato Yellow L...	17	2	0	0	1	15	2	819	1	0
Tomato mosaic v...	0	0	0	0	0	3	0	0	57	0
healthy...	0	0	0	0	0	7	6	0	0	242

[Fig.6: Confusion Matrix of VGG16 on the Validation Dataset]



	Bacterial spot...	Early blight...	Late blight...	Leaf Mold...	Septoria leaf s...	Spider mites Tw...	Target Spot...	Tomato Yellow L...	Tomato mosaic v...	healthy...
Bacterial spot...	244	22	7	1	8	3	40	3	7	6
Early blight...	13	59	4	3	12	7	29	6	16	11
Late blight...	18	25	117	12	28	13	44	1	33	14
Leaf Mold...	12	4	2	53	11	5	11	1	46	7
Septoria leaf s...	19	8	3	7	105	6	33	1	50	51
Spider mites Tw...	1	1	0	0	1	166	32	0	11	56
Target Spot...	1	3	0	0	0	31	99	0	4	87
Tomato Yellow L...	65	29	4	4	2	87	12	541	106	7
Tomato mosaic v...	0	1	0	1	0	3	2	0	48	5
healthy...	0	0	0	0	0	0	1	0	0	254

[Fig.7: ResNet50 Confusion Matrix on the Validation Set]

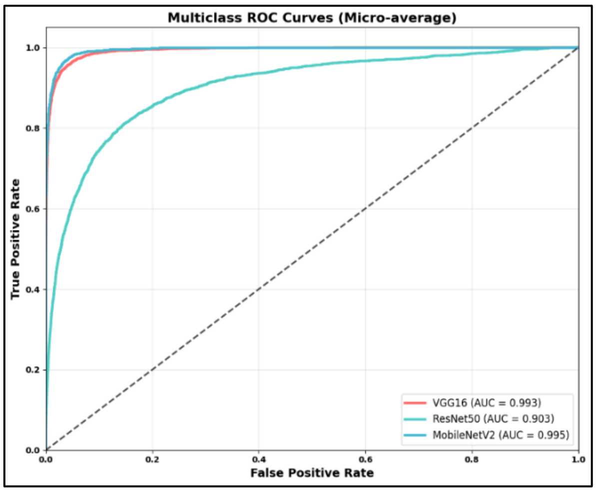


	Bacterial spot...	Early blight...	Late blight...	Leaf Mold...	Septoria leaf s...	Spider mites Tw...	Target Spot...	Tomato Yellow L...	Tomato mosaic v...	healthy...
Bacterial spot...	315	2	2	1	16	1	1	3	0	0
Early blight...	8	97	16	7	15	2	12	0	3	0
Late blight...	1	7	272	15	1	0	7	0	1	1
Leaf Mold...	0	1	0	137	3	4	3	2	2	0
Septoria leaf s...	9	2	1	4	250	1	7	0	7	2
Spider mites Tw...	1	0	0	2	1	217	32	1	10	4
Target Spot...	5	1	3	2	2	11	185	0	0	16
Tomato Yellow L...	14	2	1	3	3	1	3	813	16	1
Tomato mosaic v...	0	0	0	0	0	0	0	0	60	0
healthy...	0	0	0	0	0	1	4	0	0	250

[Fig.8: MobileNetV2 Confusion Matrix on the Validation Set]

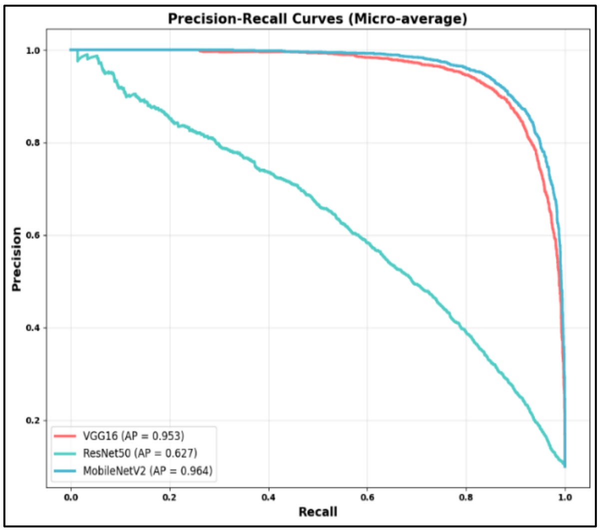
[Fig. 9](#) shows the Receiver Operating Characteristic (ROC) curves, in which the AUC values under micro-averaging quantify the models' discriminative power. VGG16 obtained the highest AUC (≈ 0.97), followed by MobileNetV2 (≈ 0.94), while ResNet50 had a significantly lower AUC (≈ 0.65), illustrating that ResNet50 was not very discriminative, and its sensitivity-specificity trade-off was very low. The AUC metric measures the model's discriminative power by optimising the True Positive rate (TPi) and minimising the false positive rate (FPi).

Therefore, a higher AUC indicates an optimal Precision-Recall balance, which is vital for disease classification in plant pathology.



[Fig.9: ROC Curve Comparison Between VGG16, ResNet50, and MobileNetV2]

[Fig. 10](#) below shows the Precision-Recall (PR) curves, especially helpful when there is an imbalance between data classes. Model VGG16 achieved the best average precision (AP ≈ 0.96), indicating that it can effectively detect diseases from minority classes. The improvement in precision at higher recall indicates that the model is effective in reducing false positives without sacrificing recall.



[Fig.10: Precision-Recall (PR) Curves for VGG16, ResNet50, and MobileNetV2]

V. CONCLUSIONS

In this research, three pre-trained lightweight CNNs (VGG16, ResNet50, and MobileNetV2) were evaluated for multiclass classification of tomato plant diseases by combining the Plant Village tomato leaf database with image data collected from farms in Namibia. This experiment was an attempt to address some problems in the rapid identification of diseases based on local conditions. Three models were fine-tuned on a blended dataset. Experimental results showed that the proposed MobileNetV2 achieved the highest validation accuracy of approximately 90%, outperforming VGG16 (88.33%) and ResNet50 (58.02%), while also being computationally efficient with only 2.42 million parameters. Whereas VGG16, though slightly less accurate, demonstrated stable convergence and excellent class-wise discrimination, as evidenced by ROC and Precision-Recall curves. ResNet50 underperformed due to its deeper architecture, which was less adaptable to the dataset size, leading to higher misclassification rates among visually similar disease classes. Table 2 shows that VGG16 (14.78 million parameters) and ResNet50 (23.81 million parameters) are not efficient enough, which limits their suitability for edge devices. Comparative metrics confirmed that lightweight architectures such as MobileNetV2 are well-suited for resource-constrained environments while maintaining robust disease-classification performance. The performance of the considered DL models is slightly lower (approximately 5%) than in the literature. This is due to the blended dataset, which contains real-time tomato leaf images from farms in the Oshana region. Overall, the findings highlight that transfer learning with lightweight CNN models can effectively enable real-time, scalable, and cost-effective tomato disease monitoring. In particular, MobileNetV2 offers a practical solution for deployment on mobile or embedded devices, empowering Namibian farmers through digital technologies. Future work could explore advanced YOLO based models for faster and more efficient disease classification in real-world field settings.

DECLARATION STATEMENT

As the article's author, I must verify the accuracy of the following information after aggregating input from all authors.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted objectively and without external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** Each author has individually contributed to the article. **Iita Anatolia** contributed to the development and fine-tuning of lightweight CNN models, the implementation of the blended-dataset pipeline, and data analysis for multiclass classification of tomato leaf diseases. **Srinu Sesham** conceptualized the study, designed the experiments, data analysis, and contributed to writing and revising the manuscript. **Mateus Abisai** was responsible for model performance validation, comparative analysis of VGG16, ResNet50, and MobileNetV2, and revising the manuscript. **Kennth Gideon:** Responsible for Model performance validation and proofreading.

REFERENCES

1. I. Pacal et al., "A systematic review of deep learning techniques for plant diseases," Artificial Intelligence Review, 2024, Vol. 57, article 304. DOI: https://doi.org/10.1007/s10462-024-10944-7
2. Namibian Agronomy Board, Agronomy and Horticulture Development Division, An analysis of the impacts of climate change on the Namibian crop industry, Research Report, Mar. 2024. DOI: https://share.google/qdpXYtNG5aWwYuC6K
3. A. Fortunato and S. Enciso, "Food for Growth: A Diagnostics of Namibia's Agriculture Sector," *CID Research Fellows and Graduate Student Working Paper Series*, no. 2023.154, Harvard University, Cambridge, MA, Oct. 2023. https://dash.harvard.edu/entities/publication/41235843-1410-41e7-8839-c967d2cb55e9
4. S. Chaitanya P. *et al.*, "Plant Disease Detection Using CNN," *Data Science and Intelligent Computing Techniques*, 2023. ISBN: 978-81-955020-2-8. DOI: https://doi.org/10.56155/978-81-955020-2-8-5
5. B. S. Hamed, M. M. Hussein and A. M. Mousa, "Plant Disease Detection Using Deep Learning," *International Journal of Intelligent Systems and Applications*, 2023, Vol. 15, No. 6, pp. 38-50. DOI: https://doi.org/10.5815/ijisa.2023.06.04
6. A. H. Wajid *et al.*, "Tomato Leaf Disease Detection and Classification Using Convolutional Neural Network and Machine Learning," *Journal of Computing & Biomedical Informatics*, 2024. DOI: https://jcbi.org/index.php/Main/article/view/348
7. D. Andrianto, R. Prathivi and M. Liu, "Comparison of MobileNet and CNN Methods for Identifying Tomato Leaf Diseases," *Journal of Teknik Informatika (Jutif)*, 2024. DOI: https://doi.org/10.52436/1.jutif.2024.5.6.3477
8. H. Sun *et al.*, "Efficient deep learning-based tomato leaf disease detection through global and local feature fusion," *BMC Plant Biology*, 2025, Vol. 25, 311. DOI: https://doi.org/10.1186/s12870-025-06247-w
9. K. Dhakyanaikabdel and A. J, "A Hybrid CNN–Vision Transformer Framework for Optimized Leaf Disease Detection Using Feature Fusion and Transfer Learning," *Journal of Robotics and Control (JRC)*, vol. 6, no. 6, pp. 3002–3012, Jan. 2026. https://journal.umy.ac.id/index.php/jrc/article/view/28426
10. B. Sowmya and S. Guruprasad, "Deep learning-based plant health disease detection in tomatoes using Inception v4 convolutional neural network and YOLO v8," *Discover Artificial Intelligence*, 2025, Vol. 5, article 278. DOI: https://doi.org/10.1007/s44163-025-00540-1
11. A. G. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017. DOI: https://doi.org/10.48550/arXiv.1704.04861
12. A. Picon, *et al.*, "Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild," *Computers and Electronics in Agriculture*, vol. 161, pp. 280–290, 2019. DOI: https://doi.org/10.1016/j.compag.2018.04.002
13. Z. Zhong, Z. Teng, and M. Tong, "LightMixer: A novel lightweight convolutional neural network for tomato disease detection," *Frontiers in Plant Science*, vol. 14, p. 1166296, 2023. DOI: https://doi.org/10.3389/fpls.2023.1166296
14. Prama, T. T. and Oni, M. K., "Optimized Custom CNN for Real-Time Tomato Leaf Disease Detection," *ICCS 2025, Lecture Notes in Computer Science*, Vol. 15909, Springer, Cham. DOI: https://doi.org/10.1007/978-3-031-97564-6_6
15. Y. Gulzar, "Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique," *Sustainability*, 2023, Vol. 15, No. 3, article 1906. DOI: https://doi.org/10.3390/su15031906
16. J. Adebisi, S. Srinu, and V. Mitonga, "Deep learning algorithm analysis of potato disease classification for system-

on-chip implementation," *Journal of Digital Food, Energy & Water Systems*, vol. 5, no. 1, 2024. DOI: https://doi.org/10.36615/4gv6hf98

17. E. Zhang, N. Zhang, F. Li, and C. Lv, "A lightweight dual-attention network for tomato leaf disease identification," *Frontiers in Plant Science*, vol. 15, p. 1420584, 2024. DOI: https://doi.org/10.3389/fpls.2024.1420584
18. Y. Wang, P. Zhang and S. Tian, "Tomato leaf disease detection based on attention mechanism and multi-scale feature fusion," *Frontiers in Plant Science*, 2024, Vol. 15, article 1382802. DOI: https://doi.org/10.3389/fpls.2024.1382802
19. E. C. Too *et al.*, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, 2019, Vol. 161, pp. 272–279. DOI: https://doi.org/10.1016/j.compag.2018.03.032
20. I. Anatolia, S. Sesham, M. Abisai and K. Gideon, "MobileNetV2 Model for Multiclass Tomato Disease Prediction Using Blended Dataset for Smart Farming," *Second IEEE International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*, India, 2026, pp. 1-6. DOI: https://doi.org/10.1109/ICISCoIS62701.2026.11447556
21. Attallah, O. Tomato Leaf Disease Classification via Compact Convolutional Neural Networks with Transfer Learning and Feature Selection. *Horticulturae* **2023**, *9*, 149. DOI: https://doi.org/10.3390/horticulturae9020149
22. L. K. Ndovie and E. Masabo, "Leveraging MobileNetV3 for in-field tomato disease detection in Malawi via CNN," *SAIEE Africa Research Journal*, vol. 115, no. 3, pp. 74–85, 2024. http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1991-16962024000300001
23. M. Akay *et al.*, "Deep Learning Classification of Systemic Sclerosis Skin Using the MobileNetV2 Model," in *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 2, pp. 104-110, 2021, https://ieeexplore.ieee.org/document/9380371
24. M. Tariq *et al.*, "Corn leaf disease: insightful diagnosis using VGG16 empowered by explainable AI," *Frontiers in Plant Science*, 2024, Vol. 15, article 1402835. DOI: https://doi.org/10.3389/fpls.2024.1402835
25. M. Alruwaili and M. Mohamed, "An integrated deep learning model with EfficientNet and ResNet for accurate multi-class skin disease classification," *Diagnostics*, 2025, Vol. 15, No. 5, article 551. DOI: https://doi.org/10.3390/diagnostics15050551
26. M. Sandler *et al.*, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520. DOI: https://doi.org/10.1109/CVPR.2018.00474
27. S. Farhadpour, T. A. Warner, and A. E. Maxwell, "Selecting and interpreting multiclass loss and accuracy assessment metrics for classifications with class imbalance: Guidance and best practices," *Remote Sensing*, vol. 16, p. 533, 2024. DOI: https://doi.org/10.3390/rs16030533

AUTHOR'S PROFILE

Iita Anatolia is an M.Sc. student in the Department of Electrical and Computer Engineering, Faculty of Agriculture, Engineering and Natural Sciences, at the University of Namibia, Namibia. She received her B.Sc. Honours degree in Electronics and Computer Engineering from the same institution in Namibia. Her research interests include smart agriculture systems, artificial intelligence, deep learning model development, embedded systems, Internet of Things (IoT) applications, and intelligent data-driven solutions for precision farming. She is actively involved in research focused on applying AI and modern engineering technologies to improve agricultural productivity, automate operations, enhance sustainability, and develop real-time monitoring systems for smart farming applications in Namibia.

Srinu Sesham is a Senior Lecturer in the Department of Electrical and Computer Engineering, Faculty of Agriculture, Engineering and Natural Sciences, at the University of Namibia. He is a PhD recipient from UOH, India, and a postdoctoral research fellow at the University of Cape Town for 2 years. He is an accomplished researcher and author with publications in peer-reviewed journals and international conferences. He also serves as a reviewer for IEEE, Elsevier, Springer Nature journals, and international conferences. His research interests include physical layer challenges in Cognitive radio networks, digital signal and image processing, UAV airspace management, AI and machine learning models, Quantum communications, and embedded system design using FPGA-based platforms.

Abisai Mateus is a Lecturer in the Department of Electrical and Computer Engineering within the Faculty of Agriculture, Engineering and Natural Sciences at the University of Namibia. He is a self-motivated professional with over 13 years of experience in computer engineering, including strong backgrounds in computer programming, database systems, software engineering, high-performance computing, and operating systems. His research interests include database systems, mobile and web application development, machine learning, and computer networks.

Kennth Gideon is a Lecturer in the Department of Electrical and Computer Engineering, Faculty of Agriculture, Engineering and Natural Sciences, at the University of Namibia, Namibia. He received both his B.Sc. Honours degree and Master's degree in Electronics and Computer Engineering from the University of Namibia, where he graduated with distinction in his Master's studies. He is actively involved in teaching, research, and academic development in engineering and emerging technologies. His research interests include artificial intelligence applications, embedded systems design, robotics development, automation technologies, Internet of Things (IoT) systems, and intelligent control systems. He also serves as a reviewer for renowned international IEEE conferences and peer-reviewed journals.
