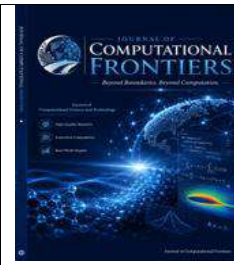
GYRES
Scholarly Press

Journal of Computational Frontiers
Volume 01, Issue 01, 2026
Journal Web Page: https://gyrespress.com/jcf



Comparative Evaluation of Deep Learning Models for Maize Leaf Disease Classification Using ResNet50, EfficientNet-B0 and MobileNetV3

Darius Tienhua Chinyio, *Orseer Coliins Nor, Chat Darius Chinyio
Department of Computer Science of the Nigerian Defence Academy, Nigeria

Article Info

Received: 10 July 2026
Received in revised: 25 August 2026
Accepted: 02 September 2026
Available Online:05 September 2026

Keywords

Maize leaf disease, Deep learning, EfficientNet-B0, Image classification, Transfer learning

Abstract

Maize production is significantly affected by various pests and diseases, making early and accurate disease identification important for improving crop productivity. This study evaluates the performance of transfer-learning-based deep learning models for automated maize leaf disease classification. A publicly available maize disease dataset containing multiple disease classes was used, and the images were preprocessed and divided into training, validation, and testing datasets. Three convolutional neural network architectures, ResNet50, EfficientNet-B0, and MobileNetV3, were trained and evaluated using accuracy, precision, recall, F1-score, classification reports, and confusion matrices. The results showed that EfficientNet-B0 achieved the best overall performance, with 97% accuracy, 0.98 precision, 0.97 recall, and 0.97 F1-score. MobileNetV3 achieved 95.05% accuracy, while ResNet50 achieved 94% accuracy. The confusion matrices further indicated that EfficientNet-B0 produced fewer major classification errors and demonstrated better generalization. The findings indicate that EfficientNet-B0 is the most suitable of the evaluated models for automated maize leaf disease classification and provides a promising foundation for intelligent agricultural disease-diagnosis systems.

1.INTRODUCTION

Maize (*Zea mays L.)* is one of the most vital food and industrial crops for human beings and ranks among the most essential cereal crops across the globe. It is widely recognized as a leading cereal by production volume and plays a central role in food security, particularly in regions where it serves as a staple [1,2]. In addition to its edible value, maize serves as a major raw material for industrial products such as starch, ethanol, and other bio-based derivatives, as well as a primary source of animal fodder and feed [1,2].

However, maize is highly susceptible to a wide range of pests and diseases. These biotic stresses substantially undermine crop yield and quality, with recent estimates indicating annual yield losses of 15–20% attributable to pests and diseases [2,3]. Such losses pose ongoing challenges to agricultural productivity and economic returns for farmers worldwide.

Early detection is a critical strategy for limiting the spread of pest and disease outbreaks and minimizing associated yield damage. Conventional expert-based identification, however, is time-consuming, labour-intensive, subjective, and costly, particularly when continuous or large-scale field monitoring is required [1,2, 4]. These limitations have stimulated substantial research interest in automated solutions.

In recent years, computer vision and machine learning (including deep learning) techniques have attracted considerable attention for the accurate, rapid, and scalable detection of infected plants. A growing body of work has demonstrated the effectiveness of convolutional neural networks, YOLO-based detectors, hybrid architectures, and related approaches for maize pest and disease identification under both controlled and complex field conditions [1-5]. These advances highlight the potential of intelligent systems to overcome the shortcomings of traditional diagnostic methods and support timely, precision crop protection.

Building on these developments, the present study aims to develop, evaluate and improved deep-learning models for the detection of maize pests and diseases.

Corresponding author: nororseercollins@gmail.com

2.Related works

Mapamela et al. [6] investigated maize disease detection using deep learning techniques, proposing the use of ResNet-9 and EfficientNet-B4 for image-based classification of maize leaf conditions. The researchers developed their models using 16,729 maize images representing common rust, northern leaf blight, lethal necrosis, streak virus, Cercospora leaf spot and healthy leaves. The images were obtained from multiple repositories and subjected to preprocessing, resizing and augmentation before being divided into training, validation and testing datasets. ResNet-9 achieved an accuracy of 98.2%, compared with 94.3% for EfficientNet-B4, demonstrating the effectiveness of deep learning for automated maize leaf classification. The study highlights the value of heterogeneous field-like imagery and model stability for practical deployment in West-African settings, although training was limited to five epochs.

Zhu et al. [7] developed a deep learning model for maize leaf disease detection using a state-space attention mechanism and multi-scale feature fusion. They utilized images representing six major maize diseases and compared the proposed model with established deep learning architectures, including AlexNet, GoogLeNet, ResNet, EfficientNet and Vision Transformer. The proposed model achieved a precision of 0.95, while recall, accuracy and F1-score reached 0.94, outperforming the comparative models. The authors attributed the improved performance to the ability of the state-space attention mechanism to capture spatial and temporal relationships and the multi-scale fusion module to integrate features at different scales. However, the study concentrated on disease detection rather than nutrient-deficiency classification, and the authors acknowledged limitations associated with computational complexity, environmental variability, data diversity and multimodal information integration. Consequently, further research is required to investigate whether machine learning and deep learning techniques can effectively identify and classify nutrient deficiencies in maize leaves, particularly under diverse field conditions.

Diaz-Larios et al. [8] developed a machine-learning framework for detecting maize foliar diseases from images captured under real field conditions. Their study compared EfficientNetB0, DenseNet121 and Enhanced K-Nearest Neighbors (EKNN), incorporating preprocessing, data augmentation, segmentation and feature extraction. EKNN achieved the highest classification accuracy of 94.33%, outperforming EfficientNetB0 and DenseNet121, which achieved 89.70% and 88.04%, respectively. The findings demonstrate the potential of image-based machine learning for automated maize foliar diagnosis. Diaz-Larios et al. [8] emphasized the importance of image preprocessing and segmentation in maize leaf classification. Their approach used data augmentation, SLIC-based segmentation and enhanced feature extraction to reduce background interference and improve the representation of disease-specific regions. The resulting EKNN model achieved 94.33% accuracy, suggesting that appropriate preprocessing can significantly enhance image-based plant diagnosis, particularly under heterogeneous field conditions. Although Diaz-Larios et al. [8] demonstrated the effectiveness of machine-learning techniques for maize foliar disease detection using field-acquired images; their investigation was restricted to fewer dataset and fewer classes of diseases. This creates an opportunity for further research into automated image-based classification of maize leaf diseases detection using a dataset that has more data and classes of diseases using deep learning models.

Theerthagiri et al. [9] investigated deep learning-based classification of maize leaf diseases using SqueezeNet, VGG16, ResNet34 and ResNet50. The study employed a dataset containing maize leaf images representing common rust, grey leaf spot, northern leaf blight and healthy leaves. Following image preprocessing, augmentation and class balancing, the models were evaluated using accuracy, precision, recall and F1-score. SqueezeNet achieved the best overall performance, with an accuracy of 97%, outperforming VGG16 (92%), ResNet34 (93%) and ResNet50 (94%). The findings demonstrate the potential of lightweight deep-learning architectures for automated maize disease identification, particularly where computational efficiency and classification performance are important. Despite these contributions, several gaps remain. First, the study considered only three diseases and a healthy class, thereby leaving an opportunity to develop a more comprehensive classification framework involving multiple nutrient-deficiency categories. Finally, there is an inconsistency in the reported dataset size, with 4,188 images described in the methodology and 3,852 images reported in the conclusion. These gaps provide opportunities for further research into robust machine-learning and deep-learning models for automated maize leaf nutrient-deficiency classification.

Bachhal et al. [10] developed a hybrid deep learning and machine learning framework, termed PRFSVM, for maize leaf disease recognition. The proposed framework integrated Pyramid Scene Parsing Network (PSPNet) for semantic segmentation, ResNet50 for deep feature extraction, and Fuzzy Support Vector Machine (Fuzzy SVM) for disease classification and severity assessment.

Using maize leaf images obtained from the PlantVillage dataset the study examined common rust, southern rust, grey leaf spot, maize leaf blight, northern corn leaf blight, and healthy leaves. Following data augmentation, the dataset comprised 8,443 images, which were divided into training and testing subsets.

The proposed PRFSVM model achieved an average accuracy and a mean average precision values as 96.67% and 0.81 respectively, demonstrating the effectiveness of combining segmentation, deep feature extraction and fuzzy classification for maize disease recognition. However, the study relied on a dataset characterized by relatively controlled imaging conditions, which may limit the generalizability of the model to images acquired under diverse field conditions. The computational complexity resulting from the integration of PSPNet, ResNet50 and Fuzzy SVM also presents a challenge for deployment on resource-constrained devices. Therefore, further research is required to develop lightweight and computationally efficient models using diverse maize leaf images.

Dube and Mambudzi [11] developed a convolutional neural network-based system for detecting maize leaf diseases and integrated the model into a React Native mobile application to facilitate disease diagnosis among smallholder farmers. The study used 10,933 images representing Gray Leaf Spot, Healthy maize, Northern Leaf Blight, Common Corn Rust, and a non-maize class. A custom CNN implemented using TensorFlow/Keras achieved 92.4% training accuracy and 88.0% validation accuracy. The system was further enhanced with multilingual support in English, Shona, and Ndebele, enabling farmers to obtain disease symptoms and treatment recommendations through a mobile device. However, the study was limited by its relatively modest validation performance, restricted disease classes, and the need for more diverse images representing different lighting conditions, occlusions, and field environments. Furthermore, the study did not provide an extensive comparative evaluation of contemporary deep-learning architectures or investigate explainable AI techniques. These limitations indicate the need for further research involving diverse field datasets, comparative deep-learning models, robust performance evaluation, and explainable disease classification.

Recent advances have increasingly explored hybrid deep learning architectures for maize disease recognition. Shandilya et al. [12] , for instance, proposed a hybrid CNN-ViT model that combines the local feature extraction capabilities of convolutional networks with the global contextual modelling capabilities of Vision Transformers. The model achieved 99.15% validation accuracy and an average cross-validation accuracy of 99.06%, while obtaining 95.93% accuracy on an independent CD&S dataset. These results demonstrate the potential of combining convolutional and transformer-based architectures for maize disease classification. Nevertheless, the model's approximately 87 million trainable parameters raise concerns regarding computational cost and deployment on resource-constrained agricultural devices. Although recent studies such as Shandilya et al. [12] have demonstrated that hybrid CNN-ViT architectures can achieve high accuracy in maize leaf disease classification, comparatively less attention has been given to the automated identification and severity classification of maize nutrient deficiencies. Existing disease-oriented models may also not adequately distinguish nutritional disorders from pathological symptoms. Furthermore, the computational complexity of large hybrid architectures presents challenges for deployment in resource-constrained agricultural environments. Therefore, there is a need for research that investigates efficient machine and deep learning approaches for maize nutrient deficiency classification, particularly using larger maize images dataset and evaluation criteria that consider both predictive performance and computational efficiency.

Previous research has demonstrated the applicability of transfer learning for maize disease detection. Paul et al. [13] , for example, employed a pre-trained VGG16 CNN to classify three maize leaf diseases and achieved 93% testing accuracy. Although the study demonstrated the feasibility of deploying a deep learning model in a mobile application, its evaluation was restricted to a limited number of disease classes and a relatively small dataset. Therefore, further investigation is required to determine whether alternative deep learning architectures can provide improved classification performance across a more diverse maize disease dataset.

Recent research has increasingly focused on developing accurate and computationally efficient deep learning models for maize leaf disease diagnosis. Alpsalaz et al. [14] proposed a lightweight CNN architecture for classifying maize leaves into healthy, gray leaf spot, common rust, and northern leaf blight categories. Their model achieved 94.97% classification accuracy and a micro-average AUC of 0.99 while utilizing only 1.22 million parameters, demonstrating a favorable balance between predictive performance and computational efficiency. The study further distinguished itself by incorporating LIME and SHAP explainable artificial intelligence techniques, with SHAP achieving an IoU of 0.82 in identifying disease-relevant regions. Despite these contributions, the study relied on a relatively limited set of disease classes and acknowledged that controlled image conditions could restrict generalization to real-world environments. Furthermore, the investigation focused on disease identification rather than quantifying disease severity. Consequently, there remains a need for studies using more diverse maize leaf datasets and models capable of identifying both disease type and severity, particularly under varying environmental and geographic conditions.

Nakatumba-Nabende and Murindanyi [15] developed deep learning models for in-field maize leaf disease diagnosis using images collected with digital cameras and smartphones across Uganda, Tanzania, Ghana and Namibia. The study evaluated custom CNN, MobileNetV2, InceptionResNetV2, Vision Transformer and classical machine-learning models, while an enhanced YOLOv10 architecture was developed for disease detection. The study focused on four maize health conditions: Maize Leaf Blight, Maize Lethal Necrosis, Maize Streak Virus and Fall Armyworm damage. MobileNetV2 achieved the highest classification accuracy of 97%, whereas the enhanced YOLOv10 achieved a mAP of 0.995. The study also incorporated Grad-CAM and LIME for explainability and deployed the classification model in a mobile application for field diagnosis. However, the study was limited by its coverage of only four maize disease/health classes and the scale of its dataset, which may not adequately represent the wider diversity of maize leaf conditions encountered across different production environments. This creates a research gap for developing and evaluating models using a larger dataset with a broader range of maize leaf disease or health classes. Consequently, further research is needed to determine whether machine-learning and deep-learning models can maintain high classification performance when trained and evaluated on larger and more diverse maize-leaf datasets.

3.Methodology

Data Collection and Dataset Description

This study used the publicly available Maize Crop Disease (Leaf) dataset developed by Adege [16] and hosted on Mendeley Data. The dataset was selected because it was specifically curated for supervised maize disease classification and contains eight maize disease classes, providing a multiclass problem suitable for evaluating deep learning models. The dataset was compiled from multiple image sources with a particular focus on African maize varieties, thereby providing greater importance to maize disease recognition in African agricultural environments.

The dataset contains images representing different maize disease conditions and incorporates variations in image resolution, disease symptoms, and visual conditions. Such variability is useful for deep learning research because maize disease symptoms can vary considerably according to disease stage, environmental conditions, image acquisition conditions, and maize variety. The diversity of the images therefore provides an opportunity to evaluate the ability of deep learning models to learn discriminative visual features and generalize across heterogeneous image conditions.

The dataset was designed primarily for maize disease classification, although the dataset is also being extended to support object-detection applications for real-time disease identification. In this study, the images were used for image-level multiclass classification. Prior to model development, the images were organized according to their respective disease classes and subjected to appropriate preprocessing. The processed images were subsequently divided into training, validation, and testing subsets for model training, hyperparameter selection, and independent performance evaluation.

Deep learning models

ResNet50: ResNet50 is a deep convolutional neural network (CNN) architecture consisting of 50 layers, developed using residual learning to facilitate the training of deep networks. It employs residual or skip connections, which allow information to bypass one or more convolutional layers and facilitate gradient propagation during training [17] ,. ResNet50 has been widely used for image classification and transfer learning because of its ability to learn hierarchical and discriminative visual features. In maize leaf disease classification, a pre-trained ResNet50 can be fine-tuned using maize leaf images to learn disease-specific visual patterns, including lesions, discoloration, spots, and texture characteristics, thereby enabling the classification of different disease categories. Recent maize disease research has demonstrated the applicability of ResNet50 to maize leaf disease identification, including classification using field-collected maize leaf images [17].

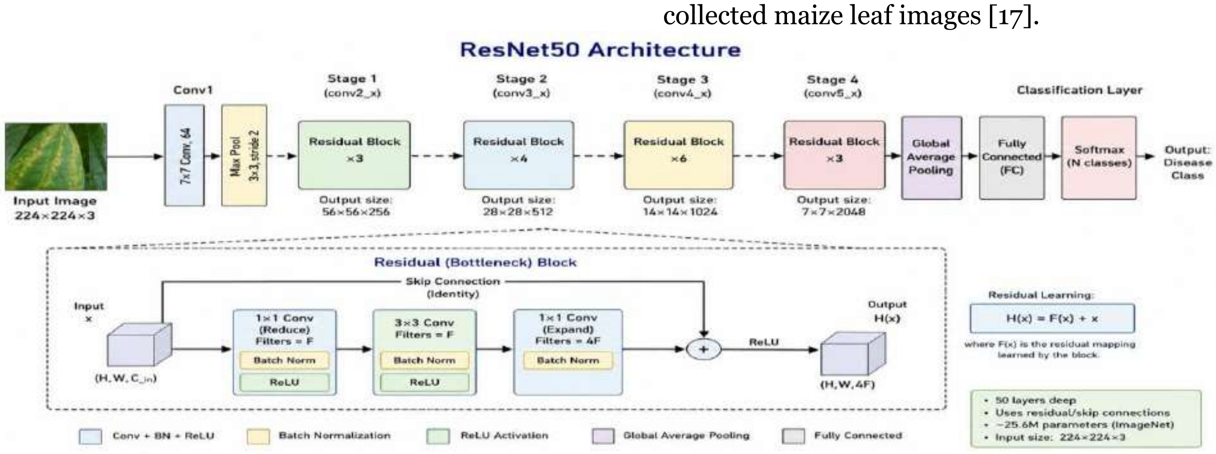


Figure 1: ResNet50 Architecture[17]

EfficientNet-B0 is a lightweight convolutional neural network (CNN) and the baseline model of the EfficientNet family introduced by Tan and Le [18]. The architecture employs mobile inverted bottleneck convolution (MBConv) blocks, depth wise separable convolutions, squeeze-and-excitation modules, and Swish activation to achieve efficient feature extraction. Its major innovation is compound scaling, which jointly balances network depth, width, and input-image resolution rather than scaling a single dimension. EfficientNet-B0 therefore provides a favorable trade-off between classification performance and computational cost, making it suitable for image classification and transfer-learning applications, including maize leaf disease recognition [18].

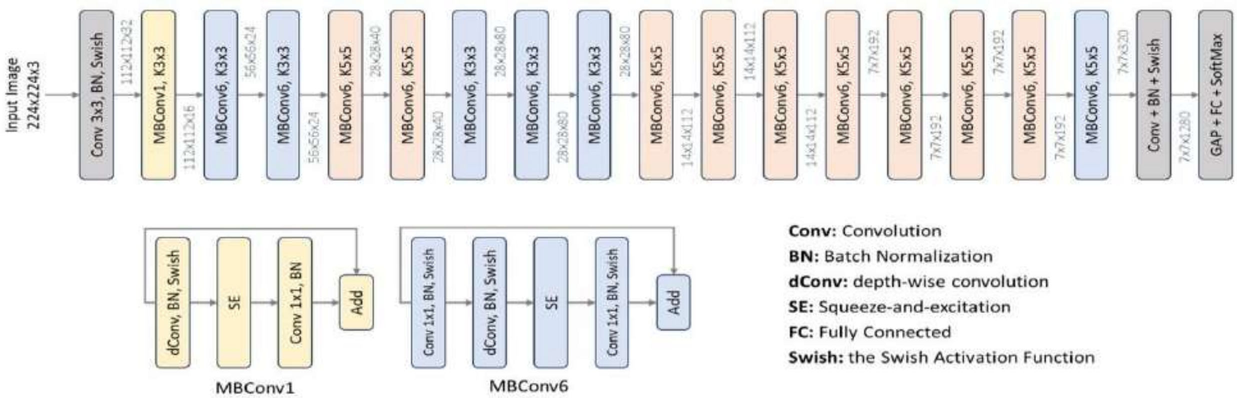


Figure 2: Architecture of EfficientNet-B0.

The architecture consists of an initial convolutional layer followed by a sequence of MBConv blocks, with feature-map resolution decreasing from 224 × 224 toward 7 × 7 while the channel depth increases. The final layers perform convolution, global pooling, and classification.

MobileNetV3

MobileNetV3 is a lightweight convolutional neural network (CNN) architecture proposed by Howard et al. [19] for efficient computer vision applications, particularly on mobile and resource-constrained devices. It was developed using hardware-aware neural architecture search (NAS) together with the NetAdapt algorithm, followed by architectural improvements designed to enhance accuracy and computational efficiency.

The architecture incorporates inverted residual bottleneck blocks, depth wise separable convolutions, squeeze-and-excitation (SE) attention modules, and the h-swish activation function. These components enable MobileNetV3 to learn effective visual representations while reducing computational complexity compared with conventional CNN architectures. MobileNetV3 was introduced in two main variants: MobileNetV3-Large, intended for applications with relatively greater computational resources, and MobileNetV3-Small, designed for more resource-constrained applications.

For maize leaf disease classification, MobileNetV3 can be employed as a transfer-learning model to extract discriminative features from leaf images and classify them into different disease categories. Its relatively low computational requirements make it suitable for potential deployment in mobile or edge-based agricultural disease-diagnosis systems. Howard et al.[19] reported that MobileNetV3-Large improved ImageNet classification accuracy while reducing latency compared with MobileNetV2, demonstrating its effectiveness as an efficient deep-learning architecture.

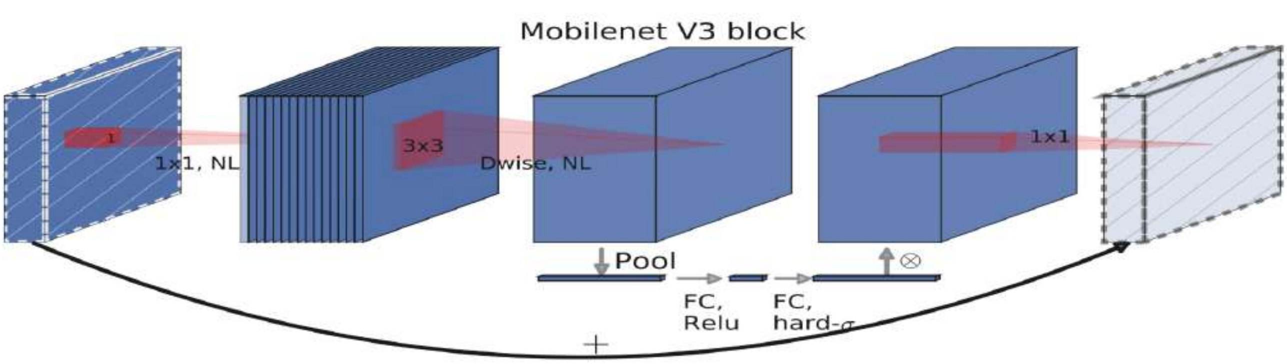


Figure 3 : MobileNetV3 [20]

Figure 3 illustrates the simplified feature-extraction and classification flow of MobileNetV3. The actual architecture varies between the MobileNetV3-Large and MobileNetV3-Small configurations.

Performance Evaluation Metrics

Evaluation metrics are quantitative measures used to assess the performance and effectiveness of a statistical or machine learning model. These metrics provide insights into how well the model is performing and help in comparing different models or algorithms.

When evaluating a machine learning model, it is crucial to assess its predictive ability, generalization capability, and overall quality. Evaluation metrics provide objective criteria to measure these aspects. The choice of evaluation metrics depends on the specific problem domain, the type of data, and the desired outcome.

Confusion Matrix: - Confusion matrix is a simple table having two dimensions: ‘actual’ and ‘predicted’, and for both the dimensions it has same sets of classes. In a simple classification problem, there are two classes: positive (1) and negative (0) [21].

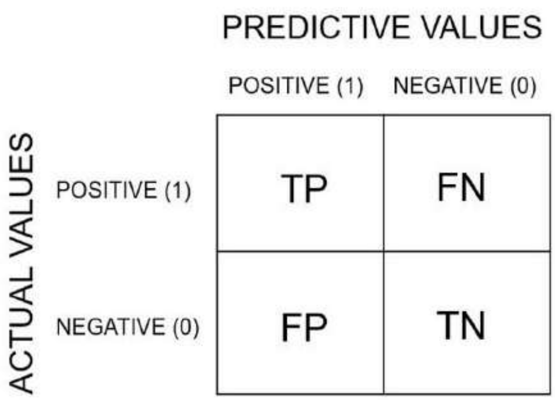


*Figure 4: **Confusion Matrix** [21]*

There are four quadrants in the confusion matrix and they represent the following:

True Positive (TP): The number of cases that were positive (+) and correctly classified as positive (+)

False Positive (FP): The number of occurrences that were negative (-) and incorrectly classified as (+). This also known as Type 1 Error.

False Negative (FN): The number of occurrences that were positive (+) and incorrectly classified as negative (-). It is also known as Type 2 Error.

True Negative (TN): The number of cases that were negative (-) and correctly classified as (-).

For a perfect classifier there should be no error that is, Type 1 or Type 2 errors should be equal to zero, therefore, False Positives = False Negatives = 0. So, to improve the model’s performance we should minimize these errors, False Positives and False negatives. Confusion matrix forms the base for other performance metrics [21].

Accuracy:

Accuracy is a commonly used metric for evaluating the performance of a classification model. It indicates the proportion of observations that are correctly classified among all observations in the dataset. Accuracy can be determined from the confusion matrix using

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (1)$$

Precision:

Precision, also referred to as the **Positive Predictive Value (PPV)**, measures the proportion of instances predicted as positive that are actually positive. It is particularly useful when evaluating classification performance on datasets with class imbalance. Precision is calculated using Equation (2):

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

Recall:

Recall, also known as the **True Positive Rate (TPR)** or **Sensitivity**, measures the ability of a classification model to correctly identify positive instances. It represents the proportion of actual positive cases that are correctly predicted by the model. Recall is expressed as:

$$sensitivity = \frac{TP}{TP + FN} \quad (3)$$

Error Rate (ERR):

The error rate represents the proportion of incorrectly classified instances in the dataset. It is obtained by dividing the total number of false positive (FP) and false negative (FN) predictions by the total number of observations, as given in Equation (4):

$$ERR = \frac{FP + FN}{TN + FP + FN + TP} \quad (4)$$

False Negative Rate (FNR):

The False Negative Rate indicates the proportion of actual positive instances that are incorrectly classified as negative. It is calculated as the number of false negatives divided by the total number of actual positive cases:

$$FNR = \frac{FN}{FN + TP} \quad (5)$$

Where **TP** represents True Positive, **TN** represents True Negative, **FP** represents False Positive, and **FN** represents False Negative.

A classification report serves as a machine learning performance metric, displaying the precision, recall, F1 score, and support of a trained classification model [22].

	precision	recall	f1-score	support
ham	0.99	0.99	0.99	1587
spam	0.93	0.92	0.92	252
accuracy			0.98	1839
macro avg	0.96	0.95	0.96	1839
weighted avg	0.98	0.98	0.98	1839

Figure 5: Classification Report in Machine Learning [22]

Figure 5 shows an example of a classification report of a machine learning model, in the figure, the precision, recall, F1-score, accuracy, Macro avg, weighted avg and support are clearly stated.

4.RESULT AND DISCUSSION

A comparative analysis of the selected deep learning models indicates that EfficientNet-B0 demonstrated superior reliability in predicting maize leaf diseases. This section summarizes and compares the performance obtained from the three evaluated models: ResNet50, EfficientNet-B0, and MobileNetV3.

For ResNet50, Figure 6 presents the precision, recall, F1-score, and accuracy results. The corresponding confusion matrix is shown in Figure 7, while Figure 8 illustrates the training and validation accuracy and loss curves.

Similarly, the performance metrics of EfficientNet-B0, including precision, recall, F1-score, and accuracy, are presented in Figure 9. Figure 10 provides its confusion matrix, and Figure 11 shows the training and validation accuracy and loss curves.

For MobileNetV3, Figure 12 presents the precision, recall, F1-score, and accuracy results. Its confusion matrix is depicted in Figure 13, whereas Figure 14 illustrates the training and validation accuracy and loss curves.

Overall, the comparative results indicate that EfficientNet-B0 achieved more consistent and reliable performance for maize leaf disease classification than the other evaluated models.

ResNet50 Result

Classification Report of ResNet50 Model on Testing Set:

	precision	recall	f1-score	support
Abiotic_disease-D	0.88	0.78	0.82	9
Aphids-P	1.00	1.00	1.00	2
Curvularia-D	1.00	0.90	0.95	30
Healthy_leaf	1.00	0.94	0.97	68
Helminthosporiosis-D	0.86	0.83	0.84	23
Rust-D	1.00	1.00	1.00	4
Spodoptera_frugiperda-P	1.00	1.00	1.00	55
Stripe-D	0.91	1.00	0.95	217
Virosis-D	1.00	0.25	0.40	16
accuracy			0.94	424
macro avg	0.96	0.86	0.88	424
weighted avg	0.94	0.94	0.93	424

Figure 6: ResNet50 Classification Report

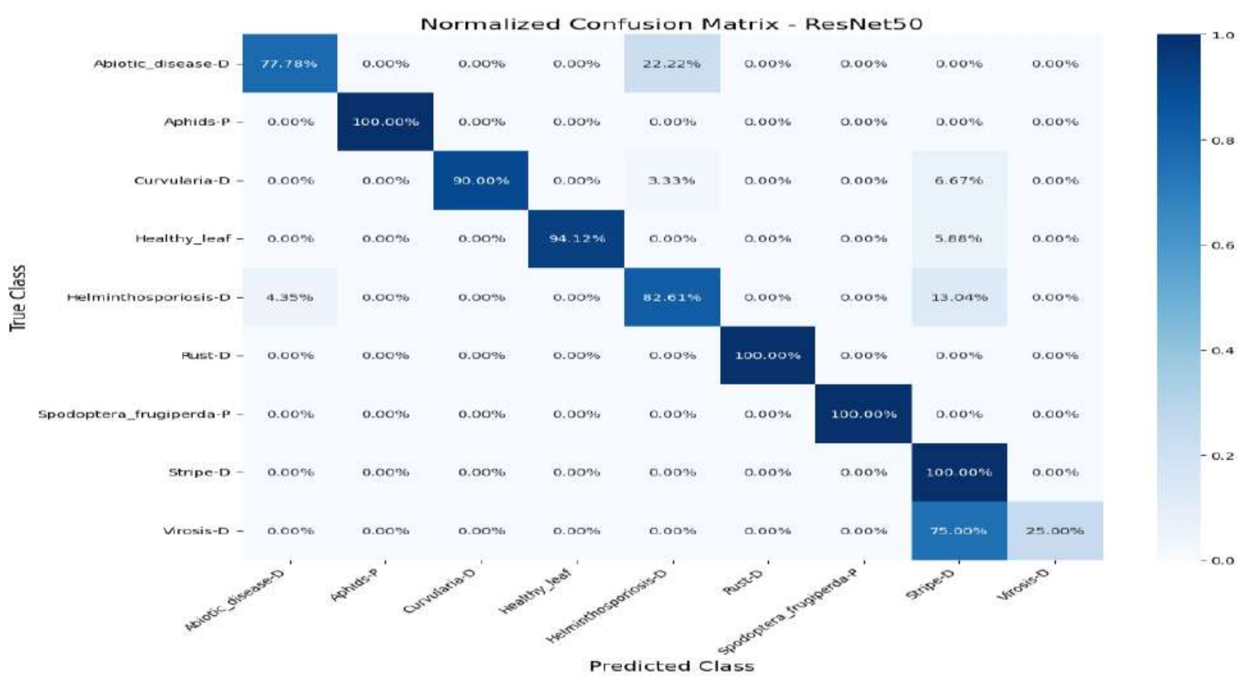


Figure 7: ResNet50 Confusion Matrix

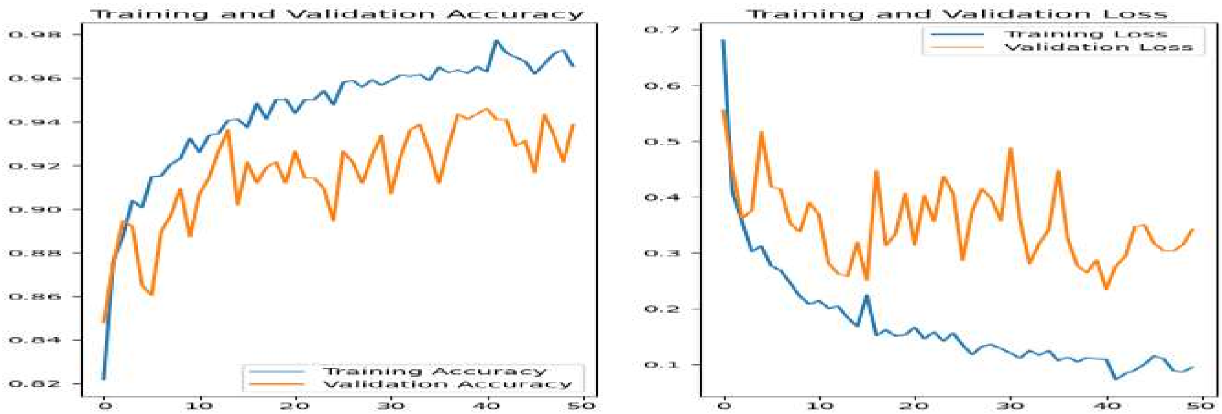


Figure 8: ResNet50 Training and Validation Accuracy/Training and Validation Loss

During the experiment as shown in figure 6, the data used for the training was 70% while the test dataset from testing was 30% which amounted to 94%. ResNet50 achieved an overall accuracy of 94%, with weighted precision, recall, and F1-score of 0.94, 0.94, and 0.93, respectively. The lower recall for Virosis-D reduced the overall performance. The ResNet50 confusion matrix shown in figure 7 shows strong classification for Aphids-P, Rust-D, and *Spodoptera frugiperda*-P, with 100% correct classification. However, Virosis-D was poorly recognized, with only 25% correctly classified and 75% misclassified as Stripe-D. As shown in figure 8, the training accuracy increased to approximately 97%, while validation accuracy remained around 92–94%.

Classification Report of EfficientNet-B0 Model on Testing Set:

	precision	recall	f1-score	support
Abiotic_disease-D	1.00	1.00	1.00	9
Aphids-P	1.00	1.00	1.00	2
Curvularia-D	0.97	0.93	0.95	30
Healthy_leaf	0.97	0.97	0.97	68
Helminthosporiosis-D	1.00	0.96	0.98	23
Rust-D	0.57	1.00	0.73	4
Spodoptera_frugiperda-P	1.00	1.00	1.00	55
Stripe-D	0.98	0.99	0.99	217
Virosis-D	0.92	0.75	0.83	16
accuracy			0.97	424
macro avg	0.93	0.96	0.94	424
weighted avg	0.98	0.97	0.97	424

EfficientNet-B0

Figure 9: EfficientNet-B0 Classification Report

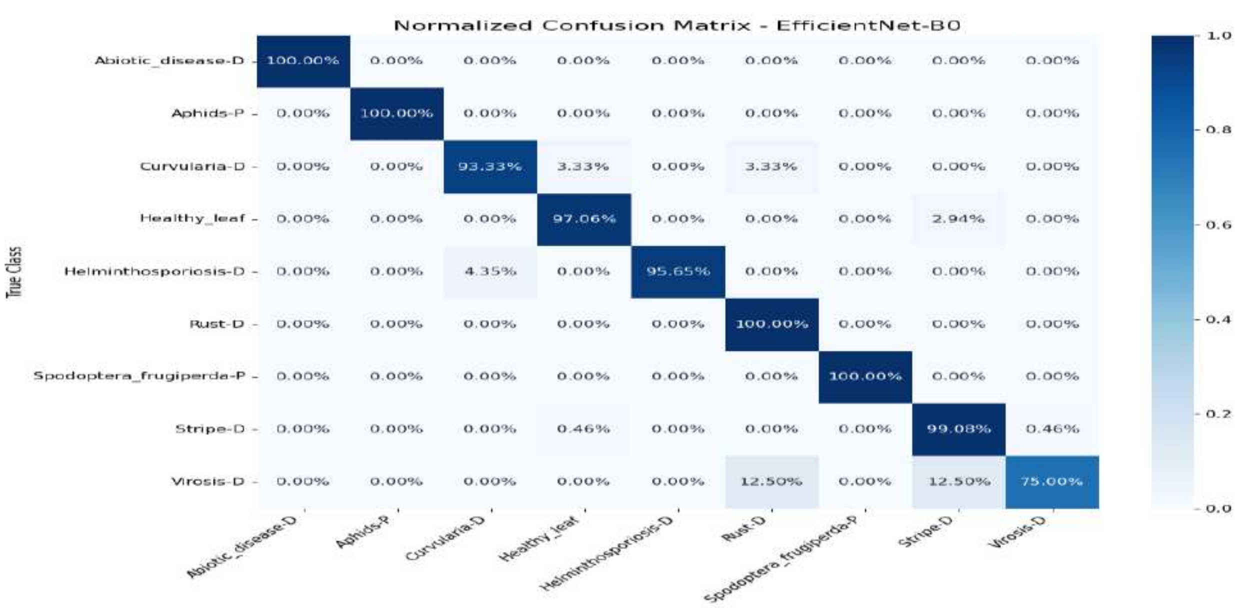


Figure 10: EfficientNet-B0 Confusion Matrix

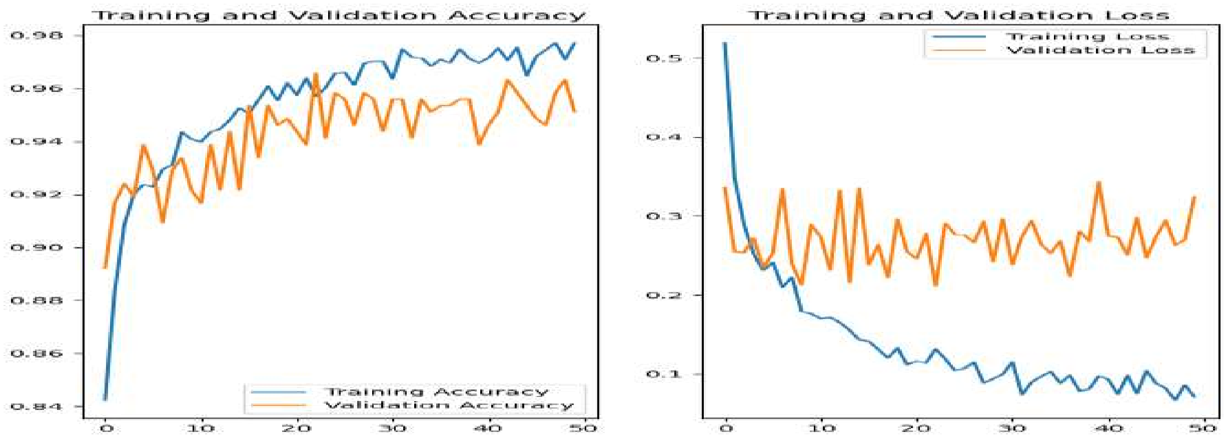


Figure 11: EfficientNet-B0 Training and Validation Accuracy/Training and Validation Loss

During the experiment as shown in figure 9, the data used for the training was 70% while the test dataset from testing was 30% which amounted to 97%. EfficientNet-B0 achieved the highest accuracy of 97%, with weighted precision, recall, and F1-score of 0.98, 0.97, and 0.97. This demonstrates its superior and relatively balanced classification performance. EfficientNet-B0 produced the strongest confusion matrix, with high correct classification across most classes. Stripe-D achieved about 99% correct classification, while Virosis-D achieved 75% as shown in figure 10. As shown in figure 11, the training accuracy increased steadily to approximately 98%, while validation accuracy remained around 95–96%. The relatively small gap indicates good generalization compared with ResNet50.

MobileNetV3 Results

Classification Report - MobileNetV3

	precision	recall	f1-score	support
Abiotic_disease-D	1.0000	1.0000	1.0000	11
Aphids-P	0.0000	0.0000	0.0000	0
Curvularia-D	0.9756	0.9524	0.9639	42
Healthy_leaf	0.9298	0.9464	0.9381	56
Helminthosporiosis-D	1.0000	0.8333	0.9091	18
Rust-D	0.9000	0.8182	0.8571	11
Spodoptera_frugiperda-P	0.9857	0.9718	0.9787	71
Stripe-D	0.9458	0.9796	0.9624	196
Virosis-D	0.8235	0.7368	0.7778	19
accuracy			0.9505	424
macro avg	0.8401	0.8043	0.8208	424
weighted avg	0.9504	0.9505	0.9498	424

Figure 12: MobileNetV3 Classification Report

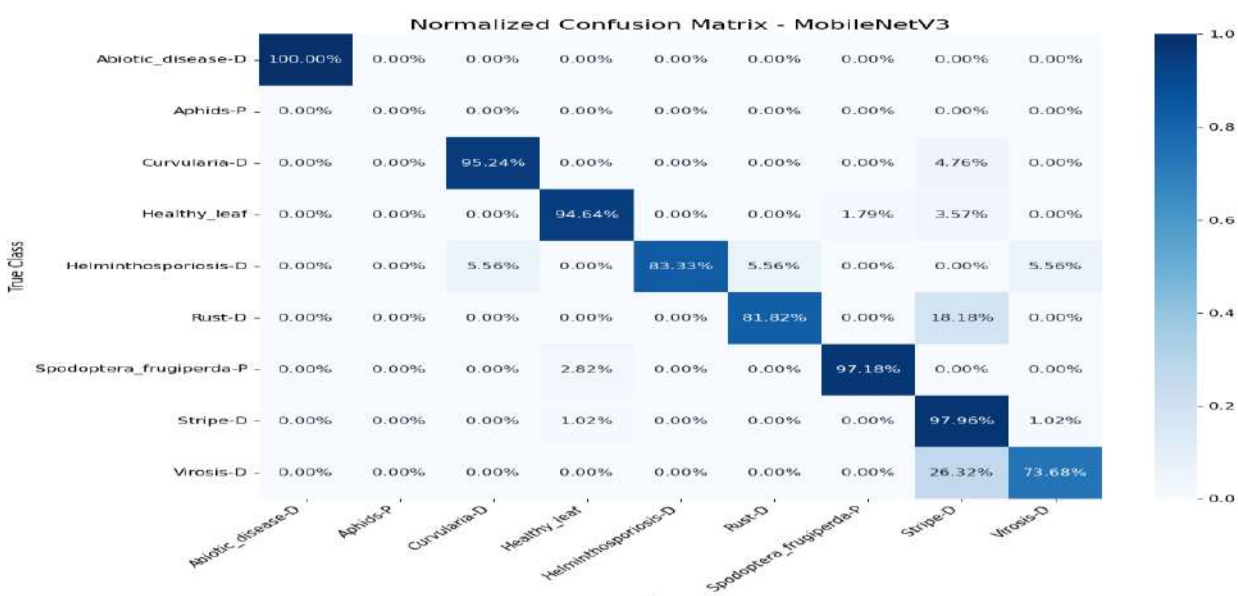


Figure 13: MobileNetV3 Confusion Matrix

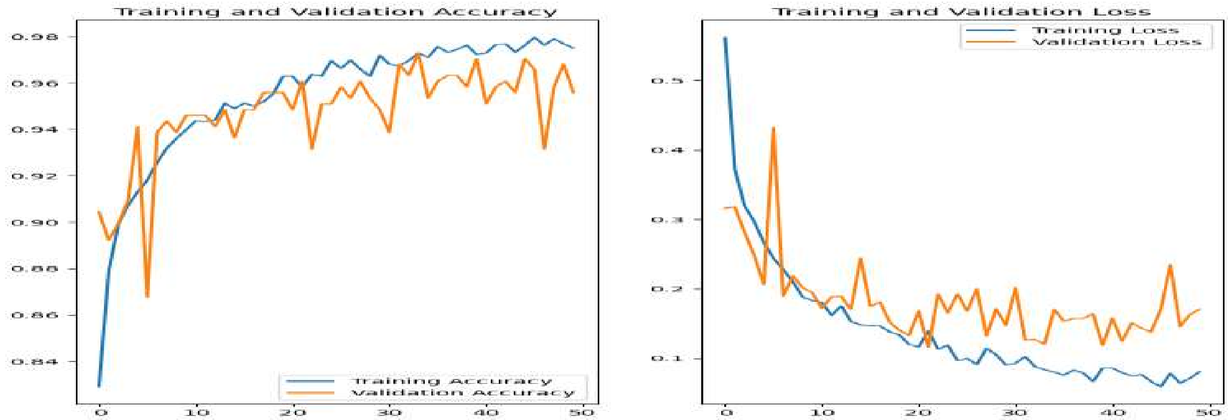


Figure 14: MobileNetV3 Training and Validation Accuracy/Training and Validation Loss

During the experiment as shown in figure 12, the data used for the training was 70% while the test dataset from testing was 30% which amounted to 95.05%. MobileNetV3 achieved 95.05% accuracy, with weighted precision of 0.9504, recall of 0.9505, and F1-score of 0.9498. This indicates strong overall performance, although some individual classes were less accurately recognized. MobileNetV3 performed well on most classes, particularly Abiotic_disease-D and Spodoptera frugiperda-P. However, Virosis-D showed some confusion with Stripe-D, with approximately 26% incorrectly classified as Stripe-D as shown in 14. As shown in figure 11, Training accuracy approached 98%, while validation accuracy remained around 95–96%. The curves indicate good learning and generalization, although validation loss fluctuated during training.

Among the three models, EfficientNet-B0 performed best (97%), followed by MobileNetV3 (95.05%) and ResNet50 (94%). EfficientNet-B0 also showed the most consistent class-level performance and fewer major misclassifications, making it the most suitable model for the maize leaf disease classification task.

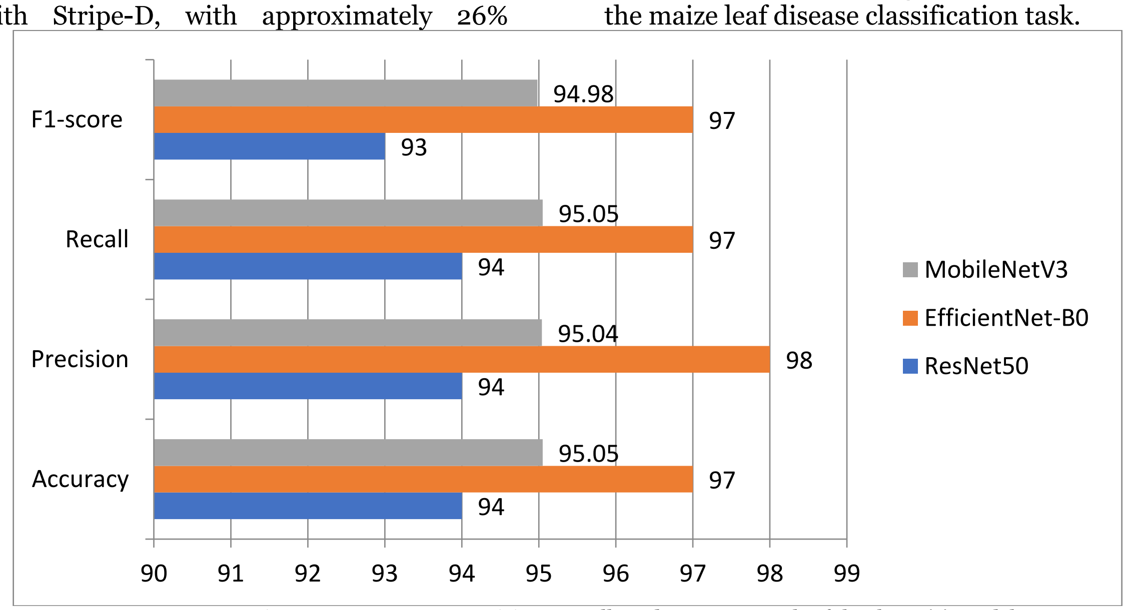


Figure 15: Accuracy, Precision, Recall, and F1-score result of the three (3) models

Figure 15 depicts the summary of the Accuracy, Precision, Recall, and F1-score result of the three (3) models used in the prediction of the maize leaf disease. The result shows that EfficientNet-B0 have the highest accuracy of 97%, followed by MobileNetV3 95.05% and ResNet50 94% respectively.

5.CONCLUSION

This study demonstrated the effectiveness of deep learning for automated maize leaf disease classification. Among the three evaluated models, EfficientNet-B0 achieved the best performance with 97% accuracy, followed by MobileNetV3 with 95.05% and ResNet50 with 94%. EfficientNet-B0 also produced the most balanced precision, recall, and F1-score and showed fewer major classification errors.

Therefore, EfficientNet-B0 can be considered the most suitable model among those evaluated for this maize disease classification task. Future studies should investigate larger and more diverse field-based datasets and consider model optimization for deployment in practical agricultural and mobile disease-diagnosis applications.

REFERENCES

[1]. G, J., Dhole, A., & D, D. (2026). Artificial intelligence–driven detection and decision support system for precision management of maize downy mildew. *PLoS One*, 21(3), e0343517. https://doi.org/10.1371/journal.pone.0343517.

[2]. Li, B., Yu, L., Zhu, H., & Tan, Z. (2025). YOLO-FDLU: A lightweight improved YOLO11s-based algorithm for accurate maize pest and disease detection. *AgriEngineering*, 7(10), 323. https://doi.org/10.3390/agriengineering7100323.

[3]. Yuan, L., & Wang, F. (2026). MSTA-YOLOv11: An improved detection model for maize pests and diseases under complex scenes. *Journal of Real-Time Image Processing*, 23, 22. https://doi.org/10.1007/s11554-025-01821-9.

[4]. Gao, X., He, L., Liu, Y., Wu, J., Cao, Y., Dong, S., & Jia, Y. (2025). DSTANet: A lightweight and high-precision network for fine-grained and early identification of maize leaf diseases in field environments. *Sensors*, 25(16), 4954. https://doi.org/10.3390/s25164954.

[5]. Murugan, T., Badusha, N. A. N. M., Musa, N. S., Alahbabi, E. M. M., Ahmed Alyammahi, R. A., Adege, A. B., Abdi, A., & Megersa, Z. M. (2026). Research advances in maize crop disease detection using machine learning and deep learning approaches. *Computers*, 15(2), 99. https://doi.org/10.3390/computers15020099.

[6]. Mapamela, W., Radiokana, B., Moore, S. E., Modupe, A., & Marivate, V. (2025). *Maize disease detection using deep learning models*. *Research Square***.** https://doi.org/10.21203/rs.3.rs-5763601/v1.

[7]. Zhu, T., Yan, F., Lv, X., Zhao, H., Wang, Z., Dong, K., Fu, Z., Jia, R., & Lv, C. (2024). A deep learning model for accurate maize disease detection based on state-space attention and feature fusion. *Plants, 13*(22), 3151. https://doi.org/10.3390/plants13223151.

[8]. Diaz-Larios, J. M., Luna-Flores, P. A., Bances-Saavedra, D. E., & Arcila-Diaz, J. (2026). Detection of diseases in maize plants by analyzing foliar images using machine learning techniques. *AgriEngineering, 8*(8), 343. https://doi.org/10.3390/agriengineering8080343.

[9]. Theerthagiri, P., Ruby, A. U., Chandran, J. G. C., Sardar, T. H., & Shafeeq, A. B. M. (2024). Deep SqueezeNet learning model for diagnosis and prediction of maize leaf diseases. *Journal of Big Data, 11*, 112. https://doi.org/10.1186/s40537-024-00972-z.

[10]. Bachhal, P., Kukreja, V., Ahuja, S., Lilhore, U. K., Simaiya, S., Bijalwan, A., Alroobaea, R., & Algarni, S. (2024). Maize leaf disease recognition using PRF-SVM integration: A breakthrough technique. *Scientific Reports, 14*, 10219. https://doi.org/10.1038/s41598-024-60506-8

[11]. Dube, S. S., & Mambudzi, H. (2025). Maize plant disease detection using convolutional neural network. *International Journal of Research and Innovation in Applied Science, 10*(7), 157–166. https://doi.org/10.51584/IJRIAS.2025.100700013.

[12]. Shandilya, G., Gupta, S., Mohamed, H. G., Bharany, S., Rehman, A. U., & Hussen, S. (2025). Enhanced maize leaf disease detection and classification using an integrated CNN-ViT model. *Food Science & Nutrition, 13*(7), e70513. https://doi.org/10.1002/fsn3.70513

[13]. Paul, H., Udayangani, H., Umesha, K., Lankasena, N., Liyanage, C., & Thambugala, K. (2025). Maize leaf disease detection using convolutional neural network: A mobile application based on pre-trained VGG16 architecture. *New Zealand Journal of Crop and Horticultural Science, 53*(2), 367–383. https://doi.org/10.1080/01140671.2024.2385813.

[14]. Alpsalaz, F., Özüpak, Y., Aslan, E., & Uzel, H. (2025). Classification of maize leaf diseases with deep learning: Performance evaluation of the proposed model and use of explicable artificial intelligence. *Chemometrics and Intelligent Laboratory Systems, 262*, 105412. https://doi.org/10.1016/j.chemolab.2025.105412.

[15]. Nakatumba-Nabende, J., & Murindanyi, S. (2025). Deep learning models for enhanced in-field maize leaf disease diagnosis. *Machine Learning with Applications, 20*, 100673. https://doi.org/10.1016/j.mlwa.2025.100673.

[16]. Adege, A. (2026). *Maize Crop Disease (Leaf)* (Version 1) [Data set]. Mendeley Data. https://doi.org/10.17632/6w6gsvghfw.1

[17]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

[18]. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 6105–6114.

[19]. Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V., & Adam, H. (2019). Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 1314–1324. https://doi.org/10.1109/ICCV.2019.00140

[20]. https://pytorch.org/blog/torchvision-mobilenet-v3-implementation/

[21]. Aryan, L. (2021). Performance Evaluation Metrics. https://medium.com/analytics-vidhya/performance-evaluation-metrics-5c9104b2a407.

[22].Aman, K. (2021). Classification Report in Machine Learning. https://thecleverprogrammer.com/2021/07/07/classific.ation-report-in-machine-learning/#google_vignette.