International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(19s) 2026 pp. 1207-1220
https://doi.org/10.70917/ijcisim-2026-5119
Article

An Attention-Enhanced Lightweight Deep Learning Framework for Banana Leaf Disease Detection Using CBAM-MobileNetV2 with Grad-CAM Visualization

Ashish Patil1*, Deepak Kumar Yadav[2]

[1]Research Scholar, Department of Computer Science & Engg., SAGE University, Indore, India
[2]Associate Professor, Department of Computer Science & Engg., SAGE University, Indore, India
*ashish.patil1389@gmail.com[1], deepak_ku_yadav@outlook.com[2]

Abstract: Fungal diseases like Sigatoka and nutrient deficiencies like potassium deficiency greatly affect farmer income and yield of bananas. The traditional visual inspection methods are subjective, labor-intensive and subject to errors. An improved deep learning model is proposed to identify banana leaves diseases efficiently by introducing a Convolutional Block Attention Module (CBAM) and MobileNetV2 architecture. The proposed method is based on a small labeled dataset consisting of 2,801 images across three classes (healthy, Sigatoka infected and potassium deficiency) and pre-processing techniques and incorporates visual explanations in the form of Gradient-weighted Class Activation Mapping (Grad-CAM). The training set is small but an independent test set of 1307 images from different agricultural farms has been used only for the final evaluation of the primary training set. The independent set was not used for training or validation providing an independent assessment of the model's ability to generalize to unseen real-world data. Overall, the developed system attained an accuracy of 94.4%, recall value of 97% for Sigatoka and 89% for potassium deficiency. The model's light weight stress and attention guided feature selection allows for accurate real-time diagnosis in low resource smallholder farms for timely interventions and better crop management.

Keywords: Deep Learning (DL); Convolutional Block Attention Module (CBAM); MobileNetV2; Gradient-weighted Class Activation Mapping (Grad CAM)

1. Introduction

Food has been essential for human survival. Agriculture supports both the economy and the nation. It produces basic food items like grains, vegetables, fruits and animal products all of which are important for ensuring that people have enough to eat. It is well known that food and nutritional security is a food system disrupter and has a negative impact on the socioeconomic situation of populations all over the world [1].

According to food and agriculture organization (FAO) India produces banana up to 26% of the global production. Indian state conditions like Maharashtra, Tamilnadu, Gujarat, Andhra Pradesh and Karnataka are found to be conducive for banana cultivation. Its high yield and short growing cycle as well as its ability to be harvested throughout the year are profitable and sustainable. Banana production is beset with several biotic and abiotic stresses which adversely affect both fruit quality and yield. Fungal diseases are especially devastating. Black sigatoka is the fungal disease and it reduces production up to 50%. In addition to impacting on yield this disease outbreaks have economic consequences of boosting production costs and curtailing farmers access to export markets which in turn impacts the income of millions of farmers who rely on bananas as a cash crop. The traditional approach to disease diagnosis is mainly through farmers visual assessment of bananas and there are many important limitations to this approach. These include the inability to identify diseases in real time with accuracy, low prediction accuracy, limited coverage of field surveys and important lag times in disease identification. The diagnosis of visual symptoms is subjective without consistency from farm to farm and even within the same farm and even among experienced farmers which can result in the misapplication of fungicides or fertilizers. These errors use up precious resources and increase environmental contamination this result into encourage pathogen resistance. Many banana growing areas are hampered by low agricultural extension service which leads to farmers receiving no advice when disease symptoms are first noticed.

This study has objectives: (1) To develop a lightweight deep learning model called MobileNetV2-CBAM for efficient banana disease identification (2) To evaluate its performance using accuracy, precision, recall and F1 score on a real-world dataset. (3) Use the enhanced Grad-CAM technique for interpreting disease localization. The remainder of this paper is structured as follows. In the section 2 literature survey of the research is discussed. Section 3 presents the methodology, covering the dataset description, the proposed CBAM-MobileNetV2 architecture and the evaluation metrics employed. Section 4 provides the experimental results along with a discussion of the findings which covers per-class performance, confusion matrix and Grad-CAM visualizations.

2. Literature Survey

Recent global challenges have spurred the development of artificial intelligence technologies with diverse practical applications. In the domain of agricultural monitoring Liao et al. explored the effectiveness of several deep learning models including Convolutional Neural Networks (CNN), YOLO and Vision Transformers for assessing banana crop health. Their work evaluated these architectures across multiple deployment settings such as mobile devices, unmanned aerial vehicles (UAVs) and Internet of Things (IoT) platforms. While their findings highlighted persistent difficulties related to data quality and the financial burden of implementation the authors also put forward adaptable strategies for improving pest and disease surveillance [1]. In a broader contribution Riyanto et al. carried out an extensive systematic review of 111 peer-reviewed studies published from 2019 to 2023 synthesizing current knowledge on the use of machine learning (ML), deep learning (DL) and transfer learning approaches for diagnosing plant diseases. They categorised models, datasets and evaluation metrics and showed that mobile applications are a potential way to boost agricultural productivity [2]. The Plant disease detection using ML and DL technique was introduced by P. Sajitha in which he explained the types of data sources, the algorithm types and the few successful case studies and current challenges in the field of automation [4]. Andrew and colleagues trained several CNN architectures using the PlantVillage dataset which contains 54,305 images spanning 38 disease categories. Among the models tested DenseNet-121 achieved the highest performance with an accuracy of 99.81%. This result highlights the strong potential of deep learning techniques for precise plant disease identification. [5]. Seven thousand two hundred and twelve leaf images were used to identify eight nutrient deficiencies using a deep learning method called IDWC-MSN proposed by C. K. Sunil et al. Their model successfully solved the confusion among similar symptoms like manganese and zinc deficiencies and achieved 91.34% accuracy which is higher than ResNet-50, VGG-16, MobileNet V3, EfficientNet-B0 and SqueezeNet [7]. T. Domingues et al. surveyed ML techniques used to detect, identify and predict plant diseases and pests in agriculture, focusing on tomato. They promoted smart farming methods that reduce pesticide consumption without compromising crop quality and productivity to ensure food security in the future for the world [8]. A hybrid CNN with 96% accuracy in detecting and classifying banana diseases for early fertilizer application guidance that reduces economic losses and increase the productivity of the crop was presented by K. Lakshmi Narayanan et al. [9]. Neema Mduma et al. collected 16,092 images of banana leaves and stems with Black Sigatoka and Fusarium Wilt from Tanzania establishing the largest publicly available dataset for such ML tasks as detection, classification and segmentation to boost food security in Africa [10]. A CNN based mobile application was developed for early detection of Black Sigatoka and Fusarium Wilt by Christian A. Elinisa and Neema Mduma with a total of 27,360 images of banana and non-banana plants and achieved 91.17% accuracy. They showed how deep learning can be used for fast and accurate disease classification to provide mitigation advices and multilingual support [11]. Deep Convolutional Neural Network (14-DCNN) was proposed by J. A. Pandian et al., which detected 14 plant diseases with 99.96% accuracy for 147,500 augmented images which outperforms the transfer learning models [12]. Sangeetha and colleagues developed an agro-deep-learning model designed to predict the severity of

Panama wilt disease in banana plants. Their approach relies on analyzing changes in leaf color and shape and the model achieved a prediction accuracy of 91.56%. [13]. To enable large scale crop monitoring Selvaraj et al. (2020) designed a surveillance system for banana diseases that integrates M with imagery captured by UAV and satellites. Using Random Forest and RetinaNet models their system achieved accuracy rates exceeding 90%. [15]. Sanga et al. created a DL model to detect banana diseases early in Tanzania for smallholder farmers using the VGG16, ResNet18, ResNet50, ResNet152 and InceptionV3 architectures. The best accuracy was obtained by the algorithm InceptionV3 of 95.41% and it was decided to deploy it on the mobile device because of its low memory consumption and in-time detection ability [16]. K. Patel and A. Patel have explored image processing, ML and DL methods for detecting and classifying plant diseases and concluded that CNN based techniques were giving better classification accuracy than traditional techniques. They also mentioned the limitation of data, overlapping of disease symptoms with other diseases and requirement of compact real time model suitable for mobile and embedded agriculture applications as challenges [17]. Mohanty and colleagues introduced a smartphone-based deep learning system for identifying crop diseases. Their deep CNN model was tested on a dataset of 54,306 images covering 14 crop species and 26 disease types achieving an accuracy of 99.35%. This finding underscores the promise of AI-driven tools for large-scale plant disease detection and their potential contribution to global food security [18]. CNN models are shown to be feasible to develop early warning and advisory systems for plant diseases using 87,848 images of 25 plant species and 58 diseased classes, with accuracy rate of 99.53% [19]. Transfer learning of deep CNN InceptionV3 for cassava disease detection in field images was applied in Tanzania with accuracy of 95% to 98% for multiple cassava disease and pest damages which proved to be cost-effective and scalable mobile disease diagnosis [20]. A. Picon et al. presented a DRNN that was used to detect multiple wheat fungal diseases in real-field environments on 8178 images from Spain and Germany with a balanced accuracy score of 87% [21]. Keceli and colleagues developed a Multi Input Multi Task Neural Network designed to simultaneously predict both plant species and disease type. Their results indicated that this multi-task approach was more efficient and converged faster than traditional single-task models. [22]. A. M. S. Kheir et al. showed that the MobileViTv2 model compared to EfficientNetB7 and hybrid DL showed 94% accuracy and was successfully deployed in a web-based real-time diagnosis application for classification of plant leaf diseases [23]. Ashurov and colleagues designed a Depth wise Convolutional Neural Network that incorporates squeeze-and-excitation blocks and residual skip connections for automatic plant disease detection. Their model achieved 98% accuracy and an F1 score of 98.2% while also demonstrating superior feature extraction capabilities and computational efficiency compared to other approaches. [25].

Although great strides have been made there are still areas of research that need to be addressed. This is because many of the models in use are very accurate in controlled conditions or with small data sets but not on data collected on the field where the lighting conditions, background clutter and environment vary. Mobile and web-based disease detection systems have been suggested but a number of these systems require high computational resources or are not lightweight enough for deployment in the rural agricultural environment [11, 16, 18, 23]. Moreover, the majority of frameworks are limited to a few banana diseases for instance Black Sigatoka, Fusarium Wilt or Panama Wilt [11, 13] and there is a lack of comprehensive framework for multiple diseases under a range of conditions. Another important drawback is the absence of extensive, varied and balanced datasets to support diagnosis of banana diseases [10, 17]. Furthermore, there is limited use of advanced data augmentation, IoT sensors, UAV data, and predictive disease severity models for banana disease surveillance systems [1, 15]. To address these challenges this study proposes an improved DL model for banana disease detection that integrate the CBAM with the MobileNetV2 architecture. The CBAM component applies both channel-wise and spatial attention enabling the model to focus on disease-related visual features while filtering out irrelevant background information. Grad-CAM visualization makes heat maps and Bounding boxes understandable so that farmers can know why the model predicts certain things. The proposed system is specifically designed in the context of resource poor environments characteristic of smallholder farming systems.

3. Methods and Materials

3.1 Banana leaf disease dataset

In this study an image database was also developed comprising both symptomatic and symptom-free banana leaves. Healthy leaf samples were collected from banana plantations located in the Jalgaon district of Maharashtra, India. A digital camera was used to capture all images under natural lighting conditions to ensure real-world relevance. Prior to model training the images were uniformly resized to 224 × 224 pixels for consistency. There are three classes in the dataset: leaves that are healthy, leaves with sigatoka disease and leaves with potassium deficiency symptoms. In this collection there were 1,075 images in the Sigatoka class, 896 images in the potassium deficiency class and 830 images in the healthy class. The images in each of the diseased classes were preprocessed (resized) to the consistent input size for the model building processes. Figure 1 shows representative samples from the data set.

a. Healthy leaf a. Potassium Deficiency b. Sigatoka

Figure 01: Banana Leaf Disease Dataset

3.2 Proposed Methodology

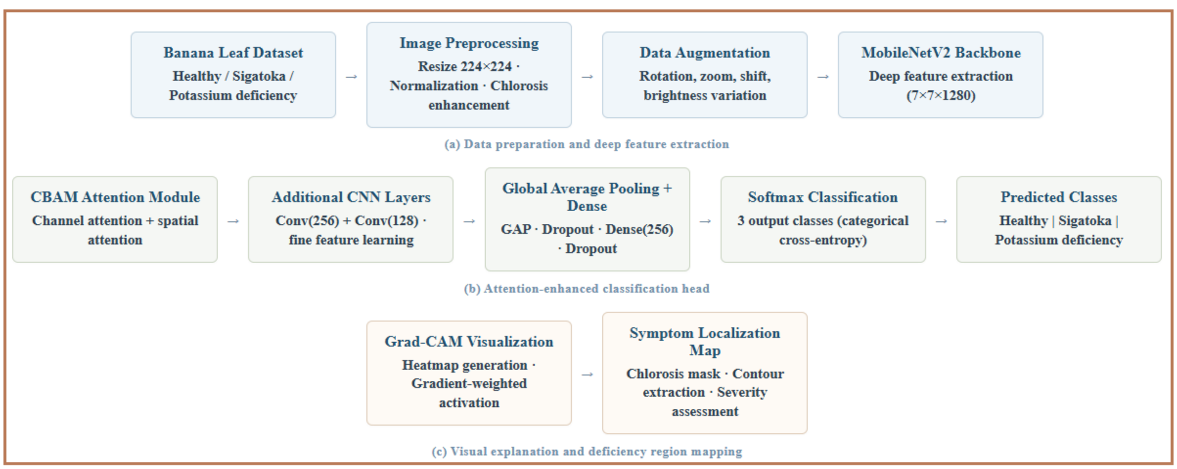

Figure 2: Three Stages of Proposed Methodology

The above figure 2 demonstrates the three stage proposed methodology that has an attention-enhanced deep learning architecture to automatically classify banana leaf image into three classes: Healthy, Sigatoka, and Potassium Deficiency.

In the first step we perform the pre-processing on the dataset. This includes resizing of the images in to 224 x 224 pixels and then perform the normalization. To highlight the nutrient deficiency chlorosis enhancement are performed. The Data augmentation technique such as rotation, zooming, shifting and brightness adjustment is applied to boost the robustness of the model and reduce the risk of overfitting. The resulting images are then fed into a pre-trained MobileNetV2 network which extracts high level features that are discriminative for leaf texture coloration and pathological symptoms. Following this a CBAM is added to improve the representation of the features by using channel-wise and spatial attention to focus on regions that suggest the disease. These features are further refined by adding additional convolutional layers then followed by Global Average Pooling layers, dense layers, dropout regularization and a Softmax classifier. The data are divided into 80% training and 20% testing. At last step Grad-CAM is employed to produce heat maps to localize symptoms and evaluate the severity, providing an interpretable and explainable solution for the detection of banana leaf diseases and nutrient deficiency.

The proposed framework addresses the challenge of effectively distinguishing between healthy leaves, Sigatoka-affected leaves and potassium deficient leaves because they have similar visual symptoms such as yellowing and discoloration as well as leaf damage. Conventional methods depend on visual inspection by human specialists a process that is often time-intensive, subjective and susceptible to mistakes. The proposed system addresses these limitations by integrating MobileNetV2 with a CBAM attention module. This combination enhances detection accuracy while directing the model focus toward disease-specific regions and nutrient deficiency symptoms. The key novelty is providing a seamless integration of chlorosis enhancement with attention guided feature extraction and explainable visualizations using Grad-CAM in a single architecture. This approach not only boosts classification accuracy but also enables the localization and severity assessment of symptoms. As a result the system becomes more reliable and interpretable making it well-suited for deployment in precision agriculture contexts.

3.3 MobileNetV2 Architecture

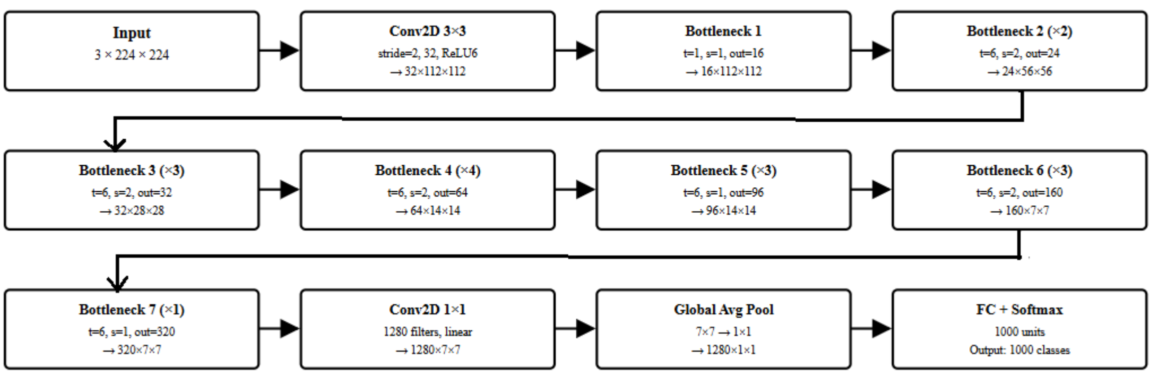

Figure 3:MobileNetV2 Feature Extractor

The architecture used to predict the diseases in banana plants is given below in Figure 3. The MobileNetV2 is a lightweight and efficient DL model that has the ability to accurately classify banana leaf diseases with lower computational requirements. The process starts with an input image of a banana leaf in the size of 224 × 224 × 3 pixels. The first convolutional layer captures basic visual information such as the edges of the leaves, the different color patches and the texture patterns. Then, the picture is fed into a series of bottleneck blocks that are the fundamental structural units of the MobileNetV2. There are three operations in each bottleneck block: expansion layer, depth wise convolution and projection layer. These components allow the model to incorporate disease-relevant features while keeping the number of parameters low. MobileNetV2 starts with a 3×3 convolution to shrink the spatial size from 224×224 to 112×112 in terms of architectural progression. There are then seven blocks of bottlenecks. The first one of the blocks is the size of the spatial size, while the later are progressively smaller and increasingly deeper. They are 24, 32, 64, 96, 160, and 320 spatial sizes deep respectively. Spatial filtering is applied to the channel dimension of the bottleneck with 3×3 depth wise convolutions which do not incur a significant computational cost followed by an expansion with ReLU6 activations. Lastly, a linear 1×1 projection layer (without ReLU activation) shrinks the channel size again allowing for the preservation of the important feature information. As more of these networks are developed, they start to represent more complex patterns such as lesions, necrotic spots, discoloration and regions of infected tissue on banana leaves. The number of features in each channel increases and the spatial resolution of each feature gradually decreases allowing the network to get more detailed information from the disease symptoms.

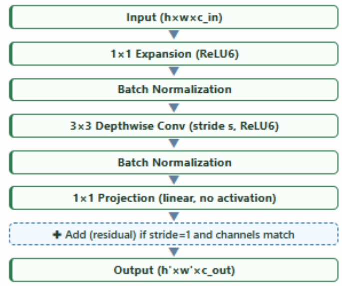

Figure 4: Bottleneck Layer (Inverted Residual)

The inverted residual block also known as the bottleneck layer is the key building block of MobileNetV2 as shown in Figure 4. The aim of this component is to be able to extract the relevant features efficiently with very little computational cost. The block starts with the 1×1 expansion convolution to increase channel dimensionality to enable the network to learn finely granular and disease-relevant information. Followed by a batch normalization to stabilize training. This is followed by a 3×3 depth wise convolution which encodes spatial patterns with low computational impact such as discoloration, leaf spots or lesions. A 1×1 projection convolution follows this which reduces the channel width so generating a compact feature representation. If the input and output dimensions are the same a residual connection is added to the output as the input itself. This skip connection enables the retention of crucial information and aids in the propagation of gradients more effectively during training. These designs allow for the MobileNetV2 network to be accurate and efficient in the classification of banana plant diseases.

3.4. Convolution Block Attention Module Architecture

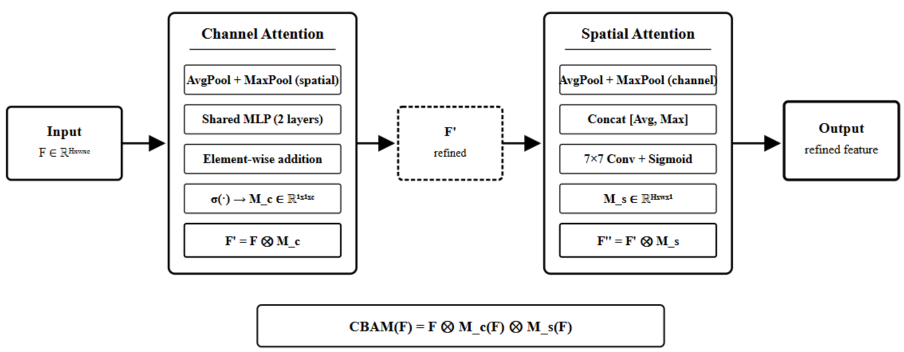

Figure 5: Convolution Block Attention Module Architecture

The above figure 5 demonstrates the architecture of the CBAM which perform channel wise attention followed by spatial wise attention to most important features. As shown in above figure 5 input feature map F (H, W, C) where H indicates height, W indicates width and C indicates Channels as an input. This input goes into the channel attention submodule that detects the most discriminative feature channels. In this submodule, the average pooling and max pooling is done over the spatial dimensions. These merged descriptors are supplied to a shared two layer MLP. The outputs of the MLP are added together and fed to a sigmoid activation function for generating the channel attention map M_c. This map is then multiplied to the original feature map F element-wise resulting in a channel-refined feature representation F'.

The refined feature map denoted as F' is then passed from the channel attention module to the spatial attention submodule. In this stage both average pooling and max pooling are applied along the channel dimension. The resulting feature maps are concatenated and fed into a 7×7 convolutional layer followed by a sigmoid activation function to generate the spatial attention map M_s. The overall CBAM process can be summarized as follows:

$$CBAM(F) = F \otimes M\_c(F) \otimes M\_s(F)$$

By applying channel and spatial attention in sequence the network is able to prioritize the most informative channels as well as the spatial regions that carry the most relevant information. Thus the discriminability and classification performance is greatly improved.

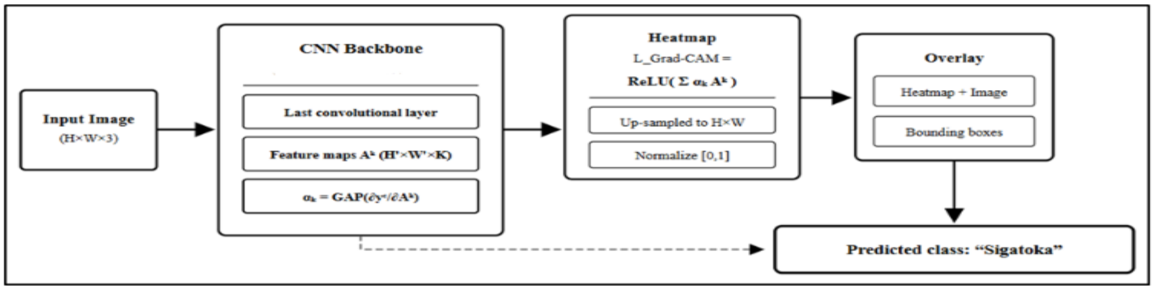

Figure 6: Grad-CAM pipeline

Figure 6 illustrates the working principle of Grad-CAM method used to interpret the decision making process of a CNN. The first step in the process is to get an H×W×3 sized image as an input to a CNN backbone (like MobileNetV2) to get H×W×C size feature maps AK from the final CNN convolutional layer.

For the given target class c the corresponding class score yc is first calculated. The gradients of this score with respect to each feature map are then computed using backpropagation. These gradients are globally averaged to calculate the neuron importance weights αk which quantify the contribution of each feature map to the final prediction. The Grad-CAM heatmap LGrad CAM is then generated by computing a weighted combination of the feature maps using Ak applying the ReLU activation function to retain only the positively contributing regions. The resulting heatmap is up-sampled to match the original image resolution and normalized to the range [0,1].

Finally the heatmap is superimposed on the original image optionally with bounding boxes to visually emphasize the areas that had the greatest impact on the model's decision. In the context of banana disease detection, Grad-CAM provides visual interpretability by identifying infected leaf areas responsible for classifying diseases such as Sigatoka, thereby improving model transparency, reliability and trustworthiness.

The model performance was assessed using four standard evaluation metrics: accuracy, precision, recall and F1 score

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

In this context a true positive (TP) refers to instances where the model correctly identifies the positive class such as correctly detecting diseased or healthy plants. A true negative (TN) occurs when the model correctly rejects the negative class. A false positive (FP) represents cases where the model incorrectly predicts the positive class while a false negative (FN) indicates situations where the model fails to identify the positive class and incorrectly predicts the negative class. For example: a plant with Sigatoka being identified as being in good condition would allow the plant to spread the disease and cause substantial loss of crop.

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

Table 1. Hyperparameter

Category	Hyperparameter	Value	Description
Input	Batch Size	32	No. of Samples per gradient update
	Epochs	15	Total Training Epoches
Model Architecture	Dense Layers	512 (ReLU)	Fully Connected head

		0.3 (Dropout)	
		(3, softmax)	
Optimizer	Type	Adam	Adaptive moment estimation
	Initial learning rate	0.001	Starting step size
	Learning rate scheduler	ReduceLROnPlateau	Factor = 0.2, patience = 3 epochs

The above equation 2 indicates the precision metric for the model evaluation. High precision value indicates that it has few false positives. When the model predict Sigatoka, then it is almost always correct. Precision measures the reliability of a positive prediction for each class.

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

The above equation 3 indicates the recall as evaluation metric. The recall indicates the sensitivity of detection for each class.

$$\text{F1 score} = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

The above equation 4 shows the F1 score as evaluation metric. It summarizes the model accuracy for each class. Precision indicates how many of the predicted disease labels were correct; recall indicates how many of the actual disease cases were detected. The model was trained for 15 epochs. An epoch means one complete pass through all training images. Each batch contained 32 images, meaning the model updated its internal parameters after every 32 images. The optimizer used was Adam, which is a popular method for adjusting learning during training. The initial learning rate, which controls how big a step the model takes when correcting errors, was set to 0.001. The loss function was categorical cross entropy this measures how far the predicted class probabilities are from the true one-hot encoded. The F1-score balances both, providing a robust evaluation for each banana leaf condition

4. Results and Discussions

The results of the proposed CBAM enhanced MobileNetV2 model for banana leaf disease detection are reported in this section. In this section the results of the proposed CBAM enhanced MobileNetV2 model used for Banana Leaf Disease Detection are reported. All experiments were performed with Python on the Google Colab platform with a GPU T4. The main objective of this study is to analyse the effectiveness of the proposed model. This method responds to the need for light systems in smallholder farming in rural areas [1, 17]. The test accuracy of the model was 94.4%. The enhanced Grad-CAM visualization is effective in creating heatmaps that highlight specific areas of the leaf relevant to the disease such as Sigatoka streaks and chlorotic patches thus providing interpretable disease localization. This interpretability is very useful to establish trust among farmers and severity assessment, which is not the case in black-box models [17].

In this study the data set used is 2,801 banana leaf images which were validated for classification of the disease. The images can be divided into three groups: potassium deficiency (896 images), healthy leaves (830 images) and Sigatoka disease (1075 images). Stratified splitting method was used to keep class size even. The dataset was split into a training and validation set with an 80:20 ratio, giving 2241 training images for training and 560 validation images for testing. In addition, 1,307 images were taken in an independent test set for the model assessment. This independent test dataset is also collected from different agricultural farm area so that our model can be tested on the real time dataset. This external independent dataset was not involved in either training or validation phase. This external dataset exclusively used for final performance evaluation. This ensure that unbiased assessment of the proposed model on unseen real time dataset. To fit deep learning architecture input all images were resized to a uniform size of 224 × 224 pixels with three RGB channels. This normalization process guarantees the uniformity of the data sets and facilitates the efficient extraction of features and classification for training and testing.

Table 2. Training and Validation accuracy per epoch

Epoch	Training Accuracy (%)	Validation Accuracy (%)	Gap (%)
1	84.5	96.5	8
3	87.0	81.5	5.5
5	88.1	83.0	4.9
7	90.1	87.1	3.0
9	94.5	89.0	5.5
11	95.5	92.0	3.5
13	96.5	93.5	3.0
14	98.1	94.5	3.5

The table 2 above displays the training and validation accuracy with respect to epoch. The above table shows that training accuracy from the epoch 1 to epoch 14 is steadily increasing. The 14th epoch is used as the early stopping point. The highest accuracy of 98% is achieved in the last epoch. The accuracy of the trained model is also improving from epoch 1 to epoch 14. The difference between training and validation accuracy is reaches to 3% at 13th epoch. This indicates the model shows the minimal overfitting and model generalised well on unseen data. With a regularization dropout rate of 0.3 the accuracy of the validation set is not decreased during the dropout process.

As show in bellow table 3 per-class classification report was calculated on the test dataset containing 1,307 images. The results show that the CBAM-MobileNetV2 model generalizes well to unseen data. Sigatoka detection achieves a high recall of 0.97, meaning very few infected leaves are missed. Potassium deficiency performs acceptably with recall 0.89 and precision 0.92. Healthy leaves are classified with an F1-score of 0.95. Weighted averages of precision 0.94, recall 0.95 and F1-score 0.94 confirm that the model is well suited for real-world banana disease diagnosis.

Table 3. Per class classification report

Class	Precision	Recall	F1Score	Support
Potassium deficiency	0.92	0.89	0.90	229
Healthy	0.96	0.95	0.95	182
Sigatoka	0.94	0.97	0.95	896
Weighted Average	0.94	0.95	0.94	1307

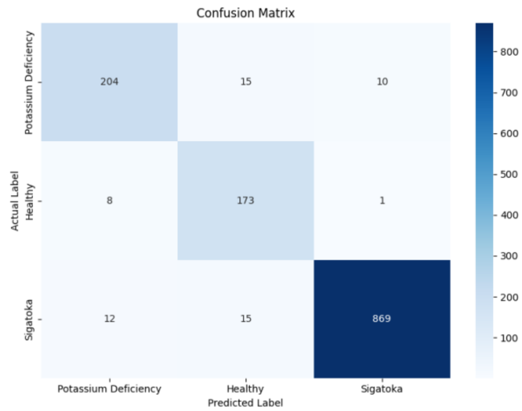

Figure 7: Confusion Matrix

The above figure 7 indicates the confusion matrix and get result from the 1307 test sample images. Sigatoka shows only 27 misclassifications that mostly with healthy leaves confirming excellent sensitivity. The model made 25 errors for potassium deficiency. 15 cases were misclassified as healthy and 10 as Sigatoka. This suggests some overlap in symptoms, yet the model still performed well and correctly identifying 89% of the actual potassium deficiency cases. The bellow figure demonstrates the result graph for the accuracy and loss per epoch.

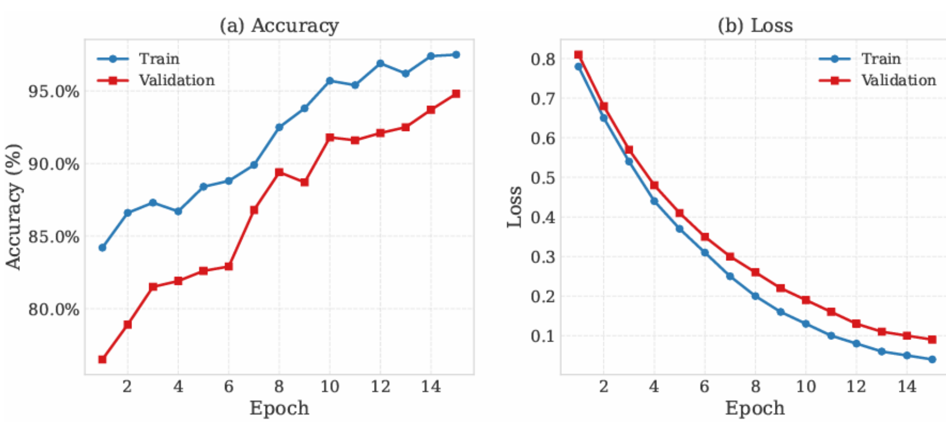

Figure 8: Accuracy and Loss graph per epoch

4.1 Sigatoka Detection

The model has a recall of 97 % and a precision of 94 % with an F1 score of 0.95 when applied to the test set to detect Sigatoka disease on the trees. The best validation accuracy was 94.5% upon which the difference from the accuracy during training was 3.5% at the 14th epoch. The MobileNetV2 model's capability to capture significant features from the banana leaf images effectively contributes to this good performance while the CBAM model's attention mechanism allows the model to concentrate on the disease symptoms like leaf streaking and color change while minimizing background information from the images. The small difference in the accuracy of the training and validation sets suggest that the model learnt meaningful disease patterns and generalized well to unseen data as opposed to memorizing training data a common problem in agricultural deep learning studies [2,17].

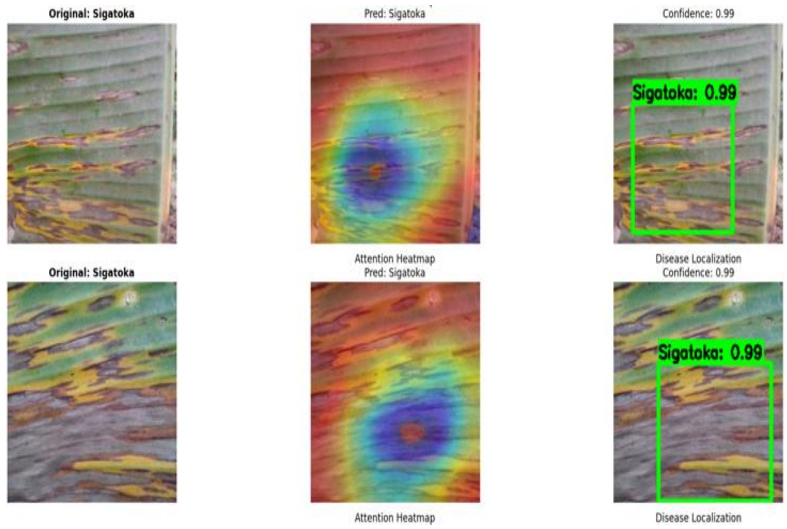

Figure 9: Attention Heatmap and Bounding box for Sigatoka disease detection

Figure 9 demonstrates the ability of the model to detect Sigatoka on banana leaves through Grad CAM. The original leaf images are shown on the left, the middle column shows heatmaps which identify which parts of the leaf had the most impact on the model's decision warmer colours correspond to more impactful regions and the right column shows regions of the leaf that were identified as diseased with bounding boxes around them. The model was able to predict Sigatoka with 99% confidence in both instances. The findings validate that the MobileNetV2 CBAM model is efficient in capturing the disease features like streaks, lesions and colour changes with high accuracy and explainability.

Table 4. Comparison with previous study

Study	Dataset & Methodology	Sigatoka Result and analysis
Elinisa & Mduma [11]	27360 images used with CNN Model	91.17% overall accuracy
Sanga et al. [16]	Tanzania field images with InceptionV3 & ResNet152 model	95% with mobile deployment
Lakshmi Narayanan et al. [9]	Hybrid CNN Model	96% overall accuracy
Proposed Methodology	CBAM with MobileNetV2 on 2805 real field images	97% recall with lightweight model and build interpretable model that can build trust for farmers

The above table 4 shows the comparison of the proposed methodology with the existing models. The farmers use this model as reliable first line screening tools for sigatoka disease detection. The high recall of the sigatoka detection reduces the risk of untreated infections.

4.2 Potassium Deficiency

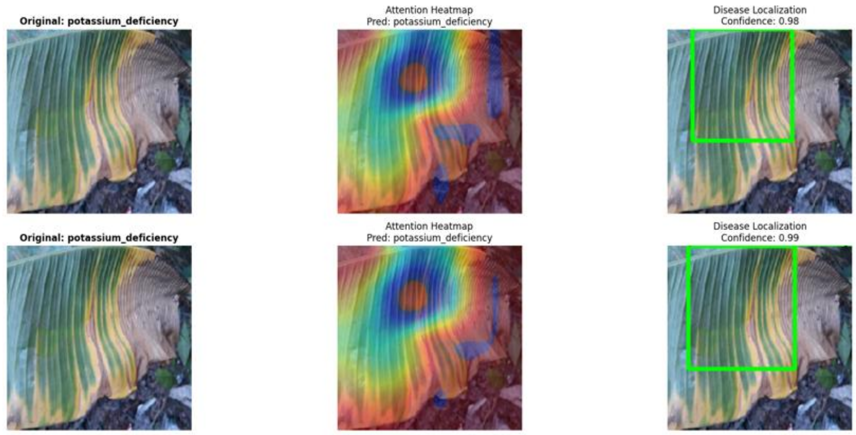

Figure 10: Attention Heatmap and Bounding box for Potassium Deficiency

The attention heatmaps and the disease localization results of banana leaves for Potassium deficiency are shown in figure 10. These original leaf images have characteristic yellowing and chlorosis along the leaf veins as shown in the left column. The attention heatmaps in the middle column show where the model concentrated its attention on the image, with warmer colors representing areas that had a significant impact on the prediction made. The areas that are highlighted indicate the areas where potassium deficiency is evident. The localization results in the right are presented with the presence of bounding boxes with a high confidence score of 0.98 and 0.99 respectively which accurately pinpoint the affected leaf areas. The results show that the proposed MobileNetV2-CBAM model can effectively capture nutrient-deficiency regions which not only boosts the accuracy of classification but also increases the interpretability of the model.

On a dataset of 7,214 images Sunil et al. [7] performed well in identifying eight different nutrient deficiencies using the IDWC-MSN model with an accuracy of 91.34% for Potassium deficiency. The proposed model achieved a potassium deficiency detection rate of 89% compared to the other model. This is an important result as both potassium deficiency and Sigatoka disease have similar symptomology and may be difficult to differentiate. In contrast to other studies, this paper tackles the problem of nutrient deficiency and fungal disease classification together. The results highlight the need for accurate validation and class-wise assessment since overall accuracy cannot be an estimate of the model performance. On the practical side correct diagnosis can aid in the proper timing of fertilizers, prevent unnecessary fungicide applications, cut costs and prevent damage to the environment.

While the proposed model gave good performance, there were a few constraints to be taken into account. Only three classes of banana diseases and many other nutrient deficiencies were studied. The number of images in the data set, 2,801 is relatively small when compared to larger agricultural data sets. Images were taken at one growth stage with only RGB cameras which would not allow for temporal analysis and advanced spectral evaluation. Proposed methodology did not include external validation and variations in lighting may impact real-world performance.

5. Conclusions

In this research a light-weight DL model combined with attention mechanism was successfully designed to detect diseases and nutrient deficiencies in banana leaves. The proposed CBAM-MobileNetV2 model obtained an overall accuracy of 94.4%, a high recall of 97% in Sigatoka detection and moderate recall of 89% in identifying potassium deficiency. Although the primary dataset contains only 2,801 images this limitation was rigorously addressed by collecting an independent test set of 1,307 images from different agricultural farms which was used exclusively for final evaluation. The strong performance on this unseen independent dataset confirms that the model generalizes well despite the modest training data size. These outcomes directly respond to the requirement for accurate and real-time diagnostic tools to be suitable for resource constrained smallholder farming systems. Grad-CAM visualizations offer interpretable heat maps and bounding boxes to bolster farmer confidence and aid in the assessment of symptom severity. The main novelty of this work is the successful integration of channel wise and spatial attention mechanisms with the MobileNetV2 backbone that is the ability to accurately detect both fungal diseases and nutritional disorders even in case of visual symptoms resembling one another. The approach should be extended to other disease categories with temporal monitoring at various growth stages as well as adding to the imaging using such an unmanned aerial vehicle and multispectral data and by testing the model in real-time field testing.

7. Acknowledgements

The author declares no conflict of interest. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. J.-R. Liao, Ed., AI-driven banana pest and disease management: methods, applications, challenges, and future directions, vol. 5, no. 110. Liao Discover Internet of Things, 2025, doi.org/10.1007/s43926-025-00175-9
2. Verry Riyanto, Sri Nurdiati, Marimin, Muhamad Syukur, Shelvie Nidya Neyman, Ed., Deep Learning Approaches for Plant Disease Diagnosis Systems: A Review and Future Research Agendas, vol. 9, no. 2-185–201. Journal of Applied Agricultural Science and Technology, 2025 doi.org/10.55043/jaast.v9i2.308
3. Hong P., Luo Xi, Bao L., “Crop disease diagnosis and prediction using two-stream hybrid convolutional neural networks,” Crop Protection, vol. 184, no. 106867, 2024. doi.org/10.1016/j.cropro.2024.106867
4. P. Sajitha, Diana Andrushia, N. Anand, M.Z. Naser, “A Review on Machine Learning and Deep Learning Image-based Plant Disease Classification for Industrial Farming Systems,” Journal of Industrial Information Integration, vol. 38, no. 100572, Mar. 2024. DOI:10.1016/j.jii.2024.100572
5. A. J. Andrew, J. Eunice, D. E. Popescu, M. K. Chowdary, and J. Hemanth, “Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications,” Agronomy, vol. 12, no. 10, p. 2395, Oct. 2022. doi.org/10.3390/agronomy12102395
6. J. D. Thiagarajan, S. V. Kulkarni, S. A. Jadhav, A. A. Waghe, S. P. Raja, S. Rajagopal, H. Poddar, and S. Subramaniam, “Analysis of banana plant health using machine learning techniques,” Scientific Reports, vol. 14, no. 1, p. 15041, Jul. 2024, DOI: 10.1038/s41598-024-63930-y.
7. C. K. Sunil, R. Anusha, C. K. Raghavendra, and G. Dai, “A multi-scale deep learning approach for nutrient deficiency detection in banana plants using adaptive feature pooling,” Neural Computing and Applications, vol. 38, no. 4, p. 71, Feb. 2026, doi.org/10.1007/s00521-025-11726-0
8. T. Domingues, T. Brandão, and J. C. Ferreira, “Machine learning for detection and prediction of crop diseases and pests: A comprehensive survey,” Agriculture, vol. 12, no. 9, p. 1350, Sep. 2022. doi.org/10.3390/agriculture12091350
9. K. Lakshmi Narayanan, R. Santhana Krishnan, Y. Harold Robinson, E. Golden Julie, S. Vimal, V. Saravanan and M. Kaliappan, “Banana plant disease classification using hybrid convolutional neural network,” Computational Intelligence and Neuroscience, vol. 2022, pp. 1–10, Feb. 2022. doi.org/10.1155/2022/9153699
10. N. Mduma and J. Leo, “Dataset of banana leaves and stem images for object detection, classification and segmentation: A case of Tanzania,” Data in Brief, vol. 49, p. 109322, Aug. 2023. doi.org/10.1016/j.dib.2023.109322
11. C. A. Elinisa and N. Mduma, “Mobile-Based convolutional neural network model for the early identification of banana diseases,” Smart Agricultural Technology, vol. 7, p. 100423, Mar. 2024. https://doi.org/10.1016/j.atech.2024.100423
12. J. A. Pandian, V. D. Kumar, O. Geman, M. Hnatiuc, M. Arif, and K. Kanchanadevi, “Plant Disease Detection Using Deep Convolutional Neural Network,” Applied Sciences, vol. 12, no. 14, p. 6982, Jul. 2022. doi.org/10.3390/app12146982
13. R. Sangeetha, J. Logeshwaran, J. Rocher, and J. Lloret, “An Improved Agro Deep Learning Model for Detection of Panama Wilts Disease in Banana Leaves,” AgriEngineering, vol. 5, no. 2, pp. 660–679, 2023. doi.org/10.3390/agriengineering5020042
14. M. G. Selvaraj, A. Vergara, H. Ruiz, N. Safari, S. Elayabalan, W. Ocimati, and G. Blomme, “AI-powered banana diseases and pest detection,” Plant Methods, vol. 15, no. 1, p. 92, Aug. 2019. doi.org/10.1186/s13007-019-0475-z.
15. M. G. Selvaraj, A. Vergara, F. Montenegro, H. A. Ruiz, N. Safari, D. Raymaekers, W. Ocimati, J. Ntamwira, L. Tits, A. B. Omondi, and G. Blomme, “Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 169, pp. 110–124, Nov. 2020. doi.org/10.1016/j.isprsjprs.2020.08.025
16. S. L. Sanga, D. Machuve, and K. Jomanga, “Mobile-based Deep Learning Models for Banana Disease Detection”, Eng. Technol. Appl. Sci. Res., vol. 10, no. 3, pp. 5674–5677, Jun. 2020. https://doi.org/10.48084/etasr.3452
17. K. Patel and A. Patel, “Plant disease diagnosis using image processing techniques A review on machine and deep learning approaches,” Ecology, Environment and Conservation, vol. 28, no. 2, Jan. 2022. 10.53550/EEC.2022.v28i02s.057
18. S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” Frontiers in Plant Science, vol. 7, p. 1419. doi.org/10.3389/fpls.2016.01419
19. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” Computers and Electronics in Agriculture, vol. 145, pp. 311–318, Feb. 2018. 10.1016/j.compag.2018.01.009

20. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. Hughes, “Deep learning for image-based cassava disease detection,” Frontiers in Plant Science, vol. 8, p. 1852, Oct. 2017. doi.org/10.3389/fpls.2017.01852
21. A. Picon, A. Alvarez-Gila, M. Seitz, A. Ortiz-Barredo, J. Echazarra, and A. Johannes, “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, Jun. 2019. 10.1016/j.compag.2018.04.002
22. A. S. Keceli, A. Kaya, C. Catal, and B. Tekinerdogan, “Deep learning-based multi-task prediction system for plant disease and species detection,” Ecological Informatics, vol. 69, Art. no. 101679, Jul. 2022. doi.org/10.1016/j.ecoinf.2022.101679
23. A. M. S. Kheir, A. Koubaa, V. Kolluru, S. Mungara, and T. Feike, “Smart plant disease diagnosis using multiple deep learning and web application integration,” Journal of Agriculture and Food Research, vol. 21, Art. no. 101948, Jun. 2025. DOI:10.1016/j.jafr.2025.101948
24. G. UshaDevi and B. V. Gokulnath, “A survey on plant disease prediction using machine learning and deep learning techniques,” Inteligencia Artificial: Revista Iberoamericana de Inteligencia Artificial, vol. 23, no. 65, pp. 136–154, Jul. 2020. DOI:10.4114/intartif.vol23iss65pp136-154
25. A. Y. Ashurov, M. S. A. M. Al-Gaashani, N. A. Samee, R. Alkanhel, G. Atteia, H. A. Abdallah, and M. S. A. Muthanna, “Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections,” Frontiers in Plant Science, vol. 15, 2025. https://doi.org/10.3389/fpls.2024.1505857
26. W. Shafik, A. Tufail, C. L. De Silva, and R. A. A. H. Mohd Apong, “A novel hybrid inception-xception convolutional neural network for efficient plant disease classification and detection,” Scientific Reports, vol. 15, no. 1, pp. 1–18, Jan. 2025. https://doi.org/10.1038/s41598-024-82857-y