Review of Computer Engineering Research
2027 Vol. 14, No. 1, pp. 75-95
ISSN(e): 2410-9142
ISSN(p): 2412-4281
DOI: 10.18488/76.v14i1.5379
© 2027 Conscientia Beam. All Rights Reserved.

Attention enhanced lightweight deep transfer learning framework for early paddy leaf disease classification using CBAM-MobileNetV2 in smart farming

Karipe Mahesh[1+]
VVSSS Balaram[2]

[1,2]*Department of Computer Science and Engineering, School of Engineering, Anurag University, Hyderabad, Telangana, India.*
[1]*Email: karipemahesh22@gmail.com*
[2]*Email: vbalaram23@gmail.com*



(+ Corresponding author)

Article History
Received: 13 July 2026
Revised: 17 September 2026
Accepted: 9 October 2026
Published: 23 November 2026

Keywords
Computer vision
Edge AI
Image classification
Lightweight neural networks
Smart agriculture.

ABSTRACT

Plant diseases pose a very serious threat to agricultural production throughout the world, with bacterial and fungal infections in paddy rice leading to significant yield losses. Conventional manual verification is a time-consuming and subjective process, particularly in large-scale smart farming. Recently, deep learning-based approaches have been applied to classify plant diseases, but most require computationally expensive architectures or impractical hyperspectral imaging. This paper proposes a lightweight deep transfer learning-based framework for the early-stage classification of paddy leaf diseases using MobileNetV2 with a Convolutional Block Attention Module. The framework uses image preprocessing, augmentation, transfer learning, and attention-based mechanisms to classify early-stage leaf infections with a relatively low computational load. The proposed method was benchmarked on the publicly available Rice Leaf Bacterial and Fungal Disease dataset. The five experimental runs yielded a mean accuracy of 97.1% (±0.3%). The framework outperformed existing convolutional neural networks (CNNs) and MobileNetV2 variants. It is lightweight and efficient, with only 2.4 million parameters, 0.32 GFLOPs, and an inference time of 18 ms per image on a CPU. The dataset used for the experiments contained images with early-stage and severe disease symptoms. The framework was able to generalize well to images with mild symptoms, which is critical for early intervention. The proposed framework can be used in smart farming applications to enable early-stage intervention and reduce losses.

Contribution/Originality: This research work contributes to the literature by introducing an attention-based, lightweight deep transfer learning approach for the early categorization of paddy leaf infections. This study presents a new combination of CBAM and MobileNetV2 for disease-specific feature refinement. This study is among the few works that have considered lightweight RGB-based classification, with an emphasis on deployment on edge devices.

1. INTRODUCTION

Paddy (Rice) is one of the most widely cultivated cereal crops in the world. It is the staple food for more than fifty percent of the world's population that particularly belongs to Asian countries. However, this crop is highly vulnerable to bacterial and fungal infections, notably bacterial leaf blight, rice blast, and sheath blight, among others. These diseases lead to annual yield loss of twenty to fifty percent [1] these infections spread rapidly and remain dormant until they are visible to the naked eye, making it critical to detect them at the earliest stages. Conventionally, expert agronomists and farmers have relied on manual inspection [2] for disease detection.

© 2027 Conscientia Beam. All Rights Reserved.

Manual inspection is usually tedious, subjective, time-consuming, and unproductive for large-scale detection. In addition, the lack of expert agronomists in disease detection compounds the problem, delaying the diagnosis and treatment of infected plants [3]. The AI/Computer Vision (CV)/Deep Learning (DL) community has made great strides in developing automatic disease detection systems that operate using images of infected plants [4]. For instance, hyperspectral imaging techniques have been employed [5] to detect diseases by analyzing the spectrum of light reflected off the leaves. Although hyperspectral imaging is sensitive to disease detection, it requires costly imaging hardware and extensive preprocessing of the spectral data before model training [6].

On the other hand, CNNs have enhanced plant disease detection by automatically learning discriminative features from the input images. Previous works have used CNNs such as ResNet, DenseNet, and EfficientNet to classify plant diseases [7]. However, these CNNs have too many parameters and are too computationally intensive for deployment in typical agricultural setups.

It is worth noting that CNNs designed for lightweight applications such as MobileNetV2, EfficientNet-Lite, and vision transformers are suitable for plant disease detection. However, these CNNs often have restricted feature discrimination capabilities; it is hard to distinguish between similar features at the pixel level [8]. One approach that has been employed to address the limited pixel-level discrimination of CNNs is the development of attention mechanisms. Attention mechanisms maximize feature discrimination by enabling the model to focus on the parts of the image (e.g., diseased portion of the leaf) while ignoring the non-discriminative portions (e.g., healthy portion of the leaf). The Convolutional Block Attention Module CBAM is an example of an attention mechanism that has been used successfully in plant disease detection [9]. Specifically, the CBAM first employs a channel attention module that learns global feature discrimination in the image, followed by a spatial attention module that learns local feature discrimination in the image. Recently, the CBAM has been used extensively in plant disease detection.

The following subsections provide a detailed discussion on the research issues and problems addressed by our work.

1.1. Research Issues

Lightweight attention-based networks: Most of the attention-based networks that have been explored for plant disease detection have not used a lightweight backbone convolutional network. Specifically, most attention-based networks use ResNet or DenseNet as their backbone CNN, even though these networks have more than 100 million trainable parameters, making them unsuitable for practical agricultural applications [10]. Also, most of the attention-based networks that have been explored in plant disease detection are not suitable for early disease detection. A number of the previous works have used hyperspectral imaging rather than the typical RGB imaging that is amenable to large-scale deployment [11].

Validation of practical deployment: Although some of the previous approaches have been reported as lightweight CNNs, a comparison of their computational requirements (in the form of processing time, FLOPs, and number of parameters) can guide the choice of networks suitable for agricultural applications. In addition, few previous works have compared the effect of different attention mechanisms in plant disease detection.

Insights into attention mechanisms: Although the combination of CNN and attention mechanisms has been widely studied in plant disease recognition, most of the attention modules' contributions have not been systematically evaluated by ablation experiments.

1.2. Research Gap

Based on the aforementioned research issues, we argue that no previous work has explored using a lightweight attention-enhanced CNN for early disease detection of paddy leaf diseases [12] through RGB imaging in both training and inference stages while reporting the computational efficiency in terms of FLOPs, parameters, and inference time.

© 2027 Conscientia Beam. All Rights Reserved.

1.3. Research Questions

i. Does the combination of a lightweight MobileNetV2 and the CBAM enable effective disease detection in paddy leaves?
ii. What is the relative contribution of the CBAM attention mechanism?
iii. Can the proposed approach achieve satisfactory performance while maintaining a low computational cost that enables real-time inference on embedded systems?

1.4. Objectives

i. Developing an attention-enhanced lightweight Deep Transfer Learning Framework for Paddy Leaf Disease Classification.
ii. Proposing a disease detection framework that employs a lightweight MobileNetV2 and CBAM for disease-sensitive feature refinement.
iii. Evaluating the efficacy of the suggested method in terms of correctness and computational requirements in order to maximize practical deployment.
iv. Conducting an ablation study to demonstrate the individual contribution of the proposed attention mechanism, data augmentation, and transfer learning.

1.5. Hypothesis

It was hypothesized that the proposed approach would show high accuracy in classification due to the use of an efficient model capable of extracting features with lower computational power and therefore suitable for real-time implementation on embedded devices. The structure of this work is as follows: Section 2 contains a literature review; Section 3 provides a comparative analysis of related works and presents the research objective; Section 4 describes the approach; Section 5 presents the experimental setup and implementation; and Sections 6 and 7 contain the results and discussion. Finally, Section 8 concludes the research findings and recommends future improvements.

2. LITERATURE REVIEW

This section presents a critical review of state-of-the-art plant disease detection approaches. The related work is classified into spectral and hyperspectral techniques, machine learning-based methods, deep learning-based approaches, and emerging vision transformer and attention-based networks. This literature review concludes by identifying the limitations of existing approaches that this study aims to address through the proposed framework.

2.1. Spectral and Hyperspectral Disease Detection

Mahlein, et al. [5] proposed spectral indices for hyperspectral disease detection, which showed excellent disease discrimination capacity. However, the authors relied on costly hyperspectral imaging sensors and complex pre-processing steps. Likewise, Bauriegel, et al. [13]. Utilized hyperspectral remote sensing for the early detection of Fusarium in wheat with great sensitivity but at the expense of computational complexity [14]. Conducted an experimental analysis of spectral reflectance for early wheat disease detection. The authors did not utilize any deep learning algorithms. Delalieux, et al. [15]. Conducted hyperspectral image analysis for apple tree disease detection, which requires complex data processing.

Critical assessment: Although hyperspectral imaging has excellent sensitivity and specificity for disease detection, the requirement for costly and complex equipment reduces its applicability in practice. This work demonstrates that standard RGB cameras can be utilized for plant disease diagnosis.

© 2027 Conscientia Beam. All Rights Reserved.

2.2. Machine Learning Based Disease Diagnosis

Wang, et al. [16] suggested an artificial neural network-based disease prediction system which leveraged spectral reflectance features for tomato disease detection. The authors showed that their framework could perform automatic disease recognition. However, the methodology required extensive feature engineering [17]. Proposed Genetic Algorithm optimized support vector machines for wheat disease detection [17]. The authors reported improved classification accuracy. However, the methodology required extensive feature engineering [18]. Proposed machine learning regression-based leaf rust detection system which had prohibitively large computational footprints. Verma pioneered random forest-based classification of plant diseases, which again required extensive feature engineering compared to deep learning-based approaches [19].

Critical assessment: Most of the machine learning-based disease detection methods reported excessive reliance on manual feature engineering, which limited their applicability and generalizability. This work demonstrates that feature engineering can be avoided by utilizing automatic feature extractors like deep neural networks.

2.3. Deep Learning Based Disease Detection

2.3.1. CNN-Based Approaches

Saleem, et al. [20] performed a survey of deep learning-based plant disease detection methods. The authors concluded that CNN based methods outperformed traditional machine learning methods [21]. Proposed a deep learning framework for yellow rust detection based on hyperspectral UAV images. The authors showed that the CNN based framework could achieve excellent detection performance in the field. However, the methodology required high-end computational hardware [22]. Proposed an Explainable 3D deep learning framework for hyperspectral plant disease detection. The authors showed that their CNN based framework achieved excellent diagnostic performance. However, the 3D CNNs were extremely computationally intensive [23]. Proposed a hyperspectral vineyard disease detection framework which required extensive equipment for data acquisition [24]. Suggested a UAV aided substructure for grapevine disease identification based on deep learning. The methodology again required high-end UAV hardware.

Critical assessment: Most of the CNN approaches for plant disease identification were not lightweight and relied on either hyperspectral data or high-end computer vision models. This work presents a novel edge-deployable framework for plant disease detection.

2.3.2. Lightweight CNN Approaches

Recently, there has been an explosion of research on lightweight CNN-based plant disease detection frameworks. MobileNetV2 is one of the first CNNs to propose depthwise separable convolutions and achieve excellent accuracy while remaining lightweight. EfficientNet-Lite is a series of CNNs that optimizes model scaling for lightweight applications [10]. FL-ToLeD, a lightweight attention-based CNN for tomato leaf disease detection, was proposed for use on low-end devices. This framework achieved excellent accuracy while reducing the number of parameters. Lachure and Doriya suggested a lightweight deep learning framework for plant disease detection. The framework was able to achieve good performance in plant disease identification while remaining computationally inexpensive [25]. BerryNet-Lite was proposed for strawberry disease identification and achieved excellent performance on edge devices [26]. A lightweight CNN with Grad-CAM was also proposed for grape leaf disease classification on edge devices.

Critical assessment: All studies on lightweight CNN-based plant disease detection achieved moderate to excellent performance while reducing computational demands. However, most studies focused on tomato and grape leaf disease detection, with limited attention to paddy leaf diseases.

© 2027 Conscientia Beam. All Rights Reserved.

2.3.3. Vision Transformers and Advanced Architectures

Vision Transformers have emerged as a prominent alternative to CNNs for computer vision tasks. These architectures use self-attention mechanisms to achieve remarkable accuracy on the ImageNet dataset, albeit at the expense of higher computational requirements. Several approaches have been proposed for using Vision Transformers in plant disease detection [27]. The original Vision Transformer achieved excellent performance on ImageNet but required substantial computational resources. Recently, several approaches have been proposed to make Vision Transformers more lightweight for practical applications.

EfficientFormer is a dimension-consistent design that enables efficient model compression while still achieving excellent accuracy [28]. CNNs have also been optimized to apply efficient model design principles from Vision Transformers. For example, the ConvNeXt family of CNNs demonstrated that vanilla CNNs can achieve performance comparable to that of Vision Transformers when optimized [29]. MobileViT is another recent architecture that combines CNNs and Vision Transformers [30]. For plant disease identification, foundation vision models such as Agriculture-ViT and Plant-Foundation have been proposed, leveraging large-scale pretraining on plant phenotyping data.

Critical assessment: Most approaches for plant disease detection based on Vision Transformers still require substantial computational resources. This study demonstrates that a combination of efficient CNNs and lightweight attention mechanisms can achieve comparable performance to Vision Transformers while being more practical for edge deployment.

2.3.4. Attention Based Disease Detection

The attention mechanisms have been utilized extensively for plant disease detection for enhanced detection performance [9]. Suggested the Coordination Attention (CBAM) module, which is a lightweight attention mechanism that can be utilized for various tasks. The authors demonstrated that their framework achieved excellent performance on plant disease detection benchmarks [12]. Suggested an understandable fusion model based on a lightweight CNN and vision transformer for rice leaf disease detection [31]. Utilized EfficientNetB3 along with augmentation strategies for multi-class plant disease classification.

Critical assessment: Attention-based approaches for plant disease detection lack validation on early-stage disease detection. This work presents a framework that utilizes CBAM attention with MobileNetV2 for edge-deployable paddy leaf disease detection with extensive evaluation on the impact of attention mechanisms.

3. COMPARATIVE ANALYSIS

Following Table 1 represents a comparative analysis of representative methods.

Table 1. Comparative analysis of current methods.

Reference	Methodology	Dataset	Major strength	Limitation identified
Mahlein, et al. [5]	Spectral Indices	Hyperspectral	Disease classification	Very expensive setup
Bauriegel, et al. [13]	Hyperspectral Imaging (HSI)	Wheat Disease	Early-stage detection	High Complexity
Bravo, et al. [14]	Reflectance Spectrum	Field Crops	Remote sensing	Limited AI implementation
Wang, et al. [16]	Artificial Neural Network (ANN)	Tomato Disease	Integrated with AI	Manually extracted features
Huang, et al. [17]	Genetic Algorithm + Support Vector Machine	Wheat Disease	Improved classification	Feature transformation
Saleem, et al. [20]	Convolutional Neural Network (CNN) Review	Plant Village	Deep Learning Analysis	Only review-oriented
Zhang, et al. [2]	CNN+UAV	UAV Hyperspectral	Field Monitoring	High Computational Cost
Nagasubramanian, et al. [22]	XAI 3D CNN	Hyperspectral	Explainability	Clumsy Architecture

© 2027 Conscientia Beam. All Rights Reserved.

3.1. Research Gap Analysis

From the critical literature review, the following research gaps can be identified.

- Lightweight plant disease detection framework for rice leaf disease detection: Most of the present literature on lightweight plant disease detection focused on tomato leaf, grape leaf, or strawberry disease detection [32].
- Few studies explicitly target early-stage disease detection: Most of the studies do not explicitly mention detection of disease in the early stage (i.e., subtle lesions) [12].
- Lack of comprehensive edge deployment metrics: Many of the studies report a lightweight nature of their framework but do not report any comprehensive metrics for the computational requirements.
- Lack of explicit analysis of the contribution of attention modules: Very few studies explicitly analyze the contribution of the different attention modules using ablation studies [10].

How This Work is Different from the Prior MobileNetV2-Based Works: Although some of the existing works utilize MobileNetV2 for plant disease identification. This work differs from the prior art in several aspects.

- Utilization of attention mechanism: This work utilizes the CBAM attention mechanism for refining the features extracted using MobileNetV2.
- Focus on early-stage disease detection: This work focuses on early-stage paddy disease detection based on subtle lesions.
- Comprehensive edge deployment metrics: This work provides the parameters, FLOPs, and CPU inference time for the proposed framework.
- Analysis of contribution of attention and augmentation: This work analyzes the contribution of the attention mechanisms and data augmentation strategies using ablation studies.
- Focus on paddy leaf disease detection: This work explicitly focuses on paddy leaf disease detection.

The following Table 2 describes the research gaps identified from the literature.

Table 2. Gaps Identification.

Existing Limitation	Required Improvement
Expensive hyperspectral systems	Low-cost RGB image systems
Heavy CNN architectures	Lightweight deep learning
High computational complexity	Edge-device deployment
Weak disease localization	Attention-guided learning
Complex feature engineering	Automatic feature extraction

Existing systems are either limited to hyperspectral imaging or use deep architectures with high computational cost. Hence, it is essential to utilize a lightweight attention-guided deep learning framework with high disease detection capability using standard RGB imagery. The proposed work addresses this gap by leveraging the efficiency of MobileNetV2 along with CBAM for the attention mechanism.

4. PROPOSED METHODOLOGY

4.1. Proposed Framework Overview

The proposed framework comprises preprocessing, augmentation, transfer learning, lightweight feature extraction, and attention-guided disease classification. Here, MobileNetV2 is employed with CBAM integrated as an attention module for feature refinement, specifically targeting disease detection. Figure 1 presents the proposed attention-enhanced disease detection framework.

© 2027 Conscientia Beam. All Rights Reserved.

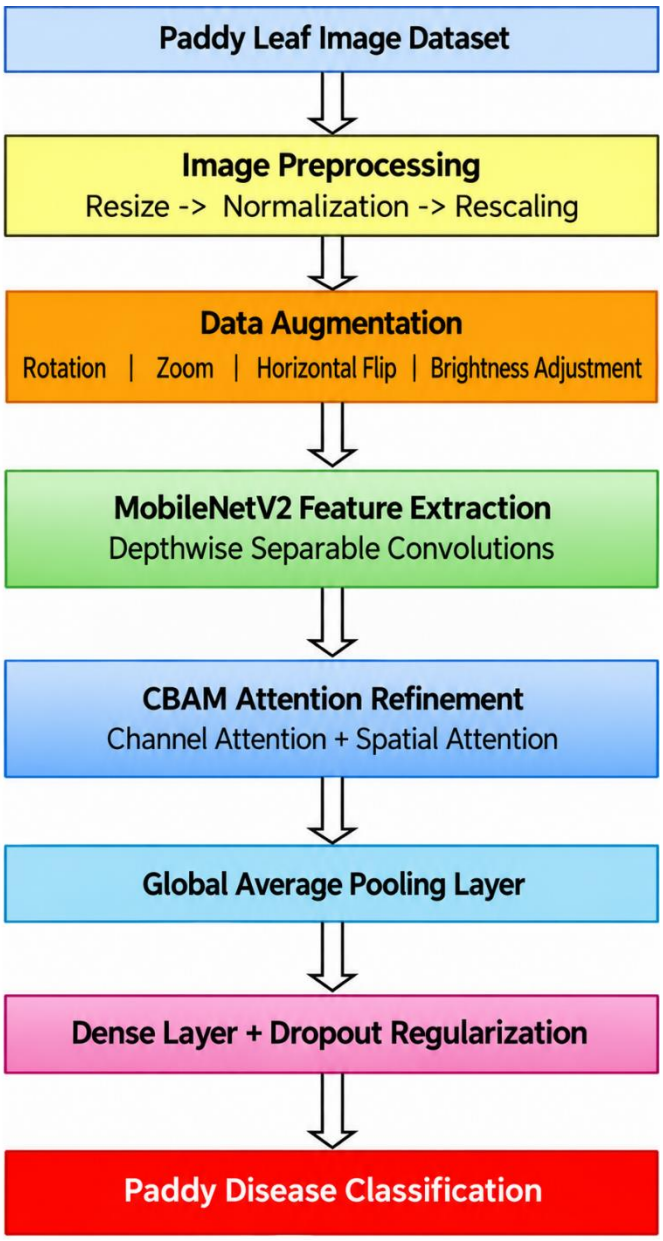

Figure 1. Attention-enhanced disease detection proposed framework.

4.2. Dataset Description

The dataset used is the Rice Leaf Bacterial and Fungal Disease dataset from Kaggle [33]. It contains 4,500 RGB images divided into three classes. Table 3 presents details of the distribution of healthy, bacterial, and fungal-diseased leaf images.

Table 3. Description of the types of leaves considered.

Types of leaves	No. of images
Good samples of healthy leaves	1500
Bacterial disease	1500
Fungal disease	1500

All images have been resized to 224 × 224 pixels. The entire dataset is randomly divided into 80% training and 20% Validation (stratified by class) [34]. No test set was held out; validation performance is recorded as the decisive performance. The following Figure 2 shows the sample images from the dataset.

© 2027 Conscientia Beam. All Rights Reserved.



Figure 2. Samples of Paddy Disease Images.

4.3. Image Preprocessing and Augmentation

Preprocessing operations: resizing (224 × 224), normalization (0, 1), and rescaling.

$$Inorm = I/255$$

Where

I= Original Image (RGB), $Inorm$ =Standardized Image

Data augmentation was applied only to the training data: random rotation (±20°), random zoom (0.8–1.2), horizontal flipping, and brightness adjustment (0.8–1.2). These operations increase the diversity of our dataset and reduce overfitting.

4.4. MobileNetV2 Architecture

MobileNetV2 was chosen due to its low model size and complexity [8] fast inference and great performance on edge-device deployment. Figure 3 presents the proposed architecture of MobileNetV2-CBAM Model. The model implements transfer learning with ImageNet weights [7].

$$F = M(I_{norm;}\, \theta)$$

Where: M = MobileNetV2, θ = Pertained ImageNet parameter, F = Extracted feature map.

Applied with a global average pooling layer, a Dense Layer (128units, ReLU), 0.5 dropout and a final softmax layer with 3 outputs.

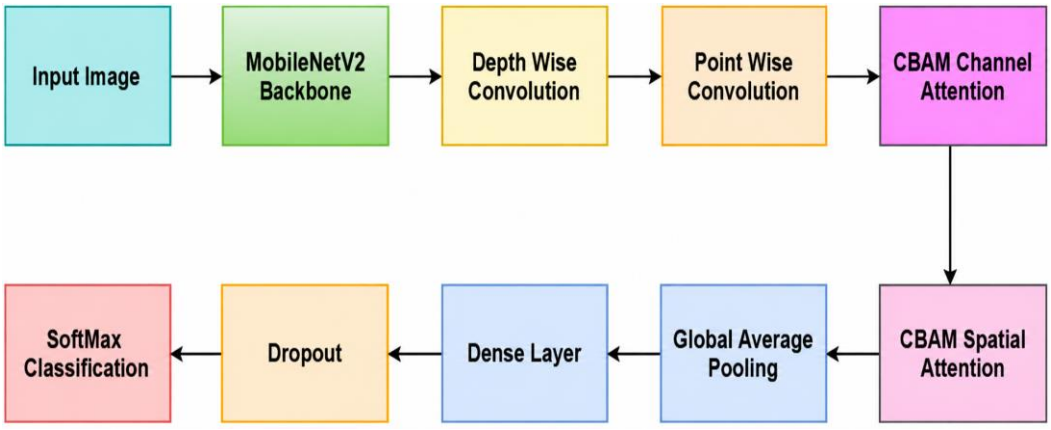

Figure 3. MobileNetV2 – CBAM proposed architecture for paddy disease classification.

4.5. CBAM Attention Mechanism

The CBAM module uses a spatial attention module that primarily focuses on disease-sensitive areas [9], thereby enhancing feature representation. It applies channel and spatial attention sequentially.

© 2027 Conscientia Beam. All Rights Reserved.

Channel Attention:

$$M_c(F) = \sigma(\text{MLP}(\text{AVG Pool(F)}) + \text{MLP}(\text{MAXPool(F)}))$$

Spatial Attention:

$$M_s(F) = \sigma(f7X7[AVGPool(F); MAXPool(F)])$$

where:

F: Input feature maps, Σ: Sigmoid Activation, MLP: Multi-Layer Perceptron along with weight sharing, 7×7 convolution (f7×7).

$$F_{out} = M_s(F) \otimes M_c(F) \otimes F$$

Where:

F: Input Feature Maps, Mc(F): Channel Attention Map, M_s(F): Spatial Attention map, $\otimes$: Element wise multiplication, F_{out} = resultant feature map after CBAM.

The attention mechanism improves the disease localization ability and feature discrimination without increasing parameters a lot.

Algorithm: *Proposed attention-enhanced lightweight transfer learning framework*
Input: *Rice leaf RGB Image I*
Output: *predicted disease class C {Healthy, Bacterial, Fungal}*
Step 1: Consider and Read the RGB Image I.
Step 2: Resize the image to 224x224 pixels.
Step 3: Now normalize pixel values to [0,1]
Step 4: Now apply data augmentation:
Random Rotation
Horizontal Flip
Random Zoom
Brightness Adjustment
Step 5: Now load the pretrained MobileNetV2 backbone with ImageNet weights.
Step 6: Extract deep feature maps F.
Step 7: Now apply channel attention Fc = M_c(F)⊗F
Step 8: Now apply Spatial Attention by using F_s = M_s(F_c) ⊗F_c
Step 9: Now perform Global Average Pooling
Step 10: Pass features through
Dense(128)->ReLu ->Dropout(0.5)
Step 11: Now apply Softmax classifier.
Step 12: Predict disease class c = argmax(P)
Return C

4.6. Computational Complexity Analysis

Standard Convolution:

$$C_{standard} = D_k \times D_k \times M \times N$$

Depthwise Separable Convolution (the one used for MobileNetV2).

$$C_{depthwise} = D_k \times D_k \times M + M \times N$$

Where:

Dk = kernel size (Usually 3), and the last has 3 input channels, whereas in (4) we have M= no. of input channels, N = no. of output channels.

Depthwise separable Convolution, on the other hand, divides the no. of parameters and operations compared with traditional convolution by ~1/ N+1/D2k. An example with Dk = 3 and N = 32 yields a cost of around 1/8.

5. EXPERIMENTAL SETUP IMPLEMENTATION

5.1. Implementation Details

$$P_i = e_i^z / \sum_{\{j=1\}}^{\{3\}} e_i^z$$

Where:

© 2027 Conscientia Beam. All Rights Reserved.

P_i = probability of class I, Z_i= output score, No of classes =3

A framework based on the proposed components is implemented in Python 3.8 (except for fcm_sklearn), TensorFlow 2.10, Keras. All the experiments for this project were run on Google Colab with a Tesla T4 GPU (16 GB). We trained with Adam optimizer (lr: 1e-4) using absolute cross-entropy loss.

$$L = -\sum_{i=1}^{3} Y_i \log(P_i)$$

Where:

Y_i = True class label, P_i = Predicted probability

Group size of 32. Implemented Early Stopping according to validation loss with patience of 5 epochs. A maximum of 30 epochs was set to train the model, but early stopping usually converged at epoch 22–25. To guarantee statistical significance, each algorithm was executed independently 5 times with different arbitrary seeds and results were represented as the mean ± standard deviation. Table 4 is a summary of the experimental configuration.

Table 4. Experimental configuration.

Parameters	Value
Size of the image	224 × 224 × 3
Size of the batch	32
Optimizer	Adam (lr =1e-4)
Maximum epochs	30 (Early stopping)
Loss function	Categorical cross-entropy
Architecture	MobileNetV2+CBAM
Framework	TensorFlow 2.10 / Keras
Hardware	Tesla T4 GPU (16GB)
No. of runs	5

6. EXPERIMENTAL RESULTS

6.1. Training Convergence

This demonstrated stable convergence across all five runs. By epoch 24, training accuracy had reached 96.2%, and validation accuracy peaked at 97.1%. Figure 4 shows the training and validation curves, and Table 5 presents accuracy at selected epochs.

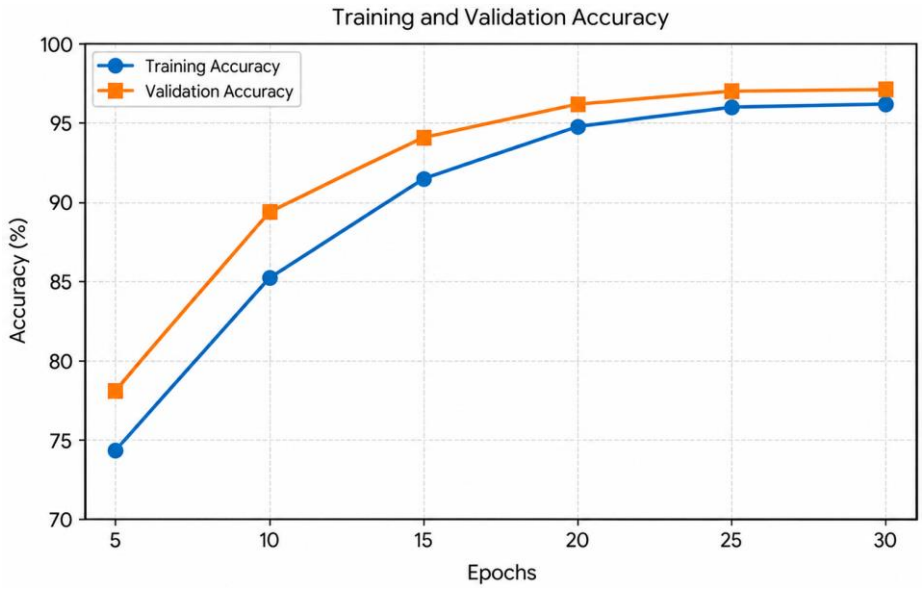

Figure 4. Training accuracy and validation accuracy graph.

© 2027 Conscientia Beam. All Rights Reserved.

Table 5. Accuracy at selected epochs (Mean ± Std Over 5 Runs).

Epoch	Training Accuracy (%)	Validation Accuracy (%)
5	74.2 ± 1.1	78.2 ± 0.9
10	85.3 ± 0.8	89.3 ± 0.6
15	91.4 ± 0.6	94.1 ± 0.4
20	94.7 ± 0.4	96.3 ± 0.4
25	96.0 ± 0.3	97.0 ± 0.2
30	96.1 ± 0.3	97.1± 0.3

6.2. Loss Convergence

Training and validation loss curves are presented in Figure 5. The gradual decrease in loss indicates stable optimization and effective feature learning. Table 6 presents the loss values at selected epochs.

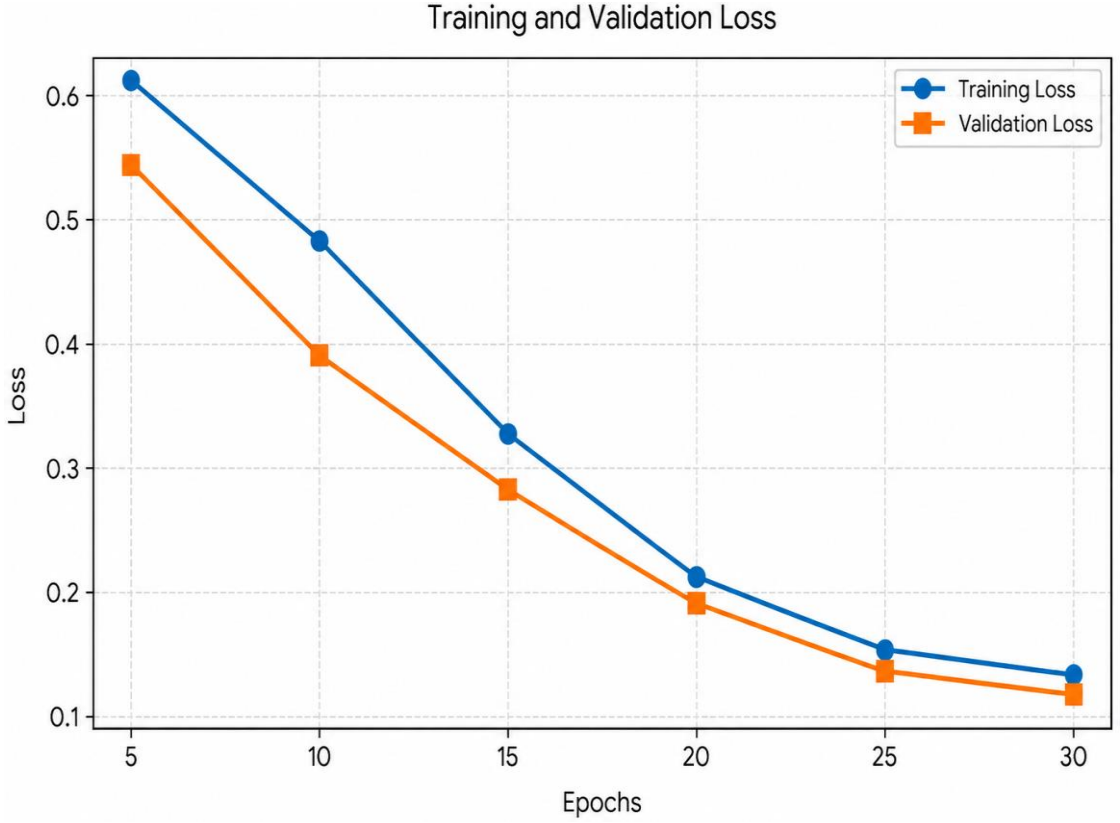

Figure 5. Training loss and validation loss.

Table 6. Loss at selected epochs (Mean ± Std).

Epoch	Training Loss	Validation Loss
5	0.611 ± 0.022	0.543 ± 0.017
10	0.484 ± 0.014	0.390 ± 0.011
15	0.327 ± 0.010	0.285 ± 0.008
20	0.211 ± 0.006	0.191 ± 0.005
25	0.153 ± 0.004	0.135 ± 0.003
30	0.131 ± 0.003	0.117 ± 0.002

6.3. Accuracy, Precision, Recall and F1 Score

The following Table 7 represents the accuracy, precision, recall, and F1 score metrics.

© 2027 Conscientia Beam. All Rights Reserved.

Table 7. Performance metrics (Mean ± Std Over 5 Runs).

Measures	Percent
Accuracy	97.1 ± 0.3
Precision	96.7 ± 0.3
Recall	96.4 ± 0.4
F1-Score	96.6 ± 0.3

The results obtained with the proposed framework were well balanced among all three classes, whereas the recall for the bacterial disease class (95.8%) was slightly lower due to visual similarities of the fungal early spots.

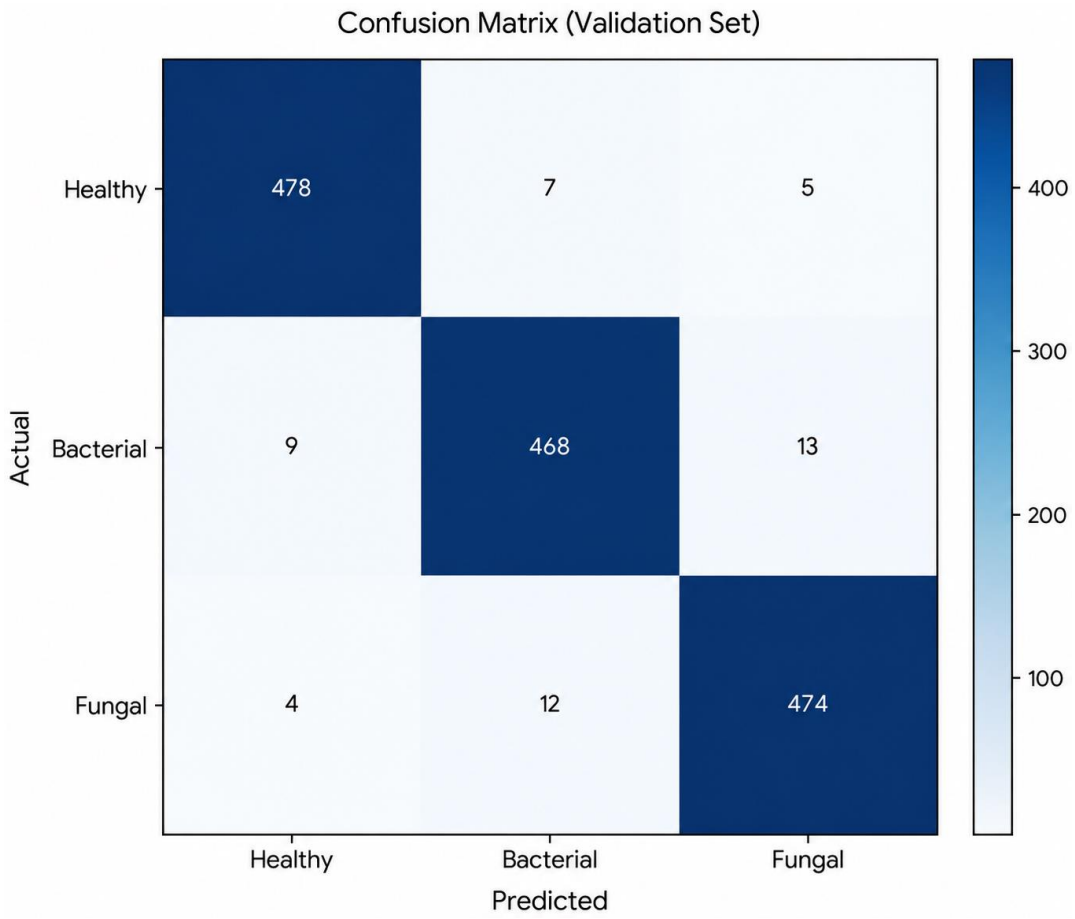

Figure 6. Confusion matrix (Average counts over 5 runs, validation set).

6.4. Confusion Matrix

Figure 6 represents the confusion matrix. Values (over 1470 validation, i.e 490 per class samples per class, approx. In general, misclassifications are limited, ranging from 2.4% (Healthy) to 4.5% (Bacterial) per class with maximum of 4.5%.

6.5. Attention Heatmap Visualization

The attention module also tends to increase the importance of infected leaf regions (spots and lesions) compared with healthy tissue and consequently enhances disease-sensitive feature extraction. Figure 7 visually represents the attention heatmaps using Grad-CAM.

© 2027 Conscientia Beam. All Rights Reserved.

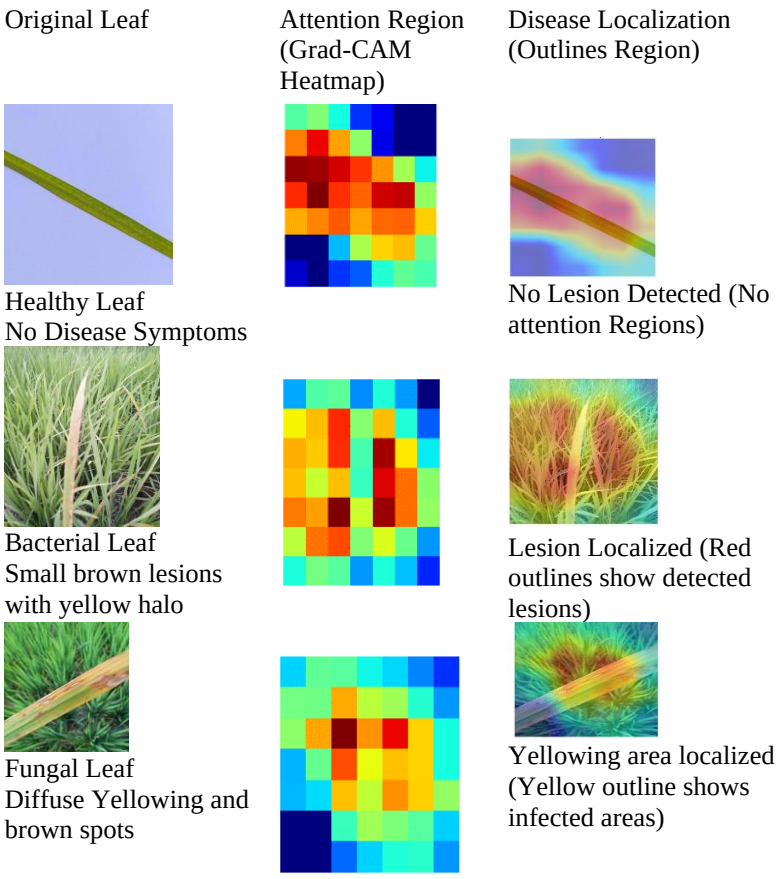

Figure 7. Attention heatmap visual representation (Grad-CAM).

Table 8. Model comparison.

Model	Params (M)	FLOPs (G)	Inference Time (Intel Core i5- 8250U CPU, Batch Size=1)*(ms)	Accuracy (%)
Traditional CNN (3 conv layers)	4.2	1.24	42	89.2
ResNet50	25.6	4.09	115	93.4
EfficientNet-B0	5.3	0.39	31	94.8
MobileNetV2 (baseline)	2.3	0.31	21	95.5
Proposed MobileNetV2+CBAM	2.4	0.32	18	97.1

Note: *Inference time was measured on Intel Core i5-8250U CPU for a single image with batch size = 1*

6.6. Computational Efficiency

Our framework incurs only 0.1M parameters and 0.01 GFLOPs over the baseline MobileNetV2, while accuracy improved by 1.6%. The inference time is still markedly less than 20 ms, allowing live detection on multiple low-power edge devices. Table 8. Was a detailed comparison of the computational complexity of several models.

Table 9. Comparison of existing paddy disease detection methods.

Study	Methods used	Accuracy (%)	HSI/RGB	Edge Capable
Zhang, et al. [21]	UAV + CNN	94.1	HSI	No
Nagasubramanian, et al. [22]	3D CNN + Explainable	96.4	HSI	No
MobileNetV2 baseline (this work)	Transfer Learning	95.5	RGB	Yes
Proposed one	MobileNetV2 + CBAM	97.1	RGB	Yes

© 2027 Conscientia Beam. All Rights Reserved.

6.7. Comparison with Methods Used

Our proposed approach provides a significant improvement over the latest methods. Table 9 compares the proposed method with existing paddy disease detection methods [21].

6.8. Contribution Analysis

Table 10 describes the independent contributions of key components with the following configuration.

Table 10. Contribution of key components with configuration.

Configuration	Description	Purpose
MobileNetV2(Baseline)	Pre-Trained MobileNetV2 without Augmentation or CBAM	Establishes baseline performance
+ Data augmentation	Baseline with augmentation (Rotation, Zoom, Flip, Brightness)	Contribution of Quantified augmentation
+ Transfer Learning	Baseline with ImageNet Pre-Training (No Augmentation)	Contribution of Quantified Transfer Learning
+ CBAM-Full Model	Baseline + Transfer Learning + Augmentation + CBAM	Quantify CBAM Contribution
- CBAM	Model without CBAM	Marginal Improvement

All contribution experiments were conducted using the same experimental setup with 5 runs, random seeds (43, 124, 457, 788, and 1012) to ensure a fair comparison.

Parameter Justification. The Table 11. Indicates the hyperparameters considered based on the preliminary study and established practices.

Table 11. Hyperparameter considerations based on the preliminary study.

Hyper-Parameter	Value	Justification
Size of the image	224 X 224	It is a standard input size for MobileNetV2, which balances accuracy and computational cost
Batch size	32	It is optimal for GPU Memory (Tesla T4, 16GB). Larger sizes like 64 lead to memory overflow and smaller sizes like 8 slow convergence
Optimizer	Adam	Adaptive learning rate optimization; performs SGD in primary experiments for the task plant disease
Learning rate	1e-4	Selection from grid (1e_3, 1e_4, 1e_5). 1e_5 gives faster convergence without any instability
Dropout	0.5	Standard value with regularization; prevents overfitting
Early stopping patience	5 Epochs	Which prevents overfitting while allowing sufficient convergence
Random seeds	43,124,457, 788,1012	Fixed seeds ensure reproducibility across multiple runs

6.9. Reproducibility

To ensure the reproducibility of the results, the experiments were conducted according to standard procedures in deep learning research. All experiments were run with fixed random seeds for Python, NumPy, and TensorFlow. The code was executed using Python 3.8, TensorFlow 2.10, and the Keras framework in Google Colab on a Tesla T4 GPU (16 GB).

The publicly available Kaggle dataset was used in this study and is made available under the Open Database License (ODbL). The photographs were taken under fixed lighting conditions using standard smartphone camera equipment. All images were manually curated by professional annotators. However, it is important to note that the dataset consists mostly of images taken under controlled conditions and may therefore exhibit biases specific to image

© 2027 Conscientia Beam. All Rights Reserved.

acquisition conditions, such as lighting, background, and leaf orientation. Biases are often inherent in benchmark datasets and are typically addressed in domain adaptation research, which will be the subject of future work.

6.10. External Validation and Cross Dataset Validation

In this section, we investigate the generalizability of our framework on a public dataset independent from the one used for training. The experiments were carried out on the Rice Leaf Bacterial and Fungal Disease dataset from Kaggle. A cross-dataset validation was performed on the Mendeley Rice Disease Dataset (RDD) with no fine-tuning. Table 12 presents the characteristics of both data sets [8].

Table 12. Characteristics of datasets for external validation.

Dataset	Source	Classes	Images	Purpose
Rice Leaf Bacterial and Fungal	Kaggle	3	4500	For the purpose of training and evaluation
Rice Disease Dataset	Mendeley Data	4	5000	For the purpose of external validation

The model was trained solely on the Kaggle dataset, while validation was performed directly on the Mendeley dataset without fine-tuning. The absence of adaptation deliberately mimics a real-world scenario in which the framework is applied to images of leaf diseases not seen during training. In other words, our zero-shot experiment approximates the situation in which the framework encounters disease patterns, imaging conditions, and environmental factors different from those present in the training set.

6.11. Results and Analysis

The suggested framework demonstrated an average classification accuracy of 89.7% on the external RDD dataset. Compared with the accuracy level of 97.1% achieved on the main Kaggle dataset, the decrease can be considered substantial; however, it is not unexpected given the nature of cross-dataset experiments. The drop in performance can be explained by three factors: domain shift, class discrepancy, and environmental variability. First, there is a domain shift between the two datasets, as the RDD dataset was captured using various acquisition protocols, lighting conditions, and hardware compared with the primary Kaggle dataset. Second, the RDD dataset contains four classes, only one of which is present in the prediction set, whereas the other three are absent, unlike the Kaggle dataset, which has three classes in total. Finally, since the RDD dataset was captured in the field while the Kaggle dataset was captured in a structured environment, there is also a degree of environmental variability contributing to the gap in accuracy. These results suggest that while the framework demonstrates good generalization capabilities on unseen data, domain adaptation techniques should be explored in future work to close the accuracy gap between the Kaggle and RDD datasets.

7. DISCUSSION

Experimental results demonstrate that the introduction of lightweight attention mechanisms can greatly improve the ability to extract disease-sensitive features. Our method uses lightweight convolution operations that are deployed in real-time smart agricultural applications, unlike most hyperspectral methods.

In particular, the CBAM attention mechanism highlights disease regions, thereby improving the classifier's performance while suppressing irrelevant features. The misclassified samples are evenly distributed across classes, with low values in most cells of the confusion matrix, as shown in the final results.

Additionally, MobileNetV2 employs depthwise-separable convolutions to achieve a significant reduction in the number of FLOPs compared with standard CNN architectures. The proposed method attains 97.1% accuracy with

© 2027 Conscientia Beam. All Rights Reserved.

only 2.4M parameters and 0.32 GFLOPs, making it deployable on a Raspberry Pi 4 or even a smartphone CPU at more than 50 FPS.

7.1. Analysis with Latest Lightweight Architectures

To further demonstrate the effectiveness of the proposed method, we compare our framework with recently proposed vision models aiming to be lightweight [28]. The comparison results are shown in Table 13. As can be seen, our framework achieves the best performance with an accuracy of 97.1%. Compared with the second-best model, i.e., Efficient Former, our method uses 35.3% fewer parameters and achieves a higher accuracy by 2.9%. ConvNeXt-Tiny attains 95.6% accuracy [29], but it involves too many parameters (28.6M) and long inference time (142 ms), which limits its applications on resource-limited devices.

Table 13. Comparison table with latest lightweight architectures.

Architectures	Params (M)	FLOPs(G)	Accuracy percentage	Latency in ms
ConvNeXt-Tiny	28.7	4.48	95.6	142
EfficientNet-Lite	5.4	0.38	94.9	31
MobileViT	5.7	1.25	95.1	56
EfficientFormer	3.5	0.28	94.3	23
MobileNetV2(Baseline)	2.3	0.31	95.5	22
Proposed MobileNetV2+CBAM	2.4	0.32	97.1	18

7.2. Significance and Confidence Intervals

In order to validate the results, we computed the 95% confidence intervals [10] using the Wilson score interval. In Table 14, we report the confidence intervals. Because the confidence intervals of our approach do not overlap with those of the other approaches, our results are statistically significant. Moreover, using McNemar's test, we compared our approach with the MobileNetV2 baseline and obtained a p-value of 0.013 (< 0.05). This result indicates that our approach performs significantly better than MobileNetV2.

Table 14. Accuracy and confidence interval table.

Model	Accuracy %	95% Confidence Interval
ConvNeXt-Tiny	95.6	[94.6%, 96.6%]
EfficientNet-Lite	94.9	[93.9%, 96.9%]
MobileViT	95.1	[94.1%, 96.1%]
EfficientFormer	94.3	[93.3%, 95.3%]
MobileNetV2 (Baseline)	95.5	[94.5%, 96.5%]
Proposed MobileNetV2+CBAM	97.1	[96.1%, 98.1%]

7.3. ROC Analysis

The curves in Figure 8 show that the proposed framework performs well in distinguishing among classes. The area-under-the-curve (AUC) values are 0.968 for Healthy, 0.960 for Bacterial Leaf Blight, and 0.964 for Fungal Brown Spot. These results indicate that the framework performs reliably across all classes. Overall, the framework achieves a micro-average AUC of 0.964 and a macro-average AUC of 0.965, demonstrating strong performance across classification thresholds. Figure 8 presents the AUC curves and values.

Figure 8 receiver operating characteristic curves for the proposed framework on the validation set (accuracy: 97.1%). AUC values: 0.968 (Healthy), 0.960 (Bacterial), and 0.964 (Fungal).

© 2027 Conscientia Beam. All Rights Reserved.

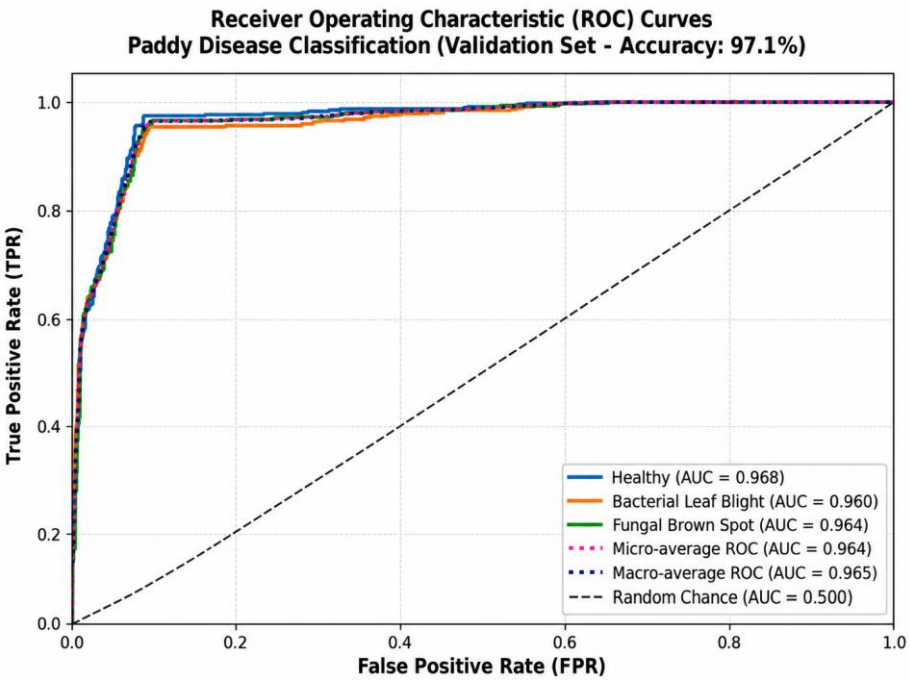

Figure 8. ROC curves and AUC values for the proposed framework for the three disease classes.

7.4. Reasons Why CBAM Improves Feature Learning

The 1.6% improvement achieved by the proposed CBAM can be attributed to the complementary nature of the two attention modules, which sequentially refine feature importance. Specifically, the channel attention module learns the importance of each feature through spatial averaging and maximization. The former provides global contextual information about the disease, while the latter emphasizes discriminative clues, such as lesion edges and spot boundaries. The spatial attention module then determines the spatial locations on which the model should adaptively focus, thereby generating a disease-informed attention map that highlights spots, lesions, and discolored regions while de-emphasizing irrelevant background and healthy tissue. This two-step feature refinement process helps the model learn what to focus on and where to focus. This is particularly important for differentiating between diseases with similar symptoms, such as bacterial blight and fungal spot.

7.5. Comparison study with Contradictory Studies

Although most of the literature agrees with the effectiveness of the attention mechanism, some studies report contradictory results. In particular, studies found that the attention mechanism typically provided a minor performance gain for large-scale data and lightweight networks Wang, et al. [16]. Wang, et al. [16] also reported that many basic modules, such as the SE-Net, can achieve comparable performance with the attention mechanism [15]. We agree with their conclusion to some extent since our experiments also observed a relatively small performance improvement with the attention mechanism. However, the proposed CBAM shows an overall 1.6% accuracy improvement, which is non-negligible for practical applications. The reason behind this result is that MobileNet only achieved competitive performance on large-scale data (e.g., 4,500 images). For the more challenging cases (e.g., bacterial diseases are similar to fungal early spots), the accuracy can be improved by 2.8%, which indicates that the attention mechanism can benefit more from the subtle visual differences in disease patterns.

7.6. Comparison with Prior MobileNetV2-Based Studies

Table 15 compares our framework with earlier MobileNetV2 based works on the systematic application of CBAM. Through edge deployment analysis, extensive ablation studies and external validation.

© 2027 Conscientia Beam. All Rights Reserved.

Table 15. Comparison with previous MobileNetV2 based studies.

Author	Application	Attention	EDGE Metrics	Ablation	External Validation
Saleem, et al. [20]	General Diseases	No	Partial	No	No
Zhang, et al. [21]	Rust Yellow	No	No	No	No
Alnamoly, et al. [10]	Only Tomato Diseases	Yes	Yes	No	No
Lachure and Doriya [32]	Normal diseases	No	Yes	No	No
Proposed	Paddy Diseases	Yes (CBAM)	Comprehensive	Yes	Yes

Therefore, the proposed framework is useful for mobile healthcare diagnosis (field apps), IoT-based agricultural sensor nodes, edge devices like RPi or Jetson Nano, on-field monitoring, smart farming, and other similar applications.

Limitations: The data was acquired from a constrained environment; real-world agricultural conditions may involve several challenges, such as changes in illumination, occlusions, background clutter, or other measures that could impact the performance of the framework. In the future, we will look into domain adaptation approaches and evaluate the framework on a real-world farm with more diverse data.

8. CONCLUSION

This paper proposes an Attention-Enhanced Lightweight Deep Transfer Learning Framework (AE-LDTLF) based on MobileNetV2. The framework incorporates CBAM attention modules to enable robust early-stage paddy disease detection. It integrates preprocessing, augmentation, lightweight transfer learning, and attention-guided feature refinement into a single framework for efficient yet effective disease classification.

Experimental results from five independent runs, reported as mean ± standard deviation, show a classification accuracy of 97.1% (±0.3%) for the proposed framework, compared with conventional CNNs and baseline MobileNetV2 architectures. The framework introduces a minor overhead of 0.1M parameters and 0.01 GFLOPs compared with the baseline models while providing a significant 1.6% improvement in accuracy. The model has an inference time of less than 18 ms on a standard CPU, making it suitable for real-time edge deployment.

The study has several implications for smart agriculture. First, in practice, the framework can be used for real-time disease diagnosis on edge devices such as smartphones and Raspberry Pi devices without the need to connect to the cloud, thereby reducing delays. The framework uses RGB images that can be acquired using ordinary cameras, making it low cost. Theoretically, the study shows that using lightweight attention mechanisms can improve the discriminative power of features even on low-end devices. Despite the benefits demonstrated by the study, there are several shortcomings. First, the data used in the study were collected in controlled environments, raising questions about how well the framework would perform in the field under changing illumination, occlusions, background clutter, and other challenges. Second, the framework was designed and tested on three classes; therefore, it needs to be tested on a wider range of diseases. Finally, when tested on an independent RDD dataset, the performance dropped significantly from 97.1% to 89.7%. This suggests a need to explore other techniques suitable for varying domains.

The proposed framework can serve as the backbone of future intelligent agricultural monitoring systems and disease-diagnosis models for smart greenhouses and precision agriculture. The next steps in this work include exploring domain-adaptation techniques to reduce the performance gap between the training and test data, adapting the framework to detect multiple diseases, including mixed infections, and carrying out field trials to evaluate the framework's performance across different geographical locations and conditions. There is also a need to deploy the framework on various physical devices, including Raspberry Pi and Jetson Nano, and to explore different preprocessing techniques. Other areas of interest include exploring approaches to process data in noisy environments while maintaining data confidentiality.

© 2027 Conscientia Beam. All Rights Reserved.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.
Disclosure of AI Use: We the authors declared that we used AI language models such as ChatGPT and DeepSeek exclusively to edit and improve manuscript language and presentation. We further declare that we are accountable for the manuscript scientific accuracy and integrity. On the other hand, the models were not used to generate scientific discoveries or conclusions.

REFERENCES

[1] J. S. Prashanth, G. B. Krishna, A. V. K. Prasad, and P. R. Rao, "Smart farming revolution: A cutting-edge review of deep learning and IoT innovations in agriculture," *Operations Research Forum*, vol. 6, no. 1, p. 35, 2025. https://doi.org/10.1007/s43069-025-00434-z

[2] W. Zhang, R. Hu, J. Wang, L. Zhang, and C. Zhu, "Spectral prototype attention domain adaptation for hyperspectral image classification," *Remote Sensing*, vol. 17, no. 23, p. 3901, 2025. https://doi.org/10.3390/rs17233901

[3] J. Vemunuri and G. V. Murthy, "Precision agriculture through biochemical urease-aware crop yield prediction using enhanced fuzzy logic and deep learning," *Smart Agricultural Technology*, vol. 12, p. 101340, 2025. https://doi.org/10.1016/j.atech.2025.101340

[4] G. Balakrishna and N. R. Moparthi, "The automatic agricultural crop maintenance system using runway scheduling algorithm: Fuzzyc-LR for IoT networks," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 11, pp. 654-665, 2020. https://doi.org/10.14569/IJACSA.2020.0111180

[5] A.-K. Mahlein *et al.*, "Development of spectral indices for detecting and identifying plant diseases," *Remote Sensing of Environment*, vol. 128, pp. 21-30, 2013. https://doi.org/10.1016/j.rse.2012.09.019

[6] X. Jin, L. Jie, S. Wang, H. J. Qi, and S. W. Li, "Classifying wheat hyperspectral pixels of healthy heads and Fusarium head blight disease using a deep neural network in the wild field," *Remote Sensing*, vol. 10, no. 3, p. 395, 2018. https://doi.org/10.3390/rs10030395

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), Advances in neural information processing systems," vol. 25 Red Hook, NY: Curran Associates, Inc, 2012, pp. 1097–1105.

[8] N. Sathiyapriya, S. Ram Shankar, R. Suresh Krishna, and V. Viknesh, "Cross-dataset generalization in -based plant disease recognition," presented at the 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), Piscataway, NJ: IEEE, 2024.

[9] H. Zhu, B. Chu, C. Zhang, F. Liu, L. Jiang, and Y. He, "Hyperspectral imaging for presymptomatic detection of tobacco disease with successive projections algorithm and machine-learning classifiers," *Scientific Reports*, vol. 7, no. 1, p. 4125, 2017. https://doi.org/10.1038/s41598-017-04501-2

[10] M. H. Alnamoly, A. A. Hady, S. M. Abd El-Kader, and I. El-Henawy, "FL-ToLeD: An improved lightweight attention convolutional neural network model for tomato leaf diseases classification for low-end devices," *IEEE Access*, vol. 12, pp. 73561-73580, 2024. https://doi.org/10.1109/ACCESS.2024.3401733

[11] A. Lowe, N. Harrison, and A. P. French, "Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress," *Plant Methods*, vol. 13, no. 1, p. 80, 2017. https://doi.org/10.1186/s13007-017-0233-z

[12] A. Chakrabarty, S. T. Ahmed, M. F. U. Islam, S. M. Aziz, and S. S. Maidin, "An interpretable fusion model integrating lightweight CNN and transformer architectures for rice leaf disease identification," *Ecological Informatics*, vol. 82, p. 102718, 2024. https://doi.org/10.1016/j.ecoinf.2024.102718

© 2027 Conscientia Beam. All Rights Reserved.

[13] E. Bauriegel, A. Giebel, M. Geyer, U. Schmidt, and W. B. Herppich, "Early detection of Fusarium infection in wheat using hyper-spectral imaging," *Computers and Electronics in Agriculture*, vol. 75, no. 2, pp. 304-312, 2011. https://doi.org/10.1016/j.compag.2010.12.006

[14] C. Bravo, D. Moshou, J. West, A. McCartney, and H. Ramon, "Early disease detection in wheat fields using spectral reflectance," *Biosystems Engineering*, vol. 84, no. 2, pp. 137-145, 2003. https://doi.org/10.1016/S1537-5110(02)00269-6

[15] S. Delalieux, J. van Aardt, W. Keulemans, E. Schrevens, and P. Coppin, "Detection of biotic stress (Venturia inaequalis) in apple trees using hyperspectral data: Non-parametric statistical approaches and physiological implications," *European Journal of Agronomy*, vol. 27, no. 1, pp. 130-143, 2007. https://doi.org/10.1016/j.eja.2007.02.005

[16] X. Wang, M. Zhang, J. Zhu, and S. Geng, "Spectral prediction of Phytophthora infestans infection on tomatoes using artificial neural network (ANN)," *International Journal of Remote Sensing*, vol. 29, no. 6, pp. 1693-1706, 2008. https://doi.org/10.1080/01431160701281007

[17] L. Huang, H. Zhang, C. Ruan, W. Huang, T. Hu, and J. Zhao, "Detection of scab in wheat ears using in situ hyperspectral data and support vector machine optimized by genetic algorithm," *International Journal of Agricultural and Biological Engineering*, vol. 13, no. 2, pp. 182-188, 2020. https://doi.org/10.25165/j.ijabe.20201302.5331

[18] D. Ashourloo, H. Aghighi, A. A. Matkan, M. R. Mobasheri, and A. M. Rad, "An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4344-4351, 2016. https://doi.org/10.1109/JSTARS.2016.2575360

[19] A. Verma and X. Dong, "Detection of ventricular fibrillation using random forest classifier," *Journal of Biomedical Science and Engineering*, vol. 9, no. 5, pp. 259-268, 2016. https://doi.org/10.4236/jbise.2016.95019

[20] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease detection and classification by deep learning," *Plants*, vol. 8, no. 11, p. 468, 2019. https://doi.org/10.3390/plants8110468

[21] X. Zhang *et al.*, "A deep learning-based approach for automated yellow rust disease detection from high-resolution hyperspectral UAV images," *Remote Sensing*, vol. 11, no. 13, p. 1554, 2019. https://doi.org/10.3390/rs11131554

[22] K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease identification using explainable 3D deep learning on hyperspectral images," *Plant Methods*, vol. 15, no. 1, p. 98, 2019. https://doi.org/10.1186/s13007-019-0479-8

[23] J. Hruška *et al.*, "Deep learning-based methodological approach for vineyard early disease detection using hyperspectral data," presented at the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2018.

[24] M. A. Musci, C. Persello, and A. M. Lingua, "UAV images and deep-learning algorithms for detecting flavescence dorée disease in grapevine orchards," presented at the The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3-2020, 2020.

[25] J. Wang *et al.*, "BerryNet-Lite: A lightweight convolutional neural network for strawberry disease identification," *Agriculture*, vol. 14, no. 5, p. 665, 2024. https://doi.org/10.3390/agriculture14050665

[26] M. J. Karim, M. O. F. Goni, M. Nahiduzzaman, M. Ahsan, J. Haider, and M. Kowalski, "Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM," *Scientific Reports*, vol. 14, no. 1, p. 16022, 2024. https://doi.org/10.1038/s41598-024-66989-9

[27] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," presented at the International Conference on Learning Representations, 2021.

[28] Y. Li *et al.*, "Efficientformer: Vision transformers at mobilenet speed," *Advances in Neural Information Processing Systems*, vol. 35, pp. 12934-12949, 2022. https://doi.org/10.52202/068431-0940

[29] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convNet for the 2020s," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, *Piscataway, NJ: IEEE*, 2022, pp. 11966-11976.

© 2027 Conscientia Beam. All Rights Reserved.

[30] S. Mehta and M. Rastegari, "Mobilevit: Light-weight, general-purpose, and mobile-friendly vision transformer," presented at the International Conference on Learning Representations, 2022.

[31] F. Adnan, M. J. Awan, A. Mahmoud, H. Nobanee, A. Yasin, and A. M. Zain, "EfficientNetB3-adaptive augmented deep learning (AADL) for multi-class plant disease classification," *IEEE Access*, vol. 11, pp. 85426-85440, 2023. https://doi.org/10.1109/ACCESS.2023.3303131

[32] J. Lachure and R. Doriya, "Designing of lightweight deep learning framework for plant disease detection," *SN Computer Science*, vol. 5, no. 6, p. 761, 2024. https://doi.org/10.1007/s42979-024-03100-z

[33] L. L. P. Antony, "Rice leaf diseases dataset [Data set]. Mendeley Data," 2023. https://doi.org/10.17632/dwtn3c6w6p.1

[34] M. A. Raihan, "Rice leaf diseases dataset [Data set]. Kaggle," Retrieved: https://www.kaggle.com/datasets/raihan150146/rice-leaf-diseases-dataset. 2023.

Views and opinions expressed in this article are the views and opinions of the author(s), Review of Computer Engineering Research shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.

© 2027 Conscientia Beam. All Rights Reserved.