Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Convolutional Fusion Architecture for Processing Smart Phone based Images to Detect Banana Leaf Diseases

[1*]Jonalee Barman Kakati, [1]Tapashi Kashyap Das
[1]*Department of Computer Science, Krishna Kanta Handiqui State Open University, Khanapara, Guwahati - 781022, Assam, India*
bkjonalee.research@gmail.com, tapashi.kashyap@gmail.com

ARTICLE INFO

Received: 10 Nov 2024

Revised: 25 Dec 2024

Accepted: 22 Jan 2025

ABSTRACT

Banana diseases cause significant loss to the cultivators and it is essential to automate the diagnostic process to prevent inconsistencies and delays in disease detection by human experts. This study proposes a convolutional fusion architecture called RAttnNet that integrates separable convolution in traditional convolution with attention mechanism within a unified architecture to enhance feature extraction while utilizing support vector machine (SVM) for better classification accuracy. Although in this field research is going on, most of the earlier works relied on publicly available datasets and computational complexity of these models presents opportunity for improvement in both accuracy and computational costs. A dataset of banana leaves is created for this study by capturing banana leaves images, with smartphones, from different parts of Assam in India. The performance of RAttnNet is evaluated with seven other state-of-the-art models and it revealed that RAttnNet outperformed others achieving an impressive accuracy of 99%. This study showcases the potential of deep learning tools in agriculture and encourages further research and application in this domain.

Keywords: Agriculture, banana, convolutional fusion, plant disease, residual network

1. INTRODUCTION

The state of Assam, in the North-Eastern region of India, is primarily an agrarian state with around 70% of its people directly dependent on agriculture. Assam is one of the major banana producing state and accounts for about 2.4% total production in India. However, banana yield suffers a lot because of disease occurrences and there can be loss even up to 100%. Banana planters who grow bananas for commercial purpose suffer a heavy economic loss due to the diseases. Apart from reduction in the yield, diseases also cause a substantial reduction in the quality of the fruit. Early detection of the diseased banana plants is crucial in order to implement appropriate preventive measures. To ensure overall health and vitality of the plant, early diagnosis and classification of diseases is necessary to successfully treat them. For the purpose of this work, we have considered two most dreaded fungal diseases of banana, Yellow Sigatoka and Panama. The banana diseases are identified traditionally by human experts, which is a time-consuming process. Availability of plant pathologist or agronomist in the rural areas is another difficulty in identification of the diseases. Visual evaluation is susceptible to intellectual and cognitive experience and may also lead to biases.

Banana diseases were recognized through color, texture and shape information using SVM classification techniques [1]. An average accuracy of 85% was demonstrated to classify the four diseases namely, Sigatoka, Bacterial wilt, CMV and Panama. The SVM classification achieved accuracy of 84%, 86%, 85% and 85% for four diseases. In a study [2], KNN was used as a classifier to classify banana diseases from its leaves achieving an accuracy of 96.87%. Healthy and unhealthy banana leaves were classified in [3] using a hybrid model, SVM infused CNN, achieving an accuracy of 92.8%. Classification of banana plant diseases, healthy-black sigatoka, and healthy-cordana leaf spot using SVM and KNN algorithms are discussed in [4]. The texture feature extraction was based on LBP (local binary pattern) achieving an accuracy of 89.1% and 90.9%. SVM, artificial neural network and minimum distance algorithms are used in [5] to monitor yellow sigatoka in banana. SVM achieved the best performance of 99.28% overall accuracy. Singh et.al. [6] used

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Support Vector Machine (SVM) algorithm on banana, beans, lemon and rose leaf diseases and reported an overall accuracy of 95.71%. Mrunalini R. et.al [7] proposed an approach using K-Means clustering algorithm and neural network for automatic detection of different plant diseases. In a separate study [8] focused on plant disease detection and classification, a methodology involving image segmentation was proposed utilizing K-means clustering to identify the infected portions of the leaves. Feature extraction was performed with GLCM (Gray-level Co-occurrence Matrix) and Random Forest was used as a classifier to obtain an accuracy of 98%. Banana bunchy top, xanthomonas wilt of banana, healthy banana and individual banana plants were classified in a study [9] using random forest algorithm. An accuracy of 99.4%, 92.8%, 93.3% and 90.8% was achieved for the four classes in the study. A lightweight CNN model, Banana Squeeze Net, was proposed by Bhuiyan et.al. [10] for the diagnosis of three banana leaf diseases. The model demonstrated an accuracy of 96.25%. A CNN model using hybrid segmentation called total generalized variation fuzzy C means (TGVFCMS) was proposed in [11] to classify five classes of banana leaf diseases achieving an accuracy of 93.45%. Amara et.al. [12] proposed an approach to classify the banana diseases, black sigatoka and black speckle, with the help of CNN LeNet architecture on the Plant Village dataset. A learning rate of 0.001 was used with SGD optimizer algorithm on both color and grayscale images. Experiments were conducted with different training and test sets having 80-20, 40-60, 50-50, 60-40 and 20-80 ratios. Accuracy of 99% (50-50) in color and 97% in grayscale domain (40-60) was reported. An optimized Capsule Network model (CapsNet) was proposed in [13] to classify banana leaf diseases. The model classified the diseases, banana bacterial wilt, black sigatoka and healthy leaves with an accuracy of 95%. In an attempt to detect Fusarium wilt in banana crop, the authors in their study [14], built a CNN using transfer learning approach with ResNet50 as the base model. The model was able to achieve an accuracy of 98%.

2. RESEARCH GAPS AND MOTIVATION

Though various methods have been proposed, there are still notable challenges to address in detecting the banana leaf diseases.

- One of the primary challenges in the classification of banana leaf diseases is that it is difficult to acquire significant features using traditional image processing techniques. The performance of machine learning models is largely influenced by hand-made features chosen manually.

- Much of the previous work in deep learning has used computationally expensive models creating challenges to develop algorithms suitable for resource-constrained devices for real-time detection of healthy and diseased banana leaves. Computationally efficient lightweight models, in terms of model size, would be more useful for practical deployment in the real field. Also, the traditional feed-forward CNN models suffer from vanishing gradient problem with the increase in the number of layers in the network.

- The images used for the studies were captured by good quality digital cameras and phones in a controlled environment. In real-time scenarios, the camera used for capturing the images by farmers may be of different resolutions and the images may be under various luminous conditions. The classification model should be able to handle banana leaf images with background clutters, different leaf orientations, variant image quality and also the lighting conditions under which the images have been taken.

- The time-consuming process of diagnosing diseases in banana plants and its susceptibility to intellectual and cognitive experience of the expert justifies the need for a system that could automate classification.

2.1 Research Contribution

The proposed study employed a deep learning architecture to address the identified research gaps in the detection of diseases in banana leaves. The main contributions of our study are:

- ✓ This research proposes a lightweight deep learning architecture that is capable of accommodating variations in image quality, orientations, and diverse luminous levels. The main rationale for building this model is its smaller size, which entails fewer parameters and lower computational requirements

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

as compared to existing models. The research contribution from this study is the achievement of better accuracy with added benefit of reduced computational resource usage compared to the previous approaches.

- ✓ A banana leaves dataset is constructed by capturing the banana leaves under varied lighting conditions, orientations, weather variations and at different times of the day. The images are collected using mobile phones having different resolutions.
- ✓ A comparison of the result of the model in terms of accuracy and other performance matrices is conducted with Xception, InceptionNet V3, InceptionResNetV2, EfficientNet-B0, EfficientNet-B1, EfficientNet-B2 and EfficientNet-B3.

3. MATERIALS AND METHODS

The materials and methods used in this work are discussed in this section.

3.1 DataSet

We have selected banana plants of Assam in India for the purpose of our study. To achieve the research objective of this study, healthy and diseased images of the banana leaves are collected from Kamrup, Chhaygaon, Boko, Palashbari, Nagarbera, Rangia, Nalbari, Kamalpur, Tezpur, Golaghat and Jorhat areas of Assam. During different time periods of the day, the images have been captured under different weather conditions, angles and orientations and under natural settings. The dataset consists of 3 classes namely, Healthy, Sigatoka and Panama. Total number of banana leaves sample in the dataset is 3425 representing real environmental images from fields. A sample of the images in the dataset that is prepared for the purpose of this study is shown in the table 1. The sample distribution in the dataset is shown in the figure 1.

Table 1: Images from the dataset

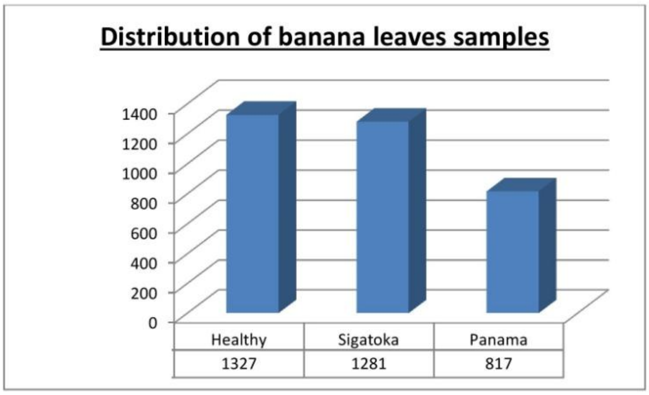


Figure 1: Dataset image distribution

3.2 Image preprocessing

The conventional cropping method is used to remove the irrelevant information as much as possible. All the images are in the RGB color and .jpeg format. Images captured by the smart phones had different resolutions and sizes. To ensure compatibility, all images are normalized and resized. To increase the number of similar images in the dataset, augmentation has been performed. Algorithm for augmentation (Image data generator) is as below.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Input : Images of banana leaf from dataset
Output : Augmented images
Step 1. Rotate images in range = 40
Step 2. Images width shift range = 0.2
Step 3. Do image height shift range = 0.2
Step 4. Do image shear range = 0.2
Step 5. Zoom range of image = 0.2
Step 6. Do horizontal flip = True
Step 7. Do fill-mode = 'nearest'

3.3 Proposed Method

3.3.1 Architecture of the proposed model:

The proposed work is to create a cost-effective and precise learning model for the detection of healthy and diseased banana leaf. The existing CNN based models for identifying banana diseases rely on the traditional feed-forward structure to extract features for detection of banana leaf diseases. Such models, characterized by the numerous deep layers, frequently face the problem of vanishing gradient, which hinders the model's feature-learning capability. Moreover, they are computationally expensive and require significant resources for training and inference. Our proposed work introduces a novel deep learning architecture called RAttnNet specifically tailored for the detection of diseases in banana leaves. This model fuses convolution techniques with convolutional attention mechanism (CBAM) and separable convolution within the same architecture to enhance feature extraction, and SVM is used for classification accuracy. Standard convolution is used in the initial layers, and finally separable convolution is used for subsequent layers. The proposed architecture of the model is shown in the figure 2.

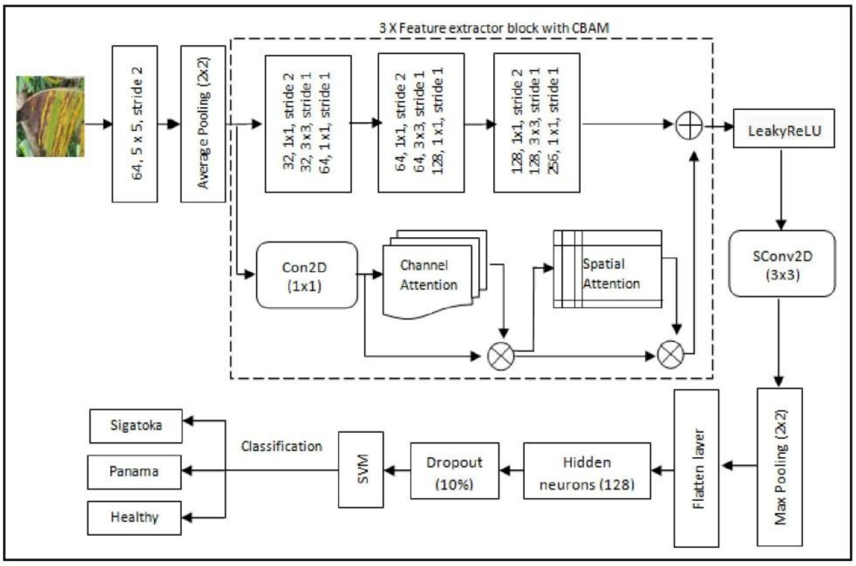


Figure 2: Architecture of proposed RAttnNet model

The motive behind the building of RAttnNet model is to build a lightweight deep neural network. Inspired by ResNet [15] and MobileNet architectures [16], RAttnNet takes advantage of residual networks with skip connections and separable convolutions to build a lightweight model. Integration of SVM with L2 regularization in the classification part of RAttnNet involves combining the concepts of CNN and machine learning to enhance the performance and robustness of the proposed model.

Convolutional layer: In RAttnNet, RGB image of size 224x224 is fed as an input to a convolutional layer having 64 filters having kernel size 5 x 5. A stride of 2 is used to downsample the feature map by a factor of 2. It also allows the kernel to cover a larger receptive field in the input feature map. This helps RAttnNet model to learn more global

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

features and capture larger patterns in the input data. For a single convolution layer, the total number of addition and multiplication operations can be calculated as

$$O_t = C_i \times H \times W \times k \times k \times C_o$$

where,
O_t = number of operations in traditional convolution
C_i = number of input channels
H = height of the input feature map
W = width of the input feature map
k = kernel size
C_o = number of output channels

Pooling : In RAttnNet, average pooling having window size 2 x 2 and a stride of 2 is applied to each of the non-overlapping regions of the input feature map, and the average value within each window is computed to obtain the corresponding value in the output feature map.

Feature extractor block: Our proposed architecture consists of three residual feature extraction blocks in sequence, each designed to learn increasingly complex features. As we move deeper into the network, the number of output channels increases from 64 to 256. The output of the pooling layer is the input to the first feature extractor block and it traverses through two functional paths in the block - a main path and a shortcut path. In the main path, the input goes through a series of standard convolutions, batch normalization and activation functions, which refine the feature representation. Through the other path, the input is passed through a 1x1 convolution, a CBAM mechanism, and batch normalization, allowing it to feed back into the network. The 1x1 convolution is applied to match the dimensions of the output from the main path. Attention mechanism is added in the shortcut path to help the network focus on important features. After processing through their respective paths, the outputs of both paths are summed up to get the final feature maps. Suppose p(x) denote the operations of the main path and q(x) represent the skip connection, then output of the feature extractor blocks can be expressed by the following equation.

$$f(x) = p(x) + q(x)$$

LeakyReLU activation function : To enhance the performance of RAttnNet in identifying banana leaf images, LeakyReLU activation function is used in this research as it can lead to faster convergence, better generalization, and improved accuracy in the image classification tasks. The slope is introduced to prevent the "dying ReLU" problem, where neurons can become stuck in a state of inactivity, resulting in dead neurons that do not contribute to the learning process. In RAttnNet, LeakyReLU with an alpha value 0.1 is used as the activation function. Mathematically, LeakyReLU activation function is defined as follows:

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ \alpha \cdot x, & \text{if } x < 0 \end{cases}$$

Here, alpha is a small positive constant. If input x is greater than or equal to zero, the function behaves like standard ReLU, producing the input value as the output. When the input x is negative, the function introduces a small negative slope by multiplying the input with the alpha constant. The small negative slope introduced by LeakyReLU helps in making the network more robust to such variations, as it allows the network to learn from both positive and negative inputs. This enhances the ability of RAttnNet to generalize and classify banana leaf images accurately, even in the presence of variations.

Residual skip connection : Residual skip connections are used for proper gradient flow. In the feature extraction block, 1 x 1 convolutions are performed on the feature maps. LeakyReLU with alpha value as 0.1 is used as the activation function. Assuming x represents the input to the feature block or a layer in the network, and f(x) represents the transformation that is applied on x by the block or layer, the mathematical expression of residual connection can be given as

$$y = f(x) + x$$

where y represents the final output of the block, which is the sum of the transformed input and the original input. In this way, the network can effectively propagate the gradients and preserve the important features information from the earlier layers. The residual connection allows the network to learn the residual mapping, f(x) instead of directly learning the complete mapping from the input. This contributes to improve the overall network performance.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Convolutional Block Attention Module (CBAM): In this research work, an attention mechanism is added in each feature block of RAttnNet model to boost the performance of our model and to improve its ability to capture disease-specific patterns in the banana leaf images. The attention module is added after 1x1 2D convolutional layers. The attention mechanism enhances the robustness of the RAttnNet model by enabling it to adaptively attend to different disease symptoms or variations in banana leaves.
Separable Convolutions: To reduce computational complexity of the model, particularly in the resource-constrained environment or on devices with limited computational power, separable convolutions are inserted in RAttnNet. This lead to faster training and the efficiency of the network improves as SConv (Separable convolutions) introduce fewer parameters and have less capacity to memorize the training data. In depthwise convolution, the total number of addition and multiplication operations for a single convolution layer, denoted by O_d can be calculated as

$$O_d = C_i \times H \times W \times k \times k$$

To combine the information from different channels, pointwise convolution is performed where 1x1 convolution is applied to the intermediate feature maps. In pointwise convolution, the total number of operations, denoted by O_p can be calculated as

$$O_p = C_i \times H \times W \times C_o$$

Therefore, the total number of operations, O_S, for a separable convolution layer can be expressed as:

$$\begin{aligned} O_S &= C_i \times H \times W \times k \times k + C_i \times H \times W \times C_o \\ &= C_i \times H \times W \times (k \times k + C_o) \end{aligned}$$

In RAttnNet, image of size 224×224 is fed into a traditional convolution layer with a kernel size of 5×5 and 64 output channels. The output feature map of this layer is processed and fed into separable convolution layer with a kernel size of 3×3 and also 64 output channels. The layer with a 5×5 kernel capture larger patterns, while the separable convolution layer with a 3×3 kernel can focus on finer details. As such, the number of operations is

$$O_t = C_i \times H \times W \times 5 \times 5 \times 64$$
$$O_S = C_i \times H \times W \times (3 \times 3 + 64)$$

Thus, there is significant number of computational reductions in the proposed model.

3.3.2 Implementation

The experiments are conducted in Google Colab with Keras and TensorFlow as its backend and having 12 GB NVIDIA Tesla K80 GPU. The dataset is randomly divided into training, validation and test sets by 80%, 10% and 10% respectively. RAttnNet model is trained using the training set and the images in it are used to update the model parameters through the Adam optimization algorithm with a learning rate of 10^{-3}. The flow diagram of RAttnNet is shown in the figure 3. The hyper-parameters for the different models are chosen empirically after a series of experiments to achieve the best classification results. Categorical cross-entropy loss function is used for evaluation with number of iterations as 30 epochs. The performance of the model is evaluated using the performance matrices such as accuracy, precision, recall, F1-score and confusion matrices. Classification report is produced to analyze the results per class.
In the next phase of our study, a comparative evaluation of the RAttnNet model is carried out with state-of-the-art models - InceptionV3, Xception, InceptionResNetV2, EfficientNet-B0, EfficientNet-B1, EfficientNet-B2 and EfficientNet-B3. The analysis of the experimented results is provided in section 4.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

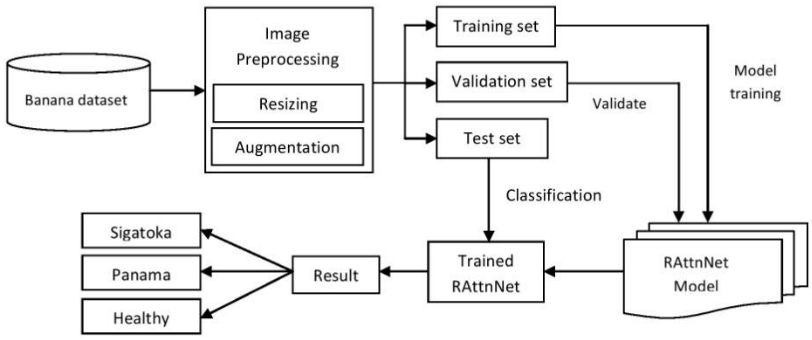


Figure 3: Flow diagram of proposed model

4. RESULTS AND DISCUSSION

4.1 Proposed methodology results

This section presents the experimental results of RAttnNet model that we obtained. Figure 4 shows the loss and accuracy graph. It can be seen from figure 4 that the accuracy initially exhibits rapid progression but after a certain number of iterations, the progress starts to stabilize and then achieves its optimum performance of 99%. The loss function initially shows an increase and after the 3rd epoch it decreases sharply. The loss function almost stabilizes after the 12th epoch and reaches minimum possible value. As in this research work, we are concerned about identifying three classes of leaves; a 3x3 dimension confusion matrix is obtained for our RAttnNet model. Figure 5 gives the confusion matrix obtained.

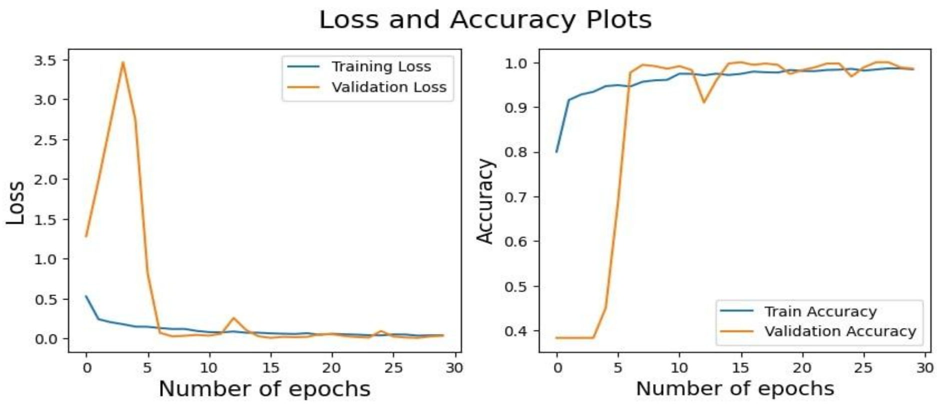


Figure 4: Loss and accuracy graph obtained from RAttnNet

The rows in matrix represent the true class labels and the columns represent the predicted class labels. The elements on the main diagonal (131, 78, 130) represents the true positive counts, indicating the number of instances that belong to the particular class and were correctly predicted as belonging to that class. The elements outside the main diagonal represents the false positive (FP) and false negative (FN) counts.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

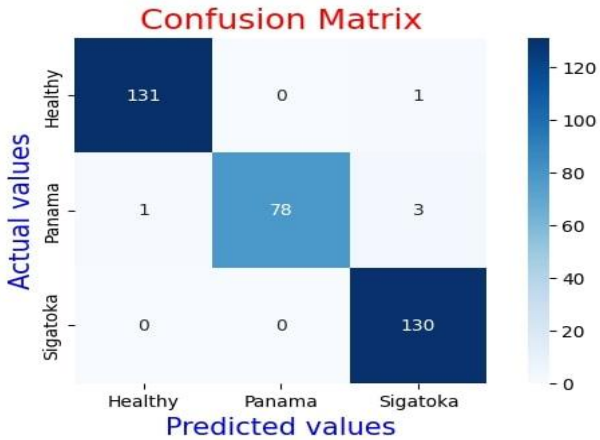


Figure 5: Confusion matrix of RAttnNet

To assess the accuracy and effectiveness of the classification model, and to gain insights into its behavior for the different classes in the dataset, a classification report is obtained. The classification report depicting the class-wise analysis of precision, recall and F1 score, for the different classes is shown in figure 6. Class 0 indicates the healthy class; 1 indicates class Panama and 2 is class Sigatoka. The F1-score offers a balanced evaluation of the model's performance, taking into account the precision and recall.

	precision	recall	f1-score
0	0.99	0.99	0.99
1	1.00	0.95	0.97
2	0.97	1.00	0.98
accuracy			0.99
macro avg	0.99	0.98	0.98
weighted avg	0.99	0.99	0.99

Figure 6: Classification Report

4.2 Comparative analysis of the selected model with other deep neural models

The performances of the well known architectures along with the RAttnNet model are presented in table 2. Transfer learning approach using Xception and InceptionResNetV2; and without transfer learning approach using EfficientNet-B0, EfficientNet-B1, EfficientNet-B2 and EfficientNet-B3 models is used for comparison with RAttnNet. These models are trained and tested on the same dataset created for this research study.

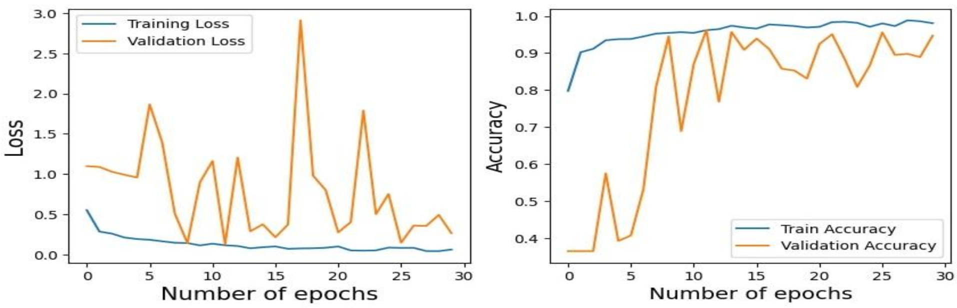
Table 2: Class-wise performance comparison with other deep learning models

Models	Class	Precision %	Recall %	F1 score	Validation Loss %	Validation Accuracy %
Xception	Healthy	100	85	92	33	91
	Panama	79	96	87		
	Sigatoka	91	95	93		
InceptionV3	Healthy	100	85	92	30	91
	Panama	83	96	89		
	Sigatoka	88	95	91		
InceptionResNetV2	Healthy	98	92	95	37	90
	Panama	100	71	83		

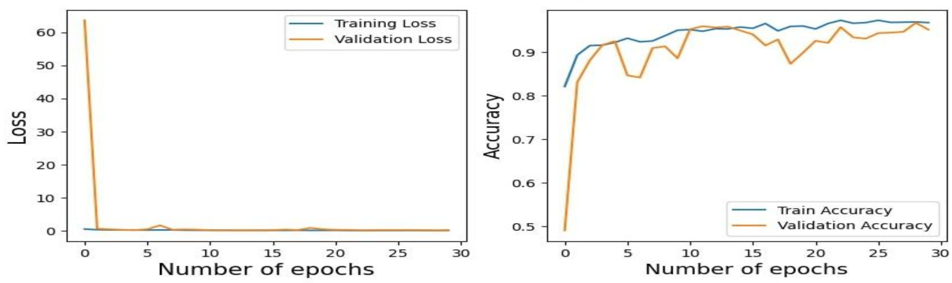
Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

	Sigatoka	79	98	88		
EfficientNet-B0	Healthy	92	97	95	27	93
	Panama	96	91	93		
	Sigatoka	92	90	91		
EfficientNet-B1	Healthy	92	97	94	26	90
	Panama	83	87	85		
	Sigatoka	92	84	88		
EfficientNet-B2	Healthy	85	85	85	30	93
	Panama	85	73	78		
	Sigatoka	83	91	87		
EfficientNet-B3	Healthy	97	98	97	18	95
	Panama	92	91	92		
	Sigatoka	94	94	94		
RAttnNet	Healthy	99	99	99	05	**99**
	Panama	100	96	97		
	Sigatoka	97	100	98		

As seen from table 2, RAttnNet outperforms all the architectures on our self created dataset of field images of banana leaves. The validation accuracy and loss of 99% and only 5% respectively reveals that RAttnNet performed far better than the others in feature extraction and classification. EfficientNet-B3 has attained the next highest validation accuracy of 95% with a validation loss of 18%. It is followed by the EfficientNet-B0 and EfficientNet-B2 models showing an impressive accuracy of 93% with validation loss of 27% and 30% respectively. The Xception and InceptionNetV3 model architectures have both demonstrated their capabilities by achieving an accuracy rate of 91% with validation loss of 33% and 30% respectively. Interestingly, both InceptionResNetV2 an EfficientNet-B1 has accomplished the same remarkable accuracy rate of 90% with loss of 37% and 26% respectively. Figure 7 and 8 presents the loss and accuracy plots. Figure 9 [a-g] shows the confusion matrices of the deep architectures. Figure 10 displays the test accuracies of the different models.



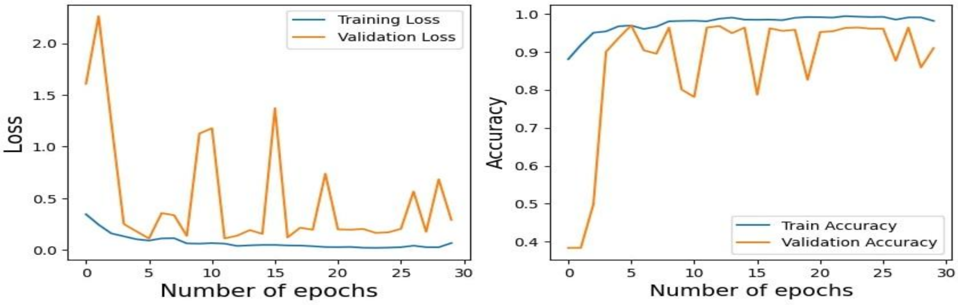
(a) Xception



(b) InceptionV3

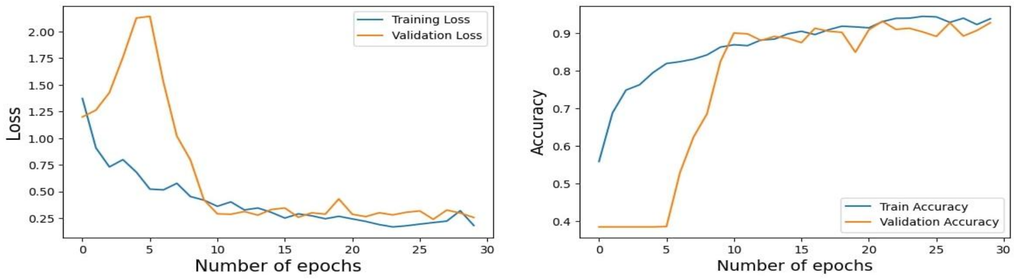
Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

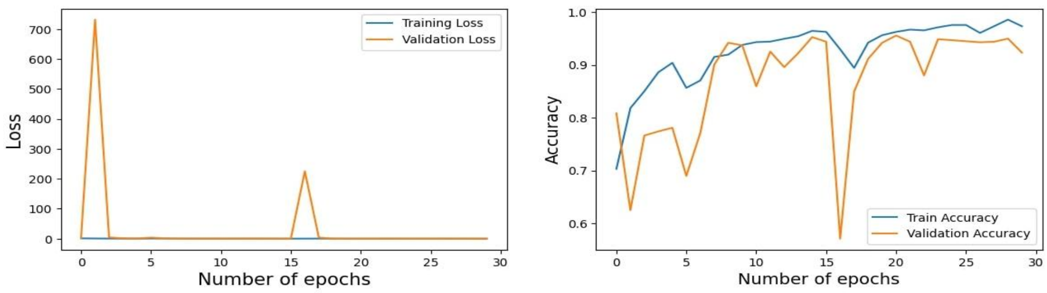


(c) InceptionResNetV2

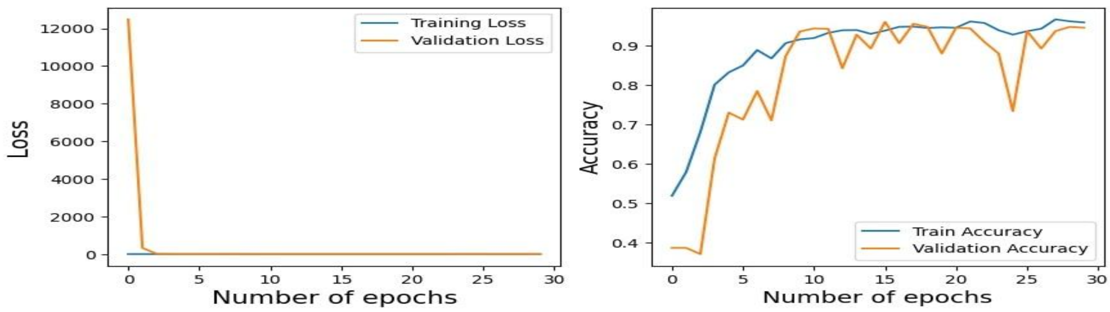
Figure 7: Loss and accuracy plots of CNN models



(a) EfficientNet-B0



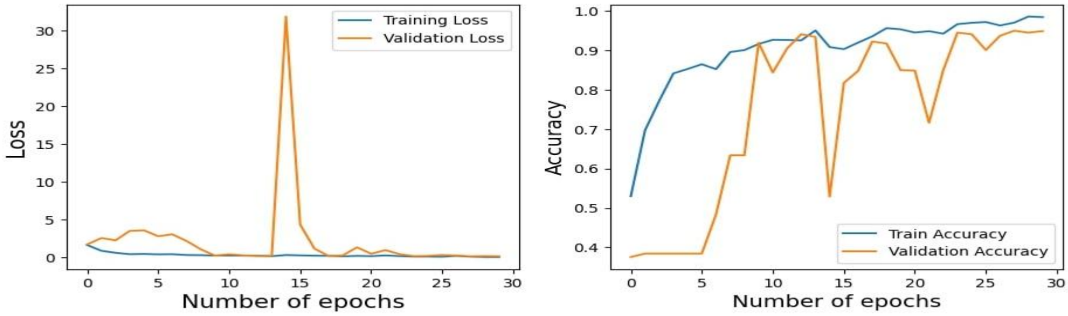
(b) EfficientNet-B1



(c) EfficientNet-B2

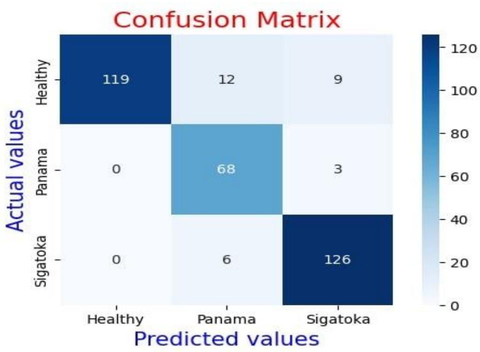
Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

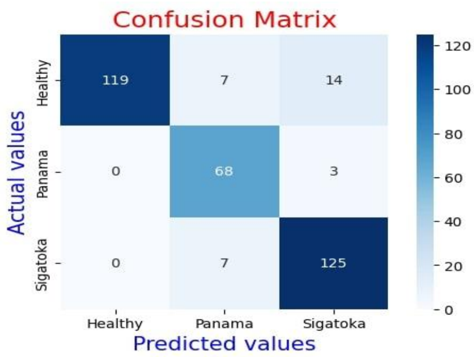
Research Article



(d) EfficientNet-B3

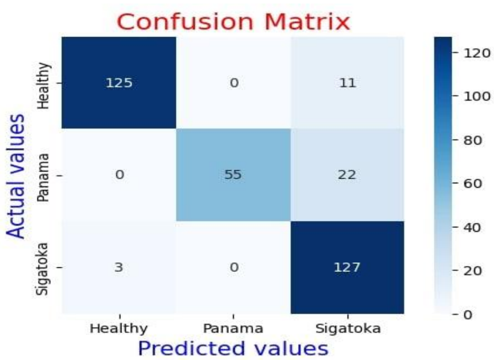
Figure 8: Loss and accuracy plots of the EfficientNet models

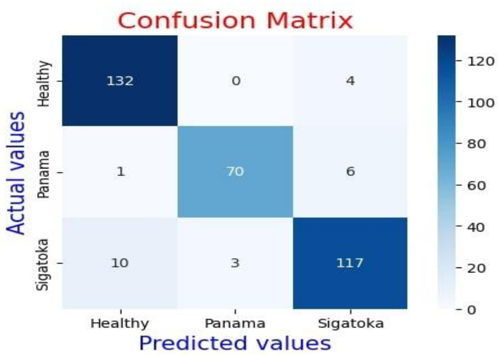




(a) Xception

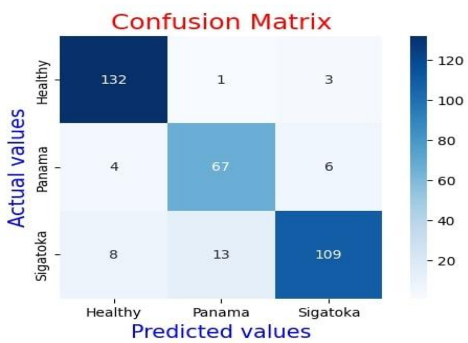
(b) InceptionV3

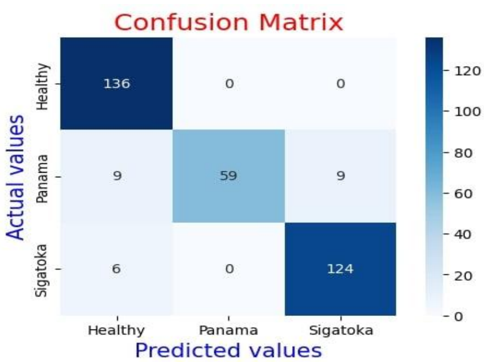




(c) InceptionResNetV2

(d) EfficientNet-B0

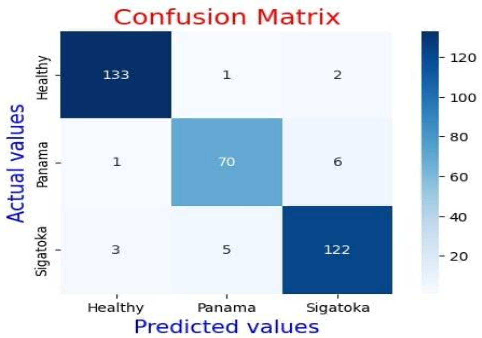




(e) EfficientNet-B1

(f) EfficientNet-B2

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article



(g) EfficientNet-B3
Figure 9: Confusion matrices

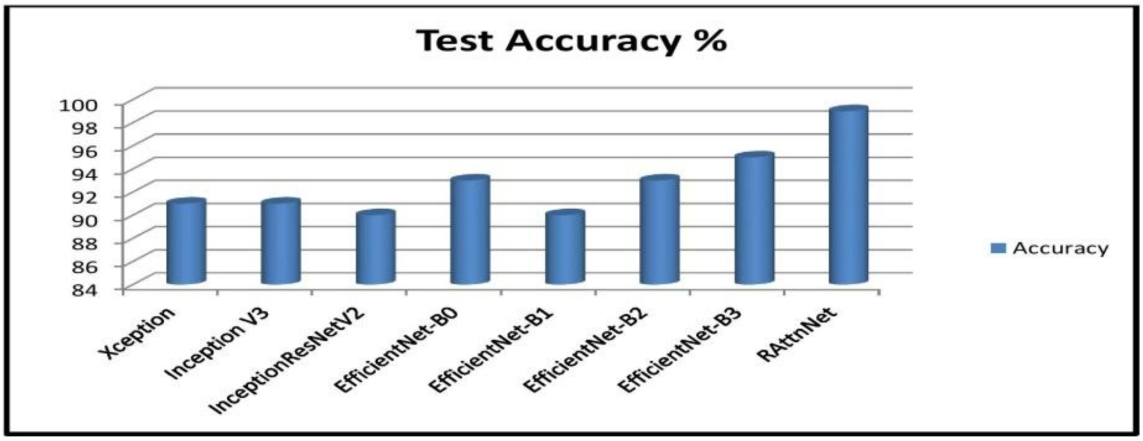


Figure 10: Test accuracies of different models

From table 2, it can be observed that the F1 score achieved by RAttnNet for the healthy, panama and sigatoka classes are impressively high reaching 99%, 97% and 98% respectively. Considering the number of parameters of the models, RAttnNet architecture boasts an impressive parameter count of 3.6 million. Table 3 provides a comparison of the parameter counts of the different deep learning models that are considered in this study. The graph for the same is shown in the figure 11. A comparison of the accuracy achieved by our proposed method with other convolutional neural network architectures for banana leaf disease classification is presented in table 4.

Table 3: Parameter count comparison of different CNNs with proposed model

Model	No. of parameters (in millions)
Xception	22.8
InceptionV3	24
InceptionResNetV2	56
EfficientNet-B0	5.3
EfficientNet-B1	7.8
EfficientNet-B2	9.2
EfficientNet-B3	12
RAttnNet	**3.6**

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

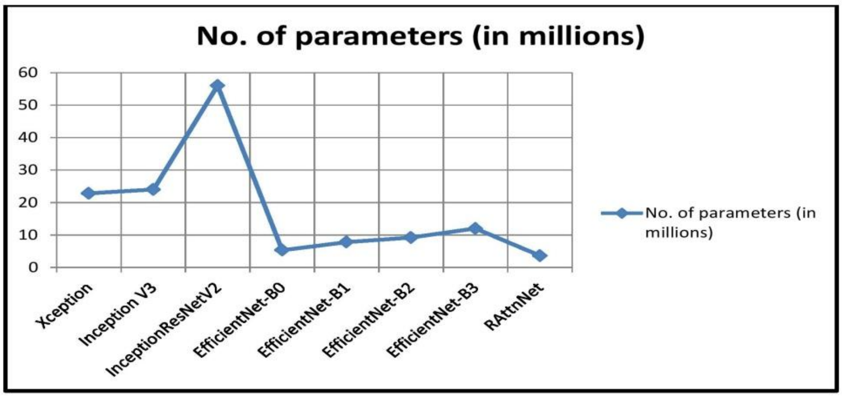


Figure 11: Parameter comparison

Table 4: Accuracy comparison of different CNNs with proposed method

Ref	Classes	Dataset/Source	Algorithm	Accuracy%
[24]	4	BananaLSD dataset	EfficientNet-B0	87.50
[24]	4	BananaLSD dataset	MobileNetV3	86.25
[24]	4	BananaLSD dataset	ResNet-101	90
[24]	4	BananaLSD dataset	SqueezeNet	96.25
[25]	5	CIAT image library	CNN with TGVFCMS	93.45
[25]	5	CIAT image library	CNN with FCM	89.69
[25]	5	CIAT image library	CNN with Autoencoder	85.05
[25]	5	CIAT image library	CNN with ANN	73.54
[26]	2	PlantVillage	CNN LeNet	97
[27]	3	Camera	Capsule Network	95
[27]	3	Camera	CNN	89
[27]	3	Camera	LeNet5	82.5
[27]	3	Camera	ResNet50	96
[28]	1	PSFD-Musa	CNN with ResNet50	98
[29]	1	PlantVillage	ADLM	91.56
Proposed method	3	Smartphone	**RAttnNet**	**99**

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

5. CONCLUSION

There are several challenges with the rising population. It is paramount to address these challenges with sustainable solutions. The most promising solution is the adoption of automation in farm and as such agriculture has become one of the important research topics. This paper addresses the problem of identifying the diseased and healthy banana leaves based on smart-phone images with a new lightweight and cost-effective approach based on a fusion deep learning architecture. From the experimental results it is seen that it produced an impressive accuracy of 99% in identifying the dreaded fungal sigatoka and panama diseases along with the healthy ones. From a comparative analysis with other models in this study, we found our model to be optimum for our research problem. The results confirm that this lightweight model could extract and learn the features of the banana leaves effectively. This research work is limited to three classes of banana leaves. Further studies would comprise of extending the model to classify more classes of banana leaf diseases and also to look in to the other crops. Drone captured aerial images of banana plantations for detecting diseases can also be studied in future.

REFERENCES

[1] V. Chaudhari, M. Patil, Banana leaf disease detection using k-means clustering and feature extraction techniques, in: 2020 International Conference on Advances in Computing, Communication & Materials (ICACCM), IEEE, 2020, pp. 126–130.

[2] W. Liao, D. Ochoa, Y. Zhao, G. M. V. Rugel, W. Philips, Banana disease detection by fusion of close range hyperspectral image and high resolution rgb image, in: IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 2018, pp. 1744–1747.

[3] J. B. Kakati, T. K. Das, Classification of healthy and unhealthy banana leaves using deep learning approach: A comparative assessment, in: 2023 4th International Conference on Computing and Communication Systems (I3CS), IEEE, 2023, pp. 1–5.

[4] A. Aruraj, A. Alex, M. Subathra, N. Sairamya, S. T. George, S. V. Ewards, Detection and classification of diseases of banana plant using local binary pattern and support vector machine, in: 2019 2nd International Conference on Signal Processing and Communication (ICSPC), IEEE, 2019, pp. 231–235.

[5] V. B. C. Calou, A. dos Santos Teixeira, L. C. J. Moreira, C. S. Lima, J. B. de Oliveira, M. R. R. de Oliveira, The use of uavs in monitoring yellow sigatoka in banana, biosystems engineering 193 (2020) 115–125.

[6] V. Singh, A. K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques, Information processing in Agriculture (1) (2017) 41–49.

[7] M. R. Badnakhe, P. R. Deshmukh, An application of k-means clustering and artificial intelligence in pattern recognition for crop diseases, in: International conference on advancements in information technology, Vol. 20, 2011, pp. 134–138.

[8] P. Panchal, V. C. Raman, S. Mantri, Plant diseases detection and classification using machine learning models, in: 2019 4th international conference on computational systems and information Technology for Sustainable Solution (CSITSS), IEEE, 2019, pp. 1–6.

[9] M. G. Selvaraj, A. Vergara, F. Montenegro, H. A. Ruiz, N. Safari, D. Raymaekers, W. Ocimati, J. Ntamwira, L. Tits, A. B. Omondi, et al., 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 169 (2020) 110–124.

[10] M. A. B. Bhuiyan, H. M. Abdullah, S. E. Arman, S. S. Rahman, K. Al Mahmud, Bananasqueezenet: A very fast, lightweight convolutional neural network for the diagnosis of three prominent banana leaf diseases, Smart Agricultural Technology 4 (2023) 100214.

[11] V. G. Krishnan, J. Deepa, P. V. Rao, V. Divya, S. Kaviarasan, An automated segmentation and classification model for banana leaf disease detection, Journal of Applied Biology and Biotechnology 10 (1) (2022) 213–220.

[12] J. Amara, B. Bouaziz, A. Algergawy, A deep learning-based approach for banana leaf diseases classification, Datenbanksystemef¨ur Business, Technologie und Web (BTW 2017)-Workshopband (2017).

[13] B. F. Oladejo, O. O. Ademola, Automated classification of banana leaf diseases using an optimized capsule network model, in: CS & IT Conference Proceedings, Vol. 10, CS & IT Conference Proceedings, 2020.

Journal of Information Systems Engineering and Management
2025, 10(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[14] K. Yan, M. K. C. Shisher, Y. Sun, A transfer learning-based deep convolutional neural network for detection of fusarium wilt in banana crops, AgriEngineering 5 (4) (2023) 2381–2394.
[15] S. R. J. S. Kaiming He, Xiangyu Zhang, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[16] E. E. S. S. Khasoggi, B., Efficient mobilenet architecture as image recognition on mobile and embedded devices, Indonesian Journal of Electrical Engineering and Computer Science 16 (1) (2019) 389–394.