Journal of Agricultural Sciences (Tarim Bilimleri Dergisi) 2025, 31 (2) : 558 – 576 DOI: 10.15832/ankutbd.1552013

Journal of Agricultural Sciences
(Tarim Bilimleri Dergisi)

J Agr Sci-Tarim Bili
e-ISSN: 2148-9297
jas.ankara.edu.tr

JAS

Diagnosis of Paddy Diseases Using Pre-Trained Architectures and a Proposed Enhanced EfficientNetB3 Model

B Johnson[a*], Thangavel Chandrakumar[a]

[a]*Department of Applied Mathematics and Computational Science, Thiagarajar College of Engineering, Madurai 625015, Tamil Nadu, INDIA*

ARTICLE INFO
Research Article
Corresponding Author: B Johnson, E-mail: johnson.vaigai@gmail.com
Received: 18 September 2024 / Revised: 26 November 2024 / Accepted: 23 December 2024 / Online: 25 March 2025

Cite this article
Johnson B, Chandrakumar T (2025). Diagnosis of Paddy Diseases Using Pre-Trained Architectures and a Proposed Enhanced EfficientNetB3 Model. *Journal of Agricultural Sciences (Tarim Bilimleri Dergisi)*, 31(2):558-576. doi: 10.15832/ankutbd.1552013

ABSTRACT

Rice is an important crop in India and is often affected by pests and diseases, which can lead to a significant drop in production. This research investigates advanced deep learning approaches for accurate paddy disease diagnosis, focusing on comparing several transfer learning models. The study specifically targets diseases such as Tungro, Dead Heart, Hispa, Blast, Downy Mildew, Brown Spot, Bacterial Leaf Blight, Bacterial Panicle Blight, and Bacterial Leaf Streak. The base EfficientNetB3 model attains approximately 95.55 % accuracy during training and 95.12% during evaluation on unseen data. However, it encounters challenges when applied to domain-specific tasks such as diagnosing paddy diseases, frequently experiencing issues such as overfitting and inadequate convergence. To overcome these issues, an Enhanced EfficientNetB3 model was developed, incorporating batch normalization, dropout, and data regularization techniques. The training was conducted using the 'Paddy Doctor' dataset, featuring 10,407 high-resolution images of paddy leaves. It reached an accuracy of 98.92 % during training with a loss rate of 0.1385. For validation, the model reached an accuracy of 98.20 % and a loss rate of 0.1450. On an independent test set, the accuracy 98.50 % obtained with a test loss of 0.1505. With remarkable accuracy and a training time of just 68 minutes, the model demonstrates its significant potential for precise paddy disease diagnosis. Its impressive performance plays a crucial role in advancing disease management and boosting crop yields.

Keywords: Paddy Disease Detection, Transfer Learning, Enhanced EfficientNetB3, Deep Learning in Agriculture, Precision Agriculture, Image Classification

1. Introduction

Rice is a primary food globally; it is very important in ensuring the cultivation of rice with food security. However, paddy crops face numerous diseases that can severely impact both their yield and quality. Understanding and identifying the paddy crop diseases is vital for managing and control the paddy crop. Paddy leaves are susceptible to several significant diseases, such as hispa, blast, tungro, brown spot, downy mildew, dead heart, bacterial leaf streak, bacterial leaf blight, and bacterial panicle blight.

Blast is a highly damaging disease affecting rice, resulting from the fungal pathogen Magnaporthe oryzae. Under optimal conditions for its proliferation, it can cause substantial reductions in crop yield (Rahman et al. 2020; Dubey et al. 2024). Hispa, a pest-related issue, results in leaf damage that can reduce photosynthesis, ultimately affecting crop growth. Dead Heart is a symptom commonly associated with stem borers, which damage the stem and disrupt nutrient flow, leading to dead tillers (Deb et al. 2021). Tungro, a viral disease spread by green leafhoppers, causes stunted growth, reduced tillering, and yellow-orange leaf discoloration, significantly impacting rice yield (Yakkundimath et al. 2022). Brown Spot is another fungal disease that affects Paddy leaves, leading to lesions that can merge and cause extensive damage (Shah et al. 2023). Diseases such as Bacterial Leaf Blight, Downy Mildew, Bacterial Panicle Blight, and Bacterial Leaf Streak are also major concerns for rice cultivation, posing substantial risks to crop yields. These diseases are often characterized by leaf spots, streaks, and blight symptoms that reduce photosynthetic efficiency and weaken plants (Dogra et al. 2023). Timely identification and control of these diseases are essential to maintaining rice yields and safeguarding food security.

Deep learning approaches are now widely utilized for identifying and categorizing paddy diseases due to their effectiveness in analysing and learning from large datasets. Several models, such as Convolutional Neural Networks (CNNs), VGG-16, VGG-19, Inception-v1, ResNet-50, Inception-v3, DenseNet-121, Xception, along with the EfficientNetB2 and EfficientNetB3 architectures, have been explored for various tasks, each exhibiting unique performance levels and contributing to the advancements in the field (Liang et al. 2022; Yakkundimath et al. 2022; Simhadri et al. 2024). These models can examine paddy leaf images to detect disease symptoms, making them highly effective for prompt diagnosis and management in advanced farming practices.

EfficientNetB3, a newer advancement within deep learning, provides notable enhancements in both accuracy and computational efficiency compared to conventional models. It employs a scaling technique that proportionally adjusts the depth, width, and resolution, resulting in improved performance for image recognition tasks (Li et al. 2022; Verma et al. 2024). To enhance the efficiency, EfficientNetB3's architecture is crafted to obtain high accuracy, making it well-suited for use in environments with restricted computational capacity (Bhujel & Shakya 2022).

This research contributes significantly to the field of paddy disease classification through several advancements:

- The study leverages pre-trained architectures, focusing particularly on the Enhanced EfficientNetB3 model, which effectively balances accuracy with computational efficiency.

- A comprehensive dataset comprising 10,407 images across 10 distinct paddy disease classes is utilized, addressing the limitations in dataset diversity found in prior studies.

- The classification model demonstrates notable performance, achieving an accuracy rate of 98.50% in disease detection, thereby establishing a new benchmark for precision in the field.

Given its advantages, EfficientNetB3 has been adapted to the task of predicting paddy leaf diseases, demonstrating superior accuracy and robustness in disease classification. Studies have shown that this model outperforms other deep learning architectures in terms of both speed and precision, particularly when identifying subtle differences between healthy and diseased paddy leaves (Li et al. 2022; Bhujel & Shakya 2022; Ganesan & Chinnappan 2022). This makes EfficientNetB3 an excellent option for building effective and dependable systems aimed at identifying and managing paddy diseases.

This study investigates leveraging pre-trained architectures and introduces an Enhanced EfficientNetB3 model for identifying paddy diseases. By harnessing the capabilities of EfficientNetB3, this approach seeks to offer a robust and efficient solution for the early identification of diseases, thereby assisting in the management and reduction of crop losses caused by rice diseases.

2. Material and Methods

2.1 Dataset

The dataset used in this research is the widely recognized Paddy Doctor Dataset, which is sourced from (https://www.kaggle.com/competitions/paddy-disease-classification), consisting of 30,000 images (Patil et al. 2023). This dataset includes 10 distinct classes of paddy leaf diseases, such as Dead Heart, Downy Mildew, Bacterial Leaf Streak, Brown Spot, Bacterial Panicle Blight, Tungro, Normal, Hispa, Bacterial Leaf Blight, and Blast. This dataset encompasses a broad spectrum of disease conditions, representing various stages of disease development, ranging from early to later stages. It also covers a variety of environmental conditions in which the images were captured, including differences in lighting, image angles, and the growth conditions of the plants. These factors contribute to the dataset's robustness, making it suitable for building models that are capable of generalizing to real-world agricultural scenarios. For visual illustration, Figure 1 presents sample images of different paddy diseases.

The initial dataset of 30,000 images underwent significant pre-processing steps, which involved cleaning, resizing to a uniform 480 x 640 pixels for enhanced computational efficiency, and removing images that were blurred or noisy. Following these pre-processing steps, the dataset was reduced to a more refined set of 10,407 high-quality images. From this processed dataset, 8,324 images were assigned for training, while 2,083 images were kept for testing purposes (Kumar et al. 2023). Table 1 outlines the specific distribution of images between the training and test sets.

In addition, the dataset is evenly distributed across various disease categories, ensuring that no single disease type dominates. This balanced distribution is vital for preventing model bias and guarantees that the trained model can effectively classify a wide range of paddy diseases. Such diversity and balance improve the model’s generalization capability, making it well-suited to handle varying disease conditions that may be encountered in real-world agricultural settings.

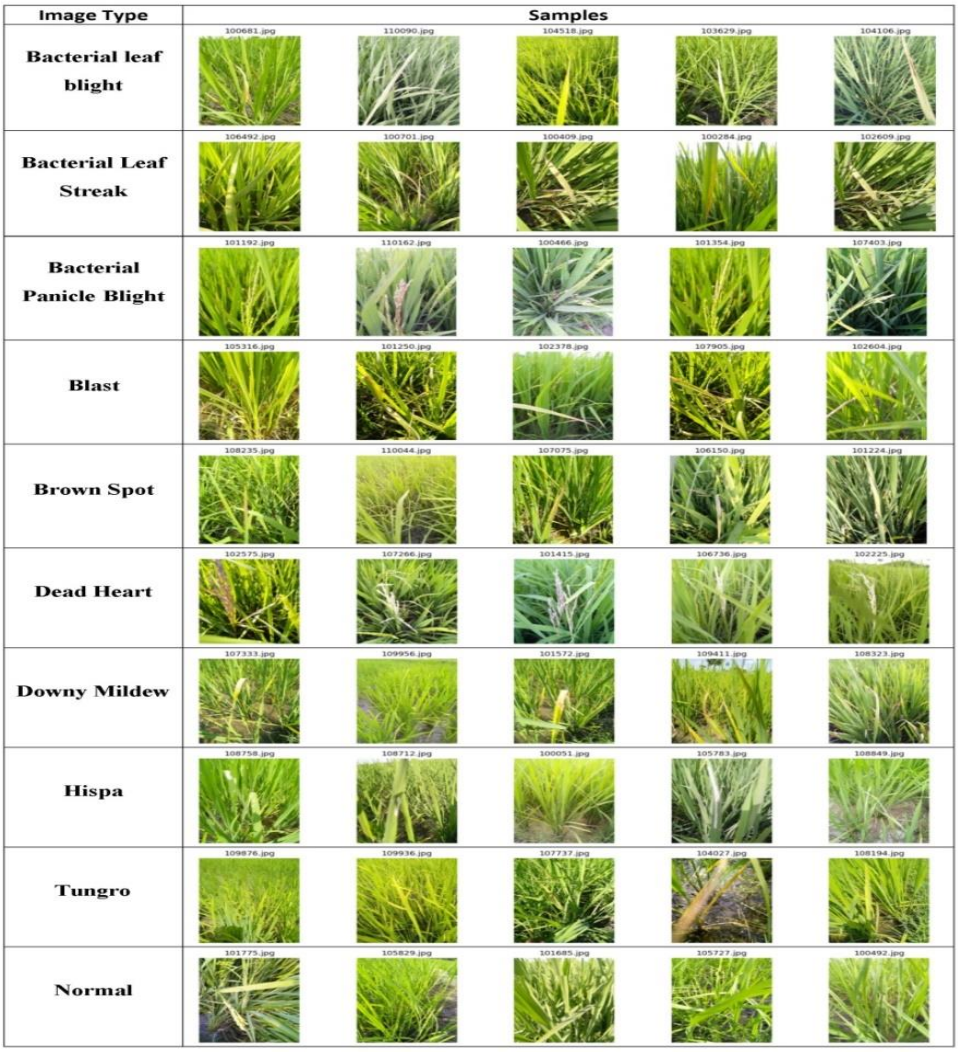

Figure 1- Sample Images of Paddy Diseases

Table 1- Training and Test Dataset Distribution

Disease Category	***Training Images***	***Test Images***
Normal	1411	353
Blast	1390	348
Hispa	1275	319
Dead Heart	1153	289
Tungro	870	218
Brown Spot	772	193
Downy Mildew	496	124
Bacterial Leaf Blight	383	96
Bacterial Leaf Streak	304	76
Bacterial Panicle Blight	270	67
Total	**8324**	**2083**

2.2 Models utilizing convolutional neural networks

The Convolutional Neural Network (CNN) is a widely recognized approach for tasks in natural language processing and image analysis, such as classifying paddy diseases. Its strength lies in its capability to autonomously detect and extract important features using convolutional and pooling layers. This process reduces the complexity of the data while retaining crucial information. This process allows CNNs to manage complex patterns and enhance computational efficiency (Ozdemir 2024; Malvade et al. 2023; Ganesan et al. 2023)**.** When training CNNs for detecting paddy diseases, the model is exposed to images of healthy and diseased leaves, ultimately providing classifications based on the visual features observed. By leveraging such advanced techniques, farmers can more accurately diagnose and manage diseases affecting their crops, leading to improved agricultural practices and crop yields.

For classifying plant diseases, numerous studies were carried out using Convolutional Neural Networks (CNNs). For instance, Shah et al. (2023) conducted a comparative analysis of CNNs with other models, including Inception V3, VGG16, VGG19, and ResNet50. The authors previously discovered that CNNs achieve strong results in the timely detection of rice plant disorders and effectively distinguish between different leaf conditions. In another study, Liang et al. (2022) introduced an enhanced, lightweight CNN based on VGG16, tailored for paddy disease detection and classification, which achieved notable gains in both accuracy and efficiency. Yakkundimath et al. (2022) also demonstrated the application of CNN models for classifying paddy diseases, showcasing the adaptability of these models to different agricultural contexts. Dogra et al. (2023) also used CNN architecture to diagnose brown spot paddy disease.

Within the scope of this research work, the Enhanced EfficientNetB3 deep learning model was developed to diagnose paddy crop diseases. The effectiveness of this model was assessed by comparing it with several other pretrained deep learning models, including Xception, DenseNet-121, ResNet-50, Inception v1, Inception v3, VGG16, VGG19, as well as EfficientNetB2 and EfficientNetB3. This assessment aimed to gauge its accuracy in detecting and categorizing paddy diseases (Ozdemir et al. 2024).

2.2.1 VGG16

The deep convolutional neural network VGG16 has been effectively employed for image categorization and paddy disease diagnosis. This Oxford Visual Geometry Group model, featuring 16 layers with over half being convolutional, is recognized for its simplicity in architecture and effectiveness in feature extraction due to its use of 3 x 3 extension filters throughout the model. This design allows for the capture of small visual features, making VGG16 particularly suitable for paddy disease detection. Numerous studies have demonstrated VGG16's proficiency in paddy disease discrimination compared to other deep learning models. Notable works include Shah et al. (2023), Sun et al. (2023), Liang et al. (2022), and Gerdan et al. (2023), which highlight the model's efficacy across various agricultural datasets (Liang et al. 2022; Sun et al. 2023; Shah et al. 2023; Gerdan et al. 2023). Figure 2 depicts the VGG16 architecture.

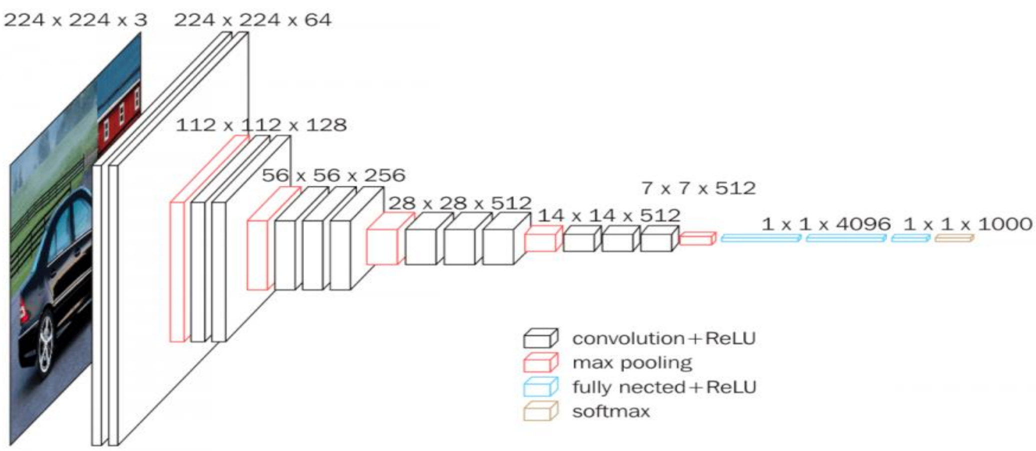

Figure 2- VGG16 architecture

2.2.2 VGG19

VGG19 is an advanced variant of the VGG16 model, incorporating 19 layers for enhanced image classification. This deeper architecture allows VGG19 to capture more intricate features from input images, which is especially advantageous for identifying subtle details in paddy leaf diseases. The model retains the use of small receptive fields with 3×3 filters, which helps preserve the spatial resolution throughout the network. This characteristic is crucial for applications in agriculture, such as precise disease identification in crops (Simonyan & Zisserman 2015). VGG19's effectiveness in paddy disease detection has been demonstrated in various studies applying deep learning techniques to agricultural datasets (Shah et al. 2023; Sun et al. 2023). The architecture of VGG19 is shown in Figure 3.

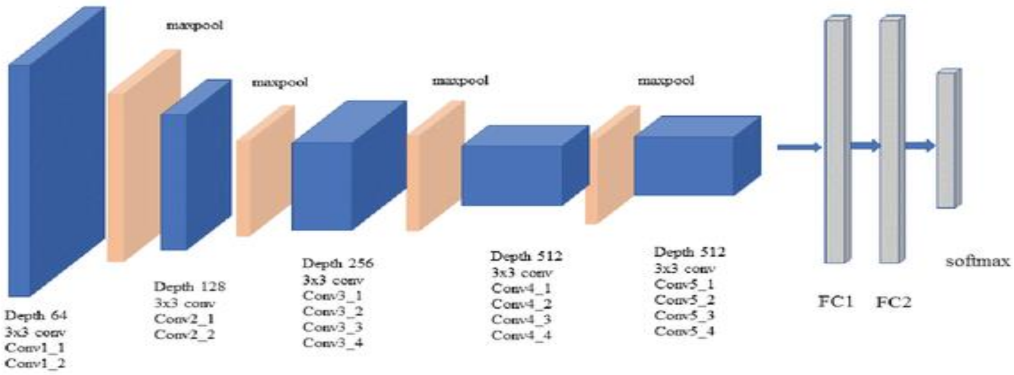

Figure 3- VGG19 architecture

2.2.3 Inception V1

The Inception V1 model is specifically designed to enhance performance in various image recognition and classification challenges. To improve the Inception V1 model's effectiveness with a plant disease image dataset, Particle Swarm Optimization (PSO) techniques are used to adjust and fine-tune the model's hyperparameters. This optimization process helps in achieving better accuracy and efficiency in disease identification (Liang et al. 2022). Using PSO has been demonstrated to markedly enhance the model's performance on unfamiliar datasets, including those related to plant diseases (Rahman et al. 2020). Additionally, deep learning models like Inception V1 have demonstrated their effectiveness in agricultural applications by enhancing disease classification and detection (Sun et al. 2023). Figure 4 shows the Inception V1 architecture.

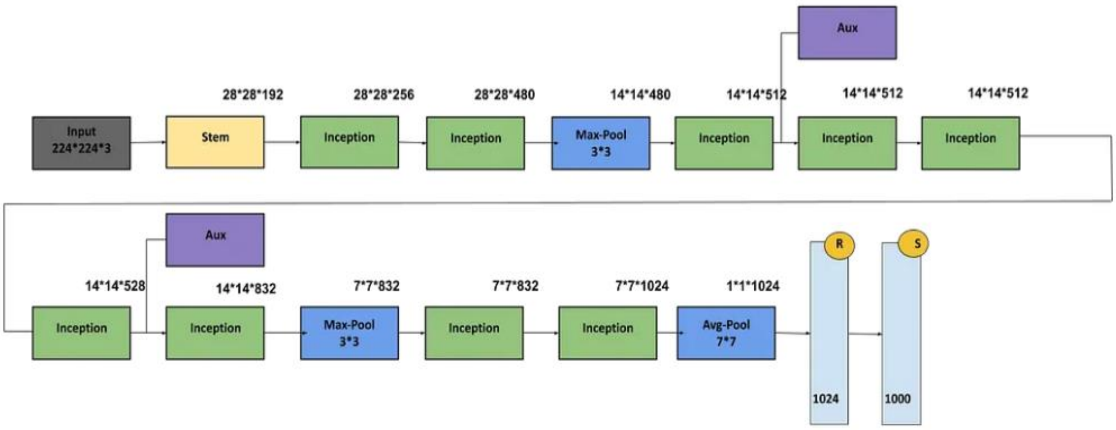

Figure 4- Inception V1 architecture

2.2.4 ResNet-50

The ResNet50 model, where "ResNet" stands for Residual Network, is based on a well-established design that includes fifty layers. This advanced image classification model excels in training with large datasets and achieving leading-edge results. The deep residual learning framework of ResNet50 enhances both feature extraction and classification performance, making it a favoured option for complex image recognition tasks. Research by Shah et al. (2023) and Razavi et al. (2024) has highlighted ResNet50's effectiveness in agricultural and plant disease classification, demonstrating its capability to manage detailed image data efficiently.

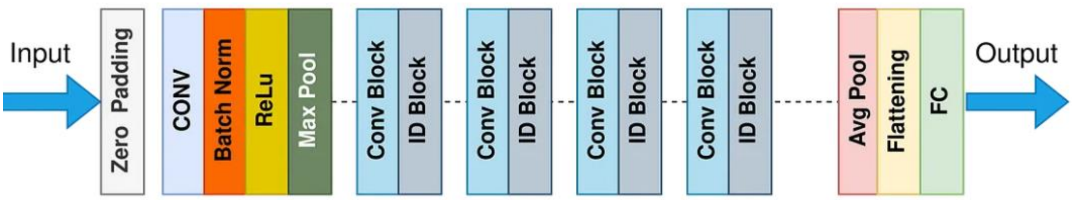

Figure 5- ResNet-50 architecture

2.2.5 Inception V3

The Inception V3 model is a pertained convolutional neural network (CNN) model trained with an extensive image dataset. To enhance its performance on plant disease images, the Particle Swarm Optimization (PSO) techniques are used to adjust and fine-tune the model’s hyperparameters. It’s a model from Google's Inception CNN series that features several techniques such as label smoothing, factorized 7x7 convolutions, BatchNorm in auxiliary classifiers, and the RMSProp Optimizer. This model is commonly used as a foundational architecture and for transfer learning in disease prediction research. The effectiveness of Inception V3 in plant disease classification is demonstrated by Shah et al. (2023) and Liang et al. (2022), highlighting its value in agricultural research. Figure 6 outlines the architecture of the Inception V3 model.

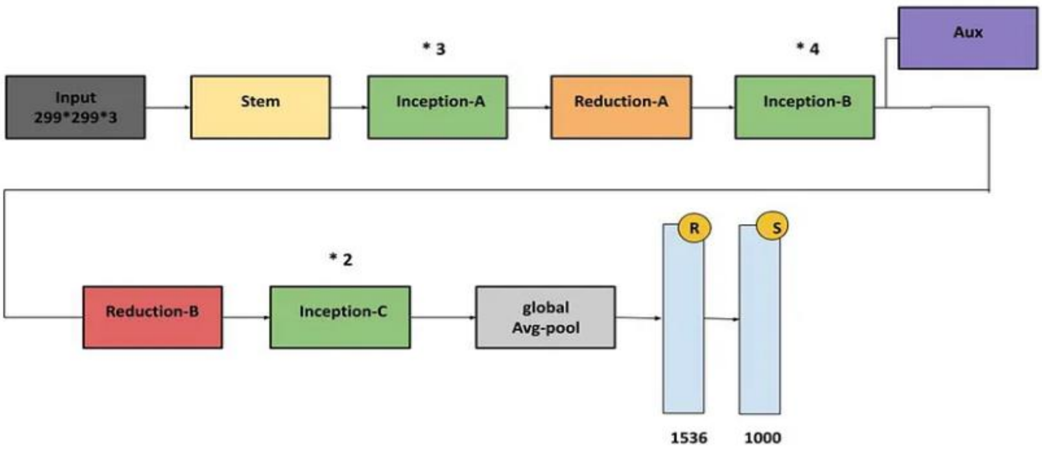

Figure 6- Inception V3 architecture

2.2.6 Densenet-121

The DenseNet121 variant, a notable version of this architecture, includes four dense blocks consisting of 6, 12, 24, and 16 layers in sequence. This architecture's dense connectivity enhances its ability to capture and learn complex features, which proves highly effective for paddy disease detection. The dense connectivity in DenseNet121 supports the extraction and learning of intricate features from paddy leaf images, facilitating the accurate identification of different paddy diseases and improving diagnostic precision. This capability is emphasized by Rahman et al. (2020) and Liu et al. (2022), who showcased how DenseNet models can be highly effective for detecting and classifying paddy diseases. Figure 7 depicts the architecture of DenseNet-121.

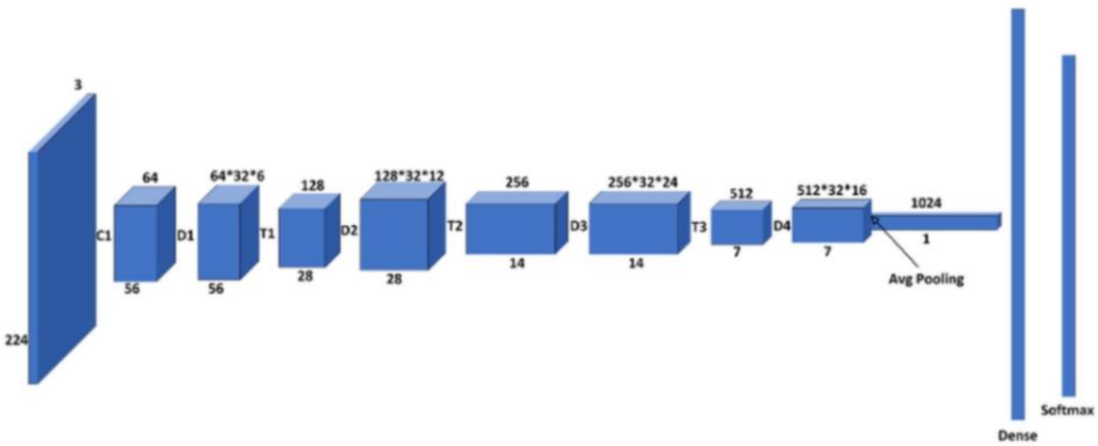

Figure 7- Densenet-121 architecture

2.2.7 Xception

The Xception model is an advanced deep convolutional neural network that utilizes depth wise separable convolutions to enhance the feature extraction process. Building on the Inception model, Xception aims to improve computational efficiency and model performance. Its application in agriculture, especially for detecting paddy diseases, is recognized for efficiently capturing intricate details from images. This capability is enhanced by its depth wise separable convolution layers, which diminish the number of parameters, thereby minimizing the risk of overfitting—especially beneficial in resource-constrained environments such as smart agriculture systems (Meena et al. 2024). Additionally, the model's effectiveness in various agricultural applications has been documented by Rahman et al. (2020) and Sun et al. (2023), highlighting its robustness and adaptability. Figure 8 illustrates the architecture of Xception.

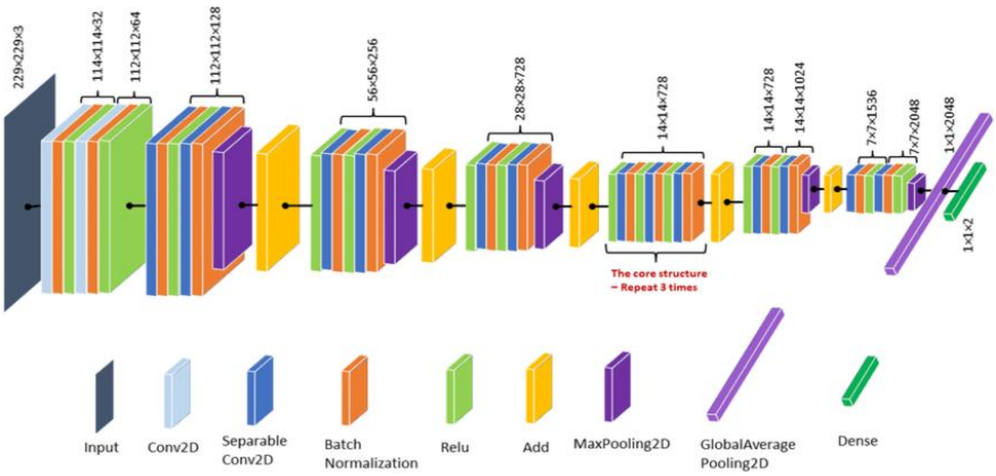

Figure 8- Xception architecture

2.3. Proposed EfficientNetB3 (Enhanced) CNN Model

With reference to (Mingxing and Quoc, 2019) the proposed CNN model EfficientNetB3 was developed. EfficientNetB3 stands out for its capability in feature extraction, attributed to the scaled-up enhancements from the original EfficientNetB0 architecture.

Model Overview

The proposed EfficientNetB3 model is known for its balance between accuracy and efficiency; it requires modifications for specific tasks like diagnosing paddy diseases. The original classification layer was omitted (include top=False), and a Global Average Pooling (GAP) layer was introduced to condense the spatial information from the feature maps into a compact vector. This modification is represented as

$$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i$$

Where; X_i is the input feature map, and N represents the number of spatial locations.

However, the base EfficientNetB3 model has limitations when applied to domain-specific tasks like paddy disease diagnosis. It typically reaches a training accuracy of about 95.55% and a testing accuracy of approximately 95.12%. Nonetheless, the model often experiences problems with overfitting and may not achieve optimal convergence. When a model adapts too closely to the training data, it leads to overfitting. This makes it worse at handling new data. This issue manifests as a significant discrepancy between the accuracy achieved during training and that observed during testing. Additionally, the model's performance can plateau during training, indicating that it does not fully utilize its learning capacity. Figure 10 illustrate the architecture of base EfficientNetB3 model.

To address these issues, the Enhanced EfficientNetB3 model incorporates several key modifications. Batch Normalization was employed to increase both the stability and efficiency of training by standardizing the inputs at each layer. This method lessens the effects of internal covariate shift, thereby boosting the overall performance of the model. This technique is expressed as

$$\hat{x}^{(k)} = \frac{x^{(k)} - \mu^{(k)}}{\sqrt{\sigma^{2k} + \epsilon}}$$

Where; $x^{(k)}$ is the input to the k[th] layer, $\mu^{(k)}$ and σ^{2k} denotes the mean and variance of the batch, while ϵ is a small constant used to ensure numerical stability.

Dense layers with L2 regularization were also added to enhance learning capacity by incorporating more trainable parameters. L2 regularization, given by $\lambda\Sigma_i w_i^2$ in the loss function, penalizes large weights, helping to mitigate overfitting and encourage better generalization.

To further regularize the model, dropout was applied, randomly deactivating a proportion of neurons during training through a Bernoulli (p) process. This method keeps the model from relying too much on certain neurons, which helps avoid over fitting. It was finally finished by adding a Dense Output layer with a Softmax activation function where the Softmax function $\sigma(z)_i = \frac{e^{z_i}}{\Sigma_j e^{z_j}}$ converts raw scores into probabilities, allowing for clear probabilistic interpretation in multi-class classification.

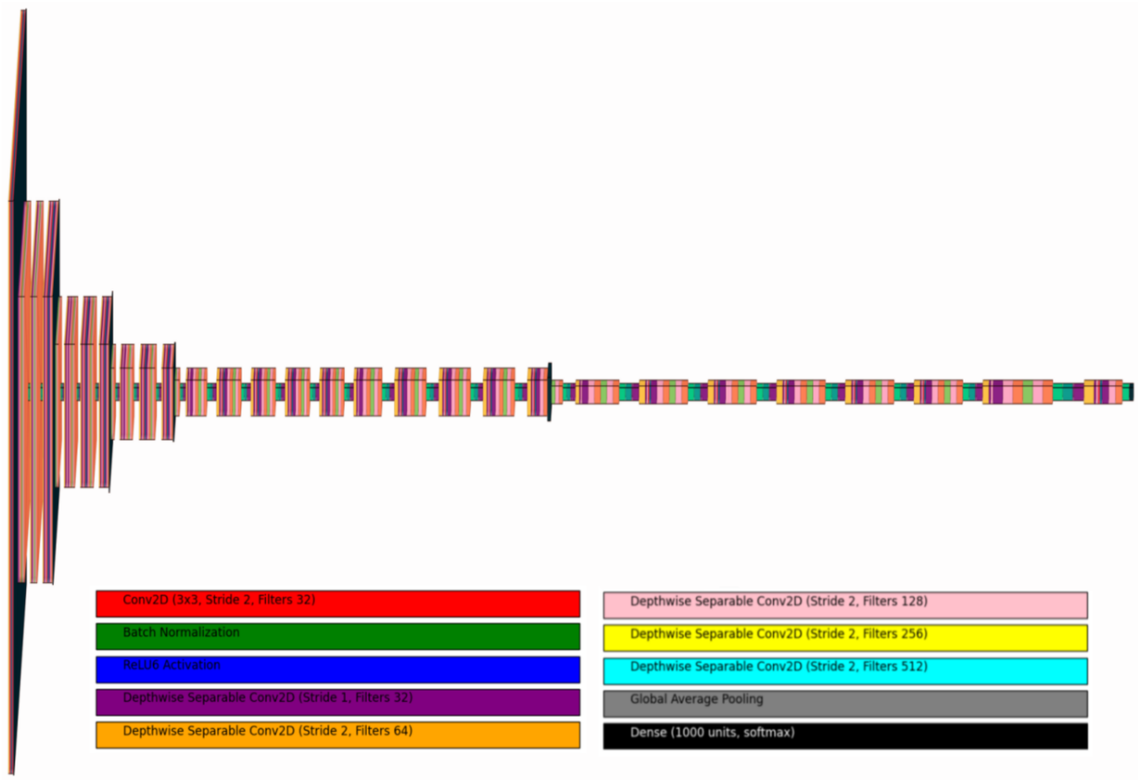

Figure 10- Architecture of Base EfficientNetB3 Model

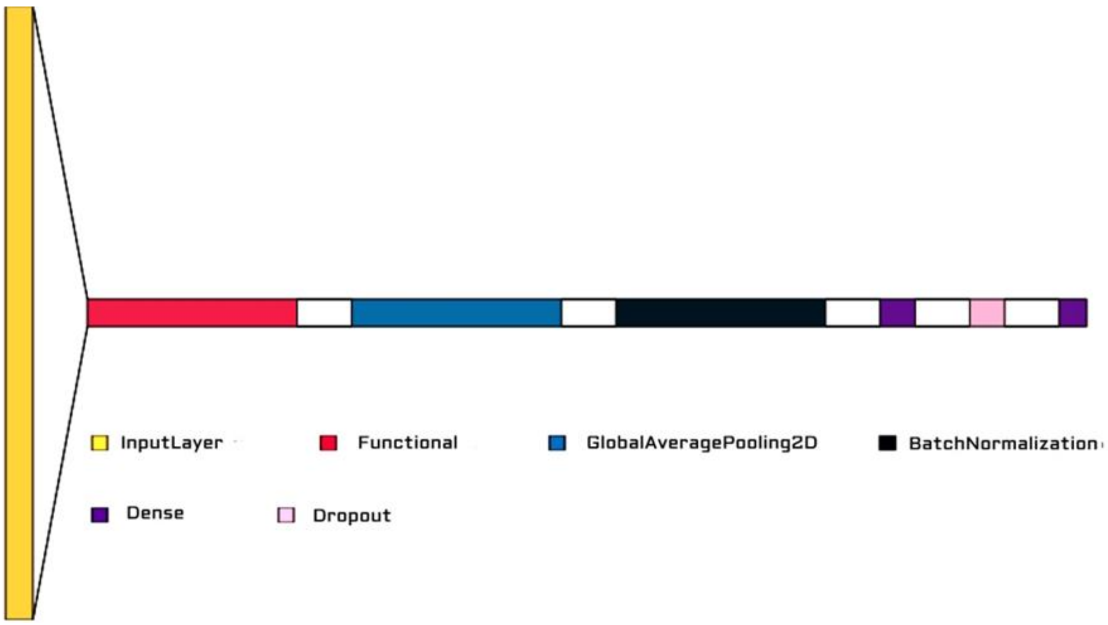

Figure 11- Architecture of the Enhanced EfficientNetB3 Model

Table 2 outlines the parameters of the base EfficientNetB3 model, whereas Table 3 details the updated parameters for the Enhanced EfficientNetB3 model, which aims to enhance both performance and efficiency. Figure 11 illustrates the architecture of the Enhanced EfficientNetB3 model, highlighting key components such as Batch Normalization, Dense layers with L2 regularization, Dropout, and a final Dense layer equipped with a Softmax activation function.

Table 2- EfficientnetB3 Model Architecture

Layer Type	***Input Shape***	***Output Shape***	***Parameters***
Conv2D	(224, 224, 3)	(112, 112, 32)	Kernel size: 3x3, Stride: 2, Filters: 32
Batch Normalization	(112, 112, 32)	(112, 112, 32)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Activation (ReLU6)	(112, 112, 32)	(112, 112, 32)	ReLU6 Activation
Depthwise Separable Conv2D	(112, 112, 32)	(112, 112, 32)	Kernel size: 3x3, Stride: 1, Depthwise, Filters: 32
Batch Normalization	(112, 112, 32)	(112, 112, 32)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Depthwise Separable Conv2D	(112, 112, 32)	(56, 56, 64)	Kernel size: 3x3, Stride: 2, Depthwise, Filters: 64
Batch Normalization	(56, 56, 64)	(56, 56, 64)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Depthwise Separable Conv2D	(56, 56, 64)	(28, 28, 128)	Kernel size: 3x3, Stride: 2, Depthwise, Filters: 128
Batch Normalization	(28, 28, 128)	(28, 28, 128)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Depthwise Separable Conv2D	(28, 28, 128)	(14, 14, 256)	Kernel size: 3x3, Stride: 2, Depthwise, Filters: 256
Batch Normalization	(14, 14, 256)	(14, 14, 256)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Depthwise Separable Conv2D	(14, 14, 256)	(7, 7, 512)	Kernel size: 3x3, Stride: 2, Depthwise, Filters: 512
Batch Normalization	(7, 7, 512)	(7, 7, 512)	Axis: -1, Momentum: 0.99, Epsilon: 1e-3
Global Average Pooling	(7, 7, 512)	(512)	
Dense (Fully Connected Layer)	(512)	(1000)	Units: 1000, Activation: Softmax

Table 3- Enhanced Part of EfficientNetB3 Model Architecture

Layer Type	***Input Shape***	***Output Shape***	***Parameters***
Batch Normalization	(512)	(512)	Axis: -1, Momentum: 0.9, Epsilon: 0.001
Dense Layer 1	(512)	(256)	Units: 256, Activation: ReLU, Kernel Regularizer: L2(0.016)
Dropout Layer	(256)	(256)	Dropout Rate: 0.5, Seed: 123
Dense Output Layer	(256)	(10)	Units: 10, Activation: Softmax

Furthermore, a custom callback function was utilized to adjust the learning rate dynamically in response to training and validation metrics, thereby enhancing the optimization process. Data generators were employed to handle both the training and validation datasets. Several measures, such as loss, accuracy, confusion matrices, and classification reports, were hired to evaluate the model's effectiveness. The trained model and its weights were preserved for future applications, allowing for potential reuse or additional fine-tuning. These modifications ensure that the Enhanced EfficientNetB3 model effectively overcomes the limitations of the base model, providing improved performance in the task of paddy disease diagnosis. The algorithm for diagnosing paddy diseases using the Enhanced EfficientNetB3 model is detailed below.

Algorithm for diagnosing paddy diseases using the Enhanced EfficientNetB3 model

Input:

Dataset D = (X, Y), where X is the set of images and Y are the corresponding disease labels.

Output:

- Trained Enhanced EfficientNetB3 Model $M_{enhanced}$
- Evaluation Metrics E
- Classification Report R

1 Import libraries:

L ← {TensorFlow, Keras, NumPy, Pandas, Matplotlib};

2 Load and Preprocess Dataset:

(X, Y) ← LoadAndPreprocessDataset(D);

3 Data Splitting:

$(X_{train}, Y_{train}), (X_{val}, Y_{val}), (X_{test}, Y_{test}) \leftarrow$ StratifiedSplit(X,Y)

4 Data Augmentation Setup:

DataAugmentation ← ConfigureAugmentation(horizontal_flip=True)

5 Create Data Generators:

TrainingData ← CreateGenerator(DataAugmentation, X_train, Y_train)
ValidationData ← CreateGenerator(None, X_val, Y_val)
TestData ← CreateGenerator(None, X_test, Y_test)

6 Initialize Base Models:

M_base ← EfficientNetB3(include_top=False, weights='imagenet')

7 Add Feature Extraction Layer:

M_base ← M_base + GlobalAveragePooling2D()

8 Enhance Model:

M_enhanced ← M_base
M_enhanced ← M_enhanced + BatchNormalization(axis=-1)
M_enhanced ← M_enhanced + Dense(units, activation='relu')
M_enhanced ← M_enhanced + L2Regularization(strength)
M_enhanced ← M_enhanced + Dropout(rate)
M_enhanced ← M_enhanced + Dense(number_of_classes, activation='softmax')

9 Compile Model:

M_enhanced ← Compile(optimizer, loss_function, evaluation_metrics)

10 Define Training Parameters:

Params ← {batch_size, epochs, learning_rate}

11 Configure Callbacks:

Callbacks ← SetupCallbacks(monitor='val_loss', patience, save_best=True)

12 Train the Model:

M_enhanced ← TrainModel(M_enhanced, TrainingData, ValidationData, Params, Callbacks)

13 Evaluate Model:

E, R ← EvaluateModel(M_enhanced, TestData)

14 Save Trained Model:

SaveModel(M_enhanced, 'model_path.h5')

2.4. Training-Testing data and model evaluation

In this study, the data is divided into three different parts: 80% for training, 10% for validation, and 10% for testing. The choice of these proportions is based on established practices in machine learning to ensure a well-balanced approach. Allocating 80% of the data for training provides a substantial amount of samples for the model to learn from, which is crucial for developing a robust and effective model (Shah et al. 2023). A segment of the 10% validation set is used to assess the model's effectiveness during training and adjust its hyperparameters. This intermediate evaluation assists in reducing overfitting by providing ongoing feedback (Sun et al. 2023; Kiratiratanapruk et al. 2020). The remaining 10% of the dataset is reserved for testing, offering an unbiased assessment of the model's accuracy with new data. This step ensures that the assessment of its ability to generalize is precise and reflective of real-world conditions (Li et al. 2022; Rahman et al. 2020). Table 4 details the parameters used during the model training phase.

Table 4- Model Training Parameters

Parameter	***Value***
Batch Size	64
Epochs	32
Momentum	0.9
Learning Rate	0.0002
Metric	Categorical Crossentropy
Patience	3
Factor	0.2
Verbose	2
Optimization Method	AdamW
Dropout Rate	0.5
Image Augmentation	Random rotations, flips, brightness adjustments
Regularization	L2 regularization

The parameters in Table 4 were selected to improve model performance and minimize overfitting. A batch size of 64 balances computational efficiency with training stability. Training the model for 32 epochs allows adequate learning without risking overfitting. A momentum of 0.9 speeds up training by incorporating previous gradients, while a learning rate of 0.0002 ensures steady progress. The Categorical Crossentropy metric measures accuracy for multiple classes. The patience of 3 reduces the learning rate if no improvement occurs, using a reduction factor of 0.2 for gradual changes. Verbose level 2 offers detailed training feedback. The AdamW optimizer effectively manages large datasets and prevents overfitting. A dropout rate of 0.5 randomly disables neurons to improve generalization. L2 regularization prevents overly complex models by penalizing high weights. Image augmentation, including rotations, flips, and brightness changes, diversifies training data, enhancing the model's adaptability to different scenarios.

2.4.1 Model evaluation metrics

To evaluate the paddy disease classification model, several metrics are used:

Precision measures the accuracy of positive predictions:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Where; TP denotes true positives and FP denotes false positives.

Recall evaluates the model's ability to identify all relevant positive cases:

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

Where; FN represents false negatives.

F1-Score balances precision and recall:

$$\text{F1-Score} = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision+Recall}}$$

Support refers to the number of actual occurrences of each class, providing context for other metrics.

Training Loss and **Training Accuracy** are calculated to assess performance during training:

$$\text{Training Loss} = -\frac{1}{N} \cdot \sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

And

Training Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$

Validation Loss and **Validation Accuracy** evaluate performance on the validation set, while **Test Loss** measures effectiveness on the test set.

3. Results

Table 5 presents a comparative evaluation of the Enhanced EfficientNetB3 model alongside several other CNN architectures, emphasizing critical metrics like accuracy and loss during both training and validation phases.

The Enhanced EfficientNetB3 model accomplished an impressive training accuracy of 98.92% and recorded a low training loss of 0.1385. During the validation phase, the system achieved a performance level of 98.20% accuracy and a validation loss of 0.1450. It is also found that a test accuracy of 98.50% with a test loss of 0.1505 for the independent test dataset. These metrics demonstrate the model's robust performance across various evaluation stages, outpacing several other models as shown in Table 5. Figure 12 displays the trends in both training and validation loss and accuracy for the enhanced model.

Table 5- Performance Metrics and Training Time of CNN Models

Model		***Training***		***Validation***		***Test***		***Training Time***
Architectures	**Input Size**	**Loss**	**Accuracy**	**Loss**	**Accuracy**	**Loss**	**Accuracy**	**Minutes**
VGG16	224, 224, 3	0.2239	92.12%	0.2050	91.00%	0.1835	93.88%	120
VGG19	224, 224, 3	0.2105	93.01%	0.1980	92.10%	0.1798	94.12%	130
Inception v1	224, 224, 3	0.1984	93.50%	0.1860	92.75%	0.1759	94.45%	150
ResNet-50	224, 224, 3	0.1893	93.87%	0.1755	93.20%	0.1702	94.80%	160
Inception v3	299, 299, 3	0.1785	94.20%	0.1680	93.50%	0.1651	95.02%	180
DenseNet-121	224, 224, 3	0.1709	94.53%	0.1605	94.00%	0.1604	95.23%	200
Xception	299, 299, 3	0.1627	94.95%	0.1550	94.50%	0.1580	95.40%	210
EfficientNetB2	260, 260, 3	0.1594	95.10%	0.1835	93.90%	0.1928	94.85%	90
EfficientNetB3	300, 300, 3	0.1528	95.55%	0.1752	94.60%	0.1864	95.12%	80
Enhanced EfficientNetB3	224, 224, 3	0.1385	98.92 %	0.1450	98.20%	0.1505	98.50%	68

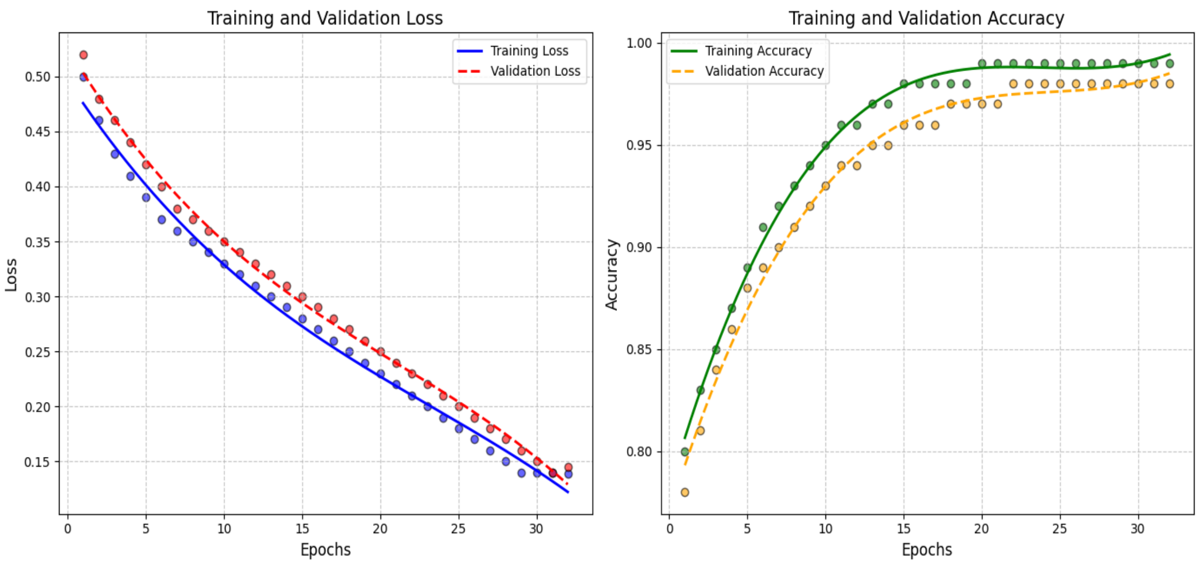

Figure 12- Training and Validation Loss and Accuracy for Enhanced EfficientNetB3

3.1 Confusion matrix

The performance of the Enhanced EfficientNetB3 model was assessed through a confusion matrix, which provided an in-depth evaluation of its ability to classify various paddy diseases. The model achieved remarkable success in accurately identifying both healthy plants and several disease types, such as blast. It showed especially strong performance in recognizing Dead

Heart, where it correctly classified 289 instances, and Blast, with 345 correct predictions. Additionally, the model effectively identified Bacterial Leaf Blight with 94 correct classifications and Bacterial Leaf Streak, correctly identifying all 76 instances. Other diseases, including Brown Spot, Downy Mildew, and Hispa, were also accurately classified with only minor misclassifications. The model performed well in identifying Normal (healthy plants), with 342 correct classifications. The confusion matrix highlights the model's consistent and reliable ability to distinguish between various paddy diseases, reinforcing its potential for practical use in disease diagnosis. The confusion matrix depicted in Figure 13 illustrates the model's proficiency in recognizing different paddy diseases.

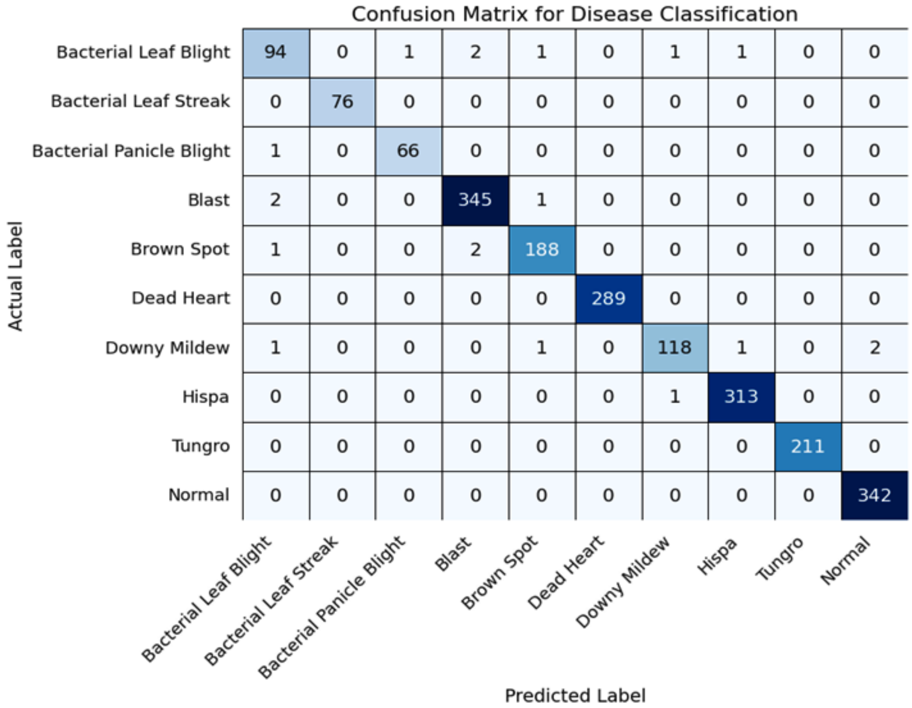

Figure 13- Confusion Matrix for Paddy Disease Classification Model Performance

3.2 Model disease classification

Table 6 displays the Precision, Recall, F1-Score, and Accuracy metrics for the Enhanced EfficientNetB3 model across various paddy disease categories. The model proven exceptional performance, with perfect Precision and Recall of 1.00 for bacterial leaf streak. Additionally, it achieved high accuracy in identifying paddy diseases such as dead heart and blast.

Table 6- Performance Metrics for the Enhanced EfficientNetB3 Model

Disease	***Precision***	***Recall***	***F1-Score***	***Accuracy***
Bacterial Leaf Blight	0.99	0.94	0.96	98.0%
Bacterial Leaf Streak	1.00	1.00	1.00	100%
Bacterial Panicle Blight	0.98	0.98	0.98	98.0%
Blast	0.99	0.98	0.99	98.5%
Brown Spot	0.96	0.98	0.97	97.5%
Dead Heart	1.00	1.00	1.00	99.5%
Downy Mildew	0.95	0.95	0.95	95.0%
Hispa	0.98	0.98	0.98	98.0%
Tungro	0.97	0.99	0.98	98.0%
Normal	0.97	0.99	0.98	98.0%

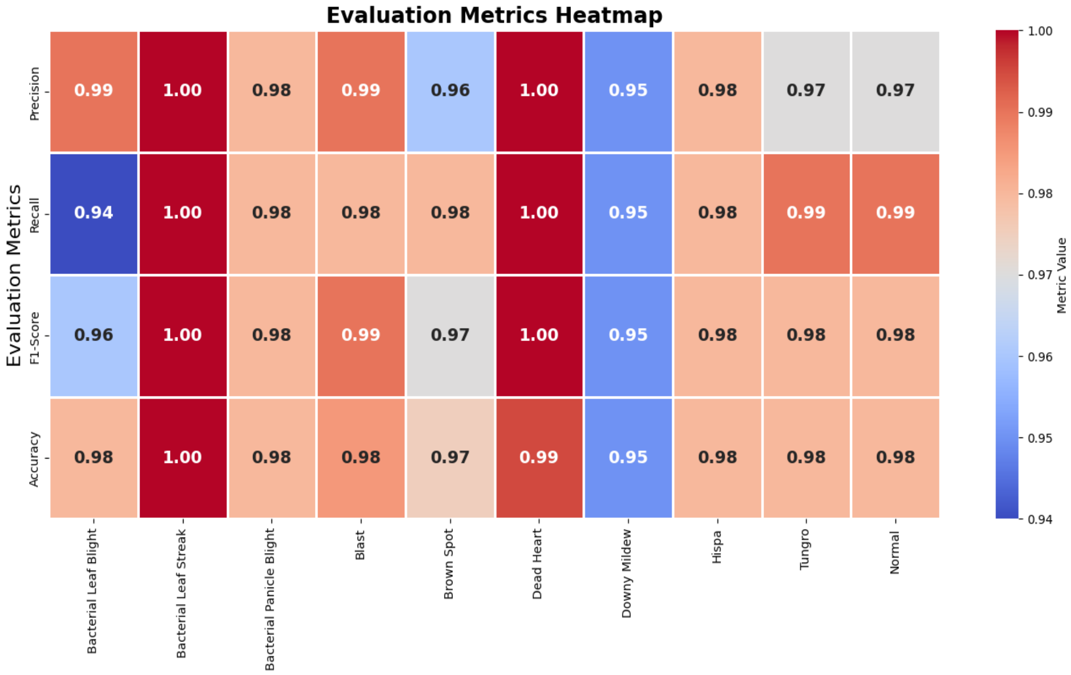

Figure 14- Evaluation Metrics Heatmap

Similar results are observed from the Evaluation Metrics Heatmap in Figure 14 thus ensuring that the model excels most in categorizing paddy diseases. The disease prediction performance of the Enhanced EfficientNetB3 model is compared to other models is shown in Figure 15(a) and 15(b).

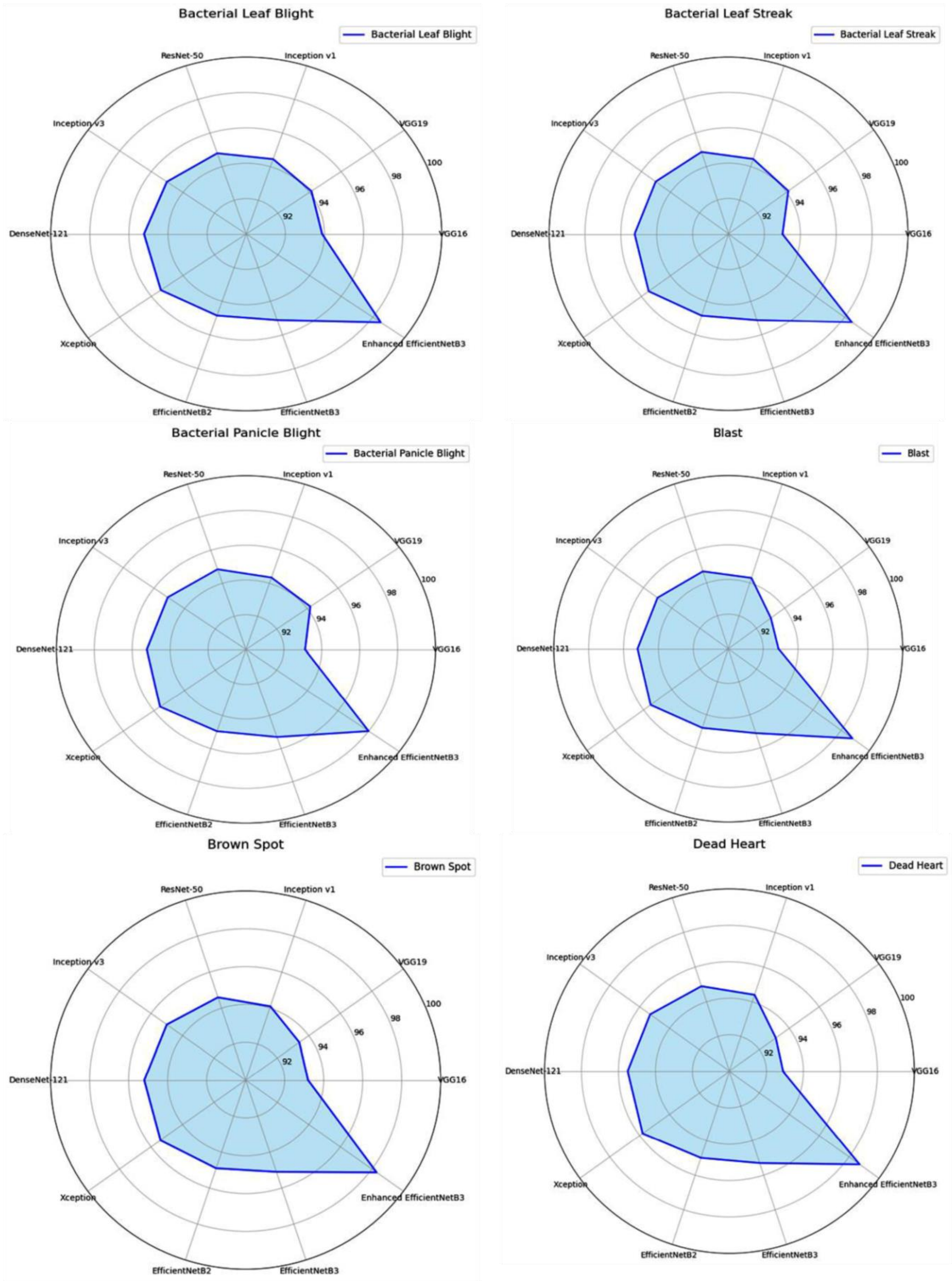

Figure 15 (a)- Predictive Performances of Disease Prediction Models

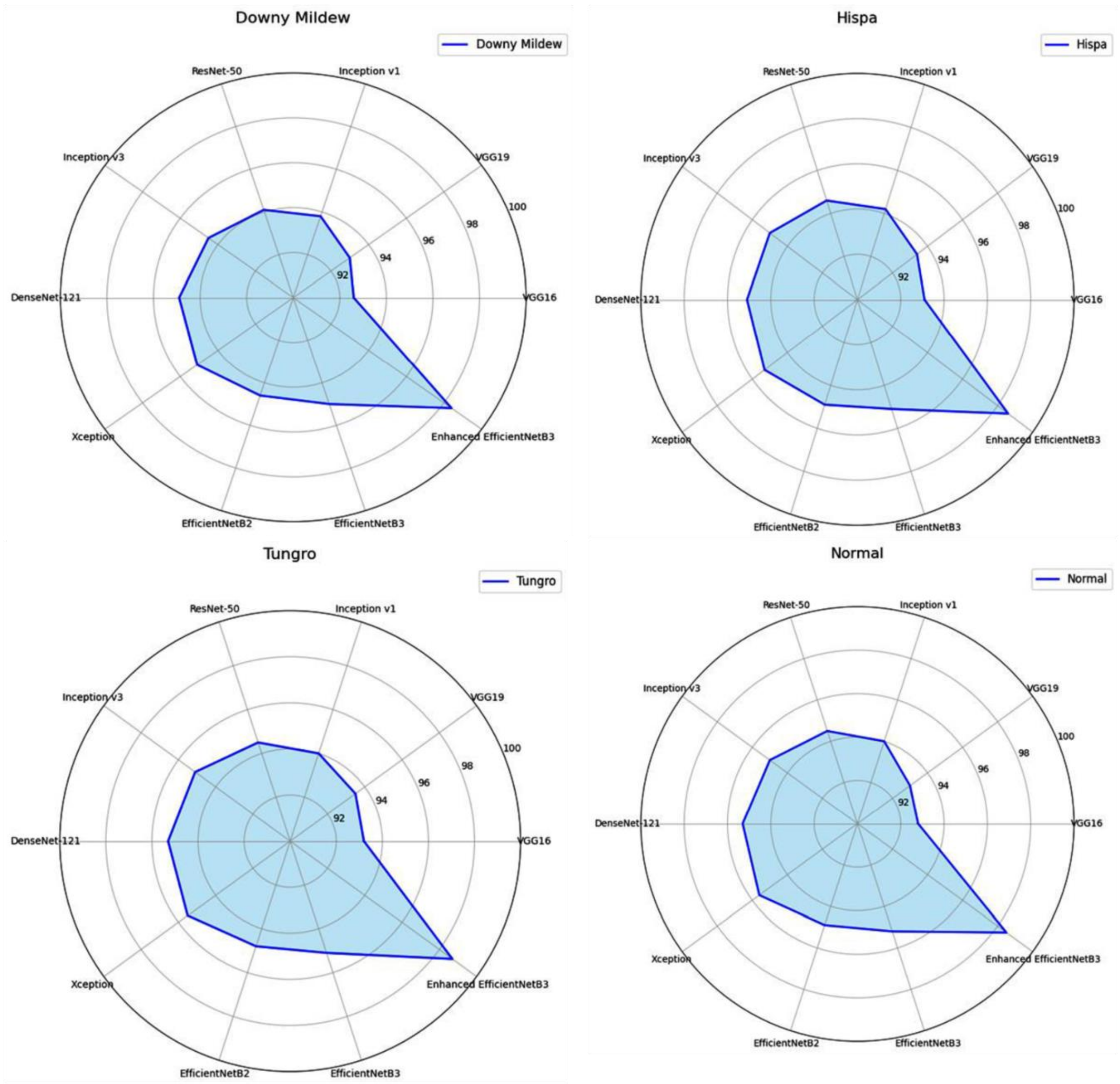

Figure 15 (b)- Predictive Performances of Disease Prediction Models

4. Discussion

The comparative analysis of paddy disease classification models emphasizes the effectiveness of different deep learning architectures for diagnosing paddy diseases. Even simpler models, such as MobileNet, achieve a commendable accuracy of around 90%. In contrast, more advanced models like ResNet50, EfficientNet, and hybrid architectures—such as ResNet-YOLO and DenseNet-UNet—exhibit even higher accuracy, often exceeding 95%. For example, Ganesan & Chinnappan (2022) demonstrated a system that achieved an impressive accuracy of 97.1% using ResNet-YOLO for identifying paddy leaf diseases, showcasing the effectiveness of these advanced architectures in disease detection.

Models incorporating self-attention mechanisms, such as the Self-attention-based ResNet studied by Stephen et al. (2023), achieved an accuracy of 96.7%. This underscores the benefit of self-attention techniques, which enhance model efficiency by focusing on the most critical elements of the dataset. The Enhanced EfficientNetB3 model discussed in this study also aligns with these findings, achieving a test accuracy of 98.50%. While it showed a lower test loss compared to some other models, it performed well on different data samples, indicating robust generalization.

On the other hand, simpler models, including conventional structures like VGG16, basic CNNs, as well as EfficientNetB2 and EfficientNetB3, also demonstrated solid performance with accuracy ranging from 90% to 95%. This range of results suggests that even models with lower complexity can deliver effective results when applied to moderately complex datasets.

Despite the promising performance of the Enhanced EfficientNetB3 model, there are still some limitations that need to be addressed in future research. One limitation is the model's dependency on a large amount of labeled training data, which may not always be available, especially in regions with limited data for paddy disease classification. Additionally, the model's performance can be impacted by noise and variation in data quality, which may affect the accuracy in real-world applications. Future work could explore techniques such as semi-supervised learning or data augmentation to reduce the dependence on large labeled datasets and improve robustness (Ozdemir et al. 2024). Furthermore, incorporating domain-specific knowledge or hybrid models that combine deep learning with expert systems may help improve model accuracy and generalization in diverse field conditions. Another avenue for future research is exploring the deployment of these models on edge devices with limited computational resources. This could involve further optimization of the model to maintain high performance while reducing model size and computational requirements.

4.1. Detailed model performance comparison

Table 6- Comparison of Accuracy Rates of Paddy Disease Classification Models Using Deep Learning

Study & Reference	*Number of Images*	*Diseases Diagnosed*	*Models Used*	*Accuracy Rates (%)*
Shah et al. (2023)	2000	Blast, Brown spot, Bacterial blight	Inception V3, VGG16, VGG19, CNN, ResNet50	Inception V3: 92.4, VGG16: 90.2, VGG19: 91.5, CNN: 89.8, ResNet50: 93.7
Sun et al. (2023)	1500	Seed defects	Rice-VGG16	96.3
Liang et al. (2022)	1200	Multiple paddy diseases and pests	Improved CNN based on VGG16	94.6
Rahman et al. (2020)	3000	Brown spot, Leaf blast, Sheath blight	CNN	91.2
Yakkundimath et al. (2022)	2500	Blast, Sheath blight, Brown spot	Multiple CNN models	90.8
Dogra et al. (2023)	1800	Brown spot	Deep learning model	94.0
Razavi et al. (2024)	2200	Rice cultivar classification	ResNet models	95.8
Ganesan & Chinnappan (2022)	1400	Paddy leaf diseases	Hybrid ResNet-YOLO	97.0
Stephen et al. (2023)	1600	Paddy leaf diseases	Self-attention based ResNet	96.7
Liu et al. (2022)	2100	Paddy leaf diseases	Hybrid DenseNet-UNet	95.4
Li et al. (2022)	1300	Rice germ integrity	Improved EfficientNet	93.5
Bhujel & Shakya (2022)	1700	Paddy leaf diseases	EfficientNet	95.9
Deng et al. (2021)	2000	Multiple Paddy diseases	Deep learning models	93.2
Kaur & Sivia (2024)	1600	Leaf blast	Deep and machine learning convolutional networks	94.8
Latif et al. (2022)	1800	Paddy diseases	Improved CNN	94.2
Hukkeri et al. (2024)	2400	Various plant diseases	Pretrained CNN on ImageNet	92.9
Meena et al. (2024)	1700	Paddy diseases	Xception model	96.0
Proposed Work (2024)	**10407**	**10 Paddy diseases including normal leaves**	VGG16, VGG19, Inception v1, ResNet-50, Inception v3, DenseNet-121, Xception, EfficientNetB2 EfficientNetB3 **Enhanced EfficientNetB3**	93.88%, 94.12% 94.45% 94.80% 95.02% 95.23% 95.40% 94.85% 95.12% **98.50%**

In assessing the training performance of various models, including VGG16, VGG19, Inception v1, ResNet-50, Inception v3, DenseNet-121, Xception, EfficientNetB2, EfficientNetB3, and Enhanced EfficientNetB3, the Enhanced EfficientNetB3 model achieved the lowest training loss at 0.1385 and reached a training accuracy of 98.92%. The results indicate that the model has effectively learned the patterns from the training data, and the validation loss shows consistent performance, reflecting the model's ability to generalize well to new, unseen data.

The Enhanced EfficientNetB3 model showed a validation loss of 0.1450 and a test loss reaching 0.1505, both lower than those observed in other models. It attained a validation accuracy of 98.20% and a test accuracy of 98.50%, demonstrating strong generalization to new data. These observations suggest that the model demonstrates strong performance with new data, with the loss values staying within an optimal range.

Moreover, Enhanced EfficientNetB3 outperformed models such as Xception, DenseNet-121, EfficientNetB2, and EfficientNetB3 in validation accuracy. Its ability to maintain an effective balance between accuracy and loss highlights Enhanced EfficientNetB3 as a highly efficient model, making it an excellent choice for dataset classification.

Additionally, the Enhanced EfficientNetB3 model required only 68 minutes of training time, significantly faster than models such as Xception, which took 210 minutes, DenseNet-121, which took 200 minutes, and EfficientNetB2, which took 90 minutes. This efficiency highlights that Enhanced EfficientNetB3 is not only accurate but also computationally cost-effective compared to other architectures. This aspect makes Enhanced EfficientNetB3 a suitable choice when computational resources or training time are limited, while still maintaining high performance. Table 6 provides a comparison of accuracy rates among various deep learning models for paddy disease classification, including Enhanced EfficientNetB3. Given the lower computational cost and faster training time, Enhanced EfficientNetB3 is recommended when the task requires high accuracy but with limited time or computational resources.

5. Conclusions

The proposed Enhanced EfficientNetB3 model sets a new benchmark in paddy disease identification with its impressive performance, achieving a training accuracy of 98.92%, which exceeds the accuracy of previous models. It shows strong validation and test accuracies of 98.20% and 98.50%, respectively, indicating effective generalization to new data. While the model maintains excellent accuracy overall, it also achieves efficient training in just 68 minutes. The validation and test losses are consistent with the training loss, further emphasizing the model's strong performance. The Enhanced EfficientNetB3 model excels in precision, recall, and F1-scores, making it highly effective for classifying paddy diseases. It also supports real-time inference and resource optimization, providing significant advantages for precision agriculture and offering potential applications in managing diseases across various crops and climates.

Data availability: Indicates the availability of data upon request from the corresponding author.

Authorship Contributions:

Johnson B contributed to the study conception, design, methodology, conceptualization, and resources. Material preparation, investigation, and analysis were also conducted by Johnson B, along with drafting the initial manuscript. Chandrakumar Thangavel supervised the research and provided critical revisions to the manuscript. Both authors read and approved the final manuscript.

Conflict of Interest: The authors declare that they have no conflict of interest.

Acknowledgments

The authors express their gratitude to the Thiagarajar College of Engineering (TCE) for Supporting us to carry out this research work. Also, the financial support from TCE under Thiagarajar Research Fellowship scheme (File.no:TRF/Jan-2023/10) is gratefully acknowledged.

References

Bhujel S & Shakya S (2022). Rice leaf diseases classification using discriminative fine tuning and CLR on EfficientNet. *Journal of Soft Computing Paradigm* 4(3): 172-187. doi.org/10.30534/jscp/2022/05162022

Deb M, Dhal K G, Mondal R & Gálvez J (2021). Paddy disease classification study: A deep convolutional neural network approach. Optical Memory and Neural Networks 30: 338-357. doi.org/10.1134/S1063775821030077

Deng R, Tao M, Xing H, Yang X, Liu C, Liao K & Qi L (2021). Automatic diagnosis of rice diseases using deep learning. Frontiers in Plant Science 12: 701038. doi.org/10.3389/fpls.2021.701038

Dogra R, Rani S, Singh A, Albahar M A, Barrera A E & Alkhayyat A (2023). Deep learning model for detection of brown spot rice leaf disease with smart agriculture. Computers and Electrical Engineering 109: 108659. doi.org/10.1016/j.compeleceng.2023.108659

Dubey R K & Choubey D K (2024). Reliable detection of blast disease in rice plant using optimized artificial neural network. *Agronomy Journal* 116(3): 1099-1111

Ganesan G & Chinnappan J (2022). Hybridization of ResNet with YOLO classifier for automated paddy leaf disease recognition: An optimized model. *Journal of Field Robotics* 39(7): 1085-1109. doi.org/10.1002/rob.22085

Ganesan S, Sinha N & Sundararajan V (2023). A new hybrid deep learning model for classifying rice plant diseases. *Journal of Computational Biology and Bioinformatics* 17(2): 99-110. doi.org/10.1016/j.jcb.2023.103122

Gerdan D, Koç C & Vatandaş M (2023). Diagnosis of tomato plant diseases using pre-trained architectures and a proposed convolutional neural network model. *Journal of Agricultural Sciences* 29(2): 618-629

Hukkeri G S, Soundarya B C, Gururaj H L & Ravi V (2024). Classification of Various Plant Leaf Disease Using Pretrained Convolutional Neural Network on Imagenet. *The Open Agriculture Journal* 18(1). doi.org/10.2174/1874331502414010113

Kaur G & Sivia J S (2024). Development of deep and machine learning convolutional networks of variable spatial resolution for automatic detection of leaf blast disease of rice. Computers and Electronics in Agriculture 225: 109210. doi.org/10.1016/j.compag.2023.109210

Kiratiratanapruk K, Temniranrat P, Sinthupinyo W, Prempree P, Chaitavon K, Porntheeraphat S & Prasertsak A (2020). Development of paddy rice seed classification process using machine learning techniques for automatic grading machine. *Journal of Sensors* 2020(1): 7041310. doi.org/10.1155/2020/7041310

Kumar Y, Singh R, Moudgil M R & Kamini (2023). A systematic review of different categories of plant disease detection using deep learning-based approaches. Archives of Computational Methods in Engineering 30(8): 4757-4779. doi.org/10.1007/s11831-022-09783-x

Latif G, Abdelhamid S E, Mallouhy R E, Alghazo J & Kazimi Z A (2022). Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. Plants 11(17): 2230. doi.org/10.3390/plants11172230

Li B, Liu B, Li S & Liu H (2022). An improved EfficientNet for rice germ integrity classification and recognition. Agriculture 12(6): 863. doi.org/10.3390/agriculture12060863

Liang K, Wang Y, Sun L, Xin D & Chang Z (2022). A lightweight-improved CNN based on VGG16 for identification and classification of rice diseases and pests. In The international conference on image, vision and intelligent systems (ICIVIS 2021) (pp. 195-207). Singapore: Springer Nature Singapore. doi.org/10.1007/978-981-16-7487-8_19

Liu W, Yu L & Luo J (2022). A hybrid attention-enhanced DenseNet neural network model based on improved U-Net for rice leaf disease identification. Frontiers in Plant Science 13: 922809. doi.org/10.3389/fpls.2022.922809

Malvade N N, Yakkundimath R, Saunshi G B & Elemmi M C (2023). Paddy variety identification from field crop images using deep learning techniques. *International Journal of Computational Vision and Robotics* 13(4): 405-419. doi.org/10.1504/IJCVR.2023.129934

Meena R, Joshi S & Raghuwanshi S (2024). Xception model for disease detection in rice plant. *Journal of Intelligent & Fuzzy Systems* (Preprint): 1-18. doi.org/10.3233/JIFS-230473

Ozdemir C, Dogan Y & Kaya Y (2024). RGB-Angle-Wheel: A new data augmentation method for deep learning models. Knowledge-Based Systems, 291: 111615

Ozdemir C (2024). Adapting transfer learning models to dataset through pruning and Avg-TopK pooling. Neural Computing and Applications 36(11): 6257-6270

Ozdemir C, Dogan Y & Kaya Y (2024). A new local pooling approach for convolutional neural network: local binary pattern. Multimedia Tools and Applications 83(12): 34137-34151

Razavi M, Mavaddati S & Koohi H (2024). ResNet deep models and transfer learning technique for classification and quality detection of rice cultivars. Expert Systems with Applications 247: 123276. doi.org/10.1016/j.eswa.2023.123276

Rahman C R, Arko P S, Ali M E, Khan M A I, Apon S H, Nowrin F & Wasif A (2020). Identification and recognition of rice diseases and pests using convolutional neural networks. Biosystems Engineering 194: 112-120. doi.org/10.1016/j.biosystemseng.2020.03.003

Shah S R, Qadri S, Bibi H, Shah S M W, Sharif M I & Marinello F (2023). Comparing Inception V3, VGG 16, VGG 19, CNN, and ResNet 50: A case study on early detection of a rice disease. Agronomy 13(6): 1633. doi.org/10.3390/agronomy13061633

Simhadri C G, Kondaveeti H K, Vatsavayi V K, Mitra A & Ananthachari P (2024). Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques. Information Processing in Agriculture.

Sun J, Zhang Y, Zhu X & Zhang Y D (2023). Enhanced individual characteristics normalized lightweight rice-VGG16 method for rice seed defect recognition. Multimedia Tools and Applications 82(3): 3953-3972. doi.org/10.1007/s11042-022-12954-2

Verma S, Kumar P & Singh J P (2024). A unified lightweight CNN-based model for disease detection and identification in corn, rice, and wheat. *IETE Journal of Research* 70(3): 2481-2492

Yakkundimath R, Saunshi G, Anami B & Palaiah S (2022). Classification of rice diseases using convolutional neural network models. *Journal of The Institution of Engineers* (India): Series B 103(4): 1047-1059. doi.org/10.1007/s40031-022-00732-0