

Transfer Learning-Based Convolutional Neural Network for Accurate Detection of Rice Leaf Disease in Precision Agriculture

Bety Wulan Sari ^{1)*} , Donni Prabowo ²⁾ , Yoga Pristyanto ³⁾ , Afrig Aminuddin ⁴⁾ 

¹⁾²⁾³⁾⁴⁾Department of Information System, Faculty of Computer Science, Universitas AMIKOM Yogyakarta, Yogyakarta, Indonesia.

¹⁾bety@amikom.ac.id, ²⁾donniprabowo@amikom.ac.id, ³⁾yoga.pristyanto@amikom.ac.id, ⁴⁾afri@amikom.ac.id

Abstract

Background: Traditional approaches to rice disease identification depend mainly upon visual examination, which is quite labor-intensive and generally demands a certain skill level from people engaged in this activity. However, these approaches suffer from high time costs and potential errors and are impractical for large-scale daily monitoring. The recent rise of deep learning has offered opportunities for automated detection process improvement, which needs to be fast-accurate as good farmer-centric.

Objective: This study aims to enhance the accuracy of image rice leaf disease classification via feature extraction for rice leaf disease in four instances of pre-trained CNN models and provide an automated solution for early detection ahead of timely care by obtaining insights into crop production through precision agriculture.

Methods: This study combined transfer learning with four pre-trained CNN models - InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16.

Results: The outcome of this research enables the identification of the optimal model to relate datasets where DenseNet121 achieved the highest accuracy, i.e. 99.10%, followed by MobileNetV2, having a precision of 97.10%.

Conclusion: The new framework results in a highly accurate and high-throughput early disease detection element in precision agriculture, better than state-of-the-art approaches based on traditional techniques.

Keywords: Deep Learning, DenseNet121, Image Classification, Rice Leaf Diseases, Transfer Learning

Article history: Received 22 October 2024, first decision 15 May 2025, accepted 9 September 2025, available online 28 October 20XX

I. INTRODUCTION

Detection of crop disease in the early stage with high accuracy is important to ensure more productivity and achieve global food security, particularly as fast growth status correlates with precision agriculture. Rice (*Oryza sativa*) is one of the most economically important food staples in the world, but is extremely vulnerable to several forms of disease that cause massive reductions in yield and literacy [1], [2]. When leaf disease is treated at the initial stage of development, yield can reach a maximum level with and without special assistance. In agriculture, one of the toughest areas to study is crop disease prediction, which has been widely recognized and classified using rehabilitation programs. The widespread classification is attributed to growing public awareness of agriculture's importance to food security [3]. Rice is affected by bacteria leaf blight (BLB), tungro, brown spot, and blast, which are more serious globally. The causal agents include bacteria, viral, or fungal pathogens such as *Xanthomonas oryzae pv. oryzae* (XOO), which continues to pose a great threat to the production rate.

Xanthomonas oryzae pv. oryzae (XOO) causes a fatal disease of rice BLB, affecting many varieties [4]. The infection caused by XOO reduces the efficiency of metabolism and photosynthesis, leading to an 80% decrease in production [5]. The target area is often leaf and causes considerable damage in terms of yield. Symptoms of this disease first appear as water-soaked streaks, which will become yellow and brown starting from leaf tips and margins. Management of BLB, including proper treatment of resistant varieties, clean and sanitation practices in the field, combined with suitable chemicals. Early warning is also essential for controlling moths to reduce congestion in healthy rice production.

Tungro disease occurs due to a complex infection with *rice tungro spherical virus* (RTSV) and *rice tungro bacilliform virus* (RTBV), which is often transmitted through green leafhoppers [6], [7], [8]. These viruses cause various symptoms that appear 6 to 15 days after the infection, including stunting, yellow to orange leaf discoloration,

* Corresponding author

and reduced grain quality [9]. The major symptoms include severe stunting and leaf yellowing, poor root development, low tiller production, as well as small sterile panicles with black blotches on rice grains [10].

Bipolaris oryzae is the primary cause of brown spot (BS) disease that damages crop both qualitatively and quantitatively [11], [12]. This disease compromises the photosynthetic capacity of leaf due to the formation of elongated spot lesions in the green tissues, which progressively expand after infection [13], [14]. When not properly managed, it can lead to considerable reductions in yield, with BS potentially causing an average yield loss of 52% in rice [15].

Magnaporthe oryzae (*M. oryzae*) is another disease, recognized as the source of rice blast that threatens production globally and causes significant output losses every year [16]. Compared to BS, blast can lead to an average yield loss of approximately 30% when not effectively managed [17]. The frequency of rice blast may be increased by several variables, including high humidity (>80%), gloomy and damp weather, low temperatures (15–25%), and heavy fertilizer treatment [18]. At any stage of rice growth, *M. oryzae* can infect airborne components and cause neck, node, grain, and seedling blast [19]. Therefore, there is a need to identify rice disease as an early process of implementing a control strategy.

The conventional methods of identifying rice disease are highly visual, which often require skill and knowledge to be well-trained professionals. However, these methods are extremely slow and labor-intensive for incidental monitoring. To address the limitation, artificial intelligence and deep learning have been identified to accelerate the process of crop disease detection with high accuracy. Deep learning model, particularly convolutional neural network (CNN), has shown promising results in image classification with various applications, such as crop disease detection. CNN is extremely powerful in image processing, operating through a multi-layer framework that includes input (visible) convolutional, pooling, normalization, and fully connected Layers. These learnable parameters (weights) allow CNN to capture spatial relationships in input information and adapt effectively to different classification tasks [20].

Despite the advancement in technology, detection models for leaf disease still have a limited focus [21][22], lacking extensive hyperparameter optimization to maximize performance [23]. The majority works with curated dataset that do not extend to large-scale and field-specific data, limiting practical re-usability in real agricultural settings [24]. Several solutions have focused on reasonably versatile and localized problems, although some models cannot generalize well to broader or more diverse agricultural conditions [25]. There is also limited information on the use of modern pre-trained models to investigate full-scale transfer learning for better accuracy and efficiency in classification [26]. Therefore, an appropriate innovation is needed to develop intelligent automated systems that can be robust and scalable for precisely detecting multiple rice leaf disease in any environmental condition. In this context, transfer learning, where pre-trained models on large dataset are adapted for specific tasks, presents a very attractive solution to improve performance and efficiency, as well as reduce computational costs and training time [27], [28], [29].

Numerous similar studies have been conducted on rice leaf disease using CNN models, one of which applied InceptionResNetV2 [30], VGG16 [20], VGG19 [31], and SVM Classifier [32]. A similar study by Narasimha applied the InceptionResNetV2 model to predict three types of rice leaf disease, namely leaf blast, brown spot, and BLB [30]. In comparison, this study predicts four types, such as tungro, brown spot, blast, and BLB, by comparing the performance of four models, namely InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16. Although previous reports achieved relatively high accuracy, further optimization with additional hyperparameters is recommended using alternative method to achieve improved accuracy.

This study proposes a deep learning-based method for detecting rice leaf disease using transfer learning with four pre-trained CNN models, namely InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16. The objective is to fill the knowledge gap by exploring the use of CNN through feature extraction capability provided to improve rice leaf disease classification accuracy in an automated pattern in precision agriculture. Several rice leaf images need to be used, preprocessed, and reshaped to a size appropriate for uniformity. Augmentation methods such as rotations, translations, and flips are used to create more variation in dataset and prevent overfitting [33], [34]. Pre-trained models, trained on ImageNet, are used as feature extractors that have convolutional layers frozen and keep only general features. Specifically, DenseNet121 uses a *conv_base* set to False to retain pre-trained weights from ImageNet to accelerate the training process and reduce the risk of overfitting. By forking the features into several dense layers tailor-made for classifying rice leaf disease, this model detects numerous disease categories effectively and accurately. The final layer produces softmax outputs that stratify samples into four disease categories. Subsequently, model performance is evaluated using accuracy, precision, recall, and F1 score matrices. The results show the advantages of transfer learning in improving the accuracy and speed of classification [35], [36], which can provide a powerful, large-scale solution for precision agriculture.

The selection of four models is due to the effectiveness in previous studies. For example, InceptionResNetV2 performs multiscale feature extraction with residual connections [30], MobileNetV2 is both lightweight and computationally efficient [37], DenseNet121 fosters efficient gradient flow and feature reuse [38], while VGG16, a strong and simplistic baseline, is well-known for robustness [20]. These architectures are well balanced in their performance and computational cost, showing suitability for use in resource-constrained agricultural settings.

This study focuses on the application of transfer learning frameworks to evaluate the detection of rice leaf disease using deep learning in precision agriculture. The analysis includes estimating InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16 by proposing a custom convolutional block design that improves model generalization and feature extraction for diverse classes of disease. The results support the constructive role of deep learning in agricultural diagnostics and practical application in farming environments, showing the need for timely and precise disease identification. Additionally, this study contributes to the development of more efficient and robust systems for crop disease recognition in precision agriculture by analyzing the strengths and weaknesses of the examined architectures, providing methodological recommendations. The structure of this study is as follows, Section II covers the related works, and Section III defines methods. Section IV discusses the results, Section V shows the performance of model compared to existing method, and Section VI consists of conclusions and recommendations for future studies.

II. LITERATURE REVIEW

Dogra et al. [31] presented the significance of rice as a crucial crop, feeding over half of the world's population. In the study, CNN-VGG19 model was introduced with transfer learning for detecting brown spot rice leaf disease, which achieved 93.0% accuracy. The model performance metrics included 89.9% sensitivity, 94.7% specificity, 92.4% precision, and 90.5% F1 score. The proposed model outperformed the existing baseline model in identifying and classifying rice leaf disease. Another study by Mannepalli et al. [20] presented VGG16 CNN to identify BLB, Leaf Smut (LS), and Brown Spot in rice leaf. During the experiment, VGG16 model was used for feature extraction and classification. The results showed that the implementation of VGG16 contained a high ability to differentiate rice leaf disease with 97.77% classification accuracy. The early and accurate disease diagnosis helped in taking necessary measures that reduced crop loss and the declining economic power of farmers.

Narasimha et al. [30] presented InceptionResNetV2 model with deep neural networks in transfer learning to predict rice leaf disease. A proprietary dataset of rice leaf images was used for the experiment, with image pre-processing and augmentation performed before classifying using a fine-tuned InceptionResNetV2. The model achieved an accuracy of 95.67% in detecting rice leaf disease, which was better than a basic CNN model. This study proved the potential use of transfer learning to increase classification accuracy and proposed future investigation on other crops and optimization for hyperparameters.

Shrivastava and Pradhan [32] showed an image-based method that used color cues to categorize disease affecting rice. The study assessed seven classifiers by observing 14 different color spaces and extracting 172 characteristics. Based on the results, Support Vector Machine (SVM) had the best accuracy, at 94.65%. A dataset of 619 images, divided into four classes such as Rice Blast, BLB, Sheath Blight, and Healthy Leaf, was used to train and evaluate the model. The results showed color cues in an automated rice disease classification system that could help farmers improve crop quality and output.

Chen et al. [39] presented an automated detection and diagnosis system for food safety and quality. Pre-trained deep CNN, VGGNet, and Inception modules, and transfer learning were used to improve accuracy. This solution improved performance by a large margin using pre-trained weights of model trained for ImageNet. The model achieved validation accuracies of 91.83% on the common dataset and 92.00% on rice planting images in challenging backgrounds, despite using computationally inefficient backbone layers. This proved the method efficacy and efficiency in real-time plant disease detection.

Bari et al. [40] identified the main threat to sustainable production and food security, suggesting early detection strategies to secure crop health. One of the reasons for this transition was that classical diagnosis systems used to be sluggish, erroneous, and pricey, producing inefficient outcomes-focused machines. However, the system performance has been affected by challenges in terms of various image backgrounds (making images harder to find schema-related patterns), fuzzy borders of symptoms, inconsistent weather conditions, and co-occurring disease. These challenges have been recently explored using Faster R-CNN algorithm to build a real-time rice leaf disease detection system. Combined with public and field-specific dataset, high accuracy was obtained in diagnosing rice blast (98.09%), brown spot (98.85%), and hispa damage at the seedling stage, along with healthy leaf (99.25%). Faster R-CNN model has shown good results in identifying rice leaf disease accurately and quickly.

Nandhini and Bhavani [41] presented machine learning methods for classifying crop disease, using SVM, K-NN, and decision trees. The experiment used feature extraction methods to obtain the infected area from leaf image. This method enhances the contrast between healthy and diseased leaf areas for model to learn more distinctive representations from weakly labeled data (color and geometry), while effectively extracting diverse features from the leaf's structural patterns. The other two methods were less accurate at classification using the algorithms than SVM.

III. METHODS

This section presents the methodology for classifying rice leaf disease using a deep learning pipeline, mainly focusing on transfer learning with InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16. The process includes several critical stages, namely image acquisition, pre-processing, augmentation, feature extraction, model evaluation, and visualization. Each stage is described below, and relevant equations are introduced to formalize key concepts, showing a six-stage method for identifying rice leaf disease, as presented in Fig. 1.

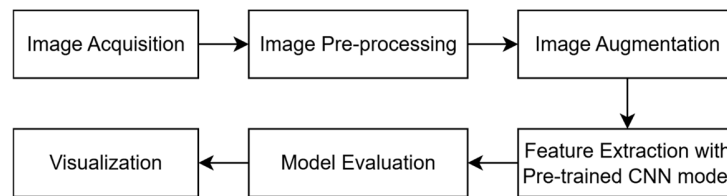


Fig. 1 Framework for rice leaf disease identification using pre-trained model.

A. Image Acquisition

The experiment was conducted using 5932 rice leaf images sourced from Kaggle, which were divided into four categories. These included tungro, blast, brown spot, and BLB, with each image containing only one disease, as shown in Fig. 2 [42]. Training and testing sets of the data were split 90:10 from each other. The images were classified using InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16. Subsequently, dataset was divided into training and testing set using a ratio of 90:10 based on the Pareto principle of 80:20. In this study, 90:10 was selected because training required more data.

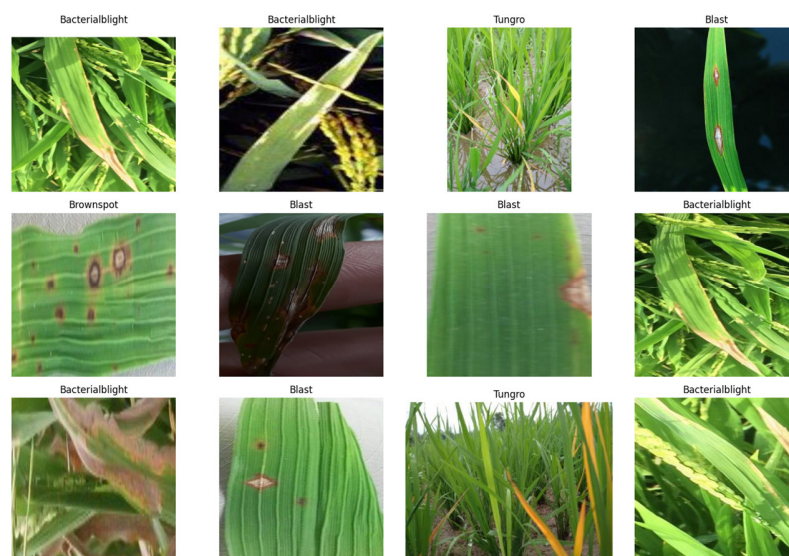


Fig. 2 Samples of rice leaf disease [42]

Deep learning model such as InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16 perform better with larger training sets to learn more robust and generalizable features. Since the dataset used in this study was large, allocating 10% for testing was considered feasible and sufficient to evaluate reliable performance. A function called `tf.keras.preprocessing.image_dataset_from_directory` was used to load and manage images. By the folder names, this

program pulls images from folders, resizes, and labels. The primary settings for loading the data included resizing all images to 150×150 pixels, thereby ensuring a consistent input size. The images obtained were processed in batches of 32 during training to optimize memory usage and computational efficiency. Furthermore, the dataset was shuffled to avoid model learning unintended patterns from the order of the data.

Mathematically, in Eq. (1), the input image data I is resized, where $I_{resized}$ presents the resized image, and the operation ensures all images I conform to the same dimensions of 150×150 pixels.

$$I_{resized} = \text{Resize}(I, (150, 150)) \quad (1)$$

B. Image Pre-Processing

Image pre-processing prepared the data for effective model training. The first step scales the image pixel values to a range from 0 to 1 using Eq. (2).

$$x_{normalized} = \frac{x_{original}}{255} \quad (2)$$

where $x_{original}$ stands for the original pixel values. This scaling helps to make the training process more stable and faster. Additionally, dataset splits into training and validation sets. 10% of the data goes to the validation set. This split allows to check how well the model works on new data during training.

C. Image Augmentation

Data augmentation methods enhanced the dataset and improved model generalization ability by introducing variability through transformations like rotation, flipping, and scaling. This helped prevent overfitting and allowed the model to perform better on unseen data, particularly when addressing limited or imbalanced dataset. ImageDataGenerator class was used to perform various data augmentation transformations. Images were randomly rotated by approximately 30 degrees, as described by the rotation matrix in Eq. (3). Horizontal and vertical translations were applied in a range of 20% of the image dimensions, while random shear transformations were introduced as represented in Eq. (4). Additionally, random zooming of 20% was applied to further enhance variability. To increase robustness, images were randomly flipped horizontally. Finally, pixels generated in the augmented regions were filled using the nearest pixel values through the fill mode setting.

$$R\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

$$S = \begin{bmatrix} 1 & shear_factor & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

D. Feature Extraction with Pre-Trained CNN Models

Feature extraction was carried out using pre-trained CNN as a basis for transfer learning. The model used comprised InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16, which were pre-trained on extensive dataset, providing a robust set of learned features. As shown in Fig. 3, DenseNet121 connected each layer to others in a feedforward fashion [43]. This suggested that each layer received inputs from all preceding layers and passed output to others through concatenation.

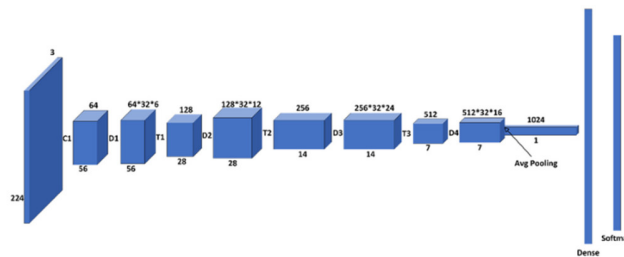


Fig. 3 DenseNet121 Architecture [43]

The convolutional base (*conv_base*) of pre-trained model was frozen by setting `conv_base.trainable = False`. This prevented the weights of the convolutional layers from being updated during training, preserving the learned features from the initial training. The convolutional blocks used in this study were the same as the initial architectures of pre-trained model (DenseNet121). These blocks were not changed but still used in transfer learning mode with frozen weights (`conv_base.trainable = False`) to ensure that the feature extraction ability remained intact. The originality was not in creating a new conv block but in how the architectures were comparatively assessed, improved, and combined with personalized dense layers for the classification of a domain-specific task to overcome rice leaf disease. Custom classification layers were added to model for the final prediction task. Initially, a flattened layer was used to convert the two-dimensional feature maps into one-dimensional vectors. This was followed by a fully connected dense layer with 256 units and a ReLU activation function, as described in Eq. (5), where W represented the weight matrix and b served as the bias term. Finally, the output layer consisted of a dense layer with four units and a softmax activation function for multi-class classification, as shown in Eq. (6), where z_i is the logit or score for class i , and $\sum_j e^{z_j}$ is the sum of the exponentials of the logits across all classes.

$$output = ReLU(W \cdot input + b) \quad (5)$$

$$\hat{y}_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (6)$$

The model is compiled with the Adam optimizer (learning rate of 0.001) and categorical cross-entropy loss function in Eq. (7).

$$L = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (7)$$

where y_i is the true label and $\hat{y}_i$ is the predicted probability for class i . Accuracy is used as the evaluation metric in Eq. (8).

$$Accuracy = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \quad (8)$$

Compared to DenseNet121, InceptionResNetV2 integrates multi-scale to extract the feature and allows layers to be skipped. Meanwhile, MobileNet splits the convolution operation into two, reducing the number of parameters significantly, making processing lighter and faster. VGG16 consists of only 3x3 convolution layers in a row, followed by a pooling layer. This model has many parameters, which leads to high memory usage and slower computation. VGG16 is still frequently used due to the stability and good performance in feature extraction.

In this study, the experiment was conducted on a high-performance computational platform with an NVIDIA GPU to accelerate the training process. The proposed model was implemented using TensorFlow and Keras, with Python as the programming language. Rice leaf disease dataset was split into training (90%) and test (10%) sets. The training process used data augmentation methods to prevent overfitting, and early stopping was applied to monitor validation loss and halt training when performance plateaued. Model hyperparameters and training configuration are shown in Table 1.

TABLE 1
MODEL HYPERPARAMETERS AND TRAINING CONFIGURATION

Parameter	Value
Batch size	32
Image size	150x150
Epochs	10
Optimizer	Adam
Learning rate	0.001
Loss function	Categorical Crossentropy
Validation split	10%
Metrics	Accuracy

E. Model Evaluation

The validation dataset was used to assess model performance by determining the loss and accuracy with the *model.evaluate* function. Performance metrics were also used, including precision, recall, and F1-score from scikit-

learn's *classification_report*. The confusion matrix using *confusion_matrix* provided a detailed breakdown of how model classified across different classes.

Evaluation of the suggested model was carried out with basic metrics such as the accuracy, precision, recall, F1 score, and confusion matrix. Precision corresponded to true positive outcomes divided by the total predicted positive cases. Accuracy captured the fraction of images that were correctly classified, and recall was the fraction of true positive cases. F1 score served as the harmonic mean of precision and recall. All of these measures became increasingly important for a multi-class classification, particularly in medical diagnoses, due to the asymmetrical costs of false alarms and misses. Furthermore, a confusion matrix was constructed to analyze and visualize class-wise distributions of the predictions. These methods of evaluation were based on the definitions provided in Scikit-learn library [44]. Accuracy was defined as the total proportion of images that were correctly classified. Precision, for every disease class, represented the ratio of true positives to the total of true and false positives. Recall represented the proportion of true positives relative to the sum of true positives and false negatives. F1 score, the harmonic mean of Precision and Recall, provided a balanced measure of both metrics. Confusion Matrix provided a comprehensive overview of the classification outcomes that display each class's true positives, false positives, true negatives, and false negatives. Finally, Training and Validation Loss/Accuracy track the model's accuracy and loss evolution over epochs, which were visualized to ensure a proper learning curve.

F. Visualization

Visualization methods were used to interpret and present model performance effectively. These methods provide a better understanding of model prediction behavior across many classes, assist in spotting trends during training, and identify possible problems like overfitting. Therefore, useful insights into the strengths, weaknesses, and reliability of model could be obtained by viewing confusion matrices, accuracy/loss curves, and sample forecasts. Training and validation curves were plotted for both accuracy and loss to observe the learning process and detect potential issues such as overfitting. A heatmap of the confusion matrix was shown using seaborn, offering a visual representation of classification performance across different classes. Model architecture was visualized using the *plot_model* function, providing a graphical overview of the network structure, including layers and their connections. Finally, sample predictions are shown alongside true labels to qualitatively assess model ability to generalize and perform well on new data.

IV. RESULTS

The results provide the experimental data of the proposed method described in Chapter 3 based on InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16-based CNN for classifying rice leaf disease. This section started by explaining the experimental configuration and assessing results in terms of performance metrics. The performance of different model was also compared, concluding with a discussion about how well the model might work on new, unseen data, along with the potential limitations.

This study used four advanced pre-trained CNN models, InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16, to identify the most effective. The experiment used accuracy (%) and training time as evaluation metrics, with the results shown in Table 3.

TABLE 3
EXPERIMENTS WITH VARIOUS CNN MODELS

CNN Model	Accuracy	Training Time
InceptionResNetV2	96.45	00:08:53
MobileNetV2	97.10	00:07:48
DenseNet121	99.10	00:09:27
VGG16	96.11	00:09:49

Table 3 compares InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16, based on their accuracy and training time. In line with the results, DenseNet121 achieved an accuracy of 99%, although the complexity led to longer training time. Meanwhile, MobileNetV2 was recognized for efficiency, reaching an accuracy of 97% while requiring the shortest training time. InceptionResNetV2 delivered results with an accuracy of 96%, coupled with a training time. Despite the accuracy rate of 96%, VGG16 demanded extensive training time due to high computational requirements. These results varied significantly, showing the balance between reaching precision and handling training effectiveness.

The evaluation metrics offered an in-depth knowledge of model capacity to categorize every rice leaf disease accurately. Fig. 4 presents InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16 model. The confusion matrix provides a detailed view of model performance across the four classes of rice leaf disease. Table 2 shows the details of each model's performance metrics. Most classes were classified with high accuracy, with minor

misclassifications between visually similar disease. Figure 5 shows the accuracy and loss trends across ten epochs, indicating an enhancement in both metrics without signs of overfitting.

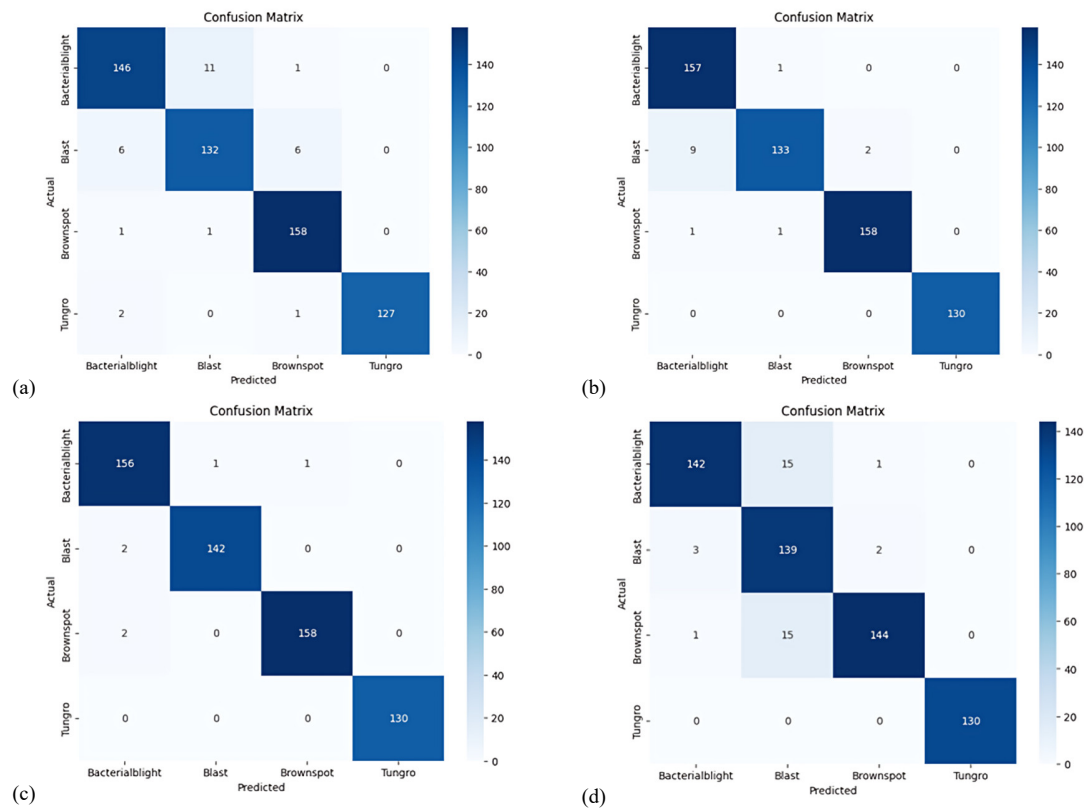


Fig. 4 Illustration of the confusion matrix of (a) InceptionResNetV2; (b) MobileNetV2; (c) DenseNet121; (d) VGG16

TABLE 2
 THE DETAILS OF EACH MODEL'S PERFORMANCE METRICS

CNN Model	Precision	Recall	F1-Score	Support
Leaf Disease				
Inception-ResNetV2				
BLB	0.97	0.90	0.93	158
Blast	0.90	0.92	0.91	144
Brownspot	0.94	0.99	0.96	160
Tungro	1.00	1.00	1.00	130
MobileNetV2				
BLB	0.98	0.94	0.96	158
Blast	0.90	0.98	0.94	144
Brownspot	0.99	0.95	0.97	160
Tungro	1.00	1.00	1.00	130
DenseNet121				
BLB	0.99	0.99	0.99	158
Blast	0.99	0.97	0.98	144
Brownspot	0.99	1.00	0.98	160
Tungro	1.00	1.00	0.99	130
VGG16				
BLB	0.96	0.97	0.97	158
Blast	0.93	0.94	0.93	134
Brownspot	0.97	0.93	0.95	160
Tungro	0.98	1.00	0.99	130

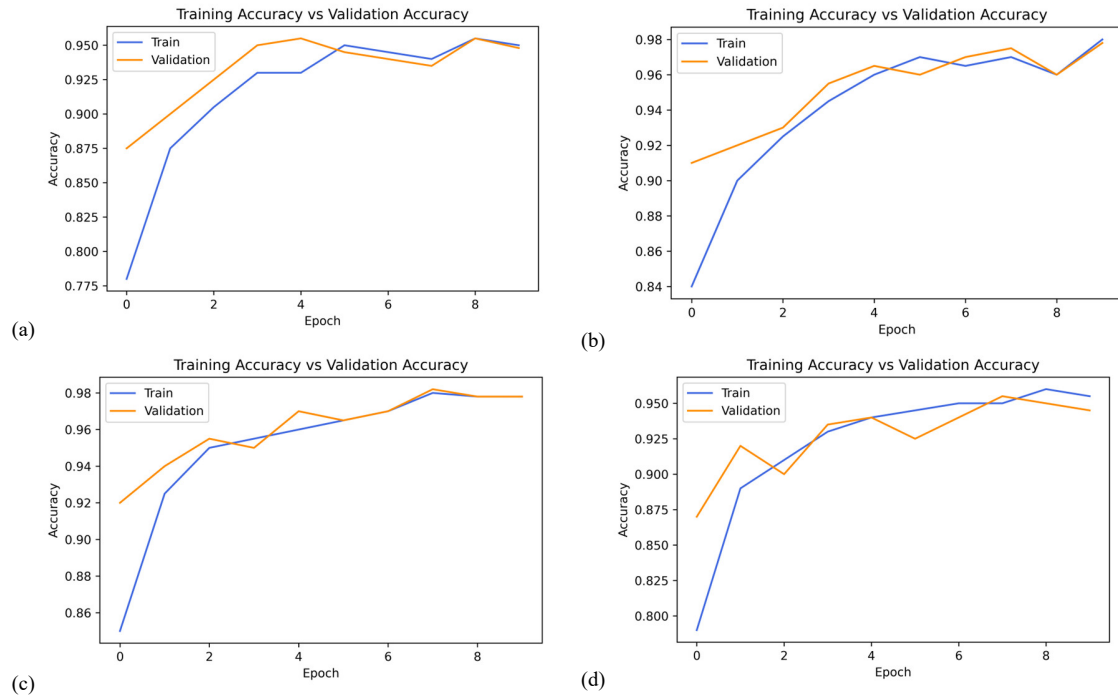


Fig. 5 Training and Validation Accuracy Curves of (a) InceptionResNetV2; (b) MobileNetV2; (c) DenseNet121; (d) VGG16

The experiment was carried out to evaluate the accuracy of a category classification system by showing random test images next to their predicted and actual class labels in a grid layout. This consisted of 24 images, each shown with color-coded titles indicating correct (green) and incorrect (red) predictions. The objective was to present an easy-to-understand display of model precision to help distinguish between accurate and incorrectly classified samples in a brief and informative pattern, as shown in Fig. 6 [42].

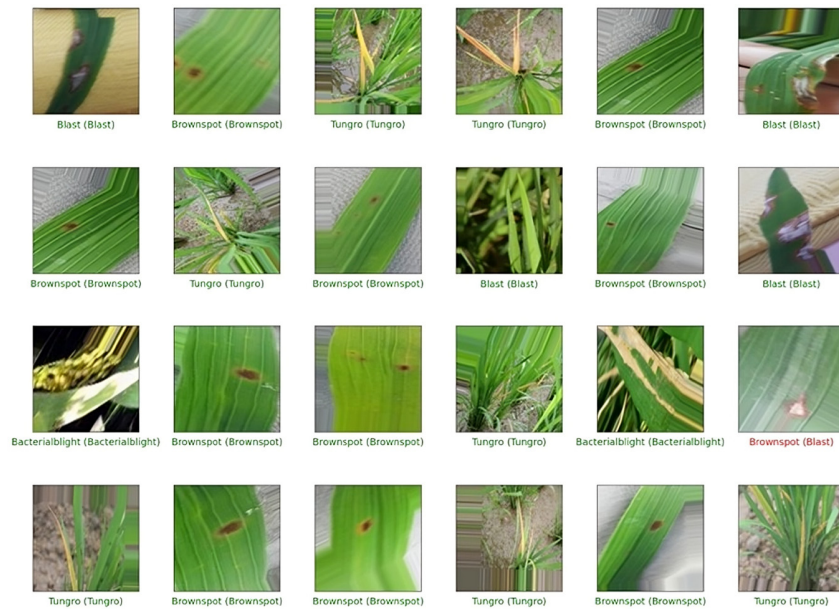


Fig. 6 Random Test Images [42]

The image shows a grid displaying 24 rice leaf disease classification model test samples. Each tile consists of image showing a disease leaf with a heading that carries the predicted label and the true one. The predicted label gets the upper hand, while the true label is in brackets. Any title in green font indicates that the prediction was correct, suggesting appropriate allocation of label. For instance, in the top-left tile, model could correctly predict that the true label provided was Brownspot, which led to the green title 'Brownspot (Brownspot)'. The red titles indicated the incorrect class, showing that model output did not match the true label. For example, red title like 'Tungro (Brownspot)' indicates that the model incorrectly predicted 'Tungro' when the true label was 'Brownspot'. This color-coding system resolves queries adequately, where correct and incorrect predictions are placed green and red, respectively.

V. DISCUSSION

The performance of baseline model trained from scratch, such as InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16, was compared using the same dataset and experimental setup to assess the efficacy of the suggested method. DenseNet121 had the best classification performance and the highest accuracy (99.10%) but required a longer time (9:27), as indicated in Tables 2 and 3. With the lowest training time (7:48) and the highest accuracy (97.10%), MobileNetV2 showed outstanding efficiency. VGG16, an older architecture, showed lower efficiency and the longest training time (9:49), while InceptionResNetV2 provided a good trade-off between computational cost and performance. These results showed how transfer learning performed well with contemporary CNN architectures to achieve accuracy and efficiency, as presented in Table 4.

TABLE 4
 COMPARATIVE PERFORMANCE OF BASELINE AND PROPOSED MODEL

Author	Method	Accuracy
Dogra et al. [31]	CNN-VGG19	93.00%
Mannepalli et al. [20]	CNN-VGG16	97.77%
Narasimha et al. [30]	InceptionResNetV2	95.67%
Shrivastava and Pradhan [32]	Color cue-based method with SVM classifier	94.65%
Chen et al. [39]	VGGNet and Inception	92.00%
Proposed Method	DenseNet121 with conv base	99.10%

Table 4 shows the analysis of image classification methods for rice disease detection. Based on the results, the proposed method with DenseNet121 and convolutional base offered better feature extraction and classification, achieving a higher accuracy of 99.10% among the classification methods. A high accuracy of 97.77% was recorded by CNN-VGG16 model [20], 95.67% by InceptionResNetV2 [30], and 94.65% by the color cue-based SVM method [32]. The hybrid model combining VGGNet and Inception networks produced 92.00% accuracy [39], and CNN-VGG19 method achieved 93.00% [31]. The results emphasized the performance that advanced deep learning architecture, and particularly DenseNet121, could achieve in terms of classification accuracy.

In a recent report, Rahman et al. [45] introduced an explainable deep transfer learning ensemble to rice leaf disease diagnosis. However, this study did not develop ensemble model, but compared the individual capabilities of the modern CNN architectures (InceptionResNetV2, MobileNetV2, DenseNet121, and VGG16) with the unified transfer learning pipeline. The results showed that DenseNet121, with dense connectivity and a lower parameter count, was the best performer among others, regarding accuracy and efficiency. A simple framework that could be run on limited computational environments without the loss of accuracy was also developed, which was important for the application in agricultural areas with fewer resources.

Despite the significant contribution, this study has several contributions. First, the comprehensive analysis of four pre-trained CNN models for the multi-class classification of rice leaf disease is a novel contribution. Second, the analysis shows the powerful performance of DenseNet121, which tends to do well in the same training conditions due to dense connectivity and feature reuse. The performance is mainly due to the transfer learning used, along with data augmentation and the custom dense classification layers, which increase feature extraction and generalization. Lastly, the results offer a foundation that can be advanced towards the design of automated disease detection for use in various agricultural technology.

After conducting the comparison shown in Table 4, this study presents several limitations, particularly regarding the variability in the image dataset. Factors such as lighting conditions, different stages of disease progression, and leaf damage unrelated to disease are not fully represented. There is no real-world validation, and the evaluation is only performed on the testing part of the dataset used. Therefore, future studies should include testing in different real-life conditions to find the applicability of this model. The excellent classification accuracy, more hyperparameter adjustment is also required to enhance functionality.

VI. CONCLUSIONS

In conclusion, this study shows the efficiency of pre-trained CNN models in rice leaf disease identification. Among the four models evaluated, DenseNet121 shows the best accuracy of 99.10% using CNN power in image classification and feature extraction to enhance model generalization through data augmentation. Furthermore, four kinds of rice leaf disease, including tungro, brown spot, blast, and bacterial leaf blight, are detected by accuracy, precision, recall, and F1-score metrics. The results indicate that this method outperforms previous reports, offering a reliable tool for precision agriculture to support rapid and accurate decision-making in plant disease management. This study has some drawbacks, which include the inability to test under real field conditions and the limited representation of dataset. Therefore, future investigations are recommended to evaluate this model under different real-world conditions and optimize hyperparameters.

Author Contributions: *Bety Wulan Sari:* Conceptualization, Methodology, Writing - Original Draft. *Donni Prabowo:* Software, Methodology, Review & Editing. *Yoga Pristyanto:* Review & Editing, Supervision. *Afrig Aminuddin:* Review & Editing, Supervision.

All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency.

Acknowledgments: This research is supported by Department of Information System, Faculty of Computer Science, Universitas AMIKOM Yogyakarta

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: The data used is open to the public and available at this following address: <https://www.kaggle.com/datasets/maimunulkjisan/rice-leaf-dataset-from-mendeley-data>.

Informed Consent: There were no human subjects.

Institutional Review Board Statement: Not applicable.

Animal Subjects: There were no animal subjects.

ORCID:

Bety Wulan Sari: <https://orcid.org/0009-0001-5275-7553>

Donni Prabowo: <https://orcid.org/0009-0006-8078-589X>

Yoga Pristyanto: <https://orcid.org/0000-0002-1358-4864>

Afrig Aminuddin: <https://orcid.org/0000-0003-0963-5419>

REFERENCES

- [1] M. R. Jamal, P. Kristiansen, M. J. Kabir, and L. Lobry de Bruyn, "Challenges and Adaptations for Resilient Rice Production under Changing Environments in Bangladesh," *Land*, vol. 12, no. 6, p. 1217, Jun. 2023, doi: 10.3390/land12061217.
- [2] M. Huang, "The decreasing area of hybrid rice production in China: causes and potential effects on Chinese rice self-sufficiency," *Food Secur.*, vol. 14, no. 1, pp. 267–272, Feb. 2022, doi: 10.1007/s12571-021-01199-z.
- [3] H. Pallathadka et al., "Application of machine learning techniques in rice leaf disease detection," in *Materials Today: Proceedings*, 2022. doi: 10.1016/j.matpr.2021.11.398.
- [4] S. Singh, S. Chand, N. K. Singh, and T. R. Sharma, "Genome-Wide Distribution, Organisation and Functional Characterization of Disease Resistance and Defence Response Genes across Rice Species," *PLoS One*, vol. 10, no. 4, p. e0125964, Apr. 2015, doi: 10.1371/journal.pone.0125964.
- [5] C. Ma et al., "Nanomaterials in agricultural soils: Ecotoxicity and application," *Curr. Opin. Environ. Sci. Heal.*, vol. 31, p. 100432, Feb. 2023, doi: 10.1016/j.coesh.2022.100432.
- [6] G. Agrios, "Plant Pathology 5th Edition," San Diego Acad. Press, 2005.
- [7] H. Bunawan, L. Dusik, S. N. Bunawan, and N. M. Amin, "Rice tungro disease: From identification to disease control," *World Appl. Sci. J.*, vol. 31, no. 6, 2014, doi: 10.5829/idosi.wasj.2014.31.06.610.
- [8] J. Shim et al., "Rice tungro spherical virus resistance into photoperiod-insensitive japonica rice by marker-assisted selection," 2015. doi: 10.1270/jsbbs.65.345.

- [9] V. de P. Bigirimana, G. K. H. Hua, O. I. Nyamangyoku, and M. Höfte, "Rice sheath rot: An emerging ubiquitous destructive disease complex," 2015. doi: 10.3389/fpls.2015.01066.
- [10] E. Ibrahim, N. Rosida, Khaerana, and M. Haiqal, "Growth and result performance of 10 promising resistant lines of tungro disease in Lanrang South Sulawesi," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1230, no. 1, p. 012132, Sep. 2023, doi: 10.1088/1755-1315/1230/1/012132.
- [11] Y. M. Shabana, G. M. Abdel-Fattah, A. E. Ismail, and Y. M. Rashad, "Control of brown spot pathogen of rice (*Bipolaris oryzae*) using some phenolic antioxidants," *Brazilian J. Microbiol.*, vol. 39, no. 3, 2008, doi: 10.1590/S1517-83822008000300006.
- [12] R. Chhabra, R. Sharma, M. S. Hunjan, V. K. Sharma, P. Sharma, and S. K. Chauhan, "Microstructural and metabolic variations induced by *Bipolaris oryzae* inciting brown spot disease of rice," *Cereal Res. Commun.*, vol. 51, no. 4, 2023, doi: 10.1007/s42976-023-00351-z.
- [13] N. K. Chakrabarti, "Epidemiology and Disease Management of Brown Spot of Rice in India," in *Major Fungal Diseases of Rice*, 2001. doi: 10.1007/978-94-017-2157-8_21.
- [14] T. R. Sharma, A. K. Rai, S. K. Gupta, V. Vijayan, B. N. Devanna, and S. Ray, "Rice Blast Management Through Host-Plant Resistance: Retrospect and Prospects," 2012. doi: 10.1007/s40003-011-0003-5.
- [15] M. K. Barnwal et al., "A review on crop losses, epidemiology and disease management of rice brown spot to identify research priorities and knowledge gaps," 2013. doi: 10.1007/s10658-013-0195-6.
- [16] M. U. Younas et al., "Approaches to Reduce Rice Blast Disease Using Knowledge from Host Resistance and Pathogen Pathogenicity," 2023. doi: 10.3390/ijms24054985.
- [17] L. Nalley, F. Tsioboe, A. Durand-Morat, A. Shew, and G. Thoma, "Economic and environmental impact of rice blast pathogen (*Magnaporthe oryzae*) alleviation in the United States," *PLoS One*, vol. 11, no. 12, 2016, doi: 10.1371/journal.pone.0167295.
- [18] H. L. Chen, B. T. Chen, D. P. Zhang, Y. F. Xie, and Q. Zhang, "Pathotypes of *Pyricularia grisea* in rice fields of central and southern China," *Plant Dis.*, vol. 85, no. 8, 2001, doi: 10.1094/PDIS.2001.85.8.843.
- [19] J. M. Bonman, "Leaf and Neck Blast Resistance in Tropical Lowland Rice Cultivars," *Plant Dis.*, vol. 73, no. 5, 1989, doi: 10.1094/pd-73-0388.
- [20] P. K. Mannepal, A. Pathre, G. Chhabra, P. A. Ujjainkar, and S. Wanjari, "Diagnosis of bacterial leaf blight, leaf smut, and brown spot in rice leaves using VGG16," in *Procedia Computer Science*, Elsevier B.V., 2024, pp. 193–200. doi: 10.1016/j.procs.2024.04.022.
- [21] C. G. Simhadri, H. K. Kondaveeti, V. K. Vatsavayi, A. Mitra, and P. Ananthachari, "Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques," *Inf. Process. Agric.*, vol. 12, no. 2, pp. 151–168, Jun. 2025, doi: 10.1016/j.inpa.2024.04.006.
- [22] H. Pallathadka et al., "Application of machine learning techniques in rice leaf disease detection," *Mater. Today Proc.*, vol. 51, pp. 2277–2280, 2022, doi: 10.1016/j.matpr.2021.11.398.
- [23] T. Kavzoglu and A. Teke, "Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost)," *Bull. Eng. Geol. Environ.*, vol. 81, no. 5, p. 201, May 2022, doi: 10.1007/s10064-022-02708-w.
- [24] H. Haapala, L. Pesonen, and P. Nurkka, "Usability as a Challenge in Precision Agriculture – Case Study: an ISOBUS VT," *Agric. Eng. Int. CIGR e-journal*, vol. VIII, 2006.
- [25] A. Nanavaty et al., "Integrating deep learning for visual question answering in Agricultural Disease Diagnostics: Case Study of Wheat Rust," *Sci. Rep.*, vol. 14, no. 1, p. 28203, Nov. 2024, doi: 10.1038/s41598-024-79793-2.
- [26] R. Ahmed, "Leveraging Convolutional Neural Network and Transfer Learning for Cotton Plant and Leaf Disease Recognition," *Int. J. Image. Graph. Signal Process.*, vol. 13, no. 4, pp. 47–62, 2021, doi: 10.5815/ijigsp.2021.04.04.
- [27] A. W. Salehi et al., "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope," *Sustainability*, vol. 15, no. 7, p. 5930, Mar. 2023, doi: 10.3390/su15075930.
- [28] S. Asif, W. Yi, Q. U. Ain, J. Hou, T. Yi, and J. Si, "Improving Effectiveness of Different Deep Transfer Learning-Based Models for Detecting Brain Tumors From MR Images," *IEEE Access*, vol. 10, pp. 34716–34730, 2022, doi: 10.1109/ACCESS.2022.3153306.
- [29] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach," *Electronics*, vol. 10, no. 12, p. 1388, Jun. 2021, doi: 10.3390/electronics10121388.
- [30] K. N. L. V. Narasimha Prasad, C. S. Pavan Kumar, B. Subedi, H. B. Abbraha, and V. E. Sathishkumar, "Rice leaf diseases prediction using deep neural networks with transfer learning," *Environ. Res.*, vol. 198, 2021, doi: 10.1016/j.envres.2021.111275.
- [31] R. Dogra, S. Rani, A. Singh, M. A. Albahar, A. E. Barrera, and A. Alkhayyat, "Deep learning model for detection of brown spot rice leaf disease with smart agriculture," *Comput. Electr. Eng.*, vol. 109, 2023, doi: 10.1016/j.compeleceng.2023.108659.
- [32] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: a machine learning paradigm," *J. Plant Pathol.*, vol. 103, no. 1, 2021, doi: 10.1007/s42161-020-00683-3.
- [33] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, 2022, doi: 10.1016/j.gltp.2022.04.020.
- [34] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, p. 46, Feb. 2023, doi: 10.3390/jimaging9020046.
- [35] B. Neupane, T. Horanont, and J. Aryal, "Real-Time Vehicle Classification and Tracking Using a Transfer Learning-Improved Deep Learning Network," *Sensors*, vol. 22, no. 10, p. 3813, May 2022, doi: 10.3390/s22103813.
- [36] S. Faisal et al., "Deep Transfer Learning Based Detection and Classification of Citrus Plant Diseases," *Comput. Mater. Contin.*, vol. 76, no. 1, pp. 895–914, 2023, doi: 10.32604/cmc.2023.039781.
- [37] A. Arya and P. K. Mishra, "MobileNetV2-Incep-M: a hybrid lightweight model for the classification of rice plant diseases," *Multimed. Tools Appl.*, vol. 83, no. 33, pp. 79117–79144, Mar. 2024, doi: 10.1007/s11042-024-18723-w.
- [38] V. Sahasranamam, T. Ramesh, R. Rajeswari, A. Karthikkumar, and D. Muthumanickam, "AI-Driven Paddy Leaf Disease Classification and Prediction using DenseNet-121," 2025, pp. 272–292. doi: 10.2991/978-94-6463-740-3_24.
- [39] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," *Comput. Electron. Agric.*, vol. 173, p. 105393, Jun. 2020, doi: 10.1016/j.compag.2020.105393.

- [40] B. S. Bari et al., "A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework," *PeerJ Comput. Sci.*, vol. 7, 2021, doi: 10.7717/PEERJ-CS.432.
- [41] N. Nandhini and R. Bhavani, "Feature extraction for diseased leaf image classification using machine learning," in 2020 International Conference on Computer Communication and Informatics, ICCCI 2020, 2020. doi: 10.1109/ICCCI48352.2020.9104203.
- [42] "Rice Leafs Disease Dataset," <https://www.kaggle.com/datasets/maimunulkjisan/rice-leaf-dataset-from-mendeley-data>. Accessed: Aug. 31, 2024. [Online]. Available: <https://www.kaggle.com/datasets/maimunulkjisan/rice-leaf-dataset-from-mendeley-data>
- [43] R. S. Tiwari, L. Dandabani, T. K. Das, S. B. Khan, S. Basheer, and M. S. Alqahtani, "Cloud-Based Quad Deep Ensemble Framework for the Detection of COVID-19 Omicron and Delta Variants," *Diagnostics*, vol. 13, no. 22, 2023, doi: 10.3390/diagnostics13223419.
- [44] F. Pedregosa FABIANPEDREGOSA et al., "Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot," 2011. [Online]. Available: <http://scikit-learn.sourceforge.net>.
- [45] M. M. Rahman et al., "Explainable Deep Transfer Learning Framework for Rice Leaf Disease Diagnosis and Classification," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 12, pp. 862–884, 2024, doi: 10.14569/IJACSA.2024.0151287

Publisher’s Note: Publisher stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.