JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Deep Transfer Learning with MobileNetV2 for Automated Classification of Healthy and Diseased Rice Plant Leaves

[1] S. KOKILA, [2] DR. S. ABIRAMI, [3] DR. D. NAGARAJU
[1] Research Scholar, [2]Assistant Professor, [3] Professor
[1,2] Department of CSE, Annamalai University, Tamil Nadu,
[3] Department of CSE, Sri Venkateswara Perumal College of engineering &Technology, Andhra Pradesh

Abstract: India is one of the leading productions of Paddy. Compared to previous year Gross Domestic Product (GDP) Export rate of Paddy in the year 2021 has increased to around 33%. Paddy is the Major food production crop in India. Every crop is prone to many diseases throughout their lifespan. The disease can affect the crop at any stage of their growing phase. Early detection of disease is the only solution to reduce the damage. Early detection of plant diseases is vital for sustainable agriculture. In this study, we propose a deep learning–based approach for classifying healthy and diseased rice leaves using MobileNetV2 with transfer learning. A dataset of 2,314 rice leaf images (2,126 for training, 188 for testing) was used. Data augmentation and class weighting techniques were applied to address dataset imbalance. The proposed model achieved 96% accuracy, 0.96 F1-score, and a ROC AUC of 0.996 on the test set, demonstrating strong generalization ability. These results suggest that lightweight transfer learning models such as MobileNetV2 can provide scalable and accurate disease diagnosis solutions for precision agriculture.

IndexTerms - Plant disease detection, Rice leaf, MobileNetV2, Transfer learning, Deep learning, Precision agriculture

I. Introduction

India is home to some of the largest rice fields in the world and is widely recognized for its vast rice cultivation. Rice farming plays a vital role in shaping the country's economy. About 60% of agricultural land is dedicated to paddy farming. However, overall yield declines due to damage caused by both abiotic and biotic factors. The infection usually begins on the leaves and gradually spreads to other parts of the crop across successive seasons. Environmental fluctuations play a significant role in triggering these infections. Among the primary agents of disease transmission are fungi, bacteria, and viruses. Early detection of plant diseases is essential to minimize their negative impact on crop productivity. In [1] proposed an advanced method for rice disease identification using a computer vision–based system. The study incorporated various modern techniques such as neural networks, machine learning, and image processing to effectively recognize and classify plant disorders. [2] This study introduces a visual approach for identifying various rice diseases by analyzing the unique texture patterns of leaves. The primary aim is to demonstrate how image processing techniques can be applied to classify plant diseases effectively. A major limitation of current classification systems is their poor generalizability, as their performance is often restricted to the specific datasets used during training [3]. For instance, although a rice disease recognition model initially achieved 90% accuracy, its performance significantly declined when tested under different conditions, raising concerns about its reliability. Such variability greatly reduces the precision of results. At present, there remains a clear need for a robust deep learning–based approach that can accurately detect and classify multiple rice diseases across diverse conditions. Disease identification plays a vital role in agricultural research, as it supports precise diagnosis and effective monitoring of crop health in the field. [4] This study aims to improve the performance of image-based detection methods by focusing on the classification of different rice leaf diseases. In [5] A machine learning framework was developed using a Convolutional Neural Network (CNN) to identify two of the most common rice diseases—Rice Blast and Bacterial Blight—along with healthy leaf samples. In addition to disease detection, the system incorporated a feature that recommends suitable pesticides or insecticides based on the identified condition. The model was subsequently trained and evaluated to assess its effectiveness.

Rice is one of the most important staple food crops worldwide, serving as the primary dietary source for billions of people. However, its productivity is often threatened by various diseases that can cause substantial yield losses, thereby impacting both food security and the agricultural economy [6]. Traditional methods of disease detection typically rely on manual observation by farmers or experts, which is not only time-consuming but also prone to human error and subjectivity.With recent advancements in artificial intelligence, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated significant potential in the field of plant disease recognition due to their strong feature extraction and classification capabilities [7]. Among these models, MobileNetV2 stands out as a lightweight yet powerful architecture, optimized for computational efficiency without compromising accuracy. Its design makes it especially well-suited for deployment on mobile and edge devices, offering practical solutions for real-time disease detection in resource-constrained agricultural settings.

In light of these advantages, this study explores the application of MobileNetV2-based transfer learning for the classification of rice leaf health status. The goal is to evaluate its effectiveness in distinguishing between healthy and diseased leaves, while also highlighting its potential for scalable, field-ready disease monitoring systems that support precision agriculture.

1.1. Motivation

Rice is a staple food for a large portion of the global population, and its productivity is often threatened by various plant diseases. Traditional disease detection methods are time-consuming, labour-intensive, and often require expert knowledge, making them less practical for large-scale farming. With the rapid growth of precision agriculture, there is a strong need for automated, accurate, and scalable solutions that can enable farmers to detect plant diseases at an early stage and reduce crop losses. Deep learning, particularly lightweight models, offers an efficient way to address this challenge while being deployable on resource-constrained devices.

1.2. Objective

The primary objective of this study is to develop and evaluate a deep learning–based framework for rice leaf disease classification using MobileNetV2 with transfer learning. Specifically, the work aims to:

1. Build a reliable classification model that distinguishes between healthy and diseased rice leaves.
2. Apply data augmentation and class weighting to handle dataset imbalance and improve generalization.
3. Assess the performance of the proposed model through accuracy, F1-score, and ROC AUC metrics.
4. Demonstrate the potential of lightweight deep learning models as scalable, accurate, and efficient tools for precision agriculture and early disease detection.

II. RELATED RESEARCH

Existing works employ CNNs like VGG, ResNet, Inception for plant disease detection. Most use large, computationally heavy models. Few explore lightweight architectures such as MobileNetV2 for real-world agricultural deployment.

In order to highlight the benefits of hybrid designs, this [8] study compared three models for the categorization of plant diseases. Using the EfficientNetB0 backbone, the Hybrid Model outperformed the CNN and KNN classifiers by achieving 99.77% accuracy. Its resilience and dependability are demonstrated by its nearly flawless precision, recall, and F1-scores across all classes.

The two-stage convolutional neural network architecture was proposed in [9], compared the outcomes with CNN classifiers such as NasNet Mobile, MobileNet, and SqueezeNet. The best accuracy of 93.3% was attained by this model. 1426 photos of both healthy and diseased rice plant leaves that were gathered from actual fields were used. In [10] suggested the Deep Neural Network-based model using a transfer learning methodology. One of the Convolutional Neural Network models, InceptionResNetV2, uses a transfer learning technique to automatically identify rice leaf disease. The highest accuracy of 95.67% was attained by their model. A few photos from the internet and the dataset were gathered via Kaggle. The crucial challenge of early detection and management of rice leaf diseases in natural settings was the main emphasis [11]. They introduced rE-

GoogLeNet, a CNN model based on the Google Net architecture, to address this issue. Their results showed that rE-GoogLeNet performed better than both conventional and sophisticated multiscale models, averaging an astounding 99.58% accuracy rate. This study demonstrated the efficacy of this model for rice leaf disease identification and control, showing a noteworthy 1.72% improvement over the original Google Net model. [12] compared the effectiveness of many CNN paradigms for identifying and localizing rice illnesses, incorporating elements of ensemble models and transfer learning. DenseNet121 and Inceptionv3 were two of the six CNN-based deep learning architectures that were assessed in their study. Additionally, they included a new ensemble model and applied transfer learning to a number of these models. In addition to focusing on categorizing rice leaf illnesses, this [13] study could be very helpful to farmers and the agricultural community as a whole. This study opens the door for technologically driven improvements in agricultural operations by demonstrating the benefits of customized CNN models for accurate and efficient rice leaf disease categorization.

III. MATERIALS AND METHDOS

The goal of the endeavor is to accurately diagnose diseases at an early stage. Data collection is the first step. We gathered pictures from the Plant village and Kaggle datasets [14].

3.1. Proposed Methodology

The five steps of the research approach include data collection, pre-processing, transformation, modelling, fine-tuning, evaluation, and outcomes, as seen in Figure 1. Getting a dataset from Kaggle that includes photos gathered from various sources is the first phase, known as data gathering. To guarantee a clean picture dataset, the second step, data pre-processing, entails image reclassification. Data augmentation, image resizing, and dataset separation into train and validation datasets are all part of the third stage, data transformation. In order to prepare the dataset for modelling, it is reshaped and reconfigured during data augmentation. Training and fine-tuning models are part of the fourth step, data modelling. MobilenetV2 has been utilized. The performance of each deep learning model is assessed based on its accuracy, precision, loss, and size in the fifth stage, Evaluation and Results.

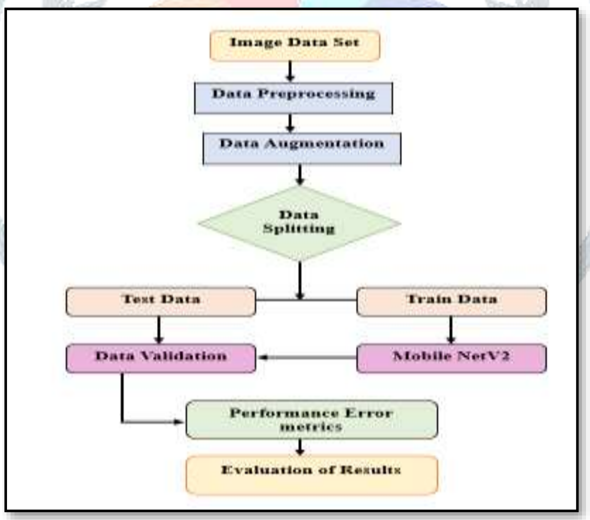

Figure 1: Proposed Methodology

3.2. Dataset Description

This is a rice leaf disease dataset collected via the internet and independently. The dataset which totals 2627 images consists of 6 rice leaf diseases in the train and validation folders. This dataset includes:

- Bacterial Leaf Blight
- Brown Spot
- Healthy
- Leaf Blast
- Leaf Scald
- Narrow Brown Spot

3.3. Data Preprocessing: This is the initial stage of data processing. The information gathered from various sources appears disorganized. It is necessary to clean up those data for subsequent processing. Images resized to 224×224 pixels (input size of MobileNetV2). Normalized using preprocess input (scales pixel values appropriately).

3.4. Data Augmentation: We can use data augmentation if there is not enough data to build and develop our model. It is a method for producing more data for model training. Every created image will have the same label as the original image that The model has produced more. To avoid overfitting, training images are randomly transformed: Rotation (±20°), Zoom (up to 20%), Horizontal flip

3.5. MobileNetV2 Transfer Learning

Initially, a MobileNet-V2 model that had already been trained on ImageNet was used for transfer learning. The architecture of the base model was kept, but in order to stop additional training and maintain the previously learnt characteristics, its weights were frozen. The output of the base model was enhanced with custom layers. These layers comprised global average pooling, a dense layer with 512 units and ReLU activation, a second dense layer with 256 units and ReLU activation, dropout with a regularization rate of 0.4, and, lastly, an output layer with a softmax activation function set up to have as many units as the designated number of classes. To facilitate future training of the model on the relevant dataset with the designated optimization and Loss settings, the model was assembled using the Adam optimizer [15], a categorical cross-entropy Loss function (fit for multi-class classification (4)), and Accuracy as the evaluation metric. The final model was an end-to-end neural network for picture categorization that combined custom layers with the foundational MobileNetV2 architecture is shown in figure below.

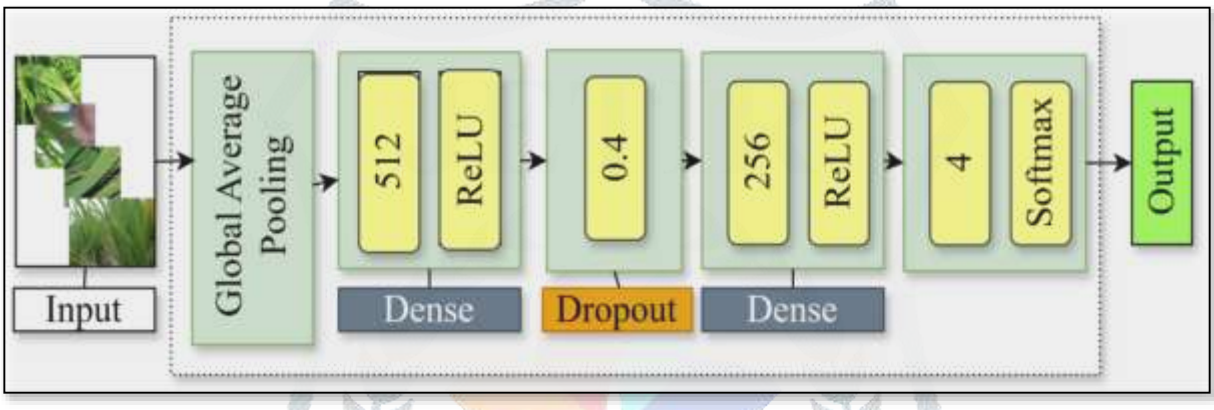

Figure 2: The MobileNetV2 model architecture.

3.6. Performance Error Metrics

Precision: Precision measures how many of the positively predicted instances are actually positive.

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (1)$$

Recall: Actual negative instances to total cases of disease ratio is known as recall.

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (2)$$

F1-Score: The harmonic mean of the model's recall and accuracy is known as the F1 score.

$$\text{F1-Score} = 2\left(\frac{Precision * Recall}{Precision + Recall}\right) \quad (3)$$

Support: The number of instances (samples) of a given class in the dataset.

IV. RESULTS AND ANALYSIS

The aim of this experiment is to classify the rice leaf disease as healthy or unhealthy instead of 4 class classification problem. This will still help farmers to detect if there is disease with higher accuracy.

Table 1: Classification Report

Metrics	Precision	Recall	F1-score	Support
Healthy	0.98	0.94	0.96	95
Unhealthy	0.94	0.98	0.96	93
Accuracy			0.96	188
Macro avg	0.96	0.96	0.96	188
Weighted Avg	0.96	0.96	0.96	188

From the above table 1 the classification report shows that the model performs very well in distinguishing between Healthy and Unhealthy cases. For the Healthy class, it has a precision of 0.98 (meaning almost all predicted Healthy cases are correct) and a recall of 0.94 (it successfully identifies 94% of actual Healthy cases). For the Unhealthy class, precision is 0.94 and recall is 0.98, indicating that most actual Unhealthy cases are correctly detected, though with a few false positives. Both classes have an F1-

score of 0.96, showing a good balance between precision and recall. The overall accuracy is 96% across all 188 samples. Since the dataset is nearly balanced, the macro average and weighted average are also 0.96, confirming consistent performance across both classes.

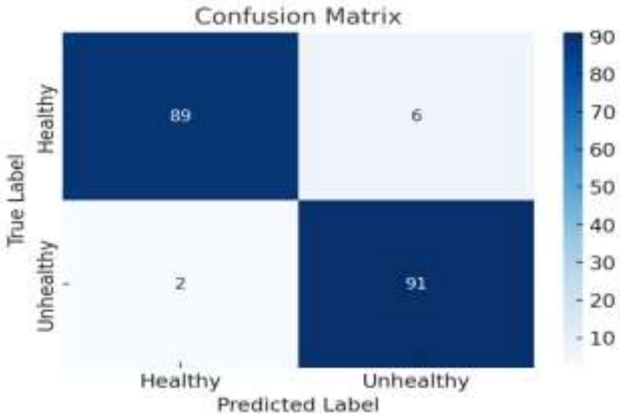

Figure 3: Confusion Matrix

The confusion matrix shows the model's classification performance in detail. Out of 95 actual Healthy cases, the model correctly predicted 89 as Healthy (true positives) but misclassified 6 as Unhealthy (false negatives). For the 93 actual Unhealthy cases, the model correctly identified 91 as Unhealthy (true positives) while misclassifying only 2 as Healthy (false positives). This indicates that the model is highly accurate in detecting both classes, with very few misclassifications. The ROC AUC score of 0.9960 further confirms that the model has excellent discriminative ability, meaning it can almost perfectly separate Healthy and Unhealthy cases.

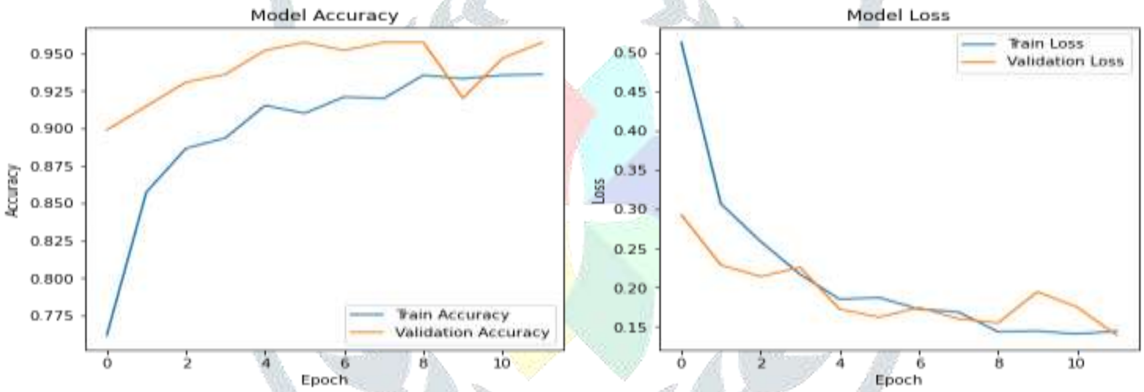

Figure 4: Model accuracy and Model Loss

The training curves show that the model is learning effectively. In the accuracy graph, both training and validation accuracy increase steadily over epochs, with validation accuracy consistently higher and reaching about 95%, indicating strong generalization. In the loss graph, both training and validation loss decrease smoothly, showing the model is reducing errors during learning. Although there are minor fluctuations in validation loss, the overall trend suggests the model is well-trained without signs of severe overfitting, achieving both high accuracy and stable performance.

V. CONCLUSION

The proposed classification model demonstrates robust performance in differentiating between Healthy and Unhealthy cases. The confusion matrix indicates a minimal number of misclassifications, with high recall and precision values achieved across both classes. The overall classification accuracy of 96% and a ROC AUC score of 0.996 highlight the strong discriminative capability of the model. Furthermore, the training and validation learning curves exhibit steady improvements in accuracy alongside decreasing loss values, suggesting effective convergence and generalization with negligible overfitting. Collectively, these findings confirm that the model is both reliable and efficient, underscoring its potential applicability in real-world diagnostic or predictive scenarios.

References:

1. Tian, H., Wang, T., Liu, Y., Qiao, X., Li, Y. (2020). Computer vision technology in agricultural automation—A review. Information Processing in Agriculture, 7(1): 1-19. https://doi.org/10.1016/j.inpa.2019.09.006.

2. Habib, M.T., Majumder, A., Jakaria, A.Z.M., Akter, M.,Uddin, M.S., Ahmed, F. (2020). Machine vision based papaya disease recognition. Journal of King Saud University-Computer and Information Sciences, 32(3):300-309. https://doi.org/10.1016/j.jksuci.2018.06.006
3. Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R.,Kateris, D., Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. Sensors,21(11): 3758. https://doi.org/10.3390/s21113758
4. Nigam, S., Jain, R. (2019). Plant disease identification using Deep Learning: A review. Indian Journal of Agricultural Sciences, 90(2): 249-257. https://doi.org/10.56093/ijas.v90i2.98996.
5. Eunice, J., Popescu, D.E., Chowdary, M.K., Hemanth, J.(2022). Deep learning-based leaf disease detection in crops using images for agricultural applications.Agronomy, 12(10): 2395.https://doi.org/10.3390/agronomy12102395.
6. Too, E. C., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, 161, 272–279.
7. Kokila, M., Abirami, D., & Nagaraju, D. (2025). Fast Mask Region And Entropy Based Histogram Equalized Segmentation Of Rice Plant Diseases. *Journal of Theoretical and Applied Information Technology*, *103*(4).
8. Sukhwinder Kaur, Saurabh Sharma (2025) " A Comparative Study of Plant Disease Classification Using a Hybrid EfficientNetB0-Based Deep Learning Model" Journal of Emerging Technologies and Innovative Research (JETIR) Volume 12, Issue 1.
9. Yang Lua, Shujuan Yi, Nianyin Zeng , Yurong Liu, Yong Zhang (2017),"Identification of rice diseases using deep convolutional neural networks",Neurocomputing, Elsevier.
10. Krishnamoorthy N a , L.V. Narasimha Prasad (2021), "Rice leaf diseasesprediction using deep neural networks with transfer learning,Environmental Research, 2017 Elsevier.
11. Yang, L.; Yu, X.; Zhang, S.; Long, H.; Zhang, H.; Xu, S.; Liao, Y. (2023) "GoogLeNet based on residual network and attention mechanism identification of rice leaf diseases". Comput. Electron. Agric., 204, 107543.
12. Ahad, M.T.; Li, Y.; Song, B.; Bhuiyan,(2023) "T. Comparison of CNN-based deep learning architectures for rice diseases classification" .Artif. Intell. Agric. 2023, 9, 22–35.
13. Dutta, M.; Islam Sujan, M.R.;Mojumdar, M.U.; Chakraborty, N.R.;Marouf, A.A.; Rokne, J.G.; Alhajj, R.Rice Leaf Disease Classification (2024) "A Comparative Approach Using Convolutional Neural Network (CNN), Cascading Autoencoder with Attention Residual U-Net (CAAR-U-Net), and MobileNet-V2 Architectures. Technologies, 12, 214. https://doi.org/10.3390/technologies12110214.
14. https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset?resource=download