Classification of rice leaf diseases using transfer learning based on MobileNetV2 architecture

Achmad Yasid[1*], *Ana* Tsalitsatun Ni'mah[2], *Rima* Tri Wahyuningrum[1], *Muhammad* Nur Hidayat[3], *Mochammad* Mahreza Rizky Fahrozi[3], and *Retno* Hestiningrum[3]

[1]Lecturer, Department of Informatics Engineering, Faculty of Engineering, University of Trunodjoyo Madura, Indonesia
[2]Lecturer, Department of Informatics Engineering Education, Faculty of Education, University of Trunodjoyo Madura, Indonesia
[3]Undergraduate Student, Department of Informatics Engineering, Faculty of Engineering, University of Trunodjoyo Madura, Indonesia

Abstract. Rice leaf disease is a serious problem in the agricultural sector that can significantly reduce productivity. Early detection is necessary to minimize losses, but manual identification by agricultural experts is time-consuming and prone to error. This study aims to develop an automatic classification system for rice leaf disease using a Convolutional Neural Network (CNN) with a transfer learning approach on the MobileNetV2 architecture. The training process was carried out using several scenarios, namely Scratch, Fixed Feature, Fine Tuning First, Fine Tuning Middle, and Fine Tuning Last, with training parameters of 50 epochs, batch size 32, Adam optimizer, and learning rate 0.0001. The results indicated that the Fine Tuning First scenario worked best, with an accuracy of 78%, a Macro Average F1-score of 76%, and a Weighted Average F1-score of 78%. These findings indicate that the application of transfer learning on the MobileNetV2 architecture is capable of automatically detecting rice leaf diseases with fairly satisfactory results. In addition, this method has the potential to be further developed to support artificial intelligence-based smart farming systems.

1 Introduction

Agriculture remains pivotal to Indonesia's food security, with rice as the staple that underpins national diets and policy planning. Yet paddy systems face persistent headwinds: continued conversion of agricultural land and shifts in the economic structure threaten cultivated area and production, while seasonal shocks from weather, pests, and diseases depress yields and grain quality. Recent official statistics provide concrete baselines: harvested-area and production figures for 2024/2025 released by Statistics Indonesia (BPS) frame both progress and vulnerability at the national level and in key provinces [1], [2]. Concurrently, Indonesia's national pest outlook for the 2024/2025 season highlights geographically distributed risks from major rice diseases and insect pests, underscoring the need for earlier detection and response in the field [3].

Plant diseases typically manifest as distinctive lesions, color changes, and textural patterns on multiple organs, and leaves are the most accessible for visual monitoring and imaging. Recent rice-focused work documents canonical symptomology, for example elongated yellow-orange streaks for bacterial stripe, irregular gray-white lesions with brown margins for sheath blight, and elliptical spots with pale centers for brown spot. These findings reinforce leaves as a practical target for early screening and precision-agriculture interventions [4]. When captured consistently, leaf imagery enables timely diagnoses that translate into more effective control and improved harvest outcomes [4], [5], [6].

Manual diagnosis at scale is difficult because symptoms can be subtle, visually similar across pathogens, and confounded by lighting or background, which yields inconsistent judgments in real-world field conditions. Broad reviews show that traditional machine-learning pipelines relying on handcrafted features often plateau in accuracy and robustness, whereas modern image processing with deep neural networks provides stronger generalization on diverse plant-disease datasets [5], [6]. These results motivate automated classification systems tailored to rice leaves that can operate reliably on heterogeneous imagery from farms and extension services [4], [5], [6].

In this landscape, deep learning extends classical machine learning by learning hierarchical visual representations directly from pixels. Convolutional neural networks are the backbone for image tasks and, with transfer learning, allow practitioners to adapt pretrained encoders to domain-specific data efficiently [5], [6]. For deployment on resource-constrained devices common in agricultural settings, lightweight backbones are attractive. MobileNetV2 balances accuracy and efficiency, which makes it a strong candidate for rice leaf disease classification when

* Corresponding author: ayasid@trunojoyo.ac.id

coupled with transfer learning and careful preprocessing [5], [7], [8].

Evidence from adjacent vision tasks supports this choice. In large-scale fruit recognition, a transfer-learned MobileNetV2 variant achieved state-of-the-art accuracy while retaining computational frugality, which illustrates that compact encoders can match or exceed heavier baselines when heads are adapted appropriately [7]. More broadly in agricultural product classification, an improved MobileNetV2 architecture surpassed the original MobileNetV2 and compared favorably with classic models like VGG-16 and ResNet on public benchmarks while remaining efficient, an encouraging signal for mobile and edge deployments in agriculture [8]. In biodiversity classification, for example butterflies, transfer-learning studies report MobileNet-family models as competitive and efficient, though the highest top-1 accuracy may still come from larger backbones such as InceptionV3 on some datasets, which highlights a trade-off landscape where MobileNetV2 remains compelling for on-device use [9].

For rice specifically, recent international studies confirm that CNN-based classifiers and transfer learning are effective on curated rice-leaf datasets, and performance can be pushed further with judicious architectural choices. A 2024 comparative study showed MobileNetV2 and a compact CNN achieving high accuracy and near-perfect ROC for several rice diseases, with slight variability across classes [10]. Scientific Reports papers in 2025 further demonstrate advances through fusion pipelines and large annotated sets, including hybrid feature-extractor plus gradient-boosting schemes and end-to-end CNN ensembles for multi-disease recognition with rigorous splits. These results indicate that deep models remain state of the art and that carefully engineered lightweight backbones can remain competitive [11], [12]. Together with up-to-date reviews, this body of work situates MobileNetV2 as a practical baseline for rice leaf disease classification with room for targeted enhancements [4], [5], [6], [10], [11], [12].

Building on these insights, this study implements a MobileNetV2-based transfer-learning pipeline for classifying rice leaf diseases. The objective is to improve accuracy and efficiency while addressing data limitations typical of agricultural settings by leveraging pretrained visual knowledge, standardized augmentation, and domain-aligned preprocessing. The expected contributions include a compact and reproducible baseline that is deployable on modest hardware, empirical evidence on class-wise performance for common rice diseases, and a discussion of error patterns that inform subsequent improvements, for example lightweight attention heads or fusion classifiers, without sacrificing the deployability advantages of MobileNetV2 [5], [7], [8], [11], [12].

2 Literature review

A Research on plant leaf disease classification has surged over the past five years alongside advances in deep learning and transfer learning. In grapes, end to end detection and classification pipelines based on lightweight CNNs deliver strong real time performance in challenging orchard conditions such as clustered fruits and dense canopies, which confirms that efficient architectures are reliable for field use [13]. In citrus, recent studies highlight the importance of fine grained grading and multi class learning through transfer learning and ensembling, which raises accuracy across symptom variation and orchard backgrounds [14]. In tomato, ensembles that combine lightweight backbones such as MobileNetV2 with deeper networks such as ResNet50 achieve test accuracy near state of the art while keeping computation moderate [15].

Another research line emphasizes deployability on mobile devices. For corn, mobile first CNN models integrated into Android or iOS apps perform on field detection and classification of corn leaf diseases with low latency and throughput suitable for smart scouting [16]. Frontier work that designs phone based diagnosis systems also stresses a data to model to UX chain that remains robust to lighting changes and leaf pose [17]. To shrink the computational footprint further, several studies design lightweight architectures such as ShuffleNetV2 with residual or attention modules and show that competitive accuracy is achievable on multi commodity leaf images even under complex real backgrounds [18].

The MobileNet family remains prominent thanks to a strong balance between accuracy and efficiency. MobileNetV2 variants enriched with attention show favorable trade offs between model size and precision on several horticultural leaf datasets [19]. In tomatoes, MobileNetV2 based hybrids that add temporal gating also called neural controlled primitives or that insert inception or dense style blocks reduce intra class ambiguity and improve generalization in complex scenes [20], [21]. In cassava, comprehensive deep learning studies show that transfer learning on lightweight architectures remains competitive when facing orchard images with high texture and color variability [22]. Meanwhile, diffusion based research for soybean shows that modern object level detectors can complement or rival classic CNNs on fine label sets, which indicates a widening architectural toolkit for smart farming [23].

Beyond horticulture, other cereals also serve as testbeds for lightweight CNNs. In wheat, systematic evaluations of MobileNetV3, ShuffleNetV2, EfficientNetV2, and MnasNet in complex scenes report accuracy above 97 percent with small parameter counts, which underscores the value of training strategy and learning rate scheduling for extracting maximum performance from light backbones [24]. Returning to tomato, applied reviews affirm that field oriented data augmentation including GAN based methods and appropriate regularization are key to maintaining high accuracy without sacrificing deployability on constrained devices [25].

In summary, three lessons emerge across crops, transfer learning on lightweight CNNs, especially the MobileNetV2 family, is effective for moderately annotated images with complex backgrounds [15], [19], [20], [21], mobile centric design and latency optimization open the door to practical on device inference in orchards and fields [16], [17], and modularization through attention, ensembling, or integrated detection and classification helps overcome visual similarity across diseases and variation in imaging conditions [13], [14], [18], [22], [23]. These findings provide a strong foundation for selecting a light backbone such as MobileNetV2 and for designing pipelines that are ready for deployment within precision agriculture ecosystems.

3 Methodology

3.1 MobileNetV2 architecture

The MobileNetV2 architecture was used as the main model in this study because it has an efficient and lightweight structure and is suitable for transfer learning implementation on limited-sized datasets. MobileNetV2 is an extension of MobileNetV1 that adopts two main components, namely inverted residual blocks and linear bottleneck layers. Each block consists of three stages:

1. 1×1 convolution to expand the channel dimension (expansion layer),
2. 3×3 depthwise convolution for spatial feature extraction, and
3. 1×1 projection layer to reduce the dimension.

The ReLU6 activation function is used in each block to maintain numerical stability during training. This model utilizes pretrained weights from ImageNet, which are then modified through five test scenarios to determine the best configuration for rice leaf disease classification. The five scenarios are visualized in Fig 1 to 5, each depicting the frozen and trainable layers:

3.1.1 Model without transfer learning (Scratch)

In the Scratch scenario Fig 1, all MobileNetV2 layers are trained from scratch without using pretrained weights. The 224×224×3 rice leaf image input is processed through a series of inverted residual blocks and fully trainable bottleneck layers. The classification layer consists of global average pooling and dense layers with softmax activation for multi-class prediction. This approach evaluates the ability of the MobileNetV2 model to learn directly from target data but requires longer training time and risks overfitting if the dataset size is limited.

#

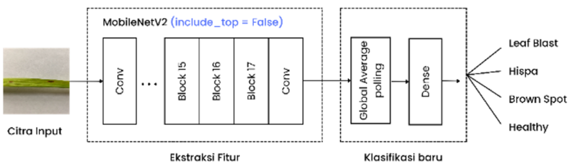

Fig. 1. Model without transfer learning.

3.1.2 Transfer learning fixed feature extraction

In the Fixed Feature Extraction scenario Fig 2, the MobileNetV2 architecture utilizes pre-trained weights from ImageNet by freezing all feature extractor layers. Only the classification layer at the end is retrained using the rice leaf dataset. This approach speeds up the training process and reduces the risk of overfitting while maintaining the visual feature generalization capabilities of the pre-trained model.

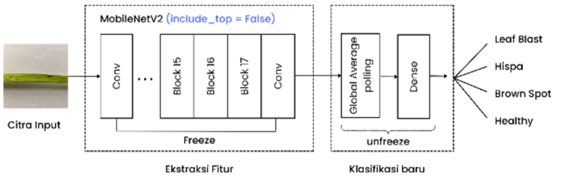

Fig. 2. Transfer learning fixed feature extraction.

3.1.3 Transfer learning with fine-tuning first layer

Fig. 3 shows the MobileNetV2 architecture used in this study with a transfer learning approach based on pre-trained weights from ImageNet. The strategy applied is fine-tuning on the first layers, where only some of the first layers are reactivated (unfrozen) for retraining, while the other layers remain frozen.

Specifically, the fine-tuning process was performed on the first 50 layers, which included the initial convolution layers and several initial bottleneck blocks (approximately blocks 1–4). These layers play a role in extracting basic features from images, such as edges, colors, and surface textures on rice leaves. By unfreezing some of the early layers, the model can adjust the basic feature representations to the specific visual characteristics of the new domain without changing the intermediate and high-level feature representations that have been stabilized in the ImageNet pre-trained weights.

Meanwhile, the subsequent bottleneck layers (blocks 5–17) and the final convolutional layer remain frozen so that the weights that have matured in complex feature representations are not distorted during the retraining process. The extracted features are then passed to a Global Average Pooling (GAP) layer to reduce the spatial dimensions, followed by a 1024-unit Dense layer activated with the ReLU function and equipped with 0.5 Dropout to prevent overfitting. Finally, an output dense layer with softmax activation is used as a classification layer for the four target classes.

This fine-tuning approach on the initial layers is designed to achieve a balance between domain adaptation and training efficiency. By retraining only a portion of the initial layers, the model can adapt to new data characteristics with faster training time, better convergence stability, and lower risk of overfitting, especially when the amount of training data is limited.

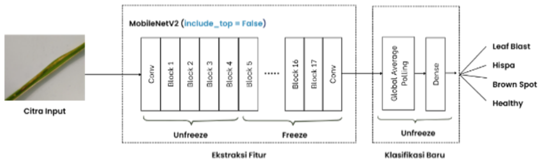

Fig. 3. Transfer learning with fine-tuning first layer.

3.1.4 Transfer learning with fine-tuning middle layer

Fig. 4 shows the MobileNetV2 architecture design with a transfer learning strategy using a fine-tuning approach on the middle layers. All initial layers of the model are first frozen to maintain stable general feature representations. Next, selective fine-tuning is performed on the middle layers, namely around layers 70 to 120, which structurally cover bottleneck blocks 5 to 12 of the MobileNetV2 architecture. These layers play a role in extracting mid-level features such as complex texture patterns and object shapes, so that the retraining process in this section allows the model to adapt to the specific visual characteristics of rice leaf images without changing the basic and high feature weights that are already optimal.

The extracted features are then processed through a Global Average Pooling (GAP) layer to reduce spatial dimensions, followed by a 1024-unit Dense layer with ReLU activation and Dropout 0.5 to prevent overfitting. The final layer is a dense layer with softmax activation used to generate predictions on four target classes.

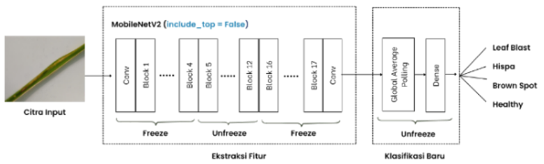

Fig. 4. Transfer learning with fine-tuning middle layer.

3.1.5 Transfer learning with fine-tuning last layer

Fig. 5 shows the MobileNetV2 architecture design with a transfer learning approach through fine-tuning on the last layers. All initial layers of the model are first frozen to maintain the stability of weights that have been trained on general feature representations. Next, selective fine-tuning is performed on the last 50 layers of the base model, which roughly covers the 13th bottleneck block to the final convolution layer (Conv_1). These layers play a role in forming high-level features that represent complex patterns and semantic characteristics of objects, so that updating the weights in this section allows the model to adapt to new data distributions more specifically without having to train the entire network from scratch.

The extracted features are then passed to the Global Average Pooling (GAP) layer to reduce the spatial dimensions, followed by a 1024-unit Dense layer with ReLU activation that serves as an additional non-linear mapping. To reduce the risk of overfitting, Dropout of 0.5 is applied before finally passing the results to the output Dense layer with Softmax activation, which classifies the images into four target classes.

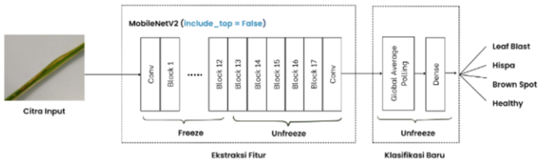

Fig. 5. Transfer learning with fine-tuning last layer.

3.2 Research flow

The research flow for the rice leaf disease classification system is shown in Fig 6. The research process begins with the initial stage, which involves preparing all data processing steps and model training. The dataset used consists of a collection of images of healthy rice leaves and leaves infected with diseases such as leaf blast, hispa, and brown spot as the main data source. Before being used for training, each image undergoes a preprocessing stage that includes resizing to 224×224 pixels to match the MobileNetV2 architecture input. In addition, oversampling is performed to balance the amount of data in each class, and image augmentation is performed to increase data variation and strengthen the model's resistance to overfitting.

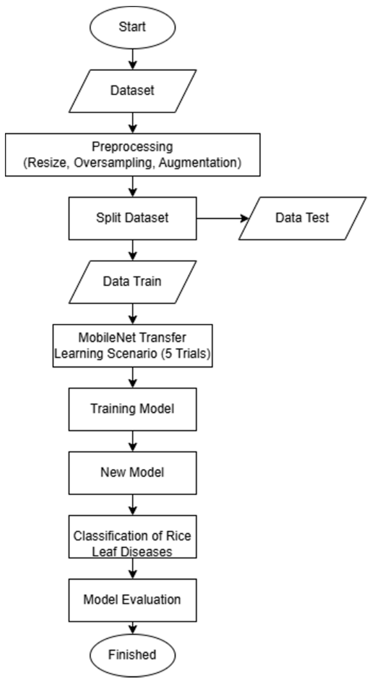

Fig. 6. Research flow.

Next, the dataset is divided with a ratio of 80:10:10, where 80% is used as training data, 10% as test data, and 10% for validation data. The next stage is the application of transfer learning with the MobileNetV2 architecture, where five test scenarios are carried out by setting the number of frozen and unfrozen layers, as well as adjusting parameters such as learning rate and optimizer to obtain the best training configuration.

The model training process was carried out using training data until optimal weights were obtained that produced the best classification performance. The resulting training model was then saved as a new model used in the rice leaf disease classification stage. At 2 this stage, the model predicted the class of each rice leaf image to determine whether the leaf belonged to the healthy category or was infected with a specific disease. The final stage is model evaluation, which is carried out using accuracy, precision, recall, and F1-score metrics to assess the effectiveness of the model in performing classification. After all scenarios have been tested and the best model has been obtained, the research process is declared complete. This approach is designed to produce an accurate, efficient, and reliable model for automatically detecting rice leaf diseases.

3.3 Dataset

The dataset used in this study consists of rice leaf images representing four classes, namely, leaf blast, hispa, brown spot, and healthy. The main dataset was obtained from the Kaggle public repository titled "Rice Leafs" https://www.kaggle.com/datasets/shayanriyaz/riceleafs), which contains 3,355 images. In detail, this dataset includes 1,488 healthy leaf images, 779 leaf blast images, 565 Hispa images, and 523 brown spot images. To improve the validity and generalization ability of the model, 200 additional images were taken directly in the Telang rice fields, Bangkalan Regency, using the main camera (50 megapixels) on a Tecno Pova 5 smartphone. The additional images consisted of four classes with a balanced distribution of 50 images per class and were used specifically as test data. All images collected were manually checked to ensure the quality of the dataset. Images showing more than one type of disease on a single leaf were excluded from the dataset because this study focused on identifying one type of disease per image.

The next step is preprocessing to standardize and improve data quality before model training. In addition, to overcome the imbalance in the amount of data between classes and expand the variety of training data, image augmentation techniques were applied. The augmentation techniques used included random rotation of up to ±15°, random zoom in/out of 10%, horizontal and vertical shifts of up to 10%, and horizontal image flipping. Empty areas that appear as a result of the transformation are filled using the edge pixel reflection method (fill mode = reflect). This augmentation process is carried out mainly on minority classes so that their numbers are balanced with the majority classes.

Through these preprocessing and augmentation stages, the dataset becomes more balanced, diverse, and representative of real-world conditions in the field. The final result of this process is a dataset that is ready for use in the transfer learning and fine-tuning stages using the MobileNetV2 architecture for accurate and efficient classification of rice leaf diseases.

4 Experimental results

4.1 Model training results

The experiment was conducted with five model training scenarios to evaluate the effect of fine-tuning strategies on model performance. Each scenario had a different approach to freezing and retraining convolutional network layers. The Table 1 shows the training results for each scenario, including the number of epochs, training accuracy, training loss, validation accuracy, and validation loss.

Table 1. Training results.

Scenario	Epoch	Train Accuracy	Train Loss	Val accuracy	val_loss
Scratch	18	0.5702	1.0089	0.2486	1.3703
Fixed Feature	29	0.6911	0.7473	0.7238	0.7459
First Layer	50	0.8572	0.3739	0.7735	0.5777
Middle Layer	24	0.8853	0.3009	0.7707	0.6704
Last Layer	8	0.7819	0.5816	0.5912	2.0361

From these results, it can be seen that the model with fine-tuning on the first layer gave the best results on the validation data, with a validation accuracy of 0.7735 and a validation loss of 0.5777. This shows that retraining a small portion of the first layer can improve the model's ability to recognize relevant features without losing its generalization ability.

The Middle Layer scenario also showed competitive performance with a validation accuracy of 0.7707, despite having a slightly higher loss than the First Layer. Meanwhile, the model trained from scratch produced the lowest validation accuracy, namely 0.2486, indicating that without the initial weights from the pre-trained model, the model requires more epochs and data to achieve optimal performance.

The fixed feature scenario (convolutional layers are frozen and only the classification layer is retrained) produces fairly good results with a validation accuracy of 0.7238, indicating that the feature representation of the pre-trained model is still very relevant to the dataset used. In contrast, the Last Layer scenario shows unstable results with a high validation loss (2.0361), indicating the possibility of overfitting in the last layer only.

4.2 Model testing results

The results of testing the model on the test data are shown in Table 2.

Table 2. Testing results.

Scenario	Accuracy Test (%)
Scratch	23
Fixed Feature	70
First Layer	78
Middle Layer	69
Last Layer	44

From the test results, the model with the first-layer fine-tuning scenario again showed the best performance with a test accuracy of 78%, followed by the fixed-feature scenario with 70% and the middle-layer with 69%. These results are consistent with the training phase, where fine-tuning part of the early layers provided the best balance between feature adaptation and generalization to the new dataset.

Meanwhile, the model trained from scratch only achieved an accuracy of 23%, reinforcing that training without leveraging pre-trained weights is less effective, especially on datasets with limited size or high complexity. The Last Layer scenario achieved a test accuracy of 44%, which is significantly lower compared to other transfer learning configurations. This suggests that updating only the final classification layer is insufficient for adapting feature representations to the new data, causing the model to struggle in learning deeper distinguishing patterns and resulting in suboptimal performance.

Overall, these results confirm that the transfer learning approach with fine-tuning on the initial layers is the most effective strategy for improving model performance without requiring full training from scratch.

4.3 Confusion matrix analysis

Fig. 7 shows the confusion matrix of the best model, which is the model with the first-layer fine-tuning scenario. This visualization is used to analyze the distribution of model predictions against actual labels in the test data.

From Fig. 7, it can be seen that the model is able to recognize most classes well. The main diagonal values show the number of correct predictions for each class. For example, in Class 1, the model successfully classified 141 samples correctly, indicating that the model has excellent ability to recognize feature patterns in that class.

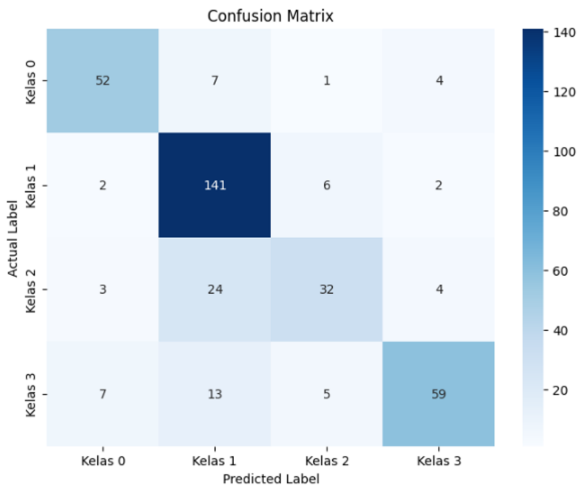

Fig. 7. Confusion matrix first layer.

However, there are still a number of classification errors between classes, such as several samples from Class 2 that were incorrectly predicted as Class 1 and a small portion of Class 3 that were classified as Class 1 or Class 2. This indicates that some classes have similar visual features, causing overlap in the feature space.

Overall, the pattern in the confusion matrix indicates that the model with fine-tuning on the initial layer has good generalization capabilities and a relatively balanced distribution of predictions among the existing classes.

To evaluate the performance of the proposed method in a broader research context, a comparison was conducted with previous studies that applied deep learning techniques for leaf disease classification but achieved lower accuracy. As shown in Table 3, the proposed method obtained a higher accuracy (78%) compared to earlier models that utilized traditional CNN or transfer learning approaches with limited fine-tuning optimization.

Table 3. Comparison with other studies.

Author and year	Methods	Object	Accuracy (%)
Askale G et al., 2025 [16]	ResNet50	Maize leaf diseases	72%
Khairun Saddami et al., 2024 [26]	*ShuffleNet* (Lightweight CNN)	Rice leaf diseases	66.51%
Proposed Method (this study)	MobileNetV2 + Fine-Tuning (First Layer)	Rice leaf diseases	78%

From the comparison results in Table 4, it can be concluded that the proposed method provides a balanced combination of accuracy and model efficiency, surpassing earlier architectures that performed below 78%. This confirms that early-layer fine-tuning in MobileNetV2 can improve feature generalization in small agricultural datasets while maintaining a lightweight computational design suitable for practical smart-farming applications.

4.4 Strengths and limitations

This study offers several strengths that contribute to the growing body of research on deep learning applications in agriculture. The experimental design involving five distinct training scenarios provides a systematic comparison of fine-tuning strategies within the MobileNetV2 architecture, allowing a comprehensive understanding of how different layer adaptations affect model performance. The results demonstrate the strength of early-layer fine-tuning in achieving high accuracy and generalization while maintaining computational efficiency. The inclusion of locally sourced images alongside the Kaggle dataset also strengthens the model's relevance by incorporating real-world variability, reflecting the visual diversity of rice leaf conditions encountered in the field.

However, several limitations should be acknowledged. The dataset size remains relatively small and geographically limited, which may reduce the model's ability to generalize to different rice varieties or environmental conditions. The study focuses solely on the MobileNetV2 backbone, without direct benchmarking against other lightweight or hybrid CNN architectures such as EfficientNet or DenseNet. Additionally, hyperparameter settings were held constant across experiments, which might have constrained the optimization potential of certain configurations. Another limitation is that the study did not evaluate inference speed or hardware efficiency, which are critical factors for real-world deployment on mobile or embedded agricultural systems. Addressing these aspects in future work could further validate the model's practicality and robustness across diverse operational contexts.

5 Conclusion

Based on the experimental results obtained from five training scenarios, this study concludes that the transfer learning approach with fine-tuning on the first convolutional layers yields the best overall performance for rice leaf disease classification using the MobileNetV2 architecture. This configuration achieved the highest validation accuracy (0.7735) and test accuracy (78%), indicating that retraining a limited portion of early layers effectively adapts pre-trained features to new data while maintaining generalization capability. In contrast, the Scratch model produced the lowest accuracy (23%), demonstrating that training from random initialization is inefficient for small or moderately complex datasets. The Fixed Feature and Middle Layer configurations also performed well (70% and 69% test accuracy, respectively), confirming that pre-trained visual representations remain highly transferable across similar agricultural domains.

The confusion matrix analysis further reinforces these results, showing balanced classification performance across all classes, although a few misclassifications occurred between visually similar disease categories. These findings emphasize that fine-tuning the initial layers provides an optimal trade-off between learning efficiency, model stability, and classification accuracy. Thus, this approach contributes new empirical evidence supporting early-layer adaptation as a robust and resource-efficient transfer learning strategy for agricultural image classification tasks.

Nevertheless, this study has several limitations. The dataset size was relatively limited and sourced from specific regions, which may restrict model generalization to other environmental or varietal conditions. Furthermore, only one lightweight CNN architecture (MobileNetV2) and fixed hyperparameters were explored, leaving room for broader architectural comparisons and optimization strategies. In future research, larger and more diverse datasets should be collected, including multi-season and multi-location images, to evaluate the model's robustness. The integration of attention mechanisms, multi-modal features, or hybrid architectures may also enhance detection precision and generalization. Additionally, deployment in mobile or edge-based smart farming systems can be explored to assess inference speed and usability in real agricultural settings.

In summary, this research contributes to the ongoing development of AI-assisted plant pathology by empirically demonstrating that selective fine-tuning of early convolutional layers within a pre-trained MobileNetV2 can significantly improve classification performance while maintaining computational efficiency. These results not only validate transfer learning as a practical approach for agricultural image recognition but also establish a foundation for future adaptive and field-ready disease diagnosis systems.

This research was supported by the Research and Community Service Institute of Trunodjoyo Madura University through the Research Group program. [grant numbers: 306/UN46.4.1/PT.01.03/RISMAN/2025].

References

1. Badan Pusat Statistik (BPS - Statistics Indonesia), Pada 2024, luas panen padi mencapai sekitar 10,05 juta hektare dengan produksi padi sebanyak 53,14 juta ton gabah kering giling (GKG). Badan Pusat Statistik (BPS - Statistics Indonesia). (2025)
2. Badan Pusat Statistik (BPS - Statistics Indonesia), Luas panen padi di Indonesia pada Mei 2025 diperkirakan mencapai 0,98 juta hektare dengan total produksi sekitar 4,98 juta ton gabah kering giling (GKG). Badan Pusat Statistik (BPS - Statistics Indonesia). (2025)
3. BBPOPT, Prakiraan Serangan OPT Utama Padi, Jagung, Kedelai, Aneka Kacang dan Umbi di Indonesia MT 2024/2025. Balai Besar Peramalan Organisme Pengganggu Tumbuhan. (2024)
4. P. Pai et al., Deep learning-based automatic diagnosis of rice leaf diseases using ensemble CNN models. Sci Rep. **15**, 1, 27690 (2025). https://doi.org/10.1038/s41598-025-13079-z
5. M. Shoaib et al., An advanced deep learning models-based plant disease detection: A review of recent research. Front Plant Sci. **14**, 1158933 (2023). https://doi.org/10.3389/fpls.2023.1158933
6. J. Zhao et al., A review of plant leaf disease identification by deep learning algorithms. Front Plant Sci. **16**, 1637241 (2025). https://doi.org/10.3389/fpls.2025.1637241
7. Y. Gulzar, Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique. Sustainability. **15**, 3, 1906 (2023). https://doi.org/10.3390/su15031906
8. H. Chen, G. Zhou, W. He, X. Duan, and H. Jiang, Classification and identification of agricultural products based on improved MobileNetV2. Sci Rep. **14**, 1, 3454 (2024). https://doi.org/10.1038/s41598-024-53349-w
9. F. Rajeena P. P. et al., A Novel Method for the Classification of Butterfly Species Using Pre-Trained CNN Models. Electronics. **11**, 13, 2016

(2022). https://doi.org/10.3390/electronics11132016

10. M. Dutta et al., Rice Leaf Disease Classification—A Comparative Approach Using Convolutional Neural Network (CNN), Cascading Autoencoder with Attention Residual U-Net (CAAR-U-Net), and MobileNet-V2 Architectures. Technologies. **12**, 11, 214 (2024). https://doi.org/10.3390/technologies12110214
11. Z. Wang, Y. Wei, C. Mu, Y. Zhang, and X. Qiao, Rice Disease Classification Using a Stacked Ensemble of Deep Convolutional Neural Networks. Sustainability. **17**, 1, 124 (2024). https://doi.org/10.3390/su17010124
12. B. Naresh kumar and S. Sakthivel, Rice leaf disease classification using a fusion vision approach. Sci Rep. **15**, 1, 8692 (2025). https://doi.org/10.1038/s41598-025-87800-3
13. Md. J. Karim et al., Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM. Sci Rep. **14**, 1, 16022 (2024). https://doi.org/10.1038/s41598-024-66989-9
14. B. Kaur et al., Precision diagnosis of citrus leaf diseases using image enhancement and nonlinear fuzzy ranking ensemble approach NLFuRBe. Sci Rep. **15**, 1, 32296 (2025). https://doi.org/10.1038/s41598-025-16923-4
15. J. Sharma et al., Deep learning based ensemble model for accurate tomato leaf disease classification by leveraging ResNet50 and MobileNetV2 architectures. Sci Rep. **15**, 1, 13904 (2025). https://doi.org/10.1038/s41598-025-98015-x
16. G. T. Askale et al., Mobile based deep CNN model for maize leaf disease detection and classification. Plant Methods. **21**, 1, 72 (2025). https://doi.org/10.1186/s13007-025-01386-5
17. F. Khan et al., A mobile-based system for maize plant leaf disease detection and classification using deep learning. Front Plant Sci. **14**, 1079366 (2023). https://doi.org/10.3389/fpls.2023.1079366
18. H. Zhou et al., Identification of leaf diseases in field crops based on improved ShuffleNetV2. Front Plant Sci. **15**, 1342123 (2024). https://doi.org/10.3389/fpls.2024.1342123
19. H. Wang, S. Qiu, H. Ye, and X. Liao, A Plant Disease Classification Algorithm Based on Attention MobileNet V2. Algorithms. **16**, 9, 442 (2023). https://doi.org/10.3390/a16090442
20. A. T. Le, M. Shakiba, I. Ardekani, and W. H. Abdulla, Optimizing Plant Disease Classification with Hybrid Convolutional Neural Network–Recurrent Neural Network and Liquid Time-Constant Network. Applied Sciences. **14**, 19, 9118 (2024). https://doi.org/10.3390/app14199118
21. L. Li, B. Wang, Y. Li, and H. Yang, Diagnosis and Mobile Application of Apple Leaf Disease Degree Based on a Small-Sample Dataset. Plants. **12**, 4, 786 (2023). https://doi.org/10.3390/plants12040786
22. G. Sambasivam et al., A hybrid deep learning model approach for automated detection and classification of cassava leaf diseases. Sci Rep. **15**, 1, 7009 (2025). https://doi.org/10.1038/s41598-025-90646-4
23. J. Yin et al., A Diffusion-Based Detection Model for Accurate Soybean Disease Identification in Smart Agricultural Environments. Plants. **14**, 5, 675 (2025). https://doi.org/10.3390/plants14050675
24. X. Wen, M. Zeng, J. Chen, M. Maimaiti, and Q. Liu, Recognition of Wheat Leaf Diseases Using Lightweight Convolutional Neural Networks against Complex Backgrounds. Life. **13**, 11, 2125 (2023). https://doi.org/10.3390/life13112125
25. A. Guerrero-Ibañez and A. Reyes-Muñoz, Monitoring Tomato Leaf Disease through Convolutional Neural Networks. Electronics. **12**, 1, 229 (2023). https://doi.org/10.3390/electronics12010229
26. K. Saddami, Y. Nurdin, M. Zahramita, and M. S. Safiruz, Advancing Green AI: Efficient and Accurate Lightweight CNNs for Rice Leaf Disease Identification. (2024)