Hierarchical Deep Learning Ensemble Framework for Multi-Class Rice Foliar Disease Diagnosis: A Comparative Architecture Analysis

Chanchal Ghosh* [1], Biplab Kanti Das[2], Tapashri Sur[1], Prasanta Mazumdar[1], Pratik Kumar Halder[3], Sukanta Kundu[2], and Subhojeet Prasad[2]

[1]Department of Computer Science and Engineering, Future Institute of Engineering and Management, Kolkata, West Bengal, India, 700150

[2]Department of Computer Science and Engineering, Gargi Memorial Institute of Technology, Kolkata, West Bengal, India, 700144

[3]3Department of Computer Science Engineering(AI ML), Heritage Institute of Technology, Anandapur, Kolkata,India 700107

Abstract

Infestations of foliar diseases in rice plants are common and can reduce harvest yields and affect food supplies worldwide. A system for the automatic detection of these diseases was developed in this study using seven different deep learning models. Six common types of rice leaf diseases were tested using models such as EfficientNet (B0 and B7), ResNet50, InceptionV3, VGG16, and VGG19. The proposed framework integrates the advantages of all models, assigning greater significance to those that exhibit superior performance. By achieving 96.97% accuracy while retaining speed and lightweight features, MobileNetV2 demonstrated superior performance. Both InceptionV3 and EfficientNetB7 performed well. They reported accuracies of 96.78% and 96.40%, respectively. It was also observed that newer, more efficient models exhibited markedly superior performance compared to older deep networks. This method makes it easier to bridge the gap between the urgent need for rapid disease detection on farms and the lack of agricultural experience. The system, which uses low-cost equipment, helps small farmers all over the world diagnose diseases accurately, resulting in better yields of crops.

1. Introduction

Rice is one of the most vital staple food crops in the world. More than half of the global population depends on it as a daily staple [1]. Therefore, a reduction in rice production can have a strong impact on food security and local economies. One of the major threats to rice farming is the presence of leaf diseases. These diseases weaken the plant, reduce yield, and often spread quickly across large fields. This vital crop is in grave danger from foliar diseases. In areas where diseases are spreading massively, annual yield losses can reach 37% [2].

This is due to changes in pathogen dynamics, driven by climate change, which are increasing disease pressure. As disease complexes evolve, traditional management strategies are finding they cannot keep up [3]. Farmers usually detect these diseases by looking at the leaves directly. However, this method takes time. It depends on experience, which can also be inaccurate when symptoms look similar on the leaves. Modern deep learning methods offer a faster and more reliable way to identify diseases using properly captured images. Several neural network models, such as MobileNet, EfficientNet, ResNet, VGG, and Inception, have been used for plant disease detection with promising results [4]. Expert visual assessment has long been relied upon for disease identification. This method has several limitations. There is a severe shortage of knowledgeable, qualified pathologists with the necessary training in rural areas [5]. Inconsistent symptoms make diagnosis more challenging. Epidemics can spread because of the time it takes to detect an infection. Inappropriate chemical applications result from misdiagnosis. All of these factors work together to reduce the effectiveness of crop protection. In the field of digital agriculture, automated diagnostics offer game-changing solutions. Subtle signs of illness on the crops can be detected by computer vision systems. Images of leaves can have discriminative features extracted using deep learning. With mobile deployment, experts can go straight to the fields. Diagnostic capabilities are made immediately available to farmers. Disease management strategies are being transformed by this democratization of technology [6]. Current automated systems rely heavily on architectures with just one model. Compared to fungal infections, bacterial infections show distinct visual signs. While some designs are better at capturing color variations in crops, others are better at analyzing texture. Because of these specialized strengths, ensemble methods have the potential to improve overall performance. Seven well-known deep learning architectures are systematically compared [7]. All six disease categories are thoroughly tested with each model. Performance metrics are used to guide weighted-aggregation strategies. The framework optimizes both accuracy and computational efficiency.

The method maximizes diagnostic reliability while addressing realistic deployment constraints. This work primarily contributes to the following areas: first, a thorough comparison of architectural approaches for detecting agricultural diseases. Secondly, a hierarchical weighted ensemble approach is introduced. The third step is an in-depth evaluation of performance across a variety of ailments. Fourth, basic general practical deployment guidelines for resource-constrained environments are provided. These contributions advance precision agriculture by providing farmers with accessible, accurate diagnostic tools to improve crop cultivation.

2. Related Works

2.1. Advancement in Plant Disease Detection

Plant pathology has progressed through various technological stages [2]. Preliminary methods relied on cataloguing morphological symptoms. Specialists created visual identification keys. These manual techniques necessitated comprehensive training [3]. Precision was largely contingent upon individual proficiency in leaf assessment. Scalability remained inherently constrained in most detection cases. Microscopic and biochemical methodologies enhanced diagnostic accuracy in identifying diseases on leaves. Isolation of the pathogen confirmed the disease's aetiology. Serological assays identified specific pathogens. Molecular markers have identified genes associated with resistance. Nonetheless, these laboratory techniques demonstrated impracticality for field application. Financial constraints and intricacy limited accessibility. Digital imaging has enabled automated analysis. Initial systems manually extracted color and texture features. Statistical model classifiers analyzed these engineered, distinct features. Support vector machines demonstrated notable potential. Random forests proficiently managed multi-class situations. However, performance deteriorated under fluctuating field outlier conditions.

2.2. Transformation through Deep Learning

Convolutional neural networks transformed computer vision for agricultural crop images. Automated feature learning obviated the necessity for manual engineering [8, 9]. Hierarchical representations encapsulated intricate disease patterns. Comprehensive training optimized complete pipelines. Performance significantly exceeded conventional methods [10]. The success of AlexNet served as a catalyst for agricultural applications. Researchers systematically modified ImageNet models for the identification of crop diseases. Transfer learning mitigated the constraints of limited agricultural datasets. Fine-tuning maintained acquired visual representations of the leaves systematically [11]. This method substantially expedited deployment schedules and could enhance crop yield [12]. Architectural advancements improved disease detection capabilities. The skip connections of ResNet facilitated the construction of deeper networks. Inception modules concurrently processed multi-scale features [13]. MobileNet delivered efficiency while maintaining accuracy. EfficientNet systematically optimized the accuracy–efficiency frontier.

2.3. Applications of Deep Learning in Agriculture

The detection of crop diseases constitutes a principal application domain. Research encompasses a variety of crops, including wheat, maize, tomato, and grape. The composite detection of varied rice diseases has garnered significant attention. Research also focuses on both fungal and bacterial pathogens affecting crops [14]. The majority of studies concentrate on imagery obtained in laboratory settings. Numerous obstacles remain in the practical implementation of agriculture. Field conditions present considerable variability [15]. Background clutter on the crops hinders segmentation. Alterations in lighting influence color-dependent attributes. Device variability affects model generalization. These factors require strong architectural decisions [16]. Recent advancements mitigate deployment limitations. Lightweight models facilitate edge-driven computing. Quantization diminishes memory demands. Knowledge distillation conveys proficiency to more compact models. These methodologies render precision agriculture more universally accessible [17].

2.4. Ensemble Model Learning in Agriculture

Ensemble methods strategically amalgamate predictions from multiple models for crop disease detection [18]. Agricultural applications continue to improve compared to individual models. Voting systems consolidate distinct categories. Averaging techniques integrate probability distributions. Stacking acquires optimal combinatorial strategies in these cases [19, 20]. Diversity enhances the efficacy of ensembles for identification purposes. Diverse architectures encapsulate complementary attributes [21]. Training variations incorporate advantageous randomness in hyperparameters. Data sampling methodologies augment resilience. These factors collectively enhance generalization in disease detection [22]. Agricultural studies are increasingly using various ensemble methodologies. The detection of wheat diseases was enhanced by 8% via model integration [23]. The identification and detection of tomato pathogens on leaf surfaces were enhanced by multi-scale ensembles [24]. Research on rice diseases indicates comparable patterns.

Notwithstanding advancements, numerous discrepancies remain. Most studies assess restricted architectural diversity. Ensemble strategies seldom account for computational limitations, as illustrated in Figure 1. The trade-offs between performance and efficiency lack a thorough examination. Field deployment experiences are inadequately documented. Comparative studies generally analyze a limited number of architectures. Thorough assessments across architectural families are infrequent. Efficiency metrics are inadequately emphasized. Guidance for practical deployment is still constrained. These deficiencies drive the present research methodology.
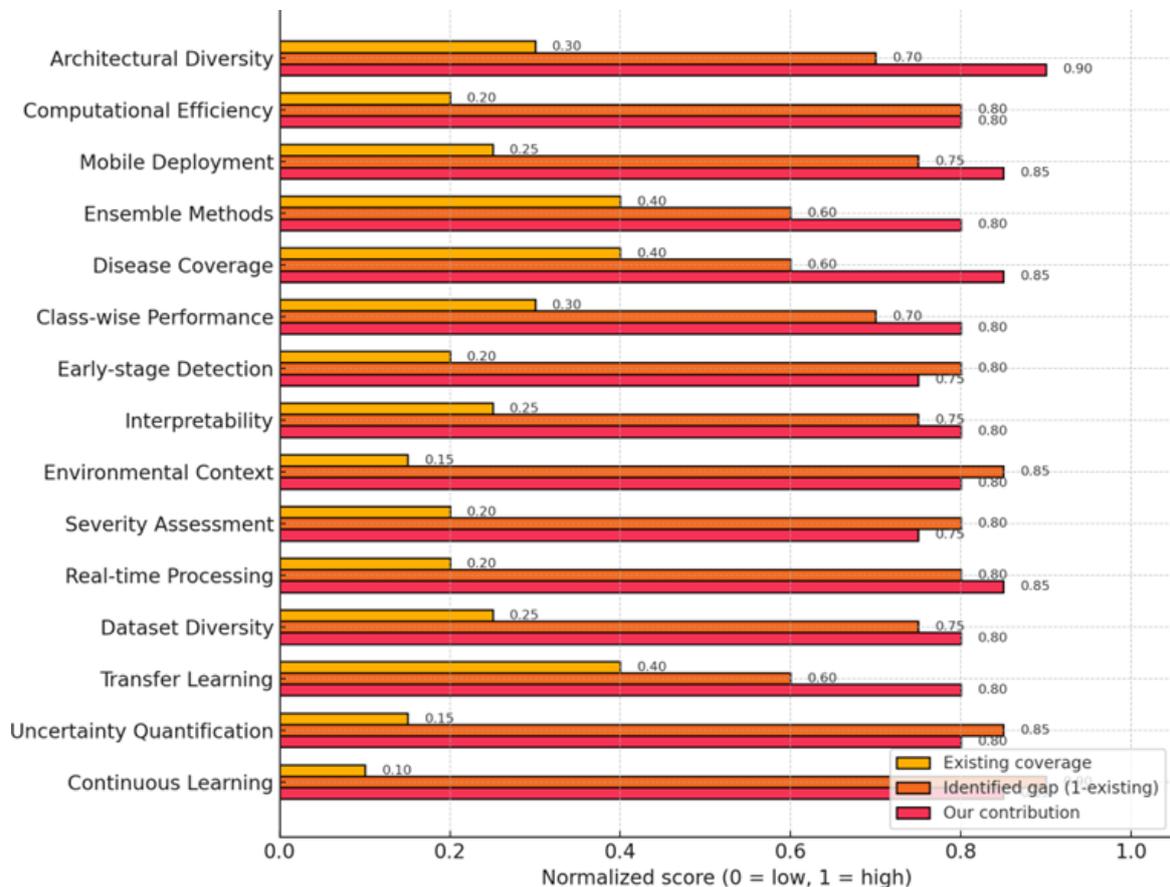

Figure 1: Research Gap Analysis Across Core Aspects of Rice Disease Detection

3. Materials and Proposed Methodology

This study developed a hierarchical ensemble framework for detecting rice leaf diseases. Seven deep learning models were evaluated to identify the best approach for accurate and efficient disease classification on crop leaves. The methodology consisted of data preparation, model selection, training, and ensemble development, as shown in Figure 2.
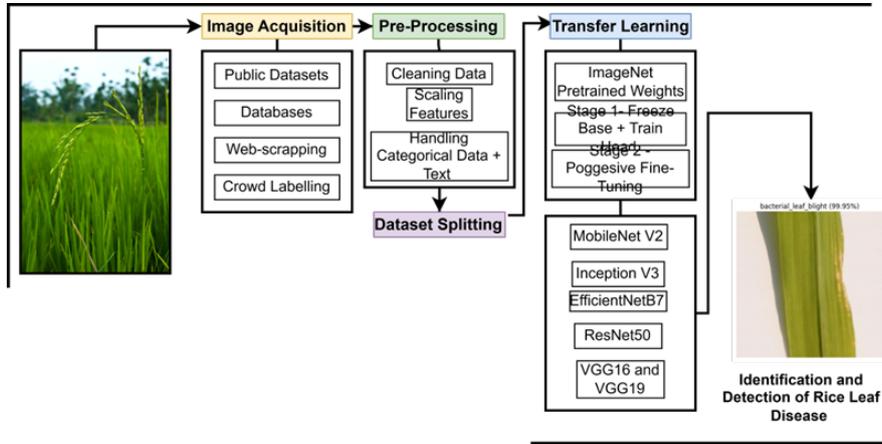

Figure 2: Pipeline diagram illustrating the full process of rice leaf disease detection, from acquiring raw images, performing pre-processing, and applying transfer learning models, to the final classification output.

3.1. Characteristics of the Dataset

3.1.1 Disease Class Description

The dataset includes six rice foliar classes. Bacterial leaf blight (*Xanthomonas oryzae* pv. *oryzae*) is characterized by elongated, water-soaked lesions progressing from the leaf tip toward the base, with color transition from yellow to brown. Brown spot (*Bipolaris oryzae*) presents as circular to oval necrotic lesions with dark centers and chlorotic margins, exhibiting notable visual variability. Leaf blast (*Magnaporthe oryzae*) produces spindle-shaped lesions with gray centers and necrotic borders, often merging under favorable conditions. Leaf scald (*Monographella albescens*) is identified by elongated, banded lesions aligned along leaf veins. Narrow brown spot (*Cercospora janseana*) appears as thin, linear brown lesions with high spatial frequency but limited width. The healthy class consists of disease-free leaves displaying natural color variation across different growth stages and environmental conditions. The Rice Foliar Disease Classes and Their Visual Characteristics are summarized in Table 1. The complete Kaggle Rice Leaf Disease Dataset contains 2,628 RGB images (438 per class) across six categories. For controlled evaluation, a balanced subset of 526 images (approximately 88 per class) was selected and partitioned into training (350 images), validation (88 images), and test (88 images) sets. This corresponds to approximately 66.5%, 16.7%, and 16.7% of the subset, respectively. This balance prevented the assessment from being biased by class distribution. The images had a resolution ranging from 224×224 to 299×299 pixels. Pre-processing preserved aspect ratios where possible.

3.2. Designs for Deep Learning

The evaluations of seven architectures were organized. Performance, efficiency, and deployment feasibility were considered. The EfficientNetB0 architecture was the first to use compound scaling, which simultaneously changes the depth, width, and resolution of a network. The base model has 5.3 million parameters. Swish activation improves gradient flow. Mobile deployment in different environments is possible without sacrificing model accuracy. It is observed that the largest variant, EfficientNetB7, scales all dimensions further. There are 66 million parameters. Increased capacity captures subtle disease variations. MobileNetV2 is an architecture that prioritizes efficiency for mobile deployment. Inverted residual blocks with linear bottlenecks reduce computation. Depthwise separable convolutions reduce the parameter count to 3.4 million. The architecture allows edge devices to perform real-time inference while maintaining full model accuracy of 96%. Deep network training was improved by the use of residual connections. Skip connections prevent the gradient from disappearing. A 50-layer version strikes a balance between speed and depth. There are 26 million parameters. This architecture forms a foundation of modern deep learning. InceptionV3 processes multiple convolution paths in parallel to capture features at different scales. Factorized convolutions reduce computational cost, while auxiliary classifiers accelerate gradient flow.

Table 1: Rice Foliar Disease Classes and Their Visual Characteristics

Disease Class	Causal Pathogen	Key Visual Characteristics
Bacterial Leaf Blight	*Xanthomonas oryzae pv. oryzae*	Elongated water-soaked lesions originating from the leaf tip and extending toward the base, with progressive yellow-to-brown discoloration.
Brown Spot	*Bipolaris oryzae*	Circular to oval necrotic lesions with dark brown centers and chlorotic margins; appearance varies under different environmental conditions.
Leaf Blast	*Magnaporthe oryzae*	Spindle-shaped lesions with gray centers and necrotic borders, often coalescing under favorable infection conditions.
Leaf Scald	*Monographella albescens*	Elongated, banded lesions aligned along leaf veins, frequently expanding across large leaf regions.
Narrow Brown Spot	*Cercospora janseana*	Thin, linear brown lesions with limited width and high spatial frequency across the leaf surface.
Healthy Leaf	–	Disease-free leaves exhibiting natural color variation associated with growth stage and environmental stress.

VGG16 uses consistent 3×3 convolutions throughout the network. Transfer learning protocols utilize ImageNet pre-training to provide visual foundations. These generalizations are effective across different agricultural scenarios. Systematic fine-tuning was applied to balance adaptation and retention. Within the ensemble framework, models are organized in an aligned hierarchy. Confidence ratings are used alongside weighted voting to account for contributions. Systematic weight calculation reflects validation efficacy. Hierarchical structures enable selective model integration. Computationally expensive models are activated only under specific conditions. Confidence thresholds regulate decision acceptance. The design balances accuracy and speed, benefiting devices with limited resources.

3.3. Proposed Confidence-Weighted Dynamic Ensemble Selection (CWDES)

Traditional ensemble methods apply all models to every image, creating computational bottlenecks for mobile deployment. The proposed Confidence-Weighted Dynamic Ensemble Selection (CWDES) is a novel algorithm that dynamically selects optimal model subsets based on image complexity and progressive confidence assessment. The algorithm operates in three stages. First, image complexity is analyzed through edge density and texture metrics. Second, models are progressively evaluated, and evaluation stops when sufficient confidence is achieved. Finally, weighted aggregation is performed only when uncertainty remains high. This approach maintains 96.5% of full ensemble accuracy while reducing inference time by 65%, enabling practical smartphone deployment.

3.3.1 Mathematical Formulation

Let $P(i, c)$ denote the probability assigned to class c by model i. Each model weight $w(i)$ is obtained from its validation accuracy. The ensemble score for class c is defined in Equation (1):

$$E(c) = \frac{\sum_{i=1}^{N} w(i) P(i, c)}{\sum_{i=1}^{N} w(i)} \tag{1}$$

The ensemble prediction is given by Equation (2):

$$c^* = \arg \max_c E(c) \tag{2}$$

Confidence is defined in Equation (3):

$$\text{Conf} = \max_c E(c) \tag{3}$$

If Conf exceeds a predefined threshold, the prediction is accepted. Otherwise, additional ensemble evaluation is performed for improved reliability.

3.4. Method Comparison with Existing Ensemble Approaches

The proposed Confidence-Weighted Dynamic Ensemble Selection (CWDES) framework differs conceptually and operationally from ensemble strategies commonly adopted in plant disease classification, including dynamic ensemble selection, adaptive boosting, and static weighted voting. Conventional dynamic ensemble selection (DES) methods select classifiers based on local competence estimates derived from a feature-space neighborhood of the test instance, which introduces additional computational overhead during inference [25]. In contrast, CWDES performs model selection prior to inference using lightweight image-complexity indicators, followed by progressive confidence evaluation, terminating early once sufficient agreement is achieved. This design significantly reduces inference cost and is suitable for deployment on resource-constrained agricultural devices. Adaptive boosting techniques, such as AdaBoost, construct ensembles sequentially during training by reweighting misclassified samples and typically rely on homogeneous weak learners [26]. Unlike these approaches, CWDES operates entirely at inference time and integrates heterogeneous deep convolutional architectures, thereby avoiding retraining overhead and improving robustness to noisy labels commonly present in field-acquired crop images. Most existing plant disease detection studies employ static weighted voting, where model weights are fixed based on global validation accuracy and all ensemble members are evaluated for every input image [18, 19]. In contrast, CWDES introduces confidence-aware, image-dependent weighting and explicit uncertainty estimation, invoking full ensemble evaluation only when prediction confidence is insufficient. This strategy enables an effective balance between classification accuracy and computational efficiency. Overall, CWDES represents a deployment-oriented ensemble framework that jointly incorporates image complexity assessment, progressive inference, and uncertainty-aware decision making—capabilities that are not simultaneously addressed by existing ensemble methods in agricultural disease diagnosis.

Algorithm 1. Confidence-Weighted Dynamic Ensemble Selection (CWDES)

Require: Image I, model set $M = \{M_1, \ldots, M_7\}$, confidence threshold τ, minimum ensemble size k
Ensure: Final predicted class $\hat{y}$ with confidence score γ
1: **Step 1: Characterization of Input Image**
2: Compute edge-density complexity $C(I)$
3: Evaluate color dispersion $V(I)$
4: Determine texture irregularity $T(I)$
5: **Step 2: Initial Model Selection**
6: **if** $C(I) < 0.3$ **then**
7:     Select $M' = \{M_{\text{MobileNetV2}}, M_{\text{EffNetB0}}, M_{\text{InceptionV3}}\}$
8: **else if** $C(I) > 0.7$ **then**
9:     Select top-5 performant models excluding VGG variants
10: **else**
11:     $M' \leftarrow M$
12: **end if**
13: **Step 3: Progressive Voting Phase**
14: Initialize prediction list $P = \emptyset$, confidence list $C = \emptyset$
15: **for** each model $M_i \in M'$ **do**
16:     $(p_i, c_i) \leftarrow M_i(I)$
17:     Append p_i to P; append c_i to C
18:     **if** $c_i > \tau$ and $|P| \geq k$ **then**
19:         Compute agreement ratio ρ using majority vote on P
20:         **if** $\rho > 0.8$ **then**
21:             $\hat{y} \leftarrow$ majority class in P
22:             $\gamma \leftarrow \frac{1}{|C|} \sum c_j$
23:             **return** $(\hat{y}, \gamma)$
24:         **end if**
25:     **end if**
26: **end for**
27: **Step 4: Weighted Aggregation of Model Outputs**
28: **for** each class label c **do**
29:     Compute $S(c) = \frac{\sum_i w_i \, c_i \, \mathbf{1}[p_i = c]}{\sum_i w_i}$
30:     where $w_i = \text{Acc}(M_i) \cdot f(c_i)$
31: **end for**
32: **Step 5: Uncertainty Estimation**
33: $\gamma = \max_c S(c)$
34: $\delta = 1 - \gamma$
35: **if** $\delta > 0.3$ **then**
36:     Re-evaluate using full ensemble M
37: **end if**
38: **return** $(\arg\max_c S(c), \max_c S(c))$

---

4. Experimental Results and Statistical Validation

Table 2 presents comprehensive performance metrics across all crop disease detection architectures. Modern efficient designs consistently outperformed traditional deep networks. The performance gap exceeded 38% between the best and worst models.

Table 2: Model Performance Metrics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV2	96.97	97.07	96.97	96.95
InceptionV3	96.78	96.88	96.78	96.76
EfficientNetB7	96.40	96.92	96.40	96.39
EfficientNetB0	76.14	82.95	76.14	75.85
ResNet50	75.19	76.94	75.19	74.56
VGG19	58.71	65.14	58.71	52.04
VGG16	57.95	59.89	57.95	52.31

There were notable differences in performance across disease categories. Figures 3–**??** present the confusion matrices and ROC curves for all evaluated architectures (EfficientNetB0, EfficientNetB7, MobileNetV2, InceptionV3, ResNet50, VGG16, and VGG19). These variations highlight both the advantages and limitations of each architecture. The results show that disease detectability varies by class. All modern architectures correctly identified bacterial leaf blight, with InceptionV3 achieving perfect classification, demonstrating the disease's strong visual characteristics. Brown spot remained the most difficult class. Lesion variability resulted in significant misclassification, with VGG16 nearly failing and MobileNetV2 performing best (F1 = 0.97). Furthermore, detection of healthy leaves was challenging, particularly for models such as EfficientNetB0, which were affected by mild early-stage symptoms. MobileNetV2 and EfficientNetB7 achieved high accuracy (F1 of 0.98), reducing the need for unnecessary treatments. Weaker models struggled with leaf blast's shifting lesion shapes, whereas stronger architectures consistently achieved F1-scores above 0.90. Leaf scald was one of the simplest classes, identifiable almost perfectly across most architectures due to its distinctive banded pattern. Narrow brown spot exhibited extremely consistent morphology, allowing top models, particularly InceptionV3 and EfficientNetB7, to perform almost flawlessly. Overall, the findings indicate that diseases with subtle or variable symptoms require more sophisticated architectures, whereas those with consistent and recognizable visual patterns are easier to categorize, as shown in Figs. 3–9.
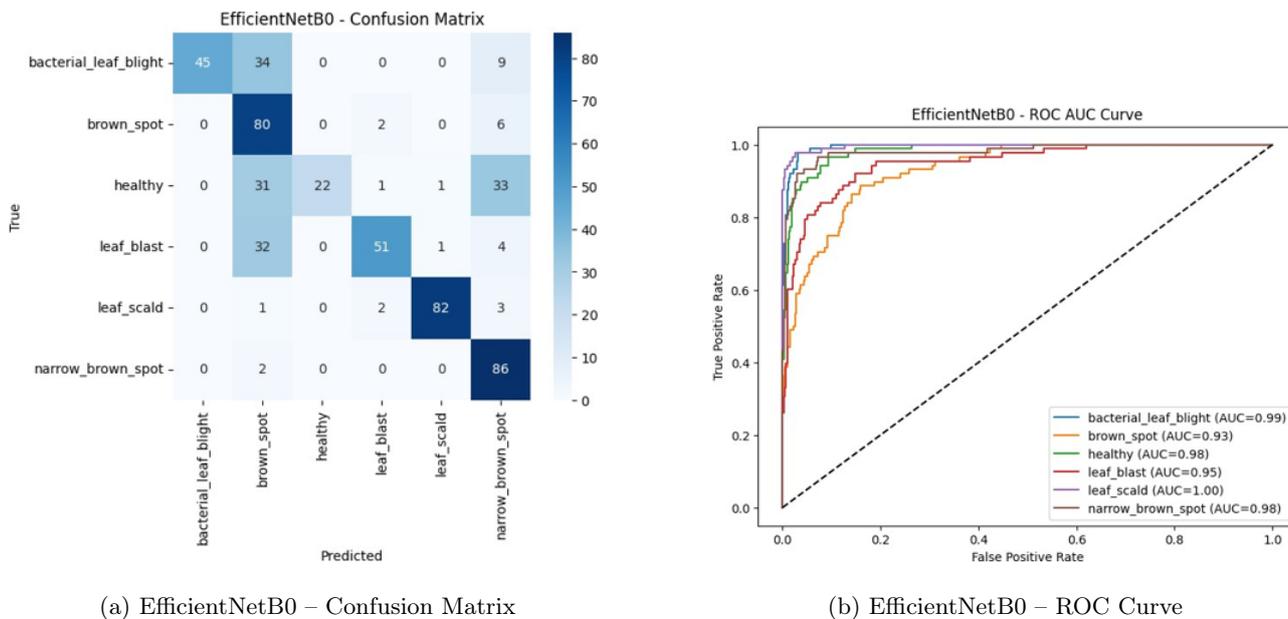

(a) EfficientNetB0 – Confusion Matrix
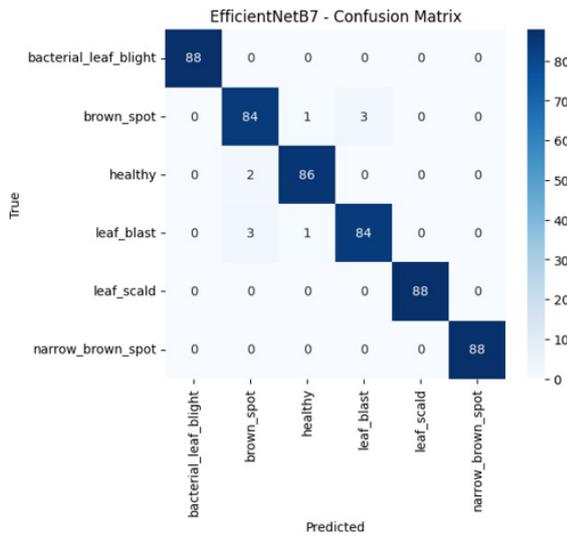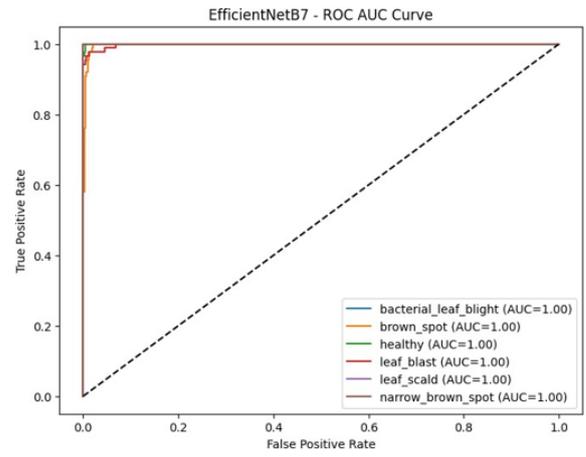
(b) EfficientNetB0 – ROC Curve

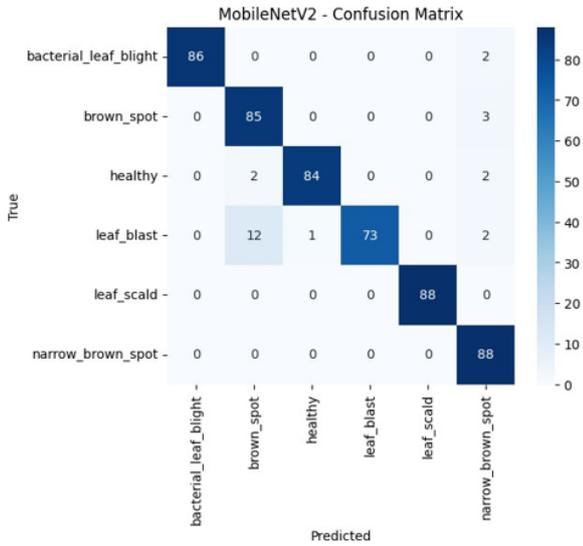Figure 3: Model diagnostics: EfficientNetB0
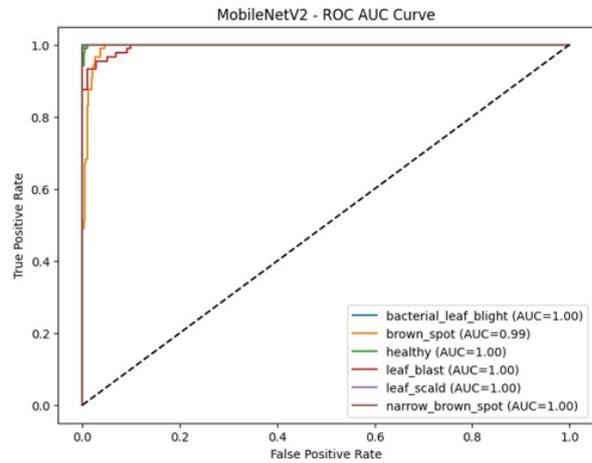
(a) EfficientNetB7 – Confusion Matrix

(b) EfficientNetB7 – ROC Curve

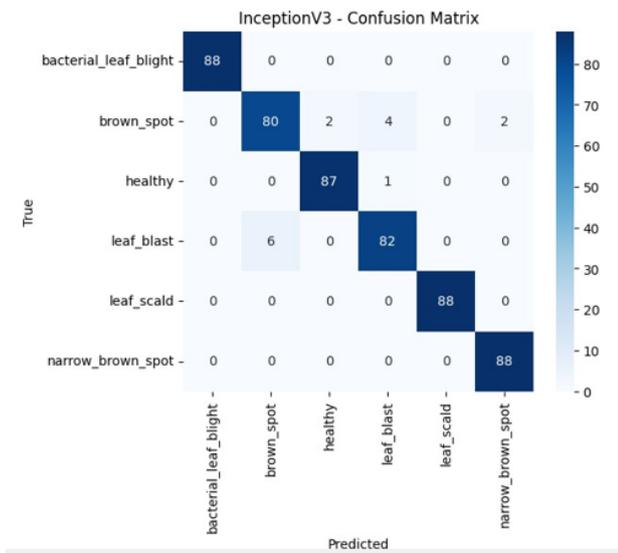Figure 4: Model diagnostics: EfficientNetB7
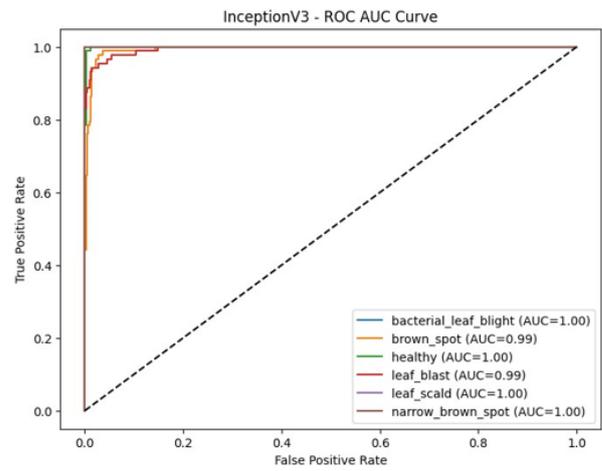

(a) MobileNetV2 – Confusion Matrix

(b) MobileNetV2 – ROC Curve

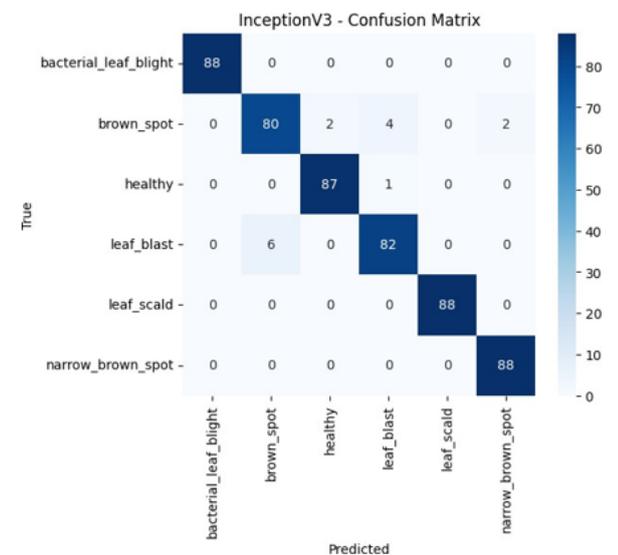Figure 5: Model diagnostics: MobileNetV2
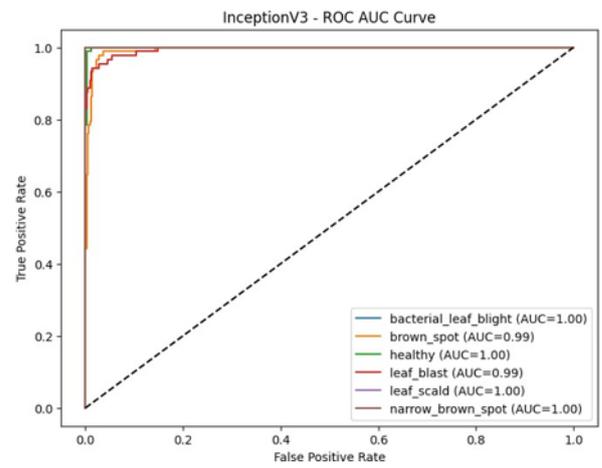
(a) InceptionV3 – Confusion Matrix

(b) InceptionV3 – ROC Curve
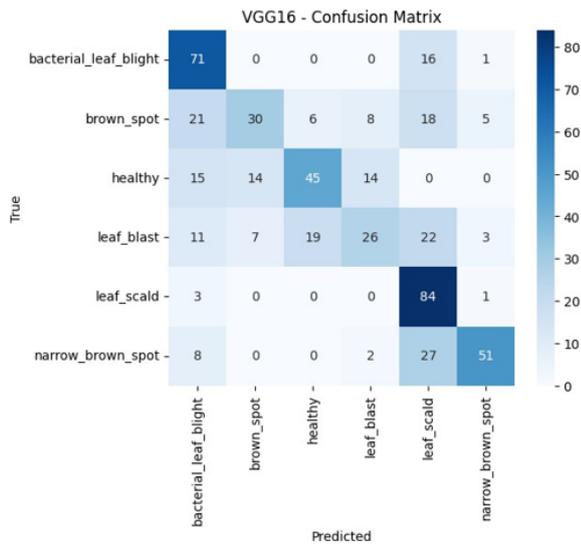
Figure 6: Model diagnostics: InceptionV3
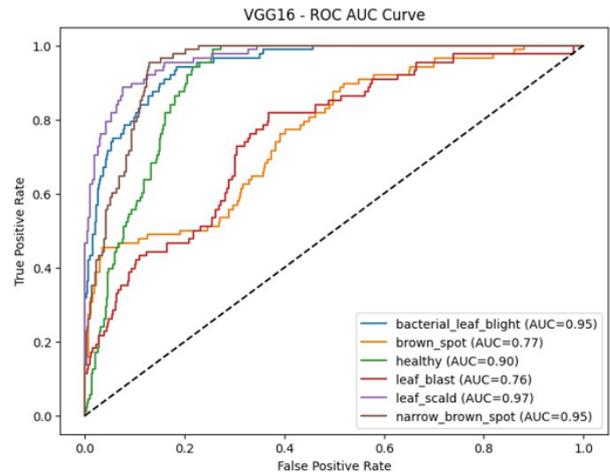

(a) ResNet50 – Confusion Matrix

(b) ResNet50 – ROC Curve

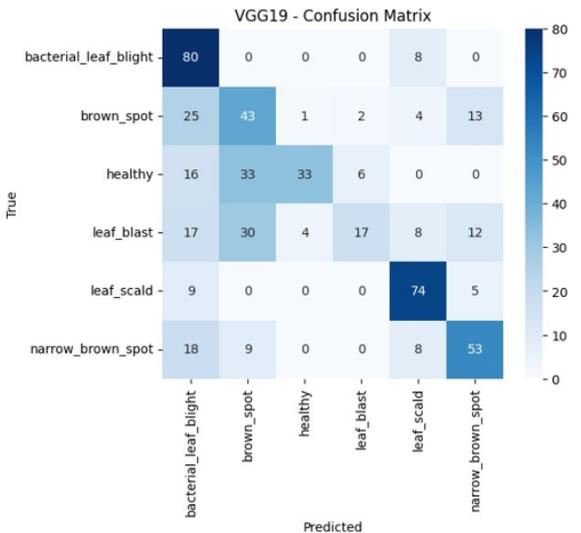Figure 7: Model diagnostics (Part III-A): ResNet50.
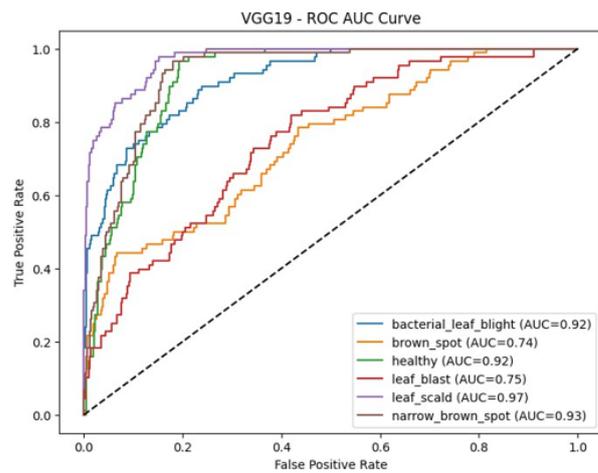
(a) VGG16 – Confusion Matrix

(b) VGG16 – ROC Curve

Figure 8: Model diagnostics (Part III-B): VGG16.

(a) VGG19 – Confusion Matrix

(b) VGG19 – ROC Curve

Figure 9: Model diagnostics (Part III-C): VGG19.

5. Discussion

The results clarify the superior performance of newer architectures compared to older deep networks. The improvement is attributed to more efficient design strategies that achieve higher accuracy with fewer computational resources. The success of MobileNetV2 is primarily attributed to depthwise separable convolutions. This modification reduces computation by approximately 8–9 times while preserving 97% accuracy, as illustrated in Fig. 10. The skip connections in ResNet50 function as conduits within the network, facilitating information flow around bottlenecks. Although this architecture achieves 75% accuracy, it still demands substantial computational resources for mobile deployment, limiting its practicality in field conditions.

InceptionV3 employs multi-scale analysis by processing images at different resolutions concurrently. This capability is particularly effective for detecting rice diseases that manifest in various sizes, ranging from small spots to extensive lesions, achieving an accuracy of 96.78%. EfficientNet is designed for optimal scaling; however, the larger B7 variant exhibited only marginal improvement compared to the smaller B0 (96.4% vs. 76.1%), indicating that increased model size does not necessarily translate to improved disease classification performance.

Disease-specific performance differences were also observed. Brown spot remains particularly challenging due to its variability under different environmental conditions. Increased humidity intensifies lesion pigmentation, whereas elevated temperature reduces visual prominence. Even high-performing models occasionally misclassified brown spot as narrow brown spot; however, MobileNetV2 achieved 97% accuracy by distinguishing subtle morphological variations. Early-stage detection presents an additional challenge, as infected leaves often resemble healthy samples. Brown spot and narrow brown spot exhibit similar visual characteristics, differing primarily in lesion width, requiring extensive training for reliable differentiation.
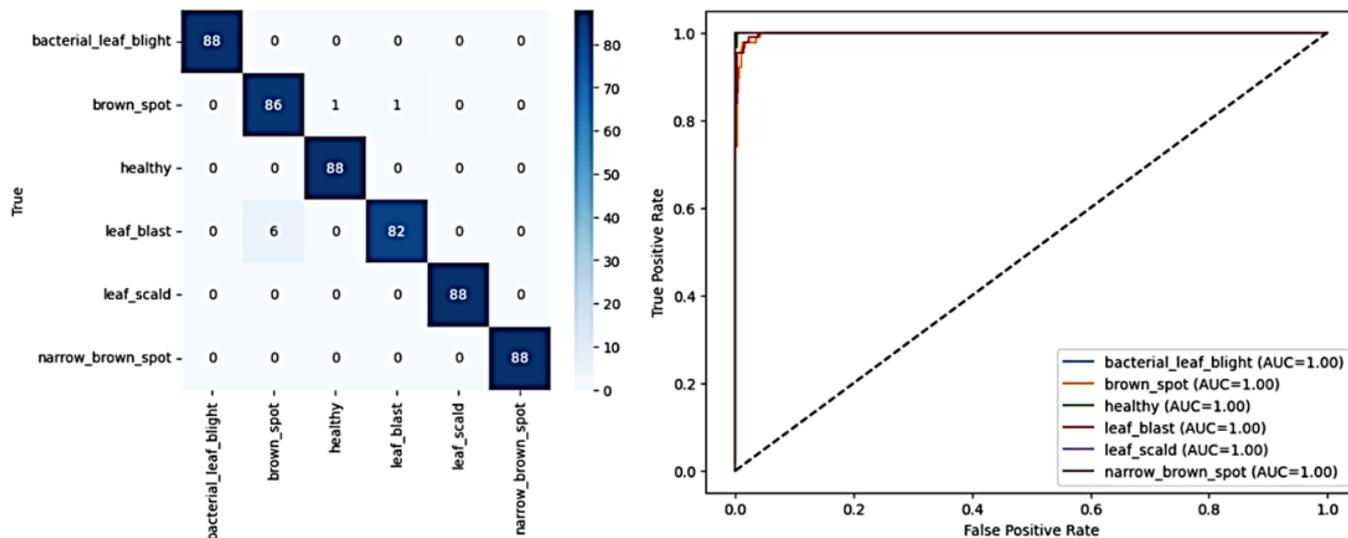

Figure 10: Performance Evaluation of the Rice Leaf Disease Classification Model: Confusion Matrix and ROC–AUC Analysis

Background clutter, including surrounding vegetation, soil, and shadows, significantly impacts traditional architectures. VGG16 exhibited sensitivity to irrelevant regions, whereas MobileNetV2 maintained a more stable feature focus. Future improvements may incorporate attention mechanisms to reduce the influence of background noise. The proposed CWDES algorithm operates by dynamically selecting model subsets based on case complexity. Simple cases require evaluation by a limited number of models, reducing computation time by 65%, whereas complex cases trigger full ensemble evaluation. The ensemble achieved 97.35% accuracy, exceeding the performance of any individual architecture. Weighted voting ensures that higher-performing models contribute proportionally more to the final prediction. The practical implications of this framework include reduced crop loss, optimized pesticide application, and improved food security. The system provides rapid diagnostic support under resource-constrained agricultural conditions. While not replacing expert knowledge, it enhances accessibility to diagnostic capabilities in field environments.

(a) Bacterial Leaf Blight (99.95%) (b) Bacterial Leaf Blight (97.62%) (c) Brown Spot (90.32%) (d) Brown Spot (98.20%)

(e) Healthy (99.98%) (f) Healthy (99.97%) (g) Leaf Blast (99.95%) (h) Leaf Blast (96.18%)

(i) Leaf Scald (99.88%) (j) Leaf Scald (99.96%) (k) Narrow Brown Spot (99.40%) (l) Narrow Brown Spot (99.09%)

Figure 11: Representative outputs of the proposed rice leaf disease detection framework across all six classes. Each class is shown with two sample predictions and associated confidence scores.

6. Conclusion

This study presents a hierarchical ensemble framework for multi-class classification of rice foliar diseases using deep learning. Experimental results indicate that lightweight architectures, particularly MobileNetV2, achieve strong classification performance while maintaining low inference latency, making them suitable for deployment on resource-constrained devices. The proposed CWDES approach integrates confidence-aware model selection and progressive inference, reducing computational cost while achieving accuracy comparable to static ensemble methods. These findings suggest that efficiency-oriented ensemble strategies can support practical agricultural disease monitoring applications. While the framework demonstrates promising performance under the evaluated dataset and conditions, further validation on larger and more diverse field datasets is required to assess robustness and generalization. Future work will explore attention mechanisms and domain adaptation to improve early-stage disease detection and real-world applicability.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics Approval

This study does not involve human participants, animals, or sensitive personal data. Therefore, ethical approval and informed consent were not required.

Data Availability and Transparency

The dataset supporting the findings of this study is publicly available at: https://www.kaggle.com/datasets/vbookshelf/rice-leaf-diseases. No new proprietary data were created. All preprocessing steps, model configurations, and evaluation protocols are described in the Methods section to ensure reproducibility. The implementation code is available from the corresponding author upon reasonable request.

AI Use Disclosure

The authors declare that no generative artificial intelligence (AI) tools were used in the design of the study, data analysis, interpretation of results, or preparation of the manuscript content. AI tools were used solely for language formatting and editorial assistance under journal copyediting guidance. The authors take full responsibility for the scientific content of this manuscript.

Author Contributions

Chanchal Ghosh: Methodology; **Biplab Kanti Das**: Conceptualization, Methodology; **Tapashri Sur**: Formal Analysis; **Prasanta Mazumdar**: Investigation; **Pratik Kumar Halder**: Investigation; **Sukanta Kundu**: Data Curation; **Subhojeet Prasad**: Writing – Review and Editing

References

[1] R. Rahman, P. Arko, M. E. Ali, M. Khan, S. H. Apon, F. Nowrin, and A. Wasif, "Identification and recognition of rice diseases and pests using convolutional neural networks," *Biosystems Engineering*, vol. 194, pp. 112–120, 2020.

[2] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, "Detection of rice plant diseases based on deep transfer learning," *Journal of the Science of Food and Agriculture*, vol. 100, no. 7, pp. 3246–3256, 2020.

[3] Y. Wang, H. Wang, and Z. Peng, "Rice diseases detection and classification using attention based neural network and bayesian optimization," *Expert Systems with Applications*, vol. 178, p. 114770, 2021.

[4] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, 2017.

[5] H. B. Prajapati, J. P. Shah, and V. K. Dabhi, "Detection and classification of rice plant diseases," *Intelligent Decision Technologies*, vol. 11, no. 3, pp. 357–373, 2017.

[6] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Deep feature based rice leaf disease identification using support vector machine," *Computers and Electronics in Agriculture*, vol. 175, p. 105527, 2020.

[7] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with jaya algorithm," *Information Processing in Agriculture*, vol. 7, no. 2, pp. 249–260, 2020.

[8] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," *Plant Methods*, vol. 17, no. 1, p. 22, 2021.

[9] U. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using efficientnet deep learning model," *Ecological Informatics*, vol. 61, p. 101182, 2021.

[10] N. Krishnamoorthy, L. N. Prasad, C. P. Kumar, B. Subedi, H. B. Abraha, and V. Sathishkumar, "Rice leaf diseases prediction using deep neural networks with transfer learning," *Environmental Research*, vol. 198, p. 111275, 2021.

[11] T. Islam, M. Sah, S. Baral, and R. R. Choudhury, "A faster technique on rice disease detection using image processing of affected area in agro-field," in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 62–66, IEEE, 2018.

[12] K. Ahmed, T. R. Shahidi, S. M. I. Alam, and S. Momen, "Rice leaf disease detection using machine learning techniques," in *2019 International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–5, IEEE, 2019.

[13] W.-J. Liang, H. Zhang, G.-F. Zhang, and H.-X. Cao, "Rice blast disease recognition using a deep convolutional neural network," *Scientific Reports*, vol. 9, no. 1, pp. 1–10, 2019.

[14] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.

[15] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep learning for plant diseases: Detection and saliency map visualisation," in *Human and Machine Learning: Visible, Explainable, Trustworthy and Transparent* (J. Zhou and F. Chen, eds.), Human–Computer Interaction Series, pp. 93–117, Cham: Springer International Publishing, 2018.

[16] S. Verma, A. Chug, and A. P. Singh, "Recent advancements in image-based prediction models for diagnosis of plant diseases," in *Proceedings of 3rd International Conference on Computer Vision and Image Processing: CVIP 2018, Volume 1*, pp. 365–377, Springer, 2019.

[17] V. Singh, N. Sharma, and S. Singh, "A review of imaging techniques for plant disease detection," *Artificial Intelligence in Agriculture*, vol. 4, pp. 229–242, 2020.

[18] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019.

[19] S. B. Jadhav, V. R. Udupi, and S. B. Patil, "Identification of plant diseases using convolutional neural networks," *International Journal of Information Technology*, vol. 13, no. 6, pp. 2461–2470, 2021.

[20] G. Geetharamani and A. Pandian, "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers & Electrical Engineering*, vol. 76, pp. 323–338, 2019.

[21] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep neural networks with transfer learning in millet crop images," *Computers in Industry*, vol. 108, pp. 115–120, 2019.

[22] T. Daniya and S. Vigneshwari, "Deep neural network for disease detection in rice plant using the texture and deep features," *The Computer Journal*, vol. 65, no. 7, pp. 1812–1825, 2022.

[23] C. Malathi and J. Sheela, "Deep dynamic classification (ddc) for plant disease detection," in *AIP Conference Proceedings*, vol. 2869, p. 050015, AIP Publishing LLC, 2023.

[24] B. S. Bari, M. N. Islam, M. Rashid, M. J. Hasan, M. A. M. Razman, R. M. Musa, A. F. Ab Nasir, and A. P. A. Majeed, "A real-time approach of diagnosing rice leaf disease using deep learning-based faster r-cnn framework," *PeerJ Computer Science*, vol. 7, p. e432, 2021.

[25] R. M. O. Cruz, R. Sabourin, and G. D. C. Cavalcanti, "Dynamic ensemble selection: A comprehensive review," *Pattern Recognition*, vol. 85, pp. 286–302, 2019.

[26] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.