JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

MACHINE LEARNING WITH LIGHTWEIGHT CNN (RESNET-18) FOR EARLY DETECTION OF RICE LEAF DISEASES

Poningsih[1*]; Suhendra[2]; Ahmad Zamsuri[3]

Department of Informatics, Master's Program [1]
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia [1]
https://stikomtunasbangsa.ac.id/[1]
poningsih@amiktunasbangsa.ac.id*

Publishing Study Program[2]
Politeknik Negeri Media Kreatif [2]
https://polimedia.ac.id/[2]
ragilsuhendra030514@gmail.com

Master's Program in Computer Science[3]
Universitas Lancang Kuning[3]
https://www.unilak.ac.id/[3]
ahmadzamsuri@unilak.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract*—Rice leaf diseases such as blast and brown spot significantly threaten rice productivity, especially in agrarian countries like Indonesia. Manual diagnosis methods remain subjective, slow, and inconsistent across field conditions, highlighting the need for an automated and reliable detection system. This study presents a lightweight deep learning framework for the automatic classification of rice leaf diseases from image data. To assess its effectiveness, four Convolutional Neural Network (CNN) architectures ResNet-18, VGG-16, Inception V3, and MobileNetV2 were evaluated. The dataset, obtained from Kaggle, consists of three classes healthy, blast-infected, and brown spot with all images preprocessed through normalization and augmentation before being split into training and validation sets. Experimental results show that ResNet-18 achieves the best overall performance, with 96.94% accuracy, 100% precision, 95.45% recall, an F1-score of 96.18%, and an AUC of 1.0000. Compared to the other architectures, ResNet-18 demonstrates higher stability, stronger generalization, and lower overfitting tendencies while maintaining computational efficiency. The findings indicate that ResNet-18 is a promising lightweight model for practical deployment in mobile or IoT-based agricultural monitoring systems, supporting early disease detection and enhancing local food security efforts.*

Keywords*: Deep Learning, Image Classification, Plant Disease Detection, ResNet-18, Rice Leaf Disease*

Intisari—*Penyakit daun padi seperti blast dan brown spot menjadi ancaman serius bagi produktivitas padi, terutama di negara agraris seperti Indonesia. Diagnosis manual masih bersifat subjektif, lambat, dan kurang konsisten, sehingga diperlukan sistem deteksi otomatis yang andal. Penelitian ini mengusulkan kerangka deep learning ringan untuk klasifikasi penyakit daun padi berbasis citra. Empat arsitektur CNN, yaitu ResNet-18, VGG-16, Inception V3, dan MobileNetV2, dievaluasi menggunakan dataset Kaggle yang terdiri atas tiga kelas: healthy, blast, dan brown spot. Seluruh citra dipraolah melalui normalisasi dan augmentasi sebelum dibagi ke dalam data pelatihan dan validasi. Hasil penelitian menunjukkan bahwa ResNet-18 memberikan performa terbaik dengan akurasi 96,94%, presisi 100%, recall 95,45%, F1-score 96,18%, dan AUC 1,0000. Dibandingkan model lain, ResNet-18 lebih stabil, memiliki generalisasi lebih baik, overfitting lebih rendah, serta tetap efisien secara komputasi. Temuan ini menunjukkan bahwa ResNet-18 berpotensi diterapkan pada sistem* *pemantauan pertanian berbasis mobile atau IoT untuk mendukung deteksi dini penyakit dan penguatan ketahanan pangan lokal.*

Kata Kunci: *Pembelajaran Mendalam, Klasifikasi Citra, Deteksi Penyakit Tanaman, ResNet-18, Penyakit Daun Padi.*

INTRODUCTION

Rice is a strategic staple commodity and serves as the backbone of national food security in Indonesia [1], [2]. According to data from the Indonesian Central Statistics Agency (BPS), over 90% of the Indonesian population consumes rice as their primary food source. Consequently, disruptions to rice production whether caused by climate variability, pests, or plant diseases—can directly affect the nation's food stability and economic resilience. One of the major challenges in rice cultivation is the prevalence of leaf diseases, such as blast (Pyricularia oryzae), brown spot (Cochliobolus miyabeanus), and bacterial leaf blight (Xanthomonas oryzae) [3], [4]. These diseases not only reduce crop productivity but also degrade the quality of the harvest. In severe cases, infections can lead to crop losses ranging from 40% to 80%, depending on the type of disease and the timeliness of intervention [5], [6], [7].

Early detection of leaf diseases is a critical component of Integrated Pest Management (IPM) and precision agriculture strategies [8], [9]. However, conventional detection methods typically involving visual inspection by farmers or agricultural officers suffer from several limitations. These include subjectivity, time inefficiency, and a lack of consistent expertise across farming regions [10]. These challenges highlight the need for intelligent and automated systems capable of accurately identifying early signs of leaf disease [11], [12]. In the domain of rice leaf disease detection via digital imagery, various Convolutional Neural Network (CNN) architectures have been widely adopted, including ResNet-18 [13], [14], [15], VGG-16 [16], [17], Inception V3 [18], [19], [20], [21], and MobileNetV2 [22], [23]. Each model offers distinct advantages and drawbacks. ResNet-18, for instance, utilizes residual connections to mitigate accuracy degradation in deeper networks, enabling more stable training and high accuracy albeit with slightly more complexity than lightweight models [17], [24], [25]. VGG-16 is known for its straightforward architecture and ease of implementation but tends to overfit and requires a large number of parameters, making it less efficient on limited datasets [26], [27]. Inception V3 leverages multi-scale feature extraction through its Inception modules, which are effective for recognizing complex patterns in infected leaves. However, the model demands longer training times and greater computational resources [21], [28]. MobileNetV2, on the other hand, is optimized for efficiency and lightweight deployment on mobile or IoT devices, though it sacrifices a degree of accuracy compared to larger models [29], [30], [31], [32]. Selecting the appropriate CNN architecture thus requires balancing accuracy, computational efficiency, and the practical constraints of deployment in the field.

Previous research by [33] proposed a technically advanced and modern approach by combining CNN and transfer learning using the ResNeXt architecture, which achieved a high accuracy of 99.22% in detecting rice leaf diseases such as blast, bacterial blight, and brown spot. The methodology included comprehensive steps such as real-world data acquisition, data augmentation, and model fine-tuning, making it both relevant and practical for agricultural applications. However, the approach heavily relies on transfer learning and large datasets, with limited generalizability since it was only trained on leaf images and did not consider other plant parts. Moreover, its applicability under varied field conditions at scale remains unexplored. In another study, [34] introduced "PlantDet: A Robust Multi-Model Ensemble Method Based on Deep Learning For Plant Disease Detection," which utilizes a powerful ensemble of architectures (InceptionResNetV2, EfficientNetV2L, and Xception).

This method effectively addresses underfitting and overfitting challenges, achieving an impressive accuracy of up to 98.53% on rice leaf datasets. The inclusion of explainable AI tools such as Grad-CAM++ and Score-CAM further enhances model interpretability and user trust. Nonetheless, the ensemble's architectural complexity results in prolonged training times and interpretability challenges, and the models have yet to be evaluated under large-scale or real-world field conditions. With the rapid advancement of Artificial Intelligence (AI), particularly in the field of Computer Vision, there is immense potential to support agricultural systems more efficiently [35], [36], [37]. One of the most promising techniques is deep learning, especially CNN-based algorithms, which excel at learning complex patterns in digital images. CNNs have proven highly effective in

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

various image classification tasks, including object detection, image segmentation, and pattern recognition applications that are particularly relevant to plant disease identification using leaf images [38], [39], [40], [41].

While several studies have demonstrated the effectiveness of CNNs in detecting diseases in crops such as tomatoes, potatoes, apples, and maize [42], their specific application to rice particularly to rice leaf disease remains relatively limited. Further development is required, especially in alignment with the unique conditions of local Indonesian agriculture. Reliable and efficient deep learning models are expected to facilitate accurate disease detection from images of rice leaves captured using conventional cameras, smartphones, or agricultural drones [35], [43]. In the context of local food security, early detection of rice diseases can enhance land management efficiency, reduce production costs by enabling more targeted pesticide use, and prevent large-scale disease outbreaks across fields. This aligns with global sustainable development efforts, especially SDG Goal 2 (Zero Hunger) and Goal 12 (Responsible Consumption and Production) [44], [45], [46].

A critical research gap remains regarding the identification of models that achieve an optimal balance between classification performance and computational efficiency especially for deployment on mobile or IoT-based agricultural devices. While large CNN architectures demonstrate high accuracy, their memory footprint and processing requirements limit practical implementation in rural farming contexts. Conversely, ultralight models may sacrifice accuracy, thus reducing reliability in real-time diagnosis. To address this gap, this study focuses on evaluating lightweight CNN architectures and proposes the following explicit hypothesis, ResNet-18 provides the most optimal balance between accuracy, generalization capability, and computational efficiency compared to VGG-16, Inception V3, and MobileNetV2 for rice leaf disease classification. This hypothesis is grounded in the residual learning mechanism of ResNet-18, which enables stable gradient flow and prevents performance degradation in deeper networks while maintaining a relatively small number of parameters. Compared to its counterparts, ResNet-18 is widely recognized for its robustness on limited datasets and its suitability for real-time or resource-constrained applications.

This study aims to evaluate and implement several deep learning architectures for rice leaf disease classification. The models examined include ResNet-18, VGG-16, and Inception V3, all of which are well-established in image classification tasks. Model performance is assessed through training accuracy, validation accuracy, and loss metrics to measure convergence and detect overfitting. By adopting an experimental approach grounded in classified rice leaf image data, this research seeks to contribute both practically and theoretically to the development of AI-powered plant disease detection systems tailored to the needs of Indonesian agriculture. Accordingly, the implementation of deep learning in this context serves not only as a technological solution but also as a critical instrument for building a smarter, more responsive, and sustainable agricultural ecosystem. This research proposes ResNet-18 as a practical and efficient model based on its consistent performance across all metrics compared to other commonly used CNN architectures.

MATERIALS AND METHODS

This section outlines the methodology employed in the study, beginning with the utilization of a rice leaf image dataset sourced from Kaggle.com. The dataset comprises three primary classes: healthy leaves, blast-infected leaves, and leaves affected by brown spot. Prior to model training, the dataset underwent preprocessing procedures, including data augmentation and normalization, to enhance data variability and ensure balanced class distribution. The processed data was then partitioned into training and validation sets. The study compares the performance of several deep learning architectures, namely VGG-16, Inception V3, and the ResNet-18. ResNet-18 was selected as the primary model due to its parameter efficiency and training stability, achieved through residual learning, which addresses the degradation problem in deeper networks. An experimental research design was adopted, encompassing model training and performance evaluation using accuracy and loss metrics. This approach was intended to identify the most optimal model for early detection of rice leaf diseases, based on image classification.

Dataset

This study utilized a rice leaf image dataset obtained from Kaggle.com, an open-access, community-driven platform widely used for research and development in artificial intelligence. The dataset, titled “Rice Leaf Diseases Dataset”, consists of three main categories: healthy rice leaves, leaves infected with blast, and leaves affected by brown spot. The dataset contains thousands of labeled images, each categorized according to its respective disease class. All images are in JPEG format, with varying resolutions, and were captured under diverse lighting conditions and backgrounds. This variability reflects real-world environmental conditions commonly encountered in agricultural fields, thereby enhancing the dataset's representativeness and applicability. An example of the dataset can be seen in Figure 1 and dataset source from link https://www.kaggle.com/code/umairinayat/rice-disease-classification-cnn/input.

Source: (Research Results, 2025)

Figure 1. Sample of Research Dataset

As illustrated in Figure 1, the initial stage involved curating all image data obtained from Kaggle to ensure consistent quality and balanced class representation. The dataset was then divided into two main subsets: a training set (80%) and a validation set (20%). This split was performed randomly while maintaining a proportional class distribution, allowing the model to learn from representative data and be evaluated on previously unseen samples. To improve model generalization and minimize overfitting, data augmentation was applied using image rotation, horizontal and vertical flipping, and random zooming. These transformations increased dataset diversity and simulated variations in leaf appearance that may occur under real-world conditions. While this research demonstrates high performance on benchmark datasets, the absence of validation using real-world field data remains a key limitation. Future work will focus on testing the trained ResNet-18 model on field-acquired images captured via mobile devices and drones under diverse environmental conditions. This step will enable evaluation of model robustness and practical deployment feasibility in precision agriculture scenarios.

Model Comparison Architecture

This study presents a comparative evaluation of four CNN architectures: ResNet-18, VGG-16, Inception V3, and MobileNetV2. Each model was trained and evaluated under the same experimental setting to enable a fair comparison in terms of classification performance, parameter efficiency, and suitability for practical deployment. Among these architectures, ResNet-18 was selected as the primary model because it provided the most balanced trade-off between validation accuracy, training stability, and computational efficiency. VGG-16 is known for its deep yet straightforward architecture, employing a consistent sequence of convolutional layers. In contrast, Inception V3 adopts a multi-scale approach, using varied filter sizes within a single module to capture features at different spatial resolutions simultaneously.

ResNet-18 was selected as the model in this study due to its notable advantage in mitigating performance degradation in deep networks through the implementation of residual learning. Residual Networks (ResNets) utilize shortcut connections or skip pathways that allow information to bypass certain layers without modification, thereby accelerating convergence and reducing the risk of vanishing gradients. In the context of rice leaf image classification, ResNet-18 effectively identifies relevant visual patterns such as lesions, discoloration, and texture anomalies associated with leaf diseases in a more stable and consistent manner. Based on experimental results, ResNet-18 outperformed the other models in terms of validation accuracy and robustness, making it a strong reference model for developing automated rice disease detection systems. A detailed comparison of the architectural parameters and performance metrics is presented in Table 1.

Table 1. Comparative Analysis and Architectural Parameters of CNN Models

Parameter	ResNet-18	VGG-16	MobileNetV2	Inception V3
Pretrained Weights	IMAGENET1K_V1	IMAGENET1K_V1	IMAGENET1K_V1	IMAGENET1K_V1
Fine-tuning Layer	model.fc adjusted to match number of classes	classifier[6] modified	classifier[1] modified	model.fc adjusted to match number of classes
Loss Function	CrossEntropyLoss	CrossEntropyLoss	CrossEntropyLoss	CrossEntropyLoss
Optimizer	Adam (learning rate = 0.0001)	Adam (learning rate = 0.0001)	Adam (learning rate = 0.0001)	Adam (learning rate = 0.0001)
Network Depth	18 layers	16 convolutional layers + 3 fully connected layers	~53 layers	~48 layers with Inception modules
Number of Parameters	~11 million	~138 million	~3.5 million	~23 million
Input Image Size	224 × 224 pixels	224 × 224 pixels	224 × 224 pixels	299 × 299 pixels
Training Time (Estimate)	Moderate (relatively efficient)	Fast but prone to overfitting	Very fast and lightweight	Considerably slow due to larger input size

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

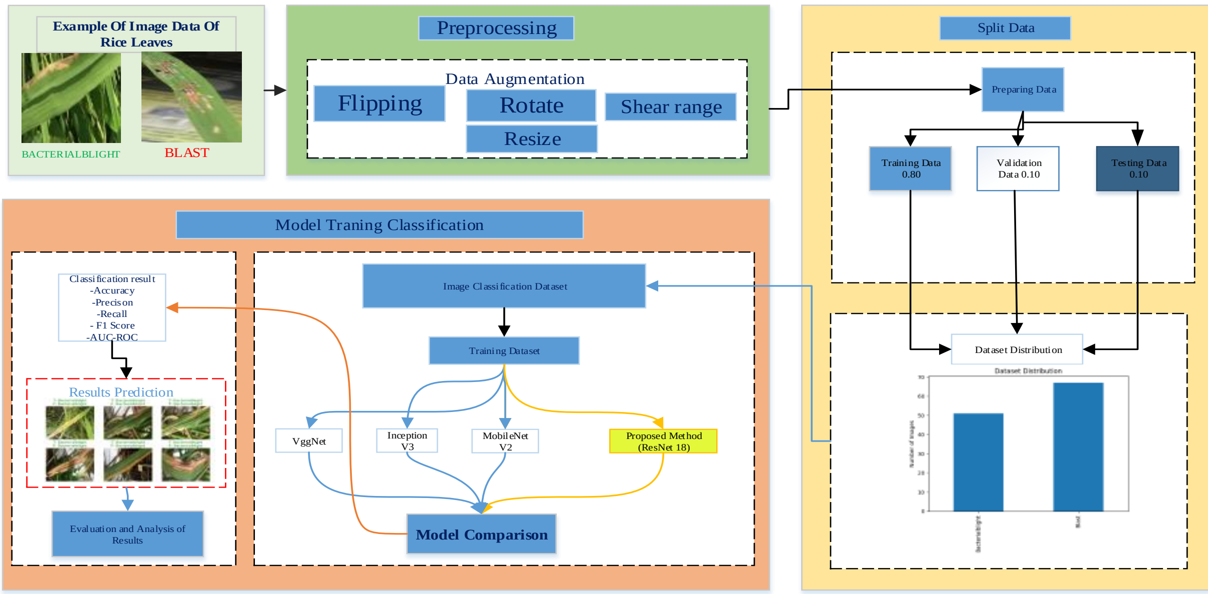
Parameter	ResNet-18	VGG-16	MobileNetV2	Inception V3
Validation Performance	Highest accuracy (97.20%), lowest loss (0.0675)	Lower accuracy (86.29%), high fluctuation in loss	Excellent validation performance (97.46% validation accuracy; 0.0445 validation loss)	High accuracy (95.97%), stable loss (0.1537)
Overfitting Potential	Minimal	High (significant)	Low	Low (stabilized after epoch 4)
Mobile Compatibility	Good	Poor suitability	Very suitable (optimized for mobile deployment)	Less efficient for resource-constrained devices
Notable Characteristics	Residual learning; stable on small datasets	Simple architecture; prone to overfitting	Lightweight and resource-efficient	Multi-scale feature extraction; strong classification performance

Source: (Research Results, 2025)

Table 1 presents a comprehensive comparison across architectural, technical, and performance aspects of the evaluated CNN models. Among them, ResNet-18 stands out for its stability and efficiency offering a relatively low number of parameters while achieving the highest validation accuracy with minimal signs of overfitting. Its residual architecture also accelerates the training process and enhances the model's resilience against performance degradation in deeper networks. These qualities make ResNet-18 highly suitable for lightweight, efficient plant disease detection systems that can be deployed on mobile devices or integrated into agricultural IoT platforms. Although VGG-16 is widely recognized for its simplicity, it tends to suffer from overfitting due to its large model size and deep sequential layer design. MobileNetV2, by contrast, is architecturally optimized for mobile and resource-constrained environments, making it well-suited for low-power applications; however, its performance in this study could not be assessed due to the lack of available evaluation data. Inception V3 demonstrated strong classification performance but comes with greater architectural complexity and higher computational demands, which may limit its practicality in real-time or embedded systems.

Research Design

This study adopts a quantitative experimental approach consisting of several key stages: data collection, preprocessing, model construction, training, performance evaluation, and result analysis. The process began by downloading a labeled rice leaf disease dataset from Kaggle, which was categorized according to disease type. During preprocessing, all images were resized to 224×224 pixels and normalized to meet the input requirements of the CNN architectures. Data augmentation was also applied to improve image variability and enhance model generalization. A schematic overview of the research design is presented in Figure 2.



Source: (Research Results, 2025)

Figure 2. Research Design Flow

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

Figure 2 illustrates the overall workflow of the study for rice leaf disease classification using deep learning. The process begins with dataset collection, followed by preprocessing, which includes image standardization and augmentation to improve data diversity and model robustness. The processed dataset is then divided into training and validation sets before model training and evaluation. Once preprocessing is complete, the dataset is split into training and validation sets, ensuring a balanced class distribution.

The core stage involves training classification models using various CNN architectures namely VGG-16, Inception V3, and the ResNet-18. Each preprocessed dataset is fed into these models, and their predictions are compared to evaluate performance based on accuracy and training stability. This evaluation is carried out to assess each model's effectiveness in automatically classifying rice leaf disease types.

The entire workflow reflects an experimental research approach, emphasizing improved accuracy in plant disease detection through image-based analysis, as part of a precision agriculture system. Following preprocessing, the models—VGG-16, Inception V3, MobileNetV2, and ResNet-18 were implemented using the PyTorch deep learning framework. Key hyperparameters, including the number of epochs (10), batch size (32), and learning rate (0.001), were chosen based on initial experiments to achieve a balance between accuracy and training efficiency.

During training, the training set was used to update model weights, while the validation set served to monitor model performance on unseen data. Each model was evaluated using four key metrics: training accuracy, validation accuracy, training loss, and validation loss. The recorded training outcomes were then analyzed to assess the consistency and stability of each model's learning process.

The primary objective of this research design is to identify the most optimal model for automatic detection of rice leaf diseases. Evaluation criteria were not limited to accuracy alone but also included training stability, resistance to overfitting, and efficiency of convergence. Through this experimental approach, the study aims to contribute practical advancements in AI-powered precision agriculture systems, particularly in the early detection of crop diseases to support local food security efforts in Indonesia.

RESULTS AND DISCUSSION

This study evaluated the effectiveness of four deep learning architectures—ResNet-18, VGG-16, Inception V3, and MobileNetV2—for early detection of rice leaf diseases using digital images. While MobileNetV2 lacked complete evaluative data, the other models were assessed across 10 epochs using four key metrics: training accuracy, validation accuracy, training loss, and validation loss. The research followed a structured experimental workflow, starting from preprocessing and classification of rice leaf images into three categories: healthy, blast-infected, and bacterial blight. These real-world images provided a representative dataset for training the CNN models in automatic disease recognition. The entire process—from data input to model evaluation—demonstrated the practical potential of AI-based approaches in supporting precision agriculture.

Preprocessing and Data Splitting Results

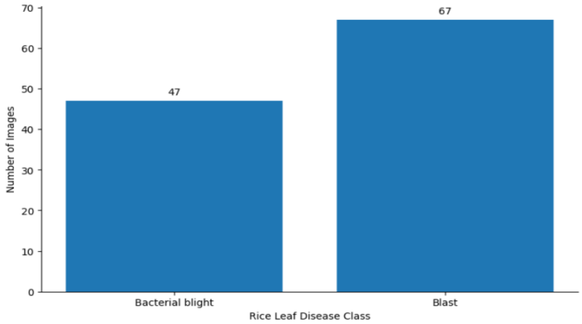
The initial stage of this study involved data preprocessing, where rice leaf images obtained from the Kaggle dataset underwent a series of preparatory steps. The primary objective of this phase was to enhance the quality and diversity of the dataset, enabling the deep learning models to more effectively recognize key features within the images. Examples of the augmented data can be seen in Figure 3.

Source: (Research Results, 2025)
Figure 3. Data Augmentation

As shown in Figure 3, various augmentation techniques, including rotation, flipping, shearing, and resizing were applied to increase image variability while maintaining correct class labels. All images were resized to 224×224 pixels, except for those used with Inception V3, which required 299×299 pixels to fit its input layer. After preprocessing, the dataset was split into 80% training and 20% validation using stratified sampling, ensuring balanced representation of all three classes: healthy, blast, and brown spot. The class distribution is visualized in Figure 4.

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

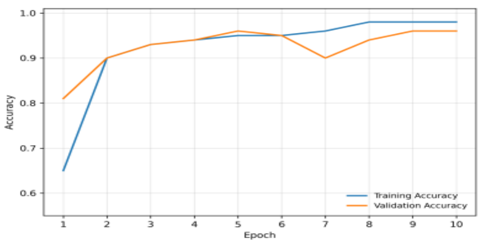


Source: (Research Results, 2025)
Figure 4. Data Distribution

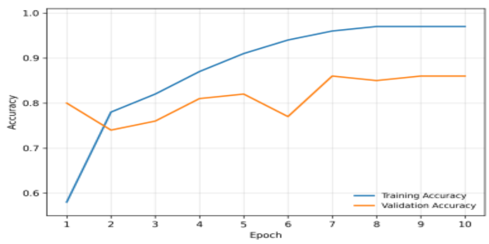
As shown in Figure 4, the dataset was split using a stratified approach to ensure that each subset maintained a balanced proportion of classes. The visualized distribution confirms that all classes are proportionally represented, which is essential for preventing classification bias during training. This step is critical for ensuring both the accuracy and generalization capability of the model when applied to real-world data scenarios in agricultural environments.

Model Training and Image Classification

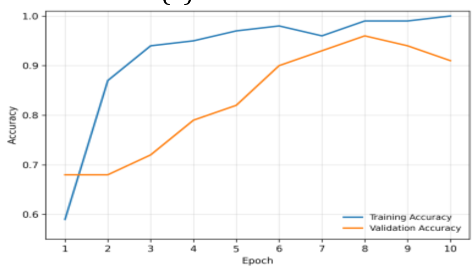
Once the data was prepared, the next step involved training the models using three CNN architectures: VGG-16, Inception V3, and ResNet-18. Each model was initialized with pretrained weights from ImageNet (IMAGENET1K_V1) to leverage transfer learning. The final classification layer of each model was modified to match the number of target classes in the dataset, which includes three categories of rice leaf conditions. All models were trained using the CrossEntropyLoss loss function and the Adam optimizer with a learning rate of 0.0001, over the course of 10 epochs. The training results are visualized in Figures 5 and 6.



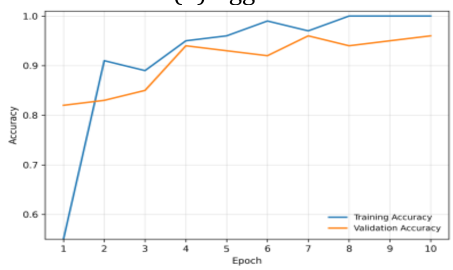
(a) ResNet 18



(b) Vgg16



(c) MobileNet V2

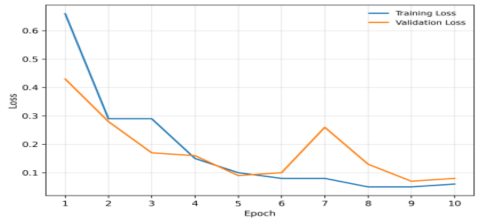


(d) Inception V3

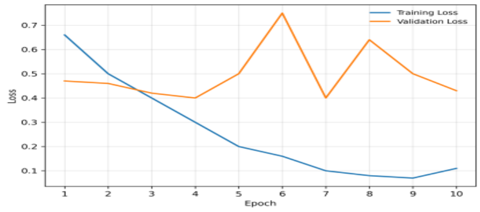
Source: (Research Results, 2025)
Figure 5. Training and Validation Accuracy of the Four Evaluated Models

Figure 5 presents the training and validation accuracy curves of the four evaluated CNN architectures across 10 epochs. Figure 5(a) demonstrated the most stable learning behavior, as the two curves remained closely aligned throughout the training process, indicating strong generalization with minimal overfitting. Figure 5(b) showed a different pattern, where training accuracy increased steadily but validation accuracy remained considerably lower and more fluctuating, suggesting a higher tendency toward overfitting. Figure 5(c) achieved rapid improvements in training accuracy and maintained strong validation performance, reflecting both efficient learning and good generalization capability, although a slight decline in validation accuracy appeared in the final epochs. Figure 5(d) also exhibited strong performance, with training accuracy approaching perfection and validation accuracy remaining high, indicating effective feature learning with only a limited generalization gap. Overall, ResNet-18 and MobileNetV2 displayed the most balanced performance, while VGG-16 showed the weakest validation behavior among the evaluated models. Loss comparisons are detailed in Figure 6.

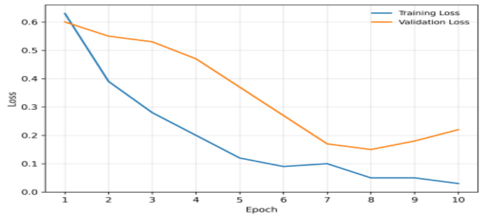
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358



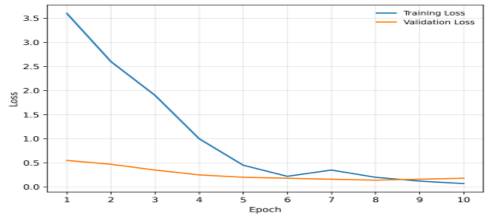
(a) ResNet 18



(b) Vgg16



(c) MobileNet V2



(d) Inception V3

Source: (Research Results, 2025)

Figure 6. Training and Validation Loss Curves for the Four Evaluated Models

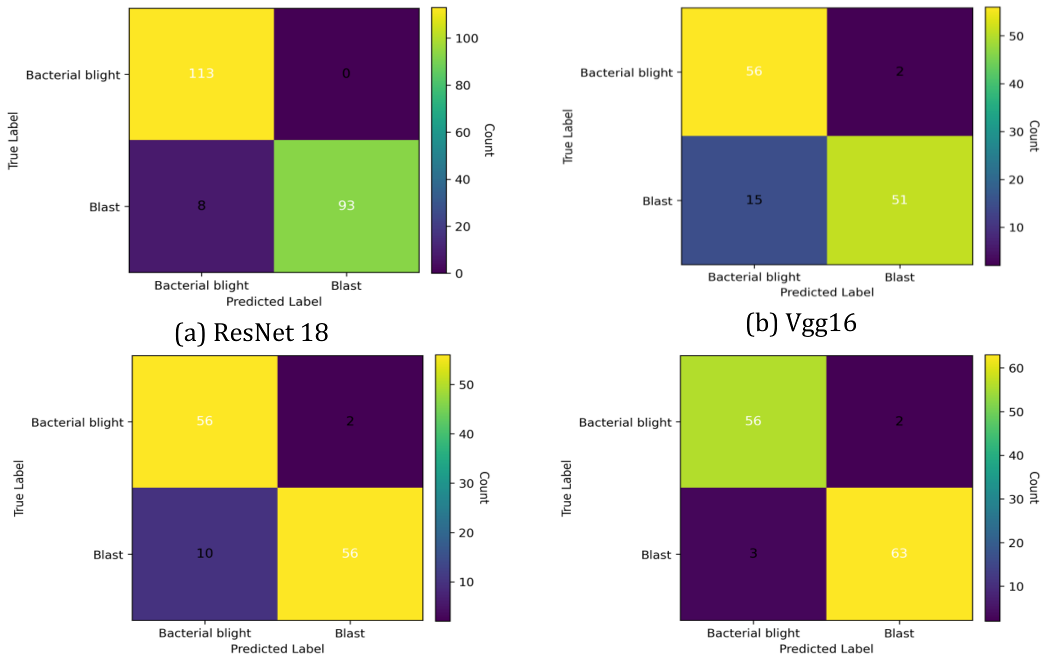
Figure 6 presents the training and validation loss curves of the four evaluated CNN architectures across 10 epochs, providing further insight into model convergence behavior and generalization performance. For figure 6(a), both training loss and validation loss decreased substantially during the training process, indicating effective learning and stable optimization. The two curves remained relatively close throughout most epochs, although a temporary increase in validation loss was observed around the middle of training. This fluctuation was not persistent, as the validation loss decreased again in subsequent epochs. Overall, the loss pattern of ResNet-18 suggests good convergence behavior and a low tendency toward overfitting. In the case of figure 6(b), the training loss consistently decreased over time, showing that the model continued to fit the training data effectively.

However, the validation loss remained considerably higher and exhibited noticeable fluctuations, including several sharp increases during later epochs. This pattern indicates that, despite good optimization on the training set, the model struggled to maintain stable performance on unseen data. Such behavior suggests weaker generalization capability and a higher susceptibility to overfitting compared with the other evaluated architectures. For figure 6(c), the training loss showed a steady downward trend, while the validation loss also decreased progressively, especially after the early epochs. Although the validation loss remained higher than the training loss, both curves followed a generally consistent pattern without severe instability. This indicates that MobileNetV2 learned the data effectively and maintained strong generalization performance. The overall loss behavior confirms that this architecture achieved efficient convergence with a relatively low overfitting tendency. For figure 6(d), the training loss decreased sharply from the first epoch onward, reflecting rapid optimization and strong fitting capability. The validation loss also showed a declining trend and remained relatively stable across epochs, despite minor fluctuations near the end of training. The substantial reduction in training loss, combined with the comparatively stable validation loss, suggests that Inception V3 was highly effective in feature learning. However, the wider gap between the two curves in some epochs indicates a slight generalization gap, although not severe enough to suggest strong overfitting. Overall, the loss curves indicate that ResNet-18, MobileNetV2, and Inception V3 demonstrated favorable convergence characteristics and relatively good generalization, whereas VGG-16 showed the least stable validation behavior. These findings are consistent with the accuracy based evaluation and further support the conclusion that lighter and more optimized architectures are better suited for rice leaf disease classification in this study.

Quantitative Model Evaluation

Quantitative evaluation was conducted based on four primary performance metrics: Training Accuracy, Validation Accuracy, Training Loss, and Validation Loss. These metrics were used to assess how well each model learned during training and how effectively it generalized to unseen data. The detailed evaluation results are visualized in Figure 7, which presents the confusion matrix for each model.

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358



(a) ResNet 18 (b) Vgg16

(c) MobileNet V2 (d) Inception V3

Source: (Research Results, 2025)

Figure 7. Confusion Matrices of the Four Evaluated Models

Figure 7 shows the confusion matrices for classifying Bacterial Blight and Blast using four CNN models. Figure 7(a) achieved the highest accuracy, with 113 correct predictions for Bacterial Blight and 93 for Blast, and only 8 false negatives—indicating excellent sensitivity and no false positives. Figure 7(b) had lower accuracy, misclassifying 15 Blast cases and 2 Bacterial Blight cases, reflecting weaker generalization. Figure 7(c) performed better, with 10 errors in Blast and 2 in Bacterial Blight, though still showing class imbalance. Figure 7(d) offered the most balanced results after ResNet-18, with only 3 misclassifications in Blast and 2 in Bacterial Blight, showing high consistency. Overall, ResNet-18 proved to be the most effective model, with Inception V3 as a strong alternative. Full performance comparisons are summarized in Table 2.

Table 2. Training Results Comparison of CNN Architectures

Model	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss	Stability	Overfitting Tendency
ResNet-18	98.60%	97.20%	0.0445	0.0675	High	Low
VGG-16	97.46%	86.29%	0.0722	0.7561	Low	High
MobileNetV2	95.97%	97.46%	0.0675	0.0445	High	Low
Inception V3	100.00%	95.97%	0.0642	0.1537	High	Low

Source: (Research Results, 2025)

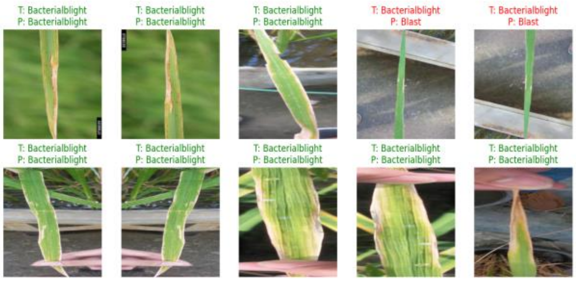
In Table 2, the qualitative indicators Stability and Overfitting Tendency were assigned based on the observed relationship between training and validation performance across epochs. Stability was evaluated from the consistency and smoothness of the training process, as reflected in the accuracy and loss curves, with models showing smaller fluctuations and more regular convergence categorized as having high stability. Overfitting tendency was assessed from the gap between training and validation accuracy, as well as the difference between training and validation loss. A smaller gap was interpreted as a lower tendency toward overfitting, whereas a larger discrepancy indicated a higher overfitting tendency.

Discussion

Beyond numerical metrics, a visual analysis was conducted on the model's predictions using validation images. The CNN models accurately classified rice leaf diseases based on distinct visual features, such as large brown spots, yellowing from green, and elongated lesions typical of blast infections. These results show the models' effectiveness in recognizing key visual symptoms,

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

confirming their potential for reliable image-based disease classification in rice crops.

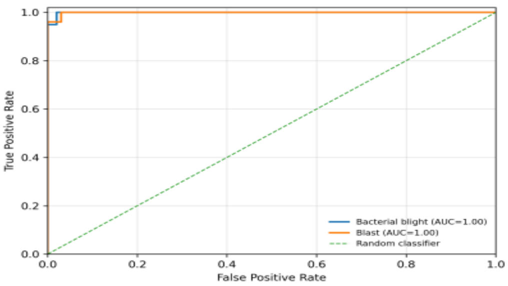


Source: (Research Results, 2025)
Figure 8. Sample Prediction Results

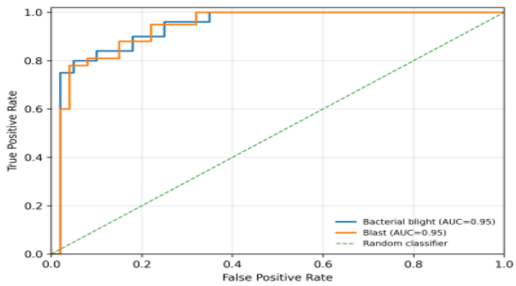
Figure 8 presents representative sample classification results for rice leaf disease prediction. Each image is annotated with the labels “T” (true class) and “P” (predicted class), where green text indicates correct predictions and red text indicates misclassifications. Most samples were classified correctly, particularly for Bacterial Blight, although several errors were still observed in cases where diseased leaves were predicted as Blast. These results support the overall effectiveness of the model in distinguishing disease patterns from leaf images.

However, these findings should be interpreted with caution. The prediction results may still be influenced by dataset-specific bias, including similarities in background, image composition, symptom appearance, and class representation within the available dataset. As a result, strong performance on the test set does not necessarily guarantee the same level of accuracy under broader field conditions involving different devices, lighting settings, leaf orientations, or disease severity levels. In addition, although ResNet-18 showed stable performance, the model remains limited in interpretability, as its decision-making process is not directly transparent and may be difficult to explain to end users such as farmers or agricultural practitioners. From a deployment perspective, hardware capability is also an important consideration, since practical implementation in mobile or low-resource environments requires not only high classification accuracy but also efficient memory use, inference speed, and computational feasibility.

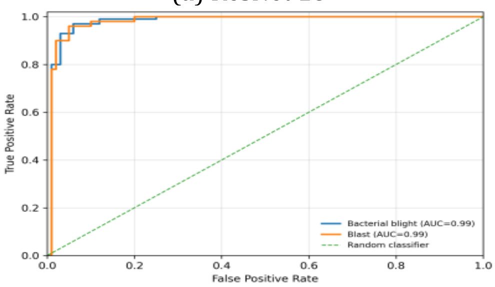
Furthermore, the robustness claim in this study is based on the available dataset and limited manual testing on dozens of samples. Therefore, while the model demonstrated promising performance under image noise and lighting variation, further validation on larger and more diverse field datasets is still needed before concluding that the approach is fully robust for real-world agricultural deployment. The AUC performance for each model is presented in Figure 8, which further illustrates the comparative discrimination capability of the four evaluated architectures.



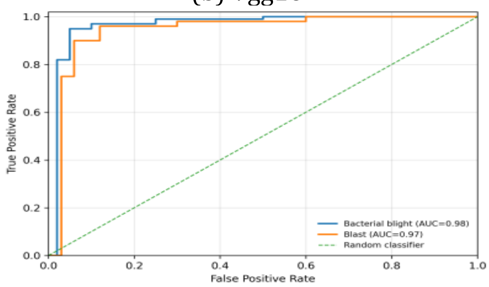
(a) ResNet 18



(b) Vgg16



(c) MobileNet V2



(d) Inception V3

Source: (Research Results, 2025)
Figure 8. ROC AUC Curves of the Four Evaluated Models

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

Figure 8 displays the ROC (Receiver Operating Characteristic) curves and AUC (Area Under the Curve) values for four CNN models in classifying rice leaf diseases. Figure 8(a) achieved perfect classification with AUC = 1.00 for both classes (Bacterial blight and Blast), indicating flawless detection. Figure 8(b) recorded a lower AUC of 0.95, suggesting more frequent misclassifications. Figure 8(c) performed nearly perfectly, with AUC = 0.99, closely matching ResNet-18. Figure 8(d) also showed high performance, with AUC = 0.98. These results confirm that ResNet-18 is the most accurate and precise model in this study, with MobileNetV2 and Inception V3 as strong alternatives. The final comparative results across all models are summarized in Table 3.

Table 3. Final Evaluation Results

Model	Accuracy	Precision	Recall	F1-Score	AUC
ResNet-18	0.9694	1.0000	0.9545	0.9618	1.0000
VGG-16	0.8333	0.9623	0.7727	0.8567	0.9500
MobileNetV2	0.8888	0.9655	0.8484	0.9027	0.9900
Inception V3	0.9513	0.9692	0.9208	0.9585	0.9800

Source: (Research Results, 2025)

Table 3 presents the final evaluation metrics for four CNN models tested on rice leaf disease classification. The ResNet-18 in this study, consistently outperformed the others across all metrics. ResNet-18 achieved the highest accuracy (96.94%), perfect precision (1.0000), and AUC (1.0000), indicating that it classified all predicted positive cases correctly and distinguished between disease classes flawlessly. Its recall (95.45%) and F1-score (96.18%) also reflect excellent balance between sensitivity and precision. Inception V3 followed closely with strong results: accuracy (95.13%), precision (96.92%), and recall (92.08%), making it a reliable alternative, though slightly behind in sensitivity and overall F1 performance compared to ResNet-18. MobileNetV2 demonstrated solid generalization with accuracy (88.88%) and AUC (0.9900). Its relatively high precision (96.55%) but lower recall (84.84%) suggests it was more conservative—effective at avoiding false positives but missed more true positives than ResNet-18 and Inception V3. VGG-16 performed the weakest, with the lowest accuracy (83.33%), recall (77.27%), and the highest disparity between precision and recall. This imbalance indicates a tendency toward overfitting and poor generalization, consistent with its earlier training and validation loss trends. Overall, ResNet-18 emerges as the most accurate, balanced, and robust model, making it well-suited for real-world implementation in automated rice disease detection systems.

CONCLUSION

This study conducted a comparative evaluation of four deep learning architectures ResNet-18, VGG-16, Inception V3, and MobileNetV2 to support early detection of rice leaf diseases through digital image classification. Using a dataset comprising three classes (healthy, blast-infected, and brown spot), each model was trained and validated within a structured experimental framework. Among the evaluated models, ResNet-18 consistently delivered the best performance, achieving an accuracy of 96.94%, perfect precision of 100%, recall of 95.45%, an F1-score of 96.18%, and an AUC of 1.0000. These results highlight its robustness, training stability, and superior generalization compared to the other architectures tested.

Based on these findings, ResNet-18 is recommended as the optimal lightweight CNN architecture for automated rice leaf disease classification. Its ability to accurately detect disease symptoms from images makes it a strong candidate for integration into intelligent agricultural monitoring systems. Ultimately, this research contributes to the advancement of AI-powered precision agriculture technologies that aim to enhance crop health management and strengthen local food security through timely and reliable disease detection.

REFERENCES

[1] M. Mavaddat, “Plant Diseases Classification Using Pre-trained and Transfer Learning Models: A Study on Rice Leaves,” *Aut Journal of Electrical Engineering*, vol. 57, no. 2, pp. 263–282, 2025, doi: 10.22060/eej.2024.23015.5580.
[2] A. R. Khan, “Automatic Diagnosis of Rice Leaves Diseases Using Hybrid Deep Learning Model,” *Journal of Advances in Information Technology*, vol. 14, no. 3, pp. 418–425, 2023, doi: 10.12720/jait.14.3.418-425.
[3] S. M. Rallapalli, “A contemporary approach for disease identification in rice leaf,”

International Journal of Systems Assurance Engineering and Management, 2021, doi: 10.1007/s13198-021-01159-y.

[4] A. Suting, "Rice Leaf Diseases Classification Using Deep Learning Algorithms for Smartphone Application Development: An Empirical Study," *International Journal of Image and Graphics*, 2024, doi: 10.1142/S0219467826500300.

[5] P. T. Pavithra, "A Hybrid ViT-CNN Model Premeditated for Rice Leaf Disease Identification," *International Journal of Computational Methods and Experimental Measurements*, vol. 12, no. 1, pp. 35–43, 2024, doi: 10.18280/ijcmem.120104.

[6] S. Sangeetha, "IoT Based Wireless Communication System for Smart Irrigation and Rice Leaf Disease Prediction Using ResNeXt-50," *International Journal on Artificial Intelligence Tools*, vol. 33, no. 2, 2024, doi: 10.1142/S0218213024500040.

[7] Y. Liu, "Detection method of rice leaf disease based on DeepLabv3—Faster R—CNN," *Journal of Chinese Agricultural Mechanization*, vol. 46, no. 4, 2025, doi: 10.13733/j.jcam.issn.2095-5553.2025.04.016.

[8] R. K and M. S. K, "Breast Cancer Prediction by Leveraging Machine Learning and Deep learning Techniques with Different Imaging Modalities," in *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, 2022, pp. 1–6. doi: 10.1109/I2CT54291.2022.9824749.

[9] I. Ozcan, H. Aydin, and A. Cetinkaya, "Comparison of Classification Success Rates of Different Machine Learning Algorithms in the Diagnosis of Breast Cancer.," *Asian Pacific journal of cancer prevention : APJCP*, vol. 23, no. 10, pp. 3287–3297, Oct. 2022, doi: 10.31557/APJCP.2022.23.10.3287.

[10] A. P. Windarto, T. Herawan, and P. Alkhairi, "Early Detection of Breast Cancer Based on Patient Symptom Data Using Naive Bayes Algorithm on Genomic Data," in *Artificial Intelligence, Data Science and Applications*, Y. Farhaoui, A. Hussain, T. Saba, H. Taherdoost, and A. Verma, Eds., Cham: Springer Nature Switzerland, 2024, pp. 478–484.

[11] J. Rashid, S. Batool, J. Kim, M. W. Nisar, and A. Hussain, "An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction," *Front. Public Health*, vol. 10, no. March, pp. 1–20, 2022, doi: 10.3389/fpubh.2022.860396.

[12] S. Saeedbakhsh, M. Sattari, M. Mohammadi, J. Najafian, and F. Mohammadi, "Diagnosis of Coronary Artery Disease based on Machine Learning algorithms Support Vector Machine, Artificial Neural Network, and Random Forest.," *Advanced biomedical research*, vol. 12, p. 51, 2023, doi: 10.4103/abr.abr_383_21.

[13] A. P. Windarto, I. R. Rahadjeng, M. N. H. Siregar, and P. Alkhairi, "Deep Learning to Extract Animal Images With the U-Net Model on the Use of Pet Images," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 1, pp. 468–476, 2024.

[14] S. Defit, A. P. Windarto, and P. Alkhairi, "Comparative Analysis of Classification Methods in Sentiment Analysis: The Impact of Feature Selection and Ensemble Techniques Optimization," *Telematika*, vol. 17, no. 1, pp. 52–67, 2024.

[15] P. Alkhairi, W. Wanayumini, and B. H. Hayadi, "Analysis of the adaptive learning rate and momentum effects on prediction problems in increasing the training time of the backpropagation algorithm," *AIP Conference Proceedings*, vol. 3048, no. 1, p. 20049, 2024, doi: 10.1063/5.0203374.

[16] T. A. T. K. Azaharan, "Investigation of VGG-16, ResNet-50 and AlexNet Performance for Brain Tumor Detection," *International journal of online and biomedical engineering*, vol. 19, no. 8, pp. 97–109, 2023, doi: 10.3991/ijoe.v19i08.38619.

[17] W. Al-Khater, "Using 3D-VGG-16 and 3D-Resnet-18 deep learning models and FABEMD techniques in the detection of malware," *Alexandria Engineering Journal*, vol. 89, pp. 39–52, 2024, doi: 10.1016/j.aej.2023.12.061.

[18] L. I. Kesuma, "ELREI: Ensemble Learning of ResNet, EfficientNet, and Inception-v3 for Lung Disease Classification based on Chest X-Ray Image," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 5, pp. 149–161, 2023, doi: 10.22266/ijies2023.1031.14.

[19] F. Edderbali, "Mobilenet, inception ResNet and GoogleNet for epilepsy detection using spectrogram images," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 2, pp. 870–877, 2024, doi: 10.11591/ijeecs.v34.i2.pp870-877.

[20] Z. Kaya, "Classification of brain tumors using wavelet transform and Inception V3 convolutional neural network model," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 39, no. 3,

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7358

pp. 1945–1952, 2024, doi: 10.17341/gazimmfd.1221952.

[21] I. Z. Renu *et al.*, “A Comprehensive Analysis on Skin Cancer Classification Using Transfer Learning,” *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering, ICAEEE 2024*, no. June, pp. 1–6, 2024, doi: 10.1109/ICAEEE62219.2024.10561770.

[22] L. F. de J. Silva, O. A. C. Cortes, and J. O. B. Diniz, “A novel ensemble CNN model for COVID-19 classification in computerized tomography scans,” *Results in Control and Optimization*, vol. 11, no. September 2022, p. 100215, 2023, doi: 10.1016/j.rico.2023.100215.

[23] T. R. Mahesh, “Early Predictive Model for Detection of Plant Leaf Diseases Using MobileNetV2 Architecture,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 2, pp. 46–54, 2023.

[24] Abhishek, “Classification of Colorectal Cancer using ResNet and EfficientNet Models,” *Open Biomedical Engineering Journal*, vol. 18, 2024, doi: 10.2174/011874120728070324011107575 2.

[25] S. A. El-Feshawy, “IoT framework for brain tumor detection based on optimized modified ResNet 18 (OMRES),” *Journal of Supercomputing*, vol. 79, no. 1, pp. 1081–1110, 2023, doi: 10.1007/s11227-022-04678-y.

[26] M. Kia, “Innovative fusion of VGG16, MobileNet, EfficientNet, AlexNet, and ResNet50 for MRI-based brain tumor identification,” *Iran Journal of Computer Science*, vol. 8, no. 1, pp. 185–215, 2025, doi: 10.1007/s42044-024-00216-6.

[27] H. A. Shah, F. Saeed, S. Yun, J. H. Park, A. Paul, and J. M. Kang, “A Robust Approach for Brain Tumor Detection in Magnetic Resonance Images Using Finetuned EfficientNet,” *IEEE Access*, vol. 10, pp. 65426–65438, 2022, doi: 10.1109/ACCESS.2022.3184113.

[28] S. Kiziloluk and E. Sert, “COVID-CCD-Net: COVID-19 and colon cancer diagnosis system with optimized CNN hyperparameters using gradient-based optimizer,” *Medical and Biological Engineering and Computing*, vol. 60, no. 6, pp. 1595–1612, 2022, doi: 10.1007/s11517-022-02553-9.

[29] R. Vallabhaneni, “MobileNet based secured compliance through open web application security projects in cloud system,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 3, pp. 1661–1669, 2024, doi: 10.11591/ijeecs.v35.i3.pp1661-1669.

[30] C. Mai, “MobileNet-Based IoT Malware Detection with Opcode Features,” *Journal of Communications and Information Networks*, vol. 8, no. 3, pp. 221–230, 2023, doi: 10.23919/JCIN.2023.10272350.

[31] A. Wahid, “Smart Blended Learning Framework based on Artificial Intelligence using MobileNet Single Shot Detector and Centroid Tracking Algorithm,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 364–369, 2022, doi: 10.14569/IJACSA.2022.0130543.

[32] R. Chen, “Rapid Detection of Multi-QR Codes Based on Multistage Stepwise Discrimination and a Compressed MobileNet,” *IEEE Internet of Things Journal*, vol. 10, no. 18, pp. 15966–15979, 2023, doi: 10.1109/JIOT.2023.3268636.

[33] V. Rajasekhar, G. Arulselvi, and K. Suresh Babu, “Detection of diseases in rice leaf using convolutional neural network with transfer learning based on ResNeXt,” *International Journal of Electrical and Computer Engineering*, vol. 14, no. 2, pp. 1739–1749, 2024, doi: 10.11591/ijece.v14i2.pp1739-1749.

[34] M. S. H. Shovon, S. J. Mozumder, O. K. Pal, M. F. Mridha, N. Asai, and J. Shin, “PlantDet: A Robust Multi-Model Ensemble Method Based on Deep Learning For Plant Disease Detection,” *IEEE Access*, vol. 11, no. March, pp. 34846–34859, 2023, doi: 10.1109/ACCESS.2023.3264835.

[35] S. Aboukadri, A. Ouaddah, and A. Mezrioui, “Major Role of Artificial Intelligence, Machine Learning, and Deep Learning in Identity and Access Management Field: Challenges and State of the Art BT - Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022,” A. E. Hassanien, V. Snášel, M. Tang, T.-W. Sung, and K.-C. Chang, Eds., Cham: Springer International Publishing, 2023, pp. 50–64.

[36] B. H. Hayadi, I. G. I. Sudipa, and A. P. Windarto, “Model Peramalan Artificial Neural Network pada Peserta KB Aktif Jalur Pemerintahan menggunakan Artificial Neural Network Back-Propagation,” *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21,

no. 1, pp. 11–20, 2021, doi: 10.30812/matrik.v21i1.1273.

[37] P. Alkhairi and B. H. Hayadi, "Pemodelan Artificial Neural Network Peramalan Tingkat Kenaikan Jabatan Berdasarkan Kinerja Pegawai Menggunakan Backpropagation," *Smart EDU: Buletin Education*, vol. 1, no. 2, pp. 73–84, 2022.

[38] Z. Wang, "CNN- and GAN-based classification of malicious code families: A code visualization approach," *International Journal of Intelligent Systems*, vol. 37, no. 12, pp. 12472–12489, 2022, doi: 10.1002/int.23094.

[39] Y. Zhu, H. Sekiya, T. Okatani, M. Tai, and S. Morichika, "B-CNN: a deep learning method for accelerometer-based fatigue cracks monitoring system," *Journal of Civil Structural Health Monitoring*, vol. 13, no. 4–5, pp. 947–959, Jun. 2023, doi: 10.1007/s13349-023-00690-9.

[40] T. H. Chun, "Efficacy of the Image Augmentation Method using CNN Transfer Learning in Identification of Timber Defect," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 107–114, 2022, doi: 10.14569/IJACSA.2022.0130514.

[41] P. Pankaj, "A New design of occlusion Invariant Face Recognition Using optimal Pattern extraction and CNN with GRU-Based Architecture," *International Journal of Information Security and Privacy*, vol. 16, no. 1, 2022, doi: 10.4018/IJISP.305222.

[42] D. Dablain, B. Krawczyk, and N. V. Chawla, "DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 9, pp. 6390–6404, 2023, doi: 10.1109/TNNLS.2021.3136503.

[43] K. Punithavathi, "Enhanced Breast Cancer Detection and Classification Using CNN-EfficientNet B3: A Deep Learning Approach with Segmentation and Feature Extraction from MIAS Dataset," *SSRG International Journal of Electronics and Communication Engineering*, vol. 12, no. 1, pp. 257–270, 2025, doi: 10.14445/23488549/IJECE-V12I1P120.

[44] X. Wang, "OsACA9, an Autoinhibited Ca2+-ATPase, Synergically Regulates Disease Resistance and Leaf Senescence in Rice," *International Journal of Molecular Sciences*, vol. 25, no. 3, 2024, doi: 10.3390/ijms25031874.

[45] D. Chaudhari, "Remora-CNN: A Novel and Effective Method for Rice Leaf Disease Detection and Classification," *Journal of Phytopathology*, vol. 172, no. 5, 2024, doi: 10.1111/jph.13411.

[46] N. D. Luu, "Application of deep learning for rice leaf disease detection in the Mekong Delta," *Ctu Journal of Innovation and Sustainable Development*, vol. 16, pp. 1–7, 2024, doi: 10.22144/ctujoisd.2024.316.