Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

Rice Leaf Disease Classification Using ResNet-50: A Comparative Study of Adam, SGD, and RMSProp

Bebin Paula[1*], **Muhammad Rizky Pribadi**[2]
[1,2]Universitas Multi Data Palembang, Indonesia
[1]bebinpaula11@mhs.mdp.ac.id, [2]rizky@mdp.ac.id

open access

***Corresponding Author**

Article History:
Submitted: 20-12-2025
Accepted: 02-01-2026
Published: 19-01-2026

Keywords:
ADAM; Resnet-50; Rice; RMSProp; SGD.

ABSTRACT

Rice plant diseases significantly affect crop productivity and require accurate and timely identification to support effective management. This study proposes a rice leaf disease classification approach using the ResNet-50 convolutional neural network and compares the performance of three optimization algorithms, namely ADAM, Stochastic Gradient Descent (SGD), and RMSProp. The model was trained and evaluated on a rice leaf image dataset consisting of four classes BrownSpot, Healthy, Hispa, and LeafBlast. The dataset contains visual variations in color, texture, and disease patterns that influence classification performance. Performance was assessed using training accuracy, loss, precision, recall, F1-score, and confusion matrix analysis. These evaluation metrics provide a comprehensive measurement of model effectiveness and class-wise prediction behavior. Experimental results show that the ADAM optimizer achieved the best performance with a training accuracy of 75.84%, followed by RMSProp at 74.60%, while SGD obtained the lowest accuracy of 71.34%. The differences in performance highlight the impact of optimization strategies on deep neural network training stability. Class-wise evaluation indicates that the model performed well in detecting BrownSpot and Healthy classes, but showed lower performance on the Hispa class across all optimizers. This limitation is influenced by the visual similarity of Hispa symptoms to other classes. These findings demonstrate that adaptive learning rate–based optimizers provide faster convergence and better classification performance for deep learning–based rice disease detection. The results support the use of optimized convolutional neural networks for image-based agricultural applications.

INTRODUCTION

Agriculture constitutes a strategic sector in national development, particularly regarding the management of food commodity yields. The agricultural sector serves as a crucial pillar of national food security; thus, the management and utilization of its products must be well-planned and optimized to be accessible to the entire population (Batubara & Mustika Pane, 2023). Rice (Oryza sativa L.) is one of the primary food commodities in Indonesia. As rice is the staple food for more than half of the Indonesian population, demand continues to rise in tandem with population growth. However, the Central Statistics Agency (Badan Pusat Statistik) reported that national rice production in 2024 is estimated to decline by approximately 2.45% to 52.66 million tons of Milled Dry Grain (Gabah Kering Giling) compared to the previous year (Badan Pusat Statistik, 2024). This production decline is partly attributed to a reduction in harvested area and various other factors, including plant disease attacks. This condition presents new challenges for food security and underscores the importance of research to optimize rice production.

Rice plants are susceptible to various leaf diseases caused by bacteria, fungi, and pests, which can significantly reduce harvest quality and quantity. These rice diseases can be observed or recognized through various symptoms such as spots and discoloration of the leaves (Pirnando et al., 2025). Some commonly encountered leaf diseases include Brown Spot, caused by the fungus Bipolaris oryzae (previously Helminthosporium oryzae); Leaf Blast, caused by Magnaporthe oryzae (anamorph Pyricularia oryzae); and Rice Hispa, caused by the insect Dicladispa armigera (Pramono & Wibowo, 2024). Symptoms of these attacks generally manifest as spots on the leaf surface, leaf discoloration, and distinctive visual damage patterns. These disease attacks not only reduce the healthy leaf area available for photosynthesis but can also drastically lower crop yields if not addressed promptly.

In the field, farmers often encounter difficulties in accurately and rapidly assessing the severity of damage to rice plants. Manual identification of plant conditions by farmers frequently leads to delays in the handling process, tends to be subjective, and is often inaccurate, which ultimately contributes to a decline in production yields. Limited access to technology and a shortage of agricultural extension workers, especially in rural areas, exacerbate this problem. Many farmers still rely on intuition and limited experience to determine plant conditions, risking diagnostic errors and ineffective treatments (Sutabri & Darma, 2024). Therefore, an automated and reliable assistive method is needed to detect and classify rice plant diseases, allowing for earlier and more precise disease control.

The development of deep learning technology, specifically Convolutional Neural Networks (CNN), offers a potential solution to the aforementioned problems. CNN is a digital image processing method capable of automatically extracting and recognizing visual patterns in images. Various studies have proven the effectiveness of CNNs in classifying plant disease images with very high accuracy. One such study conducted by Istiqomah and Murinto constructed a CNN model for rice leaf disease classification using 480 leaf images as test data, achieving an accuracy of 97.75% (Istiqomah & Murinto, 2024). These results indicate that CNNs can detect rice diseases at an early stage very effectively, enabling farmers to take preventive actions more quickly. A prominent advantage of CNNs is their ability to extract detailed image features without the need for manual feature intervention, allowing the model to recognize disease types and severity automatically with high accuracy. Previous research by Sahputra also demonstrated that state-of-the-art CNN architectures are capable of classifying rice plant diseases with high accuracy, where VGG-16 reached 98% accuracy and MobileNet achieved approximately 95% (Sahputra et al., 2025). Various optimizations involving hyperparameter tuning and training algorithms (optimizers) such as Adam, SGD, or RMSprop have also been proven to enhance CNN model performance in plant image classification.

One popular CNN architecture is the Residual Network (ResNet). ResNet50 is a ResNet variant consisting of 50 deep layers designed to overcome the vanishing gradient problem in very deep networks (Fauzan Novriandy et al., 2024). The uniqueness of ResNet lies in its residual connections or skip connections, which allow gradients to flow more directly, enabling the network to be trained deeper without performance degradation (Elizabeth & Tinaliah, 2023). The ResNet50 architecture consists of 5 main stages, each containing convolution blocks and identity blocks, with a total of over 23 million parameters trained on the ImageNet dataset (Berliani et al., 2023). Many studies have applied ResNet50 to plant disease image classification due to its ability to extract complex visual features and maintain high accuracy. For instance, research conducted by Ujawe on rice plant classification using ResNet50 proved superior in agricultural image classification tasks with a fairly good accuracy of 92.20%. While this study produced good accuracy, it was limited to a small number of disease classes, meaning the model was less tested on more diverse real-world conditions (Ujawe et al., 2024). Furthermore, research by Santosa on rice plant disease classification using CNN and several optimizers such as Adam, SGD, and RMSProp showed excellent results with accuracies of 99.84% for Adam, 99.56% for SGD, and 99.66% for RMSProp. The highest accuracy rate was achieved using the Adam Optimizer at 99.84% (Santosa et al., 2023). Based on previous research, with its substantial number of layers and parameters, ResNet50 demonstrates high accuracy performance and is expected to be capable of recognizing distinctive characteristics of rice leaf diseases more deeply compared to shallower architectures.

Based on the description above, this research addresses the problem of classifying rice plant disease types using deep learning techniques. Previous studies have shown that CNNs, particularly the ResNet50 architecture, are capable of producing high accuracy in disease identification. Therefore, this study aims to apply the Convolutional Neural Network method with the ResNet50 architecture to classify disease types in rice plants, as well as to compare the influence of using three different optimizers (Adam, SGD, and RMSprop) on model performance. It is hoped that the results of this research will provide a more efficient and accurate solution for rice disease identification, support modern agricultural practices, and assist farmers in making timely decisions regarding pest and disease management.

LITERATURE REVIEW

The increasing concern over crop productivity due to plant diseases has driven numerous studies that explore image-based classification techniques using Convolutional Neural Networks (CNN). Several recent studies have evaluated different CNN architectures, optimizers, and datasets to enhance the performance of plant disease classification systems. This section critically reviews the existing literature, emphasizing the similarities, differences, and research gaps, particularly focusing on CNN-based models, optimizer comparisons, and applications in rice disease classification.

Santosa proposed a CNN-based classification system to detect bacterial leaf blight, Brown Spot, and leaf smut in rice plants using 16,000 images. The study achieved an outstanding accuracy of 99.66% with four hidden layers and the Adam optimizer, demonstrating the optimizer's effectiveness in reaching high precision, recall, and F1-score (Santosa et al., 2023). However, the study did not explore other optimizer variants or CNN architectures, leaving a gap for comparative analysis.

Similarly, Nurdin evaluated the use of the Inception-V3 architecture with Adam, SGD, and RMSprop optimizers to classify tomato leaf diseases. The Adam optimizer showed the best accuracy (93.8%) and model stability. While this study contributed valuable insights into optimizer comparison, it focused solely on tomato plants, indicating a need to validate such findings on other crops such as rice (Nurdin et al., 2024).

In a more generalized context, Labhsetwar) conducted an optimizer performance analysis across RMSprop, Adam, and AMSGrad for classifying plant diseases from satellite and mobile imagery. Adam again outperformed the rest with 98% validation accuracy, reinforcing its robustness across domains (Labhsetwar et al., 2023). Yet, the reliance on aerial imagery differs significantly from leaf-level image classification tasks typically conducted in rice disease studies.

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

Saleem expanded the comparison to include 18 CNN models and various optimizers. Their findings highlighted Xception architecture trained with Adam as the most effective, achieving a 99.81% validation accuracy and F1-score of 0.9978. This large-scale comparison confirmed Adam's superiority but lacked a specific focus on rice or leaf-based classification models (Saleem et al., 2020).

Julianto explored CNN MobileNet-V2 with hyperparameter tuning for rice disease classification. Using RMSprop with optimal hyperparameter settings, they achieved 97.56% testing accuracy. This study emphasized the importance of optimizer and parameter selection but did not consider the deeper ResNet-based architectures (Julianto et al., 2022).

Pujiati and Rochmawati used CNNs to classify herbal plant images, achieving up to 94% training accuracy. Although not disease-specific, their work underlined CNN's versatility in plant classification. However, the moderate accuracy during testing (84%) suggests limitations in generalizability, possibly due to insufficient dataset diversity (Pujiati and Rochmawati, 2022).

Ramadhani and Tarigan compared EfficientNetV2-S and ResNet50 architectures for mango leaf disease classification. Both achieved up to 100% accuracy in some train-test splits, further supporting ResNet50's reliability. However, this result, while impressive, raises concerns about overfitting and generalization on more complex, real-world datasets (Ramadhani and Tarigan, 2025).

Fauzan Novriandy used ResNet50 with Adam and SGD to classify oral cancer from histopathology images. While not related to agriculture, their findings (up to 94.66% accuracy) reaffirm ResNet50's adaptability and performance with various optimizers across domains (Fauzan Novriandy et al., 2024).

Kresnawan implemented Faster R-CNN with ResNet50 for shallot disease classification. However, the accuracy (65%) was relatively low, which the authors attributed to limited dataset size and variability. This further highlights the importance of dataset quality and class diversity in performance evaluation (Kresnawan et al., 2024).

Lastly, Ujawe directly focused on rice diseases using ResNet50 and a modified TLResNet152V2. Their approach achieved 93.20–95.03% accuracy for classifying Brown Spot, Healthy, Leaf Blast, and Neck Blast. This model outperformed others like GoogLeNet and DenseNet201, confirming ResNet's potential in rice disease detection. Nonetheless, their study did not compare optimizer effects in-depth, which limits insights on training stability and convergence behavior (Ujawe et al., 2024).

The reviewed literature demonstrates that CNNs, especially deeper architectures like ResNet50, are highly effective for plant disease classification. The Adam optimizer consistently yields superior results across various architectures and datasets. However, limited studies have comprehensively compared multiple optimizers (Adam, SGD, RMSprop) using the same architecture, particularly ResNet50, on rice leaf disease datasets. Moreover, most works focus on either architecture performance or optimizer effects in isolation.

This study addresses these gaps by systematically comparing the performance of ResNet50 architecture using Adam, SGD, and RMSprop optimizers on a rice leaf disease classification task. The findings are expected to offer insights into the most effective optimizer for this domain, contributing to the development of efficient and reliable agricultural decision-support systems.

METHOD

This study followed a series of clear and well-defined stages, as shown in the flowchart in Figure 1. The process began with a literature review to understand the use of CNN Resnet50 and optimizers Adam, SGD, and RMSProp in the classification of rice disease images.

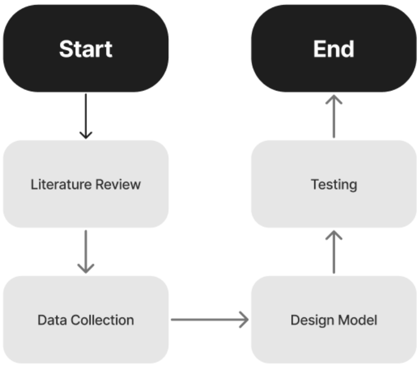

Fig. 1 Research Methodology Flowchart

This study utilized the public dataset, obtained from Kaggle named "Rice Leafs". As detailed in Table 1, the contents of the dataset are categorized into four categories, namely Brown Spot, Healthy, Hispa, and Leaf Blast. The dataset will be divided into two parts: training data and test data, with a ratio of 80% for training data and 20% for test data. In this study, the ResNet-50 model was trained from scratch using random weight initialization. We did not use pre-trained weights from ImageNet or other forms of transfer learning. All weights were updated during training on the "Rice Leafs" dataset (3355 images) with the experimental settings as reported (epoch = 50, batch size = 16).

Table 1. Details of the Dataset Used

Dataset	Dataset Image	Training (80%)	Testing (20%)	Amount of Data
Brown Spot		418	105	523
Healthy		1191	297	1488
Hispa		452	113	565
Leaf Blast		623	156	779
Total				**3355**

Before the data splitting process, the dataset will enter the Preprocessing stage, which is the process of adjusting the image size to 224x224 pixels and normalizing the pixels to the range [0,1] with the aim of making the network faster to adapt during the training process. Then, the dataset is augmented. Augmentation is a technique to enlarge the dataset size and increase data diversity by modifying existing images. The augmentation techniques used include several methods, namely rotation, which allows images to be rotated up to 40 degrees (rotation_range = 40), horizontal mirroring (horizontal_flip) to flip the image horizontally, shifting on the width and height axes up to 20% of the image dimensions (width_shift_range = 0.2 and height_shift_range = 0.2), and zooming with a variation of up to 20% (zoom_range = 0.2). In addition, a shear or angle distortion of 20% (shear_range = 0.2) is also used. Augmentation is performed with the aim of increasing the variety of training data.

After the data splitting process, the next step is model training or modeling, which is carried out using training data to train the model's performance. In the model training stage, the ResNet50 architecture is used with the help of several optimizers (Adam, SGD, and RMSprop) to improve network weights and avoid overfitting. In the ResNet50 architecture, the calculation process includes convolution, batch normalization, ReLU activation, and residual connection stages. Each residual block performs calculations, where the convolution function is the initial input that is forwarded directly to the output through a skip connection. This mechanism keeps the gradient flow stable, thereby reducing the risk of vanishing gradients. The output from the last layer is then processed through Global Average Pooling and Dense layers to produce predictions of rice disease classes. Once the model has been obtained for use in testing, the model will then be tested to obtain the classification results for rice plant quality in the form of calculations of the number of rice leaves using testing data. Once the model has been obtained for use in testing, the next step is to test the model to obtain the results of rice plant quality classification in the form of calculations of the number of rice leaves using testing data.

After all stages are complete, the model will enter the testing stage. The test results are calculated to determine the success rate of the method used with the Confusion Matrix through four main components, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) in calculating the Accuracy, Precision, Recall, and F1-score values.

Accuracy is a measure of how many predictions are correct out of all predictions made by the model. Accuracy is the most commonly used metric because it is easy to understand. Simply put, the accuracy formula can be seen in Equation (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{1}$$

Precision is a more specific metric than accuracy for measuring one label as a separate entity. Precision measures the percentage of correct predictions out of all predictions on a particular label (in this case, the positive label). It can be said that precision is the ratio between true positive predictions (TP) and total positive predictions (both correct and incorrect). Simply put, the precision formula can be seen in Equation (2).

$$Precision = \frac{TP}{TP+FP} \times 100\% \tag{2}$$

Recall is a metric for measuring how many positive cases are actually positive out of all the data. This metric is important when we want to minimize false negatives (FN), such as in disease classification where we do not want any sick patients to be classified as healthy. Technically, in a confusion matrix, recall is the ratio of correct positive predictions to the total number of data that are actually positive. Simply put, the recall formula can be seen in Equation (3).

$$Recall = \frac{TP}{TP+FN} \times 100\% \tag{3}$$

F1-Score is a metric that combines precision and recall. F1-score is very useful when we encounter class imbalance in a dataset. The F1-Score value is the harmonic mean between precision and recall, providing a balance between the two metrics. The F1-score provides a balance between precision and recall, which is useful when both are equally important. Simply put, the F1-Score formula can be seen in Equation (4).

$$F1 - Score = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right) \tag{4}$$

RESULT

This section is intended to identify obstacles in the process of classifying rice disease and to analyze the accuracy of the CNN model using Resnet-50 architecture with ADAM, SGD, and RMSProp optimizers. The testing was conducted with epochs 50, batch sizes 16 and evaluated accuracy, precision, recall, and f1-score for each architecture.

Before entering the model training process, model design is essential to ensure the model performs effectively. The design process begins with obtaining a dataset from Kaggle consisting of rice plant leaf images categorized into four classes, Brown Spot, Healthy, Hispa, and Leaf Blast. The dataset then undergoes a preprocessing stage, which includes resizing the images to 224×224 pixels and normalizing pixel values to the [0,1] range to enable faster adaptation during training. Next, data augmentation is applied using several techniques, including rotation of up to 40 degrees (rotation_range = 40), horizontal flipping (horizontal_flip), width and height shifts of up to 20% of the image dimensions (width_shift_range = 0.2 and height_shift_range = 0.2), and zoom variations of up to 20% (zoom_range = 0.2). Additionally, shear distortion of up to 20% (shear_range = 0.2) is used. These augmentation techniques are implemented to increase the variability of the training data.

The next step is dataset splitting, where the dataset is divided into training and testing sets with a proportion of 80% for training and 20% for testing. Subsequently, the model training or modeling stage is conducted using the training data to improve the model's performance. In this stage, the ResNet50 architecture is employed along with several optimizers (Adam, SGD, and RMSprop) to update the network weights and reduce the risk of overfitting.

The results of the tests that have been conducted are presented in Table 2 for Resnet-50 with ADAM optimizer, Table 3 for Resnet-50 with SGD optimizer and Table 4 for Resnet-50 with RMSProp optimizer.

Table 2. Result of Testing Resnet-50 with ADAM Optimizer

Number	Epoch	Batch Size	Train Accuracy	Loss
1	10	16	64.33%	0.8677
2	20	16	67.52%	0.7960
3	30	16	72.17%	0.7095
4	40	16	73.62%	0.6472
5	**50**	**16**	**75.84%**	**0.5996**

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

Table 2 shows the results of ResNet-50 training using the ADAM optimizer with a batch size of 16, where training accuracy gradually increases with the addition of epochs, from 64.33% at epoch 10 to a maximum value of 75.84% at epoch 50. The consistent decrease in loss value from 0.8677 to 0.5996 also indicates that the model is converging and learning well. This improvement shows that ADAM is able to optimize model parameters effectively, with the best performance achieved at epoch 50 as the highest accuracy point in this experiment.

To better understand the model's performance beyond overall accuracy, it is necessary to analyze how accurately it identifies each type of rice leaf disease. Although Table 2 presents general patterns in accuracy, F1-score, and stability across different settings, it does not reveal the specific misclassifications that occur for each class. For this reason, a classification report was created to offer a more detailed view of the model's correct and incorrect predictions, with the best results for ResNet-50 using the ADAM optimizer displayed in Table 3.

Table 3. Evaluation Result ADAM Optimizer

	Precision	**Recall**	**F1-score**	**Support**
BrownSpot	0.92	0.79	0.85	228
Healty	0.68	0.94	0.79	588
Hispa	0.70	0.33	0.45	222
LeafBlast	0.82	0.60	0.69	304
Accuracy			**0.74**	**1342**
Macro avg	**0.78**	**0.67**	**0.70**	**1342**
Weighted avg	**0.75**	**0.74**	**0.72**	**1342**

The evaluation results show that the performance of the ResNet-50 model using the ADAM optimizer varies across classes. The BrownSpot class achieved the best performance with a precision of 0.92, recall of 0.79, and an F1-score of 0.85, while the Healthy class obtained the highest recall at 0.94 but a lower precision of 0.68. The Hispa class demonstrated the weakest performance with a recall of only 0.33 and an F1-score of 0.45, indicating that many samples were not correctly detected. For the LeafBlast class, the model reached a precision of 0.82 and an F1-score of 0.69. Overall, the model achieved an accuracy of 0.74, with a macro-average F1-score of 0.70 and a weighted-average F1-score of 0.72, reflecting a reasonably good classification capability but with room for improvement in certain classes, particularly Hispa.

Based on these evaluation results, it can be concluded that the ADAM optimizer provides reasonably good performance, although certain classes still exhibit weaknesses. To obtain a more comprehensive understanding of the model's effectiveness, additional testing was conducted using other optimizers. Therefore, the next discussion presents the results of testing ResNet-50 using the SGD optimizer.

Table 4. Result of Testing Resnet-50 with SGD Optimizer

Number	**Epoch**	**Batch Size**	**Train Accuracy**	**Loss**
1	10	16	61.59%	0.9094
2	20	16	66.93%	0.8130
3	30	16	68.53%	0.7704
4	40	16	70.01%	0.7480
5	**50**	**16**	**71.34%**	**0.7103**

Table 4 presents the training results of ResNet-50 using the SGD optimizer with a batch size of 16, where the training accuracy gradually increased from 61.59% at epoch 10 to its highest value of 71.34% at epoch 50. Although the accuracy showed an upward trend, its performance remained lower than that achieved with the ADAM optimizer. The loss value also decreased from 0.9094 to 0.7103, indicating a stable but slower learning process compared to other optimizers. Overall, these results demonstrate that SGD is capable of performing optimization effectively, but its efficiency is still limited, with the highest accuracy reached at epoch 50.

Table 5. Evaluation Result SGD Optimizer

	Precision	**Recall**	**F1-score**	**Support**
BrownSpot	0.85	0.75	0.80	228
Healty	0.66	0.94	0.78	588
Hispa	0.63	0.33	0.44	222
LeafBlast	0.83	0.54	0.65	304
Accuracy			**0.71**	**1342**
Macro avg	**0.74**	**0.64**	**0.66**	**1342**
Weighted avg	**0.73**	**0.71**	**0.70**	**1342**

From the details that show in Table 5, the evaluation results using the SGD optimizer show that the model's performance varies across classes. BrownSpot achieved the best outcomes with a precision of 0.85, recall of 0.75, and an F1-score of 0.80. The Healthy class obtained the highest recall at 0.94, although its precision remained at 0.66. The weakest performance occurred in the Hispa class, which recorded a recall of 0.33 and an F1-score of 0.44, indicating the model's difficulty in identifying this class. Meanwhile, the LeafBlast class reached a precision of 0.83 and an F1-score of 0.65. Overall, the model attained an accuracy of 0.71, with a macro-average F1-score of 0.66 and a weighted-average F1-score of 0.70, reflecting reasonably good classification performance but still requiring improvement in certain classes, particularly Hispa.

Based on these evaluation results, it is evident that the SGD optimizer delivers relatively stable performance, although certain classes still present challenges. To gain a more comprehensive understanding of the effectiveness of each optimizer, the analysis is continued by testing the model using the RMSProp optimizer. Therefore, the following section presents the results of testing ResNet-50 with RMSProp.

Table 6. Result of Testing Resnet-50 with RMSProp

Number	Epoch	Batch Size	Train Accuracy	Loss
1	10	16	63.81%	0.8791
2	20	16	70.11%	0.7611
3	30	16	71.70%	0.7116
4	40	16	73.50%	0.6651
5	**50**	**16**	**74.60%**	**0.6352**

In Table 6 presents the training results of ResNet-50 using the RMSProp optimizer with a batch size of 16, where training accuracy consistently increased from 63.81% at epoch 10 to a maximum of 74.60% at epoch 50. The loss decreased as epochs progressed (from 0.8791 to 0.6352), indicating a stable learning process and parameter convergence. Overall, RMSProp demonstrated effective optimization in this experiment, with the best performance achieved at epoch 50.

Table 7. Evaluation Result RMSProp Optimizer

	Precision	Recall	F1-score	Support
BrownSpot	0.87	0.82	0.84	228
Healty	0.71	0.88	0.79	588
Hispa	0.59	0.43	0.49	222
LeafBlast	0.80	0.63	0.70	304
Accuracy			**0.74**	**1342**
Macro avg	**0.74**	**0.69**	**0.71**	**1342**
Weighted avg	**0.74**	**0.74**	**0.73**	**1342**

The evaluation results using the RMSProp optimizer indicate that the performance of the ResNet-50 model varies across classes. The BrownSpot class achieved the best performance with a precision of 0.87, recall of 0.82, and an F1-score of 0.84. The Healthy class obtained the highest recall at 0.88, although its precision was relatively lower at 0.71. The Hispa class again showed the weakest performance, with a recall of 0.43 and an F1-score of 0.49, indicating that the model still struggles to accurately classify this class. Meanwhile, the LeafBlast class achieved a precision of 0.80 and an F1-score of 0.70. Overall, the model reached an accuracy of 0.74, with a macro-average F1-score of 0.71 and a weighted-average F1-score of 0.73, reflecting reasonably good classification performance but with room for improvement in certain classes, particularly Hispa.

To deepen the analysis of model performance, we conducted a visual examination of examples of misclassification that often occur in the Hispa class. From the confusion matrix and prediction output in the previous test data, the dominant error types are Hispa but classified as Healthy and Hispa but classified as BrownSpot. Examples of these cases are shown in Figure 2.

Fig. 2 Example of Missclassification

E-ISSN : 2807-9035
Volume 6, Number 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

Figure 2 shows examples of misclassification where leaves labeled Hispa were predicted as Healthy and examples where Hispa leaves were predicted as BrownSpot. Small lesions/early stages of attack on some Hispa samples show very subtle damage or small spots, so that the visual features are not prominent and are easily read as healthy leaves by the model. In addition, the similarity in texture/color in some cases, the color or pattern of the spots resembles BrownSpot, so the model is confused between Hispa and BrownSpot. This is thought to be due to variations in lighting conditions, blurring, and angles in some images having low contrast, shadows, or blurring that obscure the distinctive features of Hispa. The signs of disease in the images appear ambiguous, so there is a possibility that the original labels (dataset) are mislabeled, and therefore cannot be ignored without manual verification.

Based on the overall training results using the three optimizers which is ADAM, SGD, and RMSProp it is evident that ADAM and RMSProp outperform SGD in terms of model performance. The ADAM optimizer achieved the highest accuracy of 75.84%, followed by RMSProp with 74.60%, while SGD only reached 71.34%. This performance difference is attributed to the distinct characteristics of each optimizer in updating model weights. ADAM combines the advantages of both Momentum and RMSProp by adaptively adjusting the learning rate for each parameter, resulting in faster and more stable convergence. Meanwhile, RMSProp also adapts its learning rate dynamically based on recent gradient magnitudes, making it more efficient than SGD. In contrast, SGD employs a fixed learning rate without adaptive adjustment, causing slower optimization and a higher tendency to get trapped in local minima. Therefore, ADAM demonstrates the best performance due to its ability to balance convergence speed and learning stability effectively.

DISCUSSION

This study investigated the performance of the ResNet-50 architecture for rice leaf disease classification using three different optimization algorithms which is ADAM, SGD, and RMSProp. The experimental results revealed clear performance differences among the optimizers in terms of training accuracy, loss convergence behavior, and class-wise evaluation metrics. ADAM consistently achieved the highest training accuracy and the lowest loss values, followed closely by RMSProp, while SGD demonstrated comparatively lower accuracy and slower convergence. These findings indicate that adaptive learning rate-based optimizers are generally more effective for training deep convolutional neural networks on complex agricultural image datasets, where feature variability and noise are common.

From a class wise evaluation perspective, all optimizers performed well in identifying the BrownSpot and Healthy classes, as evidenced by their high recall and F1-score values. In contrast, the Hispa class consistently exhibited lower recall and F1-score across all experimental settings. This suggests that Hispa disease presents visual characteristics that are more difficult for the model to distinguish, potentially due to subtle lesion patterns, early-stage symptoms, or similarities with other classes such as BrownSpot. This observation aligns with common challenges in plant disease classification, where inter-class similarity and limited discriminative features can significantly degrade performance for certain disease categories.

Beyond visual similarity, several technical factors and dataset-related limitations likely contributed to the model's overall accuracy being constrained to approximately 75%. One major factor is the limited dataset size and class imbalance. The dataset used in this study consists of 3.355 images with an uneven class distribution, where the Healthy class contains 1.488 images, while the Hispa class has only 565 samples. Such imbalance reduces the availability of representative training examples for minority classes, limiting the model's ability to learn their full feature distribution and often resulting in lower recall performance for underrepresented classes.

Another critical factor is that the ResNet-50 model was trained from scratch without utilizing transfer learning. Given the relatively small size of the dataset, training a deep architecture with millions of parameters from random initialization poses inherent challenges. Without prior feature knowledge, the model may struggle to learn robust and generalizable representations, particularly for subtle disease patterns. In contrast, transfer learning using pre-trained weights (e.g., from ImageNet) is widely recognized for improving convergence speed and generalization performance, especially in narrow-domain applications such as agricultural image analysis. The absence of transfer learning in this study therefore likely contributed to the observed performance saturation.

Additionally, the lack of in-depth hyperparameter exploration represents another limiting factor. While this study focused on comparing different optimizers, other influential training parameters such as learning rate schedules, weight decay, momentum tuning (particularly for SGD), and regularization techniques were not systematically investigated. Suboptimal hyperparameter configurations, such as an excessively high learning rate or the absence of a learning rate decay strategy, can hinder effective convergence and prevent the model from achieving higher accuracy.

Overall, the combination of dataset limitations (size and class imbalance), experimental design choices (training from scratch without transfer learning), and an unexplored hyperparameter space provides a more comprehensive and convincing explanation for why the model's accuracy remains around 75%, beyond the factor of visual similarity alone. These findings highlight important considerations for future research aiming to improve deep learning–based rice disease classification systems.

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

CONCLUSION

This study demonstrates that the ResNet-50 architecture is capable of classifying rice leaf diseases with reasonably good performance when combined with appropriate optimization algorithms. Based on the experimental evaluation, the ADAM optimizer achieved the highest overall accuracy of 75.84%, outperforming RMSProp (74.60%) and SGD (71.34%). The superior performance of ADAM can be attributed to its adaptive learning rate mechanism and momentum-based updates, which enable faster and more stable convergence during training. These results confirm that optimizer selection plays a critical role in determining the effectiveness of deep learning models for plant disease classification tasks.

Despite these promising results, the overall classification accuracy remains below 76%, indicating several inherent limitations in the current experimental setup. Factors contributing to this performance ceiling include visual similarity between certain disease classes, class imbalance within the dataset, variations in image lighting and background conditions, and the limited number of training samples for diseases such as Hispa. In addition, the use of a single deep learning architecture trained from scratch, along with fixed hyperparameter configurations, may restrict the model's ability to learn more discriminative and generalizable feature representations.

Future research should therefore focus on improving model performance through more comprehensive experimental strategies. These include systematic hyperparameter tuning particularly for the SGD optimizer by optimizing learning rate schedules, momentum, and weight decay, as well as the implementation of transfer learning using pre-trained weights from large-scale datasets such as ImageNet. Addressing data related limitations through label cleaning, automated label-noise detection, and class imbalance handling techniques (e.g., class weighting, focal loss, or oversampling) is also crucial. Furthermore, incorporating attention or localization-based approaches may help the model focus on subtle disease lesions, while explainability tools such as Grad-CAM can support active learning by prioritizing ambiguous samples for relabeling. Finally, exploring ensemble strategies and lightweight architectures such as MobileNet or EfficientNet may lead to more efficient and deployable solutions for real-world rice disease detection systems.

REFERENCES

Azhari, D. (2024). Produktivitas dan prediksi produksi padi di Provinsi Bengkulu. *Technology and Agriculture Journal Azahri*, *5*(2), 1–10. doi:10.37638/sinta.5.2.279-288

Badan Pusat Statistik. (2024). *Luas panen dan produksi padi di Indonesia 2024*. Badan Pusat Statistik.

Batubara, M., & Pane, M. M. (2023). Pengaruh pertanian terhadap pendapatan nasional. *Jurnal Penelitian Ekonomi Akuntansi (JENSI)*, *7*(1), 74–81.

Berliani, T., Rahardja, E., & Septiana, L. (2023). Perbandingan kemampuan klasifikasi citra X-ray paru-paru menggunakan transfer learning ResNet-50 dan VGG-16. *Journal of Medicine and Health*, *5*(2), 123–135. doi:10.28932/jmh.v5i2.6116

Elizabeth, T., & Tinaliah. (2023). Klasifikasi lesi benign dan malignant pada rongga mulut menggunakan arsitektur ResNet50. *Jurnal MDP*, *10*(4).

Fauzan Novriandy, M., Rahmat, B., & Junaidi, A. (2024). Klasifikasi citra penyakit kanker mulut menggunakan arsitektur ResNet50 optimasi Adam dan SGD. *Jurnal Informatika dan Teknik Elektro Terapan*, *12*(3). doi:10.23960/jitet.v12i3.4732

Hawari, F. H., Fadillah, F., Alviandi, M. R., & Arifin, T. (2022). Klasifikasi penyakit padi menggunakan algoritma convolutional neural network. *Jurnal Responsif*, *4*(2), 184–189.

Istiqomah, N., & Murinto, M. (2024). Klasifikasi penyakit tanaman padi berbasis citra daun menggunakan convolutional neural network. *JSTIE (Jurnal Sarjana Teknik Informatika)*, *12*(1), 18. doi:10.12928/jstie.v12i1.27314

Julianto, A., Sunyoto, A., Ferry, D., & Wibowo, W. (2022). Optimasi hyperparameter convolutional neural network untuk klasifikasi penyakit tanaman padi.

Kresnawan, M. I., Pamungkas, D. P., Mahdiyah, U., & Nusantara, U. (2024). Identifikasi penyakit tanaman bawang merah menggunakan Faster R-CNN dan arsitektur ResNet50. *Prosiding*, 8.

Labhsetwar, R. S., Deshpande, R., Haridas, S., Kolte, P. A., Panmand, R., & Pati, S. (2023). Performance analysis of optimizers for plant disease classification with convolutional neural networks.

Nurdin, A., Satria, D., Kartika, Y., Rezha, A., & Najaf, E. (2024). Klasifikasi penyakit daun tomat dengan metode convolutional neural network menggunakan arsitektur Inception-V3.

Pirnando, H. N. N., Petrus, J., & Tinaliah. (2025). Klasifikasi penyakit daun padi menggunakan convolutional neural network dengan arsitektur AlexNet.

Pramono, M. S., & Wibowo, A. P. (2024). Penerapan convolutional neural network untuk identifikasi penyakit pada tanaman padi dari citra daun menggunakan model ResNet-101. *Jurnal Teknologi Informasi*, *5*(3). doi:10.46576/djtechno

E-ISSN : 2807-9035
Volume 6, **Number** 1, May 2026
https://doi.org/10.47709/brilliance.v6i1.7582

Pujiati, R., & Rochmawati, N. (2022). Identifikasi citra daun tanaman herbal menggunakan metode convolutional neural network. *Journal of Informatics and Computer Science, 3*.

Ramadhan, F., & Hernadi, J. (2025). Evaluasi optimizer Adam dan RMSProp pada arsitektur VGG-19 untuk klasifikasi ekspresi wajah manusia. *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika), 10*(2), 1414–1426. doi:10.29100/jipi.v10i2.6197

Ramadhani, K. F., & Tarigan, M. (2025). Implementasi metode convolutional neural network untuk klasifikasi penyakit daun mangga menggunakan arsitektur EfficientNetV2-S dan ResNet50. *Jurnal Mahasiswa Teknik Informatika, 9*(3).

Sahputra, I., Ulva, A. F., Putri, B. A., & Eviyan, C. Y. (2025). Comparative study of VGG16 and MobileNet architectures for rice leaf disease classification using CNN. *Journal of Informatics and Telecommunication Engineering, 9*(1). doi:10.31289/jite.v9i1.15458

Saleem, M. H., Potgieter, J., & Arif, K. M. (2020). Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers. *Plants*, *9*(10), 1–17. doi:10.3390/plants9101319

Santosa, A. A., Fu'adah, R. Y. N., & Rizal, S. (2023). Deteksi penyakit pada tanaman padi menggunakan pengolahan citra digital dengan metode convolutional neural network. *Journal of Electrical and System Control Engineering, 6*(2), 98–108. doi:10.31289/jesce.v6i2.7930

Setyawan, P. A., Danuarta, A., Nur Alam, F., & Agustin, T. (2024). Analisis perbandingan optimizer pada model convolutional neural network untuk deteksi kanker paru-paru. *Prosiding SEMNASA 2024.*

Shafiq, M., & Gu, Z. (2022). Deep residual learning for image recognition: A survey. *Applied Sciences, 12*(18). doi:10.3390/app12188972

Sheila, S., Anwar, M. K., Saputra, A. B., Pujianto, R., & Sari, I. P. (2023). Deteksi penyakit pada daun padi berbasis pengolahan citra menggunakan metode convolutional neural network.

Sutabri, T., & Darma, U. B. (2024). Perancangan aplikasi pendeteksi hama tanaman padi berbasis Android menggunakan convolutional neural network.

Ujawe, Gupta, & Waghmare. (2024). An advanced CNN-based system for rice plant disease detection and classification in precision agriculture. *Journal of Electrical Systems, 20*(4).

Wardani, K. R., & Leonardi, L. (2023). Klasifikasi penyakit pada daun anggur menggunakan metode convolutional neural network. *Jurnal Tekno Insentif, 17*(2), 112–126. doi:10.36787/jti.v17i2.1130