Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

Utilizing ResNet-50 for Deep Learning-Based Rice Leaf Disease Detection

Risna Sari[1*], Hedy Leoni Asbudi[2], Fitrah Eka Susilawati[3]
[1,2,3]Univerasitas Cokroaminoto Palopo, Indonesia
[1]sarisnasari@gmail.com, [2]hedyleoni@gmail.com, [3]fitrahesusilawati@gmail.com

open access

***Corresponding Author**

Article History:
Submitted: 07-12-2025
Accepted: 26-12-2025
Published: 29-12-2025

Keywords:
ResNet50; Rice Leaf Desease Classification; Transfer Learning; Deep Learning; Data Splitting Ratio.

ABSTRACT

Rice is a primary global food commodity, yet its productivity is frequently threatened by various diseases that significantly reduce both yield quality and quantity. Traditional manual diagnosis by farmers is often subjective, time-consuming, and prone to inaccuracies, necessitating more efficient automated solutions. This research evaluates the ResNet50 architecture for the automated classification of rice leaf diseases through digital image analysis. The study specifically investigates the model's performance on a specialized dataset and analyzes how different data splitting ratios influence accuracy and stability. A public dataset comprising four classes—Hispa, Healthy, Leaf Blast, and Brown Spot—was employed. The data underwent rigorous labeling, pre-processing, and augmentation to enhance sample diversity before being partitioned into training and testing sets using three ratios: 85:15, 80:20, and 90:10. The ResNet50 model was implemented using transfer learning with pre-trained ImageNet weights and fine-tuned on the classification layers. Experimental results reveal that the 85:15 split ratio achieved the highest accuracy of 81.48%, followed by 78.77% for the 80:20 ratio and 76.21% for the 90:10 ratio. These findings suggest that ResNet50 provides competitive performance for rice disease detection. Furthermore, achieving an optimal balance between training and testing data is critical for maximizing model generalization within modern smart farming applications.

INTRODUCTION

Agriculture is a strategic sector that contributes significantly to Indonesia's national food security and economic development. Among various agricultural commodities, rice plays a vital role as the primary staple food for the population. However, rice productivity is highly vulnerable to disease attacks, particularly those affecting the leaves. Diseases such as blast, brown spot, and tungro are known to reduce crop yields by up to 50% if not properly managed. Globally, plant diseases have been reported to cause agricultural yield losses ranging from 20% to 40% annually (FOA, 2023). Unfortunately, the identification of rice leaf diseases in the field still relies heavily on the subjective experience of farmers, which often results in delayed treatment and substantial economic losses.

The challenges associated with manual identification are further exacerbated by the complexity of rice leaf disease symptoms, which often exhibit similar visual patterns across classes. For instance, lesions caused by *brown spot* frequently resemble the early-stage symptoms of *leaf blast*, making them difficult to differentiate without specialized expertise (IRRI, 2020). Uncontrolled field conditions—such as variations in lighting, inconsistent backgrounds, and the presence of visual noise—further complicate the diagnostic process and may reduce accuracy. Therefore, a more accurate, rapid, and user-friendly disease detection system is needed, one that can be widely utilized by farmers without requiring extensive technical knowledge.

With advancements in technology, the use of artificial intelligence—particularly Deep Learning—has become one of the most widely adopted approaches in modern agriculture. This technology enables automatic recognition of visual patterns through digital images, making it highly suitable for leaf disease classification tasks. *Convolutional Neural Networks* (CNNs) serve as the primary algorithms in this approach due to their ability to extract essential image features without requiring manual feature engineering (Adetunji et al., 2023). Research trends indicate that CNNs consistently deliver high performance across a variety of plant disease classification tasks, even outperforming traditional texture- and color-based methods (Putra et al., 2023; Soekarta et al., 2023). Therefore, the use of image-based classification systems is expected to accelerate disease detection and assist farmers in making timely and informed decisions.

These challenges highlight the need for a disease detection system that is more accurate, efficient, and widely accessible to farmers. One rapidly developing approach in this domain is the application of artificial intelligence, particularly Deep Learning. This technology enables the automatic recognition of visual patterns from digital images, making it highly suitable for leaf disease classification tasks. In this context, *Convolutional Neural Networks* (CNNs) serve as the primary algorithms due to their ability to extract essential image features without manual intervention. The use of image-based classification systems is also expected to accelerate the detection process and support farmers in making timely and well-informed decisions.

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

One of the most widely used CNN architectures in image classification research is the Residual Network-50 (ResNet-50). ResNet-50 offers strong capabilities in detecting complex features through its deep network structure without experiencing accuracy degradation, made possible by its effective residual learning mechanism (Falaschetti et al., 2022). Numerous previous studies have applied CNN-based approaches to leaf disease detection across various agricultural commodities. ResNet-50 is among the architectures frequently utilized in *transfer learning*–based studies due to its ability to extract deep and meaningful visual features. For example, Putra et al., (2023) implemented ResNet-50 on a corn leaf image dataset and achieved an accuracy of up to 98.4% using a conventional data-split scenario. Other studies on tomato leaf diseases similarly demonstrated that CNN architectures can yield high performance in image classification tasks, with training accuracies exceeding 98% when applied to well-structured datasets. These findings indicate that advanced CNN architectures such as ResNet-50 possess strong potential for image-based plant disease detection, particularly when the training data contain representative variations of disease symptoms. Overall, empirical evidence demonstrates that ResNet-50 delivers highly competitive performance in image-based plant disease detection tasks.

Nevertheless, to date, only a limited number of studies have specifically evaluated the performance of ResNet-50 on rice leaf images obtained from real-world environments. Most previous research has relied on clean, uniform datasets collected under ideal laboratory conditions, which do not fully represent the visual complexity encountered in actual field settings. In practice, rice leaf images captured in paddy fields often contain noise, lighting variations, complex backgrounds, and other environmental factors that may degrade model performance (Thalor et al., 2025). Therefore, this study aims to implement the ResNet-50 model directly on a more realistic rice leaf dataset to assess its capability in classifying leaf diseases. In addition, this research provides further contribution by analyzing the effects of varying data-split ratios on model accuracy and performance stability—an aspect that has received limited attention in previous studies. The findings of this study are expected to contribute to the development of plant disease detection systems that are more adaptive, reliable, and applicable for farmers in real-world field conditions.

LITERATURE REVIEW

The use of leaf imagery for plant disease detection has advanced rapidly in parallel with developments in image processing and artificial intelligence technologies. Traditional approaches based on classical image processing typically involve segmentation, texture- or color-based feature extraction, and classification using conventional machine learning algorithms. However, these approaches have inherent limitations because they rely heavily on manually designed features (*handcrafted features*) and lack adaptability to field variations such as changes in lighting, camera angles, and the presence of noise in images. These limitations have driven a shift toward deep learning–based methods that are capable of automatically learning features directly from image data.

One of the earliest and most influential works in deep learning–based plant disease detection was conducted by Mohanty et al., who utilized the PlantVillage dataset comprising 54,306 leaf images representing 14 plant species across 26 disease and healthy classes. Their *Convolutional Neural Network* (CNN) models achieved an accuracy of up to 99.35% on test images collected under controlled laboratory conditions. However, when evaluated on images from different sources, the accuracy dropped significantly, underscoring the importance of diversity and real-world relevance in training data (Thalor et al., 2025). Subsequently, Ferentinos developed and evaluated multiple CNN architectures using 87,848 images covering 25 plant species and 58 plant–disease combinations, reporting accuracies as high as 99.53%. These findings further highlight the strong potential of CNNs as a dominant approach for image-based plant disease diagnosis.

In line with these developments, several survey studies have highlighted the central role of CNNs in plant leaf disease detection. Tugrul et al. reviewed 100 articles on the application of CNNs for detecting various leaf diseases over the past five years and concluded that CNNs consistently deliver high performance, although challenges remain regarding the need for large datasets, sufficient representation of disease symptoms, and generalization in real-world field conditions (Tugrul et al., 2022). Ahmad et al., 2023 similarly reviewed a range of deep learning techniques for plant disease detection, emphasizing the necessity for systems that are robust to environmental variations while also discussing key factors such as dataset quality, augmentation strategies, and architectural model selection.

Specifically for rice, deep learning–based research has grown rapidly. Simhadri et al. conducted a systematic review of various deep learning models for rice leaf disease detection, covering both custom CNN architectures and transfer learning models such as VGG, ResNet, and DenseNet. They found that most studies relied on datasets obtained from public sources or semi-controlled environments, typically featuring relatively homogeneous image backgrounds. This highlights a research gap concerning model performance when applied to images captured directly from agricultural fields, where visual complexity is significantly higher (Simhadri et al., 2025). Other studies have also demonstrated that CNNs can achieve high accuracy in rice disease detection; for example, Hairani et al. reported an accuracy of up to 99.7% in rice leaf disease classification after employing augmentation strategies and training with a substantial number of epochs (Hairani & Widiyaningtyas, 2024).

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

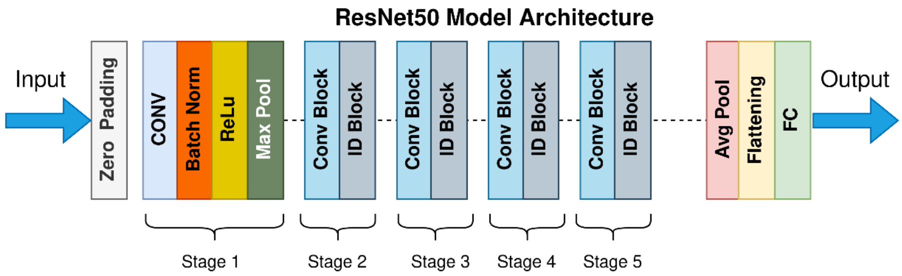

Fig. 1 Arsitektur ResNet-50

The ResNet-50 architecture holds a significant position within the *deep learning* ecosystem for image classification due to its ability to mitigate performance degradation in very deep networks through the use of residual learning mechanisms (Al-Gaashani et al., 2023). In the agricultural domain, ResNet-50 is widely employed as a backbone for image-based plant disease detection, both through transfer learning approaches and through the development of its derivative architectures. Ferentinos, for example, reported that ResNet and its variants rank among the highest-performing models across several plant disease classes (Hastari et al., 2024).

Recent studies have applied ResNet-50 and its variants specifically to rice plant disease detection. Hastari et al. classified rice leaf diseases using the ResNet-50V2 architecture on the RiceLeaf dataset obtained from Kaggle. Their study combined augmentation techniques with various optimizer configurations (Adam, RMSProp, and SGD) and three data-split scenarios (70:30, 80:20, and 90:10), resulting in exceptionally high training and testing accuracies, approaching 99% under several configurations (Hastari et al., 2024). Roseno et al., 2024 compared the performance of several CNN architectures—namely ResNet-50, VGG16, and MobileNetV3-Small—in classifying rice diseases. Their findings indicate that each architecture demonstrates distinct advantages under specific combinations of data conditions and training parameters, underscoring the importance of appropriate architectural and hyperparameter selection. However, most of the studies cited above continue to rely heavily on relatively clean and structured datasets and have not extensively analyzed the impact of varying data-split ratios on model accuracy and performance stability when applied to rice leaf images collected from real-world agricultural environments. Therefore, research that specifically evaluates the performance of the ResNet-50 architecture on realistic rice leaf images, while simultaneously analyzing the effects of multiple data-split scenarios, offers significant contributions in addressing this research gap and supporting the development of more practical rice disease detection systems for farmers.

METHOD

In this study, several stages were conducted, beginning with data collection, labeling, preprocessing, data splitting, followed by model training and model testing.

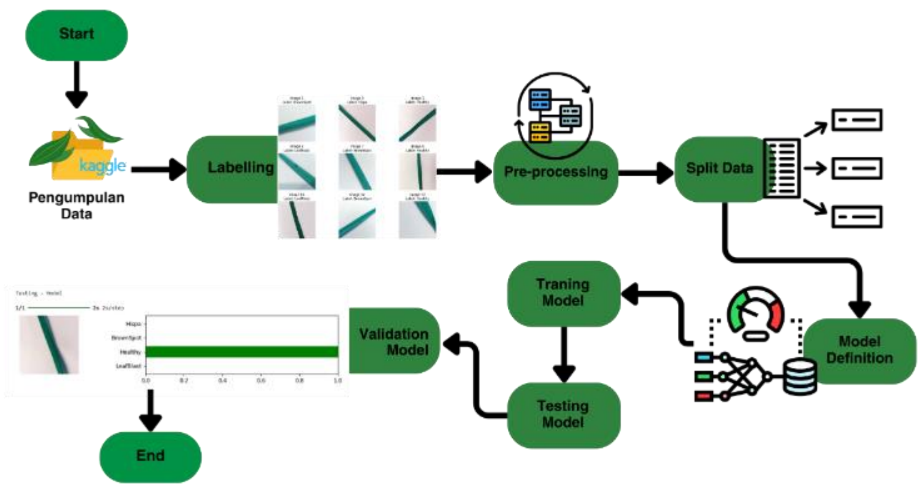

Fig. 2 Research Flow

Data Collection

The dataset used in this study was obtained from the Kaggle platform (https://www.kaggle.com/datasets/shayanriyaz/riceleafs). The dataset is already divided into two subsets, namely training data and testing data. It comprises four rice leaf categories: Brown Spot, Healthy, Hispa, and Leaf Blast.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

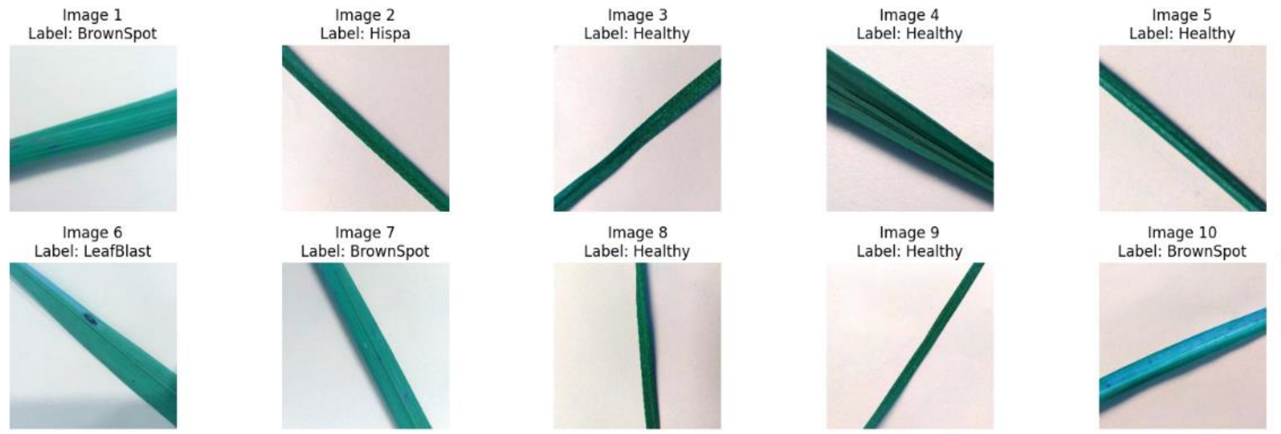

Fig. 3 Dataset

Labelling

The labeling process was conducted to categorize the images into predefined classes. Labeling is a crucial step, as it provides the foundation for the model to recognize visual patterns that distinguish one class from another. Each image was assigned a label corresponding to its category, namely *Healthy*, *Brown Spot*, *Leaf Blast*, and *Hispa*.

Pre-processing

Preprocessing was carried out to ensure that the image data met the requirements of the ResNet50 architecture. This stage began by resizing each image to 224×224 pixels, which is the standard input dimension for ResNet50. This resizing step was intended to standardize the resolution of all images so the model could process them consistently. Subsequently, pixel values were normalized to a range of 0–1 to maintain training stability and accelerate model convergence. Normalization also helps avoid large scale differences among pixel features, which could negatively affect the learning process.

In addition, *data augmentation* techniques were applied to increase the variability of the training data without adding new original samples. These techniques included rotation, flipping, and zooming, enabling the model to be exposed to various shapes and orientations of the objects. This strategy is effective in improving the model's generalization ability by making it more robust to changes in object position or viewpoint in new images. Thus, preprocessing not only prepares the data to be compatible with the ResNet50 model but also contributes to preventing overfitting and enhancing classification accuracy on the test data.

Data Splitting

Data splitting was conducted to divide the dataset into two primary subsets: the training set and the testing set, using several ratio variations, namely 90:10, 80:20, and 85:15. This separation aims to ensure that the testing data genuinely represent unseen samples, allowing the evaluation results to accurately reflect the model's generalization capability. The 90:10 ratio provides a larger portion of training data, enabling the model to learn from more information, whereas the 80:20 ratio allocates a larger portion of testing data for more rigorous evaluation. Meanwhile, the 85:15 ratio serves as a compromise between these two approaches.

In practice, the training data are used to adjust the model's weights and parameters during the training process, while the testing data are used to measure the model's final performance after training is completed. These ratio variations allow for a comparative analysis of the model's performance under different proportions of training and testing samples, enabling an assessment of how data distribution influences accuracy and generalization ability. By comparing the outcomes of all three data-splitting scenarios, a more comprehensive understanding of the performance stability of the ResNet50 model in image classification tasks can be obtained.

Model Definition

The model used in this study is ResNet50, a well-established architecture within the *Convolutional Neural Network* (CNN) family. ResNet50 consists of 50 layers designed to address the *vanishing gradient* problem commonly encountered in deep networks. The key innovation of ResNet lies in the use of *skip connections* or *shortcut connections*, which allow information to flow across layers without significant signal degradation. This mechanism enables ResNet to maintain high accuracy even as network depth increases substantially. The ResNet50 architecture comprises convolutional layers, batch normalization, ReLU activation functions, as well as global average pooling and fully connected layers for generating the final predictions.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

In this study, ResNet50 was implemented using *pre-trained weights* from the ImageNet dataset through a *transfer learning* approach, allowing the model to leverage general feature representations before being fine-tuned for the rice leaf dataset. The key parameters configured include an input image size of 224×224 pixels, the number of output classes based on the dataset labels, a *softmax* activation function in the output layer for multiclass classification, and the *categorical crossentropy* loss function. The Adam optimizer was employed with an adjusted learning rate to balance convergence speed and training stability. The combination of the ResNet50 architectural principles and the customized parameter configuration is expected to produce a model with high classification accuracy and strong generalization capability.

Training Model

The model training process utilized a fine-tuned ResNet50 architecture, in which the final layers were adjusted to match the number of classes in the research dataset. Training began with initializing the model's weights using the pre-trained ImageNet parameters, which were then adapted to learn specific patterns present in the leaf images. The loss function employed was *Categorical Crossentropy*, as this study involves multiclass classification and the function effectively measures the difference between the model's predicted probability distribution and the true labels. To optimize network weights, the Adam optimizer—known for its adaptive learning rate capability—was used, while Stochastic Gradient Descent (SGD) with momentum was considered when more stable weight updates were required.

Training parameters such as batch size and learning rate were determined through preliminary experiments to identify the optimal combination that balances training speed and model accuracy. The number of epochs was selected to ensure that the training process reached convergence, defined as the point where both training and validation losses no longer decrease significantly yet have not exhibited signs of overfitting. Throughout training, key performance metrics such as accuracy, loss, and optionally F1-score were monitored at each epoch using a validation set to ensure continuous model improvement. If signs of overfitting were detected, mitigation strategies such as early stopping, learning rate reduction, or additional regularization techniques were implemented to enhance the model's generalization ability on the test data.

Validation Model

Model validation was performed using a separate validation set to monitor the model's generalization ability throughout the training process. At this stage, the model's performance was evaluated at each epoch by observing changes in loss and accuracy, allowing detection of potential overfitting or underfitting. If overfitting was identified, corrective measures were applied, such as adding dropout layers to reduce dependency on training data, incorporating additional data augmentation to increase data variability, and adjusting the learning rate to stabilize the learning process. Through this validation procedure, the model was able to maintain a balanced performance between training accuracy and its generalization capability on unseen data.

Testing Model

Model testing was conducted after the training process was completed, using a testing set that was entirely excluded from both training and validation. The purpose of this testing phase was to obtain an objective overview of the model's final performance when classifying new images. Evaluation was carried out using several metrics, including accuracy to measure the proportion of correct predictions relative to the total test samples, and precision, recall, and F1-score to assess the accuracy and completeness of predictions for each class. Additionally, a confusion matrix was employed to visualize the distribution of correct and incorrect predictions across all classes. The results of this evaluation serve as the basis for assessing the effectiveness of the ResNet50 model and as an indicator of whether the model is ready for real-world deployment.

RESULT

The results of this study illustrate the rice leaf disease detection process using the ResNet50 architecture. Figure 4 presents a comparison of the number of images across four classes—*Hispa*, *Healthy*, *LeafBlast*, and *BrownSpot*—between the training set (blue bars) and the testing set (orange bars). Visually, it can be observed that the *Healthy* class dominates the dataset, followed by *LeafBlast* and *Hispa*, while *BrownSpot* constitutes the class with the fewest images.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

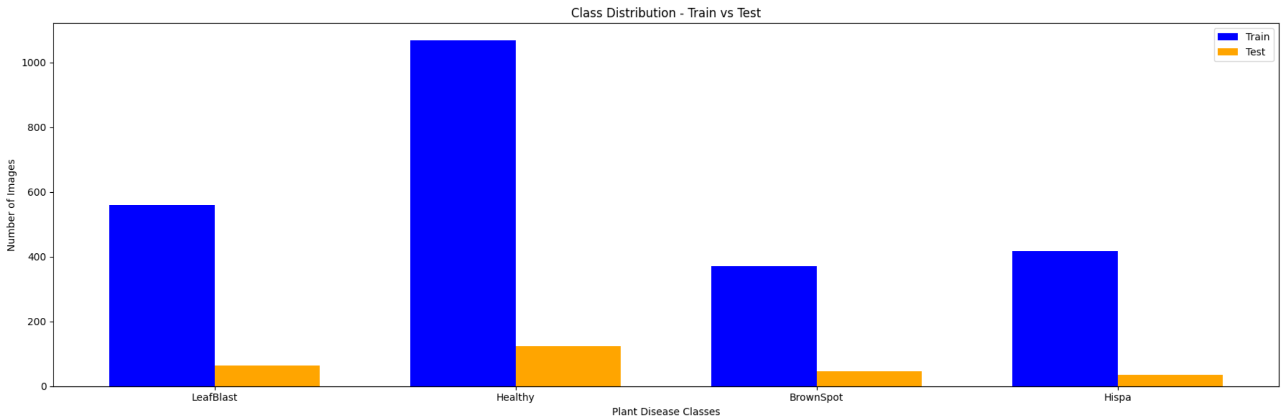

Fig. 4 Class Distribution

This ranking pattern is consistent across both the training and testing sets, indicating that the relative proportions among the classes follow the same trend. Such consistency suggests that the data-splitting process successfully preserved class representativeness—typically achieved through a stratified split—ensuring that the distribution in the testing set remains reflective of the training set population. Conversely, the disparity in sample availability across classes reveals a class imbalance, with the majority of samples belonging to the *Healthy* class, while *BrownSpot* contains the fewest examples. This condition is an important dataset characteristic, as it may influence per-class performance metrics (such as precision and recall for each label) and affect the interpretation of overall accuracy. The distribution information thus serves as an initial reference for interpreting evaluation results in subsequent sections.

This study employs the ResNet50 model with parameter details presented in Figure 5. The architecture consists of ResNet50 as the feature extractor (the ResNet50 layers), followed by a streamlined yet effective classification head: 2D Global Average Pooling (which transforms the 8×8×2048 feature map into a 2048-dimensional vector), a Dense layer (1024 units), a Dropout layer, and a Dense output layer (4 units) corresponding to the four target classes. This design leverages the advantages of Residual Learning in ResNet—namely the use of skip/shortcut connections—to maintain gradient flow in deep networks, while Global Average Pooling (GAP) is chosen over full flattening to reduce the number of parameters and mitigate overfitting without losing global information from the feature maps.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 8, 8, 2048)	23,587,712
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 2048)	0
dense_2 (Dense)	(None, 1024)	2,098,176
dropout_1 (Dropout)	(None, 1024)	0
dense_3 (Dense)	(None, 4)	4,100

Total params: 25,689,988 (98.00 MB)
Trainable params: 25,636,868 (97.80 MB)
Non-trainable params: 53,120 (207.50 KB)

Fig. 5 ResNet50

The parameter details in the classification head are clearly specified: the Dense (1024) layer contains 2,098,176 parameters (computed as 2048×1024 + 1024), the Dropout layer adds no parameters as it functions solely as a regularization mechanism, and the Dense (4) layer has 4,100 parameters (1024×4 + 4). Overall, the model comprises 25,689,988 parameters (~98 MB), with 25,636,868 trainable and 53,120 non-trainable parameters. This proportion indicates that nearly all weights—including most weights in the base model—undergo fine-tuning, whereas the non-trainable components primarily originate from the internal Batch Normalization statistics in ResNet. The output layer employs a *softmax* activation to generate probability distributions across the four classes. This configuration balances high representational capacity (enabled by ResNet50 and its 2048-dimensional feature vector) with a minimalist classification head (GAP → Dense → Dropout → Dense), resulting in efficient computation while maintaining strong generalization ability.

The model was evaluated under three training–testing split scenarios—85:15, 80:20, and 90:10—to assess the effect of data ratio on the performance of ResNet50. Each scenario produced training and validation loss and accuracy

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

curves, which were used to observe the model's convergence behavior throughout the training process.

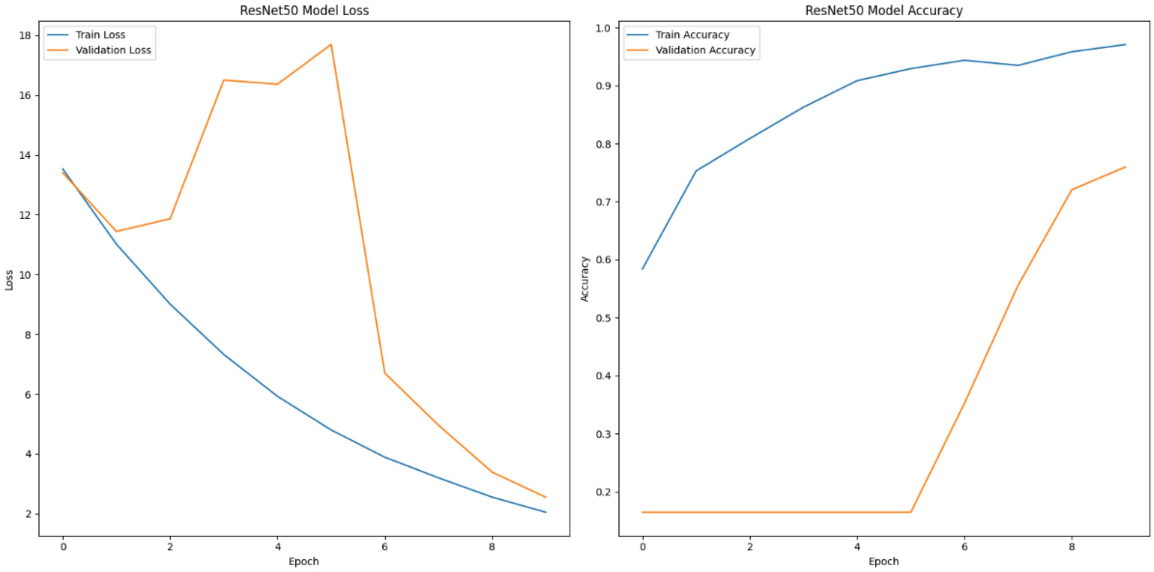

Fig. 6 Model Loss and Model Accuracy with an 85:15 Data Split

Figure 6 presents the training results for the 85:15 data split. The training loss curve exhibits a consistent decline up to the final epoch, accompanied by a significant reduction in validation loss after the fifth epoch, indicating that the model begins to adapt well to the validation data. The accuracy curves show a steady increase in training accuracy, while validation accuracy rises sharply after the sixth epoch. In this scenario, the model achieves a final accuracy of 81.48%, which is the highest among the three evaluated data-split configurations.

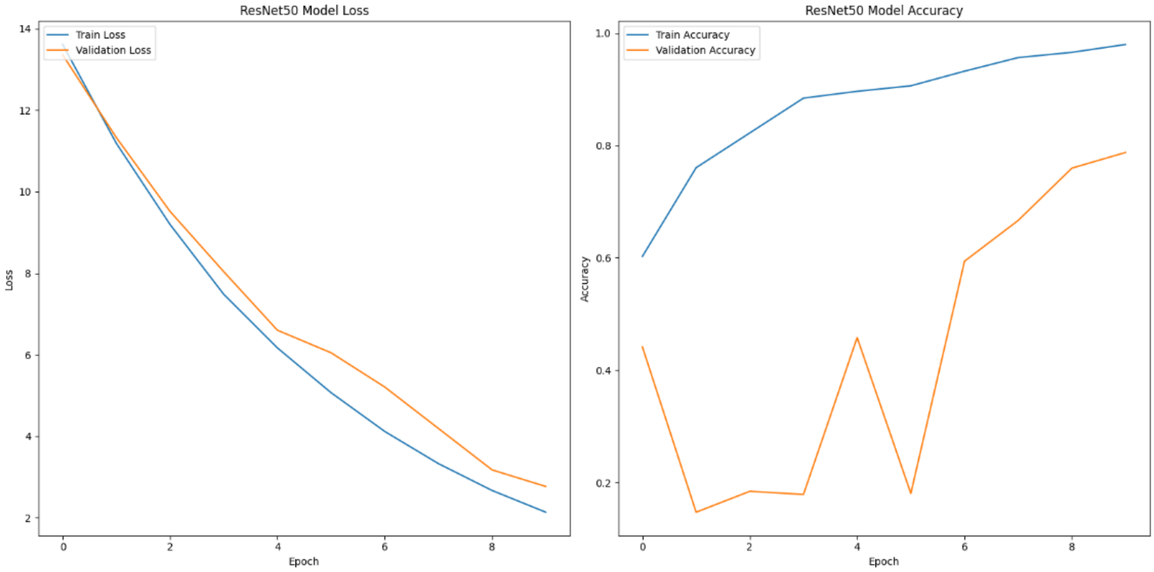

Fig. 7 Model Loss and Model Accuracy with an 80:20 Data Split

Figure 7 presents the training results for the 80:20 data split. The loss curves show a downward trend for both the training and validation sets, although fluctuations in validation accuracy are more pronounced compared to the previous scenario. This behavior may be attributed to the larger proportion of testing data, which increases the diversity of the validation set and introduces additional challenges for the model. The final accuracy achieved in this scenario is 78.77%, which is slightly lower than that obtained with the 85:15 data split.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

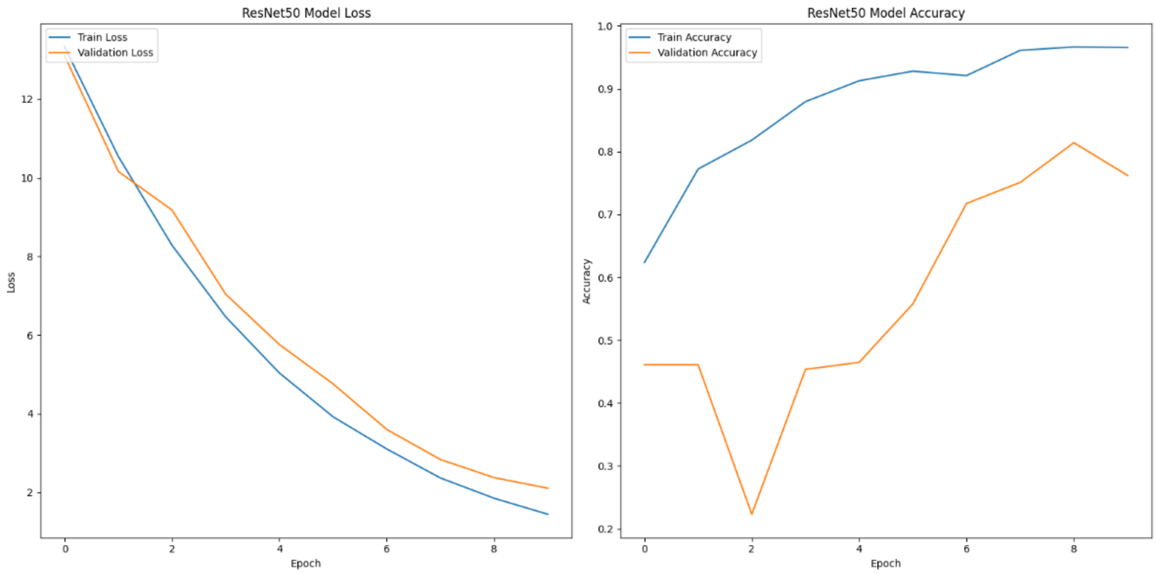

Fig. 8 Model Loss and Model Accuracy with a 90:10 Data Split

Figure 8 presents the training results for the 90:10 data split. Although the training and validation loss curves exhibit consistent decreases, the validation accuracy curve progresses more slowly compared to the other two scenarios. Due to the smaller amount of testing data, the model's validation performance tends to be less stable and does not reach the accuracy levels observed in the previous configurations. The final accuracy obtained in this scenario is 76.21%, making it the lowest among the three evaluated data-split ratios.

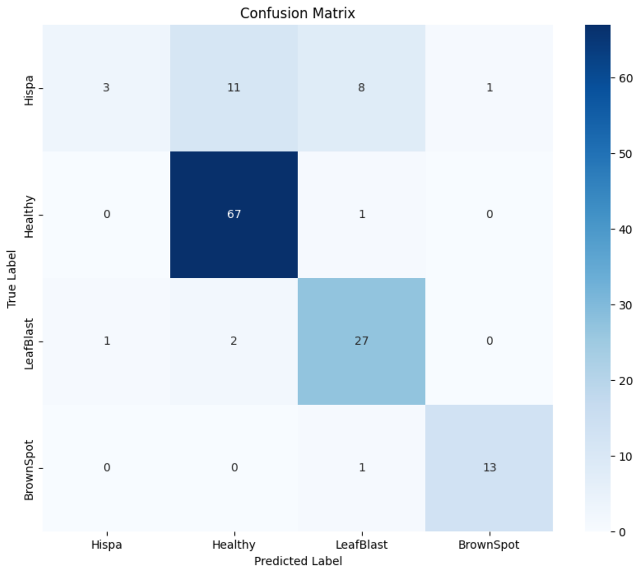

Fig. 9 Confusion Matrix Split Data 85:15

Figure 9 illustrates the confusion matrix for the 85:15 split ratio. The model demonstrates its strongest performance on the *Healthy* class, correctly classifying 67 images from all samples belonging to that category. However, the *Hispa* class shows a relatively high number of misclassifications, with many images incorrectly predicted as *Healthy* or *LeafBlast*. The *LeafBlast* class achieves 27 correct predictions, while *BrownSpot* records 13 correctly classified samples.

Based on these results, it can be concluded that the 85:15 ratio provides the best balance between training and testing data, yielding a more optimal model performance compared to the other split configurations. Overall, the model exhibits strong consistency in recognizing the *Healthy* class, yet still struggles to distinguish between visually similar categories, particularly *Hispa* and *LeafBlast* as well as *BrownSpot*. These findings serve as an important foundation for discussing potential improvement strategies in the following section.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

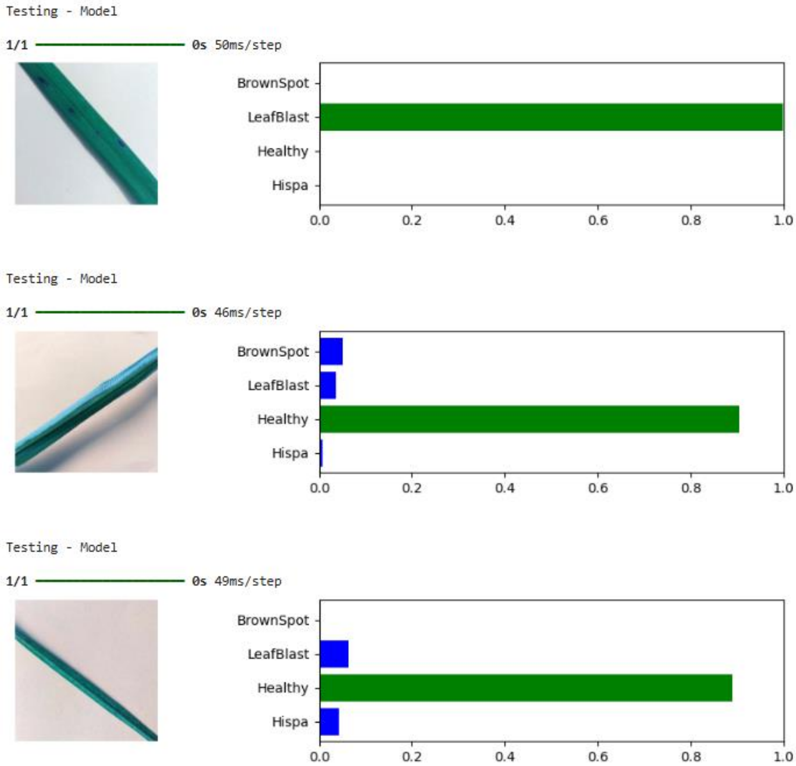

Fig. 10 Testing Model

Figure 10 presents the model's performance in detecting rice leaf diseases. In the first image, the model exhibits an almost perfect level of confidence that the leaf is infected with *LeafBlast*, with a probability approaching 100%. The bar indicator for the *LeafBlast* class is fully highlighted in green, while other classes such as *BrownSpot*, *Healthy*, and *Hispa* receive negligible probabilities. This result suggests that the visual features of the leaf closely match those typically associated with *LeafBlast*, enabling the model to classify it with high certainty. In the second image, the model predicts the leaf as healthy with a high confidence level, exceeding 90%. Nevertheless, small probabilities are assigned to the *BrownSpot* and *LeafBlast* classes, indicated by thin blue bars on the chart. These minor values may arise from subtle textures or lighting conditions that resemble early disease symptoms, although not sufficiently prominent to alter the primary classification.

In the third image, the model again identifies the leaf as healthy with a similar confidence level of over 90%. As with the previous case, small probabilities appear for *LeafBlast* and *BrownSpot*, likely due to non-disease factors such as shadows or surface reflections. Despite these minor variations, the main prediction consistently indicates that the leaf is in healthy condition. Overall, the three prediction results demonstrate that the model effectively distinguishes between healthy leaves and leaves infected with *LeafBlast*. In addition to its accuracy, the model performs efficiently, with prediction times consistently ranging between 46–50 milliseconds per step. This indicates that the model not only yields reliable classification results but also possesses strong potential for real-time deployment in rice leaf disease detection systems.

DISCUSSION

The findings of this study indicate that the performance of the ResNet50 model in detecting rice leaf diseases is strongly influenced by the characteristics of the dataset and the data-split configuration employed. The dataset exhibits class imbalance, with *Healthy* being the dominant class while *BrownSpot* is notably underrepresented. Such imbalance can affect per-class performance, especially for classes with fewer samples. This observation aligns with the claims of Ferentinos, 2018; Kamilaris & Prenafeta-Boldú, 2018, who noted that CNNs tend to more easily recognize patterns in classes with abundant representation. In this context, the model consistently identifies the *Healthy* class but struggles to distinguish *Hispa*, *LeafBlast*, and *BrownSpot*, which share strong visual and textural similarities.

The architectural analysis demonstrates that using ResNet50 as a feature extractor provides strong representational capacity due to its residual learning mechanism, which maintains gradient flow in deep layers, as described by (He et al., 2016). The use of Global Average Pooling (GAP) combined with a minimalist classification head (Dense–Dropout–Dense) offers a balanced trade-off between model complexity and the risk of overfitting. This balance is essential given that the dataset contains noise, lighting variations, and heterogeneous backgrounds—common

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

characteristics of field-collected images. Previous studies (Mohanty et al., 2016; Picon et al., 2019), similarly report that field images generally yield lower performance compared to clean laboratory datasets. The findings here are consistent with that pattern, with optimal accuracies ranging from 76% to 81%, noticeably lower than laboratory-based studies that often surpass 95%.

The influence of data-split ratios reveals an interesting relationship between training size and model generalization. In the 85:15 scenario, the model achieves the highest accuracy of 81.48%, indicating that a sufficiently large training set allows better learning of visual patterns while retaining a representative testing portion. In contrast, accuracy declines to 78.77% in the 80:20 split and further decreases to 76.21% in the 90:10 split. These results suggest that maintaining an appropriate balance between training and testing data is crucial for achieving optimal generalization.

The confusion matrix further clarifies model misclassification tendencies. ResNet50 shows high sensitivity toward the *Healthy* class but frequently misclassifies *Hispa* and *BrownSpot* samples as *LeafBlast* or *Healthy*. This challenge can be attributed to visual similarities among disease symptoms, as also highlighted by IRRI (2020), which states that early-stage symptoms of several rice diseases are difficult to distinguish without detailed visual analysis. These findings underscore the need for improved techniques in future work, such as minority-class augmentation, class weighting, or focal loss. Overall, the discussion highlights that ResNet50 delivers competitive performance in rice leaf disease detection on realistic datasets while still facing challenges related to minority classes and visually similar disease patterns. Thus, this study contributes valuable insights toward developing adaptive, efficient, and field-applicable image-based systems for rice disease detection.

CONCLUSION

Based on the findings of this study, the ResNet50 model is shown to be capable of classifying rice leaf diseases with satisfactory accuracy. The 85:15 data-split scenario achieved the highest performance with an accuracy of 81.48%, followed by the 80:20 split with 78.77%, and the 90:10 split with 76.21%. The model demonstrates strong capability in identifying the *Healthy* class, although higher misclassification rates were observed for classes such as *Hispa* and *LeafBlast*, largely due to class imbalance. These findings indicate that ResNet50—previously proven effective for disease detection in other plant species—retains competitive performance when applied to rice leaf datasets captured under real-world visual conditions with diverse textures, colors, and shapes.

Furthermore, the variations in training–testing ratios were found to influence the stability and representativeness of the model's performance. The 85:15 ratio provided the most balanced distribution of training and testing samples, resulting in the highest overall accuracy. In contrast, the 90:10 and 80:20 ratios exhibited decreased performance, attributable to a lack of representativeness in the testing subset or insufficient training samples for effective learning of visual patterns.

REFERENCES

Adetunji, O. J., Adeyanju, I. A., Esan, A. O., & Sobowale, A. A. (2023). *Flood Image Classification using Convolutional Neural Networks*. *6*(2), 113–121.

Ahmad, A., Saraswat, D., & El Gamal, A. (2023). A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agricultural Technology*, *3*(June 2022), 100083. https://doi.org/10.1016/j.atech.2022.100083

Al-Gaashani, M. S. A. M., Samee, N. A., Alnashwan, R., Khayyat, M., & Muthanna, M. S. A. (2023). Using a Resnet50 with a Kernel Attention Mechanism for Rice Disease Diagnosis. *Life*, *13*(6), 1–14. https://doi.org/10.3390/life13061277

Falaschetti, L., Manoni, L., Di, D., Pau, D., Tomaselli, V., & Turchetti, C. (2022). HardwareX A CNN-based image detector for plant leaf diseases classification. *HardwareX*, *12*, e00363. https://doi.org/10.1016/j.ohx.2022.e00363

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*(January), 311–318. https://doi.org/10.1016/j.compag.2018.01.009

FOA. (2023). *Disasters On Agriculture And Food Security Through Investment In Resilience*.

Hairani, H., & Widiyaningtyas, T. (2024). Augmented Rice Plant Disease Detection with Convolutional Neural Networks. *INTENSIF: Jurnal Ilmiah Penelitian Dan Penerapan Teknologi Sistem Informasi*, *8*(1), 27–39. https://doi.org/10.29407/intensif.v8i1.21168

Hastari, D., Winanda, S., Pratama, A. R., Nurhaliza, N., & Ginting, E. S. (2024). Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease. *Public Research Journal of Engineering, Data Technology and Computer Science*, *1*(2), 71–77. https://doi.org/10.57152/predatecs.v1i2.865

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition Kaiming. *CVPR*, *32*(5), 428–429. https://doi.org/10.1246/cl.2003.428

IRRI. (2020). *Rice Leaf Diseases Guide*. https://www.irri.org/resources/rice-knowledge-bank/rice-diseases

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in*

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7425

Agriculture, *147*(July 2017), 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, *7*(September), 1–10. https://doi.org/10.3389/fpls.2016.01419

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, *161*(October 2017), 280–290. https://doi.org/10.1016/j.compag.2018.04.002

Putra, I. P., Rusbandi, R., & Alamsyah, D. (2023). Klasifikasi Penyakit Daun Jagung Menggunakan Metode Convolutional Neural Network AlexNet. *Sudo Jurnal Teknik Informatika*, *2*(1), 28–33. https://doi.org/10.56211/sudo.v2i1.227

Roseno, M. T., Oktarina, S., Nearti, Y., Syaputra, H., & Jayanti, N. (2024). Comparing CNN Models for Rice Disease Detection: ResNet50, VGG16, and MobileNetV3-Small. *Journal of Information Systems and Informatics*, *6*(3), 2099–2109. https://doi.org/10.51519/journalisi.v6i3.865

Simhadri, C. G., Kondaveeti, H. K., Vatsavayi, V. K., Mitra, A., & Ananthachari, P. (2025). Deep learning for rice leaf disease detection: A systematic literature review on emerging trends, methodologies and techniques. *Information Processing in Agriculture*, *12*(2), 151–168. https://doi.org/10.1016/j.inpa.2024.04.006

Soekarta, R., Nurdjan, N., & Syah, A. (2023). Klasifikasi Penyakit Tanaman Tomat Menggunakan Metode Convolutional Neural Network (CNN). *Insect (Informatics and Security): Jurnal Teknik Informatika*, *8*(2), 143–151. https://doi.org/10.33506/insect.v8i2.2356

Thalor, M., Chavhan, Y., & Mate, S. (2025). *Performance Comparison of CNN Models for Tomato Disease Detection using Image-Based Data in both Controlled and Real-World Conditions*. *13*(1).

Tugrul, B., Elfatimi, E., & Eryigit, R. (2022). Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review. *Agriculture (Switzerland)*, *12*(8). https://doi.org/10.3390/agriculture12081192