Intechno Journal :
Information Technology Journal
Vol. 7 No. 2 December 2025

Homepage: https://jurnal.amikom.ac.id/index.php/intechno

INTECHNO Journal
INFORMATION TECHNOLOGY JOURNAL

e-ISSN: 2655-1438
p-ISSN: 2655-1632

.

An Advanced Deep Learning Approach for Automatic Disease Recognition and Classification in paddy leaf disease detection

Robert Marco[1], Alva Hendi Muhammad[2], Nur Aini[3], Yana Hendriana[4]

[1,2] *Magister of Informatics, Faculty of Computer Science, Universitas AMIKOM Yogyakarta, Yogyakarta, Indonesia*
[3]*Informatics, Faculty of Computer Science, Universitas AMIKOM Yogyakarta, Yogyakarta, Indonesia*
[4]*Informatics, Faculty of Information Technology, Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, Indonesia*

Article Info

Article history:

Received: 18 December 2025
Revised: 24 December 2025
Accepted: 30 December 2025
Published: 31 December 2025

Keywords:

Deep Learning
Convolutional Neural Network
Long Short-Term Memory
Attention mechanism
Paddy leaf disease detection

ABSTRACT

Purpose: Accurate detection of paddy leaf diseases is essential to ensure optimal crop yield and effective disease management.

Methods/Study design/approach: This study presents a CNN–LSTM–Attention model for improving classification performance by effectively capturing spatial, temporal, and contextual feature dependencies. The experimental results demonstrate that the proposed model consistently outperforms widely used deep learning architectures, including MobileNetV2, VGG16, and a custom CNN, across all evaluation metrics.

Result/Findings: Experimental results demonstrate that the proposed approach consistently outperforms the comparative models, achieving an accuracy of 95.5%, precision of 98.12%, recall of 98.3%, and a Macro-AUC of 0.994. These results indicate that the attention-enhanced sequential modeling significantly improves both class separability and overall robustness.

Novelty/Originality/Value: The findings confirm that the proposed CNN–LSTM–Attention model provides a reliable and effective solution for complex classification tasks, offering strong generalization performance and high suitability for real-world deployment.

Corresponding Author:

Robert Marco
Email: robertmarco@amikom.ac.id

1. INTRODUCTION

Paddy is the primary food source for more than half of the world's population [1]. It is cultivated on over three-fifths of global agricultural land, accounting for 61.54% of the total area. Indonesia is currently the fourth-largest paddy producer in the world, after China, India, and Bangladesh [2]. Unfortunately, plant pathogens and pests significantly contribute to yield reduction, with their impact largely influenced by seasonal conditions and unfavorable environmental factors, ultimately leading to economic and social losses [3]. Paddy diseases cause severe damage to agriculture, resulting in substantial yield losses and posing a threat to food security [4][5]. The urgency of this research lies in the fact that rice leaf diseases have caused yield reductions of 30–50% in several major production centers.

The determination of disease severity in paddy plants is often based on the extent and distribution of infection on the leaf surface [6]. Paddy crops are susceptible to diseases such as bacterial blight, brown leaf spot, leaf blast, and false smut [7]. Bacterial blight can inhibit photosynthesis and plant growth, while brown leaf spot reduces productivity. Leaf blast causes leaf fragility and hinders panicle formation, and false smut decreases grain quality [4][7]. The challenge is that manual detection of paddy leaf diseases is time-consuming, prone to errors, costly, and often relies on subjective visual inspection due to the complex nature of the diseases and the similarity of symptoms [8].

Previous studies have employed machine learning approaches, such as SVM, Random Forest, k-NN, and Naïve Bayes [9][10]. However, these methods often require multiple preprocessing steps and feature extraction, and they may not adapt well to different data types or specific parameters [10][11]. Deep learning (DL) approaches, such as CNNs, can automatically extract features from images, capturing low-, mid-, and high-level representations. Furthermore, transfer learning (TL) models, including VGG16, ResNet, Inception, MobileNet, and EfficientNet-V2, have shown promising results in terms of accuracy [12][13][14].

Deep learning (DL) approaches have several limitations, including dependence on large datasets, high computational costs, and limited interpretability [8]. However, these constraints can be mitigated through techniques such as eXplainable Artificial Intelligence (XAI), attention mechanisms, and hybrid models [3]. In response to these challenges, DL techniques have gained increasing attention in the field of disease detection [15].

Research on image-based identification of paddy plant diseases still faces significant challenges that have not been fully resolved. Limitations in publicly available datasets, such as the 'PlantVillage' dataset [16] which contains a restricted number of leaf disease segmentation images and suffers from class imbalance [8][3] can adversely affect model accuracy. Some diseases exhibit similar symptoms, complicating precise diagnosis [17], Misclassification can also occur in images with complex backgrounds [18] poor lighting, or noise [19][4] further increasing errors. Several previous studies have developed DL models but still report unsatisfactory results, high computational costs, and overfitting [8][18][19][20][21]. Additionally, the tendency of DL models to act as "black box" makes interpretation difficult for end users, who require results that are not only accurate but also easily understandable [18][22].

The state-of-the-art in this study focuses on strengthening CNN-based classifiers for rice leaf disease identification [18]. Study [23] proposed a deep learning approach incorporating shallow layers and residual connections, achieving an accuracy of 99.66%. Study [24] reported that a fully connected CNN provided fast and effective performance with an accuracy of 99.7%. In study [1], the Deep-CNN method achieved an accuracy of 99.81%. Research employing transfer learning, such as study [25], presented a self-attention network based on the ResNet50 architecture with a kernel attention mechanism, attaining an accuracy of 98.71%. Study [15] utilized EfficientNet-B4, reaching 99.09% accuracy in training and 96.91% in testing. Another approach combined CNN feature extraction with a self-attention mechanism using the ResNet34 architecture, resulting in 98.54% accuracy [10]. The effectiveness of VGG16 was demonstrated by its strong generalization ability to unseen images, yielding an accuracy of 99.94% [5]. Study [26] employed SqueezeNet with neural networks, obtaining an accuracy of 93.3%. Finally, study [18] introduced PlantDet, which achieved an accuracy of 98.53%.

The main contribution of this study is a hybrid CNN+LSTM+Attention Mechanism approach. CNN is employed to extract spatial features from leaf images, LSTM captures sequential dependencies among the features, and the attention mechanism enhances the model's focus on critical regions relevant for disease identification. An XAI-based approach is integrated into the classification process, allowing users to understand which areas of the leaf image most influence the model's decisions. Additionally, fine-tuning and hyperparameter optimization techniques are applied. Performance comparisons are conducted using transfer learning models, including MobileNetV2, VGG16, and a custom CNN.

The rest of this study will then review related works, Section 3 describes the proposed approach method, Section 4 presents the empirical results and discussion, the Last section describes the conclusions and future work

2. METHOD

The dataset used in this project was obtained from an ongoing Kaggle competition, namely the Paddy Disease Classification task (link: https://www.kaggle.com/datasets/imbikramsaha/paddy-doctor) he dataset contains ten classes, as described in the introduction of this report, and consists of 10,406 images. Each image has a height of 640 pixels and a width of 480 pixels, with a maximum pixel value of 255.0, a minimum value of 0.0, a mean pixel intensity of 115.9670, and a standard deviation of 71.6155. The detailed characteristics of the dataset are illustrated in Fig. 1.

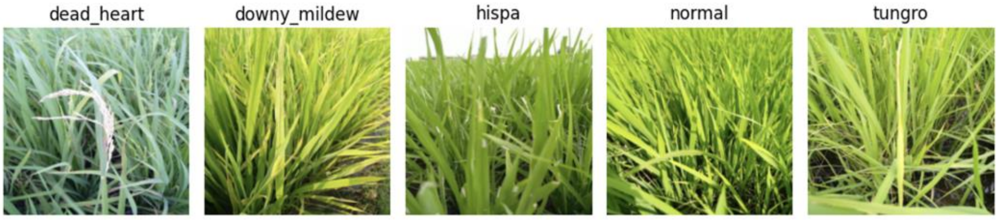

Figure 1. Rice leaf disease dataset sample.

2.1. Convolutional Neural Network (CNN)

CNN as one of the technique of feedforward neural network that uses convolution, pooling, and Rectified Linear Unit (ReLU) layers [27], CNN architecture is illustrated in Figure 2. 1D-CNN generally consists of four primary layers, which include pooling layers, convolutional layers, activation functions, and dropout layers.

The convolution layer utilizes a filter of suitable size to perform the convolution operation on data. This layer incorporates a local receptive field, allowing neurons to interpret information and extract high-level features. Furthermore, it minimizes the number of parameters by employing weight sharing, where the same set of parameters is applied across all positions of the convolution kernel [28]. The convolution layer and ReLU activation are mathematically expressed in Equation 1 and Equation 2, respectively.

$$h = f(x \otimes W + b) \tag{1}$$

$$f(x) = \max(0, x) \tag{2}$$

Where $\otimes$ denotes the convolution operation, W signifies the weights of the convolution kernel, and b represents the bias term. The function $f(\cdot)$ corresponds to the activation function, specifically the ReLU in this context. ReLU introduces nonlinearity into the network, allowing it to handle more complex problems. Additionally, ReLU forces some neuron outputs to zero, which promotes network sparsity and reduces parameter interdependence. This helps in extracting relevant features, better fitting the training data, and mitigating overfitting. The pooling layer performs a sampling operation on the output of the convolution layer, aggregating features from similar regions by selecting the maximum value. This layer retains only the most important features, reducing the amount of feature data. The features are extracted from the original data using the CNN, and the processed data are then fed into the LSTM, enhanced by the attention mechanism, for classification prediction [28].

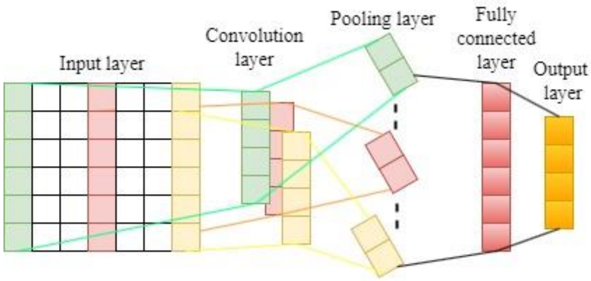

Figure 2. Structure of CNN

2.2. Long Short-Term Memory (LSTM)

LSTM was designed to address the vanishing and exploding gradient problems faced by traditional RNN [29]. It also resolves the issue of long-term dependencies that RNNs struggled with, though it still encounters overfitting challenges due to the large number of parameters that need adjustment [30].

Figure 3 illustrates the structure of the LSTM neural network. In this configuration, i_t signifies the input gate, f_t denotes the forget gate, and o_t represents the output gate. x_t is the input at the current time step, while C_{t-1} and h_{t-1} refer to the cell state and output from the previous time step. On the other hand C_t and h_t indicate the current time step's cell state and output, respectively. The LSTM leverages a unique gating mechanism to manage the forget, input, and output gates as well as the cell states, which helps in effectively capturing long-term dependencies within its memory units. The interactions between these components are mathematically described by the following equations [29][31].

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{3}$$

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{4}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{5}$$

$$\tilde{C}_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \quad (6)$$

$$C_t = f_t \times C_{t-1} + i_t + \tilde{C}_t \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

In this equation, σ represents the sigmoid activation function, W stands for the weight of the neuron, and b is the bias term of the neuron.

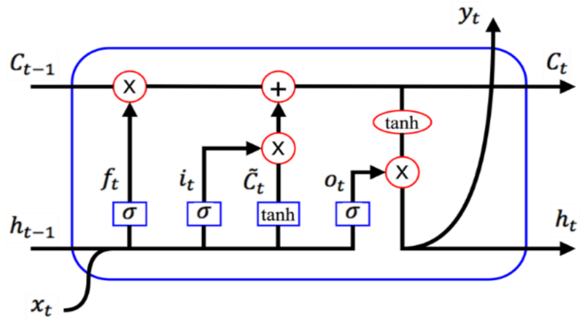

Figure 3. LSTM architecture [32]

2.3. Attention Mechanism

The attention mechanism enhances both the detection accuracy and interpretability of the model. By filtering out irrelevant disturbances, this technique allows the model to concentrate on critical information, thereby generating optimal outputs [31]. $x_1, x_2, .., x_t$ refers to the inputs of the LSTM, while $h_1, h_2, .., h_t$ signifies the outputs from the hidden layer of the LSTM, which serve as inputs to the attention mechanism to determine the distribution of attention weights. These weights indicate the significance of the state parameters. The attention mechanism formula is stated as follows [28]:

$$e_i = utanh(W_i h_i + b) \quad (9)$$

$$a_i = \frac{\exp(e_i)}{\sum_i \exp(e_i)} \quad (10)$$

$$C = \sum_i a_i h_i \quad (11)$$

Where, e_i represents the attention probability distribution for h_i at the ith moment. The symbols u and W refer to the weighting coefficients, while b denotes the bias coefficient, and C represents the weighted feature.

2.4. Proposed Algoritma

Convolutional Neural Networks (CNNs) are highly effective in extracting spatial features from images but are limited in capturing sequential dependencies. Conversely, Long Short-Term Memory networks (LSTMs) are well-suited for modeling temporal or sequential patterns [28][33][28]. To leverage the strengths of both, we propose a hybrid CNN-LSTM model with an attention mechanism to detect 10 classes of rice leaf conditions (nine disease categories and one healthy class).

In the proposed architecture, CNN layers first extract discriminative spatial features from rice leaf images through convolution and pooling operations. The extracted feature maps are then fed into LSTM layers, which capture sequential dependencies among feature representations. To enhance feature selection, an attention mechanism is applied, enabling the model to prioritize the most informative features while suppressing irrelevant noise, thereby improving classification performance.

The training and testing process of the proposed model begins by initializing all trainable parameters. Next, the training dataset $D = \{(x_i, y_i) | i = 1, ..., n\}$ is fed into the CNN through the input layer. The CNN then uses pooling layers and convolution to extract the feature matrix

The training and testing process begins with dataset preprocessing and initialization of trainable parameters. Each image input x_i with its corresponding label y_i is processed through CNN layers, generating a feature matrix $X = [X_1, ..., X_i]$ .This feature representation is passed to the LSTM layers, where temporal dependencies are modeled across sequential patterns in the data. The hidden states $H_t = [h_i, ..., h_t]$, are further refined through the attention layer, which assigns adaptive weights to highlight the most relevant features. Finally, a fully connected layer and softmax activation classify the output into one of the 10 rice leaf disease classes.

The CNN block (Conv2D) employs 64 kernels with a kernel size of 3×3, followed by ReLU activation, dropout, and max-pooling layers. A dropout rate of 0.5 is applied to prevent overfitting. The LSTM component consists of two layers with 128 and 64 hidden units, respectively, designed to balance performance and computational cost. The attention mechanism layer adaptively learns feature weights and outputs a refined feature vector. A flatten layer converts the multidimensional representation into a one-dimensional vector before passing it to dense layers. Finally, the softmax classifier generates the probability distribution over the 10 classes.

2.5. Performance Evaluation Measures

Performance metrics assess the effectiveness of the learning model defined by the following equation.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$F1 = \frac{2 \times precision \times recall}{precisio + \times recall} \quad (14)$$

$$Recall = \frac{TP}{TP + FN} \quad (15)$$

Where, TP is denoted as true positive, FP is denoted as false positive, TN is denoted as true negative, and FN is denoted as false negative

3. RESULTS AND DISCUSSIONS

The proposed CNN-LSTM-Attention model is designed to effectively extract spatial features from paddy leaf images, capture sequential dependencies among features, and emphasize relevant regions for precise disease classification. To train the model, we adopt a combined loss function including a reconstruction loss and Kullback-Leibler divergence loss, which together form the evidence lower bound objective (ELBO). This enables the model to learn informative latent representations while ensuring the predictions remain closely aligned with the true data distribution.

Figure 4 illustrates the training and validation loss over 50 epochs. The figure shows a rapid reduction in both training and validation loss during the initial epochs, with the curves stabilizing after approximately 10 epochs. The plateau in loss values indicates that the model has successfully converged, suggesting a stable and efficient training process. This demonstrates that the CNN-LSTM-Attention mechanism is capable of learning discriminative features for all 10 classes of paddy leaf diseases.

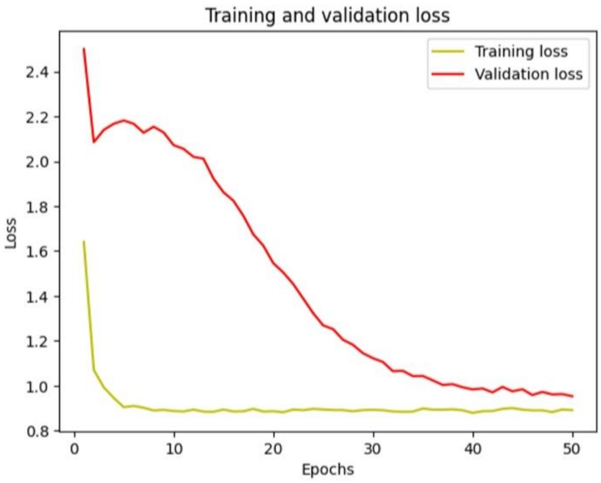

Figure 4. Validation Loss of CNN-LSTM-Attention Mechanism Models

Figure 5 depicts the training and validation accuracy across the same epochs. The training accuracy quickly rises and saturates near 0.85, while the validation accuracy steadily increases, reaching a similar level. This trend indicates that the model generalizes well to unseen data without significant overfitting.

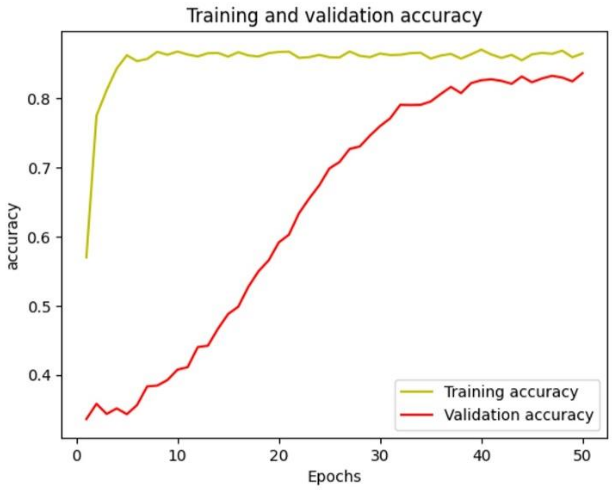

Figure 5. Validation Accuracy of CNN-LSTM-Attention Mechanism Models

Table 1 provides a detailed summary of the CNN-LSTM-Attention model architecture. The input layer processes image patches of shape (16, 32, 32, 3). The time-distributed layer extracts spatial features for each patch, followed by a bidirectional LSTM layer capturing temporal dependencies across features. The attention layer then assigns weights to highlight the most critical features for classification. Dense and dropout layers are used to produce the final output while minimizing overfitting. In total, the model contains 301,899 trainable parameters, balancing model complexity with computational efficiency.

Table 1. The Summary of CNN-LSTM based Attention Mechanism

Layer	Output Shape	Param #
Patch_input (InputLayer)	(None, 16, 32, 32, 3)	0
Time_distributed (TimeDistributed)	(None, 16, 64)	19,392
Bidirectional (Bidirectional)	(None, 1, 64)	197,632
Attention_layer (AttentionLayer)	(None, 1, 1)	16,513
Dense_2 (Dense)	(None, 1)	65,792
Dropout (Dropout)	(None, 1)	0
Dense_3 (Dense)	(None, 64, 1)	2,570

The confusion matrix for our model is computed to identify where the algorithm makes errors. Confusion matrix can be visualized using association tables in the form of heatmaps. While there are several built-in methods that can be used to display confusion matrix, visualizing them based on scores serves to improve correlation. Figure 6 shows the confusion matrix for the CNN-LSTM-based attention mechanism algorithm. The results demonstrate that our algorithm performs excellently in detecting rice leaf diseases.

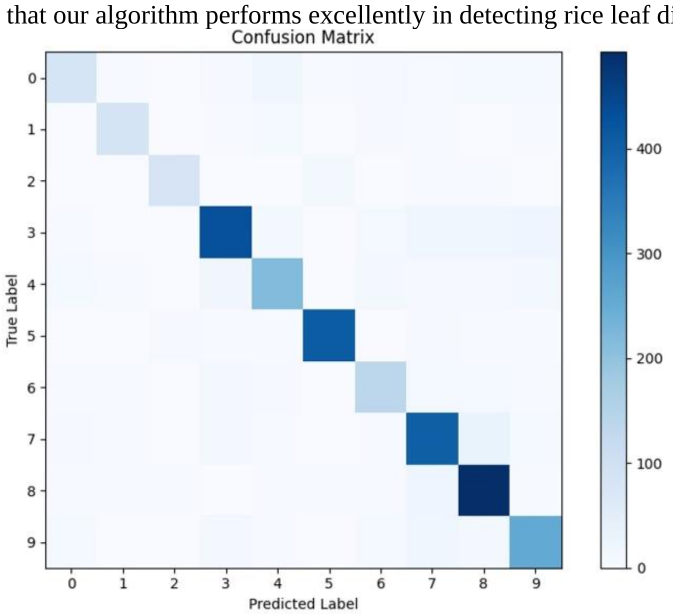

Figure 6. Confusion Metrics of CNN-LSTM based Attention Mechanism Model

Table 2 presents a comprehensive performance comparison between the proposed CNN–LSTM–Attention model and several widely used deep learning architectures, namely MobileNetV2, VGG16, and a Custom 3-layer CNN. The evaluation is conducted using four standard classification metrics: Accuracy, Precision, Recall, and Macro-AUC, which together provide a balanced assessment of overall performance, class-wise discrimination capability, and robustness.

Table 2. Performance comparison with Other Existing Algorithm

Model	Akurasi (%)	Presisi (%)	Recall (%)	AUC Macro
MobileNetV2 [14]	92.4	92.6	92.4	0.982
VGG16 [14]	93.2	93.5	93.1	0.981
Custom CNN (3-layer) [14]	88.8	89.2	88.8	0.960
CNN+LSTM+Attention	95.5	98.12	98.3	0.994

The Custom CNN (3-layer) model exhibits the lowest performance among the compared methods, achieving an accuracy of 88.8%, precision of 89.2%, recall of 88.8%, and a Macro-AUC of 0.960. This result indicates that shallow architectures have limited capacity to capture complex spatial and temporal patterns in the data, leading to suboptimal classification performance.

MobileNetV2 demonstrates improved results with an accuracy of 92.4%, precision of 92.6%, recall of 92.4%, and a Macro-AUC of 0.982. Its lightweight architecture is effective in extracting spatial features; however, its design prioritizes computational efficiency, which may limit its representational power for more complex feature dependencies. Similarly, VGG16 achieves competitive performance with an accuracy of 93.2%, precision of 93.5%, recall of 93.1%, and a Macro-AUC of 0.981. While VGG16 benefits from deeper convolutional layers and richer feature extraction, it lacks mechanisms to explicitly model sequential dependencies or to selectively emphasize informative features, which constrains its overall performance.

In contrast, the proposed CNN–LSTM–Attention model significantly outperforms all baseline methods across all evaluation metrics, achieving an accuracy of 95.5%, precision of 98.12%, recall of 98.3%, and the highest Macro-AUC value of 0.994. The superior performance can be attributed to the synergistic integration of three components: CNN, which effectively captures spatial feature representations; LSTM, which models temporal or sequential dependencies within the extracted features; and Attention mechanism, which dynamically assigns higher weights to the most relevant features, thereby enhancing discriminative capability and reducing the influence of irrelevant information.

The substantial improvement in precision and recall indicates that the proposed model not only reduces false positives but also minimizes false negatives, which is particularly critical in high-stakes classification tasks. Furthermore, the high Macro-AUC value confirms the robustness of the model across all classes, demonstrating strong generalization performance and class balance. Overall, these results validate the effectiveness of the proposed CNN–LSTM–Attention architecture in capturing complex feature relationships and highlight its suitability for deployment in real-world applications requiring high accuracy and reliability.

4. CONCLUSION

This study presents a CNN–LSTM–Attention model for improving classification performance by effectively capturing spatial, temporal, and contextual feature dependencies. The experimental results demonstrate that the proposed model consistently outperforms widely used deep learning architectures, including MobileNetV2, VGG16, and a custom CNN, across all evaluation metrics.

Specifically, the proposed approach achieves superior accuracy, precision, recall, and Macro-AUC, indicating enhanced discriminative capability, robustness, and balanced performance across classes. The integration of the attention mechanism plays a crucial role in emphasizing relevant features while suppressing less informative ones, leading to significant improvements in both precision and recall. This highlights the model’s ability to reduce false positives and false negatives simultaneously, which is essential for reliable real-world deployment.

Overall, the findings confirm that combining convolutional feature extraction with sequential modeling and attention-based feature weighting provides a powerful and effective framework for complex classification tasks. The proposed CNN–LSTM–Attention architecture offers a promising solution for applications requiring high accuracy and strong generalization performance. Future work will focus on expanding the dataset, exploring lightweight model variants for real-time implementation, and evaluating the model’s adaptability to other domains and datasets.

ACKNOWLEDGEMENTS

Funding acquisition, Kemendikbudristek program DRTPM 2025, main research contract number 126/C3/DT.05.00/PL/2025 and derivative contract 0498.21/LL5-INT/AL.04/2025.

.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Conceptualization, R. Marco; methodology, R. Marco, N. Aini, A. H. Muhammad; validation, R. Marco; formal analysis, R. Marco; investigation, R. Marco, and N. Aini; resources, R. Marco, A. H. Muhammad, and Yana H; data curation, R. Marco; writing—original draft preparation, R. Marco, and N. Aini; writing—review and editing, R. Marco, and A. H. Muhammad.

DECLARATION OF COMPETING INTERESTS

The authors declare no conflict of interest.

DATA AVAILABILITY

The dataset used in this project was obtained from an ongoing Kaggle competition, namely the Paddy Disease Classification task (link: https://www.kaggle.com/datasets/imbikramsaha/paddy-doctor) he dataset contains ten classes.

REFERENCES

[1] M. H. Bijoy *et al.*, "Towards Sustainable Agriculture: A Novel Approach for Rice Leaf Disease Detection Using dCNN and Enhanced Dataset" , *IEEE Access* , Vol. 12, No. February , pp. 34174–34191 , 2024.
[2] I. N. Widiarta *et al.*, "Rice virus disease in Indonesia: epidemiology and varietal resistance" , *Phytopathology Research* , Vol. 7, No. 1 , 2025.
[3] W. Shafik, A. Tufail, A. Namoun, L. C. De Silva, and R. A. A. H. M. Apong, "A Systematic Literature Review on Plant Disease Detection: Motivations, Classification Techniques, Datasets, Challenges, and Future Trends" , *IEEE Access* , Vol. 11, No. June , pp. 59174–59203 , 2023.
[4] A. Upadhyay *et al.*, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture" , *Artificial Intelligence Review* , Vol. 58, No. 3 , 2025.
[5] P. I. Ritharson, K. Raimond, X. A. Mary, J. E. Robert, and A. J, "DeepRice: A deep learning and deep feature based classification of Rice leaf disease subtypes" , *Artificial Intelligence in Agriculture* , Vol. 11 , pp. 34–49 , 2024.
[6] M. Tudi *et al.*, "Agriculture Development, Pesticide Application and Its Impact on the Environment" , *Environmental Rsearch and public health* , Vol. 18, No. 1112 , pp. 1–23 , 2021..
[7] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification" , *Environmental Monitoring and Assessment* , Vol. 195, No. 1 , 2023.
[8] M. Tasfe, A. Nivrito, F. Al MacHot, M. Ullah, and H. Ullah, "Deep Learning Based Models for Paddy Disease Identification and Classification: A Systematic Survey" , *IEEE Access* , Vol. 12, No. June , pp. 100862–100891 , 2024.
[9] P. Sobiyaa, K. S. Jayareka, K. Maheshkumar, S. Naveena, and K. S. Rao, "Paddy disease classification using machine learning technique" , *Materials Today: Proceedings* , Vol. 64 , pp. 883–887 , 2022.
[10] A. Stephen, A. Punitha, and A. Chandrasekar, "Designing self attention-based ResNet architecture for rice leaf disease classification" , *Neural Computing and Applications* , Vol. 35, No. 9 , pp. 6737–6751 , 2023.
[11] P. Preethi, R. Swathika, S. Kaliraj, R. Premkumar, and J. Yogapriya, "Deep Learning–Based Enhanced Optimization for Automated Rice Plant Disease Detection and Classification" , *Food and Energy Security* , Vol. 13, No. 5 , pp. 1–10 , 2024.
[12] R. Mukherjee, A. Ghosh, C. Chakraborty, J. N. De, and D. P. Mishra, "Rice leaf disease identification and classification using machine learning techniques: A comprehensive review" , *Engineering Applications of Artificial Intelligence* , Vol. 139, No. PB , p. 109639 , 2025.
[13] R. K. Rachman, D. R. Ignatius, M. Setiadi, A. Susanto, and K. Nugroho, "Enhanced Vision Transformer and Transfer Learning Approach to Improve Rice Disease Recognition" , *Journal of Computing Theories and Applications* , 2024..
[14] B. Kiruba and P. Arjunan, "Paddy Doctor: A Visual Image Dataset for Automated Paddy Disease Clas- sification and Benchmarking" , *6th Joint International Conference on Data Science & Management of Data* , No. January 2023 , 2025.
[15] J. Padhi, L. Korada, A. Dash, P. K. Sethy, S. K. Behera, and A. Nanthaamornphong, "Paddy Leaf Disease Classification Using EfficientNet B4 with Compound Scaling and Swish Activation: A Deep Learning Approach" , *IEEE Access* , Vol. 12, No. September , pp. 126426–126437 , 2024.
[16] A. A. Salamai, N. Ajabnoor, W. E. Khalid, M. M. Ali, and A. A. Murayr, "Lesion-aware visual transformer network for Paddy diseases detection in precision agriculture" , *European Journal of Agronomy* , Vol. 148, No. 126884 , 2023..
[17] P. S. Thakur, S. Chaturvedi, P. Khanna, T. Sheorey, and A. Ojha, "Vision transformer meets convolutional neural network for plant disease classification" , *Ecological Informatics* , Vol. 77, No. 102245 , 2023.
[18] 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.
[19] R. K. Dubey and D. K. Choubey, "An efficient adaptive feature selection with deep learning model-based paddy plant leaf disease classification" , *Multimedia Tools and Applications* , Vol. 83, No. 8 , pp. 22639–22661 , 2024.
[20] D. El-Shahat, A. Tolba, and A. K. Kumar, "An Improved Deep Learning Model for Detecting Rice Diseases" , *Optimization in Agriculture* , Vol. 1 , pp. 1–10 , 2024.
[21] T. H. Nhut, N. P. Hoang Khang, and N. C. Nhan, "An Automatic Diagnosis System of Rice Leaf Diseases Based on ResNet-50 Architecture" , *Proceedings - 2023 International Conference on Advanced Computing and Analytics, ACOMPA 2023* , pp. 54–58 , 2023.
[22] P. S. Thakur, P. Khanna, T. Sheorey, and A. Ojha, "Explainable vision transformer enabled convolutional neural network for plant disease identification: PlantXViT" , *arXiv preprint* , 2022. [Online]. Available: http://arxiv.org/abs/2207.07919
[23] S. M. Hassan and A. K. Maji, "Plant Disease Identification Using a Novel Convolutional Neural Network" , *IEEE Access* , Vol. 10 , pp. 5390–5401 , 2022.
[24] S. K. Upadhyay and A. Kumar, "A novel approach for rice plant diseases classification with deep convolutional neural network" , *International Journal of Information Technology (Singapore)* , Vol. 14, No. 1 , pp. 185–199 , 2022.
[25] M. S. A. M. Al-Gaashani, N. A. Samee, R. Alnashwan, M. Khayyat, and M. S. A. Muthanna, "Using a Resnet50 with a Kernel Attention Mechanism for Rice Disease Diagnosis" , *Life* , Vol. 13, No. 6 , pp. 1–14 , 2023.
[26] A. Kaur, K. Guleria, and N. K. Trivedi, "A deep learning-based model for biotic rice leaf disease detection" , *Multimedia Tools and Applications* , Vol. 83, No. 36 , pp. 83583–83609 , 2024.
[27] T. Berhane, T. Melese, A. Walelign, and A. Mohammed, "A Hybrid Convolutional Neural Network and Support Vector Machine-

Based Credit Card Fraud Detection Model" , *Mathematical Problems in Engineering* , Vol. 2023, No. 1 , 2023.

[28] L. Xiang, P. Wang, X. Yang, A. Hu, and H. Su, "Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism" , *Measurement: Journal of the International Measurement Confederation* , Vol. 175, No. January , p. 109094 , 2021.

[29] J. Li, C. Xu, B. Feng, and H. Zhao, "Credit Risk Prediction Model for Listed Companies Based on CNN-LSTM and Attention Mechanism" , *Electronics (Switzerland)* , Vol. 12, No. 7 , pp. 1–18 , 2023.

[30] D. Karmiani, R. Kazi, A. Nambisan, A. Shah, and V. Kamble, "Comparison of Predictive Algorithms : Backpropagation , SVM , LSTM and Kalman Filter for Stock Market" , *Amity International Conference on Artificial Intelligence (AICAI)* , pp. 228–234 , 2019..

[31] F. Chang *et al.*, "A CNN-LSTM and Attention-Mechanism-Based Resistance Spot Welding Quality Online Detection Method for Automotive Bodies" , *Mathematics* , Vol. 11, No. 22 , 2023.

[32] T. El Idriss, A. Idri, I. Abnane, and Z. Bakkoury, "Predicting Blood Glucose using an LSTM Neural Network" , *Proceedings of the Federated Conference on Computer Science and Information Systems* , Vol. 18 , pp. 35–41 , 2019.

[33] Z. Li, N. Qu, X. Li, J. Zuo, and Y. Yin, "Partial discharge detection of insulated conductors based on CNN - LSTM of attention mechanisms" , *Journal of Power Electronics* , 2021.

.