Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Hybrid Model for Enhanced Detection of Microbial Diseases in Rice Plants Using ResNet50 and Vision LSTM

Rishi Yadav and Dr. Sunil Patil
Rkdf Univeristy, Bhopal,
Madhya Pradesh, India (RISHIYADAV28@Gmail.Com)

ARTICLE INFO

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

In this paper, a deep learning approach is used to design a rice plant disease diagnosis model. In this approach hybridization of ResNet50 is done with vision LSTM (ViL). The high level features are extracted out of ResNet50 and fed into ViL for further classification of rice plant disease classification caused due to microbes. Cascading ResNet50 with a ViL combines the strengths of both architectures to enhance image classification. ResNet50 extracts the spatial features and patterns and then Vision LSTM shows the sequential and spatial relationships between image patches through positional embeddings and LSTM layers. This hybrid approach is designed to preserves spatial information with reduced computational complexity and higher accuracy for such computer vision applications. The result analysis shows that the proposed ResNet50+ViL shows an performance accuracy of 91% and also outperforms better over state-of-the-art methods. This shows that the proposed model is robust and efficient model for rice disease detection.

Keywords: Rice Disease, Microbes, Deep Learning, LSTM, Vision Transformer.

1. INTRODUCTION

Agriculture is a primary income source in many countries. Farmers select their crops and pesticides based on the significance of agriculture to improve plant growth in a limited time. Rice is a staple food in approx. 60% of countries which faces production challenges due to a growing population and decreasing arable land. However, rice plants face substantial challenges due to diseases or environmental factors that affect quality and quantity of rice crops. The lower production rate is attributed to the lack of expert availability in farming fields, insufficient knowledge in fertilizers and lack of awareness about diseases and pests. Such challenges will significantly cause losses in paddy production and affect rice quality [1]. Therefore, it is required to monitor regularly to provide proper care to maintain quality. To adhere this, manual detection is opted by farmers that is quite time-consuming and labour-intensive task. This manual inspection also results in costly method. The conventional visual inspection is error-prone that needs an automated system to identify and alert farmers to paddy diseases early [2]. Traditional methods of rice leaf disease detection are based on manual feature engineering but with advancements in deep neural networks presented a better solution. This will help in automatically and reliably detection of rice leaf diseases that will aid farmers and contributes for agricultural development [3]. An automated solution process images of infected leaves that will enable early disease detection. For diagnosis of diseases in rice plants, it is required to use different image processing and machine learning (ML) algorithms. Deep learning methodologies have shown promising results in image classification and have been used to examine diseases in various crops in homogeneous as well as heterogeneous backgrounds. In recent studies, researchers have significantly used advance learning approaches for detection of rice leaf diseases [4-15]. But still there are several gaps that needs to be focused and resolved to enhance their effectiveness and applicability. Early detection with environmental complexity handling is some of them. The early disease detection is dependent on high-quality and specific-background images. Computational efficiency is another critical gap because image processing is high resource demanding applications. Motivated by this the paper focused on identification of type of microbial effect on rice plant. The paper also presented a hybridization of transfer learning and vision transformer model for differentiation of these microbial diseases.

2. LITERATURE REVIEW

Patil et al. [1] proposed a model termed as “Rice Transformer” that used combined data taken from agricultural sensors as well as images simultaneously for controlling rice disease. Chen et al.[2] introduced a model named as “MobInc-Net”. The model is based on enhanced Inception module and pre-trained MobileNet as backbone extractors. The model uses two-stage transfer learning for efficient training. Patil et al. [3] proposed “Rice-Fusion” model that is a multimodal data fusion framework for rice disease diagnosis. It uses agro-meteorological sensors and a camera to collect data from 3200 samples. The framework extracts numerical and visual features from the data, fused using a concatenation and dense layer. Joshi et al. [4] presented the AI-based model termed as “RiceBioS”, for identification of biotic stress especially for rice plants. It classifies images into healthy and stressed categories, diagnoses fungal and bacterial infections, and uses an automated RoI detection workflow. Yang et al. [5] proposed the “Fully Connected Bottleneck Transformer (FCBT)” model with Yolov8 model and termed as “FCBTYOLO”. Bharanidharan et al. [6] developed “modified lemurs optimization algorithm” with incorporation with “sine cosine optimization”. The model was tested on five paddy diseases and analyzes 636 thermal images. Alshahrani et al.[7] used the “Quantum Inspired Moth Flame Optimizer” for rice disease detection. Joseph et al.[8] developed disease datasets for rice, wheat, and maize crops were build and investigated over Xception and MobileNet. Altabaji et al. [9] proposed deep learning model such as “LeafNet” for classification of rice plant diseases. Then model was tested on 2658 images. Bijoy et al. [10] proposed a “deep Convolutional Neural Network” for rice leaf disease detection that was designed for crop health monitoring system. Haridasan et al.[11] proposed a computer vision-based system for accurately detecting and classifying diseases from photographs of rice plants. This system uses image processing, machine learning, and deep learning techniques to identify and classify diseases. Pan et al.[12] proposed RiceNet is a two-stage method to identify four rice diseases, using YoloX for detection and Siamese Network for identification. Rajpoot et al.[13] presents an advanced detection method for rice diseases, including bacterial leaf blight, brown spot, and leaf smut. Using VGG-16 transfer learning and random forest, the method extracts features and categorizes them. Daniya et al. [14] proposed a hybrid model by cascading “Wave-based neural network with Rider Optimization algorithm and Water wave optimization”. In this approach, histogram equalization was used as pre-processing step. Shovon et al. [15] proposes PlantDet that is an ensemble model using InceptionResNetV2, EfficientNetV2L, and Xception. PlantDet outperforms previous models in accuracy, precision, recall, F1 and specificity for the Rice Leaf and Betel Leaf datasets. The model also outperforms existing base models, including Grad-CAM and Score-CAM, with Score-CAM slightly outperforming Grad-CAM++. Sudhesh et al. [16] used transfer-learned deep learning models to identify rice leaf diseases, identifying four categories: bacterial blight, blast, brown spot, and tungro. The DenseNet121 deep feature with Random Forest classifier outperforms other models, while the Dynamic Mode Decomposition-based attention-driven pre-processing model achieves 100% test accuracy. Moupojou et al.[17] used the “FieldPlant” dataset of 5,170 images. In this approach manual annotation was performed for individual leaves with 27 disease classes. Singh et al. [18] designed a model with CNN architecture for detection of four types of rice plant diseases. Sankareshwaran et al.[19] proposed a model termed as “crossover boosted artificial hummingbird algorithm based AX-RetinaNet (CAHA-AXRNet)”. This approach was designed for classification of healthy or unhealthy rice plants. Dorga et al.[20] presented a model based on CNN and VGG19 for detection of brown spot. Tholkapiyan et al.[21] presented an automatic diagnosis model using machine learning with meta-heuristic optimization. Akyol et al.[22] presented a rice leaf disease detection using CNN model with random forest classifier. Ahmad et al.[23] presented a comparative study of for detection and diagnosis of plant diseases.

3. METHODOLOGY USED

3.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) are one of the type of recurrent neural network (RNN) that was designed to overcome the vanishing gradient problems. LSTMs achieve this by selectively "remembering" relevant information and "forgetting" irrelevant data using different gates and maintaining an internal cell state [24]. Core Components of LSTM are:

- Memory Cell: It acts as the long-term memory of the LSTM, retaining information across time steps.
- Forget Gate (f): Decides what fraction of the previous state to retain or forget. An activation output of 1 means "remember everything," and 0 means "forget everything."

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- Input Gate (i): Determines which new information should be added to the cell state.
- Input Modulation Gate (g): Modulates the information to be added to the cell state, often introducing non-linearity and ensuring faster convergence.
- Output Gate (o): Controls the output of the LSTM unit based on the cell state and decides the hidden state for the next time step.

At each time step, the cell state is updated using the input gate, forget gate, and a candidate cell state generated by the input modulation gate as:

$$c_t = \sigma(W_i[x_t; a_{t-1}] + b_i) \odot \tanh(W_c[x_t; a_{t-1}] + b_c) + \sigma\left(W_f[x_t; a_{t-1}] + b_f\right) \odot c_{t-1} \quad (1)$$

The hidden state is calculated using the output gate and the updated cell state as:

$$a_t = \sigma(W_o[x_t; a_{t-1}] + b_o) \odot \tanh c_t \quad (2)$$

The output at each time step can be derived from the hidden state as:

$$\hat{y}_t = \emptyset_2(W_y a_t + b_y) \quad (3)$$

3.2 Vision Transformer (ViT)

The Vision Transformer (ViT) brings the transformer architecture, originally developed for natural language processing (NLP), to image recognition tasks [25]. By treating an image as a sequence of patches, ViT leverages self-attention mechanisms to process images differently from traditional convolutional neural networks (CNNs). Its core components are:

- Patch Extraction: An image is divided into fixed-size patches of 1D shape.
- Position Embeddings: Added to patch embeddings to retain spatial information.
- Transformer Encoder: Comprising layers of self-attention and feedforward neural networks, capturing global dependencies and interactions.
- Classification Token (CLS): Aggregates information from all patches and is used for the final classification.

3.3 Vision-LSTM (ViL)

Language modelling architectures like Transformers and State Space Models (SSMs) have been adapted for computer vision. The Vision Transformer (ViT) groups images into patches and processes them with language modelling techniques. The Extended Long Short-Term Memory (xLSTM) family has shown promise in language modelling, leading to the introduction of Vision LSTM (ViL) for computer vision [26]. ViL uses alternating mLSTM blocks to efficiently handle non-sequential image inputs with linear computational complexity, making it ideal for high-resolution tasks like medical imaging and segmentation. In contrast, ViT's quadratic complexity makes it less suitable for such tasks. Vision-LSTM (ViL) is designed with cascading of xLSTM blocks. As compared to ViT, ViL divides the image into non-overlapping patches by applying linear projection that adds learnable positional embeddings to each patch token. In ViL, the odd mLSTM blocks process the patch tokens from top left to bottom right whereas the even blocks are processed from bottom right to top left. The entire architecture of ViL composed of image tokenisation, positional embedding, alternating blocks, classification. In forst step, the image is divided into non-overlapping patches of pre-defined size. Entire, patch is linearly projected into a sequence and passed for positional embedding. It is added to each token for retention of spatial information. As the core architecture of ViL is composed on multiple xLSTM blocks that are arranged alternately. In this odd-numbered blocks can process tokens row-wise from top to bottom whereas the even-numbered blocks can process tokens from bottom to top. Finally for classification, output tokens from xLSTM blocks are pooled together for prediction of the final output.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

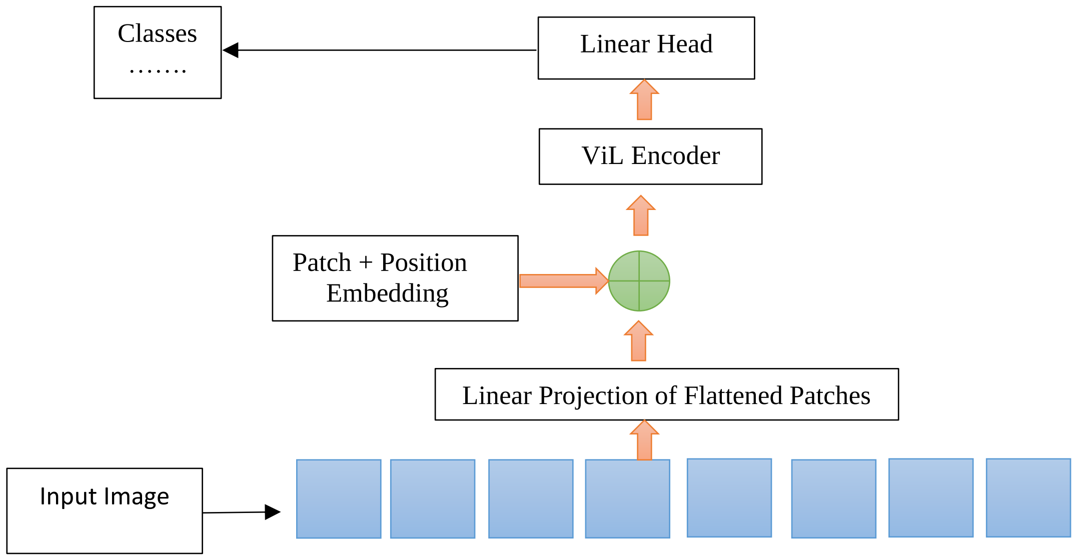

Fig 1. Architecture of ViL

3.4 Proposed Methodology

The paper proposes a three-layered model for detecting diseases in rice plants that are caused by microbes such as bacterial, viral and fungal. The entire model includes steps, pre-processing, feature extraction, and classification. The entire model is presented in fig 2. The images are pre-processed using a digital filter for enhancement of their quality. Then these pre-processed enhanced images are passed to pre-trained models such as ResNet50 for feature extraction. Finally, ViL is implemented to extract patches from high-level features extracted out by ResNet50 for accurate detection of bacterial, viral, and fungal diseases in rice plants.

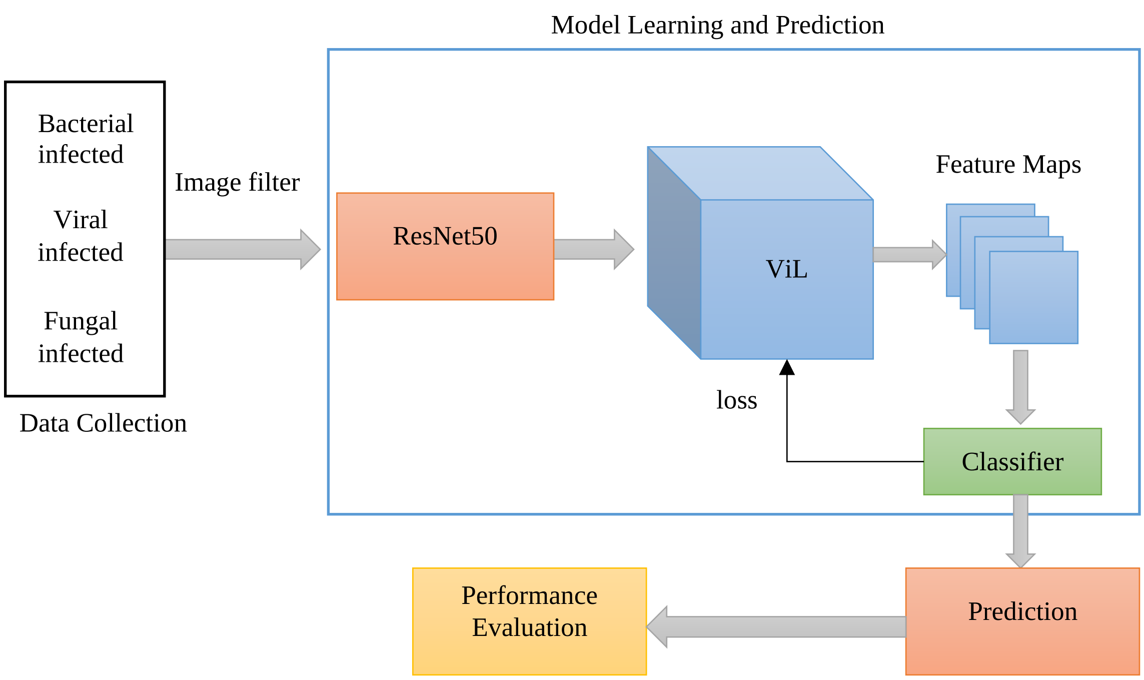

Fig 2. Proposed Architecture

Data Collection: In this step, data are collected from a publicly available resource referenced as source [27]. The dataset identifies 13 key rice diseases, which are categorized into three types: fungal, bacterial, and viral that affects different parts of the rice plants.

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Pre-Processing: The entire input rice disease images are resized in size 224×224×3. This step utilizes an adaptive bilateral filter with spatial adaptation for noise removal that will preservs the edge and texture characteristics of input images. Unlike the conventional bilateral filter, the presented filter will combines domain and range kernels to preserve edge and texture information. Mathematically, it is formulated as:

$$I_{filt}(n) = \frac{1}{N_f} \sum_{m \epsilon P} I(m) \cdot f(||n - m||) \cdot g(|I(n) - I(m)|) \tag{4}$$

Where, filtered image is considered as I_{filt} for n pixels. The neighboring pixel's (m) intensity is represented as $I(m)$ within spatial domain P with n pixels. Spatial kernel is represented as $f(||n - m||)$ that reduces the kernel distance among m and n. The range kernel is represented as $g(|I(n) - I(m)|)$.

But in the adaptive bilateral filter, the paper used the sliding window approach to identify the local adaptation features and thus combining local spatial features to generate global feature for noise removal. Mathematically, in sliding window s_p might be described as:

$$I_{local}(n) = \frac{1}{N_{flocal}} \sum_{m \epsilon P_{local}} I(m) \cdot f_{local}(||n - m||) \cdot g_{local}(|I(n) - I(m)|) \tag{5}$$

Where, filtered image in each sliding window output is considered as$I_{local}(n)$ for n pixels. The neighboring pixel's (m) intensity is represented as $I(m)$ within local spatial domain P_{local} with n pixels. Local spatial kernel is represented as $f_{local}(||n - m||)$ that reduces the kernel distance among m and n. The local range kernel is represented as $g_{local}(|I(n) - I(m)|)$. By combining all these local filtration parameters, global parameters are identified to filter out the image.

Model Learning and Prediction: For feature extraction and model learning and prediction, the proposed approach used the Vision LSTM model that will combine a pre-trained ResNet50 network with ViL for enhanced feature extraction and image classification. The ResNet50 as base model extracts high-level feature maps from input images. These feature maps are globally averaged and reshaped into patches that linearly projected into a higher-dimensional space. Given an image I of size $H \times W \times C$, it is divided into patches of size $P \times P$. The number of patches N is given by:

$$N = \frac{H \times W}{P^2} \tag{6}$$

Each patch is flattened and then projected linearly using a learnable matrix W as:

$$E_i = Linear\ (Flatten(I_i)) \tag{7}$$

Where, I_i is the i_{th} patch and E_i is its embedding. Then, positional informations are extracted by applying positional patch embeddings as:

$$Z_i = E_i + P_i \tag{8}$$

Where P_i is the positional embedding for the i_{th} patch. The ViL encoder processes the sequence of patch embeddings. It consists of multiple mLSTM blocks $X^{(l)}$. Each mLSTM block includes normalization, mLSTM layers, activation functions (SiLU), and skip connections. The output is evaluated with last mLSTM block whose data is processed by linear head for classification of disease type.

4. RESULTS AND DISCUSSION

This section outlines the implementation details, result analysis, and comparisons with state-of-the-art models for proposed model. The model was implemented using the Tesla P100-PCIE GPU on Google Colab, utilizing Keras and TensorFlow as the backend frameworks. For performance evaluation, the model was assessed based on accuracy, precision, recall, and F1-score. These metrics are critical for determining the effectiveness of the model in classifying and predicting accurately and are defined as:

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{9}$$

$$Precision = \frac{(TP)}{(TP + FP)} \tag{10}$$

$$Recall = \frac{(TP)}{(TP + FN)} \tag{11}$$

$$F1 - Score = \frac{2 * Precision * Recall}{(Precision + Recall)} \tag{12}$$

1.1 Dataset Description

In this paper, we have used rice plant diseases taken from source [27]. The dataset consists of images of 224 x 224 pixels for efficient model training for rice disease prediction in Philippines. Here fungal, bacterial, and viral agents cause diseases pose significant threats to production due to the climate's high humidity and frequent rainfall. These conditions exacerbate the spread and impact of diseases, which can severely reduce crop yield and quality. The dataset identified 13 key rice diseases divided into three categories—fungal, bacterial, and viral—that affect different parts of the plant.

1.2 Result Analysis

The fig 3 shows the accuracy graph for training duration over 100 epochs. Initial accuracy was from 65% that increase and stabilize up to 90% after 20th epoch for training and the 10th epoch for validation. Further the training accuracy continues to rise gradually up to 95% with minor fluctuations. The fig 4 shows the model loss over 100 epochs for both training and validation. Initially, losses are high approx. 0.8 and 0.7 respectively but after 10 epochs it decreases sharply that shows improvement in model performance. The ROC curves is presented in fig 5 for the multi-class classification model that show excellent performance in classification and differentiation among bacterial, fungal, and viral classes. Each class has a high area under the curve (AUC) that is approx. 98% and 99%.

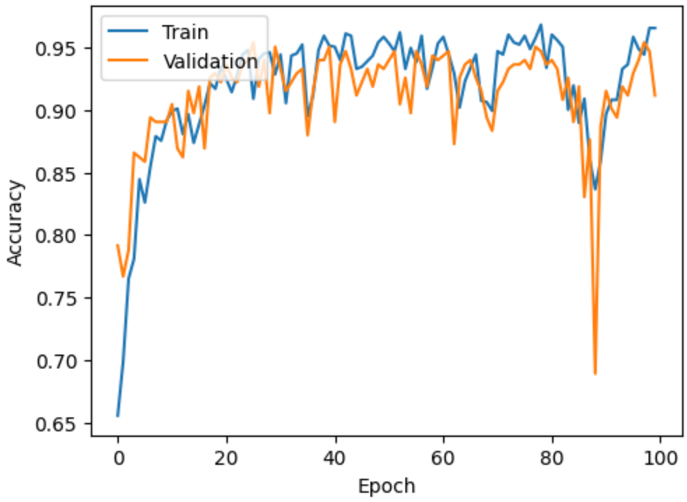

Fig 3. Training Accuracy and Validation for Rice Plant Disease Detection Caused by Microbes

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

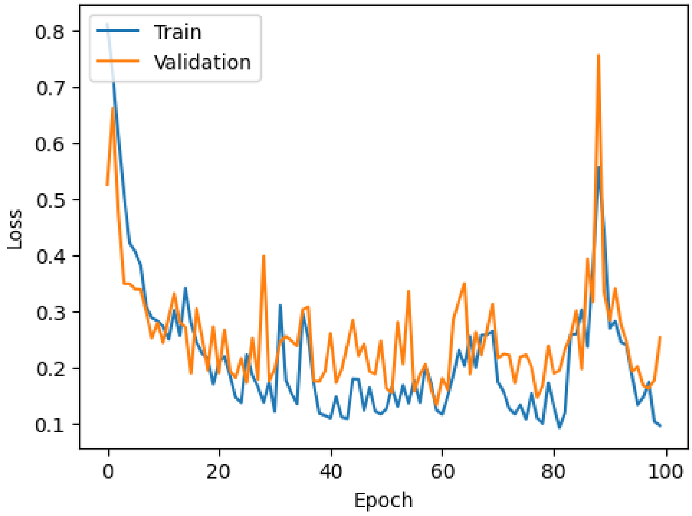

Fig 4. Training Loss and Validation for Rice Plant Disease Detection Caused by Microbes

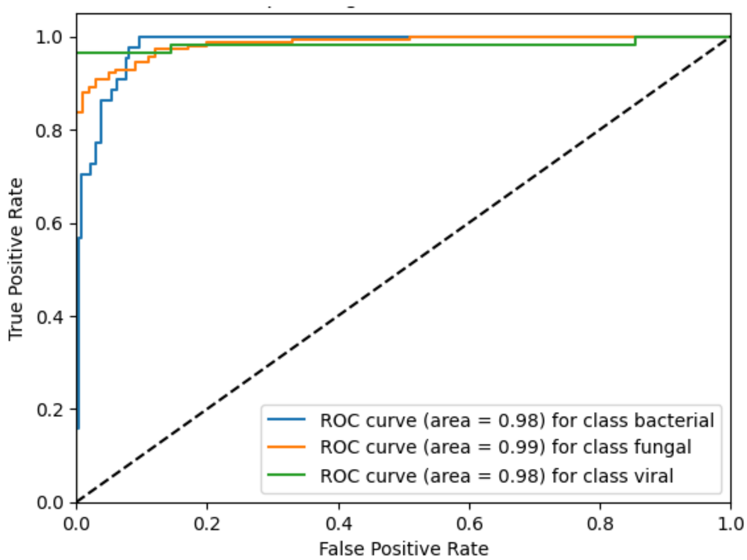

Fig 5. ROC for Rice Plant Disease Detection Caused by Microbes

Table 1. Performance Analysis

Disease Type	Model Used	Accuracy	Precision	Recall	F1-Score
Bacterial	ResNet50	90%	74	80	77
Fungal			92	92	93
Viral			97	86	91
Bacterial	ResNet50+ViL	91%	72	89	80
Fungal			95	95	95
Viral			100	82	90

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The performance analysis is presented in table 1 that compares the results of ResNet50 and ResNet50+ViL for classification of bacterial, fungal, and viral rice plant diseases. The addition of Vision LSTM (ViL) to ResNet50 generally improves the performance metrics for fungal and viral diseases. For viral diseases there is a slight decrease in precision and recall. The bacterial disease classification shows that the accuracy and recall is improved with the addition of ViL whereas the precision slightly decreases. The improvement in accuracy and F1-scores suggests that the proposed ResNet50 + ViL model is more robust and efficient in disease detection. Fig 6 shows the accuracy comparison for ResNet50 and ResNet50+ViL models. The accuracy of the ResNet50 model is 90% whereas the ResNet50 + ViL model achieves an accuracy of 91%. Fig 7 shows the precision comparison for ResNet50 and ResNet50+ViL. The precision of the ResNet50 model is 87% whereas the ResNet50 + ViL model achieves a precision of 89%. Fig 8 shows the recall comparison for ResNet50 and ResNet50+ViL models. The recall of the ResNet50 model is 86% whereas the ResNet50 + ViL model achieves an recall of 88%. Fig 9 shows the f1-score comparison for ResNet50 and ResNet50+ViL models. The f1-score of the ResNet50 model is 87% whereas the ResNet50 + ViL model achieves an f1-score of 88%. The improvement suggests that the combined model can better capture and utilize spatial and sequential features that leads to more accurate classifications.

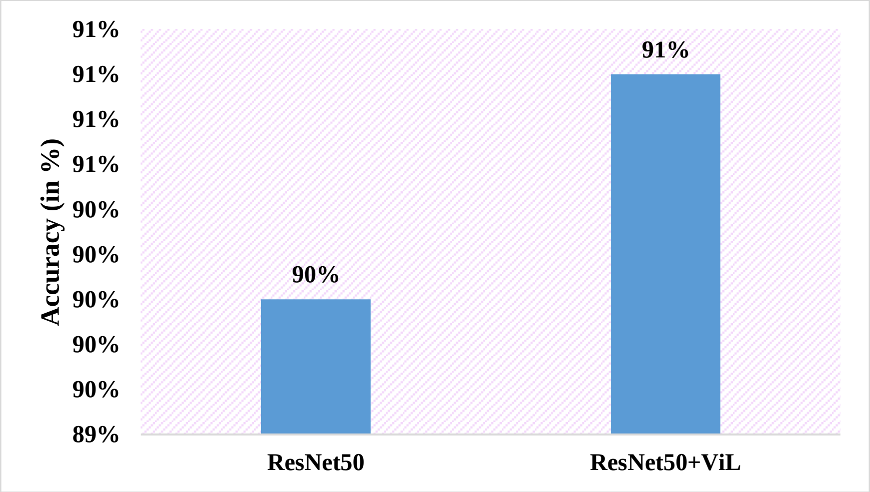

Fig 6. Accuracy Performance of Proposed Model for Rice Plant Disease Detection Caused by Microbes

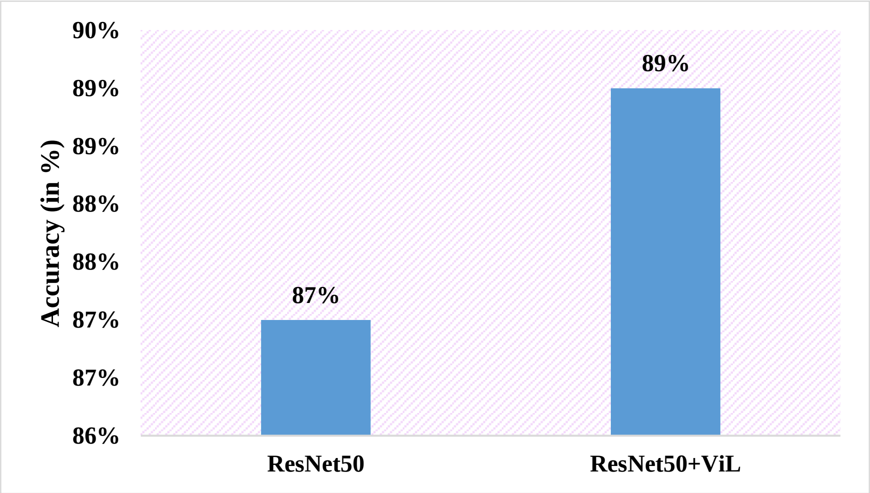

Fig 7. Accuracy Performance of Proposed Model for Rice Plant Disease Detection Caused by Microbes

Journal of Information Systems Engineering and Management
2025, 10(36s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

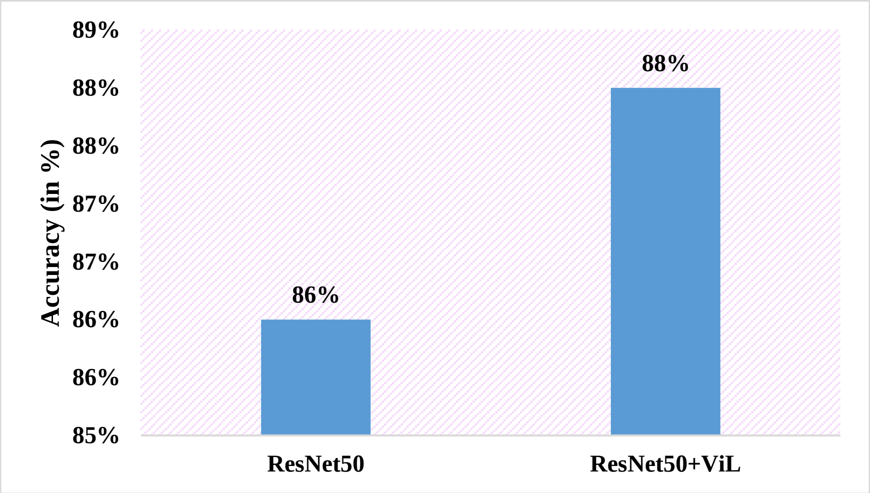

Fig 8. Accuracy Performance of Proposed Model for Rice Plant Disease Detection Caused by Microbes

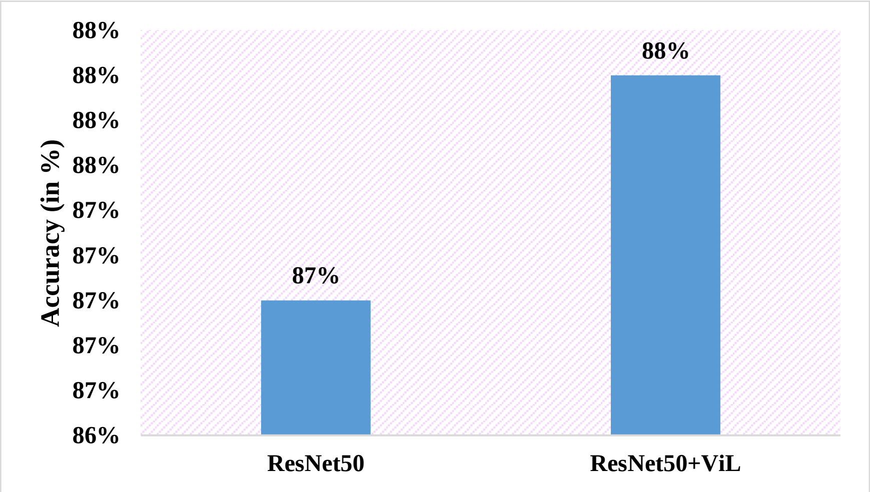

Fig 9. Accuracy Performance of Proposed Model for Rice Plant Disease Detection Caused by Microbes

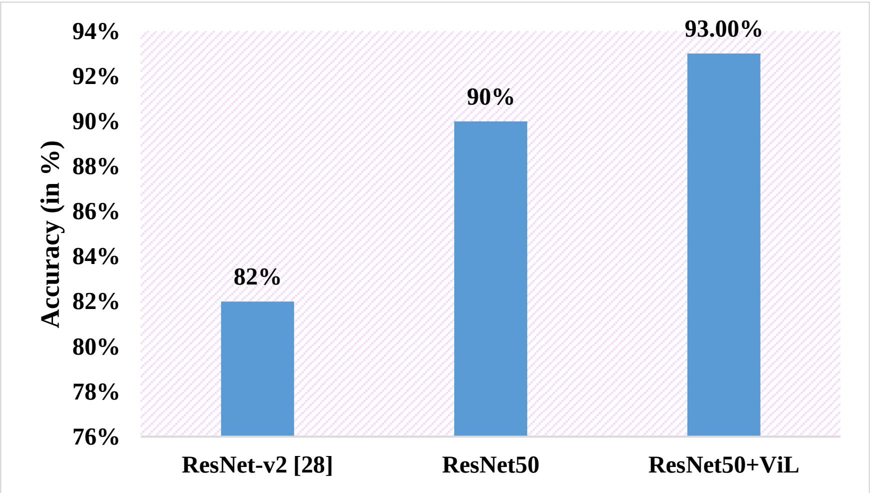

Fig 10. Comparative State-of-Art

Fig 10 presents the comparative analysis of the presented model with existing model. ResNet-v2 [28] have achieved 82% accuracy whereas the proposed learning model with ResNet50 have achieved an accuracy of 90%. The proposed hybrid approach ResNet50+ViL have achieved an accuracy of 91%. This is because the proposed model used the pre-trained model with fine-tuning and therefore achieved better detection of type of microbial infections in rice plants.

5. CONCLUSION

The paper presented a three-layered model for the diseases caused by bacterial, viral, and fungal microbes in rice plants. By employing an adaptive bilateral filter during pre-processing enhanced the quality of input images that will preserves the edge and texture information. The integration of the ResNet50 model for feature extraction and the Vision LSTM (ViL) network for classification enabled the model to effectively capture spatial and sequential features that presents the accurate disease detection. The experimental results validated on publicly available dataset that shows the significant improvement as compared to existing methods. The ResNet50+ViL model achieved an accuracy of 91%. The result shows that the proposed approach provides a robust and efficient framework that will detect the diseases earlier with minimal losses. Future work will focus on further optimizing the model and exploring its applicability to other crops and plant diseases for agricultural disease management.

REFERENCES

[1] R. R. Patil and S. Kumar, "Rice Transformer: A Novel Integrated Management System for Controlling Rice Diseases," *IEEE Access*, vol. 10, pp. 87698–87714, 2022, doi: 10.1109/ACCESS.2022.3200688.
[2] J. Chen, W. Chen, A. Zeb, S. Yang, and D. Zhang, "Lightweight Inception Networks for the Recognition and Detection of Rice Plant Diseases," IEEE Sens. J., vol. 22, no. 14, pp. 14628–14638, 2022, doi: 10.1109/JSEN.2022.3182304.
[3] R. R. Patil and S. Kumar, "Rice-Fusion: A Multimodality Data Fusion Framework for Rice Disease Diagnosis," IEEE Access, vol. 10, pp. 5207–5222, 2022, doi: 10.1109/ACCESS.2022.3140815.
[4] P. Joshi, D. Das, V. Udutalapally, M. K. Pradhan, and S. Misra, "RiceBioS: Identification of Biotic Stress in Rice Crops Using Edge-as-a-Service," IEEE Sens. J., vol. 22, no. 5, pp. 4616–4624, 2022, doi: 10.1109/JSEN.2022.3143950.
[5] Y. Yang, Y. Xiao, Z. Chen, D. Tang, Z. Li, and Z. Li, "FCBTYOLO: A Lightweight and High-Performance Fine Grain Detection Strategy for Rice Pests," IEEE Access, vol. 11, pp. 101286–101295, 2023, doi: 10.1109/ACCESS.2023.3314697.
[6] N. Bharanidharan, S. R. S. Chakravarthy, H. Rajaguru, V. V. Kumar, T. R. Mahesh, and S. Guluwadi, "Multiclass Paddy Disease Detection Using Filter-Based Feature Transformation Technique," IEEE Access, vol. 11, pp. 109477–109487, 2023, doi: 10.1109/ACCESS.2023.3322587.
[7] H. M. Alshahrani, M. K. Saeed, S. S. Alotaibi, A. Mohamed, M. Assiri, and S. S. Ibrahim, "Quantum-Inspired Moth Flame Optimizer Enhanced Deep Learning for Automated Rice Variety Classification," IEEE Access, vol. 11, pp. 125593–125600, 2023, doi: 10.1109/ACCESS.2023.3330918.
[8] D. S. Joseph, P. M. Pawar, and K. Chakradeo, "Real-Time Plant Disease Dataset Development and Detection of Plant Disease Using Deep Learning," IEEE Access, vol. 12, pp. 16310–16333, 2024, doi: 10.1109/ACCESS.2024.3358333.
[9] W. I. A. E. Altabaji, M. Umair, W.-H. Tan, Y.-L. Foo, and C.-P. Ooi, "Comparative Analysis of Transfer Learning, LeafNet, and Modified LeafNet Models for Accurate Rice Leaf Diseases Classification," IEEE Access, vol. 12, pp. 36622–36635, 2024, doi: 10.1109/ACCESS.2024.3373000.
[10] 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, pp. 34174–34191, 2024, doi: 10.1109/ACCESS.2024.3371511.
[11] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," Environ. Monit. Assess., vol. 195, no. 1, p. 120, 2022, doi: 10.1007/s10661-022-10656-x.
[12] J. Pan, T. Wang, and Q. Wu, "RiceNet: A two stage machine learning method for rice disease identification," Biosyst. Eng., vol. 225, pp. 25–40, 2023, doi: https://doi.org/10.1016/j.biosystemseng.2022.11.007.

[13] V. Rajpoot, A. Tiwari, and A. S. Jalal, "Automatic early detection of rice leaf diseases using hybrid deep learning and machine learning methods," Multimed. Tools Appl., vol. 82, no. 23, pp. 36091–36117, 2023, doi: 10.1007/s11042-023-14969-y.

[14] T. Daniya and S. Vigneshwari, "Rider Water Wave-enabled deep learning for disease detection in rice plant," Adv. Eng. Softw., vol. 182, p. 103472, 2023, doi: https://doi.org/10.1016/j.advengsoft.2023.103472.

[15] 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, pp. 34846–34859, 2023, doi: 10.1109/ACCESS.2023.3264835.

[16] S. K.M., S. V., S. K. P., and S. O.K., "AI based rice leaf disease identification enhanced by Dynamic Mode Decomposition," Eng. Appl. Artif. Intell., vol. 120, p. 105836, 2023, doi: https://doi.org/10.1016/j.engappai.2023.105836.

[17] E. Moupojou et al., "FieldPlant: A Dataset of Field Plant Images for Plant Disease Detection and Classification With Deep Learning," IEEE Access, vol. 11, no. March, pp. 35398–35410, 2023, doi: 10.1109/ACCESS.2023.3263042.

[18] S. P. Singh, K. Pritamdas, K. J. Devi, and S. D. Devi, "Custom Convolutional Neural Network for Detection and Classification of Rice Plant Diseases," Procedia Comput. Sci., vol. 218, pp. 2026–2040, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.179.

[19] S. P. Sankareshwaran, G. Jayaraman, P. Muthukumar, and A. Krishnan, "Optimizing rice plant disease detection with crossover boosted artificial hummingbird algorithm based AX-RetinaNet," Environ. Monit. Assess., vol. 195, no. 9, p. 1070, 2023, doi: 10.1007/s10661-023-11612-z.

[20] R. Dogra, S. Rani, A. Singh, M. A. Albahar, A. E. Barrera, and A. Alkhayyat, "Deep learning model for detection of brown spot rice leaf disease with smart agriculture," Comput. Electr. Eng., vol. 109, p. 108659, 2023, doi: https://doi.org/10.1016/j.compeleceng.2023.108659.

[21] M. Tholkapiyan, B. Aruna Devi, D. Bhatt, E. Saravana Kumar, S. Kirubakaran, and R. Kumar, "Performance Analysis of Rice Plant Diseases Identification and Classification Methodology," Wirel. Pers. Commun., vol. 130, no. 2, pp. 1317–1341, 2023, doi: 10.1007/s11277-023-103333.

[22] K. Akyol, "Handling hypercolumn deep features in machine learning for rice leaf disease classification," Multimed. Tools Appl., vol. 82, no. 13, pp. 19503–19520, 2023, doi: 10.1007/s11042-022-14318-5.

[23] A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," Smart Agric. Technol., vol. 3, p. 100083, 2023, doi: https://doi.org/10.1016/j.atech.2022.100083.

[24] S. Ashraf Zargar, "Introduction to Sequence Learning Models: RNN, LSTM, GRU," no. April, 2021, doi: 10.13140/RG.2.2.36370.99522.

[25] K. Han et al., "A Survey on Vision Transformer," IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 1, pp. 87–110, 2023, doi: 10.1109/TPAMI.2022.3152247.

[26] Alkin, Benedikt, et al. "Vision-LSTM: xLSTM as Generic Vision Backbone." arXiv preprint arXiv:2406.04303 (2024).

[27] https://www.kaggle.com/datasets/shrupyag001/philippines-rice-diseases

[28] Thangaraj, P., et al. "Classification of rice leaf diseases using deep learning." AIP Conference Proceedings. Vol. 2764. No. 1. AIP Publishing, 2023