Attention-Driven Deep Learning Framework for Multi-Class Rice Leaf Disease Classification

Richa Sharma[1]*, Jimmy Singla[2]

[1] School of Computer Science and Engineering Lovely Professional University, Punjab, India Email: Richa17lpu@gmail.com
[2] School of Computer Science and Engineering Lovely Professional University, Punjab, India Email: jimmy.21733@lpu.co.in

Abstract
Crop productivity can be significantly impacted by rice leaf diseases, and prompt and precise disease detection is essential for efficient crop management. Using a combination of main and secondary image data sets, this work attempts to offer an attention-based deep learning model for multi-class rice leaf disease classification. 942 photos from five classes—bacterial leaf blight, fake smut, healthy, leaf blast, and sheath blight—were included in the final dataset after cleaning. A stratified train-validation-test split was created after duplicate, cross-labeled, and inconsistent unlabelled images were eliminated. The suggested AttentionMobileNetV2-SE model combines a lightweight backbone, MobileNetV2, with a Squeeze-and-Excitation attention module to improve the representation of disease-relevant data. Class-weighted cross-entropy loss was employed to lessen the effects of class imbalance in training. The model yielded a weighted F1-score of 0.8656, validation accuracy of 90.07%, and test accuracy of 86.62%. By emphasising disease-relevant picture regions, Grad-CAM visualisation was also utilised to interpret model predictions. The outcomes show that the suggested framework can successfully differentiate between several rice leaf disease categories and that the model's judgements are visually interpretable.

Keywords - Rice leaf disease classification; deep learning; MobileNetV2; Squeeze-and-Excitation attention; transfer learning; Grad-CAM; plant disease detection.

1. Introduction

Rice is a key staple crop globally and significantly contributes to nutrition, food security, and sustainable agriculture. It serves as a primary food source for a large proportion of the global population, making its stable production essential for meeting future food demand.[1,2] However, rice production continues to be affected by several biotic stresses, including plant pathogens and pests, which can cause substantial crop losses and threaten agricultural productivity.[3] Among rice diseases, rice blast is widely recognized as one of the most destructive diseases and remains a major challenge for rice cultivation and disease management.[4]

It is essential to detect the disease in time and accurately to minimize the crop loss and to help in effective disease management. Traditionally, rice disease detection is done manually by farmers or agricultural experts by visual inspection. Manual inspection can be helpful, but is subjective, time-consuming and requires experts to be available. The field environment, growth stage, lighting and disease symptoms may also vary, making visual diagnosis difficult. Imaging-based plant disease detection has thus become a focus of attention for precision agriculture and plant phenotyping as it can enable more objective, scalable and timely disease assessment.[5]

Recent developments in computer vision and deep learning have opened up new opportunities for using leaf images to classify plant diseases. The discriminative visual features can be learned directly from images with deep learning models, which can alleviate the need of hand-crafted feature extraction. But the rice disease classification is still difficult because of the similarity of disease symptoms among classes, class imbalance in the dataset and the variability of the image quality, background and illumination in the field. The difficulties emphasize the need for models that can learn visual patterns relevant to diseases and be accurate.

Attention mechanisms can be used to enhance deep learning models to highlight more informative feature regions or channels. In this study, the Squeeze-and-Excitation attention mechanism is combined with the MobileNetV2-based architecture to improve the channel-wise feature representation.[6] Besides the classification performance, explainability is also crucial in the development of trust in deep learning-based agricultural systems. Hence, Grad-CAM is applied to visualize the regions of the image that contribute to the model prediction, and to interpret the correct and misclassified cases.[7]

The other significant driving force of this study is the employment of both primary image data as well as secondary image data. Secondary photos are publicly available photographs that contribute to expanding the dataset's size and class coverage, whereas primary images are field-level images that provide the image with practical relevance. A more diverse data collection for the multi-class categorization of rice leaf diseases can be created by combining the two sources. It is a multi-class classification challenge for rice leaf illnesses rather than a task for identifying fungal diseases because the selected classes include bacterial leaf blight, fake smut, healthy, leaf blast, and sheath blight.

The following are the main contributions of this study. First, a rice disease image data set is prepared through an image source from primary and secondary sources and cleaned and standardized. Secondly, duplicate and cross-labeled conflicting images are eliminated to enhance the reliability of the datasets. Third, an AttentionMobileNetV2-SE model is constructed, which combines a Squeeze-and-Excitation attention module with a lightweight MobileNetV2 backbone. Fourth, class weighted training method is employed to overcome the effect of class imbalance. Finally, Grad-CAM visualization is applied to interpret model predictions and provide explainability support for the proposed classification framework. The specific research objectives of this study are:

- To develop a quality-controlled multi-source rice leaf disease dataset by integrating primary field-level images with secondary public images and applying systematic label standardization, duplicate removal, and conflict filtering.
- To design an attention-enhanced transfer learning framework by incorporating a Squeeze-and-Excitation module into a MobileNetV2 backbone for efficient multi-class rice leaf disease classification.
- To assess the proposed framework through independent test-set evaluation and visual explainability analysis using classification metrics, confusion matrix-based error analysis, and Grad-CAM visualization.

2. Literature Review

Owing to deep learning's capability to autonomously identify visual patterns associated with diseases from image data, it has been extensively utilized for automatic detection of plant diseases. Mohanty et al. proved that deep learning can be used for image-based plant disease detection with leaf images, and that convolutional neural networks can be used to help in the automated disease recognition for various types of plants.[8] Kamilaris and Prenafeta-Boldú reviewed the broader role of deep learning in agriculture and highlighted its increasing use in agricultural image analysis, crop monitoring, and disease detection tasks.[9] Similarly, Ferentinos evaluated deep learning models for plant disease detection and diagnosis and showed that CNN-based models can achieve strong classification performance when trained on suitable image datasets.[10]

Some studies have particularly focused on plant and leaf disease classification using CNN. Sladojevic et al. suggested an approach to plant disease recognition using deep neural network from the images of leaves and demonstrated that CNNs can learn the disease symptoms from the images.[11] Brahimi et al. applied deep learning to tomato disease classification and also included symptom visualization, indicating that CNN-based models can support both classification and interpretation of plant disease symptoms.[12] In rice disease classification, Lu et al. used deep convolutional neural networks for identifying rice diseases from images, showing the suitability of CNN-based methods for rice disease recognition.[13] These studies provide a strong basis for using deep learning in rice leaf disease classification.

In image classification of agricultural applications, transfer learning is also gaining significance as it can be used to adapt pretrained models to plant disease specific datasets. To et al. conducted a comparative study of various fine-tuned deep learning models for plant disease identification and found that pretrained networks can enhance the classification performance in plant disease recognition tasks.[14] Efficient architectures are especially useful in agricultural applications where future field-level or mobile deployment may be required. MobileNetV2 introduced inverted residuals and linear bottlenecks to support efficient image classification with reduced computational complexity.[15] Efficient Net further demonstrated the importance of balanced model scaling for improving classification performance efficiently.[16] Based on these studies, the lightweight and transfer learning-based architectures are selected for rice disease classification.

There are still many issues in the field of plant disease classification, despite the positive findings of earlier research. Primary field-level photographs are frequently hard to get by, and many models rely on publicly accessible datasets. Less reliability may also result from class imbalance, duplicate samples, visual resemblance of the clinical symptoms, and model interpretability. Prior studies have demonstrated the efficacy of CNNs, transfer learning, and effective designs in the identification of plant diseases.

To overcome these limitations, the present study proposes a framework called AttentionMobileNetV2-SE for multi-class rice leaf disease classification. The study relies on a harmonized dataset of primary and secondary rice disease images, filters out duplicate and conflicting data, implements class-weighted training to mitigate imbalance effects, and employs visualization techniques with Grad-CAM to enhance interpretability of model predictions.

3. Materials and Methods

3.1 Dataset Description and Class Selection

Primary and secondary image sources were combined to create the rice disease picture dataset used in this investigation. The publicly accessible rice leaf bacterial and fungal disease image data set served as the secondary data set, while rice field photos gathered for the study served as the primary data set.17 Both primary and secondary sources were used to increase the dataset's diversity and create a more trustworthy rice leaf disease multi-class classification model.

There were several rice disease and healthy classes in the original data. Five picture classes—bacterial leaf blight, false smut, healthy, leaf blast, and sheath blight—were ultimately chosen based on the available image classes and the study's new goal. These classes were chosen as they are significant rice disease conditions and healthy samples that are necessary for the classification of the diseases and normal conditions.

The contribution of primary and secondary image sources is summarized in **Table 1**. The main dataset mainly included images of false smut, healthy plants, and sheath blight, while the secondary dataset included images of bacterial leaf blight, leaf blast, healthy plants, and sheath blight. This combination allowed the final dataset to comprise of both primary samples collected in the field and disease images obtained from public sources.

Table 4. Source-wise dataset contribution.

Dataset Source	Bacterial Leaf Blight	False Smut	Healthy	Leaf Blast	Sheath Blight	Total
Primary Dataset	0	17	22	0	33	72
Secondary Dataset	179	0	154	270	267	870
Total	179	17	176	270	300	942

The final data set after the dataset cleaning, duplicate removal and label standardization was 942 images in five classes. A total of 179 bacterial leaf blight images, 17 false smut images, 176 healthy rice images, 270 leaf blast images and 300 sheath blight images were finally obtained. The finalized data set was structured based on image source, disease category, and train-validation-test split for transparent reporting of experiments.

3.2 Dataset Cleaning and Preprocessing

Systematic cleaning process was carried out prior to the development of the model. The main data set was first reviewed to find unreadable images, duplicate files, images which were visually similar and labels which were inconsistent. Images that were not labelled with any text were not included in the unlabelled review category as they did not offer any reliable text label for supervised learning. Following this inspection, the remaining primary images were then mapped to standardized class labels.

The secondary data was filtered to get only the classes needed for the final classification problem. To prevent data leakage and artificially high performance, pre-augmented folders from the secondary dataset were not used in the final preparation of the dataset. Only original pictures from the chosen secondary classes were included.

The primary and secondary records were merged and duplicate detection was done. If the same image was found in more than one class, only one image was kept. If multiple images were found under the same class, the duplicate images were deleted since they might lead to label noise in the model training. This process helped to make the final dataset "cleaner" and more reliable for supervised classification.

All of the chosen photographs were converted to RGB format before to being used in the deep learning pipeline. Images were reduced to 224 × 224 and normalised using the ImageNet mean and standard deviation because the model was built on a MobileNetV2 backbone that had been pretrained on ImageNet. This preprocessing made sure that the training, validation, and test sets were consistent and compatible with the pre-trained model.

3.3 Data Splitting and Augmentation

A stratified splitting strategy was used to split the final cleaned data into training, validation and test sets. To minimize bias in assessment, the class distribution was maintained in all subsets by stratification. The final split was 659 training, 141 validation and 142 test images.

The final train, validation and test subsets class-wise distribution is shown in Table 2. The split maintained representation of each disease category, and maintained an independent test set for final model evaluation.

Table 2. Dataset distribution across training, validation, and test sets.

Class	Training Images	Validation Images	Test Images	Total
Bacterial Leaf Blight	125	27	27	179
False Smut	12	3	2	17
Healthy	123	26	27	176
Leaf Blast	189	40	41	270
Sheath Blight	210	45	45	300
Total	659	141	142	942

The model parameters were optimized using the training set. The validation set was utilized to monitor training performance and select the optimal model checkpoint. The test set was divided and utilized solely for the model's final assessment.

Data was added solely to the training set. The augmentation techniques included resizing to 224 × 224 pixels, random horizontal flipping, random rotation, and color adjustments for brightness, contrast, and saturation. These enhancements were used to improve the model's generalization by adding variations in image orientation, lighting, and color. Only the test and validation images were resized, transformed into tensors, and normalized.

3.4 Proposed AttentionMobileNetV2-SE Architecture

The proposed model was built based on a backbone feature extractor (MobileNetV2). MobileNetV2 was chosen because it is a lightweight CNN architecture that offers a compromise between computational efficiency and classification performance. This is ideal for agricultural image classification applications and possible future use in resource-limited environments.

A Squeeze-and-Excitation attention module was added after the MobileNetV2 convolutional feature extraction block to enhance the representation of features. The attention module learns the relative importance of the feature channels and recalibrates the channel-wise feature responses. This allows the network to minimize less informative activations and concentrate on more disease-relevant ones.

The pretrained MobileNetV2 feature extractor, the Squeeze-and-Excitation attention module, and a modified classification head made up the final architecture, known as AttentionMobileNetV2-SE. The classifier comprised a fully connected layer with five output neurons after a dropout layer with a dropout rate of 0.30. These output neurons fell into five classes: bacterial leaf blight, fake smut, healthy, leaf blast, and sheath blight.

3.5 Training Configuration

A two-stage transfer learning approach was used for training the model. In the first stage, the feature extraction layers of the MobileNetV2 were frozen, while the attention module and the classifier head were trained. This enabled the newly added task-specific layers to learn the discriminative features without changing the pretrained convolutional layers at the start of training.

In the second stage, the latter convolutional blocks of MobileNetV2 were used in conjunction with the attention and classifier layers. To avoid unstable weight updates, a smaller learning rate was used to fine-tune the pretrained features to the rice disease classification task.

To overcome class imbalance in the data set, class weighted cross entropy loss was used. This was of importance since the selected classes were not equally represented, particularly the class for false smut which contained the least number of samples. The assignment of higher weights to minority classes helped to minimize the bias of the model towards majority classes.

Model optimization was done using the Adam optimizer. A learning-rate scheduler was used to decrease the learning rate if the validation loss failed to improve. Checkpointing was also included in the training pipeline. After each epoch, a checkpoint was saved to resume training in case of interruption. The best model checkpoint (based on validation accuracy) was chosen for the final test evaluation.

3.6 Evaluation Metrics and Explainability

Accuracy, precision, recall, F1 score, support, and confusion matrix analysis were used to assess the tested model. The overall percentage of correctly classified test photos was used to calculate accuracy. Precision, recall, and F1-score were computed for each class to evaluate their performance. To evaluate the models' overall performance in the face of class imbalance, macro and weighted averages were computed.

A prediction-level record was created for test set. Actual class, predicted class, confidence score and correctness status of each test image were stored. This enabled a detailed examination of the correct and incorrect predictions. Misclassification analysis was also conducted to determine the disease classes which were visually confused by the model.

To give visual explainability to the trained model, Grad-CAM was used. The final convolutional feature layer was used to create heatmaps to visualize the image areas that were most influential in the model prediction. The correctly classified and misclassified samples were visualized using the grad-CAM. This was to investigate whether the model was able to highlight the disease-relevant areas and whether it offered extra interpretability for the classification results.

4. Results

4.1 Training and Validation Performance

The training behaviour of the proposed AttentionMobileNetV2-SE model was analysed through the accuracy and loss curves in both training stages. The accuracy curve was used to observe how well the model learned the discriminative rice disease features over successive epochs, and the loss curve was used to monitor the optimization stability and convergence.

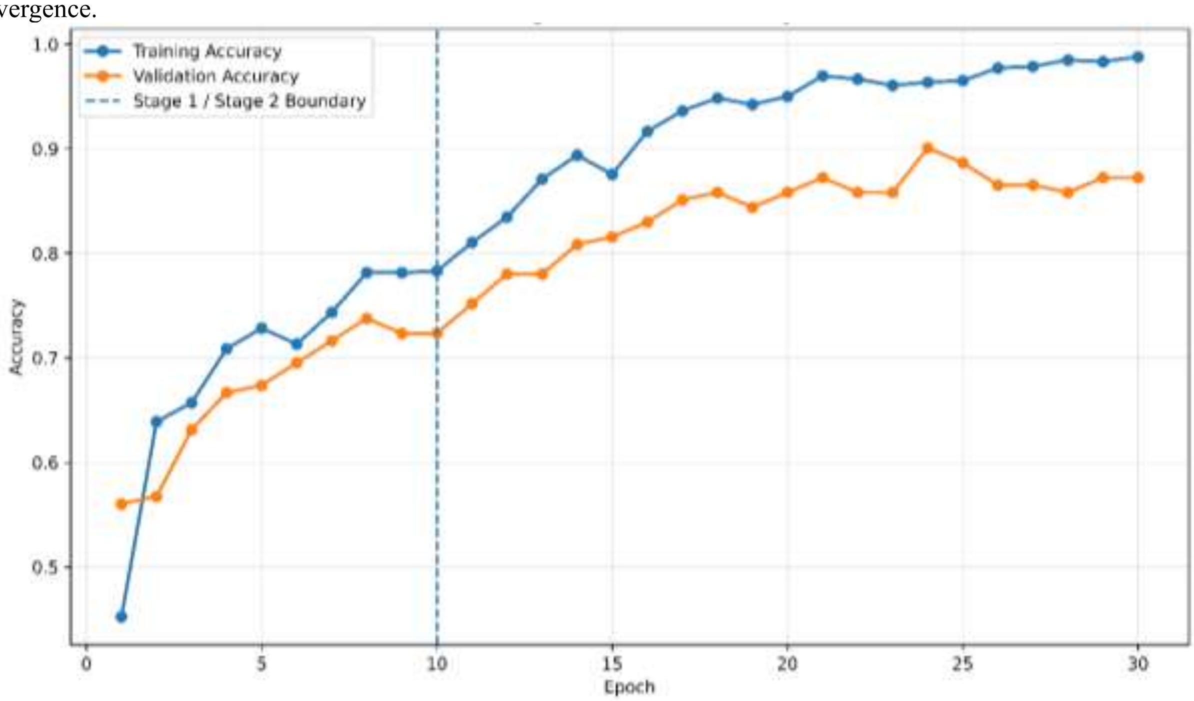

Figure 1. Training and validation accuracy curves of the proposed AttentionMobileNetV2-SE model.

During the training phase, the accuracy trend showed a steady improvement in accuracy, indicating that the model gradually picked up significant feature representations from the rice illness photos. The two-stage training strategy improved validation accuracy, as seen by the final fine-tuned model's greatest validation accuracy of 90.07%. **Figure 2** shows the loss behavior. To determine whether the model optimization was stable and whether the model was not overfitting during training, the loss curve was utilized.

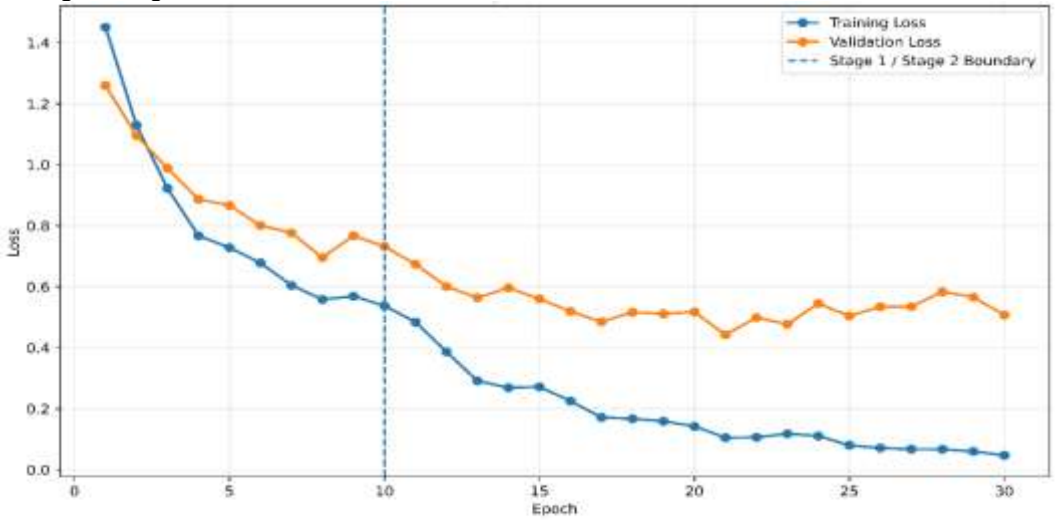

Figure 2. Training and validation loss curves of the proposed AttentionMobileNetV2-SE model.

Training and validation loss curves exhibited a general downward trend, indicating that the model reduced the classification error while training. While there were some fluctuations during fine-tuning, the validation performance stayed consistent and the best model checkpoint was chosen based on the validation accuracy.

4.2 Overall Test Performance

The overall test performance of the proposed model is summarized in Table 3. Instead of reporting all configuration details in the results table, only the key performance indicators are presented to keep the table concise and focused on model evaluation.

Table 3. Overall performance of the proposed AttentionMobileNetV2-SE model.

Model	Validation Accuracy (%)	Test Accuracy (%)	Macro F1-score	Weighted F1-score
AttentionMobileNetV2-SE	90.07	86.62	0.8005	0.8656

The proposed model achieved a validation accuracy of 90.07% and a test accuracy of 86.62%. The weighted F1-score of 0.8656 indicates that the model maintained strong overall classification performance, while the macro F1-score of 0.8005 reflects the influence of class imbalance, particularly the limited number of False Smut samples. To provide a visual summary of the overall test performance, Figure 3 presents the key evaluation metrics of the proposed model.

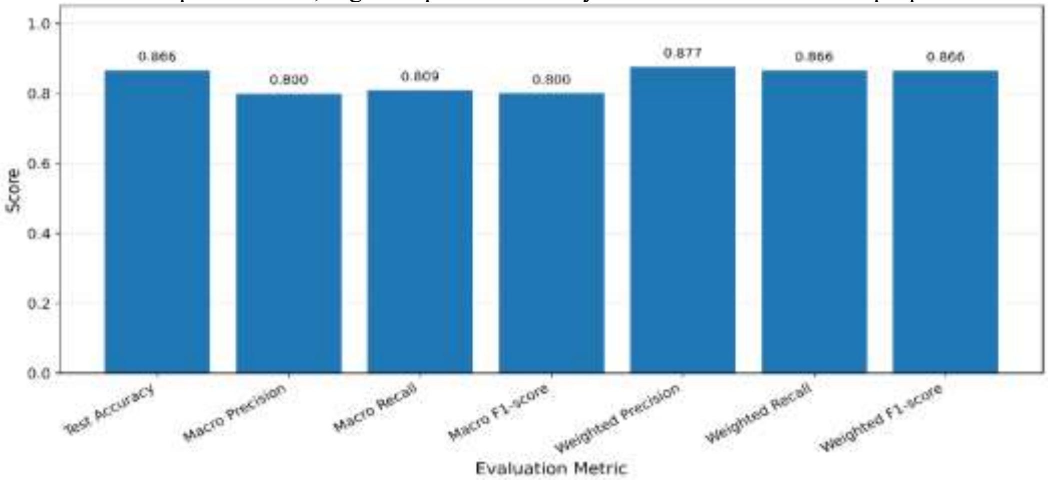

Figure 3. Overall performance summary of the proposed rice disease classification model.

The overall performance summary confirms that the model maintained balanced weighted precision, recall, and F1-score values. This indicates that the model performed consistently on the combined primary and secondary dataset, despite class imbalance in the final dataset.

4.3 Class-wise Performance

Class-wise evaluation was performed to examine the model performance for each rice disease category separately. Table 4 presents precision, recall, F1-score, and support for all five classes.

Table 4. Class-wise precision, recall, F1-score, and support on the test set.

Class	Precision	Recall	F1-score	Support
Bacterial Leaf Blight	0.7742	0.8889	0.8276	27
False Smut	0.5000	0.5000	0.5000	2
Healthy	0.9310	1.0000	0.9643	27
Leaf Blast	0.8222	0.9024	0.8605	41
Sheath Blight	0.9714	0.7556	0.8500	45

The Healthy class achieved the highest F1-score of 0.9643, indicating that the model was highly effective in distinguishing healthy rice samples from diseased samples. Leaf Blast and Sheath Blight also achieved strong F1-scores of 0.8605 and 0.8500, respectively. Bacterial Leaf Blight was able to obtain an F1 score of 0.8276, which was acceptable for classification. The F1-score for False Smut class was 0.5000, which was relatively low, due to the small number of samples in this class (2 in the test set). Precision, Recall and F1-Score values for each class are visually compared in **Figure 4**.

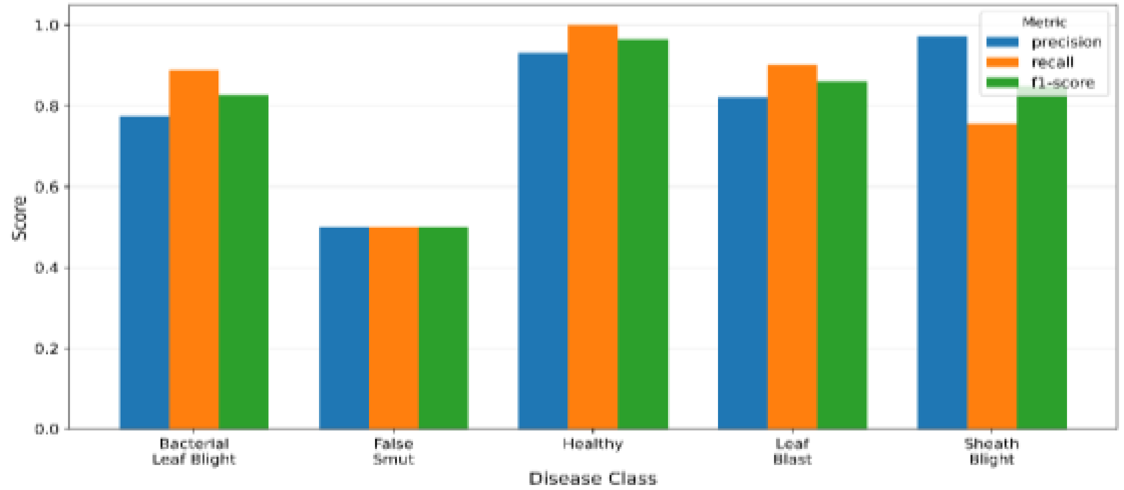

Figure 4. Class-wise precision, recall, and F1-score comparison for the proposed model.

The class-wise metric comparison shows that the model performed strongly for Healthy, Leaf Blast, Sheath Blight, and Bacterial Leaf Blight classes. The lower performance for False Smut highlights the impact of limited class samples and indicates the need for additional primary images for this disease class.

4.4 Confusion Matrix Results

To assess the proportion of accurate and inaccurate predictions across all test classes, a confusion matrix was created. The suggested model's test confusion matrix is shown in **Figure 5**.

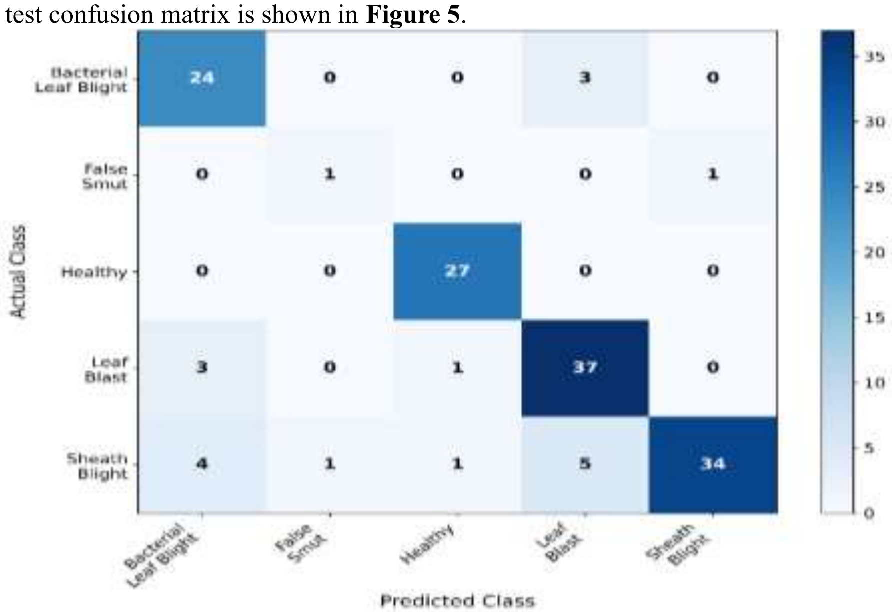

Figure 3. Confusion matrix of the proposed model on the independent test set.

The confusion matrix indicates that the model correctly classified 123 images out of 142 test images. There were no misclassified samples in the test set of the Healthy class. Recall for Leaf Blast and Bacterial Leaf Blight was good, but some misrecognition was noted between these two visually similar disease categories. Sheath Blight also exhibited high accuracy, with a few Sheath Blight samples being incorrectly identified as Leaf Blast or Bacterial Leaf Blight. These results show that the model acquired meaningful disease specific features and some visual similarity between disease symptoms still resulted in classification errors.

4.5 Misclassification Results

To further analyze model errors, a misclassification summary was prepared by grouping incorrect predictions according to actual and predicted classes. Table 5 presents the major misclassification patterns observed in the independent test set.

Table 5. Misclassification summary of incorrect test predictions.

Actual Class	Predicted Class	Misclassified Images
Sheath Blight	Leaf Blast	5
Sheath Blight	Bacterial Leaf Blight	4
Bacterial Leaf Blight	Leaf Blast	3
Leaf Blast	Bacterial Leaf Blight	3
Leaf Blast	Healthy	1
False Smut	Sheath Blight	1
Sheath Blight	False Smut	1
Sheath Blight	Healthy	1

Sheath Blight most frequently was misclassified as Leaf Blast. This indicates that some of the visual symptoms could be the same between some of these classes in terms of leaf texture, appearance of lesions, or affected areas. Some Bacterial Leaf Blight samples were also predicted as Leaf Blast, and some Leaf Blast samples were predicted as Bacterial Leaf Blight. These errors suggest that disease categories with similar looking lesions are still difficult for the model. One misclassification was made in the False Smut class, which is not surprising since the number of samples for this class is very small.

4.6 Grad-CAM Visualization Results

Grad-CAM visualization was used to interpret the regions of the image that contributed most strongly to the model predictions. Figure 6 presents Grad-CAM examples for both correct and misclassified samples.

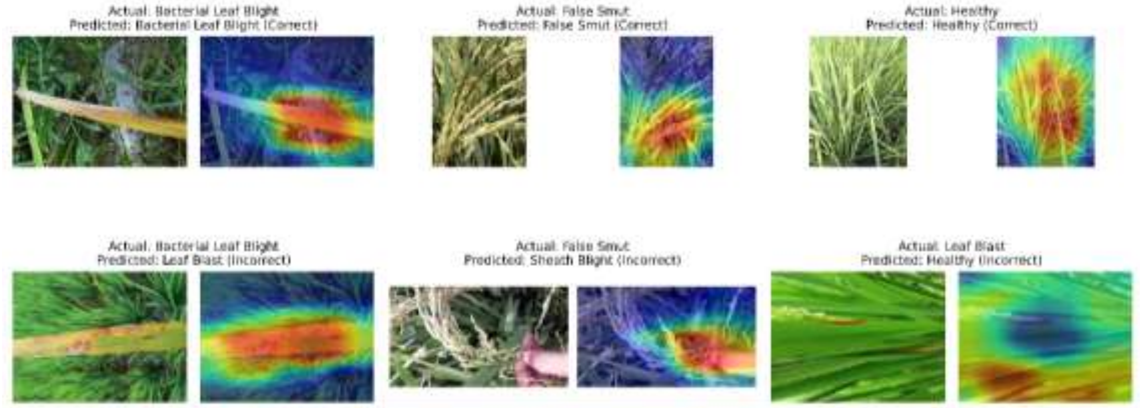

Figure 6. Grad-CAM visualization of correct and misclassified rice disease test samples.

As can be seen in the Grad-CAM results, in some of the correctly classified samples, the model was able to pay attention to the most important parts of the image, such as the affected parts of the plant, the lesion pattern, and the diseased leaf area. This indicates that the proposed AttentionMobileNetV2-SE model could learn the visual cues related to the disease instead of just background information. Heatmaps revealed that in misclassified cases, attention was focused on areas of overlapping or ambiguous symptoms, which may account for the confusion between the visually similar disease classes. Overall, the Grad-CAM analysis can be used to offer interpretability support to the model predictions and enhance the reliability of the proposed classification framework.

5. Discussion

Using the cleaned mix of main and secondary image datasets for multi-class classification, the suggested AttentionMobileNetV2-SE model successfully classified the rice leaf illnesses. The model's validation accuracy of 90.07% and test accuracy of 86.62% demonstrated that the suggested framework could extract significant visual characteristics from the pictures of rice illness. Despite the existence of class imbalance, the model performed well overall on the test set, as indicated by the weighted F1-score of 0.8656. This is important since the performance of deep learning-based plant disease recognition systems is significantly impacted by dataset quality, image acquisition conditions, and intra-class variability.[18]

When the model's performance was assessed for every class, it was found that the Healthy class had the highest F1 score. The model was able to identify distinct disease patterns for Leaf Blast, Sheath Blight, and Bacterial Leaf Blight, as evidenced by their strong classification performance. The False Smut class, on the other hand, had 17 images in the final data set with just 2 images in the test set, resulting in poorer performance. This means that minority class representation may impact class-wise performance particularly in deep learning models trained on imbalanced data. Moreover, previous studies have demonstrated that the size and diversity of the datasets have a direct impact on the performance of deep learning and transfer learning models for plant disease classification.[19]

Additionally, the misclassification results showed that certain disease categories were visually similar and difficult for the model to differentiate. The most confusing disease was Sheath Blight, which was mistaken for Leaf Blast. This was followed by Bacterial Leaf Blight and Leaf Blast. The existence of overlapping visual symptoms, similar lesion patterns, and a lack of variety in some disease categories may all contribute to some of these inaccuracies. There was one False Smut class that was incorrectly identified as Sheath Blight, which is not surprising given the limited training and test set of False Smut samples. This has been seen in image-based crop disease detection, where the actual symptom in the field, the background variation, and the diversity of images from the field can affect the reliability of the model.[20]

The Grad-CAM visualizations gave further insight into the behaviour of the model in terms of predictions. For several correctly classified samples, the heatmaps showed that the model paid attention to the areas of the leaves affected by the disease, lesion patterns and visible symptom regions. This indicates the usefulness of the proposed attention-driven approach as it means that the model was not just using irrelevant background features. In misclassified cases, however, the Grad-CAM outputs provided attention on visually ambiguous or overlapping regions, which might account for the confusion between some disease classes. This interpretability is crucial as plant disease classification systems should not only predict the disease but also give some insight into what the model is looking at. Systematic reviews have also highlighted the importance of the development of CNN-based plant disease recognition systems with strong interpretability.[21]

However, the study has a few drawbacks. First, the main dataset was quite limited, giving limited variation at field level during training. Secondly, the False Smut class had very small samples, and the performance of the classes was not very strong. Third, the model was tested on the test set prepared and was not externally tested on a separate field data set acquired under different conditions and locations. Fourth, even though class weighted loss was applied, there was still a dataset imbalance. Lastly, the model was not developed for disease severity estimation or lesion segmentation for disease category classification. The limitations suggest that although the current model can be used for classification, it needs to be further validated prior to any practical field deployment.

Future work will involve increasing the number of images of rice at the primary level that will be gathered under varying light conditions, backgrounds, growth stages, and geographic locations. More False Smut images should be captured to minimize class imbalance and provide better representation of the minority class. The framework can be expanded to include additional rice disease categories and evaluate the model in field environments not used for training. Potential future enhancements include mobile or web-based deployment for field disease screening and segmentation-based or severity estimation modules for more in-depth disease assessment. The results indicate that deep transfer learning has great potential in plant disease identification, and further research will be conducted to explore powerful transfer learning strategies to promote generalization for various real-world datasets.[22]

6. Conclusion

An AttentionMobileNetV2-SE framework for multi-class rice leaf disease classification was proposed by the authors of this study. To improve the representation of disease-relevant features, the model was created using a Squeeze-and-Excitation attention module and a lightweight MobileNetV2 backbone. A composite dataset was created using both primary field level pictures and secondary publicly accessible rice disease images. The data set was divided into a stratified train-validation-test set for equitable model evaluation after being cleaned to eliminate duplicate images, cross-labeling conflicts, and untrustworthy unlabeled samples. The suggested model's test accuracy on the independent test set was 86.62%, while its validation accuracy was 90.07%. The class False Smut did badly due to limited sample availability, while the classes Healthy, Leaf Blast, Sheath Blight, and Bacterial Leaf Blight performed well, according to the results. Grad-CAM visualizations, which emphasized areas of the image that were crucial for the model's predictions, were also used to portray the model predictions in order to improve interpretability. Finally, it is demonstrated that a cleaned primary and secondary dataset can be used to classify several rice leaf disease classes using an attention-based transfer learning model. By collecting additional primary photos, future research can improve the model's robustness in the field, adding more samples to the minority class, testing the model on other field images, and expanding the framework to include disease severity estimation or real-time deployment.

References

1. Fukagawa NK, Ziska LH. Rice: Importance for Global Nutrition. J Nutr Sci Vitaminol. 2019 Oct 11;65(Supplement):S2–3. doi:10.3177/jnsv.65.S2
2. Khush GS. What it will take to Feed 5.0 Billion Rice consumers in 2030. Plant Mol Biol. 2005 Sep;59(1):1–6. doi:10.1007/s11103-005-2159-5
3. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. Nat Ecol Evol. 2019 Feb 4;3(3):430–9. doi:10.1038/s41559-018-0793-y

4. Skamnioti P, Gurr SJ. Against the grain: safeguarding rice from rice blast disease. Trends in Biotechnology. 2009 Mar;27(3):141–50. doi:10.1016/j.tibtech.2008.12.002
5. Mahlein AK. Plant Disease Detection by Imaging Sensors – Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. Plant Disease. 2016 Feb;100(2):241–51. doi:10.1094/PDIS-03-15-0340-FE
6. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-Excitation Networks. 2018;7132–41. doi:10.1109/CVPR.2018.00745
7. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. 2017;618–26. doi:10.1109/ICCV.2017.74
8. Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci. 2016 Sep 22;7. doi:10.3389/fpls.2016.01419
9. Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture. 2018 Apr 1;147:70–90. doi:10.1016/j.compag.2018.02.016
10. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture. 2018 Feb 1;145:311–8. doi:10.1016/j.compag.2018.01.009
11. Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Computational Intelligence and Neuroscience. 2016;2016:1–11. doi:10.1155/2016/3289801
12. Brahimi M, Boukhalfa K, Moussaoui A. Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. Applied Artificial Intelligence. 2017 Apr 21;31(4):299–315. doi:10.1080/08839514.2017.1315516
13. Lu Y, Yi S, Zeng N, Liu Y, Zhang Y. Identification of rice diseases using deep convolutional neural networks. Neurocomputing. 2017 Dec 6;267:378–84. doi:10.1016/j.neucom.2017.06.023
14. Too EC, Yujian L, Njuki S, Yingchun L. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture. 2019 Jun;161:272–9. doi:10.1016/j.compag.2018.03.032
15. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition [Internet]. Salt Lake City, UT: IEEE; 2018 [cited 2026 Jun 17]. p. 4510–20. Available from: https://ieeexplore.ieee.org/document/8578572/ doi:10.1109/CVPR.2018.00474
16. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In: Proceedings of the 36th International Conference on Machine Learning [Internet]. PMLR; 2019 [cited 2026 Jun 17]. p. 6105–14. Available from: https://proceedings.mlr.press/v97/tan19a.html
17. Hasan M, Khatun S, Raihan MA, Uddin AH. Rice Leaf Bacterial and Fungal Disease Dataset [Internet]. Vol. 2. 2023 Nov 28;2. doi:10.17632/hx6f852hw4.2
18. Barbedo JGA. Factors influencing the use of deep learning for plant disease recognition. Biosystems Engineering. 2018 Aug;172:84–91. doi:10.1016/j.biosystemseng.2018.05.013
19. Barbedo JGA. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Computers and Electronics in Agriculture. 2018 Oct;153:46–53. doi:10.1016/j.compag.2018.08.013
20. Ramcharan A, Baranowski K, McCloskey P, Ahmed B, Legg J, Hughes DP. Deep Learning for Image-Based Cassava Disease Detection. Front Plant Sci. 2017 Oct 27;8. doi:10.3389/fpls.2017.01852
21. Abade AS, Ferreira PA, Vidal F de B. Plant Diseases recognition on images using Convolutional Neural Networks: A Systematic Review [Internet]. arXiv; 2020 [cited 2026 Jun 17]. Available from: http://arxiv.org/abs/2009.04365 doi:10.48550/arXiv.2009.04365
22. Chen J, Chen J, Zhang D, Sun Y, Nanehkaran YA. Using deep transfer learning for image-based plant disease identification. Computers and Electronics in Agriculture. 2020 Jun 1;173:105393. doi:10.1016/j.compag.2020.105393