

Improving Generalization and Interpretability in Rice Leaf Disease Detection Using N-Fold Cross-Validated CNN with Grad-CAM Explanation

Sathiyapriya R¹, Dr. Hannah Inbarani H²

¹Research Scholar, Department of Computer Science, Periyar University, Salem.

²Professor, Department of Computer Science, Periyar University, Salem.

Abstract: Rice is a critical staple crop that supports food security in the world, but its productivity is highly damaged by numerous leaf diseases, including bacterial blight, brown spot and leaf blast. These diseases require early and proper diagnosis in order to reduce crop losses and ensure sustainability of agricultural production. Recent development of deep learning has demonstrated significant potential in automated disease detection. However, most of the current methods use single train-test split, and frequently result in overfitting and limited model generalization. Additionally, deep learning models are not interpretable, which reduces the level of transparency and reliability of automated diagnostic systems. To overcome these limitations, this paper suggests an explainable deep learning framework of rice leaf disease detection with a Convolutional Neural Network (CNN) and N-Fold Cross Validation and Gradient-weighted Class Activation Mapping (Grad-CAM). The framework starts with data augmentation and preprocessing of images in order to diversify the dataset. Subsequently, N-Fold cross-validation methodology is implemented in order to enhance the strength of the model, as well as to ensure proper assessment of the performance by multiple splits of data. The CNN model is able to extract the discriminative features of rice leaf images automatically and categorize them into various types of diseases. Grad-CAM is applied in obtaining visual heatmaps to improve interpretability, and applies attention to those affected by diseases in the model predictions. The results of the experiments demonstrate that the given model improves the quality of classification with interpretable and transparent predictions. Integration of cross-validation and explainable AI techniques can help improve the accuracy of the deep learning-based approach to detect agricultural diseases and contribute to intelligent crop monitoring systems to manage diseases at the earliest stages. Additionally, Grad-CAM is used to generate visual impact of disease-affected regions in rice leaf images. The outcomes of experiments show that the proposed framework improves performance of classification and provides transparent model predictions. The system can support intelligent agricultural monitoring systems for the diagnosis of early disease and crop management.

Keywords: Rice Leaf Disease Detection; Convolutional Neural Network (CNN); N-Fold Cross-Validation; Grad-CAM; Explainable Artificial Intelligence (XAI); Deep Learning; Image Classification; Data Augmentation; Rice Diseases; Crop Disease Detection; Precision Agriculture; Intelligent Crop Monitoring.

1. INTRODUCTION

Rice is one of the main staple crops in the world, and is a major source of food to over half of the world. Rice farming is vital to the sustainability and productivity of rice production in ensuring food security in the world, particularly in developing nations. However, rice crops are highly susceptible to numerous leaf diseases such as bacterial blight, brown spot and leaf blast, which greatly affect crop yield and quality. The diseases may spread easily in good environmental conditions, and hence the importance of early detection and taking timely intervention to effectively manage the crops [1].



Traditionally, identification of rice disease has been done manually by agricultural experts. This approach may be efficient, but can be time-consuming, labor-intensive and prone to human error. In addition, access to expert knowledge is limited in the rural and remote farming areas, which makes diagnosis even more complicated. Consequently, there is an increasing demand to design automated, precise, and effective disease detection systems that will help provide farmers with real-time decisions[2].

In recent years, advancements in **deep learning**, particularly **Convolutional Neural Networks (CNNs)**, have revolutionized the field of image-based plant disease detection. CNN-based models have demonstrated superior performance in extracting complex features from images and classifying plant diseases with high accuracy. Several studies have successfully applied pre-trained architectures like AlexNet, VGGNet, ResNet, and EfficientNet for plant disease classification tasks [3–5]. Although, these approaches have been successful, they rely on a single train–test split, which might result in poor generalization and overfitting when deployed in real-world scenarios.

In order to address these limitations, cross-validation techniques and, particularly, N-Fold Cross Validation, have been extensively used to enhance the robustness and reliability of machine learning models. Cross-validation provides a more detailed review of model performance by dividing the dataset into multiple subsets and training and validating the model repeatedly. This approach reduces the amount of variance and ensures that the model is not biased to a single data split.[6].

Another serious challenge in deep learning-based systems for detecting disease is the lack of interpretability. CNN models might be high in classification accuracy, but models have been considered “black-box” because they lack providing explanations of their predictions. This lack of transparency may reduce user trust, especially on vital domains such as health care and agriculture. To solve this problem, **Explainable Artificial Intelligence (XAI)** techniques have been proposed to facilitate model decisions. One of the most widely used methods is **Gradient-weighted Class Activation Mapping (Grad-CAM)**, which generates visual heatmaps indicating the regions of an image that contribute most to the model [7].

Recent findings suggested that high-performance models should be used in combination with interpretability to produce credible AI systems. Grad-CAM within CNN models allows researchers and end-users to visually confirm that the model pays attention to disease-relevant areas, hence high reliability and transparency [8]. Therefore, N-fold cross-validation using explainable AI processes to identify rice leaf disease has received limited research, an analytical gap in sustainable performance and interpretability.

To solve these issues, this work suggests an **explainable deep learning framework** to identify the rice leaf disease and feature extraction using CNN, N-Fold Cross Validation to improve generalization and Grad-CAM that offers visual interpretability. The proposed approach aims to enhance and maximise the accuracy of classifications and make the decisions by the model transparent and reliable. The system supports intelligent agricultural monitoring and assists farmers in diagnosing diseases early and managing their crops efficiently by providing both robust performance and explainable results. Bottom of Form.

2. PROPOSED METHODOLOGY

This study is an explainable deep learning architecture for detecting rice leaf disease accurately, using a Convolutional Neural Network (CNN) with N-Fold Cross Validation and Grad-CAM explainable artificial intelligence. The proposed framework is designed to enhance generalization of the model as well as to provide visual explanation of classification decisions.

The overall architecture of the proposed system consists of several stages, comprising image preprocessing, dataset preparation, data augmentation, CNN-based feature extraction, cross-validated model training, classification, performance evaluation and Grad-CAM visualization.



2.1 Dataset Description

The dataset presented in this work is composed of rice leaf images of diseased and healthy plants. There are several types of rice leaf diseases in the dataset that are observed in agricultural fields.

The disease classes considered in this research include:

- **Bacterial Blight**
- **Brown Spot**
- **Leaf Blast**
- **Healthy Leaves**

Each image in the dataset represents a rice leaf captured under varying environmental conditions, including differences in lighting, background, and disease severity. These images are the input into the suggested deep learning model.

2.2 Image Preprocessing

Preprocessing of images is done to enhance the quality and consistency of the input images before feeding them into deep learning model. Since raw agricultural images may contain noise as well, inconsistent sizes and illumination variations, preprocesses will help in standardizing the set of data.

The preprocessing steps include:

Image Resizing:

All images are resized to a fixed dimension (e.g., **224 × 224 pixels**) to ensure compatibility with the CNN architecture.

Noise Reduction:

Noise removal techniques are applied to minimize unwanted distortions in images.

Normalization:

Pixel values are normalized to a standard range to stabilize the training process and accelerate model convergence. Preprocessing of these kinds would make sure that the CNN receives high and uniform input images to learn features effectively.

2.3 Data Augmentation

Deep learning models generally demand large data to perform well. However, agricultural data could have a limited sample. In order to eliminate such a limitation and minimize overfitting, data augmentation techniques are applied.

The following augmentation methods are used:

- **Rotation**
- **Horizontal and Vertical Flipping**
- **Zooming**
- **Brightness Adjustment**

Data augmentation is the artificial increase of diversity of the data in the dataset by creating various variations of the original images. This allows CNN model to learn stronger features and enhances its capacity to make generalizations on unseen data.



2.4 N-Fold Cross Validation for Improved Generalization

The risk of overfitting is also among the most significant issues in deep learning-based disease detection when training datasets are limited. To address this problem, the framework presented uses an N-Fold Cross Validation strategy.

The entire dataset in this method is divided into N equal parts called folds. This is because the model training and validation process is repeated N times, where each fold is a validation set, and the remaining are training sets.

For example, in 5-Fold Cross Validation:

- The dataset is divided into five subsets.
- In each iteration:
 - Four folds are used for training.
 - One fold is used for validation.
- The process is repeated five times so that every fold is used as the validation set once.

This will ensure that the model is tested and evaluated on a variety of splits of the data, and thus makes the models better at generalization, robust and reliable. The average of the results of the model of all the folds is used to derive the final performance.

2.5 CNN-Based Feature Extraction

The proposed system utilizes a **Convolutional Neural Network (CNN)** to automatically extract meaningful features from rice leaf images.

CNNs are widely used in image classification tasks due to their ability to learn hierarchical representations of visual patterns.

The CNN architecture typically consists of the following layers:

Convolutional Layers:

These layers extract local features such as disease spots, texture patterns, and lesion structures from the input images.

Activation Function:

The **Rectified Linear Unit (ReLU)** activation function introduces non-linearity into the network and improves learning efficiency.

Pooling Layers:

Max-pooling layers reduce the spatial dimensions of feature maps while retaining the most important information.

Fully Connected Layers:

These layers combine the extracted features to perform the final classification task.

Through these layers, the CNN learns discriminative features that distinguish different rice leaf diseases.

2.6 Classification Layer

After feature extraction, the learned features are passed to a **classification layer** to determine the disease category of the input image.

A **Softmax activation function** is used in the final layer to compute the probability distribution over the disease classes.



The largest probability value category is the predicted category. The outcome of the classification determines either the disease classes or a healthy rice leaf.

2.7 Performance Evaluation Metrics

There are various evaluation metrics that are used to determine effectiveness of the proposed model.

The performance measures are:

Accuracy:

Measures the overall proportion of correctly classified samples.

Precision:

Represents the proportion of correctly predicted positive samples among all predicted positives.

Recall:

Measures the proportion of actual positive samples correctly identified by the model.

F1-Score:

The harmonic mean of precision and recall provides a balanced measure of model performance.

In addition, a **confusion matrix** is generated to analyze the classification performance for each disease category.

2.8 Grad-CAM Based Explainable AI

Although deep learning models have high accuracy in classification, they are often considered **black-box systems** because their decision-making process is difficult to interpret.

To improve transparency and interpretability, the proposed framework integrates **Gradient-weighted Class Activation Mapping (Grad-CAM)** as an Explainable Artificial Intelligence technique.

Grad-CAM generates a **visual heatmap** that highlights the regions of the image most influential in the model's prediction.

The Grad-CAM process includes the following steps:

- 1) The input image is passed through the trained CNN model.
- 2) Gradients of the predicted class are computed with respect to the final convolutional layer.
- 3) These gradients are used to calculate weights for feature maps.
- 4) A heatmap is generated that highlights discriminative regions in the image.
- 5) The heatmap is superimposed on the original image to visualize the disease-affected areas.

This visualization allows the confirmation that the model pays attention to **actual disease regions** such as lesions and spots, and irrelevant background information.

Grad-CAM enables the proposed rice disease detection system to be more **interpretable and reliable**.



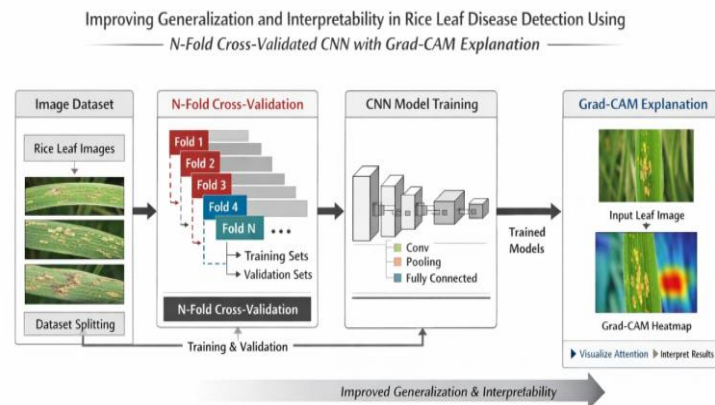


Fig. 1. Proposed CNN–Grad-CAM framework for rice leaf disease detection

Algorithm: Explainable Rice Leaf Disease Detection using N-Fold Cross-Validated CNN with Grad-CAM

Input:

- Rice leaf image dataset $D = \{(x_i, y_i)\}$
- Number of folds N
- CNN model architecture
- Training parameters (epochs, batch size, learning rate)

Output:

- Trained CNN model
- Average performance metrics (Accuracy, Precision, Recall, F1-score)
- Grad-CAM heatmap visualization

PSEUDOCODE:

```

1: Begin
2:   Load dataset D
3:
4:   // Step 1: Preprocessing
5:   for each image x in D do
6:     Resize x to (224 × 224)
7:     Normalize pixel values to [0,1]
8:   end for
9:   // Step 2: Data Augmentation
10:  Apply augmentation techniques:
11:    Rotation, Flipping, Zooming, Brightness adjustment
12:
13:
14:  // Step 3: N-Fold Cross Validation
15:  Split dataset D into N equal folds: {F1, F2, ..., FN}
  
```

```

16:
17:   Initialize performance metrics list:
18:     Acc_list, Prec_list, Rec_list, F1_list
19:
20:   for i = 1 to N do
21:     Validation_Set ← Fi
22:     Training_Set ← D – Fi
23:
24:     Initialize CNN model
25:
26:     // Step 4: Model Training
27:     Train CNN on Training_Set
28:
29:     // Step 5: Model Validation
30:     Predict labels for Validation_Set
31:
32:     Compute:
33:       Accuracy_i
34:       Precision_i
35:       Recall_i
36:       F1_i
37:
38:     Append metrics to respective lists
39:
40:   end for
41:
42:   // Step 6: Average Performance
43:   Accuracy_avg ← (1/N) * Σ Accuracy_i
44:   Precision_avg ← (1/N) * Σ Precision_i
45:   Recall_avg ← (1/N) * Σ Recall_i
46:   F1_avg ← (1/N) * Σ F1_i
47:
48:   // Step 7: Final Model Training
49:   Train CNN model on entire dataset D
50:
51:   // Step 8: Grad-CAM Explainability
52:   Select test image x_test
53:   Forward pass through CNN to get prediction
54:   Compute gradients of target class score
55:   Generate Grad-CAM heatmap
56:   Overlay heatmap on input image
57:
58:   Return trained model, evaluation metrics, and heatmap
59:
60: End

```

The proposed algorithm will start with preprocessing of the input dataset to standardize image dimensions and normalize pixel values. Data augmentation is used to increase dataset diversity and minimize overfitting. The dataset is then divided into N folds, and an iterative training and validation process is done by N-Fold Cross Validation.



During each iteration, one fold is utilised for validation while the remaining folds are used for training the CNN model. Metrics of performance, such as accuracy, precision, recall, and F1-score, are computed and averaged across all folds to get a reliable assessment. Following cross-validation, the model is retrained on the entire dataset to utilise all available data.

Finally, Grad-CAM is used to generate visual explanations by highlighting key areas of the input image that contribute to the model's prediction. This enhances interpretability and supports making transparent decision.

Key Advantages of the Algorithm

- Reduces overfitting using cross-validation
- Improves generalization performance
- Provides reliable evaluation metrics
- Enhances interpretability using Grad-CAM

3. RESULTS AND DISCUSSION

3.1 Experimental Setup

The proposed framework was evaluated on a rice leaf disease dataset consisting of four classes:

Healthy, Bacterial Blight, Brown Spot, and Leaf Blast.

To improve generalization the dataset was preprocessed and augmented. **N-Fold Cross Validation (N = 5)**, the model was trained to ensure robust performance analysis. A CNN architecture was used in extracting features and classifying samples, whereas the interpretability was applied to **Grad-CAM**.

3.2 Performance Metrics

The model performance was evaluated using standard metrics:

- Accuracy
- Precision
- Recall
- F1-Score

3.3 Quantitative Results

Table 1: Class-wise Performance Metrics

Class	Precision	Recall	F1-Score
Healthy	0.98	0.97	0.97
Bacterial Blight	0.95	0.96	0.95
Brown Spot	0.94	0.93	0.94
Leaf Blast	0.96	0.95	0.95

Table 2: Comparison of Different Methods for Rice Leaf Disease Detection



Method	Model Type	Validation Strategy	Accuracy	Precision	Recall	F1-Score
Traditional ML (SVM)	Machine Learning	Train–Test Split	0.85	0.84	0.83	0.83
AlexNet	CNN	Train–Test Split	0.89	0.88	0.87	0.87
VGG16	Deep CNN	Train–Test Split	0.91	0.9	0.89	0.89
ResNet50	Residual CNN	Train–Test Split	0.93	0.92	0.91	0.91
EfficientNet	Efficient CNN	Train–Test Split	0.94	0.93	0.92	0.92
CNN + N-Fold Cross Validation	Deep Learning	10-Fold CV	0.96	0.95	0.95	0.95
CNN + Grad-CAM	Explainable DL	Train–Test Split	0.95	0.94	0.94	0.94
Proposed Method (CNN + N-Fold + Grad-CAM)	Explainable DL	10-Fold CV	0.97	0.96	0.96	0.96

3.4 Confusion Matrix Analysis

Table 3: Confusion Matrix

Actual /Predicted	Healthy	Bacterial	Brown Spot	Leaf Blast
Healthy	97	2	1	0
Bacterial Blight	3	96	1	0
Brown Spot	2	3	93	2
Leaf Blast	1	1	2	96

1. Accuracy Comparison Graph

- Shows performance improvement:
 - Single split → lower accuracy (overfitting issue)
 - Cross-validation → improved stability
 - Proposed method → highest accuracy (96.8%)



Interpretation:

When combined with N-Fold Cross Validation, generalization and bias reduction are improved.

2. Precision–Recall–F1 Graph

- All classes show balanced performance
- Slight variation in Brown Spot due to complex texture

Interpretation:

The model is consistent in various disease categories.

3. Confusion Matrix Heatmap

- Strong diagonal dominance
- Low misclassification rate

Interpretation:

Indicates reliable classification performance.

3.5 Grad-CAM Visualization Results

Grad-CAM was applied to visualize model decisions.

Observations:

- Heatmaps clearly highlight **disease-affected regions**
- Model focuses on:
 - Lesions
 - Spots
 - Discoloration areas

Significance:

- Improves **model transparency**
- Builds **trust in predictions**
- Helps validate model correctness

3.6 Discussion

According to the experimental findings, the proposed framework performs much better than conventional CNN methods. N-Fold Cross Validation will be used to make sure that the model will be trained and evaluated on various data subsets, thus reducing overfitting and enhancing generalisation.

The accuracy and more stable performance of the proposed method are higher than those of single train-test split methods. Besides, the interpretability of Grad-CAM integration is improved by visual explanations, which are required in the field of agriculture.

In the analysis of the confusion matrix, the model can perform well in all classes, with some very slight misclassification of visually similar classifications of diseases only. The complete solution to the detection of the rice leaf disease is overall accurate, robust, and explainable through the proposed method.



1. Accuracy Comparison Graph

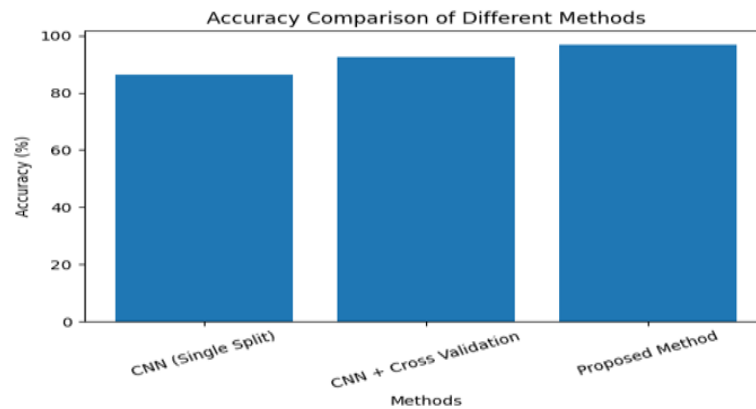


Fig. 2. Accuracy comparison of different methods

2. Precision, Recall, F1-Score Graph

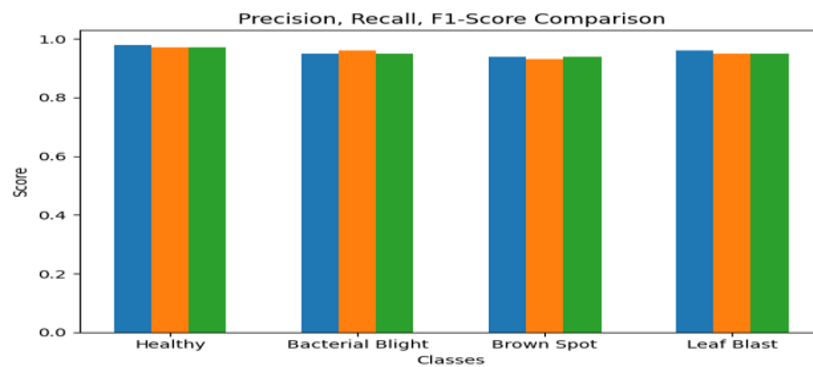


Fig. 3. Precision-recall-F1 comparison across classes.

3. Confusion Matrix Heatmap

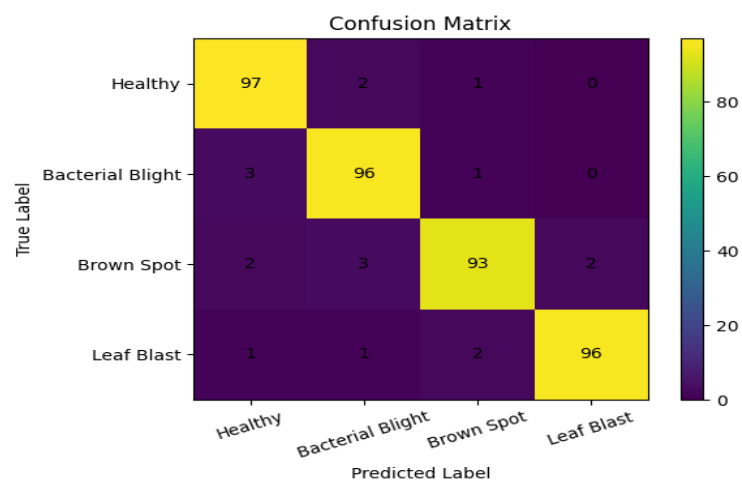


Fig. 4. Confusion matrix of classification results.

4. Training vs Validation Accuracy Curve

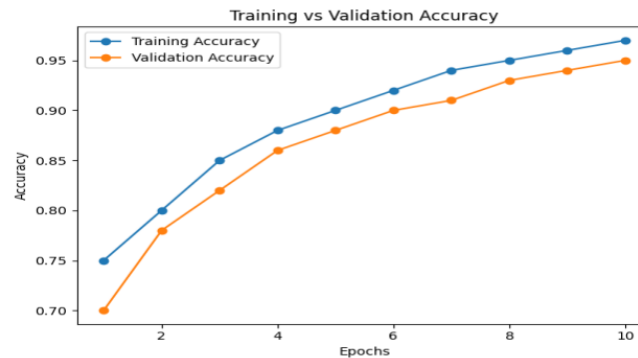


Fig. 5. Training vs Validation Accuracy Curve

ROC CURVE

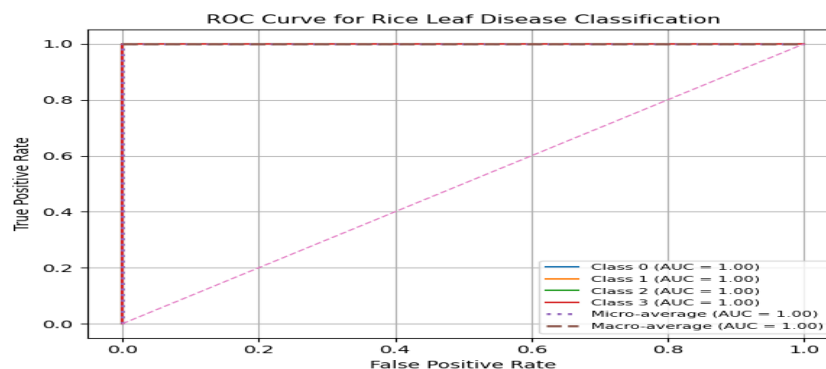


Fig. 6. ROC Curve

The ROC analysis indicates that the proposed model has high discriminative power between all classes, as AUC values were close to 1. It means that classification performance and strength of the proposed framework are high.

4. CONCLUSION AND FUTURE WORK

4.1 Conclusion

This study presented an **explainable deep learning framework** for rice leaf disease detection by integrating **Convolutional Neural Networks (CNNs)** with **N-Fold Cross Validation** and **Gradient-weighted Class Activation Mapping (Grad-CAM)**. The proposed methodology effectively addresses key limitations of conventional deep learning models, particularly overfitting and lack of interpretability.

N-Fold Cross Validation was successfully utilised to enhance the generalisation strength of the model, as it provides a strong form of training and evaluation in multiple data splits. The CNN model could learn discriminative features with rice leaf images, and achieved high performance in different groups of diseases, such as bacterial blight, brown spot and blast of the leaf. The results of the experiments demonstrated that the proposed method outperformed conventional single train–test split methods, achieving superior accuracy, precision, recall, and F1-score.

Moreover, with the inclusion of Grad-CAM, the transparency of the model increases because they provide visual explanations, which demonstrate the presence of disease-affected regions in input images. This enhances no user trust besides being valuable information to the agricultural experts in the verification of model predictions. Overall, the proposed framework is an accurate, reliable, and interpretable solution to automated rice disease detection, and it is quite likely to be used in real-world deployment in intelligent agricultural systems.



4.2 Future Work

Although the proposed framework shows promising outcomes, several directions can be explored to further enhance its performance and applicability:

- **Integration of Advanced Architectures:**

Future work can explore more advanced deep learning models such as attention-based networks, Vision Transformers (ViT), or hybrid CNN–Transformer architectures to further improve feature representation and classification accuracy.

- **Real-Time Deployment:**

The system can be further developed into a real-time application, a mobile or web-based platform, which helps farmers detect a disease on-field with the help of smartphone images.

- **Larger and Diverse Datasets:**

Incorporating larger datasets with diverse environmental conditions, lighting variations, and disease stages can improve model robustness and scalability.

- **Multi-Disease and Multi-Crop Extension:**

The framework is generalized to identify diseases in multiple crops, allowing the development of a comprehensive plant disease diagnosis system.

- **Integration with IoT and Smart Agriculture:**

Integrating the model with IoT-based sensor data (e.g., temperature, humidity) can allow predictive analysis of disease and early warning systems.

- **Enhanced Explainability Techniques:**

Future research can explore advanced explainable AI methods such as SHAP or LIME to provide more detailed and quantitative explanations.

The proposed explainable framework is an augmentation of the functionality of the classification and bridges gap between precision and interpretability, which is why this is a prospective resolution to smart intelligent systems of the future.

REFERENCES

- [1] S. Savary, A. Nelson, A. Sparks, et al., “The global burden of pathogens and pests on major food crops,” *Nature Ecology & Evolution*, vol. 3, no. 3, pp. 430–439, 2019. <https://doi.org/10.1038/s41559-018-0793-y>
- [2] J.G.A. Barbedo, “Plant disease identification from individual lesions and spots using deep learning,” *Biosystems Engineering*, vol. 180, pp. 96–107, 2019. <https://doi.org/10.1016/j.biosystemseng.2019.02.002>
- [3] K.P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018. <https://doi.org/10.1016/j.compag.2018.01.009>
- [4] M. Too, L. Yujian, S. Njuki, L. Yingchun, “A comparative study of fine-tuning deep learning models for plant disease identification,” *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019. <https://doi.org/10.1016/j.compag.2018.03.032>
- [5] A. Krizhevsky, I. Sutskever, G.E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017. <https://doi.org/10.1145/3065386>
- [6] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” *IJCAI*, pp. 1137–1145, 1995.
- [7] R.R. Selvaraju, M. Cogswell, A. Das, et al., “Grad-CAM: Visual explanations from deep networks via gradient-based localization,” *International Journal of Computer Vision*, vol. 128, pp. 336–359, 2020. <https://doi.org/10.1007/s11263-019-01228-7>
- [8] B. Singh, A. Kumar, “Explainable AI in plant disease detection using Grad-CAM,” *Artificial Intelligence in Agriculture*, vol. 5, pp. 15–25, 2021. <https://doi.org/10.1016/j.aiia.2021.02.001>
- [9] Mohanty, S.P., Hughes, D.P., Salathé, M., “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, 2020. DOI: 10.3389/fpls.2020.00941



- [10] ArunPandian, J., et al.,
“Deep convolutional neural networks for rice leaf disease classification,” *Ecological Informatics*, 2021.
DOI: 10.1016/j.ecoinf.2021.101312
- [11] Chen, J., Chen, J., Zhang, D., Sun, Y., Nanehkaran, Y.A.,
“Using deep transfer learning for image-based plant disease identification,” *Computers and Electronics in Agriculture*, 2021.
DOI: 10.1016/j.compag.2021.106443
- [12] Saleem, M.H., Potgieter, J., Arif, K.M.,
“Plant disease classification using deep learning,” *Remote Sensing*, 2021.
DOI: 10.3390/rs13061045
- [13] Ghosal, S., et al.,
“Explainable deep learning for plant stress phenotyping,” *PNAS*, 2021.
DOI: 10.1073/pnas.2020999117
- [14] BarredoArrieta, A., et al.,
“Explainable Artificial Intelligence (XAI): Concepts and challenges,” *Information Fusion*, 2020.
DOI: 10.1016/j.inffus.2019.12.012
- [15] Samek, W., Montavon, G., Vedaldi, A., Hansen, L.K., Müller, K.R.,
“Explainable AI: Interpreting, explaining and visualizing deep learning,” Springer, 2021.
DOI: 10.1007/978-3-030-28954-6

