An Explainable ResNet50-Based Framework for Paddy Crop Disease Detection Using Grad-CAM in Precision Agriculture

Ceena Mathews[1]
[1]*Prajyoti Niketan College, Pudukad, Kerala*

***Abstract*—The early and accurate detection of crop diseases is essential for improving agricultural productivity and ensuring global food security, particularly in staple crops such as rice. In this study, an explainable deep learning framework based on ResNet50 is proposed for classifying paddy crop diseases. The model is trained and evaluated on a real-world dataset comprising ten classes, including healthy leaves. Transfer learning is employed by initializing the network with pretrained ImageNet weights, followed by fine-tuning to enhance classification performance. The proposed model achieves a validation accuracy of** 90.53% **and a weighted F1-score of** 0.905**, indicating robust performance across classes. To address the lack of interpretability in deep learning models, Gradient-weighted Class Activation Mapping (Grad-CAM) is integrated to generate visual explanations of model predictions. The experimental results demonstrate that the model effectively focuses on disease-affected regions in leaf images, thereby validating its decision-making process. The proposed framework provides a balance between high accuracy and interpretability, making it suitable for real-world deployment in precision agriculture systems.**

***Index Terms*—Explainable AI, Grad-CAM, Resnet-50**

I. INTRODUCTION

Rice is one of the most important staple crops worldwide, serving as a primary food source for a significant portion of the global population. However, the productivity of rice crops is severely affected by various diseases such as bacterial leaf blight, blast, and tungro, which can lead to substantial yield losses. Early and accurate identification of these diseases is essential for effective crop management and sustainable agriculture. Traditionally, disease detection has relied on manual inspection by agricultural experts, which is not only time-consuming but also prone to subjectivity and scalability issues.

In recent years, advancements in deep learning have enabled automated image-based disease detection systems that outperform traditional methods. Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated remarkable performance in image classification tasks. However, their “black-box” nature limits their interpretability, making it difficult to understand the reasoning behind their predictions. This lack of transparency is a critical concern in applications such as agriculture and healthcare, where decision reliability and trust are essential. To address this limitation, the concept of Explainable Artificial Intelligence (XAI) has emerged, aiming to make machine learning models more transparent, interpretable, and trustworthy.

XAI techniques provide insights into how input features influence model predictions, thereby enabling users to validate and interpret the results. Among various XAI methods, Gradient-weighted Class Activation Mapping (Grad-CAM) has gained significant attention due to its effectiveness and simplicity in visualizing CNN decisions. Grad-CAM generates class-specific localisation maps by computing the gradients of the target class score with respect to the feature maps of the last convolutional layer. These gradients are used to weight the feature maps, producing a heatmap that highlights the most important regions in the input image that contribute to the prediction.

In the context of crop disease detection, Grad-CAM enables visualisation of disease-affected regions on leaf images, allowing verification of whether the model focuses on relevant areas such as lesions, discoloration, or texture changes. This not only enhances the interpretability of the model but also increases user confidence in its predictions. By integrating Grad-CAM with the ResNet50 architecture, the present work provides both accurate classification and meaningful visual explanations, thereby supporting reliable decision-making in precision agriculture.

II. RELATED WORKS

The application of deep learning in plant disease detection has been widely explored in recent years. Mohanty et al. demonstrated the effectiveness of deep CNNs on the PlantVillage dataset, achieving high classification accuracy across multiple plant species. However, their work primarily relied on controlled laboratory images, limiting its applicability in real-world agricultural scenarios [1]. Ferentinos extended this approach by applying CNNs to a broader dataset containing multiple plant diseases, showing that deep learning models can generalise well when sufficient data is available [2].

In addition, Andreas Kamilaris and Francesc X. Prenafeta-Boldu provided a comprehensive overview of deep learning applications in agriculture, highlighting the potential of CNN-based approaches for crop disease detection [3]. Too et al. conducted a comparative study of deep learning architectures, including VGG, ResNet, and DenseNet, and concluded that residual networks offer superior performance due to their ability to mitigate vanishing-gradient problems [4]. Similarly, Hughes and Salathé explored disease classification using deep learning, highlighting the importance of large datasets for improved generalization. More recent studies have focused on real-world datasets that include variations in illumination, background clutter, and leaf orientation, thereby making the problem more challenging and realistic [5].

Furthermore, J. G. A. Barbedo showed that dataset size and diversity significantly influence the performance of deep learning models, reinforcing the need for varied and representative training data [6].
Explainability in deep learning has also gained significant attention. Selvaraju et al. introduced Grad-CAM, which enables visualisation of class-specific regions in images, thereby providing interpretability to CNN-based models. Subsequent works have applied Grad-CAM in agricultural contexts to verify whether models focus on disease-relevant regions [7]. R. P. Porfírio et al. evaluate multiple XAI techniques, including Grad-CAM, LIME, and occlusion methods. It demonstrates that Grad-CAM provides clear visual explanations for CNN predictions in plant disease classification tasks [8]. S. M. Alhammad et al. applied Grad-CAM to CNN-based models for potato leaf disease classification. The study shows that Grad-CAM effectively highlights disease-affected regions and validates that the model focuses on meaningful features rather than background noise [9]. In addition, recent survey work by L. Li et al. emphasises the growing importance of integrating explainability techniques into plant disease detection systems [10].
Despite these advancements, many existing studies lack comprehensive class-wise evaluation and detailed analysis of model behaviour using explainability methods. The present work addresses these limitations by combining high classification performance with interpretability and detailed evaluation

III. MATERIALS AND METHODS

The publicly available dataset used in this study consists of paddy leaf images belonging to multiple disease classes along with healthy samples. The images were collected under real-world conditions, incorporating variations in illumination, background, and leaf orientation.

The dataset consists of a total of 2081 paddy leaf images distributed across 10 classes. The class-wise distribution includes 86 images of bacterial leaf blight, 79 images of bacterial leaf streak, 71 images of bacterial panicle blight, 336 images of blast, 196 images of brown spot, 280 images of dead heart, 121 images of downy mildew, 343 images of hispa, 339 images of normal leaves, and 230 images of tungro. This distribution indicates a moderate class imbalance, with certain classes, such as hispa, normal, and blast, having a higher number of samples compared to bacterial disease classes. Such an imbalance may influence the classification performance, and therefore, class-wise evaluation metrics are considered in this study to ensure a comprehensive assessment of the proposed model. Each image is resized to a fixed resolution of 224 × 224 pixels to match the input requirements of the ResNet50 architecture.

The proposed framework for paddy crop disease detection consists of five main stages: input image acquisition, data preprocessing and augmentation, feature extraction using a pretrained ResNet50 model, classification through fully connected layers, and model interpretability using Grad-CAM. The overall architecture of the proposed system is illustrated in Fig. 1.

To improve model generalisation and reduce overfitting, several data augmentation techniques are applied. These include random rotations, horizontal and vertical flipping, zoom transformations, and colour jittering. These operations introduce variability in the training data and help the model become robust to changes in orientation, illumination, and scale. All images are normalised using standard preprocessing corresponding to the pretrained ResNet50 model.

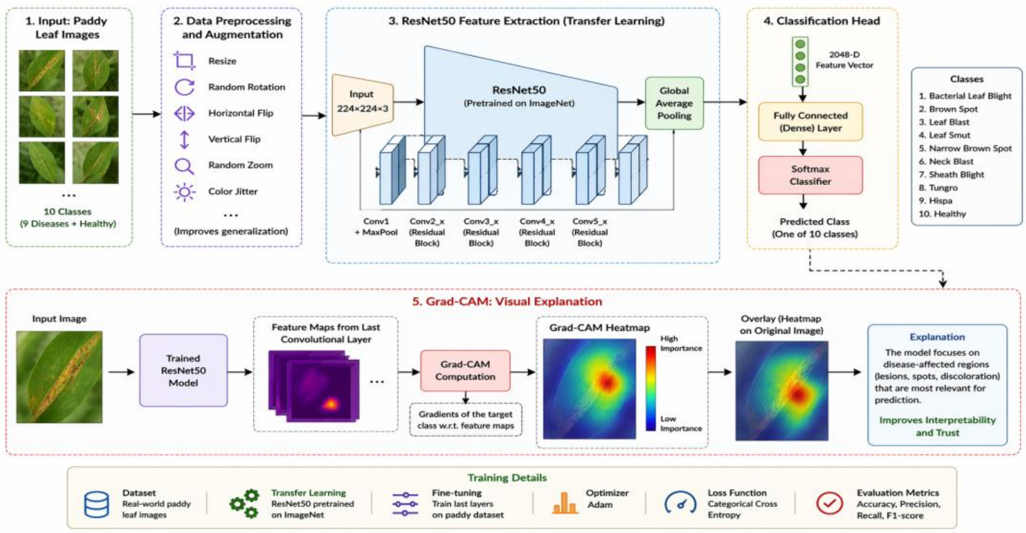


Fig 1. Proposed ResNet50-based paddy disease detection framework with Grad-CAM visualisation.

A deep convolutional neural network based on ResNet50 is employed for feature extraction. The model is initialized with weights pretrained on the ImageNet dataset, enabling transfer learning. The ResNet50 architecture consists of multiple residual blocks that facilitate the learning of both low-level features, such as edges and textures, and high-level semantic representations. The final convolutional feature maps are passed through a global average pooling layer to generate a compact 2048-dimensional feature vector.

The extracted feature vector is fed into a fully connected layer followed by a softmax classifier to perform multi-class classification. The softmax layer outputs the probability distribution across all disease classes, and the class with the highest probability is selected as the predicted label. The model is trained with categorical cross-entropy loss and optimised using the Adam optimiser.

To enhance interpretability, Grad-CAM is applied to the trained model. Grad-CAM computes the gradients of the target class score with respect to the feature maps of the last convolutional layer. These gradients are globally averaged to obtain channel-wise importance weights, which are then combined with the feature maps to generate a class-discriminative heatmap. The heatmap highlights regions in the input image that contribute most significantly to the model's prediction.

The generated heatmap is resized to match the original image dimensions and overlaid on the input image to produce a visual explanation. This allows verification of whether the model focuses on disease-relevant regions such as lesions, discoloration, or texture abnormalities. The integration of Grad-CAM thus improves the transparency and reliability of the proposed system.

IV. RESULTS AND DISCUSSION

The input image is first passed through the convolutional layers of the ResNet50 model to extract hierarchical feature representations. These feature maps capture spatial and semantic information relevant to disease characteristics. Subsequently, the gradients of the predicted class score are computed with respect to the feature maps of the last convolutional layer. These gradients are globally averaged to obtain channel-wise importance weights, which indicate the contribution of each feature map to the final prediction. The weighted combination of feature maps produces a coarse localisation map, which is then passed through a ReLU activation function to retain only positive contributions. The resulting heatmap highlights the most discriminative regions in the input image that influence the model's decision. Finally, the heatmap is upsampled and superimposed on the original image to provide a visual explanation of the prediction.

The class-wise evaluation is shown in the table. 1 reveals that the model performs exceptionally well for classes such as dead_heart and bacterial_leaf_streak, with F1-scores exceeding 0.92. However, the performance for bacterial_leaf_blight is comparatively lower, with an F1-score of approximately 0.78, indicating that this class is more challenging to distinguish. The proposed model achieves a validation accuracy of 90.53% and the weighted F1-score of 0.90, indicating consistent performance across classes.

The confusion matrix in Fig. 2 illustrates the model's performance, with strong diagonal dominance indicating a high number of correct predictions. Nevertheless, some misclassifications are observed between visually similar diseases, particularly among bacterial disease categories and between brown_spot and blast. These results suggest that while the model is effective, there is room for improvement in distinguishing closely related classes.

The Grad-CAM visualization process is illustrated in Fig.3. The highlighted regions in the figure indicate that the model primarily focuses on disease-affected areas such as lesions and discoloration, thereby validating its prediction.

Table 1. Class-wise performance of the proposed model

Class	Precision	Recall	F1-Score
Bacterial Leaf Blight	0.840	0.733	0.783
Bacterial Leaf Streak	0.947	0.899	0.922
Bacterial Panicle Blight	0.941	0.901	0.921
Blast	0.903	0.914	0.908
Brown Spot	0.886	0.872	0.879
Dead Heart	0.961	0.964	0.963
Downy Mildew	0.922	0.884	0.903
Hispa	0.884	0.889	0.887
Normal	0.898	0.935	0.916
Tungro	0.889	0.909	0.899

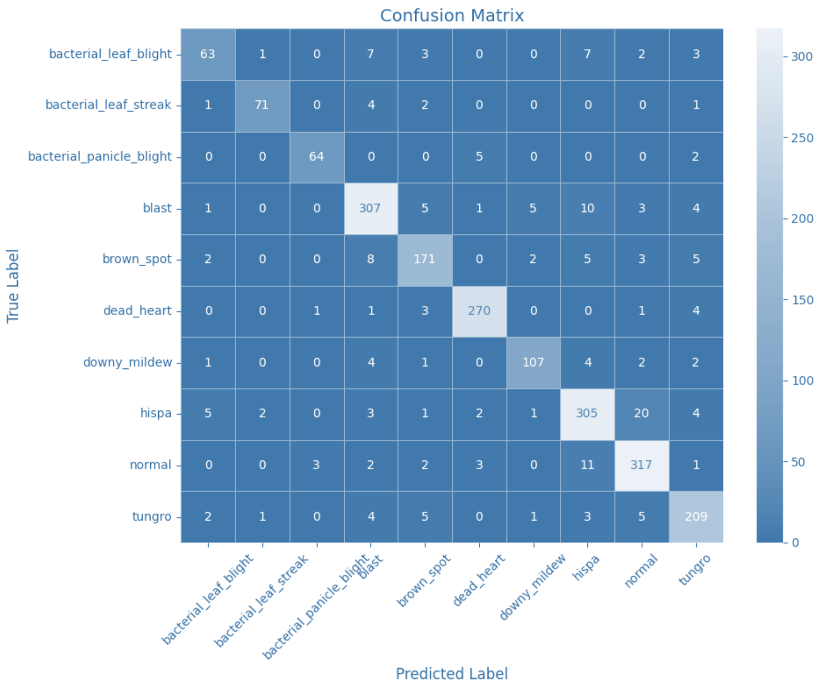
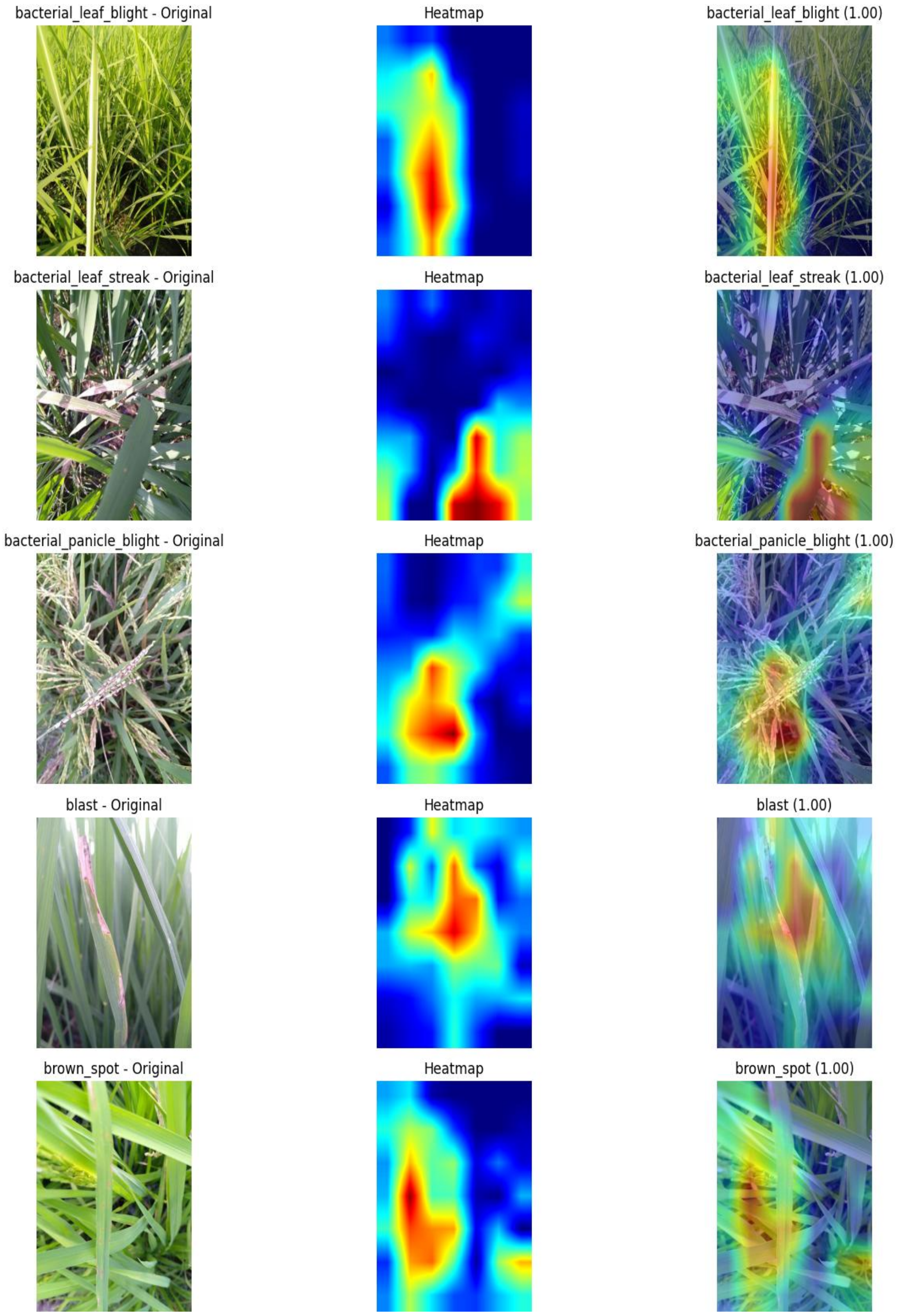


Fig. 2. Confusion matrix of the proposed ResNet50-based paddy disease classification model showing class-wise prediction performance.



a)Original b) Grad-CAM heatmap c) Overlay Visualisation

Fig. 3. *Grad-CAM visualization illustrating the explainability process of the proposed ResNet50 model.*

V. CONCLUSION

This study presents an explainable deep learning framework for paddy crop disease detection using ResNet50 and Grad-CAM. The model achieves high classification accuracy while providing interpretable visual explanations. The results demonstrate that the proposed approach is effective for real-world agricultural applications. Future work will focus on improving classification performance for challenging classes, incorporating larger datasets, and deploying lightweight models for mobile-based disease detection systems.

REFERENCES

[1] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, 2016.

[2] K. P. Ferentinos, “Deep learning models for plant disease detection,” Computers and Electronics in Agriculture, 2018.

[3] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Computers and Electronics in Agriculture, 2018.

[4] E. Too et al., “A comparative study of fine-tuning deep learning models for plant disease identification,” *Computers and Electronics in Agriculture*, 2019.

[5] D. Hughes and M. Salathé, “An open access repository of images on plant health,” arXiv, 2015.

[6] J. G. A. Barbedo, “Impact of dataset size and variety on deep learning,” *Computers and Electronics in Agriculture*, 2018.

[7] R. R. Selvaraju et al., “Grad-CAM: Visual explanations from deep networks,” ICCV, 2017.

[8] R. P. Porfírio et al., “Exploring Explainable AI Techniques for Plant Disease Classification,” Procedia Computer Science, 2025.

[9] S.M.Alhammad,et al., “Deep learning and explainable AI for classification of potato leaf diseases,” Frontiers in Artificial Intelligence, 2025.

[10] L. Li et al., “Deep learning for plant disease detection: A review,” Plant Methods, 2021.