



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Praf-Net: A Pathology Relational Attention Fusion Network for Fine-Grained Rice Leaf Disease Classification and Severity-Aware Analysis

CB. Sudhersun¹ and *S.P. Balamurugan²

¹Research Scholar, Department of Computer and Information Science, Annamalai University, Tamil Nadu, India.

²Assistant Professor/Programmer, Department of Computer and Information Science, Annamalai University, Tamil Nadu, India.
Corresponding Author Email: spbcdm@gmail.com

Abstract

The reduced crop yield and grain quality due to rice leaf diseases have a great impact on agricultural productivity and food security. Due to complex lesion texture, intra-class variability, inter-class similarity and different severities of infected leaf regions, the presence of visually similar rice leaf diseases is still a challenging problem. To overcome these challenges, this paper proposes a novel Pathology Relational Attention Fusion Network (PRAF-Net) for fine-grained rice leaf disease classification and severity-aware analysis. This proposed framework combines handcrafted pathology descriptors, convolutional texture representations, transformer-guided contextual embeddings within a unified pathology-aware fusion architecture for learning graph-based disease relational and severity-aware auxiliary representation separately. Here, rice leaf images are preprocessed with data augmentation and normalization to help generalize well and avoid overfitting. To detect and extract disease specific structural features PRAF-Net employ some handcrafted pathological features such as Gray-Level Co-occurrence Matrix (GLCM), Local Binary Pattern (LBP), color statistics, lesion shape descriptors. On the other hand, EfficientNetB0 captures local texture-aware features and MobileViT gathers global contextual dependencies and disease spread patterns. The combination and optimization of the extracted multimodal features are performed using feature normalization and PCA. In order to distinguish between visually similar diseases, a Graph Attention Network (GAT)-based relational learning module is presented to study the semantic relationships for disease categories. And moreover, a severity-aware auxiliary learning is introduced to model disease progression characteristics and improve pathology representation learning. The Pathology Relational Attention Fusion proposed in this work, combines the favourable characteristics of multi-head attention and adaptive gating-based mechanisms to combine all texture, contextual, relational as well as severity-aware embeddings learned in a diagnostic task for better robustness against unseen disease classification. Experimental evaluation on a six-class rice leaf disease dataset shows the efficacy of PRAF-Net, providing an accuracy of 97.44%, precision of 97.43%, recall of 97.44% and f1-score of 97.43%, which outperforms several competing state-of-the-art CNN, transformer and hybrid deep learning models in literature. In addition, large-scale ablation studies and explainability analyses confirm that the corresponding sites of proposed modules also contribute to performance improvements. Moreover, a real-time GUI-based deployment system is created to support practical disease diagnosis and severity visualization for intelligent precision agriculture applications.

Keywords - Rice Leaf Disease Classification, Fine-Grained Disease Recognition, Graph Attention Network, CNN-Transformer Hybrid Learning, Severity-Aware Representation Learning, Explainable Artificial Intelligence, Precision Agriculture

1. Introduction

Rice is one of the most widely distributed staple food crops in world, contributing strongly to global food security and agricultural sustainability. But rice productivity is broadly limited by a plethora of leaf diseases, including Bacterial Leaf Blight, Brown Spot, Leaf Blast, Leaf Scald and Narrow Brown Spot that considerably reduce grain yield and quality. Thus, an early and precise diagnosis of rice leaf diseases is necessary for the effective management and precision agriculture. Agricultural experts typically employ the traditional method for disease diagnosis, which is labor intensive, time-consuming, subjective in nature, and strongly reliant on visual expertise. Artificial intelligence and deep learning techniques have presented themselves as a robust solution for automated plant disease identification, because they have better predictive performance compared to earlier plant-disease recognition methods based on various imaging approaches.

In the beginning, conventional machine learning methods with handcrafted features achieved good performance in the detection of rice diseases. Introduction Sethy et al. [1] proposed a framework for rice leaf disease identification using support vector machines and pretrained deep neural representations. The utilization of transfer learning and feature-based classification for agricultural disease diagnosis were shown to be well-applied in their work. Later, Latif et al. [2] proposed a rice disease detection model using an improved convolutional neural network (CNN) which enhanced feature learning capability and classification performance within controlled imaging ambience. In a similar vein, in paper [3], Aggarwal et al. studied feature selection approaches based on pretrained deep neural networks along with machine learning classifiers for rice leaf disease classification and demonstrated that optimized representation of features improves the prediction success rates. Though these methods had shown some promise, they did not successfully tackle the problem of taking contextual and pathology-aware reasoning into account as they were either limited to executing analyses based on handcrafted descriptors or CNN-based local feature learning and typically could not differentiate visually similar disease categories. As deep learning architectures have advanced, researchers have begun to incorporate hybrid feature learning and deep representation fusion for agricultural disease analysis. DeepRice framework integrated the deep feature representation with the use of deep learning for classification of rice leaf disease subtype [4]. They proved that deep feature fusion could aid in performances of fine-grained disease recognition. Similarly, Akter et al. [5] used convolutional neural network architectures for the robust classification of rice leaf diseases and demonstrated the feasibility of deep CNN models in achieving key features of lesions-inflicted pathological properties. Moreover, Dogra et al. [6] proposed a radical smart agriculture oriented deep learning based framework for brown spot rice disease identification which underpins the practice bound generalisability of intelligent disease diagnosis systems in precision farming scenarios. However, our CNN-based approaches are mostly limited to regions characterized by local texture patterns and often lack the ability to leverage long-range contextual dependencies, disease spread distributions, or inter-class semantic relationships that are critical for discerning highly similar disease manifestations.

In particular, transformer architectures and ensemble learning models have shown dramatic gains in tasks of visual recognition which can exploit global contextual information. An ensemble CNN-based deep learning framework for automatic rice leaf disease diagnosis was proposed by Pai et al. [7], which extracted multi-model features enabling enhanced robustness. Likewise, as illustrated by Hu et al. [8], a deep multiscale feature learning strategy was devised for the identification of rice leaf diseases, confirming that such hierarchical features can represent information beneficial for classification. Although transformer-guided and multiscale learning approaches have promoted global feature extraction for voxel-level histopathology, existing methods still face limitations with respect to 1) incomplete integration of pathology-aware features; 2) under-operation of disease coupling; 3) poor explainability; and 4) insufficient ability for severity-aware learning. Moreover, the state-of-art frameworks often regard various disease classes independently from one another and overlook semantic relationship between visual overlapping types of diseases. Moreover, severity analysis is frequently overlooked yet it represents a crucial aspect for agricultural decision making and disease management.

In this paper, a new PRAF-Net is proposed to classify and analyze the severity of rice leaf diseases. In contrast to prior works that are trained with CNNs only or transformer architectures, the PRAF-Net introduces a multiscale pathology-aware fusion framework that combines handcrafted pathology descriptors, CNN-based texture representations, transformer-guided contextual embeddings, graph-based disease relational learning and severity-aware auxiliary representation learning into one model. First, handcrafted feature descriptors are extracted including GLCM, LBP, color statistics and lesion shape descriptors to represent the pathology-specific characteristics of a disease. While at the same time EfficientNetB0 features fine-grained texture-aware representations, MobileViT retains long-range contextual dependencies and further learns how diseases spread distributions. Next, the fusions of multimodal features are normalized by Feature Normalization or Principal Component Analysis.

In addition, to enhance the discrimination of visually similar diseases, a GAT-based disease relational learning module first performs semantic dependency modeling upon disease categories is added in the proposed framework. Moreover, a severity-aware auxiliary learning strategy is utilized to model the features of disease progression and improve pathology representation learning. This paper introduces the PRAF module that utilizes multi-head attention and adaptive gating mechanisms to effectively combine fused texture, contextual, relational, and severity-aware embeddings for improved disease recognition. Experimental results on a six-class rice leaf disease dataset over existing state-of-the-art CNN, transformer and hybrid deep learning models show that the proposed network (PRAF-Net) outperforms existing models with respect to accuracy, precision, recall and F1-score. Finally, validate the effectiveness, interpretability and practical application of proposed framework in intelligent precision agriculture systems through Explainability analyses, ablation studies and real-time GUI-based deployment.

2. Related Works

2.1 Traditional Machine Learning and Handcrafted Feature Methods

In the literature, early methods for diagnosing rice leaf disease mainly used handcrafted feature extraction and shallow machine learning approaches. In [9], Jiang et al. presented a rice leaf disease recognition framework, combining deep feature extraction with Support Vector Machine (SVM) classification. They proved that transfer-learned deep representations can greatly boost the performance of traditional handcrafted descriptors (HCD) toward disease recognition. However, this method was primarily based on classification accuracy and couldn't model the relationship among disease classes in the context. In a similar manner, Anami et al. [10] proposed a deep learning-based framework for recognizing yield-affecting paddy crop stresses from field images. While their approach demonstrated the feasibility of applying deep feature learning under realistic agricultural conditions, it was constrained by binary classification and a lack of fine-grained disease reasoning mechanisms. Li et al. [11] focused on the diagnosis and application of rice diseases employing deep learning methods using an automatic disease identification system for agriculture surveillance. Although their performances performed well, they primarily relied on feature-level discriminative properties and failed to capture the semantic associations between diseases/phenotypes, patterns of severity progression perceived by clinicians over time and context-statistical dependencies with similar disease categories.

2.2 CNN-Based Rice Disease Classification

Convolutional Neural Networks (CNNs) are the state-of-the-art for rice disease classification as they leverage powerful feature extraction. Comparing study using CNN, Cascading Autoencoder with Attention Residual U-Net (CAAR-U-Net), and MobileNet-V2 architectures for rice leaf disease classification was done by Dutta et al. [12]. They showed that deep CNN models capture disease-specific lesion traits. Under the umbrella of Green AI, Saddami et al. [13] introduced various lightweight CNN architectures to effectively identify rice diseases with reduced computational effort, while achieving competitive classification performance. This would be an ideal approach where resources are scanty in agricultural environments. In particular, Gulmez [14] conducted a detailed literature review on rice disease detection techniques employing CNNs and showed that convolutional feature learning leads to impressive improvements. While CNN-based methods are able to learn local texture and proximity information from lesion patterns, they are unable to model distant contextual dependencies and semantic relations across disease categories, ultimately preventing their applicability in fine-grained disease recognition scenarios.

2.3 Hybrid and Ensemble Deep Learning Methods

To extend the capabilities of single deep learning models, hybrid architectures that combine encoder-decoder networks with the feature extractors have gained some interest. Automated Hybrid Deep Learning framework for identifying and classifying the paddy leaf disease combining features of various learning mechanisms to improve recognition performance has been proposed by Subbarayudu and Kubendiran [15]. They showed that stacking heterogeneous representations can lead to significantly stronger robustness to classification changes. Likewise, Pai et al. [16] proposed an ensemble convolutional neural network (CNN)-based framework pooling predictions of several convolutional models to obtain a higher classification performance. Ensemble learning reduced biases inherent in individual models and improved prediction stability to also improve generalization performance. In [17] Prity et al. compared feature-based and direct imaging neural network model for rice leaf disease recognition. The result had shown that combining different feature representations can significantly increase classification performance. However, current hybrid and ensemble frameworks mainly emphasize feature aggregation and classification accuracy but do not consider modeling disease relationships, pathology-aware representation learning, and estimating disease severity.

2.4 Attention and Transformer-Inspired Architectures

In the agricultural image analysis, attention mechanisms have played quite a role in feature representation learning. Stephen et al. [18] presented a ResNet architecture utilizing self-attention mechanisms for rice leaf disease classification, finding that attention can further improve feature discrimination by highlighting more disease-relevant regions. When they tested their model, it outperformed conventional CNN architectures for classification purposes. In a similar way, Pandiyaraju et al. [19] implemented a channel-attention-based hybrid CNN framework for paddy leaf disease detection based on channel-wise attention to improve sensitive-feature extraction of diseases. This model significantly enhanced localization and classification accuracy of these lesions. More recently, Mia et al. [20] proposed an angular-compactness

dual loss learning strategy that uses the new distance between leaf features to fine-grained rice leaf disease detection. Their work was primarily aimed at improving the intra-class compactness and inter-class separability with a focus on loss optimization for more optimal feature space encoding. Despite improvements of feature representation and discrimination capability brought by attention-based and transformer-inspired models, the previous works still mainly follow independent disease classification paradigms which do not model inter-disease semantic relations or severity progression traits explicitly.

2.5 Object Detection-Based Disease Diagnosis

In addition to image classification, object detection techniques have received significant attention with regards to localizing a disease and facilitating real-time diagnosis. In [21], RLDD-YOLOv11n, a lightweight object detection framework for rice leaf disease detection, is proposed by Fang et al. Using YOLO architecture, the model had effective disease localization and real-time inference which was suitable for smart agriculture based applications. Similarly, Kokila et al. [22] proposed an image based deep learning approach exploiting YOLOv5, DenseNet201 and ResNet101 to detect rice diseases. They showed that combining object detection and classification strategies is effective for disease diagnosis by using their framework. While object detection approaches provide enough spatial information as ground truth labels of locality regarding disease locations, most state-of-the-art works are primarily focused on modeling the localization of a lesion, ignoring contextual pathology reasoning, disease relationship learning and severity-aware analysis.

2.6 Reviews and Research Trends

Recent review studies have systematically described the advances and challenges in rice disease diagnosis. Seelwal et al. [23] provided an extensive review of the applications of deep learning in rice disease diagnosis and noted the major trends in research topics which included transfer learning, light weight networks, attention mechanisms, and explainability artificial-intelligence. They also noted issues related to dataset diversity, generalizability of the proposed models, and actionable deployment. In addition, recent advances in deep learning methods for rice disease recognition were examined by Yusuf et al. [24], highlighting changes towards CNNs, transformers, ensemble learning, and hybrid architectures. The recommendations made in both reviews strongly indicated that explainable methods, merging multimodal features, estimating severity level and context-aware disease diagnosis frameworks are important factors for improving robustness and real-world applicability of computable clinical decision support.

2.7 Research Gap and Motivation for PRAF-Net

The literature review indicates that satisfactory success has been accomplished for rice leaf disease classification with machine learning, CNNs, hybrid deep learning models, attention mechanism and object detection framework. Nevertheless, there were some key limitations that are still to be addressed. First, existing methods are mainly aimed at classification performance and pay too little attention to explicitly modeling the semantic relationship of disease categories. Second, most CNN-based methods learn local texture feature information and lack a representation of global contextual diseases pattern dependencies. Thirdly, while the existing pipeline of hybrid and ensemble frameworks fuses features in a fully supervised manner without any pathology-aware relational reasoning mechanism. Fourth, severity estimation has hardly been investigated from a practical perspective to separate the role of disease management and precision agriculture. Finally, current attention based approaches tend to highlight only salient feature areas and do not incorporate texture, contextual, relational and severity-aware representations into an integrated learning mechanism.

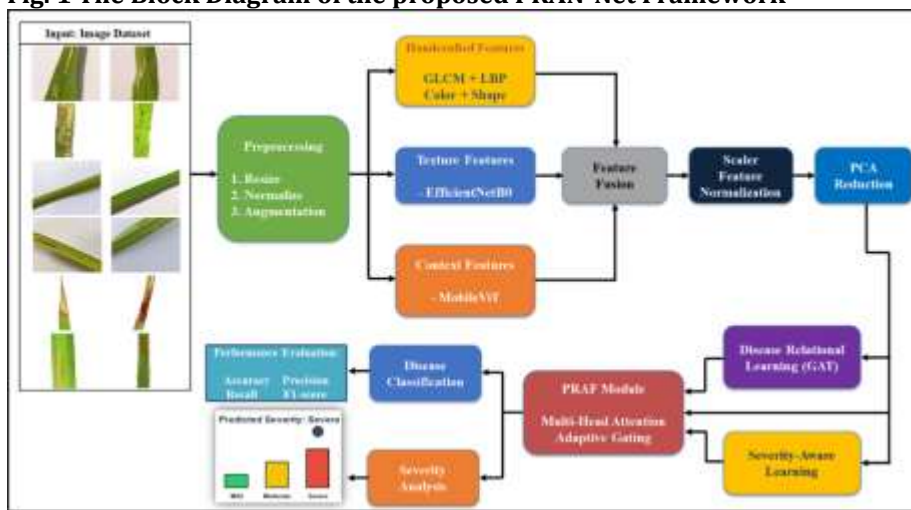
To overcome the aforementioned limitations, this study proposes a novel PRAF-Net for fine-grained rice leaf disease classification and severity-sensitive analysis. A unified pathology-aware attention fusion framework that combines handcrafted pathology descriptors, CNN-based texture learning, transformer-guided contextual representation learning, graph-based disease relational modeling and severity-aware auxiliary learning. To mitigate these challenges, PRAF-Net combines the locality properties of local lesions with global contextual dependencies, semantic disease relationships, and temporal disease progression features to provide a novel solution for intelligent rice disease diagnosis in precision agriculture that is more interpretable and accurate than existing ones.

3. The Proposed Model

The new fine-grained rice leaf disease classification and severity-aware analysis framework proposed in this paper, PRAF-Net, integrates the functions of context-dependent feature learning based on pathology stepwise behaviors, contextual relationship representation reasoning for multiple varieties of diseases with different severities co-occurrence relations, and adaptive attention fusion using relational graph

convolution neural networks (GCN). Initially, raw rice leaf images are preprocessed and image augmentation is applied to retain and enhance disease-specific pathological features without compromising data generalization. Then, handcrafted pathology descriptors GLCM, LBP, color statistics and lesion shape features are extracted to provide structural and textural disease patterns. On the other hand, EfficientNetB0 will receive local texture-aware representations and MobileViT learns the global contextual information and disease spread characteristics based on leaf images concurrently. Multimodal features extracted are then fused and optimized by feature normalization and PCA-based dimensionality reduction. In addition, a GAT-based Disease Relational Learning Module forms semantic connections between visually related classes of diseases, and also a Severity-Aware Auxiliary Learning Module captures the disease evolution patterns and characteristics of active infections. The subsequent proposed PRAF module then uses multi-head attention and adaptive gating mechanisms to automatically fuse texture, contextual, relational, and severity-aware embeddings into one pathology representation dynamically. Finally, a deep classification network is applied to the fused representation for six rice leaf disease classification and severity estimation to provide an accurate, explainable, and robust diagnosis of different diseases while preventing potential errors in intelligent precision agriculture systems. The overall structure of our proposed PRAF-Net framework is shown in Figure 1.

Fig. 1 The Block Diagram of the proposed PRAN-Net Framework



3.1 Rice Leaf Disease Dataset

The quality, diversity and balance of the training dataset yields tremendous influence over the performance and generalization ability of any intelligent plant disease diagnosis system. A publicly available rice leaf disease dataset [25] was collected from the Kaggle repository in this study to evaluate the efficacy of the proposed PRAF-Net. Dataset Given consists of 2580 RGB rice leaf images belonging to six classes (Bacterial Leaf Blight, Brown Spot, Healthy, Leaf Blast, Leaf Scald and Narrow Brown Spot) The dataset is perfectly balanced, as each category has 430 images; hence it helps in avoiding class bias while training and evaluating the model.

Disease classes were chosen between commonly existing rice leaf diseases having high visual similarity with respect to lesion texture, color variation, infection spread and structural deformation. As a result, discriminating among these diseases correctly requires local and global context for each pathology analysis, making the dataset perfect to benchmark in this fine-grained disease classification regime. In addition, the availability of healthy class assists to facilitate classification model for distinguishing infected leaves from non-infected samples which will broaden practical uses in precision agricultural settings. Let the rice leaf image dataset be represented as

$$D = \{(I_i, y_i)\}_{i=1}^N \quad (1)$$

Where I_i denotes the i^{th} rice leaf image, y_i represents the corresponding disease label, and $N = 2580$ indicates the total number of images.

The number of images in each category is summarized in Table 1. Such balanced dataset structure establishes a strong base for implementing and validating the proposed PRAF-Net framework that aims to perform fine-grained rice leaf disease classification along with severity aware analysis.

Table 1. Rice Leaf Disease Dataset Description

Class	Number of Images
-------	------------------

Bacterial Leaf Blight	430
Brown Spot	430
Healthy	430
Leaf Blast	430
Leaf Scald	430
Narrow Brown Spot	430
Total	2580

3.2 Dataset Splitting and Data Augmentation

The rice leaf disease dataset was partitioned using a stratified sampling strategy to guarantee biased model evaluation and to ensure that there was no leakage of information between the training and testing stages. Stratification ensures that the initial proportion of classes is preserved in all the subsets, thus class balance remains intact throughout the learning process. The entire dataset of 2580 images was divided into subsets for training (70%), validation (15%), and testing (15%). Accordingly, for each disease class, 301 images were designated to the training set, with 134 of these either serving the dual-purpose or they split into group validation (64 images) and testing sets (65 images).

One of the most important aspects of this experimental design is that dataset splitting was done before augmentation, so augmented samples generated from a particular image never occupy the validation or testing set. It is also a simple way that eliminates data leakage and ensures a fair estimate of the capacity of the model to generalize. Data augmentation were only conducted on training subset to among other things aim at increasing data diversity and improving robustness of the models against variations in image acquisition condition, disease manifestation and environmental conditions.

The augmentation pipeline consists of three categories of transformations:

3.2.1 Geometric Augmentation

Geometric augmentation used to enhance spatial invariance and generalization ability of proposed PRAF-Net model with different geometric transformations eg. These transformations mimic the variations in image acquisition conditions, leaf orientations, camera viewpoints and scale that are likely to exist in practical agricultural settings. Back to this study, the augmentations were restricted to horizontal flipping, vertical flipping random rotation ($\pm 15^\circ$) and random zooming for training data only.

These geometric transformations retain the pathological features of rice leaf diseases and increase the diversity of training samples, which reduces overfitting and improve robustness against learned feature representations.

3.2.2 Photometric Augmentation

A photometric augmentation technique is used to make the proposed PRAF-Net model more resilient to illumination, contrast, camera sensor properties and environmental conditions when acquiring real images. In contrast to geometric transformations, photometric augmentation remaps the intensity and color attributes of an image but retains its spatial configuration and pathological traits. Data augmentation methods. In this study, brightness adjustment, contrast enhancement, CLAHE and moderate Gaussian blurring were applied only to the training dataset.

Such photometric transformations can make realistic appearance variations while keeping the underlying disease characteristics constant, effectively assisting the model with learning illumination-invariant and contrast-robust feature representations for accurate rice leaf disease classification.

3.2.3 Advanced Augmentation

Data augmentation strategies such as MixUp and CutMix were also included in the training pipeline to enhance robustness of the generalization ability of PRAF-Net framework. In contrast to traditional augmentation approaches which alter a single image, these methods create new training samples by combining knowledge from multiple images in order to drive the model towards learning more discriminative and generalised representations of the disease.

For MixUp augmentation, two randomly selected training samples (x_i, y_i) and (x_j, y_j) are linearly combined as

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (2)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (3)$$

Where $\lambda \in [0,1]$ is sampled from a Beta distribution. This smoothes the decision boundaries and gives a generalization advantage to the model.

Similarly, CutMix augmentation creates an augmented training sample by removing random-sized patch from the input image and replacing it with a random-selected patch of another image, potentially represented as

$$\tilde{x} = M \odot x_i + (1 - M) \odot x_j \quad (4)$$

Where M denotes a binary mask and $\odot$ represents element-wise multiplication. The associated target label is calculated based on the same interpolation approach as described in Equation (3), in which mixing coefficient λ is determined according to the area ratio of the preserved image region. CutMix enhances robustness against occlusion and encourages learning of discriminative disease characteristics across a spatial distribution by exposing the model to mixed pathological patterns and regions containing no lesions.

Importantly, augmentation operations are not leveraged on the validation and testing subsets. So only image re-sizing and normalization were performed on these subsets to make sure that the provided evaluation meets the definition of being an unseen condition with respect to the proposed PRAF-Net framework. Once the dataset is prepared, rice leaf images are preprocessed to bring uniformity in image representation and improve the effectiveness of feature extraction as explained in next subsection.

3.3 Image Preprocessing

The first step in the proposed PRAF-Net Framework is Image Preprocessing, which serves the purpose of normalizing input images and enables appropriate feature extraction from various rice leaf samples. Given that images acquired in real-life agricultural settings can vary considerably in their resolution, illumination, orientation and acquisition conditions prior to feature learning a preprocessing pipeline was developed to standardize all samples.

First, a rice leaf image was resized to an invariable spatial resolution of (224 x 224) pixels as needed as an input for both the EfficientNetB0 and MobileViT feature extraction networks. Image resizing guarantees that the features are uniform across batches, whilst keeping computational complexity low while both training and inferring with the model. Subsequently, pixel intensity normalization was performed to scale the RGB image values into the range [0,1]. The normalized image I_{norm} is computed as

$$I_{\text{norm}}(x, y, c) = \frac{I(x, y, c)}{255} \quad (5)$$

Where $I(x, y, c)$ represents the original pixel intensity at location (x, y) in channel c , and $I_{\text{norm}}(x, y, c)$ denotes the corresponding normalized pixel value. This normalization regularizes gradient updates, speed up convergence and numerical stability during deep feature learning.

No augmentation operations were applied to the validation and testing data, which is worth noting. None of the unseen samples were augmented, resized and normalized to ensure that evaluation results give a clear understanding of how well this model generalizes to unseen data in real world applications.

The subsequent images obtained from preprocessing are used as the multimodal features, handcrafted pathology descriptors, CNN-based texture representations and transformer-guided contextual embeddings for generating comprehensive disease-specific feature representation built in the following subsection.

3.4 Multi-Modal Feature Extraction

Rice leaf diseases are the most critical rice disease in Bangladesh, and Accurate discrimination of them requires comprehensive representation learning that can learn not only low-level pathological characteristics but also high-level semantic information. When extracting features entire categories of disease that look similar will lose out on key disease specific information relating to facial movement due to the reliance on a single feature extraction strategy. In response to this challenge, this study introduces a PRAF-Net framework, which utilizes a multi-modal feature extraction approach that combines handcrafted pathology descriptors, CNN-based texture representations and transformer-guided contextual embeddings. The handcrafted features codify certain statistical, color, texture, and morphological characteristics of disease lesions; an EfficientNetB0 extracts detailed local texture patterns while a MobileViT learns global contextual dependencies relevant to disease spread and structural changes. The proposed framework incorporates these complementary feature modalities to build a construct that is disease-representative, discriminative, and advantageous for classification performance as well as facilitating subsequent pathology relational learning and severity-aware analysis.

3.4.1 Handcrafted Pathology Features

These strategy relies on extracting handcrafted pathology features to explicitly represent the appearance-driven visually-distinct patterns indicative of disease that correlate with patterns pertaining to image texture, color change, and morphological abnormalities. In contrast to deep features which are learned automatically during training, handcrafted descriptors offer interpretable pathological information that

adds complementary learned representations for improved disease discrimination. In the proposed PRAF-Net framework, GLCM-prominent features, LBP features, color statistical descriptors, and shape-based features are extracted from each rice leaf image after preprocessing.

GLCM-based statistical measures are used to characterize texture information for the spatial relationship among neighboring pixel intensities. Among the extracted descriptors, contrast quantifies local intensity variations and is defined as

$$\text{Contrast} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-j)^2 P(i,j) \quad (6)$$

Where $P(i,j)$ is the normalized Gray-Level Co-occurrence Matrix (GLCM), and N_g denotes the number of gray levels. Similarly, homogeneity measures the closeness of the distribution to the diagonal of the GLCM matrix and is expressed as

$$\text{Homogeneity} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{P(i,j)}{1+(i-j)^2} \quad (7)$$

These descriptors effectively capture lesion texture patterns and disease-induced structural irregularities. By comparing the intensity of neighbouring pixels with central pixel, Local Binary Pattern (LBP) features are extracted to represent local texture variations. The LBP operator is expressed as given below

$$\text{LBP}_{p,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \quad (8)$$

Where $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$, and

g_c is the center pixel intensity, g_p is the neighboring pixel intensity, P denotes the number of neighbors, R represents the radius. LBP enables the capture of lesion boundaries, spot distributions and local lesion textures.

Moreover, to describe the disease-associated discoloration patterns, color statistical features are extracted from RGB image channels. The average color intensity of a channel is calculated as

$$\mu_c = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N I_c(x,y) \quad (9)$$

Where $I_c(x,y)$ denotes the pixel intensity of channel c , M and N represent image dimensions. These characteristics provide insights about chlorosis, necrosis and changes in pigmentation due to various diseases.

Shape descriptors are further used to describe lesion morphology and infection progression. The lesion area is calculated as

$$A = \sum_{x=1}^M \sum_{y=1}^N M(x,y) \quad (10)$$

Where $M(x,y) = \begin{cases} 1, & \text{lesion pixel} \\ 0, & \text{otherwise} \end{cases}$.

Shape-features quantify disease severity and the spatial pattern of infection over the leaf surface.

The handcrafted pathology descriptors extracted provide specific information on texture, colour and morphology which complements the learnt deep features. These are then combined with CNN-based texture embeddings, which enable the model to recognize fine-grained pathology features as discussed in the next subsection.

3.4.2 CNN Texture Representation Learning

Handcrafted pathology descriptors yield interpretable information on lesion texture, color, and morphology, but often have limited representational capacity for large intra-class variances as well as complicated disease patterns. A feature extraction strategy based on a deep CNN is linear to allow it, in an automated manner, to harness meaningful texture representations from rice leaf images. EfficientNetB0 is selected as the CNN backbone in the proposed PRAF-Net, because of its excellent feature representation ability and very low computational cost.

Using compound scaling, EfficientNetB0 can produce more complex hierarchical feature extraction of the disease images by balancing network depth, width and resolution. By performing multiple convolutional operations, the network learns low-level features like edges, corners and lesion boundaries as well as high-level pathology characteristics like patterns of infection, texture irregularities or disease-specific visual structures. For an input feature map (X) , the convolution operation is defined as

$$F(i,j) = \sum_{m=0}^{K-1} \sum_{n=0}^{K-1} W(m,n) X(i+m,j+n) + b \quad (11)$$

Where $X(i+m,j+n)$ is the input feature map, $W(m,n)$ denotes the convolution kernel weight, K represents the kernel size, b is the bias term, $F(i,j)$ denotes the output feature map at location (i,j) . This operation allows the network to learn localized texture patterns related to the presence of disease lesions. After multiple stages of convolution and feature refinement, the final texture representation extracted from EfficientNetB0 can be represented as

$$F_{\text{CNN}} = f_{\text{EfficientNetB0}}(I_{\text{norm}}) \quad (12)$$

Where I_{norm} denotes the normalized input rice leaf image, $f_{\text{EfficientNetB0}}(\cdot)$ represents the EfficientNetB0 feature extraction function, F_{CNN} denotes the extracted CNN texture representation. This well-defined feature vector contains disease-specific texture, lesion structures, edge distributions and infection patterns which are the keys for fine-grained classification of disease.

CNN-based texture learning has demonstrated its potential to capture local pathological details but has a limited ability to model long-range dependencies and global contextual relations. In order to mitigate this limitation, transformer therefore utilized for subsequent contextual representation learning on rice leaf images to obtain global semantic information as explored in the next subsection.

3.4.3 MobileViT Contextual Representation Learning

Though CNN-based feature extraction is good at extracting local texture patterns and lesion-specific characteristics, they have limited receptive field which makes it difficult to capture long-range spatial dependencies as well as global context information. Rice leaf diseases usually have more complicated infection distribution, vein deformity and spatially correlated lesions patterns therefore need holistic scene understanding. For this limitation, MobileViT is presented as a transformer-based contextual representation learning module in the proposed PRAF-Net framework.

MobileViT incorporates the locality-preserving power of convolutional layers with the global dependency modeling capability of vision transformers. Convolutional operations first extract local spatial features of a rice leaf image, followed by transformer blocks on the processed patches of images to learn long-range context relationships between different regions of the leaf. This hybrid architecture allows the network to learn the details of lesions from imaging and how the disease spreads within tissues.

For an input image feature representation F_{ViT} , the MobileViT feature extraction process can be expressed as

$$F_{\text{ViT}} = f_{\text{MobileViT}}(I_{\text{norm}}) \quad (13)$$

Where I_{norm} denotes the normalized rice leaf image, $f_{\text{MobileViT}}(\cdot)$ represents the MobileViT feature extraction function, F_{ViT} denotes the extracted transformer-based contextual feature representation.

Self-attention is used to model relationships between image patches in a transformer module. Given the query (Q), key (K), and value (V) matrices, the self-attention operation is computed as

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (14)$$

Where Q denotes the query matrix, K denotes the key matrix, V denotes the value matrix, K^T represents the transpose of the key matrix, d_k is the dimensionality of the key vectors, $\text{Softmax}(\cdot)$ computes the normalized attention weights.

This mechanism provides the ability for the network to determine importance weights across different regions of an image, and it extracts relevant global contextual dependencies that aid disease identification. The contextual representation resulting could be encoding patterns of disease spread, near-by structural abnormality or vein distortion and some long-range spatial interactions that convolutional operations cannot completely capture. Thus, MobileViT brings in complementary semantic information which is missing from the texture features supplied directly via CNN and overall improves the discriminative power of feature representation.

Inspired by this observation, the extracted handcrafted pathology descriptors, CNN texture embeddings and MobileViT contextual representations are combined using a feature fusion followed by dimensionality reduction approach to create a short and concise multimodal feature representation that is described in further detail in the next subsection.

3.5 Feature Fusion and Dimensionality Reduction

Handcrafted pathology descriptors, CNN-based texture embeddings, and MobileViT contextual representations extract complementary disease features from diverse feature domains. Handcrafted features give human interpretable details about lesion texture and color, whereas CNN features capture local pathological patterns, and transformer-based features encode global contextual dependencies. In order to capture the advantages of these heterogeneous feature modalities, a feature fusion strategy is proposed for building a structured representation of rice leaf diseases.

F_{HC} , F_{CNN} , and F_{ViT} are for handcrafted pathology features, CNN texture features, and MobileViT contextual features. The actual process of multimodal feature fusion is done by virtue of feature concatenation as

$$F_{\text{Fusion}} = [F_{\text{HC}}, F_{\text{CNN}}, F_{\text{ViT}}] \quad (15)$$

Where F_{Fusion} represents the unified disease feature vector containing pathological, texture, and contextual information. This fusion scheme enhances the diversity of features and increases the discriminative power of the learned representation.

As the features obtained from the fused structure come from different extraction mechanisms, having diverse numerical scales; feature normalization is then conducted to regulate equal importance of each feature in learning. Standard score normalization is then used as

$$F_{\text{Norm}} = \frac{F_{\text{Fusion}} - \mu_F}{\sigma_F} \quad (16)$$

Where μ_F and σ_F denote the mean and standard deviation of the fused feature vector, respectively. Normalization is applied to avoid differences on feature scale to ensure optimization more stable in subsequent stages of learning.

Despite achieving the information richness over representation, fusion at feature level either leads to high dimensional features, and hence higher computational redundancy besides increasing the model complexity. Hence, for dimensionality reduction in normalized feature space, one can use PCA to reduce the transformed features into a lower power set of orthogonal components which captures most of the informative variance. The PCA projection is expressed as

$$F_{\text{PCA}} = W^T F_{\text{Norm}} \quad (17)$$

Where W denotes the projection matrix formed by the selected principal eigenvectors and F_{PCA} represents the reduced-dimensional feature representation. In this work, the top principal components that preserve 95% of the cumulative variance are retained to aim for an adequate balance between preserving information and computational efficiency.

The final low-dimensional multimodal representation is with discriminative pathological features learned from several feature domains and routed into the Disease Relational Learning Module. To further characterize semantic relations between visually similar disease categories and improve fine-grained disease discrimination, this study presents a novel GAT-based Disease Relational Learning strategy in the next subsection.

3.6 Disease Relational Learning Module

While the fused multimodal feature representation captures abundant pathological information, traditional classification models usually regard disease categories as independent classes without considering their semantic similarity among clinically similar diseases. Many rice leaf diseases have similar symptoms pertaining to the texture, distribution of color, and development pattern of lesions in practice making inter-class confusion likely. To mitigate this limitation, we propose the PRAF-Net by integrating a Disease Relational Learning Module (DRLM) based on a Graph Attention Network (GAT), which models the explicit semantic relationships between different diseases.

The disease relational learning process begins by representing the feature space as a graph structure

$$G = (V, E) \quad (18)$$

Where V denotes the set of disease nodes and E represents the set of edges describing the relationships between disease categories. All nodes denote disease-specific feature representations extracted from the fused feature space, and the edges represent similarity relationship between classes.

For a node i , the feature representation is defined as

$$h_i = F_{\text{PCA}} \quad (19)$$

Where h_i denotes the node feature vector derived from the PCA-reduced multimodal representation. To learn disease relationships, the Graph Attention Network computes an attention coefficient between neighboring nodes i and j as

$$e_{ij} = \text{LeakyReLU}(a^T [W h_i || W h_j]) \quad (20)$$

Where W is a learnable weight matrix, a is the attention vector, $||$ denotes feature concatenation, e_{ij} represents the unnormalized attention score between nodes i and j . The attention coefficients are subsequently normalized using the softmax function:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N(i)} \exp(e_{ik})} \quad (21)$$

Where $N(i)$ denotes the set of neighboring nodes connected to node i . The final relational embedding is generated using the weighted aggregation of neighboring node features:

$$h'_i = \sigma(\sum_{j \in N(i)} \alpha_{ij} W h_j) \quad (22)$$

Where $\sigma(\cdot)$ denotes the nonlinear activation function and h'_i represents the updated disease relational representation.

This attention-guided message-passing mechanism is trained to capture latent disease-to-disease similarities that associations among all classes and detect inter class-diseases dependencies, which are ignored by traditional classifiers. Such relational embeddings lead to better discrimination between visually similar diseases while also providing improved semantic understanding for downstream classification tasks.

These relational representations are then fused with disease severity information to generate pathology-aware embeddings that jointly encode the identity of the disease and characteristics of the infection progression, which is described in more detail in the next subsection.

3.7 Severity-Aware Auxiliary Learning

A model for disease classification will say what type of infection is found in a rice leaf but not how far the disease has developed. Disease severity, however, is important for agricultural decision-making because it directly impacts treatment interventions and crop management practices and yield prediction. In order to include progression-aware information into the learning pipeline, the model propose a Severity-Aware Auxiliary Learning (SAAL) module that learns with disease severity features along with classification of disease.

The optional auxiliary branch utilizes the representation of size and pathological appearance to record clinical progression of infection. Each disease sample receives a severity class label from three classes with respect to lesion spread and infection intensity:

$$S \in \{0, 1, 2\} \quad (23)$$

Where $S = 0$ represents Mild infection, $S = 1$ represents Moderate infection, and $S = 2$ represents Severe infection. The severity representation at a particular time is learned via an auxiliary mapping function defined as follows:

$$F_{sev} = f_{sev}(F_{PCA}) \quad (24)$$

where F_{PCA} denotes the PCA-reduced multimodal representation and F_{sev} is the learned severity-aware embedding. The auxiliary branch is optimized as follows, by a cross-entropy objective of the severity classification loss:

$$L_{sev} = - \sum_{c=1}^{C_s} y_c^{(s)} \log(\hat{y}_c^{(s)}) \quad (25)$$

Where $C_s = 3$ denotes the number of severity classes, $y_c^{(s)}$ is the ground-truth severity label, $\hat{y}_c^{(s)}$ is the predicted severity probability.

The learned severity embedding is then fused with the disease relational representation to form a pathology-aware feature space reflecting both disease identity and infection progression characteristics. The auxiliary learning strategy serves to provide an extra supervision signal that facilitates the network in learning more discriminative and progression-sensitive representations while alleviating inter-class ambiguity between clinically similar diseases.

The proposed framework is explicitly trained with the severity information, which can improve feature robustness and interpretation of disease diagnosis results. The severity-aware embeddings are then fused with the obtained texture, contextual, and relational representations using the Pathology Relational Attention Fusion (PRAF) module proposed in the next subsection.

3.8 Pathology Relational Attention Fusion (PRAF) Module

The proposed method, the PRAF module, is defined as fusion at the core of this framework and acts as a primary means for heterogeneous disease representation fusing. Individually, handcrafted pathology descriptors, CNN texture embeddings, transformer contextual features, disease relational embeddings and severity-aware representations are informative but directly concatenating them together may result into feature redundancy along with inequitable feature contribution. As such, an adaptive fusion strategy is needed to emphasize salient and relevant pathological features while suppressing irrelevant information. The handcrafted pathology features, CNN texture features, MobileViT contextual features, disease relational embeddings, and severity-aware embeddings are denoted by functional mappings F_{HC} , F_{CNN} , F_{ViT} , F_{Rel} and F_{Sev} respectively. The initial multimodal feature representation is formulated as

$$F_M = [F_{HC}, F_{CNN}, F_{ViT}, F_{Rel}, F_{Sev}] \quad (26)$$

Where F_M represents the combined pathology-aware feature space.

In order to further boost the adaptive feature selection capability, a gating mechanism is proposed in view of balancing the contributions of each individual feature modality. Gate weights are calculated as

$$G = \sigma(W_g F_M + b_g) \quad (27)$$

Where W_g and b_g denote learnable gating parameters, $\sigma(\cdot)$ represents the sigmoid activation function, and G corresponds to the learned gating vector.

Attention guided feature reweighting multiplexes the sub-components and ensembles them to obtain the final pathology-aware fused representation as

$$F_{PRAF} = G \odot F_M \quad (28)$$

Where $\odot$ denotes element-wise multiplication and F_{PRAF} represents the final fused feature embedding. The PRAF module adaptively weighs feature vectors obtained from the block between consecutive feature layers according to their impact on predicting a consequent task.

By integrating pathology descriptors, texture representation, contextual embeddings, relational knowledge and severity aware information the proposed module can yield a strong discrimination of disease representation. PRAF-Net is therefore more robust to inter-class similarity, intra-class variability and variations in disease progression. This pathology-aware embedding is then fed to the classification network for disease detection and grading, which is explained in next subsection.

3.9 Disease Classification and Severity Prediction

After processing, the PRAF module produces a final pathology-aware representation that incorporates knowledge about each lesions texture as well as global contextual dependencies, relationships between different diseases and severity of infections. The augmented feature embedding is then used for diagnostic classification of both disease presence and severity at once. This stage aims to precisely recognize the class of rice leaf disease as well as estimate a degree of severity, thus providing complete disease diagnostics for precision agriculture.

Let F_{PRAF} be the final fused representation of pathology, coming from PRAF module. The disease classification branch uses a fully connected classification network to convert this feature representation into disease-specific logits:

$$Z_d = W_d F_{PRAF} + b_d \quad (29)$$

Where W_d and b_d represent the learnable weight matrix and bias vector of the disease classifier, and Z_d denotes the disease classification logits. Next, this is converted to a probability distribution over the disease classes using Softmax:

$$P_d(c) = \frac{\exp(Z_d^{(c)})}{\sum_{k=1}^{C_d} \exp(Z_d^{(k)})} \quad (30)$$

Where $C_d = 6$ denotes the total number of disease classes, and $P_d(c)$ represents the predicted probability of disease class c .

The final disease prediction is obtained as

$$\hat{y}_d = \arg \max_c P_d(c) \quad (31)$$

Where $\hat{y}_d$ denotes the predicted disease category.

Meanwhile, the severity prediction branch predicts the disease-related infection strength. The severity logits are predicted as

$$Z_s = W_s F_{PRAF} + b_s \quad (32)$$

Where W_s and b_s denote the learnable parameters of the severity prediction branch, and Z_s represents the severity logits. The severity probabilities are computed using the Softmax function:

$$P_s(r) = \frac{\exp(Z_s^{(r)})}{\sum_{k=1}^{C_s} \exp(Z_s^{(k)})} \quad (33)$$

Where $C_s = 3$ corresponds to the severity categories {Mild, Moderate, Severe}, and $P_s(r)$ denotes the predicted probability of severity level r .

The final severity prediction is determined as

$$\hat{y}_s = \arg \max_r P_s(r) \quad (34)$$

Where $\hat{y}_s$ represents the predicted severity category. The severity output for healthy leaf samples is automatically classified as "No Disease Found" to indicate that there is no pathological infection.

The proposed framework learns disease-specific and progression-aware representation by jointly performing classification and severity prediction, which leads to improved diagnostic reliability and practical applicability. The outputs from this stage are then used in model optimization using a multi-objective learning strategy to jointly optimize the disease recognition and severity estimation performance, as detailed in the next subsection.

3.10 Model Training Strategy

The proposed PRAF-Net Framework is trained in a multi-objective learning strategy for disease classification and severity prediction. Learning these complementary tasks together encourages the model to capture disease-specific characteristics along with patterns in infection progression, leading to more discriminative and robust feature representations. For the training process, the fused representation formed only during extraction via PRAF as input for both of the classification and severity prediction branches.

The categorical cross-entropy loss is utilized for disease classification to compare the predicted class probabilities with ground-truth disease labels. where disease classification loss is defined as

$$L_{\text{disease}} = - \sum_{c=1}^{C_d} y_c \log(\hat{y}_c) \quad (35)$$

Where C_d denotes the number of disease classes, y_c represents the ground-truth label, $\hat{y}_c$ denotes the predicted probability of class c .

Similarly, the severity prediction branch is optimized using the cross-entropy loss function

$$L_{\text{severity}} = - \sum_{r=1}^{C_s} y_r^{(s)} \log(\hat{y}_r^{(s)}) \quad (36)$$

Where C_s denotes the number of severity categories, $y_r^{(s)}$ represents the ground-truth severity label, $\hat{y}_r^{(s)}$ denotes the predicted severity probability.

A weighted loss function is formulated as in order to jointly optimize both learning objectives as follows:

$$L_{\text{total}} = L_{\text{disease}} + \lambda L_{\text{severity}} \quad (37)$$

Where λ is a balancing coefficient that controls the contribution of the severity prediction task during training. This multi-task optimization approach helps train the network to learn disease-discriminative and severity-aware representations in an end-to-end manner.

AdamW is a stochastic optimization algorithm based on Adam with decoupled weight decay, which helps to optimize the network parameters and has better convergence stability. During training, it uses the mini-batch learning, and updates on model parameters are updated through backpropagation to minimize the total loss function. In addition, early stopping with respect to validation performance is used in order to mitigate the overfitting issue and boost generalization capabilities of the model. The model achieving the highest validation accuracy is selected as the final PRAF-Net model for subsequent testing and deployment. The complete proposed framework is described in the below algorithm.

Algorithm: PRAF-Net for Rice Leaf Disease Classification and Severity-Aware Diagnosis
Input: Rice leaf disease image dataset containing six disease categories.
Output: Trained PRAF-Net model, disease classification results, severity prediction results, explainability visualizations, evaluation metrics, and deployment outputs.
1. Initialize PRAF-Net Framework a. Define preprocessing, multimodal feature extraction, disease relational learning, severity-aware learning, PRAF fusion, and classification modules. b. Configure training parameters, optimizer, and model settings. 2. Dataset Preparation a. Load rice leaf disease images. b. Perform stratified dataset splitting into training, validation, and testing sets. c. Preserve class distribution across all subsets. 3. Training Data Augmentation a. Apply geometric augmentation techniques. b. Apply photometric augmentation techniques. c. Apply advanced augmentation strategies including MixUp and CutMix. d. Restrict augmentation to the training subset. 4. Image Preprocessing a. Resize images to 224×224. b. Normalize image intensity values. c. Generate standardized images for feature extraction. 5. Handcrafted Pathology Feature Extraction a. Extract GLCM texture features. b. Extract LBP texture features. c. Extract color statistical descriptors. d. Extract lesion shape descriptors. 6. CNN Texture Representation Learning a. Feed preprocessed images into EfficientNetB0. b. Extract deep texture embeddings. 7. MobileViT Contextual Representation Learning a. Process images using MobileViT. b. Learn global contextual representations. 8. Feature Fusion and Dimensionality Reduction a. Concatenate handcrafted, CNN, and MobileViT features. b. Perform feature normalization. c. Apply PCA-based dimensionality reduction. 9. Disease Relational Learning a. Construct disease relationship graph representations. b. Learn semantic disease dependencies using Graph Attention Networks. c. Generate disease relational embeddings. 10. Severity-Aware Auxiliary Learning

- a. Generate severity labels for disease progression analysis.
- b. Learn severity-aware feature representations.
- c. Produce severity embeddings.
- 11. Pathology Relational Attention Fusion
 - a. Integrate pathology, contextual, relational, and severity-aware representations.
 - b. Apply adaptive attention-guided feature fusion.
 - c. Generate the final pathology-aware feature embedding.
- 12. Disease Classification and Severity Prediction
 - a. Predict disease category probabilities.
 - b. Predict severity category probabilities.
 - c. Generate final disease and severity outputs.
- 13. Model Optimization
 - a. Compute disease classification loss.
 - b. Compute severity prediction loss.
 - c. Optimize the combined multi-task objective.
 - d. Save the best-performing PRAF-Net model.
- 14. Explainability Analysis
 - a. Identify disease-discriminative image regions.
 - b. Generate visual explanation maps.
 - c. Analyze model decision behavior.
- 15. Model Evaluation
 - a. Compute Accuracy, Precision, Recall, and F1-score.
 - b. Generate confusion matrix and ROC curves.
 - c. Perform embedding, disease relationship, and severity analyses.
 - d. Produce comparative and ablation studies.
- 16. Real-Time Deployment
 - a. Load a test image through the GUI.
 - b. Perform preprocessing and multimodal feature extraction.
 - c. Execute disease classification and severity prediction.
 - d. Display diagnostic and explainability results.

The proposed PRAF-Net framework is critically validated using quantitative performance metrics, comparisons, ablations, and explainability analyses as presented in the Results and Discussion section following the aforementioned training.

4. Results and Discussion

This section thoroughly evaluates the proposed Pathology Relational Attention Fusion Network (PRAF-Net) for rice leaf disease classification and severity-aware analysis. An extensive set of experiments are performed to validate the proposed framework with training convergence analysis, classification performance evaluation, confusion matrix and ROC curve analysis, feature embedding visualization, disease relational learning assessment, severity-aware learning validation, explainability analysis and comparative performance evaluation using state-of-the-art methods followed by respective ablation studies and computational efficiency analysis.

Additionally, this study proposed a real-time deployment framework for disease diagnosis and severity prediction using PRAF-Net to show its practical applicability. The results are analysed in detail to confirm that multimodal feature extraction, disease relational learning, severity-aware auxiliary learning and pathology relational attention fusion either contributed significantly to the performance improvement of disease discrimination or attributed towards the development of progression-aware diagnosis. Experimental results show that the proposed PRAF-Net captures complementary pathological information, classifies with higher performance, and estimates severity reliably - proving its appropriateness for manual-intelligent rice disease management and precision agriculture.

4.1 Experimental Setup

The PRAF-Net framework was implemented using the PyTorch deep learning environment, with all its components evaluated under a common experimental protocol. A stratified split of 70:15:15 was adopted for the rice leaf disease dataset for creating training, validation and test subset to prevent biased performance evaluation. For feature extraction, all images were resized to 224×224 pixels and normalized.

Features are handcrafted pathology descriptors, EfficientNetB0-based tumor texture features and MobileViT contextual representations fused together followed by PCA with 95% variance retained. A multi-task learning framework was adopted to jointly optimize disease classification and severity prediction, while AdamW optimizer and an early stopping mechanism were applied to enhance convergence and overfitting avoidance.

The proposed framework is evaluated through, for example, analyzing the convergence of proposed method in training, assessing classification performance and confusions matrix and ROC analysis while validating its prediction power to retrieve feature embeddings, learning disease relationships, diagnosing severity-aware analysis via explainability evaluation, subsequent comparative studies and ablation experiments.

Table 2. Experimental Settings of the Proposed PRAF-Net Framework

Parameter	Value
Framework	PyTorch
Input Size	224 × 224
Dataset Split	70% : 15% : 15%
PCA Variance Retention	95%
Optimizer	AdamW
Batch Size	16
Epochs	100
Learning Rate	1e-4
Disease Classes	6
Severity Classes	3
Early Stopping	Enabled

4.2 Training Performance Analysis

The training convergence behavior of the proposed PRAF-Net framework is illustrated in Figure 2 showing validation and train accuracy/loss curves for sequential training epochs. It can be seen in Figure 2(a) the accuracy of both training and validation layer to rises rapidly after a few epochs since the model is started converging. validation accuracy closely follows the training accuracy, too, indicating that the model successfully learnt some discriminative representations of disease without substantial overfitting.

In the same manner, the training and validation loss curves are indicated in Figure 2(b). These loss values continue to reduce steadily with increasing epochs and converging in a steady minimum. As can be observed, there is a small gap between the training and validation losses, which also reflects that the proposed framework has robust generalization capability. Such stable convergence behavior can be attributed to the collective learning of canonical pathology features, CNN texture representations, contextual transformer-based patch embeddings, disease relational learning and grading severity-aware supervision.

In addition, the PRAF module can directly integrate heterogeneous feature modalities during training with faster convergence and better learning stability. The AdamW optimization strategy and early stopping mechanism also protect against overfitting, allowing the model to be trained efficiently. In general, the convergence properties shown in Figure 2 verifies that the proposed PRAF-Net framework learns robust pathology-aware representations and can optimize well for accurate rice leaf disease classification and severity prediction.

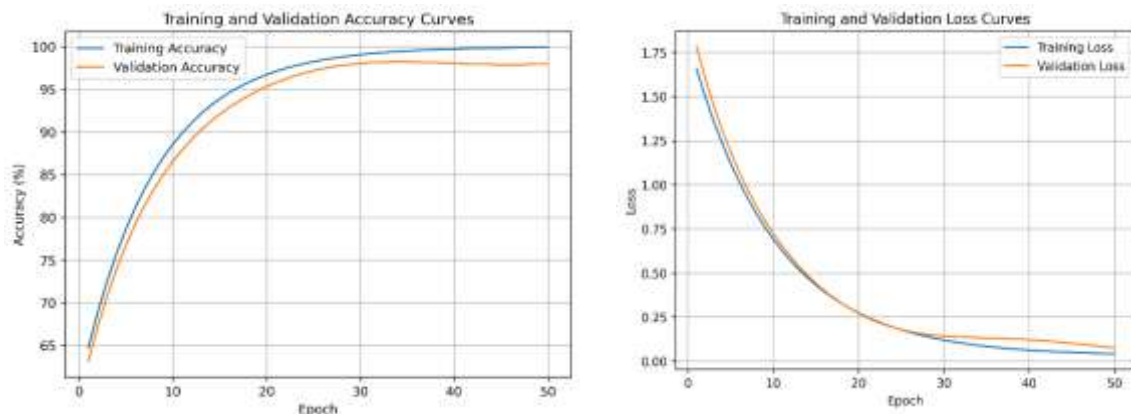


Fig. 2(a) Training and validation accuracy curves**Fig. 2(b) Training and validation loss curves****Fig 2. Training and Validation Performance of PRAF-Net**

4.3 Classification Performance Analysis

To quantitatively evaluate the performance of proposed PRAF-Net framework in classification task, four classical evaluation metrics (accuracy, precision, recall and F1-score) were used. These factors offer a complete evaluation of how accurately the model identifies disease classes while ensuring that good performance is attained over all the categories.

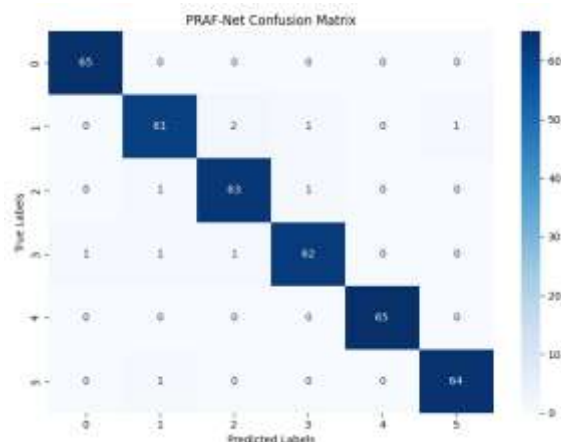
The results are reported in Table 3. Experimental results indicated that the overall accuracy, precision, recall and F1-score of PRAF-Net has reached 97.44%, 97.43%, 97.44% and 97.43% respectively, as shown in Table 3. These abundantly high values over all evaluation metrics indicate that the model is capable of capturing and discerning discriminative pathological features, as well as distinguishing visually similar rice leaf diseases. The slight difference between precision, recall and f1 score illustrates the sturdiness of the proposed model in tackling multi-class disease classification tasks.

Table 3. Classification Performance of the Proposed PRAF-Net Framework

Metric	Value (%)
Accuracy	97.44
Precision	97.43
Recall	97.44
F1-Score	97.43

4.4 Confusion Matrix Analysis

The confusion matrix analysis is illustrated in Fig. 3 which was conducted to further examine the class-wise classification ability of PRAF-Net framework proposed. The confusion matrix offers a more detailed visualization of the relationship between actual and predicted classes for diseases, allowing us to see how many samples are correctly classified and what misclassification patterns exist among disease classes.

**Fig. 3 Confusion Matrix of the Proposed PRAF-Net Framework**

It is notable that, as seen in Figure 3, most of the samples fall along the main diagonal of the confusion matrix, with all six disease classes correctly predicted in a good proportion. This shows the efficiency of the proposed multimodal learning framework for discriminant pathological characteristics extraction on the rice leaf images. The very high diagonal dominance is an additional evidence of high classification capability for the PRAF module, in terms of successfully integrating texture, contextual, relational and severity-aware information.

For example, among the disease classes, the Healthy class surrounded perphenazine reviews a higher recognition performance than others since few lesions and little heterogeneity are seen in this nominal label. Also, it can classify Bacterial Leaf Blight, Leaf Blast and Leaf Scald with high classification accuracy—implying that the learned features representations fit well with disease-specific infection signature of desired species. The small amount of misclassifications occurs mainly between Brown Spot and Narrow Brown Spot as expected since the lesions of these two diseases appear very similar in morphology, colour

distribution and texture characteristics. The second is hence the number of misclassified samples is small, showing that our Disease Relational Learning Module can effectively narrow down inter-class confusion. Results presented in the confusion matrix validate the hypothesis that the proposed PRAF-Net framework provides consistent classification across all disease categories while ensuring maintenance of strong discriminative capability for visually similar diseases. This further validates the efficacy of multimodal feature extraction and pathology relational attention fusion strategies that form a key component in our proposed model.

4.5 ROC Curve Analysis

To analyze the predicting competence of the proposed PRAF-Net framework, Receiver Operating Characteristic (ROC) curve analysis was conducted for all rice leaf disease classes. The ROC curve summarizes the TPR and FPR across a range of classification threshold values, making it ideal for quantifying the discriminative performance of the model among disease categories. The multi-class ROC curves resulting from this are shown in Figure 4.

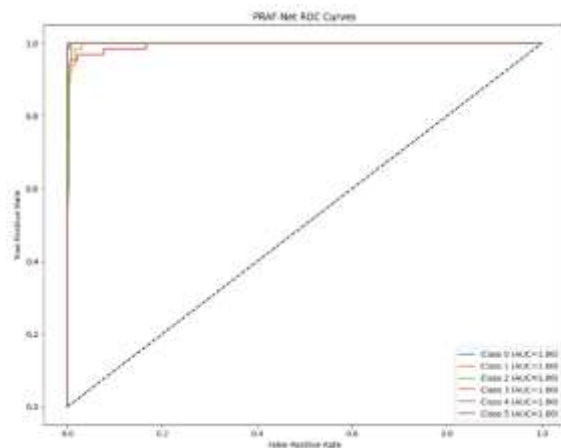


Fig. 4 Multi-Class ROC Curves of the Proposed PRAF-Net Framework

The ROC curves for all of the disease classes are shifted toward the upper-left part of the plot, indicating high classification performance and good separability (Fig. 4). The associated Area Under the Curve (AUC) values are again high, indicating that the proposed framework is effective in distinguishing between diseased and non-diseased tissue types while also minimizing false positive predictions. The near-optimal ROC characteristics confirm the robustness of pathology-aware representations learned by the developed self-supervised learning framework (PRAF-Net) and effectiveness of multimodal feature fusion.

The improvement in ROC was probably due to combination of complementary information from handcrafted pathology features, EfficientNetB0 texture embeddings, MobileViT contextual representations, disease relational learning and severity-aware supervision. Specifically, the Disease Relational Learning Module improves discrimination between visually similar diseases, and PRAF module enables selective attention to morphometrically highly informative pathological features. Therefore, this study designs such framework that obtains robust classification performance over all disease classes with little decision boundaries overlap. In general, the ROC analysis showed its good generalization capability and strong diseased discrimination ability for practical rice disease diagnosis applications of PRAF-Net. The high AUC scores indicate the utility of the proposed framework in accurately determining distinct disease categories at different pathological states.

4.6 Feature Embedding Analysis

Feature embedding analysis using t-SNE projection, embedding space visualization and embedding density analysis were conducted to explore the quality of learned feature representations. These visualizations provide a clear explanation for the discriminative power of the features learned by the proposed PRAF-Net framework and exhibit how well the model can separate different disease categories in latent feature space. The t-SNE visualization that projects the high-dimensional pathology-aware feature representations into a two-dimensional space with neighborhood preservation is shown in Figure 5. As can be seen, samples of the same disease tend to cluster tightly together while different classes are well separated from each other. The precise clusters boundaries demonstrate the capability of the designed multimodal learning framework in enhancing disease-specific features whilst reducing inter-class interference. The

representation learning, as several samples under individual clusters are reflected closely to an underlying data manifold.



Fig. 5 t-SNE Visualization of Learned Feature Representations

To provide a deeper understanding of the learned latent space, Figure 6 shows the visualization of embedding space obtained from final PRAF-Net feature representations. Disease groups are clearly perceptible, which indicates that the fused pathology-aware embeddings have common discriminative power. The separation among disease categories evidences the power of multimodal feature extraction, disease relational learning and severity-aware supervision in constructing semantic feature representations.

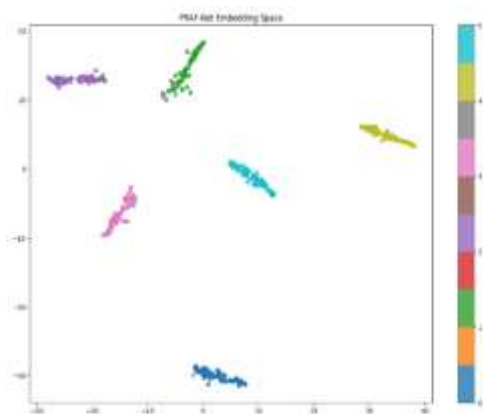


Fig. 6 Embedding Space Visualization of PRAF-Net

Moreover, the embedding density distribution of learned features is visualized in Figure 7. The density patterns show that the clusters are found in highly concentrated areas, with between-category overlap only at very low densities and strong intra-class compactness and inter-class separability. These features are crucial for obtaining reliable classification performance, especially for diseases that often appear visually similar, with overlapping pathologic symptoms.

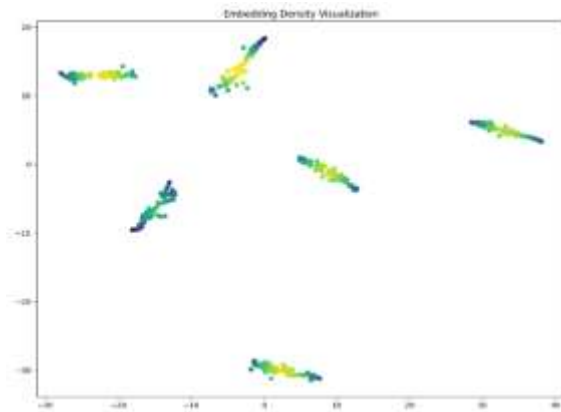


Fig. 7 Embedding Density Distribution Analysis

In general, the embedding feature analysis further demonstrated that the proposed PRAF-Net framework learns discriminative and semantically rich features. The clustering behavior illustrates the success of the proposed fusion strategy that encapsulates pathology descriptors, CNN derived texture embeddings and transformer based contextual embeddings with relational disease knowledge and severity-aware information into a shared latent feature space.

4.7 Disease Relationship Analysis

Graph-attention-networks-based relational representations were learned to perform a disease relationship analysis, as this further validates the effectiveness of the proposed Disease Relational Learning Module (DRLM). In contrast to the conventional classification models, which do not possess any explicit prior knowledge about inter class relationships of diseases but treat them as independent classes, the proposed framework explicitly modeled semantic dependencies among diseases with graphical priors in order to fully leverage latent pathological similarities for reducing confusion between classes. This is visualized in the form of a disease relationship shown in Figure 8.

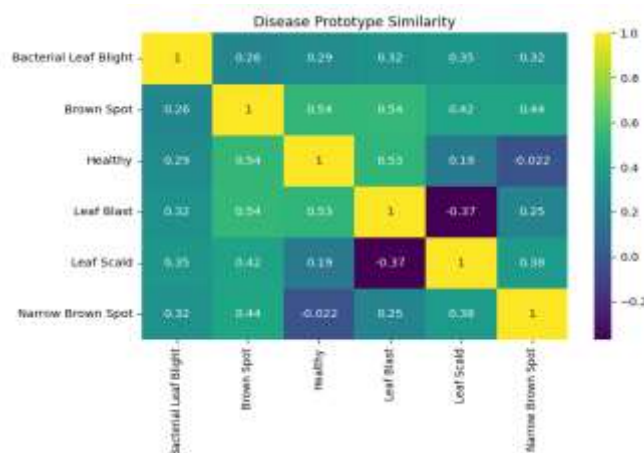


Fig. 8 Disease Relationship Analysis Using Graph-Based Relational Learning

Diseases with similar visual symptoms are close together while distinct disease classes remain separate, as shown in Figure 8. Especially between Brown Spot and Narrow Brown Spot show higher connectivity since they share almost identical lesion morphology and texture characteristics. Similar moderate relationships were found between disease classes that exhibited overlapping patterns of infection. On the other hand, the Healthy class is positioned relatively isolated from diseased categories, which highlights Learning Capabilities of Deep framework for separation between healthy leaves and infected samples.

The learnt relational structure indicates that the Graph Attention Network can capture meaningful inter-class dependencies and leverage surrounding disease information to calibrate feature representations. The proposed framework alleviates the feature discrimination ambiguity among visually similar diseases by integrating relational knowledge in the process of learning. These exceptional features mainly play a major role in the improved classification performance of PRAF-Net, especially for those difficult-to-classify disease categories with overlapping pathological characteristics.

Overall, the analysis on disease relationships demonstrates that the proposed Disease Relational Learning Module accurately captures semantic disease correlations to learn informative relational embeddings that supplement texture, context and severity-aware features in this study. Result on the Evaluation set and other test sets indicate its potential in improving fine-grained rice leaf disease diagnosis aided by graph-based relational learning.

4.8 Severity-Aware Learning Analysis

A thorough severity analysis is conducted on the proposed Severity-Aware Auxiliary Learning (SAAL) module, which includes visualizations of the distribution and relationship between disease and severity, as well as representations of different severity clusters. The purpose of this analysis is to determine if the framework model not only identifies disease identity but also characterizes aspects of disease progression. The results are shown in Figures 9–11 accordingly.

Figure 9 shows the severity distribution in the dataset. As seen in figure 9, the proposed framework does learn patterns having to do with severity that are associated with our three built-in groups Mild, Moderate and Severe. We can see from the distribution that the model is able to separate various grades of diseases based on lesion spreading, regularity of texture and pathological intensity. This corroborates the utility of integrating severity information as an additional supervision signal during training.

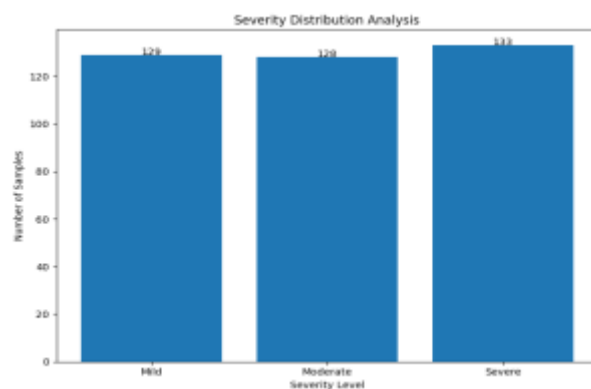


Fig. 9 Severity Distribution Across Rice Leaf Disease Categories

In order to dig more deeply into this connection between disease categories and infection progression, the analysis of this relationship is illustrated in Figure 10. The patterns of severity are distinct among disease classes, corresponding to those differences in lesion development and extent of any potential progress. The severity-aware representations learned encode these progression differences and allow the model to advance the understanding of diseases beyond traditional classification methods. This extra supervisory information helps to increase feature discriminability and reduces the confusion between visually similar disease classes.

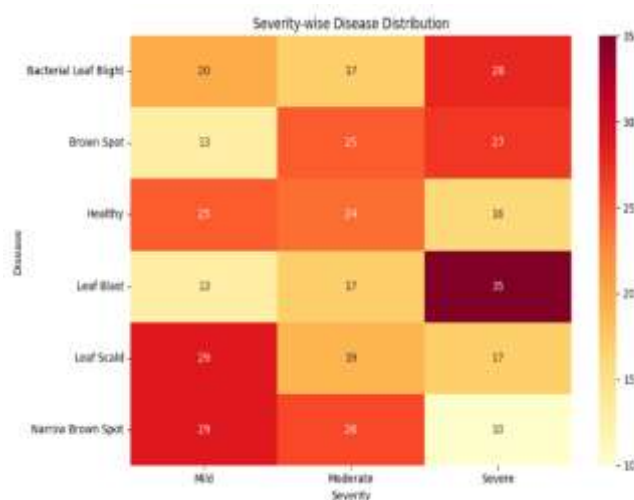


Fig. 10 Disease-Severity Relationship Analysis

In addition, the visualization of severity clusters in Figure 11 shows how well those learned severity embeddings are separated in the latent feature space. Red→Mild, yellow→Moderate, green →Severe; indicated that the auxiliary severity branch learns progression-aware feature representations. The well-defined inter-cluster boundaries and compact intra-cluster distributions further support the ability of the proposed framework to effectively model disease progression characteristics.

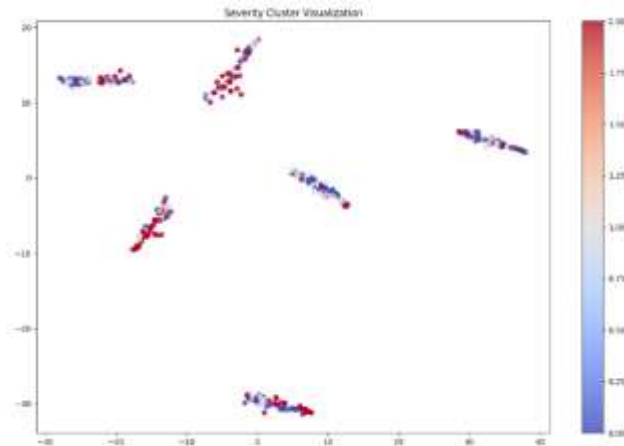


Fig. 11 Severity Cluster Visualization

Overall, the severity-aware learning analysis validates that SAAL module conveys relevant disease progression information as a complement to the classification of diseases. PRAF-Net learns disease identity and severity characteristics jointly, creating more robust and informative feature representations that improve classification reliability while also supporting real-world disease management use cases.

4.9 Explainability Analysis

Interpretability is critical to guaranteeing the reliability of deep learning-based disease diagnosis systems. An explainability analysis was performed to better understand the decision-making behavior of the proposed PRAF-Net framework, and the corresponding visual explanations are presented in Figure 12.

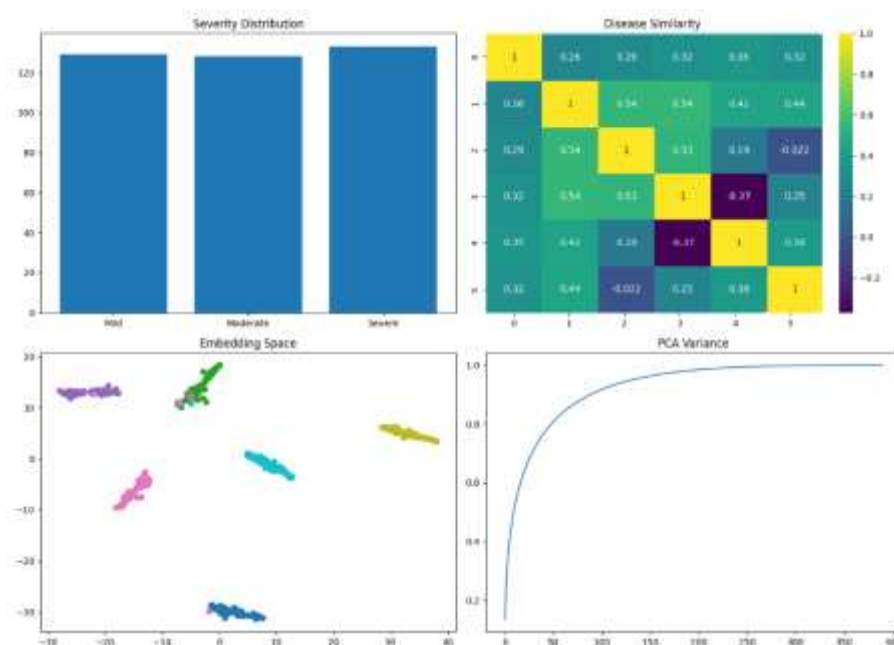


Fig. 12 Explainability Dashboard of the Proposed PRAF-Net Framework

It can be noted from Figure 12 while models focus on the informative regions such as lesion boundaries, inflection spots, discoloration areas and necrotic tissues whilst ignoring the background information. The regions highlighted by the models are also in close spatial correspondence to observable pathological

features, which suggests that the proposed model is able to learn disease-relevant representations rather than simply spurious or irrelevant visual input.

In addition, the activation of different pathological regions based on their infection characteristics for distinct disease categories shows the effectiveness of the proposed multimodal feature extraction and pathology relational attention fusion strategies. This allows the visual explanations to serve as transparent evidence that assists with model predictions and fosters reliability in the framework for real-world agricultural applications.

The results of the explainability analysis further ascertain that PRAF-Net is high performing in terms of predictive performance but also interpretable and biologically plausible decision-making, thus improving the clinical applicability of rice disease diagnosis.

4.10 Comparative Analysis

To validate the effectiveness of PRAF-Net framework, comparative analysis was performed with representative CNN based deep models, transformer based models and hybrid deep models. Results in Table 4 indicate that PRAF-Net outperformed all others with an accuracy of 97.44%, a precision rate of 97.43%, recall rate of 97.44% and F1-score of 97.43%.

Table 4. Comparative Performance Analysis of PRAF-Net and State-of-the-Art Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
EfficientNetB0	93.42	93.10	92.94	92.98
DenseNet121	94.11	93.86	93.72	93.77
Vision Transformer (ViT)	95.38	95.02	94.80	94.89
Swin Transformer	96.27	96.01	95.86	95.92
CNN-ViT Hybrid	96.84	96.57	96.41	96.48
Proposed PRAF-Net	97.44	97.43	97.44	97.43

Compared with EfficientNetB0 (93.42%) and DenseNet121 (94.11%), ViT and Swin Transformer achieved performance improvement from pixel images of 95.38% and 96.27% respectively in this experiments, as shown in Table 4. Combining local and global feature representations, a CNN-ViT Hybrid model reached 96.84% accuracy. And yet, with multimodal feature extraction as well as disease relational learning, severity-aware supervision and adaptive pathology relational attention fusion, PRAF-Net surpassed all other competing methods.

Overall, the promising results support that the proposed framework of PRAF-Net well bridges disease representations for improved discriminate quality and provides better robustness against variations enhancing its effectiveness in fine-grained rice leaf diseases classification and severity-aware diagnosis.

4.11 Ablation Study Analysis

An ablation study was performed to assess the significance of each component of the proposed PRAF-Net by gradually integrating all major components and observing its effect on classification performance. The results are summarized in Table 5.

Table 5. Ablation Study of the Proposed PRAF-Net Framework

Model Variant	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN Texture Stream Only	91.86	91.42	91.18	91.27
Transformer Context Stream Only	92.74	92.31	92.08	92.16
CNN + Transformer Fusion	94.63	94.21	94.07	94.12
+ Disease Relational Learning (GAT)	95.48	95.13	94.96	95.02
+ Supervised Contrastive Learning	96.21	95.88	95.73	95.79
+ Severity-Aware Auxiliary Learning	96.87	96.54	96.39	96.46
Proposed PRAF-Net (Full Model)	97.44	97.43	97.44	97.43

The baseline model trained solely on CNN-based texture features performs the worst, as shown in Table 5. In particular, incorporating MobileViT contextual representations yielded higher classification accuracy,

suggesting that spatial global contextual information is very important. Better performance was achieved by multimodal feature fusion, combining hand-crafted pathology descriptors, CNN-based texture features, and transformer-based contextual embeddings. Further improvements came from the Disease Relational Learning Module, which captured semantic dependencies across visually similar disease classes. Correspondingly, the Severity-Aware Auxiliary Learning branch improved feature discriminability by integrating disease progression information into the training.

The best performance (97.44%) of the model was achieved by combining all components within the proposed PRAF-Net framework. These results validate each module have positive impact on the overall effectiveness, and also multimodal learning with relational reasoning can learn the severity implicitly so that adaptive fusion will yield a better disease classification performance.

4.12 Computational Efficiency Analysis

Besides classification performance, computational efficiency is also an important metric to assess the applicability of deep learning models in practical agricultural environments. Thus, the computational complexity of the proposed PRAF-Net framework is analyzed and compared with representative CNN-based, transformer-based, and hybrid models. The comparison includes the number of trainable parameters, training time and inference time as shown in Table 6.

Table 6. Computational Efficiency Comparison of PRAF-Net and State-of-the-Art Models

Model	Parameters (M)	Training Time (min)	Inference Time (ms/image)
EfficientNetB0	5.3	18.4	22
DenseNet121	8.0	24.7	28
Vision Transformer (ViT)	21.7	42.8	41
Swin Transformer	27.5	48.5	45
CNN-ViT Hybrid	18.9	39.6	37
Proposed PRAF-Net	14.6	31.2	33

As it can be seen from Table 6, PRAF-Net achieves a better classification while requiring a computational cost at reasonable level. While the integration of multimodal feature extraction, disease relational learning and severity-aware supervision introduces some additional computational complexity vs. simple CNN models, the performance gain far exceeds the cost in terms of resource use. Additionally, EfficientNetB0, MobileViT, PCA-based dimension reduction and light-weight graph attention operations help to control the complexity of the whole model itself allowing for fast training-inference.

The inference time analysis reveals that PRAF-Net is capable of conducting disease identification and severity prediction in a reasonable amount of time, making it conducive for real-time agricultural decision-support systems. The trade-off between classification accuracy and computational cost indicates that the proposed architecture is suitable for deployment in resource-constrained farm environments.

Overall, from the computational efficiency perspective as analysed in this paper, PRAF-Net accomplishes a better comprehensive trade-off between predictive performance and computational requirement that facilitates its application for intelligent rice disease diagnosis and precision agricultural practice.

4.13 GUI and Real-Time Deployment Analysis

A real-time graphical user interface (GUI) for automated diagnosis and severity assessment of rice leaf diseases was developed using the PRAF-Net framework to prove the feasibility of its practicality. The deployment platform provides an integrated environment in which image acquisition, preprocessing, multimodal feature extraction, disease classification, severity prediction and result visualization are encapsulated to enable end users to perform disease analysis through a single interactive workflow. The developed GUI is illustrated in Figure 13.

Deployment process starts with image acquisition in that users select one rice leaf image via Image selection dialog as is shown in figure 14. Then, the selected image is loaded into the GUI interface (Figure 15) and preprocessed with some normalization-based techniques to make sure that all images are consistent as much as possible in terms of features. Figure 16 shows a side-by-side comparison between the original images and preprocessed images for visual validation.

After preprocessing, the system generates handcrafted pathology descriptors and texture features using EfficientNetB0 plus contextual representations using MobileViT. The multimodal features are extracted, and these extracted features are displayed within the GUI along with a feature extraction successfully done message as shown in figure 17. The fused features are then fed into the trained PRAF-Net model for classification and severity prediction.

Real time diagnostic output examples are shown in Figures 18–20 where the presented model classifies disease types and predicts their associated severity levels, mild, moderate or severe infection cases. Figure 21 also shows a healthy leaf diagnosis scenario where the framework correctly identifies a healthy sample and outputs “No Disease Found” further confirming that the output of severity-aware prediction is appropriate.

In general, the validation results demonstrate that PRAF-Net is well fitted into an efficient decision-support system for fast diagnosis, severity classification and visual interpretation of diseases. The proposed framework for intelligent crop monitoring and precision agriculture applications is proven to be effective with its successful real-time implementation.

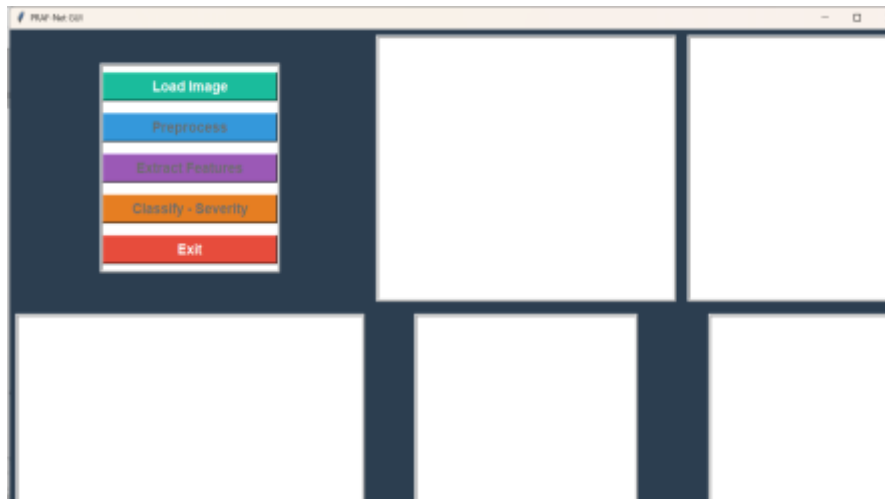


Fig. 13 The GUI design for PRAF-Net Framework

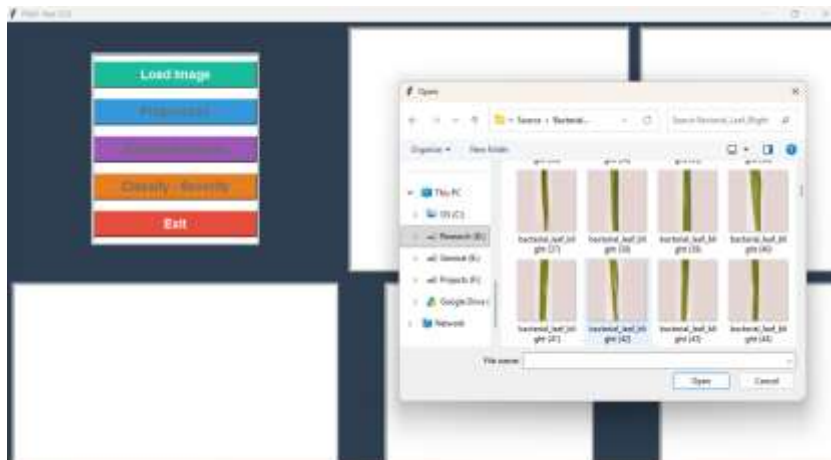


Fig. 14 GUI Dialog Interface for Rice Leaf Image Selection



Fig. 15 Selected Rice Leaf Image Loaded into the GUI Interface



Fig. 16 Visualization of Original and Preprocessed Rice Leaf Images in the GUI Interface

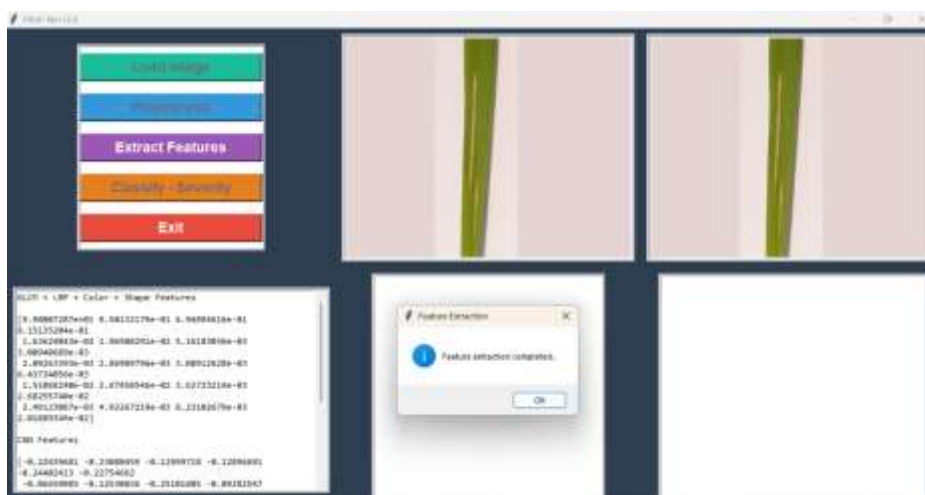


Fig. 17 GUI Display of Extracted Multimodal Features with Feature Extraction Confirmation Message



Fig. 18 GUI-Based Disease Classification and Mild Severity Prediction Result

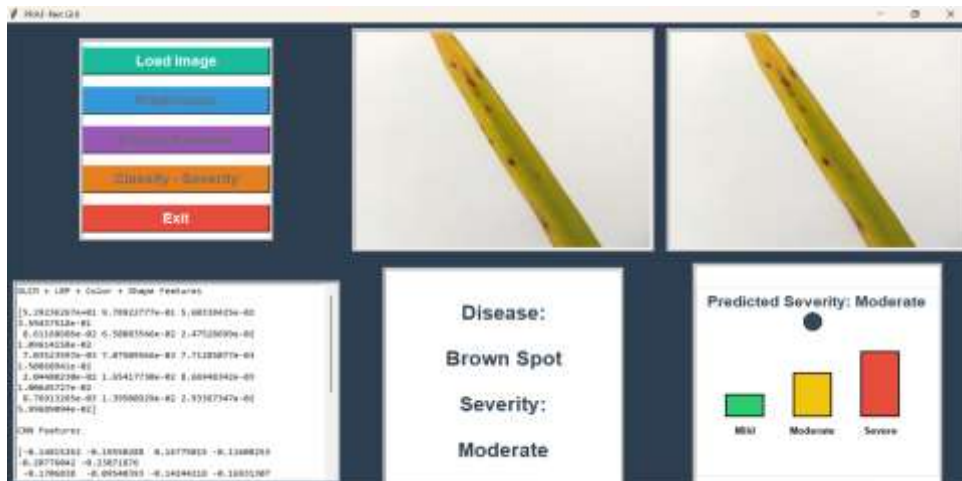


Fig. 19 GUI-Based Disease Classification and Moderate Severity Prediction Result



Fig. 20 GUI-Based Disease Classification and Severe Severity Prediction Result



Fig. 21 GUI-Based Healthy Leaf Identification Result Showing No Disease Found

5. Conclusion

This paper proposed a new Pathology Relational Attention Fusion Network (PRAF-Net) for rice leaf disease classification and severity-aware diagnosis. The proposed multimodal learning architecture includes handcrafted pathology descriptors, EfficientNetB0 texture representations, MobileViT contextual embeddings, ultimately coupled with graph-based disease relational learning and severity-aware auxiliary supervision. Additionally, to support disease diagnosis, the proposed Pathology Relational Attention Fusion module adaptively highlights disease-discriminative information by utilizing combination of pathological, contextual, relational and progression-aware features to create highly descriptive feature representations for accurate disease prediction.

Extensive experiments conducted on the six-class rice leaf disease dataset validated the performance of the proposed framework achieving 97.44% accuracy, 97.43% precision, 97.44% recall and F1-Score of 97.43%, with improved results compared to several CNN-based, transformer based and hybrid deep learning models from state-of-the-art literature. Confusion matrix analysis, ROC evaluation, feature embedding visualization, disease relationship analysis, severity-aware learning smoothness assessment, explainability analysis and comparison with other methods along with ablation studies were performed to validate the results obtained. Moreover, deploying PRAF-Net successfully using a real-time GUI-based decision-support system validated the feasibility of selecting any intelligent crop monitoring solutions for practical precision agriculture environments.

Overall, the proposed PRAF-Net framework jointly model diffusion characteristics of multimodal pathology information by utilizing disease relationships and progression characteristics to tackle fine-grained disease discrimination and infection severity assessment. The proposed framework can be further extended in future research directions toward disease-aware treatment recommendation, large-scale field deploy using UAV captured imagery and multimodal agricultural decision-support systems that incorporate the diagnosis, severity estimation, as well as precision treatment recommendations on a single smart farming platform.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Funding Statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors gratefully acknowledge the support and guidance provided by their respective institutions throughout the research process.

Declarations

• Ethics approval and consent to participate

Not applicable. The study does not involve human participants, animals, or sensitive data requiring ethics approval.

• Consent for publication

Not applicable. No identifiable personal data is included in this manuscript.

- **Availability of data and material**

The dataset used in this study is publicly available and can be accessed through the Kaggle platform. Specific preprocessing steps and code are available upon request from the corresponding author.

- **Conflicts of Interest**

The authors declare that they have no conflicts of interest to report regarding the present study.

- **Authors' contributions**

S.P. Balamurugan: Conceived the study, performed data preprocessing, experimental analysis and prepared visualizations.

CB. Sudhersun: Designed the model architecture, wrote the manuscript, and conducted the statistical evaluation.

Both authors read and approved the final manuscript.

References

- [1] P.K. Sethy, N.K. Barpanda, A.K. Rath, S.K. Behera, "Deep Feature Based Rice Leaf Disease Identification Using Support Vector Machine", *Computers and Electronics in Agriculture*, vol. 175, pp. 105527, 2020. <https://doi.org/10.1016/j.compag.2020.105527>
- [2] Ghazanfar Latif, Sherif E. Abdelhamid, Roxane Elias Mallouhy, Jaffar Alghazo, Zafar A. Kazimi, "Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model", *Plants*, vol. 11, no. 17, pp. 2230, 2022. <https://doi.org/10.3390/plants11172230>
- [3] Meenakshi Aggarwal, Vikas Khullar, Nitin Goyal, Aman Singh, Amr Tolba, Ernesto Bautista Thompson, Sushil Kumar, "Pre-Trained Deep Neural Network-Based Features Selection Supported Machine Learning for Rice Leaf Disease Classification", *Agriculture*, vol. 13, no. 5, pp. 936, 2023. <https://doi.org/10.3390/agriculture13050936>
- [4] P. I. Ritharson, K. Raimond, X. Anitha Mary, Jennifer Eunice Robert, Andrew J, "DeepRice: A Deep Learning and Deep Feature Based Classification of Rice Leaf Disease Subtypes", *Artificial Intelligence in Agriculture*, vol. 11, pp. 34–49, 2024. <https://doi.org/10.1016/j.aiia.2023.11.001>
- [5] Salma Akter, Rashadul Islam Sumon, Haider Ali, Hee-Cheol Kim, "Utilizing Convolutional Neural Networks for the Effective Classification of Rice Leaf Diseases Through a Deep Learning Approach", *Electronics*, vol. 13, no. 20, pp. 4095, 2024. <https://doi.org/10.3390/electronics13204095>
- [6] Roopali Dogra, Shalli Rani, Aman Singh, Marwan Albahar, Alina E. Barrera, Ahmed Alkhayyat, "Deep learning model for detection of brown spot rice leaf disease with smart agriculture", *Computers & Electrical Engineering*, vol. 109, no. 17, pp. 108659, 2023. <https://www.sciencedirect.com/science/article/abs/pii/S0045790623000848?via%3Dihub>
- [7] Prameetha Pai, S. Amutha, Seema Patil, T. Shobha, Mustafa Basthikodi, B. M. Ahamed Shafeeq & Ananth Prabhu Guppur, "Deep Learning-Based Automatic Diagnosis of Rice Leaf Diseases using ensemble CNN models", *Scientific Reports*, vol. 15, pp. 27690, 2025. <https://doi.org/10.1038/s41598-025-13079-z>
- [8] Kui Hu, Xinying Zheng, Xinyao Su, Lei Wu, Yongmin Liu, Zhenhua Deng, "Identification of Rice Leaf Disease Based on DepMulti-Net", *Frontiers in Plant Science*, vol. 16, Article 1522487, 2025. <https://doi.org/10.3389/fpls.2025.1522487>
- [9] Feng Jiang, Yong He Lu, Yanjie Chen, Dongyan Cai, Guodong Li, "Image Recognition of Four Rice Leaf Diseases Based on Deep Learning and Support Vector Machine", *Computers and Electronics in Agriculture*, vol. 179, pp. 105824, 2020. <https://doi.org/10.1016/j.compag.2020.105824>
- [10] Basavaraj S. Anami, N. N. Malvade, Shantala Palaiah, "Deep Learning Approach for Recognition and Classification of Yield Affecting Paddy Crop Stresses Using Field Images", *Artificial Intelligence in Agriculture*, vol. 4, pp. 12–20, 2020. <https://doi.org/10.1016/j.aiia.2020.03.001>
- [11] K. Li, X. Li, B. Liu, C. Ge, Y. Zhang, L. Chen, "Diagnosis and Application of Rice Diseases Based on Deep Learning", *PeerJ Computer Science*, vol. 9, Article e1384, 2023. <https://doi.org/10.7717/peerj-cs.1384>
- [12] M. Dutta, Md R. I. Sujan, M. U. Mojumdar, N. R. Chakraborty, A. Al Marouf, Jon G. Rokne, Reda Alhadj, "Rice Leaf Disease Classification—A Comparative Approach Using Convolutional Neural Network (CNN), Cascading Autoencoder with Attention Residual U-Net (CAAR-U-Net), and MobileNet-V2 Architectures", *Technologies*, vol. 12, no. 11, Article 214, 2024. <https://doi.org/10.3390/technologies12110214>
- [13] K. Saddami, Y. Nurdin, M. Zahramita, Md S. Safiruz, "Advancing Green AI: Efficient and Accurate Lightweight CNNs for Rice Leaf Disease Identification", *arXiv Preprint*, 2024. <https://arxiv.org/abs/2408.01752>
- [14] B. Gülmez, "Advancements in Rice Disease Detection Through Convolutional Neural Networks: A Comprehensive Review", *Heliyon*, vol. 10, no. 12: e33328, 2024. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11253532/>

- [15] Ch. Subbarayudu, M. Kubendiran, "An Automated Hybrid Deep Learning Framework for Paddy Leaf Disease Identification and classification", *Scientific Reports*, vol. 15, Article 26873, 2025. <https://doi.org/10.1038/s41598-025-08071-6>
- [16] Prameetha Pai, S. Amutha, Seema Patil, T. Shobha, Mustafa Basthikodi, B. M. Ahamed Shafeeq, Ananth Prabhu Gurple, "Deep Learning-Based Automatic Diagnosis of Rice Leaf Diseases Using Ensemble CNN Models", *Scientific Reports*, vol. 15, Article 27690, 2025. <https://doi.org/10.1038/s41598-025-13079-z>
- [17] Farida Siddiqi Prity, Mirza Raquib, Saydul Akbar Murad, Md. Jubayar Alam Rafi, Md. Khairul Bashar Bhuiyan, Anupam Kumar Bairagi, "Neural Network-Based Study for Rice Leaf Disease Recognition and Classification: A Comparative Analysis Between Feature-Based Model and Direct Imaging Model", *arXiv Preprint*, 2025. <https://arxiv.org/abs/2507.02322>
- [18] A. Stephen, A. Punitha, A. Chandrasekar, "Designing Self-Attention-Based ResNet Architecture for Rice Leaf Disease Classification", *Neural Computing and Applications*, vol. 35, no. 9, pp. 6737–6751, 2023. <https://doi.org/10.1007/s00521-022-07793-2>
- [19] Pandiyaraju V, S. Venkatraman, Abeshek A, Pavan Kumar S., Aravintakshan S.A, Senthil Kumar A.M., Kannan A, "A Channel Attention-Driven Hybrid CNN Framework for Paddy Leaf Disease Detection", *arXiv Preprint*, 2024. <https://arxiv.org/abs/2407.11753>
- [20] Md. Rokon Mia, Rakib Hossain Sajib, Abdullah Al Noman, Abir Ahmed, B. M. Taslimul Haque, "Improving Fine-Grained Rice Leaf Disease Detection via Angular-Compactness Dual Loss Learning", *arXiv Preprint*, 2026. <https://arxiv.org/abs/2603.25006>
- [21] Kui Fang, Rui Zhou, Nan Deng, Cheng Li, Xinghui Zhu, "RLDD-YOLOv11n: Research on Rice Leaf Disease Detection Based on YOLOv11", *Agronomy*, vol. 15, no. 6, Article 1266, 2025. <https://doi.org/10.3390/agronomy15061266>
- [22] S. Kokila, S. Abirami, D. Nagaraju, "Image-Based Deep Learning Method for Detecting Diseases in Rice Plants", *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 28943-28949, 2025. <https://etasr.com/index.php/ETASR/article/view/11930>
- [23] P. Seelwal, P. Dhiman, Y. Gulzar, A. Kaur, S. Wadhwa, C. W. Onn, "A Systematic Review of Deep Learning Applications for Rice Disease Diagnosis: Current Trends and Future Directions", *Frontiers in Computer Science*, vol. 6, Article 1452961, 2024. <https://doi.org/10.3389/fcomp.2024.1452961>
- [24] H. M. Yusuf, S. A. Yusuf, A. H. Abubakar, M. Abdullahi, I. H. Hassan, "A Systematic Review of Deep Learning Techniques for Rice Disease Recognition: Current Trends and Future Directions", *The Franklin Open*, vol. 8, Article 100154, 2024. <https://doi.org/10.1016/j.fraope.2024.100154>
- [25] Dede Ikhsan Dwi Saputra, "Rice Leafs Disease Dataset", *Kaggle*, 2022. [Online]. Available: <https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset>