Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **3**, September **2025**

Adaptive Vision Transformer for Accurate Paddy Leaf Disease Classification Under Cross-Infection Dynamics

Anitha G*

Dept. of CSE, Faculty of Engineering and Technology
Jain Deemed-to-be-University
Bengaluru, Karnataka, India
anithag@ubdtce.org

P. Thiruvannamalai Sivasankar

Dept. of CSE, Faculty of Engineering and Technology
Jain Deemed-to-be-University
Bengaluru, Karnataka, India
tp.sivasankar@jainuniversity.ac.in

*Corresponding author: Anitha G

Received March 5, 2025, revised May 21, 2025, accepted May 26, 2025.

Abstract. *Paddy leaf diseases, including fungal infections like Rice Blast and Brown Spot and bacterial infections like Leaf Blight and Leaf Streak, often co-exist, creating complex morphological variations. The dynamic evolution of lesions, influenced by environmental conditions and pathogen interactions, leads to unpredictable changes in texture, shape, and spectral properties. These factors cause feature misalignment, asymmetrical lesion fusion, and optical distortions, making traditional models ineffective and resulting in frequent misclassification. To address these challenges, the Vision Rotary Attention Decision Forest Perceptron Transformer is proposed for accurate fungal and bacterial disease classification. It starts with Image Patch Generation for segmenting rice leaf images for fine-grained lesion details, then using RoPE (Rotary Position Encoding) and Fourier Position Encoding for spatial and frequency representation. These enriched representations are processed using Multihead Self-Attention and GBdt-enhanced feedforward networks, gradient boosting trees, and Graph and Wavelet Patch Embeddings for precise decision-making and robust segmentation in bacterial and fungal infections. Finally, the hierarchical MLP with XGBoost and Random Forest improves classification accuracy by capturing evolving lesion structures and mitigating feature misalignment errors, while processing fungal and bacterial infections separately enhances robustness. The Experimental evaluations confirm enhanced feature extraction, reduced time complexity, and superior classification accuracy over conventional models.*

Keywords: Paddy Leaf Disease; Fungal and Bacterial Infections; Feature Misalignment; Disease Classification; Lesion Segmentation; Machine learning

1. **Introduction.** Paddy leaf plants, mainly of the *Oryza sativa* species, are some of the most important sites for rice cultivation in the world. They are the main sites of photosynthesis and nutrient absorption needed for plant growth [1]. The leaves are linear-lanceolate with parallel venation and therefore capture light more effectively. Stomatal functioning controls gas exchange and transpiration, which in turn affect yield. Leaf

architecture varies among rice cultivars due to environmental stress subjection. Drought, salinity, and extreme temperatures affect leaf morphogenesis and functioning [2]. Healthy paddy leaves are generally bright green, suggesting good chlorophyll and nutrient status. Discolored, yellowed, or curled leaves are common characteristics of deficiency symptoms or plant stress. Leaves turn pale green in plants deficient in nitrogen while they have brown edges because of potassium deficiencies [3]. The dark green to purple coloration arising from phosphorus deficiency negatively affects growth. Excess moisture or standing water weakens the roots and affects nutrient uptake and, hence, leaf health. Different soil types with unsuitable conditions, including improper balance of pH, hinder good morphological leaf development [4].

Beyond abiotic stressors, pathogenic infections, particularly viral and bacterial diseases, pose critical threats to rice production, leading to significant yield losses, weakened plant vigor, and reduced grain quality [5]. Viral infections, including Rice Tungro Virus (RTV), Rice Dwarf Virus (RDV), and Rice Grassy Stunt Virus (RGSV), are primarily transmitted by insect vectors such as green leafhoppers and brown planthoppers, causing yellowing, mosaic-like patterns, curling, and severe growth stunting in leaves, ultimately disrupting panicle development and reducing grain formation [6]. The high humidity and increasing vector populations exacerbate disease transmission, making control strategies complex. Traditional management relies on insecticides and virus-resistant cultivars, but excessive chemical use harms beneficial pollinators, soil microbiota, and biodiversity, while also accelerating vector resistance evolution, necessitating advanced detection and mitigation strategies for effective disease control [7].

Moreover, it should be noted that several bacterial infections of paddy leaves are serious problems in rice farming, causing considerable yield and quality loss [8]. Bacterial pathogens, particularly Bacterial Blight and Bacterial Leaf Streak, cause devastating yield losses in rice farming by infiltrating plant tissues through stomata, hydathodes, or mechanical wounds, proliferating rapidly under warm and humid conditions [9]. Infected leaves exhibit water-soaked lesions that turn yellowish-brown, necrotic streaks along veins, premature drying, and reduced photosynthetic efficiency, ultimately leading to poor tillering and diminished grain formation [10]. Transmission occurs through contaminated irrigation water, infected seeds, and insect vectors, making containment challenging. Conventional strategies, such as field sanitation, resistant cultivars, and bactericides, provide partial control but face limitations due to bacterial resistance evolution [11].

The detection and classification of paddy leaf diseases face significant challenges due to environmental variability, symptom overlap, model adaptability, and data limitations. Fluctuations in light conditions, humidity, and temperature affect image consistency, reducing model accuracy [12]. Many diseases, including fungal, bacterial, and viral infections, exhibit similar discoloration and lesion patterns, complicating differentiation. Conventional epidemiological models like the Susceptible-Infected-Recovered (SIR) model struggle to incorporate real-time climatic variations and sudden environmental stressors, limiting their predictive reliability [13]. Machine learning models such as CNNs, k-NN, and SVMs require large, high-resolution datasets for accurate classification, yet real-world conditions introduce shadows, occlusions, and mixed infections, hindering precise disease detection [14]. Overcoming these challenges demands robust, adaptable AI-driven models capable of real-time analysis while accounting for field complexities [15]. Hence, a need for advanced AI-driven disease detection frameworks that integrate deep learning, hyperspectral imaging, and real-time environmental monitoring to enhance accuracy and adaptability.

1.1. **The main contribution of the research.** The Vision Rotary Attention Decision Forest Perceptron Transformer effectively addresses feature misalignment and lesion evolution by leveraging advanced spatial-frequency feature extraction, ensuring that lesion morphology is preserved despite uneven infection spread. Its adaptive learning mechanisms enhance feature differentiation, enabling accurate classification of dynamic fungal and bacterial infections. To mitigate misclassification in co-existing infections, the model incorporates advanced embeddings that distinguish fungal and bacterial lesions, ensuring precise cross-infection recognition. This approach enhances feature alignment, minimizes errors, and improves disease differentiation and diagnostic accuracy, making it a robust and reliable solution for agricultural disease management.

1.2. **Organization of the Paper.** In order to address the issues with Paddy Leaf Disease Prediction, the aforementioned contributions have been taken into consideration. Section 2 of the article outlines the literature study, Section 3 outlines the suggested methodology and how it operates, and Section 4 addresses the assessment, performance, and comparative analysis of the suggested model. The article is finally concluded in Section 5.

2. **Literature review.** Ritharson et al. [16] provided a solution by accurately identifying and classifying rice leaf diseases by utilizing Deep Learning (DL) and transfer learning techniques. Together with the benchmark datasets, a comprehensive dataset of 5932 self-generated photographs of rice leaves was compiled. The suggested custom VGG16 model performed well in recognizing and classifying nine distinct rice leaf disease class labels. Data augmentation techniques were used to increase the quantity of photos after careful human labeling and dataset segmentation, which were verified by horticulture specialists. However, this model struggles to generalize to new or unseen diseases, regions with varying agricultural practices, which requires additional fine-tuning.

Chakrabarty et al. [17] developed a thorough investigation for identifying rice leaf illnesses using pre-trained CNNs (Convolutional Neural Networks) and an sophisticated artificial intelligence system known as the BEiT model (optimal bidirectional encoder representations from the transformers for pictures). For efficient disease identification, the BEiT model uses convolution and attention methods to divide the input picture into visual tokens and suppress unnecessary portions to improve emphasis on important characteristics. The transformer network used contextual linkages and visual token analysis to classify symptoms of plant diseases. By highlighting important regions involved in illness identification with superpixels and attention maps, the LIME approach was used to analyze and explain the model's decision-making process. Although LIME makes things easier to understand, it is unable to fully describe the intricate connections between the many elements that contribute to the development of illness.

Shafik et al. [18] presented two plant disease detection (PDDNet) models for effective plant disease diagnosis and classification: the lead voting ensemble (LVE) and early fusion (AE). These models were combined with nine pre-trained CNNs and refined by deep feature extraction. A more reliable feature extraction procedure was made possible by this model's integration of early fusion, which blended features from the separate CNNs at the first layers. To improve the accuracy of illness identification, this model used the capabilities of many CNN models to vote on the final diagnosis. The multi-object deep learning model's efficacy in a variety of real-world scenarios is diminished by its inability to generalize across various plant species, backgrounds, and environmental conditions.

Singh et al. [19] suggested a unique CNN architecture for identifying and categorizing prevalent diseases in rice plants. To increase efficiency and preserve high disease detection accuracy, the architecture was created with a focus on minimizing the number of parameters. The specialized CNN architecture was created to identify diseases accurately while being lightweight, efficient, and requiring fewer parameters. The model was trained using two optimization techniques Adam and Stochastic Gradient Descent with Momentum (SGDM). Both methods were evaluated both with and without the dataset of healthy rice leaves. However, only four rice plant illnesses were used to train the model, which limits its capacity to detect additional diseases not present in the dataset. The dataset must be expanded to include a greater variety of paddy leaf plant diseases.

Maheswaran et al. [20] provided a CNN-based model for image processing-based rice leaf disease detection and classification. To highlight the damaged and healthy areas of the leaves, the background of the image was removed during pre-processing based on hue values. With an emphasis on both damaged and healthy leaf parts, CNN was used to extract characteristics from the photos. After feature extraction, the model used a fully connected layer to classify the pictures into one of six categories: either healthy leaves or diseased leaves. The method helped farmers safeguard their crops by facilitating precise early disease identification, which eventually increased agricultural output and quality and contributed to food security. However, if the training dataset is not sufficiently varied, the model suffers from overfitting.

Archana et al. [21] presented an approach to enhance rice plant disease identification's classification and computational capabilities. The input image was pre-processed, after that noise and fuzzy pixels were eliminated using the Wiener filter. To determine which areas of the pre-processed photos were impacted, the modified K-means segmentation approach was used. To extract features: the area and diameter of the segmented image were used to extract shape-based features; the gray-level cooccurrence matrix (GLCM) and bit pattern features (BPF) were used to retrieve texture features; and the novel intensity-based color feature extraction (NIBCFE) method was used to extract color-based features. Lastly, the images were classified using the derived feature values using a unique probabilistic neural network based on support vector machines. However, this suffers in coordinating multiple agents to work collaboratively while handling independent sub-spaces and facilitating effective communication among them.

Wang et al. [22] suggested using photos of rice leaves to identify and categorize rice illness using an attention-based depthwise separable neural network with Bayesian optimization (ADSNN-BO). The model enhanced CNN architectures' performance for rice illness classification by using pre-trained ImageNet weights. Additionally, the model's hyperparameters were adjusted using the Bayesian optimization technique. Activation map and filter visualization were also used for feature assessment in order to demonstrate the model's performance. However, the computational and memory constraints of some mobile devices make real-time processing and large-scale deployment difficult, even if the ADSNN-OB model is suited for mobile devices.

Azim et al. [23] developed a categorization method for rice leaf diseases that can identify three distinct illnesses: leaf smut, brown spot, and bacterial leaf blight. The approach included preprocessing stages including hue thresholding for impacted region segmentation and saturation thresholding for background removal. The XGBoost model was used for the classification of the extracted features, and model parameters were adjusted to attain the best possible results. There are now just three distinct rice diseases that the model can identify: bacterial leaf blight, brown spot, and leaf smut. It would be more applicable in practical situations if the model extended to identify a wider variety of rice illnesses.

Tiwari et al. [24] introduced a DL-based technique for identifying and categorizing plant diseases from images of leaves captured at various resolutions. A deep convolutional neural network architecture was trained using a large collection of plant leaf photos from many countries. Six crops in 27 different categories were analyzed in the proposed study in both laboratory and field settings. Several intra-class and inter-class changes of images with complicated circumstances were handled by this dense neural network. The method may be applied in a range of agricultural contexts because the model was trained on a sizable and varied dataset that took into account several crop and disease categories. Nevertheless, the model does not assess the severity of the illnesses; instead, it concentrates on disease detection.

Bijoy et al. [25] suggested a deep Convolutional Neural Network (dCNN) architecture that was lightweight and optimized for the detection of rice leaf disease. The goal of this technique was to increase the precision and effectiveness of identifying common rice diseases, including brown spot, sheath blight, rice blast, bacterial leaf blight, and tungro. Additionally, the study presented a crop health monitoring system that automatically annotates fresh photos and is a useful tool for farmers and the academic community. It consists of an Android application, a website, and an open API. Only five rice leaf diseases are the focus of the system. However, additional dataset augmentation and model retraining are necessary to extend the system's detection capabilities to a wider spectrum of ailments.

The above statements expressed that [16] struggles to generalize to new or unseen diseases, requires additional fine-tuning, and [17] is unable to fully describe the intricate connections between the many elements that contribute to the development of illness. Whereas [18] multi-object deep learning model's efficacy in a variety of real-world scenarios is diminished by its inability to generalize, [19] limits its capacity to detect additional diseases not present in the dataset, [20] training dataset is not sufficiently varied, the model suffers from overfitting, [21] suffers in coordinating multiple agents to work collaboratively while handling independent sub-spaces and facilitating effective communication among them, [22] computational and memory constraints of some mobile devices make real-time processing and large-scale deployment difficult, [23] practical situations if the model extended to identify a wider variety of rice illnesses, [24] not assess the severity of the illnesses; instead, it concentrates on disease detection, and [25] dataset augmentation and model retraining are necessary to extend the system's detection capabilities to a wider spectrum of ailments. Therefore, a new approach is required to accurately predict and classify bacterial and fungal diseases in rice leaves.

3. **Problem overview.** Paddy leaf is a staple crop that sustains global food security, providing essential carbohydrates for over half the world's population. Its cultivation supports economic stability, promotes biodiversity in wetland ecosystems, and plays a vital role in sustainable agriculture through carbon sequestration and water management. However, paddy leaf disease detection faces significant challenges due to the simultaneous occurrence of fungal and bacterial infections, which induce adaptive morphological changes in plant tissue. These cross-infection interactions trigger complex biochemical modifications, altering lesion texture, color, and shape, ultimately making traditional feature extraction methods unreliable. As lesions evolve dynamically, fungal infections typically manifest as small necrotic spots, while bacterial streaks expand along veins, leading to asymmetrical lesion fusions that defy predefined segmentation models. Additionally, the metabolic byproducts released by both pathogens alter the leaf's optical properties, disrupting reflectance and fluorescence signals, which in turn diminishes the effectiveness of hyperspectral and multispectral imaging techniques. Furthermore, moisture imbalances between infections cause localized tissue swelling or shrinkage, further distorting spatial and spectral features, making conventional boundary-based segmentation techniques ineffective. These interconnected challenges necessitate the development

of real-time adaptive learning models and multimodal imaging strategies to improve the accuracy and robustness of disease detection.

In classification, environmental factors play a critical role in lesion transformations, leading to feature misalignment and increased misclassification rates. For instance, Rice Blast lesions, which typically have a distinct gray center, may lose this characteristic under high humidity, making them resemble Brown Spot, while Bacterial Leaf Streak gradually merges into blotches, mimicking Bacterial Leaf Blight. Additionally, as diseases progress, structural disruptions in plant tissue further complicate classification nutrient deficiencies cause Brown Spot lesions to lose their circular shape, while Bacterial Leaf Blight induces uneven tissue swelling, distorting vein structures and disrupting spatial feature alignment. These continuous variations blur inter-class boundaries, making it difficult for traditional models to distinguish between diseases accurately. To overcome these challenges, real-time adaptive feature extraction, multimodal imaging, and anomaly-driven classification techniques are essential for improving disease differentiation accuracy and enhancing model reliability.

3.1. **Proposed Methodology.** The Vision Rotary Attention Decision Forest Perceptron Transformer architecture combines next-generation spatial, frequency, and machine learning-based enhancements to provide strong feature extraction and disease classification of rice leaves. The model has a systematic pipeline that includes Image Patch Generation which subdivides rice leaf into size-fixed patches (16×16 pixels) to maintain the fine-grained lesion features with the ability to distinguish between bacterial and fungal infection, then Position Encoding which incorporates Rotary Position Embedding and Fourier Position Encoding to improve spatial and frequency representations by retaining lesion structure. Additionally, Multihead Self-Attention is used to learn global and local relationships while Vision Transformer (ViT) dynamically focuses on not letting overlapping infection mask disease patterns, and Gradient Boosted Decision Trees (GBDT) in the feedforward network hidden layers to improve the segmentation correctness with decision boundaries fused into them. Furthermore, Graph and Wavelet Patch Embeddings simulate lesion connectivity patterns, and wavelet embeddings assist in edge detection and texture contrasts to correctly mark lesion-induced deformations of tissues, and Hierarchical MLP Classification utilizes a Hierarchical Multi-Layer Perceptron. MLP involving XGBoost and Random Forest models to correct disease-specific routing of features as well as increasing resistance to evolving dynamic symptoms. By treating fungal and bacterial infections individually before ultimate classification, the model achieves precise one-to-one assignment of one fungal and one bacterial disease per image, which effectively counteracts real-world cross-infections. This multi-stage combination of spatial, frequency, and machine learning-based methodologies ensures better flexibility and accuracy in rice leaf disease classification, and it is thus a strong candidate for precision agriculture and plant disease surveillance.

Figure 1 illustrates the Vision Rotary Attention Decision Forest Perceptron Transformer (VRAD-FPT), a novel deep learning model for paddy leaf disease classification, integrating spatial, frequency, and decision forest-based enhancements. It begins with image patch generation, segmenting leaves into fixed-size patches to preserve fine-grained lesion details. Next, Rotary Position Embedding (RoPE) and Fourier Position Encoding enhance spatial-frequency representation, ensuring morphological consistency. The Transformer blocks, equipped with Multi-Head Self Attention (MHSA) and a Gradient Feedforward Network, extract global and local dependencies for robust feature learning. Further, Graph Wavelet Patch Embeddings refine lesion connectivity and spectral variations, enhancing classification accuracy. The extracted features are then classified using

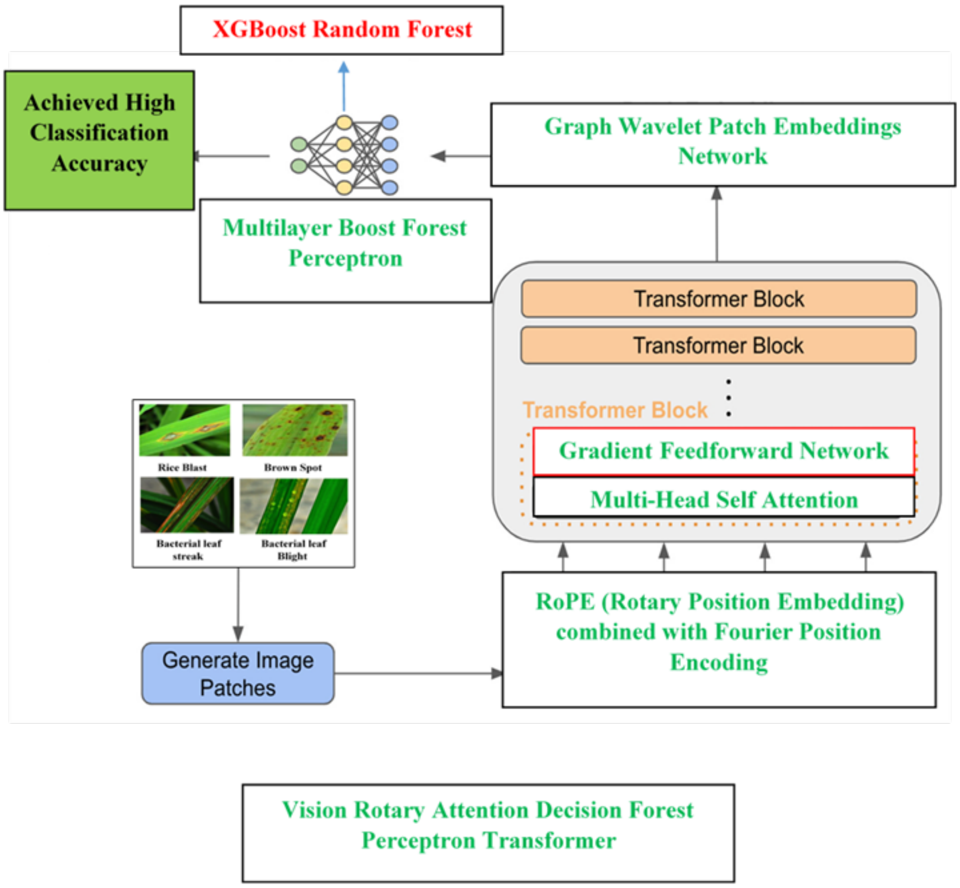


Figure 1. Block diagram of the proposed Vision Rotary Attention Decision Forest Perceptron Transformer

a Multilayer Boost Forest Perceptron, which integrates XGBoost and Random Forest to differentiate fungal and bacterial infections. This structured pipeline ensures superior classification performance, addressing feature misalignment and lesion transformation challenges in real-world agricultural scenarios.

3.2. **Image Patch Generation.** Dividing the rice leaf image into small patches is essential for preserving fine-grained lesion characteristics and improving disease detection accuracy. Processing the entire image at once can lead to the loss of crucial details due to variations in lesion size and shape, where some infections appear as small clusters while others spread along veins. Additionally, the leaf's natural texture, veins, and lighting conditions introduce complexity that may interfere with segmentation and feature extraction. Large images also require significant computational power, making direct processing inefficient. To overcome these challenges, the image is first pre-processed by resizing and normalizing to improve contrast and brightness consistency. Then, the image is divided into small, non-overlapping or slightly overlapping patches of 16×16 pixels, ensuring each patch retains localized lesion features, such as color, texture, and structural variations, without being blurred by spatial averaging.

These image patches are then tokenized and fed into a Vision Transformer (ViT) model, where each patch is treated as an independent input unit. Unlike convolutional neural networks (CNNs), which rely on local receptive fields, ViT processes these patches globally using self-attention mechanisms. This ensures that localized lesion details are captured without being suppressed by pooling layers and inter-patch relationships are preserved, allowing detection of small, clustered infections and larger, vein-spreading lesions.

3.3. **Spatial Representation with Positional embeddings.** Positional embeddings are crucial in deep learning models, especially for processing non-sequential image patches where spatial relationships must be preserved. Traditional positional encodings use absolute positions, but Rotary Position Embedding (RoPE) and Fourier Position Encoding provide more effective ways to capture lesion distribution patterns and frequency-based variations in rice leaf disease detection.

RoPE introduces a rotationally invariant encoding that ensures spatial dependencies between patches are preserved, allowing the model to track lesion spread and fusion effectively. Instead of assigning fixed positional values, RoPE applies a learned rotational transformation to each feature vector, encoding positional relationships relative to other patches rather than absolute locations. This enables the model to recognize lesion fusion patterns, ensuring that gradual changes in lesion structures are captured across adjacent patches. Additionally, it enhances robustness to variations in leaf orientation, preventing misclassification due to positional shifts in the input image. Furthermore, RoPE strengthens long-range interactions between patches, allowing the Transformer model to detect global lesion trends across the leaf. Instead of embedding fixed positional values, RoPE represents each patch as a vector in a complex plane and applies rotational transformations. This embedding method is directly integrated into the attention mechanism of ViT, improving lesion-tracking accuracy and feature alignment across different image regions.

Complementing RoPE, to further improve disease differentiation, Fourier Position Encoding enhances lesion characterization by incorporating frequency-based information, helping the model differentiate between bacterial streaks and fungal necrotic spots based on their unique spatial frequency patterns. The process begins by converting each patch's spatial data into the frequency domain using the Fourier Transform, highlighting lesion-specific periodic and irregular patterns. These frequency components are then encoded into positional embeddings, allowing the model to distinguish structural variations even when lesions appear visually similar in the spatial domain.

By integrating Fourier-based features with RoPE embeddings, ViT achieves a comprehensive understanding of both spatial connectivity and spectral differences. This fusion significantly improves disease classification by ensuring the model captures fine-grained local variations and broader lesion spread trends across the image.

3.4. **Feature Extraction.** In the feature extraction phase, Vision Transformer (ViT) uses Multihead Self-Attention (MSA) to efficiently capture both local and global dependencies between lesions, significantly enhancing rice leaf disease detection. Unlike conventional CNN-based methods with fixed receptive fields, ViT dynamically learns relationships between patches, ensuring disease-specific characteristics remain distinguishable, even in cases of overlapping bacterial and fungal infections. In contrast, MSA dynamically learns relationships between patches, ensuring that disease-specific characteristics remain distinguishable. By allowing each attention head to focus on different lesion attributes such as texture variations, color intensities, and shape structures MSA enhances the model's ability to differentiate complex lesion interactions. In Figure 2 the architecture of this process is given:

This process begins with the image being divided into 16 × 16-pixel patches, each converted into a feature vector $P \in \mathbb{R}^{s \times f}$ where S is the number of patches and f is the feature dimension. Each feature vector is projected into three learned matrices given in equation (1) [26, 27]:

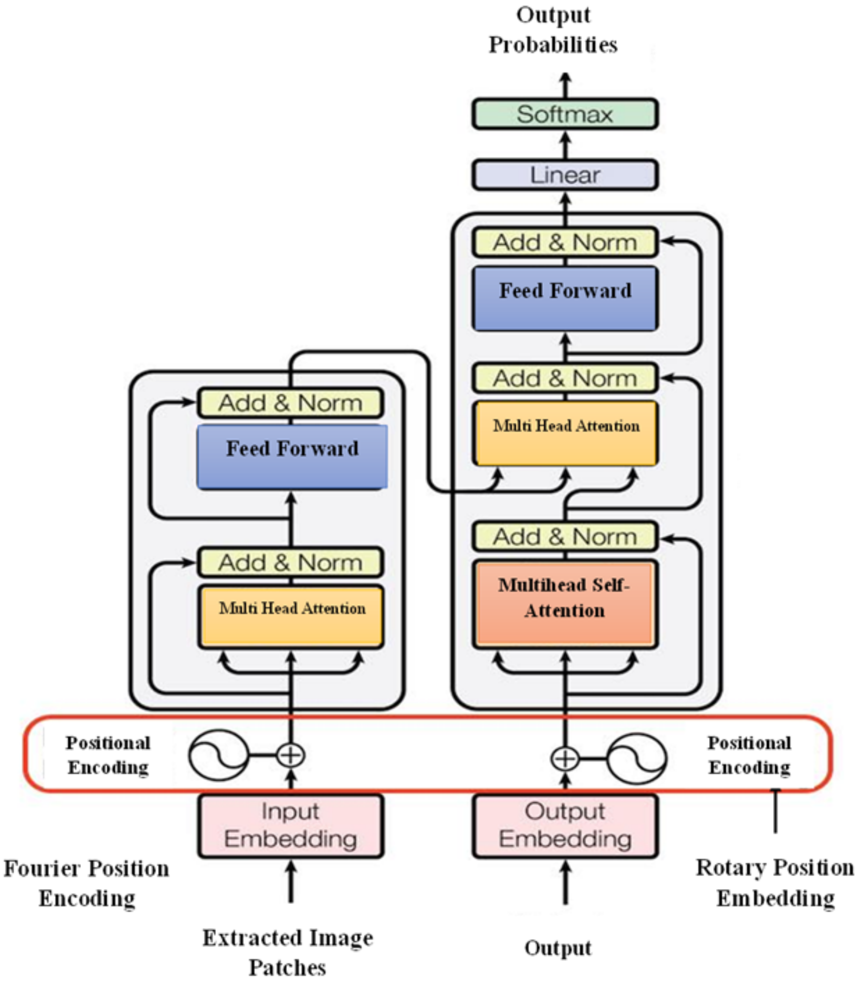


Figure 2. Architecture of Multihead Self-Attention in ViT Transformer

$$U = PW_U, \quad V = PW_V, \quad T = PW_T \tag{1}$$

Where, W_U, W_V, W_T are weight matrices that transform input features into query, key, and value representations. After that attention mechanism computes similarity scores between patches using the scaled dot-product attention which is given in equation (2)

$$Attention(U, V, T) = softmax\left(\frac{UV^T}{\sqrt{f_v}}\right)T \tag{2}$$

where f_v is the feature dimension for normalization, ensuring numerical stability. Multiple attention heads run parallel computations, capturing different lesion features given in equation (3):

$$MSA(U, V, T) = Concat(head_1, head_2, \ldots, head_h)W_Y \tag{3}$$

where W_Y is a learned projection matrix that merges insights from different attention heads. This mechanism allows ViT to simultaneously focus on different spatial lesion patterns, preserving spectral and textural characteristics. Finally, MSA-refined features are passed to the feedforward network for robust disease classification.

3.5. **Adaptive feature learning.** To enhance ViT's feature representation, Gradient Boosting Decision Trees (GBDT) are integrated within the hidden layers of the Feedforward Network (FFN). Deep learning excels at high-level feature extraction, but struggles with fine-grained decision boundaries, particularly in cases where lesion structures evolve. GBDT complements this by refining disease-specific texture patterns through an ensemble of decision trees that iteratively reduce residual errors in lesion classification. This hybrid approach ensures that misclassified patterns from the deep network are recalibrated, improving segmentation accuracy and reducing false positives by extracting features using a Transformer-based feature extractor that processes image patches, encoding spatial and spectral characteristics. The extracted feature vectors $X = \{x_1, x_2, \dots, x_n\}$ represent high-dimensional lesion attributes.

The extracted features X serve as inputs to the GBDT model, which consists of an ensemble of decision trees. Each tree in the ensemble learns to correct the errors of its predecessor using gradient-based optimization is given in equation (4) [28]:

$$F_{m+1}(x) = F_m(x) + \gamma_m h_m(x) \tag{4}$$

where: $F_m(x)$ is the model at iteration m, $h_m(x)$ is the new decision tree trained on residuals, γ_m is the learning rate.

The final prediction is obtained by aggregating the outputs of multiple decision trees, leading to more precise lesion boundaries. This results in improved disease differentiation, particularly in cases where bacterial and fungal infections overlap. By iteratively reducing residual errors, GBDT refines lesion segmentation and improves classification accuracy. Figure 3 shows the architecture of GBDT within FFNN Hidden layer.

This integration within ViT introduces a differentiable tree-based learning process, ensuring a more adaptive classification system that maintains robust performance even under varying environmental conditions.

3.6. **Refining Extracted Features.** To preserve disease-specific spatial relationships and texture variations patch embeddings are crucial. So, Graph and Wavelet Patch Embeddings are employed to enhance ViT's feature embeddings.

Graph-based embeddings provide a structured approach to analyzing lesion spread and connectivity patterns, offering a significant advantage over traditional pixel-based methods. By treating each image patch as a node and establishing edges based on spatial relationships, this method effectively captures progressive infections and interconnected lesion regions that might otherwise be overlooked. To construct the graph, each 16×16 image patch is represented as a node v_i, with spatially adjacent patches connected via edges e_{ij}, forming a structured graph representation $G = (V, E)$. Edge weights are determined based on lesion similarity metrics such as color intensity, texture patterns, and shape continuity, ensuring that visually and structurally similar regions remain connected. Once the graph is constructed, a Graph Convolutional Network (GCN) is applied to propagate lesion-related information across the graph, enhancing feature extraction. The embedding for each node (patch) is computed using the equation (5) [29]:

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in N(i)} \frac{1}{c_{ij}} W^{(l)} h_j^{(l)}\right) \tag{5}$$

Where $h_i^{(l)}$ represents the feature representation of node i at layer l, $N(i)$ denotes the set of neighboring nodes, c_{ij} is a normalization constant, and $W^{(l)}$ represents the trainable weight matrix for feature transformation. The activation function σ, typically ReLU, enables non-linear transformations. These graph embeddings effectively capture lesion

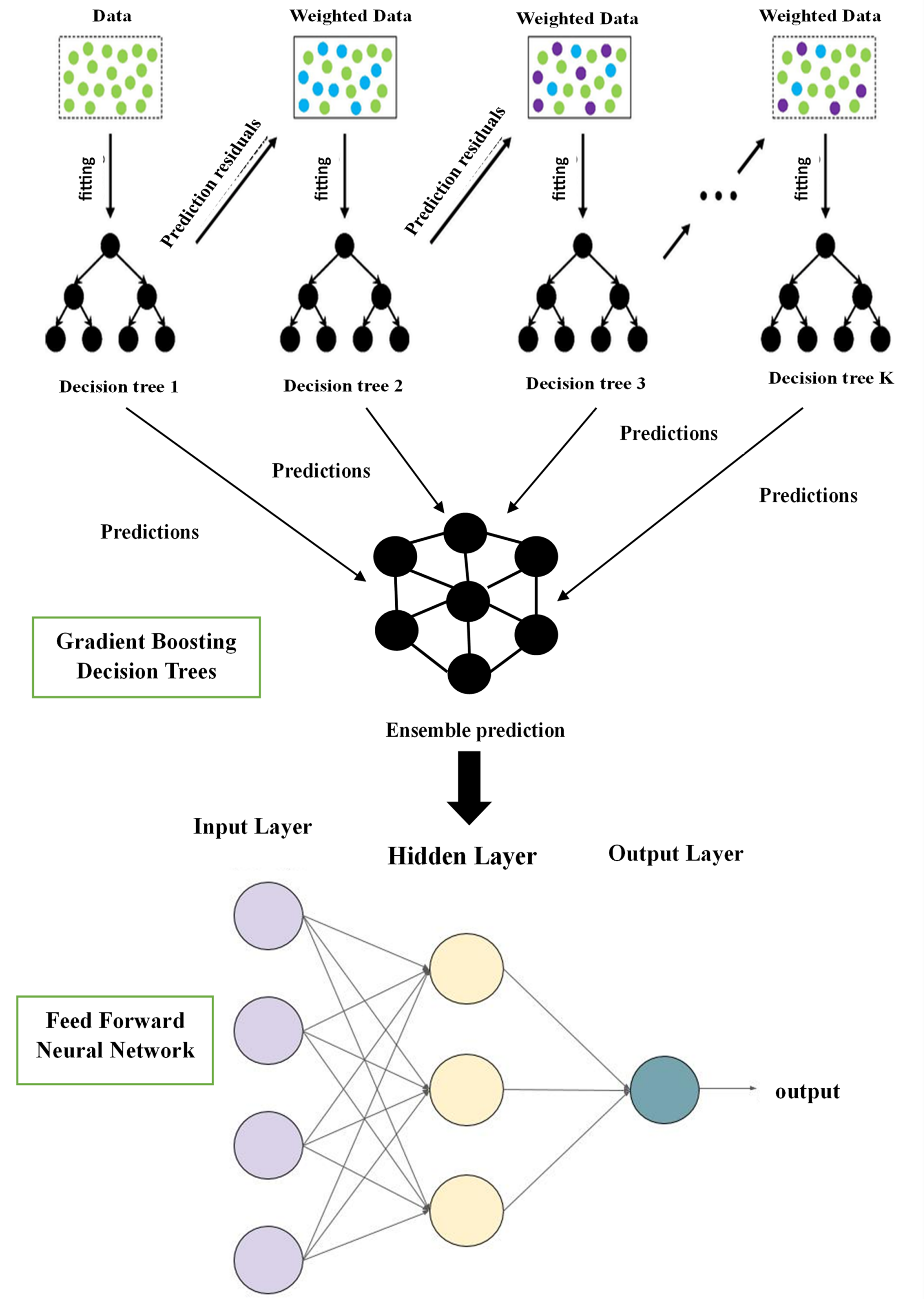


FIGURE 3. Architecture of GBDT within FFNN hidden Layer

connectivity trends, ensuring that progressive infections are recognized and preserved while maintaining long-range dependencies within the leaf structure. By integrating this graph-based modeling, the model achieves a context-aware representation, allowing it to track lesion spread dynamics and significantly enhance disease detection accuracy.

Wavelet-based embeddings enhance texture contrast, edge detection, and frequency-based lesion differentiation, making them highly effective for analyzing disease-induced tissue deformations. Since these deformations exhibit unique multi-scale frequency variations, wavelet transformation enables better lesion localization while preserving fine-grained features. The process begins with wavelet decomposition, where each 16 × 16 image patch undergoes a 2D Discrete Wavelet Transform (DWT), breaking it down into low-frequency and high-frequency components. Accurately, this decomposition is represented as equation (6) [30]:

$$W(u, v) = \sum_m \sum_n I(m, n)\psi^*_{u,v}(m, n) \tag{6}$$

where $I(m, n)$ is the image intensity at pixel (m, n), and $\psi_{u,v}$ represents the wavelet basis functions. High-frequency components highlight sharp lesion boundaries, aiding in early-stage disease detection, while low-frequency components retain global lesion structures, ensuring consistency in disease pattern analysis. After decomposition, the wavelet-extracted features undergo feature fusion, where they are concatenated with graph embeddings, ensuring that the model leverages both structural connectivity and frequency-domain lesion contrast for improved classification.

The integration of Graph and Wavelet Embeddings provides complementary advantages. Graph Embeddings preserve lesion connectivity, ensuring that spatial disease progression is effectively captured, while Wavelet Embeddings highlight sharp lesion boundaries, preventing texture loss in disease-induced tissue deformations. This combined strength enables the model to learn both spatially-aware lesion interactions and fine-grained texture details, significantly improving lesion segmentation, disease-specific feature extraction, and classification accuracy under diverse infection patterns. By leveraging both structural and frequency-domain insights, the model ensures a robust and highly precise classification of rice leaf diseases.

3.7. **Classification.** In rice leaf disease classification, lesion similarities, and symptom evolution pose significant challenges. To address this, ViT's extracted features are passed to a Hierarchical Multi-Layer Perceptron (MLP) which integrates XGBoost and Random Forest (RF) for precise disease differentiation. This hybrid approach combines deep learning with ensemble-based decision trees, ensuring fine-grained classification while maintaining robustness against variations in symptoms.

XGBoost enhances feature routing by assigning higher importance to lesion characteristics that distinguish visually similar infections, such as Rice Blast and Brown Spot. It utilizes gradient-boosted decision trees, where each tree refines the classification by correcting errors from the previous one, represented as equation (7) [31]:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{7}$$

Where $f_k(x_i)$ represents each decision tree, and F is the set of all possible trees.

Meanwhile, Random Forest (RF) enhances robustness against symptom evolution by training multiple decision trees on random subsets of the dataset. This ensemble technique prevents overfitting and ensures stable predictions across various disease stages. Classification in RF follows majority voting given as equation (8) [32]:

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \ldots, h_n(x)\} \tag{8}$$

where $h_n(x)$ represents predictions from each decision tree. The combination of XGBoost and Random Forest provides complementary advantages: XGBoost prioritizes feature importance and fine-tunes lesion differentiation, while Random Forest enhances generalization, ensuring effectiveness across different growth phases. This synergy prevents overfitting while ensuring high accuracy in disease classification across various infection stages.

Given that rice leaves often exhibit cross-infections where fungal and bacterial diseases coexist, the classification framework is designed to assign both a fungal and a bacterial disease label per infected leaf. This biologically relevant dual-disease classification follows a structured approach. The model is trained on two independent classification pipelines, one for fungal and another for bacterial infections. Each lesion is first assigned a fungal class and then a bacterial class, ensuring accurate detection of mixed infections.

For final classification, a Hierarchical MLP refines predictions from XGBoost and Random Forest, improving decision boundaries. The final output is generated using equation (9) [33]:

$$P(y|X) = \sigma(W_2 \cdot ReLU(W_1 X + b_1) + b_2) \tag{9}$$

where W_1, W_2 are weight matrices, b_1, b_2 are biases, and σ is the activation function. The separate fungal and bacterial classification approach prevents confusion between visually similar diseases and accounts for real-world co-infections, ensuring practical agricultural recommendations.

The combination of XGBoost, Random Forest, and Hierarchical MLP allows the model to capture subtle feature variations, ensure robustness across infection stages, and integrate predictions for a final, biologically relevant classification. By leveraging gradient boosting, ensemble learning, and deep neural networks, this classification pipeline achieves higher accuracy, robustness, and real-world applicability in rice leaf disease detection.

Overall, the Vision Rotary Attention Decision Forest Perceptron Transformer integrates deep learning and ensemble techniques for precise rice leaf disease detection. It segments images into 16 × 16 patches, preserving fine-grained lesion details. RoPE and Fourier embeddings enhance spatial and frequency-based lesion representations, while MSA extracts key features without losing local dependencies. GBDT within ViT improves lesion segmentation and disease-specific texture pattern learning. The Decision Forest Perceptron combines XGBoost and RF for robust classification. This structured pipeline ensures high-accuracy classification of one fungal and one bacterial infection per leaf, effectively handling complex lesion structures.

4. Results and Discussion.

In this part, the outcomes of the suggested model have been shown. When compared to other current methods, the findings shown that the suggested model offers superior fungal illness and viral infection detection and classification, as well as prediction accuracy.

4.1. Experimental Setup.

This section provides a comprehensive overview of the performance and implementation results of the suggested system, which was simulated in Python. Additionally, a proportional analysis is provided to verify that the suggested system functions as planned.

- OS: Windows 10 (64-bit)
- RAM: 8GB
- Processor: Intel i5
- Tool: Python

4.2. Dataset Description.

The dataset used in this study comprises a total of 1,400 images of paddy leaves, categorized into four distinct paddy leaf diseases namely Bacterial Leaf Blight, Brown Spot, Leaf Blast and Bacterial Leaf Streak. The dataset is organized into training and testing folders, with 80:20 split, i.e., 1,120 images are allocated for training and 280 images for testing. Each disease class is equally represented to maintain class balance during both training and evaluation. To enhance the model's robustness and generalization, data augmentation techniques were applied to each image in the training

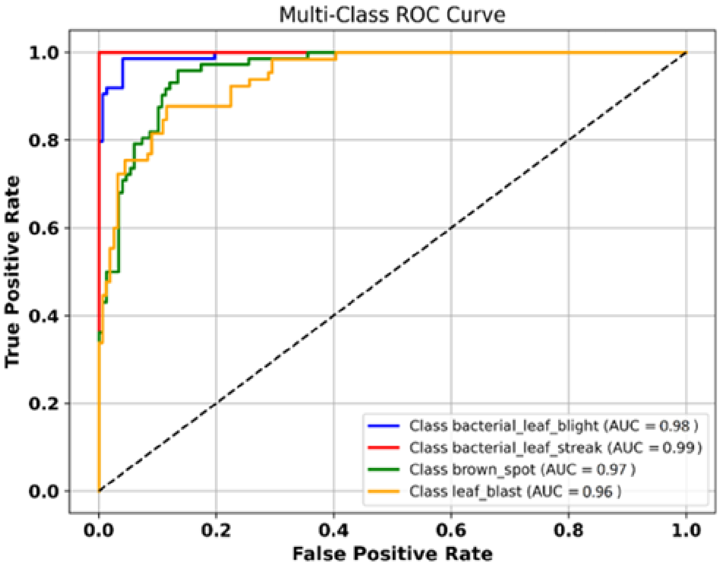


Figure 4. ROC curve for the classification of viral and bacterial diseases

set. For a substantial increase in the effective training data, five augmentation methods were used per image as Zooming, Rotation, Horizontal flipping, Brightness adjustment and Shearing transformation. These augmentations help simulate real-world variations in image capture conditions and improve the model's ability to detect diseases under diverse scenarios. The dataset is collected from `https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset`.

Dataset Preprocessing:

- All images were resized to 256×256 pixels for uniformity.
- Image normalization was applied to ensure consistent pixel intensity distributions.

Data Splitting:

- The dataset was divided into 80% training and 20% testing using stratified sampling to maintain class balance.

Model Training:

- The Vision Rotary Attention Decision Forest Perceptron Transformer (VRAD-FPT) was implemented in Python using the PyTorch framework.
- Training was conducted over 30 epochs with a batch size of 32.
- The Adam optimizer was used with an initial learning rate of 0.001 and a learning rate decay factor of 0.9 every 10 epochs.
- Early stopping was implemented based on validation loss to prevent overfitting.

Evaluation Protocol:

- Metrics such as Accuracy, Precision, Recall, F1-score, Specificity, Sensitivity, MCC, NPV, and AUC-ROC were computed on the test set.
- Each model was trained and tested five times with different random seeds to ensure result stability, and the average values are reported.
- ROC curves were plotted per class to assess classification performance.

4.3. **Simulated results of the proposed model.** In this section simulation outputs of the suggested model, are given with implemented results, which are described below.

Figure 4 displays the Multi-Class ROC Curve, which plots the True Positive Rate against the False Positive Rate at various thresholds to assess a classification model's performance. The classifier achieves high AUC values: bacterial leaf blight of 0.98, bacterial leaf streak 0.99, brown spot of 0.97, and leaf blast of 0.96. The diagonal dashed line represents a random classifier with no discriminative power. The model effectively distinguishes between disease classes, as indicated by AUC values close to 1. The high AUC scores suggest strong predictive accuracy across all categories. Overall, the classifier demonstrates robust performance in identifying bacterial and fungal infections in paddy leaves. The model handles co-existing infections by assigning one fungal and one bacterial label per image using dual classification pipelines ensuring accurate identification of co-existing infections, mitigating misclassification due to lesion overlap. Severity grading is indirectly managed through patch-based lesion segmentation and spatial-frequency embedding, which capture fine-grained lesion variations and progression patterns. This enables precise classification even in cases of overlapping or asymmetrical lesion structures. The model thus supports both cross-infection detection and nuanced symptom representation.

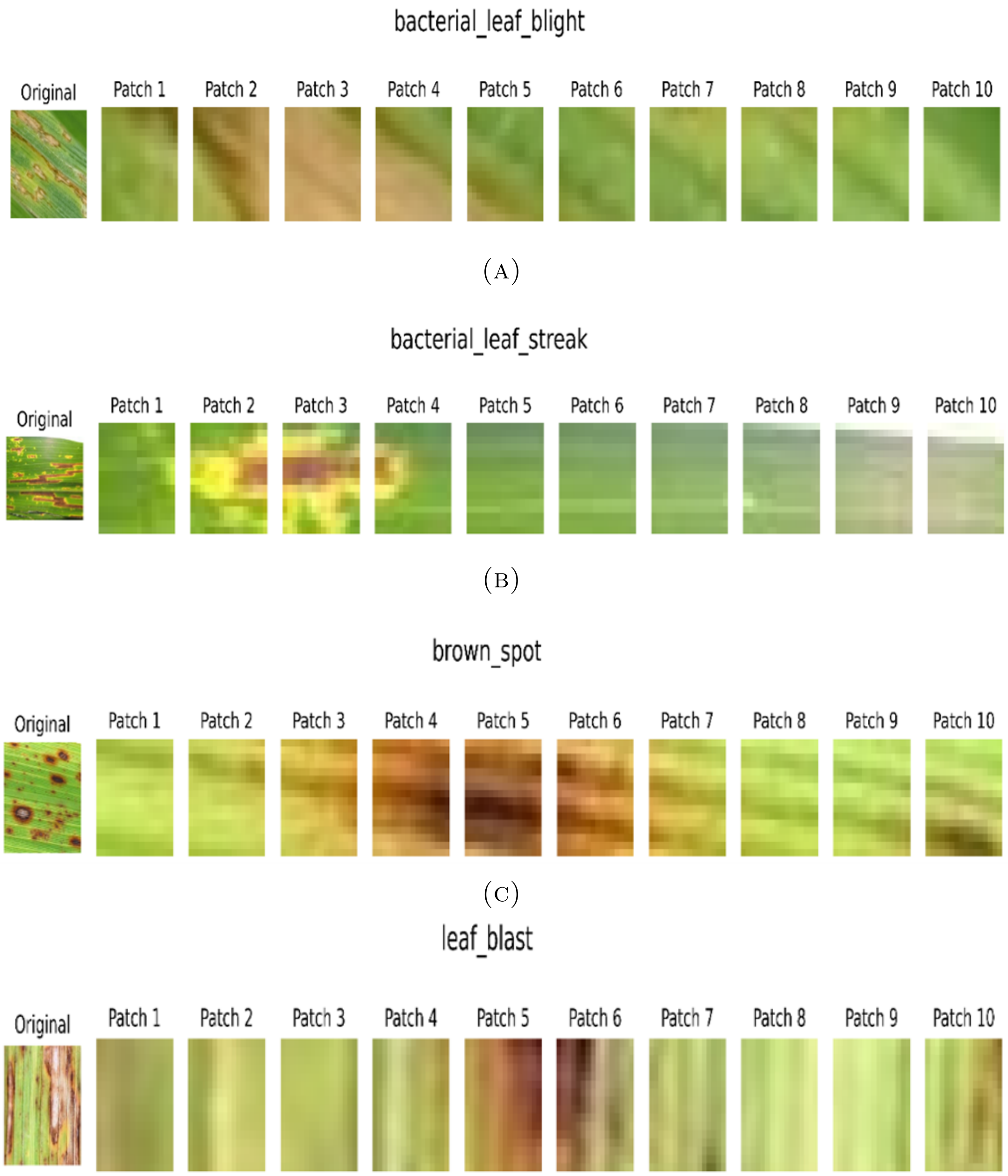
Figure 5 (a-d) showcases fungal and viral infections using one original image and ten different patches extracted from infected regions for detailed analysis. Each patch highlights distinct infection patterns, such as circular necrotic spots for fungal diseases and mosaic or streak-like discoloration for viral infections. Patch-based segmentation helps in capturing localized disease symptoms, aiding in precise classification and early detection. The variations across patches emphasize the heterogeneous nature of infection spread, which is crucial for developing robust disease identification models.

Figure 6 illustrates the segmentation results for bacterial and fungal diseases in paddy leaves, where original images show infected leaves and corresponding segmentation images highlight diseased regions in white. Bacterial infections appear as elongated streaks along veins, turning yellow-brown over time, with segmentation effectively isolating these lesions despite minor noise. Fungal infections manifest as circular lesions with brown margins and grayish-white centers, with segmentation accurately detecting their discrete distribution. While bacterial streaks show linear necrotic patterns, fungal spots are scattered and irregular, making segmentation crucial for precise classification. The results effectively differentiate bacterial from fungal infections, though challenges like over-segmentation in bacterial cases and under-segmentation in fungal lesions persist.

4.4. **Analysis of Proposed Models Performance across various metrics.** The Paddy Leaf Disease performance metrics of proposed model This section provided a detailed explanation of how the suggested method was predicted and categorized based on the results obtained.

Figure 7 depicts the Training accuracy and validation accuracy of proposed model. Initially, at the beginning training accuracy of the proposed model is 0.43 then it shows a steady increase, when the epoch is 30 training accuracy is 0.99 indicating that the model's performance on the training data improves as the number of epochs increases. MSA refines feature extraction by capturing long-range dependencies across different leaf patches, ensuring that disease patterns are accurately identified. Validation Accuracy also shows an increase, at the beginning validation accuracy is 0.69, and then epoch is 30 validation accuracy is 0.83 but with noticeable fluctuations. This suggests that while the model's performance on the validation data generally improves, it is not as consistent as the training accuracy.

Figure 8 illustrates the training and validation loss trends of the proposed model. Initially, the training loss starts at 1.74 and progressively decreases to 0.03 by epoch 30, indicating a steady reduction in model error as training progresses. Similarly, the validation loss begins at 0.75 and declines to 0.59, though with some fluctuations, suggesting



(A)

(B)

(C)

(D)

FIGURE 5. (a-d). Identifying fungal and viral infections using different patch

that while the model generalizes well, its performance on validation data is slightly inconsistent. The Hierarchical MLP with XGBoost and Random Forest enhances decision boundary optimization for complex lesion patterns, effectively minimizing both training and validation errors.

Figure 9 illustrates the precision improvement of the proposed model over increasing epochs. At 10 epochs, the precision starts at 0.36, gradually improving as training progresses, reaching 0.98 at 30 epochs. The GBDT-enhanced feedforward networks refine feature representation, enabling the model to capture intricate lesion characteristics more effectively, ultimately enhancing precision.

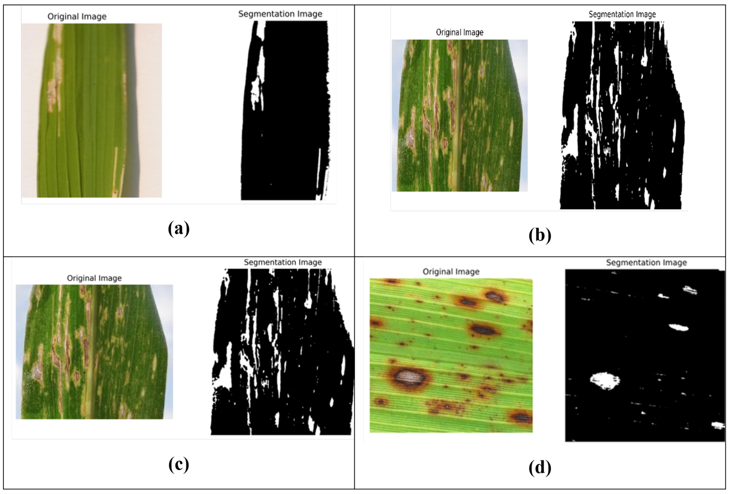


FIGURE 6. Segmentation results of Bacterial and fungal diseases

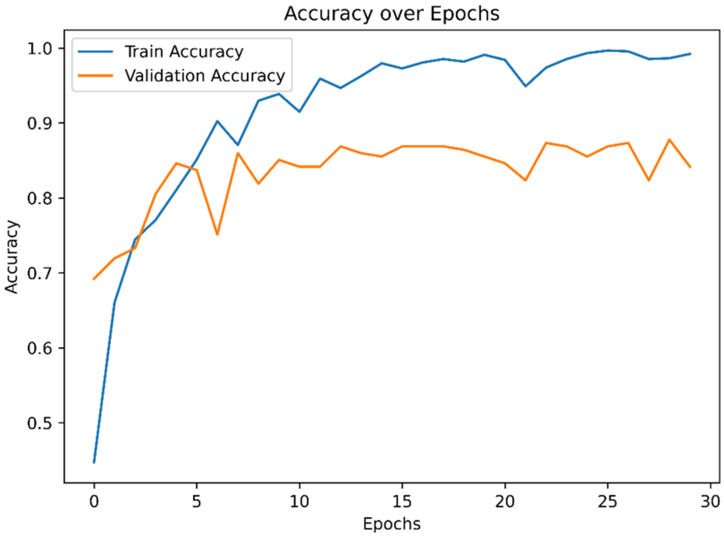


FIGURE 7. Training accuracy and validation accuracy

Figure 10 shows the process of proposed model over training epochs from 10 to 30. Initially, recall is 0.62, gradually improving and reaching 0.97 by epoch 30. A notable increase is observed at epoch 15, reflecting an accelerated learning phase. The Hierarchical MLP with XGBoost and Random Forest enhances recall by effectively capturing complex lesion variations and dynamic transformations. XGBoost refines feature importance ranking, while Random Forest improves generalization, minimizing overfitting and ensuring better lesion detection.

The proposed model's F1 score improved throughout several training epochs, as seen in Figure 11. Starting at 0.43 at 10 epochs, the score gradually increases, reaching 0.99 by epoch 30. The Hierarchical MLP with XGBoost and Random Forest enhances the F1 score by balancing precision and recall in paddy leaf disease classification. XGBoost optimizes feature selection, ensuring key lesion patterns are effectively recognized, while Random Forest strengthens generalization, reducing overfitting and improving recall for diverse lesion variations.

Figure 12 shows how the proposed model's specificity gradually improved over several training epochs.. Beginning at 0.34 at 10 epochs, specificity steadily improves, reaching

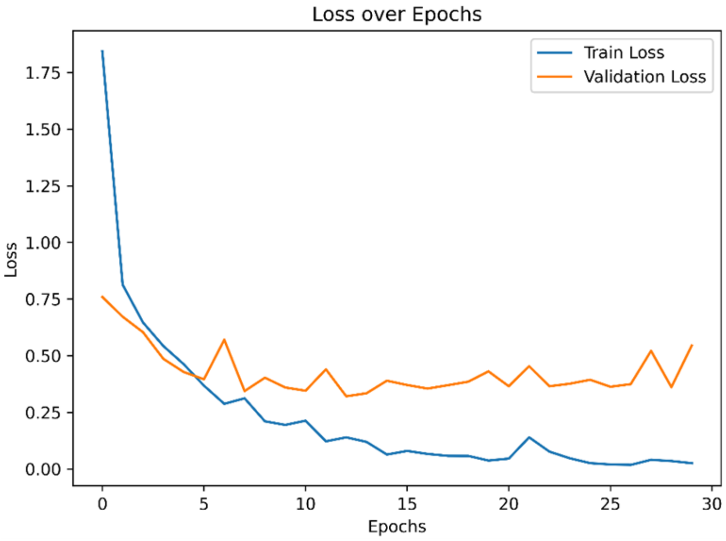


Figure 8. Training loss and validation loss

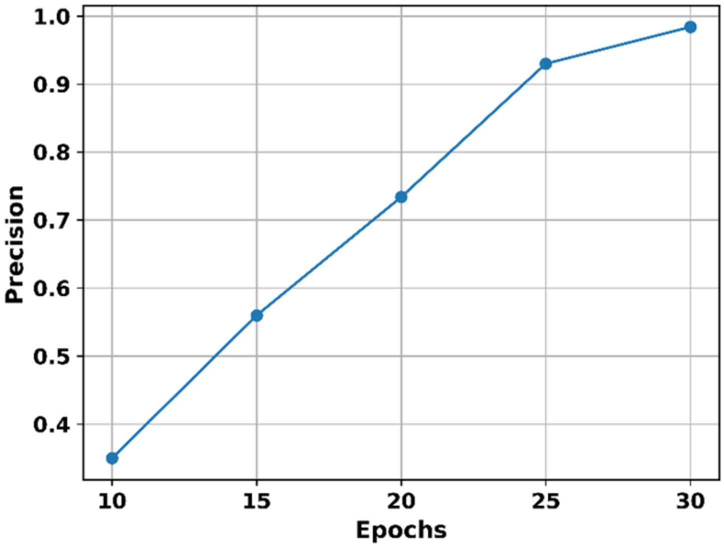


Figure 9. Analysis of Proposed Models Precision

0.98 by epoch 30. This consistent increase highlights the model's capacity for continuous learning and refinement, leading to more precise predictions over time. Optimizing decision boundaries, lowering misclassification, and enhancing the model's capacity to precisely identify non-diseased areas are all made possible by the GBDT-enhanced feed-forward networks.

Sensitivity of proposed model improves significantly over training epochs, as illustrated in Figure 13, starting at 0.42 at 10 epochs and reaching 0.97 by epoch 30. This steady increase highlights the model's ability to correctly identify diseased regions while minimizing false negatives. The Hierarchical MLP with XGBoost and Random Forest contributes to this improvement by refining feature representation and classification. XGBoost enhances feature importance ranking, ensuring the model focuses on critical lesion patterns, while Random Forest improves generalization by reducing overfitting, thereby increasing the model's ability to detect subtle disease variations accurately.

Figure 14 illustrates the progression of detection accuracy in the proposed model as training epochs increase. Initially, at 10 epochs, the precision is 0.76, and it steadily

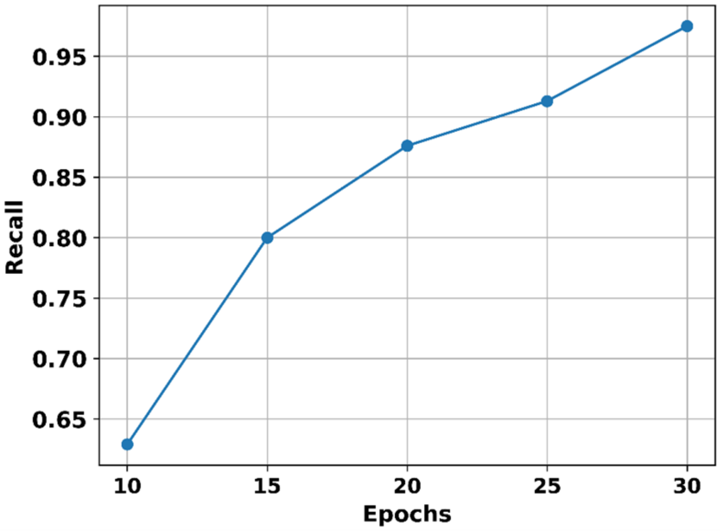


FIGURE 10. Analysis of Proposed models Recall

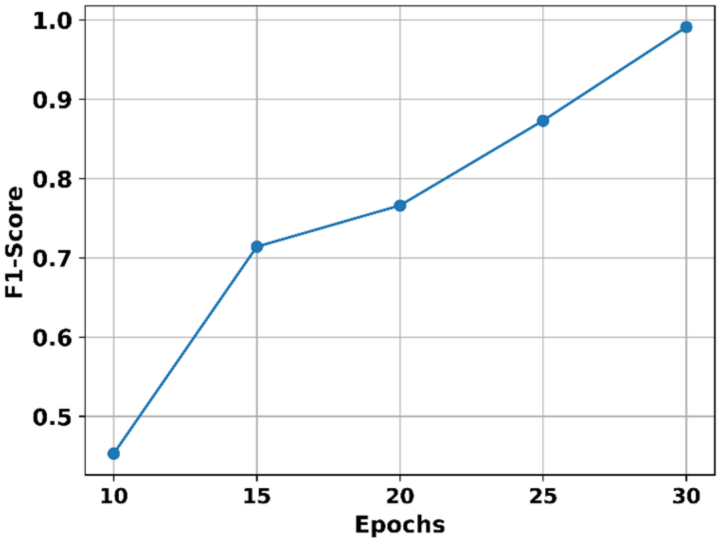


FIGURE 11. Analysis of Proposed Models F1 score

improves, reaching 0.98 by epoch 30. Graph and Wavelet Patch Embeddings enhance lesion connectivity and texture contrast, enabling the model to effectively differentiate between diseased and healthy regions. This refinement minimizes false positives and false negatives, thereby improving overall detection accuracy.

Figure 15 illustrates the progression of Matthews Correlation Coefficient (MCC) in proposed model over training epochs. Initially, at epoch 10, the MCC is 0.41, and it steadily increases, reaching 0.98 by epoch 30. This upward trend signifies the model's improving predictive capability and balanced classification. The Multihead Self-Attention & GBDT-Enhanced Feedforward Networks play a crucial role in this enhancement by refining adaptive feature learning and decision boundaries, effectively reducing false positives and false negatives. This results in a more robust and reliable classification, contributing to the continuous improvement of MCC.

Figure 16 presents the time complexity of the proposed model across training epochs. Initially, at epoch 10, the complexity is 85ms, rising to 120ms at epoch 15 before reducing to 76ms by epoch 30. This trend indicates an initial computational overhead, followed

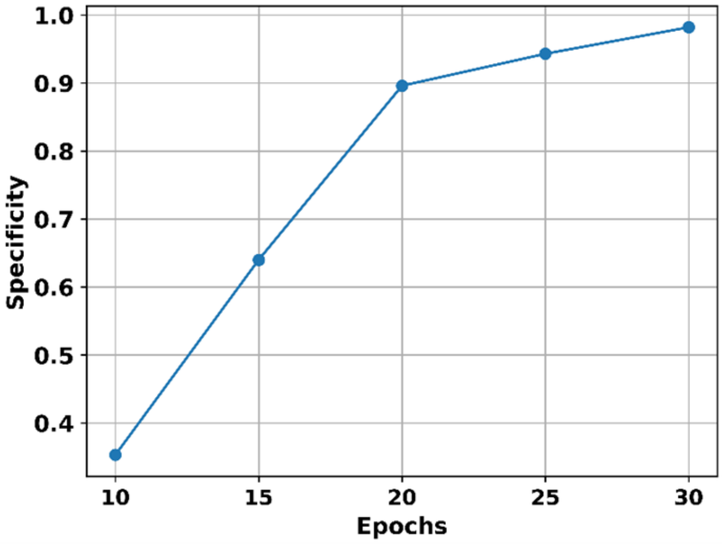


Figure 12. Proposed Models Specificity

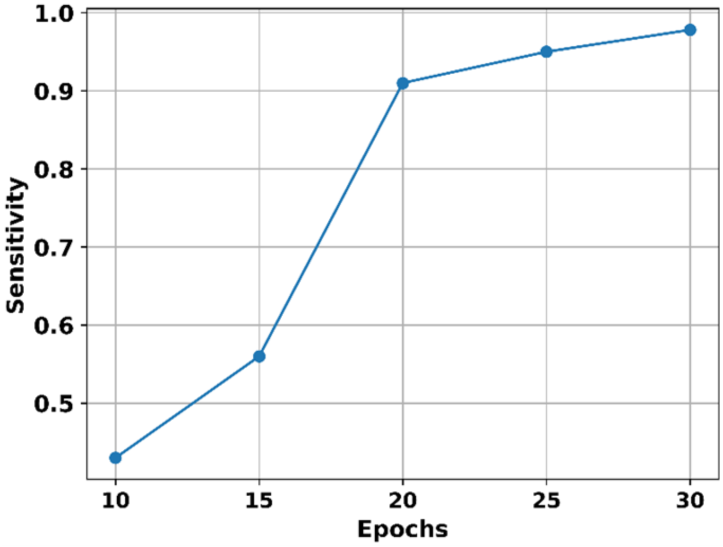


Figure 13. Proposed Models Sensitivity

by optimization as training progresses. Fourier Position Encoding enhances spatial feature representation efficiently without extensive convolution operations. RoPE preserves sequential dependencies while minimizing computational overhead.

4.5. **Analysis of Proposed model with other Approaches.** By comparing the outcomes of the proposed model with those of current methods and presenting its output based on several metrics, this part demonstrates the effectiveness of the model.

A comparison of proposed model's accuracy with other current approaches is shown in Figure 17. The proposed approach is contrasted with other current methods such as CNNIR-OWELM, VGG16, DeepNN, and ConvNN [34]. In comparison to CNNIR-OWELM's accuracy of 94.2%, VGG16's accuracy of 92.9%, DeepNN's accuracy of 90%, and ConvNN's accuracy of 93.8%, the proposed model's accuracy is 99.2%. Therefore, it appears that the suggested model's prediction accuracy for leaves is higher than that of the DeepNN model.

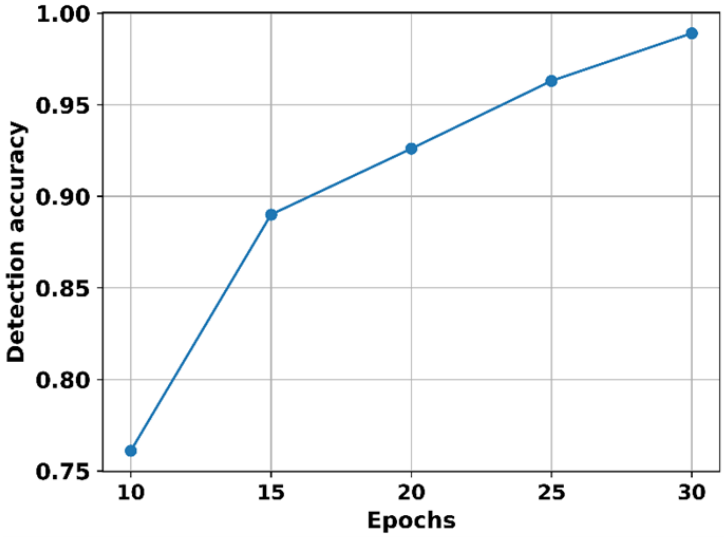


Figure 14. Detection accuracy of the Proposed Model

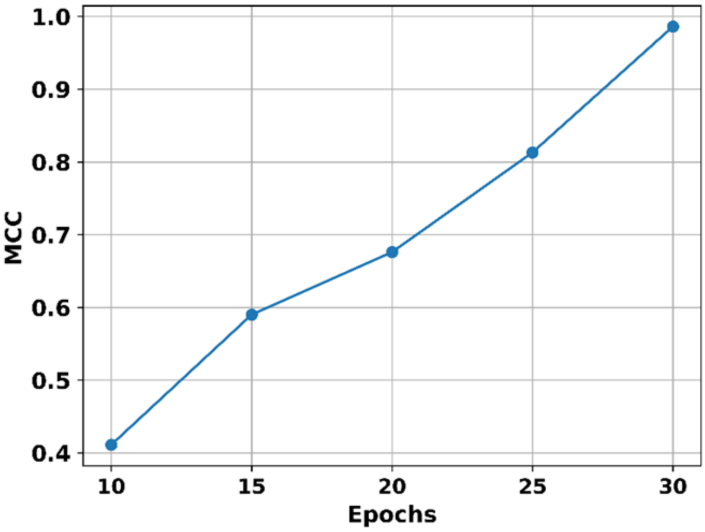


Figure 15. Proposed Models MCC

The precision comparison of the varied models presented with existing differential techniques is depicted in Figure 18 [34]. The proposed framework is compared to other methods that are currently in use, such as CNNIR-OWELM, VGG16, ConvNN, and SIFTSVM [28]. The precision achieved by the proposed model is 98.4% while the precision of CNNIR-OWELM is 92%, VGG16 is 90%, ConvNN is 94%, and SIFTSVM is 91%. The proposed model is a good classifier, greatly outranking VGG16 in this respect.

The proposed model's sensitivity comparison with other current methods is shown in Figure 19 [34]. The proposed method's sensitivity is comparatively high, at about 97.8%, in contrast to CNNIR-OWELM, VGG16, ConvNN, and SIFTSVM, which have sensitivities of 91%, 89%, 96%, and 90%, respectively.

The comparison of the suggested model's specificity with other current methods is shown in Figure 20. The suggested model's specificity is 98.2%, while CNNIR-OWELM, VGG16, ConvNN, and SIFTSVM have respective specificities of 95%, 92%, 94%, and 88%.

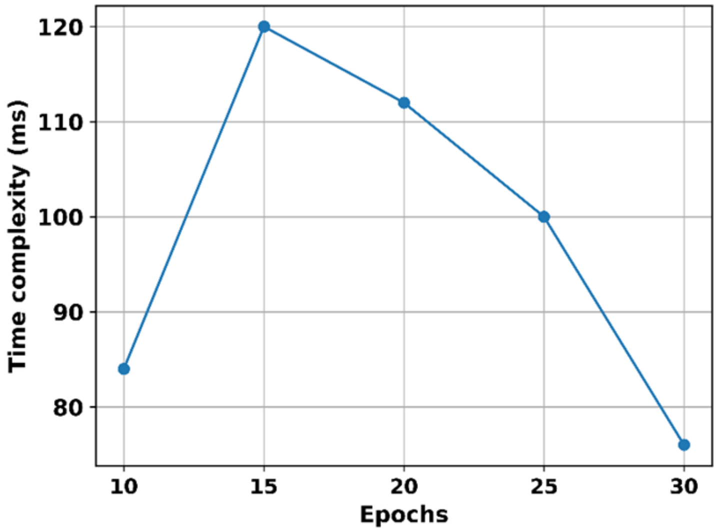


Figure 16. Time Complexity

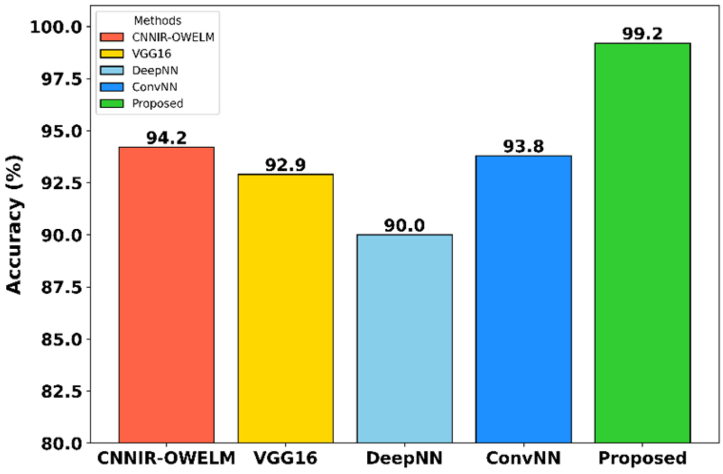


Figure 17. Comparison of Accuracy in proposed model

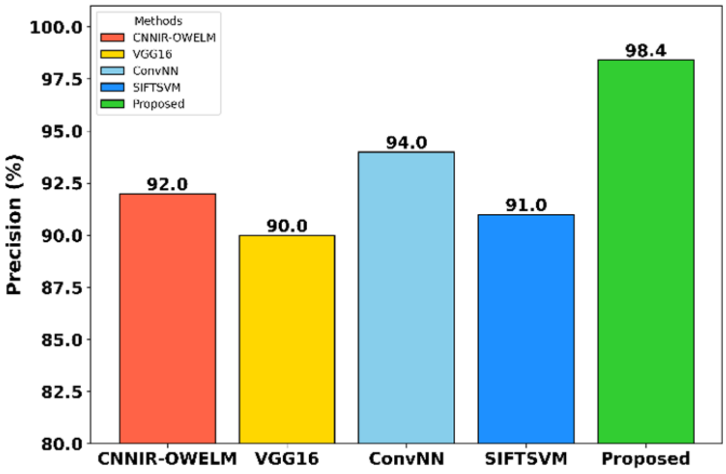


Figure 18. Comparison of precision in proposed model

The proposed model's F1 score was 99.1%, whereas CNNIR-OWELM, VGG16, ConvNN, and SIFTSVM obtained F1 scores of 91%, 89%, 94%, and 86.7%, respectively.

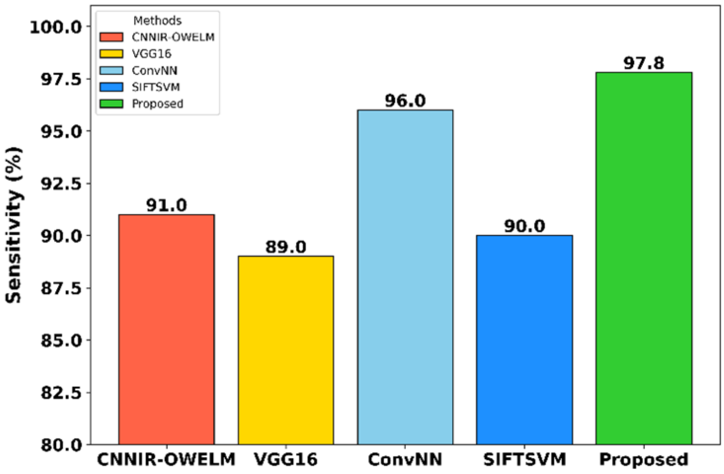


Figure 19. Comparison of sensitivity with other approaches

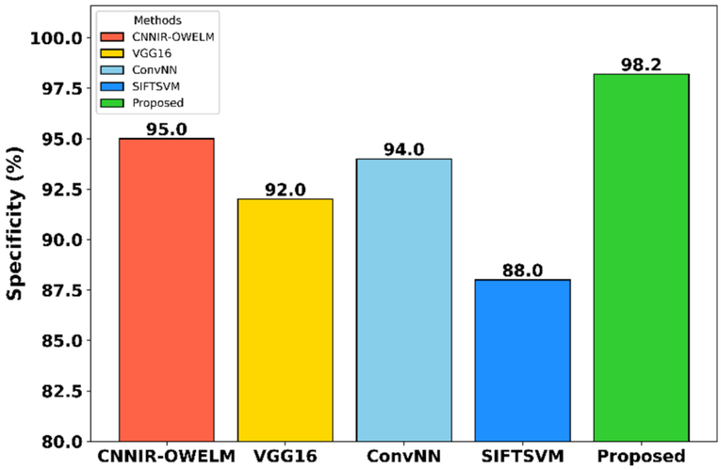


Figure 20. Comparison of specificity with other approaches

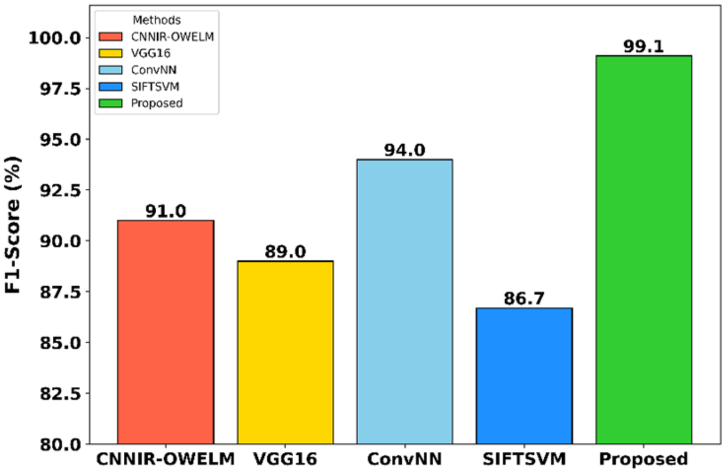


Figure 21. F1 score comparison with other approaches

The suggested model achieves a recall of 97.5%, while the corresponding values for APS-DCCNN, AlexNet, CNN, and DSGAN are 76%, 82%, 85%, and 92% [35].

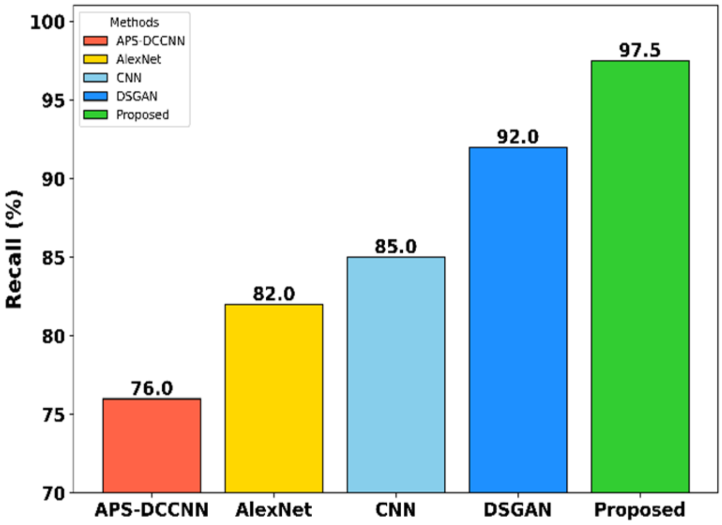


FIGURE 22. Recall comparison with other approaches

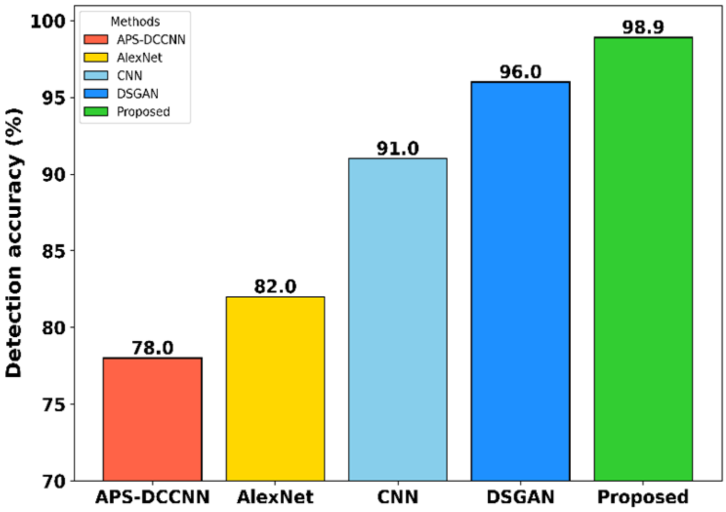


FIGURE 23. Detection accuracy Comparison

The suggested model achieves a detection accuracy of 98.9%, while APS-DCCNN, AlexNet, CNN, and DSGAN are 78%, 82%, 91%, and 96% [35].

The suggested model's temporal complexity is 76 ms, while that of APS-DCCNN, AlexNet, CNN, and DSGAN is 146 ms, 120 ms, 185 ms, and 98 ms, respectively [35].

In contrast to WOA-ANN, SVM, ANN, and KNN, which have respective MCCs of 97.7%, 83.4%, 86.5%, and 90.6%, the suggested model's MCC is 98.6% [36].

The proposed model's NPV is 99.4%, while WOA-ANN, SVM, ANN, and KNN have respective NPVs of 94.4%, 95.5%, 96.8%, and 96.8% [36].

Overall, the proposed model shows advancements in disease prediction and management in agricultural fields, particularly in classifying fungal infections and viral diseases affecting paddy leaves. By integrating techniques like Multihead Self-Attention, Gradient Boosting Decision Trees (GBDT), and Hierarchical MLP, the model achieves outstanding accuracy and performance metrics compared to existing approaches. Despite complex interactions and changing environmental conditions, the model has a high accuracy of 99.2%, precision of 98.4%, recall of 97.5%, F1-score of 99.1%, sensitivity of 97.8%, specificity of 98.2%, detection rate of 98.9%, MCC of 98.6%, and NPV of 99.4%.

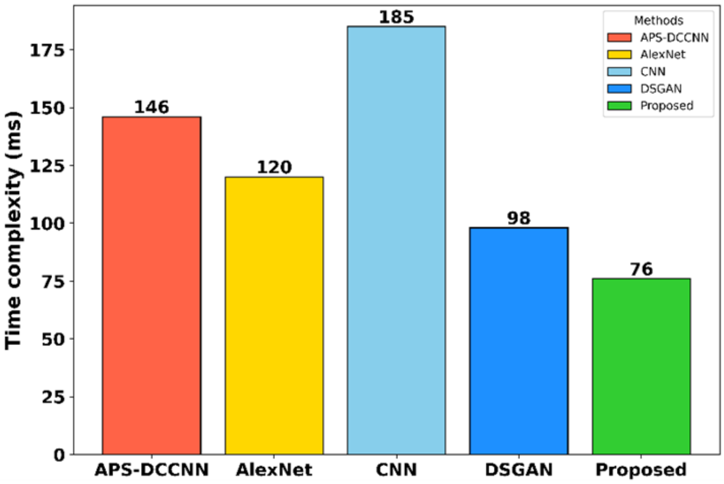


FIGURE 24. Time complexity Comparison

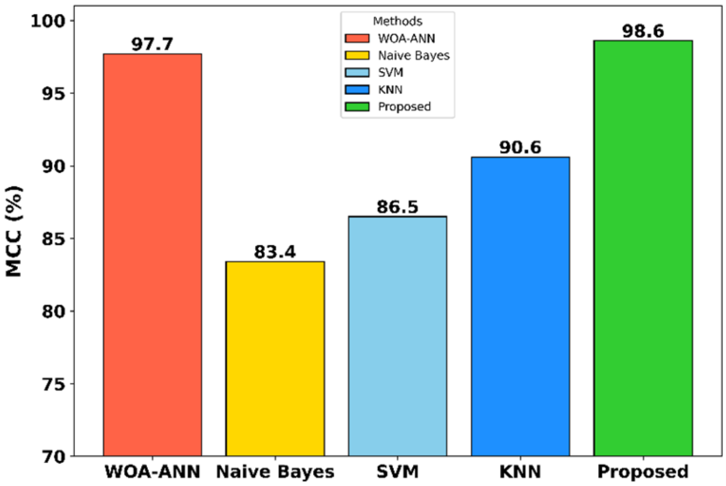


FIGURE 25. Comparison of MCC in proposed model

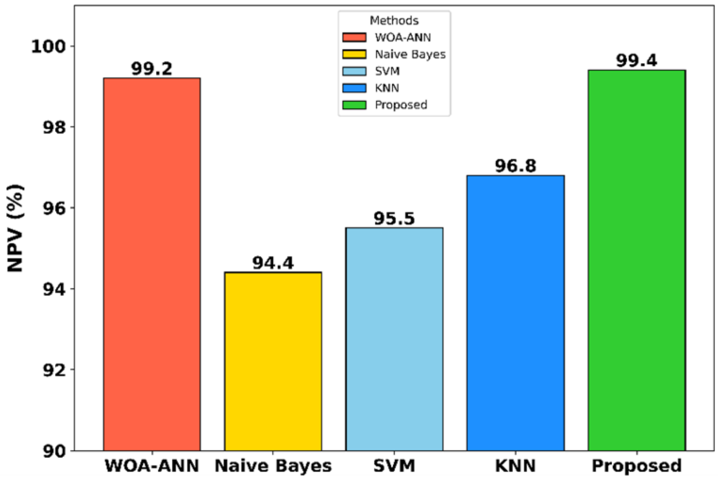


FIGURE 26. Comparison of NPV in proposed model

5. **Conclusion.** Paddy leaf disease prediction and classification are critical in image processing, so a novel Vision Rotary Attention Decision Forest Perceptron Transformer is proposed to ensure superior accuracy through a structured six-stage pipeline. By integrating

RoPE, Fourier embeddings, Multihead Self-Attention (MSA), GBDT-enhanced ViT, and Decision Forest Perceptron (XGBoost + RF), the model effectively captures fine-grained lesion details while overcoming cross-infection complexities. The patch-based image processing ensures localized disease feature retention, while MSA and GBDT-enhanced ViT optimize spatial-frequency feature extraction. The inclusion of Graph & Wavelet Patch Embeddings further refines lesion connectivity modeling, improving differentiation between fungal and bacterial infections. The Hierarchical MLP Classifier ensures precise classification through a combination of XGBoost and Random Forest, reducing misclassification errors. Experimental results validate its superiority over existing models, achieving 99.2% accuracy, 98.4% precision, 97.5% recall, 99.1% F1-score, 97.8% sensitivity, 98.2% specificity, 98.9% detection rate, 98.6% MCC, and 99.4% NPV. These outstanding performance metrics demonstrate its ability to accurately identify multiple disease types, significantly improving early disease detection and enabling timely interventions. While the proposed model shows high accuracy and robustness, the current hardware setup limit scalability for large datasets or real-time applications, particularly when processing high-resolution images. Future work could involve deploying the model on more powerful systems or optimizing it for edge devices to support broader and faster deployment. Additionally, expanding the dataset diversity could further enhance generalization across varied field conditions.

REFERENCES

[1] M.A. Islam, M.N.R. Shuvo, M. Shamsojjaman, S. Hasan, M.S. Hossain, and T. Khatun, "An automated convolutional neural network based approach for paddy leaf disease detection". *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, 2021.

[2] X. Du, B. Yuan, J. Li, G. Yin, Y. Qiu, J. Zhao, X. Duan, Y. Wu, T. Lin, and Y. Zhou, "Distribution, behavior, and risk assessment of chlorinated paraffins in paddy plants throughout whole growth cycle." *Environment International*, vol. 167, pp. 107404, 2022.

[3] N.S.H. Kurniawan, I.A.P. Kirana, A.S. Abidin, A. Jupri, S. Widyastuti, A. Hernawan, A. Nikmatullah, H. Sunarpi, and E.S. Prasedya, "Analysis of leaf chlorophyll content of paddy plants during vegetative stage grown in soil media containing macroalgae organic fertilizer." In *IOP Conference Series: Earth and Environmental Science*, vol. 913, no. 1, pp. 012025, 2021, November. IOP Publishing.

[4] R. Swathika, S. Srinidhi, N. Radha, and K. Sowmya, "Disease Identification in paddy leaves using CNN based Deep Learning." In *2021 third international conference on intelligent communication technologies and virtual mobile networks (ICICV)*, pp. 1004-1008, 2021, February. IEEE.

[5] M.M. Faizal Azizi, and H.Y. Lau, "Advanced diagnostic approaches developed for the global menace of rice diseases: a review." *Canadian Journal of Plant Pathology*, vol. 44, no. 5, pp. 627-651, 2022.

[6] Z. Wang, B. Chen, T. Zhang, G. Zhou, and X. Yang, "Rice stripe mosaic disease: characteristics and control strategies". *Frontiers in Microbiology*, vol. 12, pp. 715223, 2021.

[7] X. Tan, H. Zhang, Z. Yang, Z. Wei, Y. Li, J. Chen, and Z. Sun, "NF-YA transcription factors suppress jasmonic acid-mediated antiviral defense and facilitate viral infection in rice." *PLoS Pathogens*, vol. 18, no. 5, pp. e1010548, 2022.

[8] M.S. Ngalimat, E. Mohd Hata, D. Zulperi, S.I. Ismail, M.R. Ismail, N.A.I. Mohd Zainudin, N.B. Saidi, and M.T. Yusof, "Plant growth-promoting bacteria as an emerging tool to manage bacterial rice pathogens." *Microorganisms*, vol. 9, no. 4, pp. 682, 2021.

[9] Y. Xue, M. Hu, S. Chen, A. Hu, S. Li, H. Han, G. Lu, L. Zeng, and J. Zhou, "Enterobacter asburiae and Pantoea ananatis causing rice bacterial blight in China." *Plant Disease*, vol. 105, no. 8, pp. 2078-2088, 2021.

[10] M.A. Azim, M.K. Islam, M.M. Rahman, and F. Jahan, "An effective feature extraction method for rice leaf disease classification." *Telkomnika*, vol. 19, no. 2, pp. 463-470, 2021.

[11] S. Chen, K. Zhang, Y. Zhao, Y. Sun, W. Ban, Y. Chen, H. Zhuang, X. Zhang, J. Liu, and T. Yang, "An approach for rice bacterial leaf streak disease segmentation and disease severity estimation". *Agriculture*, vol. 11, no. 5, pp. 420, 2021.

[12] S.K. Upadhyay, and A. Kumar, “A novel approach for rice plant diseases classification with deep convolutional neural network.” *International Journal of Information Technology*, vol. 14, no. 1, pp. 185-199, 2022.

[13] E. Moupojou, A. Tagne, F. Retraint, A. Tadonkemwa, D. Wilfried, H. Tapamo, and M. Nkenlifack, “FieldPlant: A dataset of field plant images for plant disease detection and classification with deep learning.” *IEEE Access*, vol. 11, pp. 35398-35410, 2023.

[14] M. Deb, K.G. Dhal, R. Mondal, and J. Gálvez, “Paddy disease classification study: a deep convolutional neural network approach.” *Optical Memory and Neural Networks*, vol. 30, pp. 338-357, 2021.

[15] R. Yakkundimath, G. Saunshi, and S. Palaiah, “Automatic methods for classification of visual based viral and bacterial disease symptoms in plants.” *International Journal of Information Technology*, vol. 14, no. 1, pp. 287-299, 2022.

[16] P.I. Ritharson, K. Raimond, X.A. Mary, J.E. Robert, and J. Andrew, “DeepRice: A deep learning and deep feature based classification of Rice leaf disease subtypes.” *Artificial Intelligence in Agriculture*, vol. 11, pp. 34-49, 2024.

[17] A. Chakrabarty, S.T. Ahmed Islam, S.M. Aziz, and S.S. Maidin, “An interpretable fusion model integrating lightweight CNN and transformer architectures for rice leaf disease identification.” *Ecological Informatics*, vol. 82, pp. 102718, 2024.

[18] W. Shafik, A. Tufail, C. De Silva Liyanage, and R.A.A.H.M. Apong, “Using transfer learning-based plant disease classification and detection for sustainable agriculture.” *BMC Plant Biology*, vol. 24, no. 1, pp. 136, 2024.

[19] S.P. Singh, K. Pritamdas, K.J. Devi, and S.D. Devi, “Custom convolutional neural network for detection and classification of rice plant diseases.” *Procedia Computer Science*, vol. 218, pp. 2026-2040, 2023.

[20] S. Maheswaran, S. Sathesh, P. Rithika, I.M. Shafiq, S. Nandita, and R.D. Gomathi, “Detection and classification of paddy leaf diseases using deep learning (cnn).” In *International Conference on Computer, Communication, and Signal Processing*, pp. 60-74, 2022, February. Cham: Springer International Publishing.

[21] K.S. Archana, S. Srinivasan, S.P. Bharathi, R. Balamurugan, T.N. Prabakar, and A.S.F. Britto, “A novel method to improve computational and classification performance of rice plant disease identification”. *The Journal of Supercomputing*, pp. 1-21, 2022.

[22] Y. Wang, H. Wang, and Z. Peng, “Rice diseases detection and classification using attention based neural network and bayesian optimization.” *Expert Systems with Applications*, vol. 178, p. 114770.

[23] M.A. Azim, M.K. Islam, M.M. Rahman, and F. Jahan, “An effective feature extraction method for rice leaf disease classification.” *Telkomnika*, vol. 19, no. 2, pp. 463-470, 2021.

[24] V. Tiwari, R.C. Joshi, and M.K. Dutta, “Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images.” *Ecological Informatics*, vol. 63, pp. 101289.

[25] M.H. Bijoy, N. Hasan, M. Biswas, S. Mazumdar, A. Jimenez, F. Ahmed, M. Rasheduzzaman, and S. Momen, “Towards Sustainable Agriculture: A Novel Approach for Rice Leaf Disease Detection Using dCNN and Enhanced Dataset.” *IEEE Access*, 2024.

[26] Q. Zhou, S. Zhou, F. Shen, J. Yin, and D. Xu, “Hyperspectral image classification based on 3-D multihead self-attention spectral–spatial feature fusion network.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 1072-1084, 2022.

[27] A. Hernández, and J.M. Amigó, “Attention mechanisms and their applications to complex systems.” *Entropy*, vol. 23, no. 3, pp. 283, 2021.

[28] Z. Zhang, and C. Jung, “GBDT-MO: gradient-boosted decision trees for multiple outputs.” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 3156-3167, 2020.

[29] P. Zhang, C. Wang, N. Kumar, W. Zhang, and L. Liu, “Dynamic virtual network embedding algorithm based on graph convolution neural network and reinforcement learning.” *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9389-9398, 2021.

[30] H.Q. Birdawod, A.M. Khudhur, D.H. Kadir, and D.M. Saleh, “A wavelet shrinkage mixed with a single-level 2D discrete wavelet transform for image denoising.” *Kurdistan Journal of Applied Research*, vol. 9, no. 2, pp. 1-12, 2024.

[31] U.E. Kaplan, Y. Dagasan, and E. Topal, “Mineral grade estimation using gradient boosting regression trees”. *International Journal of Mining, Reclamation and Environment*, vol. 35, no. 10, pp. 728-742, 2021.

[32] L. Zhou, D. Xu, Y. Yuan, and L. Wang, "Research on transformer fault intelligent diagnosis technology based on improved random forest algorithm." In *Journal of Physics: Conference Series*, vol. 2728, no. 1, pp. 012056, 2024, March. IOP Publishing.

[33] E.A. Ruelas-Santoyo, A.J. Ríos-Lira, Y.V. Pantoja-Pacheco, J.A. Jiménez-García, S. Hernández-González, and O. Cruz-Domínguez, "Recognition of Intergranular Corrosion in AISI 304 Stainless Steel by Integrating a Multilayer Perceptron Artificial Neural Network and Metallographic Image Processing." *Applied Sciences*, vol. 14, no. 12, pp. 5077, 2024.

[34] A.S. Almasoud, A. Abdelmaboud, T.A. Elfadil Eisa, M. Al Duhayyim, A.A. Hassan Elnour, M. Ahmed Hamza, A. Motwakel, and A. Sarwar Zamani, "Artificial Intelligence-Based Fusion Model for Paddy Leaf Disease Detection and Classification." *Computers, Materials & Continua*, vol. 72, no. 1, 2022.

[35] K. Mahadevan, A. Punitha, and J. Suresh, "Automatic recognition of Rice Plant leaf diseases detection using deep neural network with improved threshold neural network." *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, vol. 8, pp. 100534, 2024.

[36] H.H. Alshammari, A.I. Taloba, and O.R. Shahin, "Identification of olive leaf disease through optimized deep learning approach." *Alexandria Engineering Journal*, vol. 72, pp. 213-224, 2023.