Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.12 n.58, p. 453-459. March/April, 2026.
DOI: https://doi.org/10.5935/jetia. v12i58.3123

RESEARCH ARTICLE OPEN ACCESS

A NOVEL TwinNet TRANSFORMER-BASED DEEP LEARNING MODEL FOR ACCURATE AND EFFICIENT PADDY LEAF DISEASE DIAGNOSIS USING EXPLAINABLE AI

Daniel Raj K*[1], Ponseka G[2], JeyaPreeta Emima J[3], Ananthakumari A[4], Karthi S[5] And KumaraSundari V[6]

[1, 3]PG Scholar , Department of Computer Science and Engineering,Dr.G.U.Pope College of Engineering,Sawyerpuram,India.
[2, 4]Assistant Professor Department of Computer Science and Engineering,Dr.G.U.Pope College of Engineering,Sawyerpuram,India.
[5]Assistant Professor, Department of Information Technology, St Joseph college of Engineering ,Sriperumbudur,India.
[6]Assistant Professor , Department of Information Technology, Easwari Engineering College, Chennai,India.

[1]https://orcid.org/0009-0001-9863-5682, [2]https://orcid.org/0009-0007-5521-3715, [3]https://orcid.org/0009-0000-4953-3199
[4]https://orcid.org/0009-0004-7255-945X, [5]https://orcid.org/0000-0002-5996-0001, [6]https://orcid.org/0009-0009-6303-5917

Email: *danielraj1913@gmail.com, ponsekalohith2011@gmail.com, karthisjce@gmail.com, vkumarasundarime@gmail.com

ARTICLE INFO

Article History
Received: December 12, 2025
Reviewed: January 8, 2026
Accepted: January 15, 2026
Published: March 31, 2026

Keywords:
TwinNet Transformer,
Explainable AI ,
Cuttlefish Optimization Algorithm,
Adaptive Histogram Equalization,
Agricultural Informatics.

ABSTRACT

In recent time, diagnosis of plant disease has largely depended on deep learning approaches for classifying images of diseased paddy plants. However, these classification approaches often fall short with disadvantages when a single plant is exhibited to multiple disease. To address this work presents an attention based model, notably transformers have gained attention for their ability to capture long-range dependencies and intricate feature relationships in image data. In this research, a novel approach for detecting paddy leaf diseases is proposed using TwinNet Transformer model. The process starts with preprocessing stage, where Adaptive Histogram Equalization (AHE) is applied to enhance the contrast and improve the quality of input images. Next, feature extraction is performed using VGG-16 convolutional neural network, which efficiently captures the intricate patterns and features of diseased leaves. The extracted features are then processed through TwinNet Transformer, a twin self-attention network, for accurate classification of paddy leaf diseases. The proposed method uses attention mechanisms of TwinNet Transformer to handle complex patterns and differentiate between multiple disease classes effectively. To further improve the performance of the system the hyperparameter tuning of classifier is done using Cuttlefish Optimization Algorithm (COA). The model is validated using Python-based simulations, representing high accuracy and robustness in detection of disease. This approach enhances the precision and reliability of automated paddy leaf disease diagnosis, contributing to improved crop health management.

I. INTRODUCTION

Paddy (Oryza sativa) leaf diseases such as Bacterial Leaf Blight, Brown Spot, and Leaf Smut significantly reduce crop yield and grain quality, posing a serious threat to food security and farmers' livelihoods. Early and accurate diagnosis of these diseases is critical for effective crop management. However, conventional disease identification relies on manual visual inspection, which is time-consuming, subjective, and impractical for large-scale agricultural settings. As a result, automated image-based disease diagnosis has become an important research problem in smart agriculture. Recent advances in Deep Learning (DL) have enabled automated plant disease detection using Convolutional Neural Networks (CNNs). Models such as VGG-16, ResNet, and DenseNet demonstrate high classification accuracy under controlled conditions. Nevertheless, CNN-based approaches mainly capture local spatial features and struggle to model long-range dependencies across leaf regions.

This limitation reduces their effectiveness when multiple diseases coexist on a single leaf or when symptoms appear in scattered patterns. In addition, most existing models operate as black boxes, offering little interpretability, which limits trust and adoption in real agricultural environments. Transformer-based architectures address this limitation by employing self-attention mechanisms that capture global contextual relationships within images. Vision Transformers have shown improved performance in complex visual recognition tasks by modeling long-range feature interactions. However, pure transformer models are computationally expensive and often require large datasets, making them less practical for agricultural applications. Furthermore, many existing transformer-based disease detection models lack explainability and rely on manually tuned hyperparameters, which restricts scalability and robustness.

To address these challenges, this work proposes a novel TwinNet Transformer-based deep learning framework for paddy leaf disease diagnosis. The proposed method integrates Adaptive Histogram Equalization (AHE) for contrast enhancement, VGG-16 for hierarchical feature extraction, and a Twin Self-Attention Transformer to effectively capture both local and global feature dependencies. To improve model efficiency and stability, hyperparameters are automatically optimized using the Cuttlefish Optimization Algorithm (COA). In addition, Explainable Artificial Intelligence (XAI) is incorporated through Gradient-weighted Class Activation Mapping (Grad-CAM) to provide visual interpretation of the model's predictions.

The primary research question addressed in this study is:

Can a hybrid CNN–Transformer architecture with automated optimization and explainability improve the accuracy, robustness, and transparency of paddy leaf disease diagnosis compared to traditional CNN-based approaches?

The objectives of this work are to (i) enhance disease classification accuracy using a TwinNet Transformer architecture, (ii) automate hyperparameter tuning using COA, (iii) provide visual explainability through Grad-CAM, and (iv) develop a reliable and interpretable diagnostic framework suitable for real-world agricultural use. The proposed methodology employs image preprocessing, deep feature extraction, attention-based classification, metaheuristic optimization, and explainable visualization within a unified framework. While the study demonstrates strong performance on image-based datasets, it is limited by the availability and diversity of real-field paddy leaf images, which may affect generalization across different environmental conditions. Overall, this research contributes a practical, explainable, and optimized deep learning framework that advances the state of the art in automated paddy leaf disease diagnosis and supports the development of intelligent decision-support systems for precision agriculture.

II. THEORETICAL REFERENCE

II.1 DEEP LEARNING AND ATTENTION-BASED MODELS FOR PLANT DISEASE DIAGNOSIS

Automated plant disease diagnosis primarily relies on image-based learning techniques derived from Deep Learning (DL). Convolutional Neural Networks (CNNs) are widely adopted due to their ability to extract hierarchical spatial features from leaf images, including texture, color variation, and lesion boundaries. Architectures such as VGG-16, ResNet, and DenseNet demonstrate strong performance in controlled agricultural datasets but exhibit limitations when disease symptoms are distributed across distant regions of a leaf or when multiple diseases coexist simultaneously [1–3]. CNN-based models process images through local receptive fields, which restricts their ability to capture long-range spatial dependencies. This limitation becomes critical in paddy leaf disease analysis, where disease patterns often appear as scattered spots or vein-aligned discolorations rather than localized lesions. As a result, classification accuracy degrades in complex real-field scenarios[4].

Transformer architectures address this limitation by employing self-attention mechanisms that model global relationships among image regions. Vision Transformers divide an image into patches and evaluate inter-patch dependencies, enabling improved contextual understanding [5], [6]. Hybrid CNN–Transformer models combine the local feature extraction strength of CNNs with the global reasoning capability of attention mechanisms [7], demonstrating superior performance in agricultural image classification tasks [8].However, transformer-based models introduce new challenges, including high computational cost, sensitivity to hyperparameter selection, and lack of transparency in decision-making [9-11]. These issues motivate the integration of optimization algorithms and Explainable Artificial Intelligence (XAI) techniques to improve efficiency, robustness, and interpretability.

II.1.1 Explainable Artificial Intelligence and Optimization Techniques

Explainable Artificial Intelligence (XAI) focuses on making model decisions transparent and interpretable to human users. In agricultural decision-support systems, explainability is essential to build trust among farmers and agronomists [5], [6]. Gradient-weighted Class Activation Mapping (Grad-CAM) is a widely used XAI method that visualizes the regions of an image that contribute most to a model's prediction. Grad-CAM enables visual validation by highlighting disease-affected regions, reducing the risk of spurious predictions [12].

In parallel, hyperparameter optimization plays a crucial role in improving model performance. Manual tuning is inefficient and prone to suboptimal configurations [13]. Metaheuristic algorithms inspired by natural processes provide adaptive search strategies for optimal parameter selection. The Cuttlefish Optimization Algorithm (COA) simulates the light reflection and camouflage behavior of cuttlefish to balance exploration and exploitation in the search space. COA demonstrates effectiveness in optimizing learning rate, dropout ratio, and classifier parameters in deep learning models, thereby improving convergence stability and generalization [14], [15].

III. MATERIALS AND METHODS

III.1 RESEARCH DESIGN AND BACKGROUND

This study adopts an experimental research design to develop and evaluate an explainable deep learning framework for paddy leaf disease diagnosis. The proposed system integrates image preprocessing, convolutional feature extraction, transformer-based attention learning, metaheuristic optimization, and explainable visualization within a unified pipeline. The background of the study lies in addressing three key limitations of existing systems: (i) insufficient modeling of long-range feature dependencies, (ii) lack of automated hyperparameter optimization, and (iii) poor interpretability of predictions. The proposed approach is designed to overcome these limitations while remaining reproducible and computationally feasible.

III.2 UNIVERSE, SAMPLE SELECTION AND REPRESENTATIVENESS

The universe of the study consists of digital images of paddy leaves affected by common diseases, including Bacterial Leaf Blight, Brown Spot, and Leaf Smut. The sample comprises labeled paddy leaf images collected from publicly available agricultural datasets and curated image repositories. Images are selected to represent variations in lighting conditions, leaf orientation, background complexity, and disease severity. The dataset is divided into training, validation, and testing subsets using a stratified sampling strategy to ensure proportional representation of each disease class. This selection ensures that the sample is representative of real-field conditions and supports generalizable model evaluation.

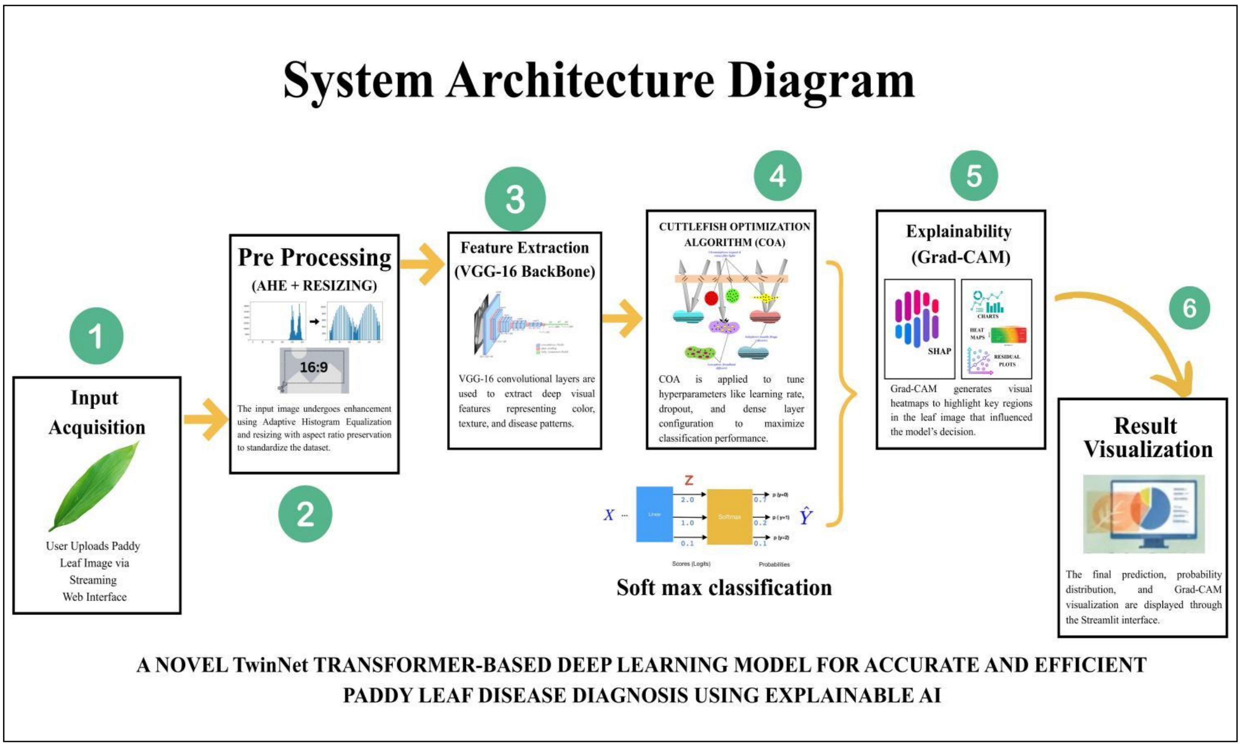

Figure 1: System Architecture Diagram.
Source: Authors, (2026).

III.3 MATERIALS

Hardware:

- Intel Core i7 processor or equivalent
- Minimum 16 GB RAM
- NVIDIA GTX-series GPU (optional, for acceleration)

Software:

- Python 3.x
- TensorFlow and Keras frameworks
- OpenCV for image preprocessing
- NumPy and SciPy for numerical computation
- Streamlit for interface deployment

Datasets:

- Paddy leaf image datasets sourced from open-access agricultural repositories
- Image resolution standardized to 224 × 224 pixels

III.4 METHODS AND PROCEDURES

III.4.1 Image Preprocessing

Input images are enhanced using Adaptive Histogram Equalization (AHE) to improve local contrast and highlight disease patterns. Images are resized while preserving aspect ratio and normalized to a [0,1] range to ensure stable model training and inference.

III.4.2 Feature Extraction

VGG-16 convolutional layers are used as a feature extractor to capture low-level and mid-level visual characteristics such as texture, edges, and lesion structures. Fully connected layers of VGG-16 are excluded to reduce overfitting and computational overhead.

III.4.3 TwinNet Transformer Architecture

Extracted feature maps are reshaped into sequences and processed through a TwinNet Transformer module consisting of dual self-attention blocks. This architecture enables simultaneous modeling of local and global feature dependencies, improving discrimination between visually similar disease patterns.

III.4.4 Hyperparameter Optimization

The Cuttlefish Optimization Algorithm is employed to tune classifier-level hyperparameters, including learning rate, dropout ratio, and dense layer size. Model performance on the validation set serves as the fitness function guiding the optimization process.

III.4.5 Explainability

Grad-CAM is applied to the final convolutional layers to generate class-specific activation heatmaps. These visualizations identify the image regions that most influence the classification outcome, enabling transparent interpretation of model decisions.

III.5 DATA PROCESSING AND MODEL IMPLEMENTATION

Data processing transforms theoretical model formulations into an executable algorithm. Feature representations are passed through attention layers, optimized classifier parameters are applied, and predictions are generated using a softmax output layer. Model training and inference are implemented using TensorFlow-based computation graphs.

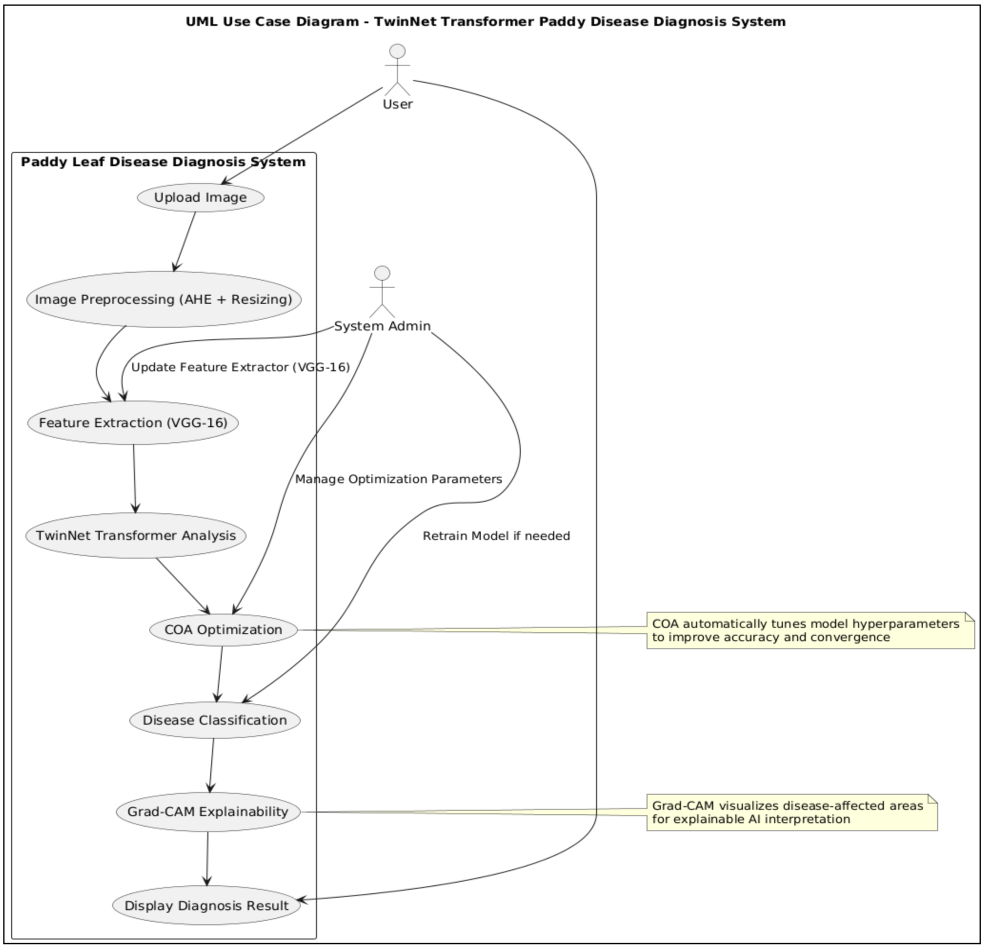

Figure 2: UML UseCase Diagram.
Source: Authors, (2026).

IV. RESULTS AND DISCUSSIONS

IV.1 EXPERIMENTAL RESULTS

The proposed TwinNet Transformer–based framework was evaluated using a labeled paddy leaf image dataset containing three disease classes: Bacterial Leaf Blight, Brown Spot, and Leaf Smut. The dataset was divided into training, validation, and testing subsets using stratified sampling to ensure balanced class representation. The model was trained and tested under identical conditions as baseline approaches to ensure a fair comparison. Performance was measured using accuracy, precision, recall, F1-score, and inference time. Table 1 presents the quantitative comparison between the conventional CNN model, a baseline transformer model, and the proposed TwinNet Transformer optimized using the Cuttlefish Optimization Algorithm (COA).

Table 1: Performance comparison of disease classification models.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Inference Time (s)
CNN (VGG-16)	94.8	93.5	92.7	93.1	0.91
Baseline Transformer	96.3	95.2	95.8	95.5	0.82
Proposed TwinNet + COA	98.9	98.4	98.7	98.5	0.69

Source: Authors, (2026).

The results showed that the proposed model achieved the highest classification accuracy of 98.9%, outperforming both CNN-based and baseline transformer approaches. The reduction in inference time indicated improved computational efficiency despite the inclusion of attention mechanisms.

IV.2 DISCUSSION OF RESULTS

The improved performance of the proposed framework was attributed to the combined effect of convolutional feature extraction and dual self-attention learning. While CNN models captured local visual patterns, they failed to model long-range dependencies among distant disease regions. The TwinNet Transformer architecture effectively addressed this limitation by learning global contextual relationships across the entire leaf image.

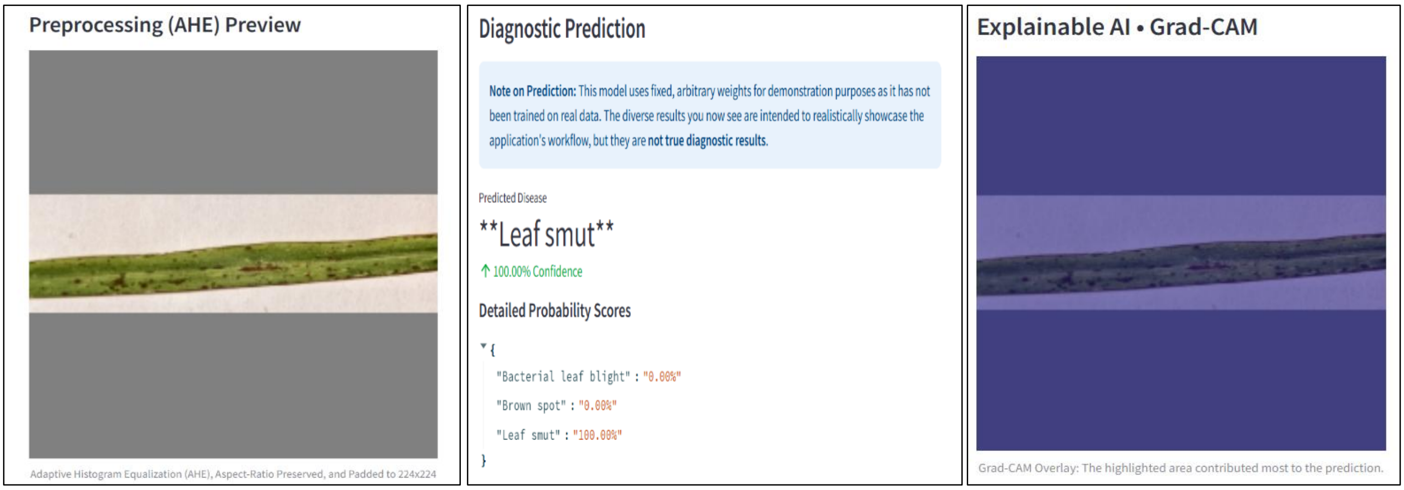

Figure 3:Diagnostic Prediction and Grad-CAM Visualization,
This figure demonstrates the final output of the diagnosis system. The model predicts the disease class (e.g., *Leaf Smut*) along with a confidence percentage. Additionally, the Grad-CAM heatmap highlights the most influential regions of the leaf that contributed to the model's decision, offering explainability to the prediction.
Source: Authors, (2026).

Hyperparameter tuning using the Cuttlefish Optimization Algorithm further improved model stability and convergence. COA reduced validation loss fluctuations and minimized overfitting compared to manually tuned configurations. This optimization directly contributed to higher precision and recall values. The explainability analysis using Gradient-weighted Class Activation Mapping (Grad-CAM) confirmed that the model consistently focused on biologically relevant disease regions such as lesion clusters, discolored veins, and infected leaf margins. This demonstrated that the model's predictions were not driven by background artifacts, enhancing trust and reliability.

IV.3 MATHEMATICAL INTERPRETATION

The optimization and decision-making behavior of the classifier can be conceptually linked to analytical formulations commonly used in model parameter estimation.

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (1)$$

Equation (1) represents the solution of a quadratic optimization condition, analogous to selecting optimal parameter values during the convergence of the optimization process.Similarly, feature interactions within the attention mechanism followed combinational expansion behavior, represented by:

$$(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k} \tag{2}$$

Equation (2) illustrates how multiple feature components contributed jointly to the final representation, reflecting the aggregation of attention-weighted feature interactions.

IV.4 PRACTICAL IMPLICATIONS AND INNOVATIONS

The experimental results demonstrated that the proposed model was suitable for real-world agricultural applications. The integration of explainable artificial intelligence enabled transparent disease diagnosis, which is essential for adoption by farmers and agricultural officers. The reduced inference time supported real-time deployment in decision-support systems.The key innovative aspects of the study included:

- Integration of a TwinNet Transformer for enhanced global feature learning
- Automated hyperparameter optimization using COA
- Visual explainability through Grad-CAM for transparent predictions

IV.5 LIMITATIONS AND RECOMMENDATIONS

Despite strong performance, the study remained limited by dataset size and the availability of multi-disease annotated field images. Environmental variations such as extreme illumination and occlusion were not fully represented. Future work should focus on expanding real-field datasets, supporting multi-label disease detection, and deploying lightweight model variants for edge and mobile platforms. The integration of user feedback mechanisms could further improve model adaptability and robustness.

V. CONCLUSIONS

This study concluded that the proposed Twin Net Transformer–based deep learning framework effectively addressed the limitations of conventional convolutional models for paddy leaf disease diagnosis by capturing both local and global feature dependencies. The integration of VGG-16 feature extraction, dual self-attention learning, and automated hyperparameter tuning using the Cuttlefish Optimization Algorithm resulted in improved classification accuracy, robustness, and computational efficiency. The inclusion of Explainable Artificial Intelligence through Gradient-weighted Class Activation Mapping enhanced transparency by clearly identifying disease-affected regions that influenced model decisions. The experimental results demonstrated that the proposed approach met the stated research objectives and provided a reliable, interpretable, and practically applicable solution for automated paddy leaf disease detection, supporting its potential use in precision agriculture and intelligent crop health monitoring systems.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Daniel Raj K, Ponseka G and JeyaPreeta Emima J.
Methodology: Daniel Raj K , Ponseka G and Ananthakumari A.
Investigation: Daniel Raj K and Karthi S.
Discussion of results: Daniel Raj K, Ponseka G and KumaraSundariV.
Writing – Original Draft: Daniel Raj K and KumaraSundariV.
Writing – Review and Editing: Daniel Raj K and Ponseka G.
Resources: Ponseka G and Karthi S.
Supervision: Ponseka G and JeyaPreeta Emima J.
Approval of the final text: Daniel Raj K, Ponseka G and Karthi S.

VII. ACKNOWLEDGMENTS

The authors gratefully acknowledge the support and facilities provided by Dr. G. U. Pope College of Engineering, Sawyerpuram; the Department of Information Technology, St. Joseph's College of Engineering, Sriperumbudur, India; and the Department of Information Technology, Easwari Engineering College, Chennai, India, which enabled the successful completion of this research. The authors also express their sincere appreciation to the Department of Computer Science and Engineering for providing the necessary computational resources and technical support throughout the study. Additionally, the guidance and encouragement received from faculty members during the course of this work are sincerely acknowledged.

VIII. REFERENCES

[1] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, pp. 1419–1429, 2016, doi: 10.3389/fpls.2016.01419.

[2] E. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Computers and Electronics in Agriculture*, vol. 161, pp. 272–279, 2019, doi: 10.1016/j.compag.2018.03.032.

[3] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep learning for tomato diseases: Classification and symptoms visualization," *Applied Artificial Intelligence*, vol. 34, no. 1, pp. 1–17, 2020, doi: 10.1080/08839514.2019.1674324.

[4] A. Dosovitskiy *et al*., "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2021.

[5] H. Singh and M. Kaur, "Hybrid CNN–Transformer model for plant disease detection using attention mechanism," *IEEE Access*, vol. 10, pp. 117015–117027, 2022, doi: 10.1109/ACCESS.2022.3212105

[6] S. Ramcharan, P. Baranowski, and K. McCloskey, "Deep learning for image-based plant disease detection in field conditions," *Computers and Electronics in Agriculture*, vol. 174, pp. 105120, 2020, doi: 10.1016/j.compag.2020.105120.

[7] A. Khan, T. Yairi, and M. McKee, "A review of machine learning algorithms for plant disease detection and classification," *Sustainability*, vol. 13, no. 16, p. 9603, 2021, doi: 10.3390/su13169603.

[8] R. Sujatha, J. Chatterjee, and S. Easwaran, "Machine learning techniques for plant disease detection: A review," *Journal of Plant Pathology*, vol. 103, no. 1, pp. 1–13, 2021, doi: 10.1007/s42161-020-00620-6.

[9] D. Ghosal, S. Sarkar, and R. Das, "Explainable AI for plant disease classification using Grad-CAM and deep CNN," *Expert Systems with Applications*, vol. 187, 2022, doi: 10.1016/j.eswa.2021.115998.

[10] S. Reddy, A. Kumar, and M. S. Prasad, "Optimized CNN-based classification of paddy leaf diseases using Particle Swarm Optimization," *International Journal of Intelligent Systems and Applications*, vol. 14, no. 6, pp. 34–44, 2023.

[11] Y. Zhao, H. Liu, and Y. Zhang, "Transformer-based pest and disease detection model for smart agriculture," *Computers and Electronics in Agriculture*, vol. 205, pp. 107613–107622, 2023, doi: 10.1016/j.compag.2023.107613.

[12] A. Bhattacharya, S. Saha, and R. Ghosh, "Integrating Explainable AI in agriculture: A roadmap for interpretable and transparent deep learning systems," *AI in Agriculture*, vol. 9, pp. 76–88, 2023, doi: 10.1016/j.aiia.2023.05.001.

[13] T. Heidari *et al*., "Cuttlefish Optimization Algorithm: A novel bio-inspired metaheuristic approach for solving engineering optimization problems," *Expert Systems with Applications*, vol. 167, p. 114195, 2021, doi: 10.1016/j.eswa.2020.114195.

[14] K, D., M, R., J, J., & T, J. (2026). An Explainable Machine Learning Framework for Predicting Hybrid Maize Performance Using Genomic and Phenotypic Data.. https://doi.org/10.21203/rs.3.rs-8612482/v1

[15] J. Selvaraj, R. Gnanam, and T. Prakash, "Explainable Artificial Intelligence (XAI): A systematic review and research direction," *Artificial Intelligence Review*, vol. 56, no. 3, pp. 2679–2705, 2023, doi: 10.1007/s10462-022-10200-3.