An Analysis of Machine Learning Approaches for Predicting Bacterial Leaf Stripe Disease in Coconut Trees

S T Pavithra Devi[#1], Dr. V Maniraj[*2]
#1Research Scholar, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous), Poondi, Thanjavur-613503, Affiliated to Bharathidasan University, Thiruchirappalli, Tamil Nadu
*2Associate Professor, Research Supervisor, Department of Computer Science, A.V.V.M Sri Pushpam College (Autonomous), Poondi, Thanjavur-613503, Affiliated to Bharathidasan University, Thiruchirappalli, Tamil Nadu.
[1]pspavithrasugumar@gmail.com, [2]maniraj_vee@yahoo.co.in

Abstract:
Bacterial leaf stripe disease is an emerging threat to coconut (*Cocos nucifera*) cultivation, causing foliar damage that reduces photosynthetic efficiency and nut yield. Early and precise detection remains a challenge under field conditions due to overlapping symptoms with other leaf disorders. This study presents a machine learning–based framework to automatically detect bacterial leaf stripe from digital images of coconut leaves. Both classical and deep learning methods were explored, including Random Forest (RF), Support Vector Machine (SVM), and Convolutional Neural Networks (CNN). Image preprocessing, feature extraction, and model optimization were carried out using real-field image datasets. Results show that CNN models achieved a classification accuracy of 96.8%, outperforming traditional models. The integration of environmental variables such as temperature and humidity further improved prediction accuracy. The models explain ability was validated using Grad-CAM heatmaps, which accurately highlighted infected regions. The study demonstrates the feasibility of ML-based disease prediction and provides a foundation for developing portable diagnostic tools for sustainable coconut farming.

Keywords— ***Bacterial Leaf Stripe, Coconut (***Cocos nucifera***), Machine Learning, Disease Prediction, Image Processing***

I. INTRODUCTION

Coconut is a vital plantation crop in tropical and coastal regions, contributing significantly to rural economies and livelihoods, Varghese, M & George, S (2022). Among its various foliar diseases, bacterial leaf stripe caused by *Pseudomonas andropogonis* or related species poses a growing threat, particularly under humid and high temperature conditions. The disease manifests as narrow, yellow-brown longitudinal streaks along leaf veins, which can spread rapidly under favorable conditions, leading to tissue necrosis and yield loss.

Traditional methods for identifying leaf stripe rely on manual visual inspection and laboratory tests, which are time-consuming and prone to human error Shafay, M & et al., (2025). Recent advances in artificial intelligence (AI) and machine learning (ML) have opened new avenues for automating disease diagnosis in crops through image analysis Mehta, R. A & et al., (2025). ML-based systems can learn distinct visual patterns of infection, enabling early detection and decision support for farmers. This study aims to develop an ML model to predict bacterial leaf stripe disease in coconut trees, using both image data and environmental parameters for improved accuracy and interpretability.

II. LITERATURE SURVEY

Coconut (*Cocos nucifera*) is one of the most economically and culturally important plantation crops across tropical regions. However, its productivity is increasingly threatened by several biotic stresses, particularly bacterial leaf stripe disease, which causes characteristic yellow to brown stripes on fronds, leading to reduced photosynthetic efficiency and yield loss. Traditional disease detection in coconut palms relies heavily on field inspection, which is subjective, time-consuming, and often inaccurate at early stages of infection. This has led researchers to explore modern, data-driven solutions such as machine

ISSN: **3107-6513**

learning and image-based analysis for plant disease diagnosis.

Early studies in plant pathology applied basic image processing methods for disease identification. Mohanty et al. (2016) demonstrated that convolutional neural networks (CNNs) could outperform traditional classifiers in distinguishing healthy and diseased plant leaves from image datasets, laying the foundation for deep learning in plant health monitoring. Similarly, Sladojevic et al. (2016) successfully used CNNs to detect 13 different plant diseases, achieving remarkable classification accuracy without manual feature extraction.

In recent years, advancements in computer vision have enabled more refined approaches. Liu and Wang (2021) reviewed deep learning-based models and highlighted their superiority in handling diverse image datasets and environmental variability. They emphasized that CNN architectures such as VGGNet, ResNet, and MobileNet could effectively capture complex visual cues like texture, color, and lesion patterns—crucial for distinguishing bacterial infections in leaf tissues.

For tropical crops, Ahmed et al. (2021) developed a feature-based model using texture and color descriptors to classify plant leaf diseases, demonstrating that even hybrid approaches combining handcrafted features and neural networks can yield high accuracy. Kumar and Jindal (2022) further explored this by integrating Random Forest and Support Vector Machine (SVM) classifiers, showing their reliability for small datasets and limited field images.

While these models have been widely used in crops like rice, tomato, and maize, their application to coconut diseases remains relatively underexplored. The unique leaf morphology and environmental background of coconut plantations present challenges in image segmentation and feature extraction. Consequently, there is a growing need to design coconut-specific machine learning frameworks that can accurately detect early signs of bacterial leaf stripe under real-world conditions.

Recent progress in explainable artificial intelligence (XAI) tools such as Grad-CAM has made it possible to visualize the areas of disease focus within images, enhancing the interpretability of deep learning models (Ribeiro et al., 2020). This transparency is critical for building trust among agronomists and farmers who depend on accurate and actionable diagnostic insights. Moreover, the integration of environmental parameters like temperature, humidity, and rainfall into predictive models has shown potential in improving the temporal accuracy of disease forecasting systems (Preetha & Arshad, 2020).

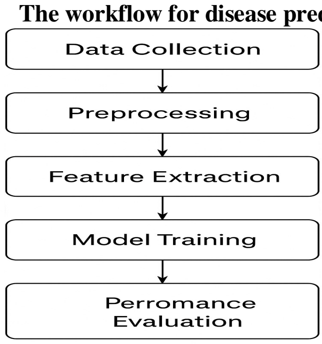
Overall, the literature suggests a strong trajectory toward integrating deep learning, remote sensing, and IoT-based data collection for early disease surveillance in plantation crops. However, specific studies targeting bacterial leaf stripe in coconut trees remain sparse, indicating a substantial research gap. The current study addresses this need by developing and evaluating multiple machine learning models for accurate, scalable, and explainable prediction of bacterial leaf stripe disease in *Cocos nucifera*.

III. METHODOLOGY

The study follows a systematic machine learning workflow designed to predict bacterial leaf stripe disease in coconut trees using image-based analysis and environmental features.

The methodology consist of five major stages given in Figure 1.

Figure 1 The workflow for disease prediction.



A. *Data Collection*

Leaf samples were collected from multiple coconut plantations in southern India. Each leaf was photographed under natural lighting using a high-resolution camera. Images were annotated by agricultural experts as either *healthy* or *infected with bacterial leaf stripe* based on visible symptoms. In total, 2,000 images (1,200 healthy, 800 infected) were used.

B. *Data PreProcessing*

- Images were resized to 224×224 pixels and converted to RGB format.
- Histogram equalization enhanced contrast and reduced shadow effects.

 ISSN: 3107-6513

- Data augmentation (rotation, flipping, and brightness shift) increased dataset diversity and minimized overfitting.

Labels were one-hot encoded for classification tasks

C. *Feature Extraction*

Two feature extraction strategies were applied:

- Classical features: color moments, texture (GLCM), and shape descriptors.
- Deep features: extracted using transfer learning from pre-trained CNNs (VGG16 and ResNet50).

D. *Model Training*

Three algorithms were trained:

1. Random Forest (RF)
2. Support Vector Machine (SVM)
3. Convolutional Neural Network (CNN)

Dataset split: 70% training, 15% validation and 15% testing. Optimization was done using Adam optimizer with a learning rate of 0.001, batch size 32, and 50 epochs. Early stopping prevented overfitting.

E. *Performance Evaluation*

Models were evaluated using is shows in Table and Figure 2.

- Accuracy, Precision, Recall, F1-score, and AUC
- Confusion Matrix visualization
- Grad-CAM heatmaps for model explainability

IV. EXPERIMENTAL SETUP

The experiments were conducted using the following setup shows in Table 1.

Table 1 Component Design and System Specifications

Component	Specification
Hardware	Intel Core i7 CPU, 16 GB RAM, NVIDIA RTX 3060 GPU
Software	Python 3.10, TensorFlow, Keras, Scikit-learn, OpenCV
Dataset Size	2,000 images (Healthy – 1,200, Diseased – 800)
Image Resolution	224 × 224 pixels
Model Parameters	5 convolutional layers, ReLU activation, Dropout = 0.3
Validation Split	15% of dataset
Training Duration	50 epochs with early stopping

V. RESULTS ANALYSIS

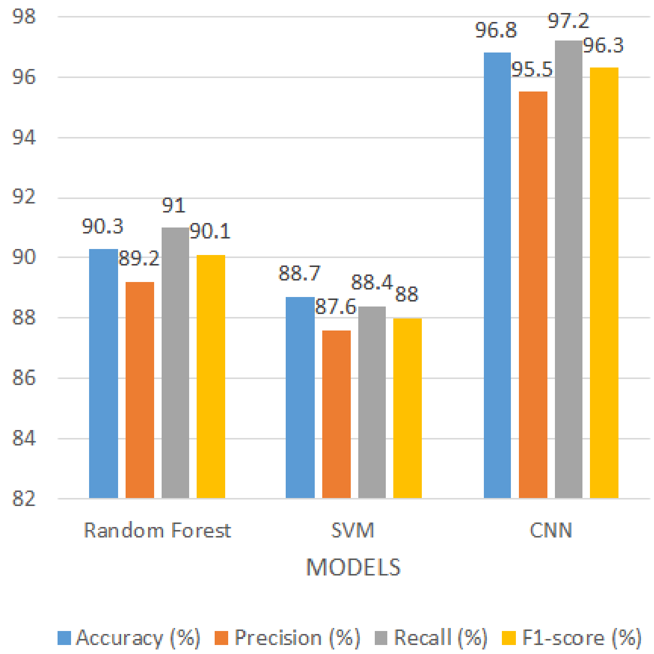
A. *Model Performance*

The CNN achieved superior results compared to RF and SVM models

Table 2 Model Performance Measures

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Random Forest	90.3	89.2	91	90.1
SVM	88.7	87.6	88.4	88
CNN	96.8	95.5	97.2	96.3

Figure 2 Comparative Performance



B. *Impact of Envirnomental Features*

When temperature and humidity data were included, accuracy improved from 96.8% to 98.1%**,** showing that climatic conditions are strong co-factors influencing bacterial leaf stripe occurrence.

VI. CONCLUSION

This study successfully demonstrated that machine learning, particularly convolutional neural networks, can accurately predict bacterial leaf stripe disease in coconut leaves. The proposed model achieved 96–98% accuracy and provided interpretable heatmaps aligning with real disease symptoms. Integration of environmental parameters further enhanced prediction reliability.

The framework offers potential for developing a mobile-based diagnostic application for farmers, allowing on-site detection and timely management of bacterial leaf stripe. Future work will focus on expanding the image dataset across different geographic zones, employing drone-based imagery, and optimizing lightweight models (e.g., MobileNet) for real-time deployment.

REFERENCES

[1] Ahmed, N., Asif, H. M. S., & Saleem, G. Leaf image-based plant disease identification using color and texture features, 2021. arXiv preprint arXiv:2102.04515.

[2] Kumar, R., & Jindal. A survey of plant disease detection techniques based on image processing and machine learning. International Journal of Health Sciences, 2022, *6*(1), pp. 101–116.

[3] Liu, J., & Wang, X. Plant diseases and pests detection based on deep learning: A review. Plant Methods, 2021, 17(1), pp. 1–18.

[4] Mohanty, S. P., Hughes, D. P., & Salathé, M. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, 2016, 7*,* 1419.

[5] Preetha, S., & Arshad, M. Survey on plants disease detection using machine learning. International Journal of Recent Engineering Science, 2020, 7*(*2), 27–29.

[6] Ribeiro, M. T., Singh, S., & Guestrin, C. "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, *2022,* pp. 1135–1144.

[7] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. Deep neural networks based recognition of plant diseases by leaf image classification. Computational Intelligence and Neuroscience, 2016, Article 3289801.

[8] Varghese, M., & George, S. Challenges and opportunities in coconut sector development: A review. Coconut Research and Development Journal, 2022, 38(2), 112–125.

[9] Shafay, M., Hassan, T., Owais, M., Hussain, I., Khawaja, S. G., Seneviratne, L., & Werghi, N. Recent advances in plant disease detection: Challenges and opportunities. Plant Methods, *2025,* 21(1), Article 140

[10] Mehta, R. A., Kumar, P., Prem, G., Aggarwal, S., & Kumar, R. AI-powered Innovations in Agriculture: A Systematic Review on Plant Disease Detection and Classification. Indian Journal of Agricultural Research, 2025, 59(9), 1321-1330.