



The Scientific Temper

JULY 2026 Vol. 17 (7): 6499-6511

ISSN: 0976-8653, E-ISSN: 2231-6396

DOI: 10.58414/SCIENTIFICTEMPER.2026.17.7.2469

<https://scientifictemper.com/>

RESEARCH ARTICLE

Predictive Analysis of Coconut Tree Diseases and Yield Optimization Using Random Forest in Precision Agriculture

S T PAVITHRA DEVI, DR. V. MANIRAJ

^{1,2}A.V.V.M Sri Pushpam College (Autonomous), Poondi, (Affiliated to Bharathidasan University, Tiruchirappalli), Thanjavur, TamilNadu, India.

Corresponding author: pspavithrasugumar@gmail.com

ABSTRACT

Coconut farming is a critical sector of the agricultural economy of most of the tropical states, where crop yields tend to be influenced by diseases, climatic and soil nutrient dearth. This is because early detection of diseases in the plants and proper prediction of the crop productivity are required to enhance productivity and sustainable agriculture. This paper proposes a predictive model of coconut tree disease monitoring and yield optimization on the basis of the Random Forest machine learning algorithm on a precision agriculture setup. The system proposed studies various agricultural parameters such as soil moisture, temperature, humidity, rainfall, nutrient concentration and visible disease symptoms so as to detect possible disease conditions and forecast yield performance. The training and evaluation of the Random Forest model is done on a structured agricultural data because it is able to deal with a high-dimensional data, minimize overfitting, and offer strong classification and regression performance. This model categorizes the health status of coconut trees and at the same time predicts yield results of the trees depending on environmental and soil parameters. Importance analysis of features is also done in order to find out the most significant variables in influence to achieve disease occurrence and crop productivity. The predictive framework guides the farmers and agricultural experts to make sound decision making on how the diseases are managed, the irrigation timeline, and the application of the nutrient. The proposed solution based on implementing machine learning methods into precision agriculture allows to detect disease at an early stage, track crops, and optimize the use of resources. As it has been shown in the experimental data, the algorithm of the Random Forest has proven to give great accuracy in its prediction and efficient operation with the agricultural decision-support system. The paper draws attention to the opportunities of machine learning-based solutions to optimize the management of coconut crops and sustainable development of agriculture.

Keywords: *Random Forest, Coconut Tree Disease Prediction, Yield Optimization, Precision Agriculture, Machine Learning.*

INTRODUCTION

Agriculture is an important aspect of maintaining the food security around the world and economy of most of the developing nations. Coconut farming is

a popular crop growing mainly in the tropical areas and a very lucrative crop among farmers. Nevertheless, the production of coconuts is normally influenced by several diseases, environmental stress, and poor practices in the

management of the crop. Root wilt, bud rot and leaf spot are among some of the diseases that greatly decrease the crop productivity and adversely affect the livelihood of farmers (Maray, M., et al., 2022). Effective management of the farm and sustainable development of agriculture depends on early detection of these diseases and proper forecasting of the crop yield.

The conventional ways of detecting diseases in coconut trees are primarily based on manual observation by the farmers or agricultural specialists. These tend to be both time consuming, subjective and usually open to human error. Also, traditional yield estimation methods rely on past records and simple statistical tools, which might not be able to reflect complicated interactions between the environmental factors, soil characteristics, and plant well-being. Consequently, farmers tend to get late or incorrect information on disease outbreak and yield potential resulting to poor utilization of resources like fertilizers, water and pesticides.

Recent developments in Machine Learning and precision agriculture have opened the door to change the situation and confront these challenges by making it possible to monitor crops and make predictions based on the data (Adivarekar, P. P., et al., 2023; Ai, Y., et al., 2020; Kathavate, P. N., & Mahant, M. A. 2025; Ramu, S. P., et al., 2025). Machine learning algorithms have the ability to process huge agricultural data helping to detect hidden trends that affect plant health and productivity. One effective algorithm among them is the Random Forest as it is capable of managing complicated data, minimize overfitting, and it also offers high accuracy of prediction. With the use of various input variables like soil moisture, temperature, humidity, rainfall and nutrient content, predictive models can help the farmers to determine the risk of diseases, and the extent of crop yields more precisely.

Even with these developments, however, there are still a number of constraints with the existing agricultural monitoring systems. Most of the current solutions are aimed at detecting disease or predicting its occurrence separately and not combining both components into a single predictive model. Also, conventional statistical frameworks are usually deficient in storing and predicting large and complex agricultural data (Rajesh, R. S., et al., 2024). These drawbacks suggest the necessity of a smart and holistic system that can attempt to simultaneously forecast the diseases of coconut trees and optimize the crop

production with the help of modern machine learning methods.

LITERATURE SURVEY

The application of machine learning methods to agriculture disease detection and crop yield forecasting has been examined by a number of papers. Decision Trees, Support Vector Machines and Neural Network machine learning have been used to process agricultural data and determine patterns of plant health and productivity.

Past studies show that machine learning models can make a major enhancement in disease detection accuracy over the traditional statistical methods. Image processing has been applied to detect plant based diseases by using leaf images and sensor based methods have been applied to predict crop growth and yield. Nevertheless, numerous known systems are aimed solely at classification of diseases or prediction of their occurrence as opposed to combining both in one predictive model.

Thite, S., et al., (2023) present a paper about the Coconut (*Cocos nucifera*) Tree Disease Dataset contains 5,798 images categorized into five disease classes: Bud Root Dropping, Bud Rot, Gray Leaf Spot, Leaf Rot, and Stem Bleeding. This dataset is designed for machine learning-based disease detection and classification, enabling researchers to develop and evaluate automated plant disease recognition models. The dataset provides diverse disease samples that support the training and testing of classification algorithms, contributing to advancements in plant pathology research and precision agriculture. By offering a structured resource for disease analysis, it helps improve early disease detection, effective crop management, and sustainable coconut plantation practices.

Pinho, R. C., et al., (2023) conduct a study analyzed the spatial distribution of the pest *E. cyparissias* in coconut plantations located in Santa Izabel, Pará, Brazil, during 2019. The research was conducted across 157 blocks of a commercial green dwarf coconut plantation, where the infestation rate was determined by the proportion of plants attacked by the pest in each block. Using the Local Moran's Index, the study examined spatial autocorrelation of infestation levels among neighboring plantation blocks. The results indicated a significant clustered distribution pattern of pest attacks, suggesting that infestations tend to occur in neighboring areas. While the study found no significant relationship between pest attacks and the age of coconut trees, a strong correlation was observed between infestation levels

and proximity to forest regions, indicating that nearby forests act as potential sources of pest infestation in coconut plantations.

DebMandal, M., & Mandal, S. (2011) portray that the coconut (*Cocos nucifera L.*) is a highly valuable tropical crop known for its wide range of nutritional, medicinal, and industrial applications. Various coconut products such as tender coconut water, copra, coconut oil, kernel, coconut cake, toddy, shell, wood products, leaves, and coir pith are extensively used in daily life, particularly in traditional coconut-growing regions. Different parts of the coconut fruit, especially the kernel and coconut water, possess several medicinal properties including antibacterial, antifungal, antiviral, antioxidant, and immune stimulant effects. Additionally, coconut contains essential micronutrients that contribute to human health and nutrition. Due to its diverse uses and health benefits, the coconut palm is often referred to as “Kalpavriksha” (the all-giving tree) in Indian literature, highlighting its importance in supporting both human health and economic sustainability.

Wijethunga, C. D., et al., (2023) presents an innovative framework to enhance the quality and sustainability of coconut farming in Sri Lanka by integrating image processing and machine learning techniques. The proposed system enables early detection and classification of pests and diseases in coconut palms, allowing timely intervention and reducing the reliance on chemical treatments, thereby promoting environmentally friendly farming practices. In addition to disease detection, the framework evaluates key factors affecting coconut growth, including soil quality, water availability, and climate conditions, to optimize crop productivity. The study also incorporates a growth prediction module using historical data and machine learning, which assists farmers in effective planning and resource management. Overall, this approach provides a comprehensive solution to improve coconut yield, product quality, and the economic and environmental sustainability of the coconut farming sector.

Singh, P., et al., (2012) develops an end-to-end system for detecting coconut tree diseases and pest infections using image processing and deep learning techniques. The framework identifies stem bleeding disease, leaf blight disease, and Red Palm Weevil infestation from coconut tree images. Different segmentation methods were evaluated, where k-means clustering showed better performance compared to thresholding and watershed techniques. Several pre-trained CNN

models were applied using transfer learning, while a custom-designed 2D CNN model achieved the highest validation accuracy of 96.94%. The trained models were also deployed in a web-based application using Flask to enable automated coconut disease detection.

Singh, P., et al., (2012) also presents an automated coconut tree disease identification framework using deep learning techniques. The proposed system allows farmers to upload coconut leaf images through a web-based interface, where image preprocessing techniques such as resizing, normalization, and noise filtering are applied. A Convolutional Neural Network (CNN) is then used to classify the images as either healthy or diseased. The model provides quick and accurate predictions along with suitable remedial suggestions, enabling farmers to take timely action to reduce crop damage. This approach minimizes the reliance on agricultural experts and encourages the adoption of artificial intelligence in smart farming practices (Paymode, S. A., & Malode, B. V. 2022; Pravitha, P. S., & Kumar, V. 2023).

PROBLEM STATEMENT

The restrictions of the traditional methods of monitoring and inability to utilize the data-based decision support systems often make coconut farmers vulnerable to the issues of early disease diagnosis and proper yield forecasting. This is caused by lack of a comprehensive predictive model to evaluate the environmental, soil and crop health parameters, which causes delayed disease management and low crop productivity. Consequently, it is necessary to elaborate the predictive framework based on machine learning so that it can be used to recognize diseases of coconut trees efficiently and assist in optimizing the yield to enhance agricultural efficiency and sustainability.

PROPOSED WORK

The proposed work will create a smart precision and agriculture system of predictive analysis of coconut tree diseases and optimal yield with the help of the Random Forest algorithm. The system is meant to study the key agriculture parameters like soil moisture, soil temperature, soil humidity, rainfall, nutrient content, PH and visible disease signs to facilitate the early detection of diseases and proper estimation of yield (Chaitra Bai, J., et al., 2026). The primary aim of the suggested model is to assist farmers and agricultural professionals to make data-driven and timely choices to enhance the condition, productiveness, and resource management of the coconut crop.

Traditional methods of coconut cultivation have the disadvantage of detection of the disease being done manually once it starts to show symptoms, and it may be very late, thus causing more loss of the crop. Equally, estimation of yield is mostly made upon past experience or mere observations, which do not necessarily represent the sophisticated correlation existing between environmental and biological conditions. To address these shortcomings, the proposed study is able to combine data preprocessing, feature analysis, disease classification, and yield prediction in one machine learning framework.

OBJECTIVE OF PROPOSED WORK

The following are the objectives on which the proposed system is built:

- The main objective is to identify the diseases of a coconut tree at an early stage with the help of a Random Forest classification model.

- Classify coconut disease pictures in less than five disease categories.
- Precondition the images and enhance their quality to analyze them.
- A meaningful visual feature, which could be: color, texture and shape.
- Forecast the input image to the desired disease classification with the help of Random Forest.
- Support early disease detection in coconuts farm set up.

ARCHITECTURE OF PROPOSED WORK

The following major stages can be suggested in the proposed work is given in Figure 1.

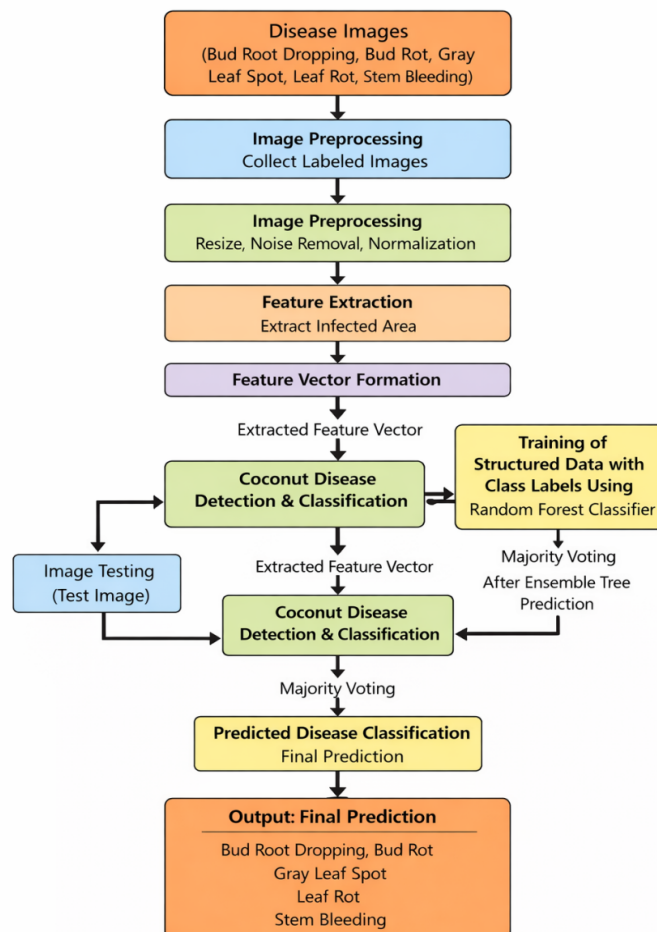


Figure 1: Workflow of proposed work

Data set

Data collection for the proposed IoT-based smart coconut farm system involves gathering real-time environmental, soil, and plant health data using sensors, cameras, and field observations. Soil moisture, temperature, humidity, pH, NPK levels, rainfall, and leaf images are collected using IoT devices such as Soil Moisture Sensor (e.g., Capacitive Soil Moisture Sensor v1.2), Temperature and Humidity Sensor (DHT11), Soil pH Sensor, NPK Sensor Module, Rain Sensor Module (YL-83/YL-69), and a Camera Module (ESP32-CAM or Raspberry Pi Camera). These sensors are connected to microcontrollers like ESP32, NodeMCU, or Arduino Uno, which collect and transmit the data to the cloud for further analysis. In addition, historical data such as irrigation schedules, fertilizer usage, and yield records are included. The collected data is cleaned, labeled, and organized to form a structured dataset that supports both real-time monitoring and machine learning-based analysis.

The dataset for this study was collected from coconut farms located in the Tanjore (Thanjavur) district of Tamil Nadu, India, during the period

from June 2025 to May 2026. Primary data was obtained through IoT sensor deployment directly in selected farm fields, while plant leaf images were captured using camera modules and farmer mobile devices. Secondary data, including past yield records and fertilizer usage, was collected through farmer interviews and local agricultural records. Regular field visits were conducted to ensure data accuracy and proper labeling of plant conditions.

For disease prediction, the dataset is categorized into three main classes: Healthy, Nutrient/Environmental Stress, and Diseased. A sample dataset may consist of a total of 1200 data samples, where the Healthy class contains 400 samples, the Nutrient/Environmental Stress class contains 350 samples (including conditions like nitrogen deficiency, potassium deficiency, and water stress), and the Diseased class contains 450 samples (including diseases such as bud rot, leaf spot, stem bleeding, and root wilt). This balanced distribution of data helps in effectively training the machine learning model for accurate classification and reliable prediction. The sample dataset is shown in Table 1 and the sample picture in input dataset is depicted in Figure 2.

Table 1: Sample dataset

Class	Disease Name	Total no of images
Healthy	Healthy	400
Nutrient/Environmental Stress	Nitrogen Deficiency	125
	potassium deficiency	120
	water stress	105
Diseased	Bud rot	124
	Leaf spot	130
	Stem Bleeding	89
	Root wilt	107
Total		1200



Figure 2: Sample data image on each class

Pre-Processing

The images gathered can have varying sizes, luminance, angles and backgrounds. Thus, the preprocessing is needed to enhance the quality of images and preprocess the data to extract features (Chandy, A. 2019). In the preprocessing phase, all images are resized to some predetermined dimension, filtered to remove noise, contrast is boosted, conversion to different color space where necessary and normalization of pixel values are achieved.

In case an image is resized to $m \times n$ then:

$$I_i = I'_i \in R^{m \times n} \quad (1)$$

For normalization of pixel intensity:

$$I_{norm}(x, y) = \frac{I(x, y)}{255} \quad (2)$$

This scales pixel values into the range [0,1], improving numerical stability.

Segmentation

In order to enhance disease detection, the infected part may be separated with the background or healthy part. This aids the model in paying attention to disease-specific trends. Segmentation can be done by various methods, in this study thresholding based method was applied. The purpose of this segmentation step is that, disease symptoms have a tendency to be localized in certain areas of images.

Let the diseased segmented region be represented as:

$$S_i = f_{seg}(I'_i) \quad (3)$$

Where,

f_{seg} is a segmentation function, S_i is segmented diseased region of the image.

Feature Extraction

The feature vector is turned into the segmented image. These extracted features are the features of the disease in visual form (Kavithamani, V., & Uma Maheswari, S. 2023). There are three significant types of image features that could be utilized in the proposed work, including color feature, texture and shape feature.

Color features

The symptoms of disease are usually manifested in the form of color change, yellowing, browning, black spots, gray spots. The mean and standard deviation values of RGB or HSV channels could be obtained.

For mean color intensity:

$$\mu_c = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N I_c(x, y) \quad (4)$$

Where, $I_c(x, y)$ color intensity of channel c , M, N -image characteristics.

Texture Features

The texture characteristics are used to record the roughness of the lesion, distribution of spots and surface irregularities. One of them is a Gray Level Co-occurrence Matrix (GLCM). Contrast, energy, homogeneity and correlation are features that can be obtained out of GLCM.

• Contrast

$$Contrast = \sum_i \sum_j (i - j)^2 P(i, j) \quad (5)$$

- *Energy*

$$Energy = \sum_i \sum_j P(i, j)^2 \quad (6)$$

- *Homogeneity*

$$Homogeneity = \sum_i \sum_j \frac{P(i, j)}{1 + |i - j|} \quad (7)$$

Where, $P(i, j)$ is the normalized GLCM.

Shape Features

Shape features characterize the geometry of the affected areas like lesion area, perimeter and eccentricity.

For region area:

$$Area = \sum_{x=1}^M \sum_{y=1}^N S_i(x, y) \quad (8)$$

To compactness the perimeter-wise:

$$Compactness = \frac{Perimeter^2}{4\pi \times Area} \quad (9)$$

Those features are used to differentiate diseases that have dissimilar lesion patterns.

Feature Vector Formation

All features obtained are summed up into a single numerical vector per image.

$$F_i = [f_1, f_2, \dots, f_k] \quad (10)$$

Where, F_i is feature vector of image I_i , k total number of extracted features.

The end result of classification is the following dataset:

$$D_f = \{(F_i, y_i)\}_{i=1}^N \quad (11)$$

This makes the problem of image classification a supervised machine learning problem.

Classification

Random Forest (RF) algorithm is an ensemble machine learning method that is employed to classify the diseases of coconut tree on the basis of features extracted out of the disease images. The RF system in the suggested system analyses visual features of images of coconut plants, including color, texture, and shape among others to identify the type of disease. Classification has various steps, which include input of features up to the final prediction.

A random Forest classifier is trained on the feature vectors. Random Forest is an ensemble learning algorithm which is created by constructing several decision trees and integrates their results.

Suppose that there are T decision trees in the forest.

$$RF = \{Tree_1, Tree_2, \dots, Tree_T\} \quad (12)$$

Each tree is also trained with a bootstrap sample of the data and a random subset of features are taken into consideration to split at each node.

Given a test image feature vector F_{new} , the trees are all used to make predictions of a class label.

$$Tree_t(F_{new}) = y_t \quad (13)$$

Majority voting is used to get the final class prediction:

$$y' = mode\{y'_1, y'_2, \dots, y'_T\} \quad (14)$$

Where, y' is the final prediction of disease type. In this manner, the image falls under one of the five disease categories.

Bootstrap Sampling

For each tree, Select data randomly in training dataset. This sampling technique is also referred to as bootstrap aggregation (bagging).

$$D_t \in D \quad (15)$$

Where D_t is the bootstrap set of the t^{th} tree.

Node Splitting using Gini Index

In the process of building the tree, the algorithm chooses the most optimal feature to divide the data. Usually, the Gini index is the impurity measure used by random forest.

$$Gini(S) = 1 - \sum_{i=1}^c p_i^2 \quad (16)$$

Where, S dataset of a node, c classes and p_i probability of class i . The split is selected to give a minimum Gini index and this way, less heterogeneous child nodes are obtained.

Decision Tree Construction

All trees divide the data recursively by using chosen features. Example rule in a tree:

When $Texture > 0.45$ then the image will be Gray Leaf Spot, it will pass to another branch.

The tree goes on to divide, until, maximum depth of tree is attained, or the samples of nodes are reduced too small. A predicted disease class is stored in each leaf node.

Prediction Phase

Upon a new image of a coconut leaf being fed in as input:

- Image processing and feature isolation is done.
- The feature vector is inputted to each decision tree in the forest.

The trees generate a separate prediction:

Majority Voting for Final Classification

The majority voting technique is used to combine predictions of all trees to give the final disease prediction.

$$y' = \text{mode}(y'_1, y'_2, \dots, y'_T) \quad (17)$$

Where, y' predicted coconut disease class final.

Multi-Class Disease Classification i.

On account of the five disease classes in the dataset, the RF model does multi-class classification. The classifier is used to classify the image into one of the following:

$$Y \in \{C_1, C_2, \dots, C_5\} \quad (18)$$

The Flow chart of classification is shown in Figure 3.

Algorithm 1: Image-Based Coconut Disease Classification Using Random Forest

Input: Coconut disease image dataset

Output: Predicted disease class

Step 1: Input labeled coconut disease image dataset.

Step 2: Preprocess all images.

Step 3: Extract color, texture, and shape features.

Step 4: Form feature matrix and label vector.

Step 5: Split the dataset into training and testing sets.

Step 6: Set Random Forest parameters.

Step 7: For each tree in the forest:

Draw a bootstrap sample from training data.

Select random subset of features at each node.

Split nodes using Gini index.

Grow the tree until stopping condition is met

Step 8: Combine all trained trees to form the Random Forest model.

Step 9: For each test sample, send the feature vector to all trees.

Step 10: Collect class predictions from all trees.

Step 11: Apply majority voting to obtain the final class.

Step 12: Evaluate the classification performance using accuracy, precision, recall, and F1-score.

Classification Algorithm

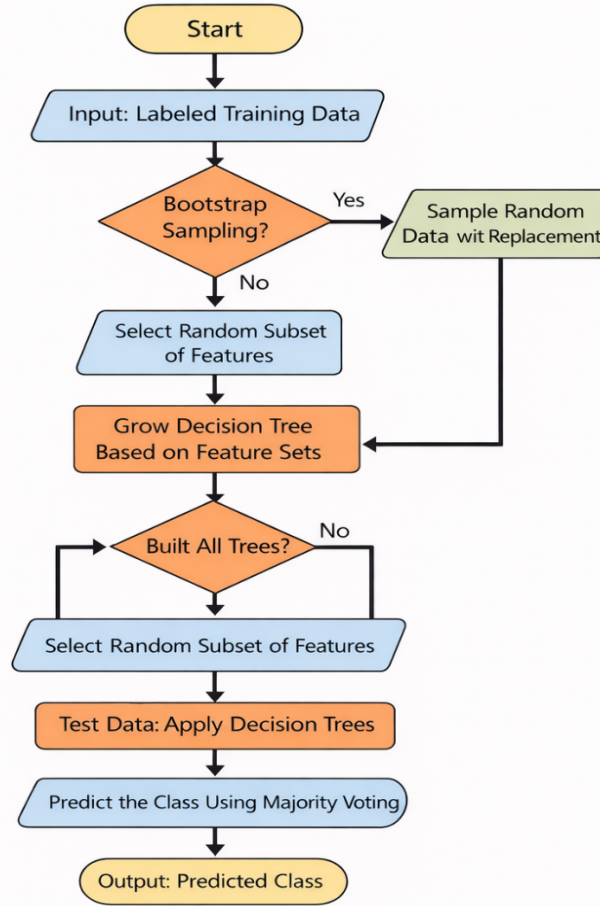


Figure 3: Flow chart of classification

RESULT ANALYSIS

The analysis of results determines the efficiency of the proposed coconut disease classification system based on the use of Random Forest (RF). The goal of this analysis is to determine the accuracy of the model in classifying the images of coconut disease into the five categories namely Bud Root Dropping, Bud Rot, Gray Leaf Spot, Leaf Rot, and Stem Bleeding. The standard classification measures based on the confusion matrix used to assess the performance of the model are accuracy, precision, recall, and F1-score.

Accuracy

Accuracy deals with the general accuracy of the model. Accuracy can also be obtained using the formula below in case of multi-class classification.

$$Accuracy = \frac{\text{Total Correct Predictions}}{\text{Total Number of Samples}} \quad (19)$$

Precision

Precision is a measure of the predicted cases of the disease which are accurate.

$$Precision = \frac{TP}{TP+FP} \quad (20)$$

Recall

Precision is a measure of the number of correct predicted cases of disease.

$$Recall = \frac{TP}{TP+FN} \quad (21)$$

F1-Score

F1-score is a harmonic mean of both precision and recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (22)$$

CONFUSION MATRIX

Multi-class classification is assessed by means of the confusion matrix. It compares the real class

names to the predicted class names. Figure 4 indicates the performance of the proposed model-based coconut disease classification model using the Random Forest in terms of the five disease categories: Bud Root Dropping, Bud Rot, Gray Leaf Spot, Leaf Rot and Stem Bleeding. The diagonal component of the matrix reflects the correct samples and off-diagonal component of the matrix reflects the misclassifications. As in the heatmap, the most accurate predictions are made with the Gray Leaf Spot (2080) and Leaf Rot (1600), which means that the model is able to learn the visual peculiarities of the diseases. On the same note, high correct prediction is recorded on Bud Root Dropping (480), Bud Rot (440) and Stem Bleeding

(990) indicating a uniform classification behavior among all classes of diseases. The fact that the off-diagonal cells have relatively small values implies the lack of any significant confusion between the categories of diseases, which proves the strength of the feature extraction and the power of the Random Forest ensemble learning method. Altogether, the general tendency to the bigger values distribution on the diagonal would indicate the high classification rates and good detection of coconut diseases in the proposed model, which means that the model can be applied practically in the sphere of precision agriculture and automated systems of crop health monitoring.

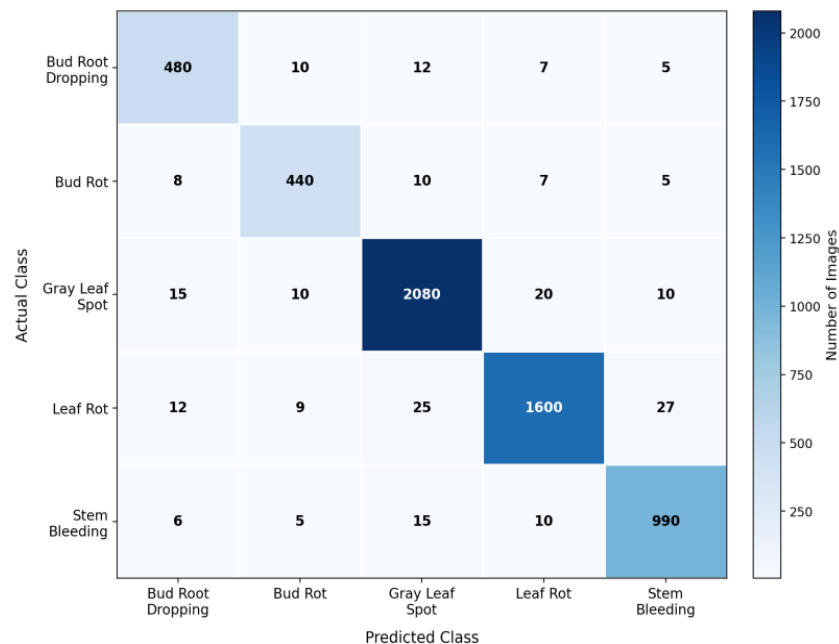


Figure 4: Confusion matrix

PERFORMANCE COMPARISON

The classification comparison of the proposed work is provided in Figure 5 and Table 2 shows the performance of the proposed classification model on five coconut disease classes, namely, Bud Root Dropping, Bud Rot, Gray Leaf Spot, Leaf Rot, and Stem Bleeding, in terms of four evaluation metrics: accuracy, precision, recall, and F1-score. As depicted in the graph, the model is relatively consistent in its performance with all types of

diseases because the accuracy stood at 85.9 to 87.2. The best performance is witnessed with Gray Leaf Spot with accuracy, precision, recall and F1-score of 87.2, 87.2, 87.4 and 87.2 respectively, showing that the model is effective to detect this disease because of its characteristic appearance. Likewise, Leaf Rot and Stem Bleeding also exhibit good classification performance with equal precision and recall values indicating a good performance in detection of the disease.

Table 2: Performance results of the proposed work

Disease Class	Accuracy	Precision	Recall	F1-Score
---------------	----------	-----------	--------	----------

Bud Root Dropping	85.9	82.1	83.4	82.7
Bud Rot	86	82.8	83.6	83.2
Gray Leaf Spot	87.2	87.1	87.4	87.2
Leaf Rot	86.6	87.3	85.6	86.4
Stem Bleeding	86.5	85.5	86.5	86

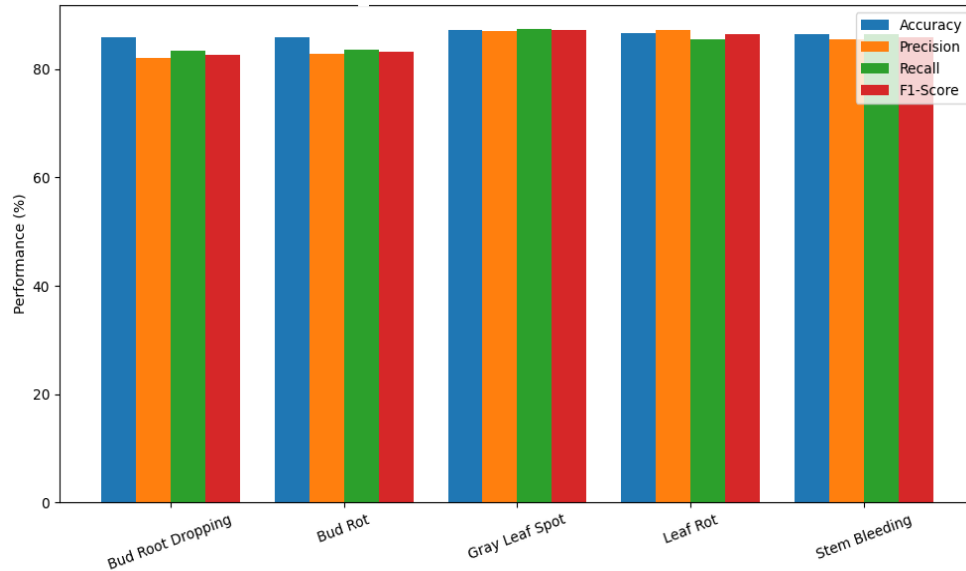


Figure 5: Overall comparison

The values of Bud Root Dropping and Bud Rot were slightly lower which means that there were slight misclassifications, that is probably due to the similarity of visual symptoms of the diseases. Altogether, the graph shows that the suggested model can achieve consistent and balanced performance in all the metrics, which proves its ability to effectively classify the diseases of coconut

trees and justifies its usage in automated plant diseases detection in precision agriculture systems. As can be seen in the comparison Table 3, the suggested Random Forest-based model performs better than a number of traditional machine learning algorithms, such as Decision Tree, KNN, SVM, Naive Bayes, and Logistic Regression.

Table 3: Performance comparison with existing methods

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	82.4	80.5	81.2	80.8
K-NN	83.7	81.6	82.3	81.9
SVM	84.5	82.9	83.7	83.2
NB	81.3	79.8	80.4	80.1
LR	83.2	81.4	82	81.7
Proposed RF	86.4	85	85.3	85.1

The accuracy of the proposed approach is about 86.4 on the whole, which is greater compared to other baseline algorithms. Also, the model has balanced precision, recall and F1-score, which implies that the model is efficient at reducing both false positives and false negatives in the classification of diseases. This enhancement of performance is explained by the fact that the

Random Forest ensemble learning is able to use multiple decision trees to increase the accuracy of prediction and decrease overfitting. Therefore, the suggested approach offers a more accurate and sounder remedy to find coconut tree disease and categorize it within the precision agriculture systems.

The performance of various machine learning algorithms used to classify coconut disease is shown in the comparison graph in terms of accuracy, precision, recall, and F1-score. Based on the graph it is possible to note that the proposed Random Forest model is better than other traditional algorithms such as Decision Tree, K-Nearest Neighbors (KNN), Support Vector

Machine (SVM), Naive Bayes and Logistic Regression. Although the current algorithms boast of ranging accuracy at around 81-84.5, the suggested algorithm boasts of the highest rate at approximately 86.4 and a better value of precision, recall and F1-score. Performance comparison with existing methods is depicted in Figure 6.

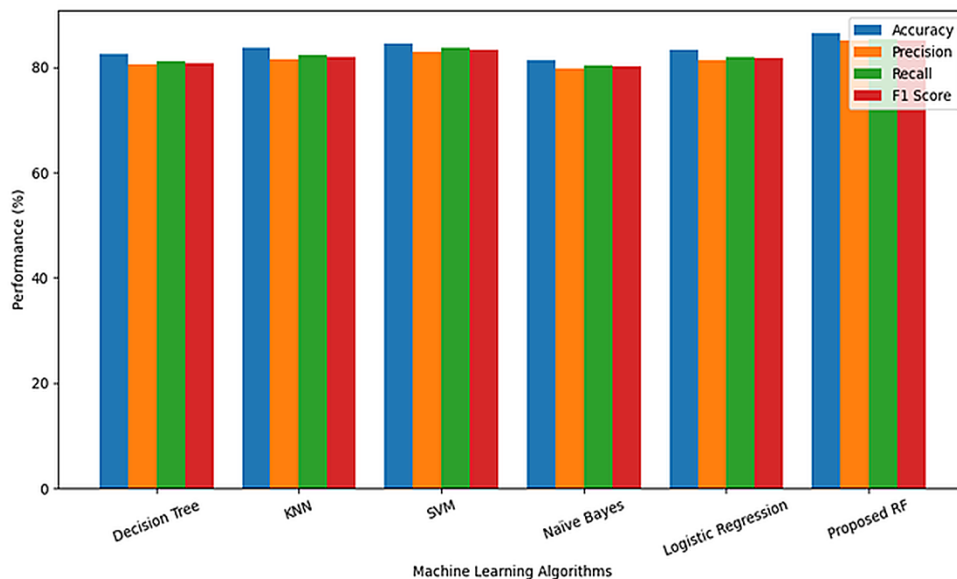


Figure 6: Performance comparison with existing methods

This is largely attributed to the fact that the Random Forest has the ability to ensemble learn by combining a number of decision trees to increase the accuracy of classification and minimize overfitting. Hence, these findings suggest that the presented method is more valid and strong in terms of detection of coconut disease to be used in the practice of the detection of this disease in precision agriculture and automated crop monitoring systems.

CONCLUSION

This paper suggested a machine learning model of classifying coconut tree disease through the random forest algorithm. The system has five major diseases known as Bud Root Dropping, Bud Rot, Gray Leaf Spot, Leaf Rot and Stem Bleeding which are identified by analyzing the images of the coconut plant. The suggested architecture comprises the main steps that should be involved which include image preprocessing, feature extraction, disease classification and performance evaluation. The experiment results reveal that the

model has a total accuracy of around 86, and all the values of the precision, recall, and the F1-score are balanced, showing that it is able to detect various categories of diseases with a high level of reliability.

In addition, the proposed model was compared to various machine learning algorithms that are currently being used, including Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes, and Logistic Regression. The outcome confirms the fact that the Random Forest model is superior to these standard approaches mostly because of the ability of an ensemble learning, which enhances accuracy of the prediction and prevents overfitting. In general, it is possible to note that the given approach has good perspectives of automated coconut disease detection and precision farming implementations, as it can assist farmers in monitoring the health of crops and taking some timely preventative actions.

ACKNOWLEDGEMENT

We would like to thank A.V.V.M Sri Pushpam College (Autonomous), Poondi, Thanjavur, which is affiliated to Bharathidasan University, Tamilnadu, for the helpful support of conducting research in an effective manner.

Source of support: None

Conflict of Interest: None

REFERENCES

- Adivarekar, P. P., A. A. P., Sharma, S., P. D., Elangovan, M., & Rastogi, R. (2023). Automated machine learning and neural architecture optimization. *The Scientific Temper*, 14(4), 1345–1351. <https://doi.org/10.58414/SCIENTIFICTEMPER.2023.14.4.42>.
- Ai, Y., Sun, C., Tie, J., & Cai, X. (2020). Deep learning based recognition framework for crop disease and pest identification under challenging field environments. *IEEE Access*, 8, 171686–171693.
- Chaitra Bai, J., Chandana, B., Geetha Priya, T., Mohana Kumari, K., Kavitha, M., & Maithri. (2026). Coconut tree disease prediction using AI. *International Journal of Advanced Research in Science, Communication and Technology*, 6(1), 1–10.
- Chandy, A. (2019). Pest infestation identification in coconut trees using deep learning. *Journal of Artificial Intelligence and Capsule Networks*, 1, 10–18. <https://doi.org/10.36548/jaicn.2019.1.002>.
- DebMandal, M., & Mandal, S. (2011). Coconut (*Cocos nucifera* L.: Arecaceae): In health promotion and disease prevention. *Asian Pacific Journal of Tropical Medicine*, 4(3), 241–247.
- Kathavate, P. N., & Mahant, M. A. (2025). Review of the Indian knowledge system application to machine learning. *The Scientific Temper*, 16 (Special Issue 1).
- Kavithamani, V., & Uma Maheswari, S. (2023). Investigation of deep learning for whitefly identification in coconut tree leaves. *Intelligent Systems with Applications*, 20, Article 200277. <https://doi.org/10.1016/j.iswa.2023.200277>.
- Maray, M., Albraikan, A. A., Alotaibi, S. S., Alabdan, R., Duhayim, M. A., Al-Azzawi, W. K., & Alkhayyat, A. (2022). Artificial intelligence-enabled coconut tree disease detection and classification model for smart agriculture. *Computers and Electrical Engineering*, 104, Article 108428.
- Paymode, S. A., & Malode, B. V. (2022). Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG. *Artificial Intelligence in Agriculture*, 6, 23–33.
- Pinho, R. C., Farias, P. R. S., Batista, T. F. V., Santos, A. V. F. D., & Bastos, L. F. (2023). Spatial patterns of the leaf crown borer *Eupalamides cyparissias* (Cramer, 1775) (Lepidoptera: Castniidae) in coconut (*Cocos nucifera* L.) tropical region. *Brazilian Journal of Biology*, 83, Article e270870.
- Pravitha, P. S., & Kumar, V. (2023). Unlocking the potential of transfer learning for leaf disease detection: A comprehensive study with CNN and pre-trained architectures. In *Proceedings of the International Conference on Computational Intelligence, Networks and Security* (pp. 1–6). IEEE.
- Rajesh, R. S., Karthikeyan, M., & Devi, S. R. (2024). Automated detection of coconut leaf diseases using CNN and image processing techniques. *International Journal of Agricultural Technology and AI*, 3(2), 45–52.
- Ramu, S. P., Daga, M., & Parikh, D. (2025). *Deep learning-based coconut tree disease identification for smart farming applications* (Technical Report pp. 1–10). VIT School of Computer Science.
- Singh, P., Verma, A., & Alex, J. S. R. (2021). Disease and pest infection detection in coconut tree through deep learning techniques. *Computers and Electronics in Agriculture*, 182, Article 105986.
- Thite, S., Suryawanshi, Y., Patil, K., & Chumchu, P. (2023). Coconut (*Cocos nucifera*) tree disease dataset: A dataset for disease detection and classification for machine learning applications. *Data in Brief*, 51, Article 109690.
- Wijethunga, C. D., Ishanka, K. C., Parindya, S. D. N., Priyadarshani, T. J. N., Buddika Harshanath, & Samantha Rajapaksha. (2023). Coconut plant disease identified and management for agriculture crops using machine learning. *International Journal of Engineering and Management Research*, 13(5), 72–78.