Hybrid Approach of Cotton Disease Detection for Enhanced Crop Health and Yield

Article *in* IEEE Access · July 2024

DOI: 10.1109/ACCESS.2024.3419906

Rahul Kumar
Amity University

Kottakkaran Sooppy Nisar
Prince Sattam bin Abdulaziz University

Siddharth Singh Chouhan
VIT Bhopal University

Anoop Tiwari
National Institute of Technology, Warangal

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

Hybrid Approach of Cotton Disease Detection for Enhanced Crop Health and Yield

Rahul Kumar[1], Ashok Kumar [2], Karamjit Bhatia [3] , Kottakkaran Sooppy Nisar[4], Siddharth Singh Chouhan[5], Priti Maratha[6], Anoop Kumar Tiwari[7,]*

[1] Department of Computer Science & Engineering, Amity School of Engineering and Technology, Amity University, Gwalior, Madhaya Pradesh, India
[2]Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Greenfield, Vaddeswaram, Guntur, Andhra Pradesh,India
[3]Department of Computer Science, Gurukula Kangri (deemed University), Haridwar, India
[4]Department of Mathematics, College of Science and Humanities in Alkharj, Prince Sattam Bin, Abdulaziz University, Alkharj 11942, Saudi Arabia
[5]School of Computing Science and Engineering, VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore, Madhya Pradesh – 466114
[6,7]Department of Computer Science and Information Technology, Central University of Haryana- 123031
*Corresponding author: (e-mail: dranoop@cuh.ac.in).

ABSTRACT The well-being of cotton crops is of utmost importance for maintaining agricultural productivity, and the early detection of diseases plays a critical role in achieving this objective. This study introduces a comprehensive approach for creating a machine learning-based system capable of identifying diseases in cotton plants through the analysis of leaf images. The research encompasses stages such as acquiring the dataset, pre-processing the data, training the model, developing an ensemble model, evaluating the models, and analyzing the results. Several machine-learning models are trained and evaluated to determine how well they can classify cotton leaves as "Healthy" or "Diseased." These models include Random Forest, Support Vector Machine (SVM), Multi-Class SVM, and an Ensemble model. This investigation yields a practical and visually informative system for disease detection, which can contribute to disease prevention, thereby enhancing both crop yield and quality. This work underscores the significance of continuous improvement by periodically updating the models and explores the potential of advanced techniques such as deep learning.

INDEX TERMS Cotton disease detection, Crop health, Disease prevention, Ensemble model, Image classification.

I. INTRODUCTION

Cotton production in India holds a pivotal position in the country's agricultural and economic landscape. As one of the globe's major cotton producers, India assumes a central role in the worldwide textile industry, catering to both domestic consumption and international trade [1]. The cotton sector in India not only provides a substantial source of livelihood for millions of farmers but also constitutes a critical component of the nation's industrial and export sectors. This introduction offers an overview of the significance of cotton production in India, its historical importance, the obstacles it encounters, and its central role in the Indian economy and rural communities [1]. Cotton cultivation in India boasts a rich historical heritage spanning millennia. The cotton plant, scientifically known as Gossypium, is believed to have been cultivated in the Indian subcontinent for thousands of years, with references to cotton textiles and trade found in ancient texts [20]. India's native cotton varieties have contributed to its status as a center for cotton cultivation, leaving an enduring imprint on the country's cultural and economic legacy [2]. Cotton production in India has served as a cornerstone of the nation's economy for several decades. The sector makes a substantial contribution to the agricultural Gross Domestic Product (GDP) and offers employment to a vast workforce, particularly in rural regions. India's cotton finds application in the production of a wide range of textile products, encompassing apparel, home textiles, and industrial textiles. The export of cotton and cotton products constitutes a significant source of revenue for the nation, positioning India as a prominent player in the global textile market [2]. While cotton production in India has flourished, it grapples with various challenges. Factors such as weather conditions, pest infestations, and market fluctuations can impact cotton yields and profitability. Furthermore, discussions surrounding sustainable practices, water usage, and genetic modification have been a focal point for the cotton sector [21]. Farmers, policymakers, and industry stakeholders continually strive to address these challenges, with an emphasis on enhancing crop yields, sustainable methods, and the income of cotton-growing communities [2]. Cotton farming profoundly influences rural livelihoods in India. Smallholder farmers and rural communities, particularly in states like Gujarat, Maharashtra, Andhra

Pradesh, and Punjab, are primarily engaged in cotton cultivation. For these communities, cotton represents more than just a cash crop; it serves as a source of sustenance, employment, and social cohesion. Understanding the dynamics of cotton production is imperative for the betterment of these rural populations [3]. Following some of the common cotton diseases found in cotton leaves in India:

A. Cotton Leaf Curl Disease (CLCuD)

This virus particularly hard hits India's cotton crops. The main cause of it is begomoviruses, which are spread by whiteflies. Symptoms include yellowing, twisted and curled leaves, and stunted growth of the plant [25].

FIGURE 1. Cotton Leaf Curl Disease [35]

B. Bacterial Blight:

Use the bacterium Xanthomonas axonopodis pv. malvacearum attacks cotton plants and causes bacterial blight [22]. It thrives in warm and humid conditions. Symptoms include the appearance of dark, water-soaked lesions on leaves, which can eventually turn brown and necrotic. In severe cases, defoliation can occur, leading to reduced yield and lower cotton quality [3].

FIGURE 2. Bacterial Bright [36]

C. Alternaria Leaf Spot:

The fungus Alternaria macrospora is the source of Alternaria leaf spot. It manifests as small, circular lesions on cotton leaves, which can develop a concentric ring pattern [23]. These lesions are initially light brown but can darken over time. While this disease may not lead to significant yield losses, it can reduce the quality of cotton lint, making it less desirable for the textile industry [4].

FIGURE 3. Alternaria leaf spot [37]

D. Ascochyta Leaf Spot:

Ascochyta gossypii [24], a fungus, causes angular necrotic spots on cotton leaves. These spots often have a dark border and can coalesce to form larger lesions. Severe infections can lead to defoliation and reduced yield. This disease is more prevalent in cooler and wetter regions [4].

E. Rust:

If Cotton rust is caused by the fungus Phakopsora gossypii. It results in the formation of rusty-orange pustules on the undersides of leaves. While it may not always cause severe yield losses, heavy infections can weaken the plant and affect cotton production [26].

F. Wilt Diseases:

Various wilt diseases, such as Verticillium and Fusarium wilt, are soil-borne pathogens that infect the plant's vascular system. Symptoms include wilting, yellowing, and, in some cases, plant death [27]. These diseases are challenging to control, and once a field is infested, it can be problematic for subsequent cotton crops [5].

G. Grey Mildew:

Grey mildew, caused by the fungus Ramularia areola, leads to the development of greyish-white powdery growth on the upper surface of cotton leaves [31]. This growth can reduce photosynthesis and overall plant health, impacting cotton yield and quality [5].

H. Powdery Mildew:

Powdery mildew is characterized by the formation of white, powdery fungal growth on the upper leaf surfaces. It is caused by various species of the fungus Erysiphe. The presence of powdery mildew can reduce photosynthesis, weaken plant vigor, and lead to yield losses.

I. Black Root Rot:

The fungus Thielaviopsis basicola attacks the roots and lower stem of cotton plants, resulting in black root rot [26]. Symptoms include blackening and rotting of the root system, which can lead to stunted growth and reduced crop yield.

J. Cotton Mosaic Virus:

Cotton mosaic virus infections result in mosaic-like patterns on cotton leaves, with mottled or streaked discoloration. The virus can cause reductions in yield and affect lint quality, making the cotton less desirable for the textile industry [6].

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

A variety of cultural methods, including crop rotation, planting disease-resistant cotton types, and putting integrated pest management strategies into effect, are frequently used in the treatment of these diseases. In some cases, chemical treatments may also be used, but these are typically reserved for situations when other methods prove inadequate. Early and accurate disease detection is crucial for effective management and mitigation of the impact of these diseases on cotton crops in India [6].

In this context, it is crucial to recognize that the success of cotton production in India extends beyond crop yields and textile manufacturing to touch the lives of millions. As this research delves into the realm of cotton disease detection using machine learning [28], it takes place within the broader framework of a vibrant and multifaceted cotton industry that continues to adapt and evolve in response to changing times and challenges.

Cotton production in India plays a vital role in the nation's economy and agricultural sector. Nevertheless, the cotton industry encounters a multitude of challenges, among which the menace of diseases looms large, capable of severely diminishing crop yields and adversely affecting the livelihoods of countless cotton farmers. Various fungal, bacterial, and viral diseases [27] pose threats to cotton plants, leading to economic setbacks and a decline in cotton quality. Timely and accurate disease detection is imperative to confront these challenges [7].

Traditional methods of disease detection often hinge on visual scrutiny, a subjective process susceptible to human errors. Machine learning and computer vision technologies [28] offer a more objective and efficient approach to identify and categorize cotton diseases based on leaf images. This research is dedicated to the development of a machine learning-based system to address this critical issue. Following are the significant contribution of the paper [7]:

a) Research on cotton disease detection in India has significant implications for the livelihoods of smallholder farmers and socio-economic challenges in rural areas.
b) Enhancing crop health through machine learning and computer vision technologies improves agricultural productivity and economic stability in India.
c) Interdisciplinary collaboration between agriculture and technology holds immense potential for advancing agricultural practices.
d) Precision and efficiency in pesticide utilization, reduced chemical waste, and eco-friendly agricultural practices contribute to environmental sustainability.
e) India's innovations and best practices in cotton disease detection can be shared globally, benefiting the broader agricultural community.
f) The fusion of traditional agriculture with cutting-edge technology has far-reaching consequences for India and the global agricultural sector [7].

Research Objective for develop and deploy a machine learning-based system for accurate and practical cotton disease detection.

a) Prepare and standardize a diverse dataset of cotton leaf images for disease classification.
b) Optimize model [29] performance through hyper parameter tuning and data segmentation.
c) Evaluate model accuracy and effectiveness using various metrics and visual representations.
d) Verify system accuracy and explore practical deployment strategies in real-world scenarios [7].
e) Foster ongoing improvement through regular model updates and incorporation of advanced techniques [30].

Cotton plants are susceptible to a range of diseases stemming from fungi, bacteria, viruses, and other pathogens, resulting in substantial yield losses that impact both the economic sustainability of cotton farming and the livelihoods of farmers. The research on cotton disease detection holds significant importance for several reasons. Firstly, it aids in preserving the economic viability of cotton cultivation by mitigating yield losses associated with diseases. Secondly, it fosters sustainable agricultural practices by facilitating precise interventions and reducing reliance on excessive pesticide application. Thirdly, advancements in disease detection technology contribute to scientific knowledge and encourage the adoption of cutting-edge solutions within farming communities. Ultimately, this research serves a crucial role in bolstering global food security by bolstering the resilience of cotton crops, which serve as a staple commodity in numerous regions.

II. LITERATURE SURVEY

Cotton is an important cash crop in India, and Bhimte, N. R., & Thool, V. R. (2018) [6] talked about how illnesses have caused large output losses. They suggested a method that makes use of image processing tools to diagnose cotton leaf diseases early. In this system, sick and healthy cotton leaves were classified using a Support Vector Machine (SVM) classifier. A digital camera was used to take pictures from cotton fields, and different pre-processing techniques were used to separate the diseased portions of the cotton leaves, including background removal, color-based segmentation, and filtering.

Kothari, J. D. (2018) [7] focused on the role of computerization and machine learning in agriculture. The paper highlighted the impact of plant diseases on agricultural outcomes and the need for efficient disease detection methods. It discussed the application of machine learning models for early-stage plant disease detection to improve crop quality and sustainability. The goal was to leverage machine learning techniques for more accurate and timely disease detection to benefit the agricultural sector.

The significance of rice for Bangladesh's economic and food security was discussed by Ahmed, K., et al. in 2019 [8]. A method for identifying the three common diseases of rice

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

plants—leaf smut, bacterial leaf blight, and brown spot—was provided in the paper. Machine learning algorithms, such as K-Nearest Neighbour (KNN), Decision Tree (J48), Naive Bayes, and Logistic Regression [31], were utilized to classify the disease from clear photographs of the afflicted rice leaves. After ten-fold cross-validation, the Decision Tree algorithm's accuracy was over 97%.

Fenu, G., & Malloci, F. M. (2019) [9] discussed the impact of Big Data Analytics and Machine Learning on the agricultural sector, particularly in mitigating the challenges posed by climate change. The paper focused on the devastating potato late blight disease in Sardinia and proposed a system for predicting disease severity using meteorological parameters. Machine learning models, including a Feed-forward Neural Network and Support Vector Machine Classification [32], were employed. The prediction accuracy for the Artificial Neural Network (ANN) was 96%, and for SVM Classification, it was 98%.

Akulwar, P. (2020) [10] highlighted the importance of intelligent agricultural systems for the Indian economy. It discussed the collaboration of recommender systems with machine learning to assist farmers in making informed decisions related to crop management, agriculture commodity prices, cultivar selection, cultivation timing, and investment prioritization. The paper also mentioned the role of machine learning in crop management, including disease detection and yield prediction.

Reddy, P. C., et al. (2021) [11] discussed the increasing productivity in agriculture and the importance of plant disease detection for enhancing crop quality. The study concentrated on the detection of plant leaf diseases through the application of cutting-edge machine learning techniques, such as Support Vector Machine (SVM) and Random Forest algorithms. It stressed that in order to boost agricultural productivity, effective, affordable, and time-saving techniques are required. Trivedi, V. K., et al. (2021) [12] presented an automatic system for the segmentation and detection of leaf diseases from plant images. Their method involved preprocessing to reduce background noise, K-means segmentation to isolate diseased regions, and the extraction of features from these regions. The disease classification system utilized a Support Vector Machine (SVM) classifier, and it attained a 90.6% accuracy rate.

Karthika, J., et al. (2021) [13] emphasized the importance of using image processing techniques for early and accurate plant disease detection, especially in the case of leaf diseases. Their proposed methodology used a subtractive pixel adjacency matrix (SPAM) for feature extraction and employed classifiers such as SVM and KNN to distinguish between healthy and diseased plant leaves. SVM achieved a better accuracy rate of 93.67%.

Prasad, A., et al. (2022) [14] discussed the challenges associated with drone-based plant disease diagnosis, including the trade-off between resolution and speed and the lack of labeled training data. They proposed a two-step machine learning approach that analyzed low-fidelity and high-fidelity images to balance efficiency and accuracy. The study included the use of generative networks to produce high-fidelity data and machine learning classifiers to identify potentially diseased plants. The results showed high accuracy for the high-fidelity system and reasonable confidence for the low-fidelity system.

Jayaramu, H. K., et al. (2022) [15] discussed the tedious and time-consuming nature of plant leaf disease detection and the importance of early detection. Their methodology utilized the subtractive pixel adjacency matrix (SPAM) for feature extraction and the exponential spider monkey optimization technique (ESMO) to select optimum features. The system used classifiers like SVM and kNN to classify plant leaf data as healthy or diseased, with SVM achieving a better accuracy of 93.67%.

Radhika et al. (2023) [16] highlighted the significant contribution of agriculture to India's GDP and the challenges posed by plant diseases. The paper discussed the use of image processing techniques to detect plant diseases early, especially those that manifest as changes in leaf or stem color. Image capture, preprocessing, spot segmentation, feature extraction, and classification utilizing thresholding and K-means clustering techniques were among the methods used. With the purpose of training classifier models for illness identification, the study concentrated on diseases such Ashen mold, Late scorch, Cottony mold, Early scorch, and Ting whiteness.

Singh, P., et al. (2023) [17] focuses on the significance of cotton production in India, where agriculture plays a crucial role in sustaining the economy and supporting around 70% of rural households. To enhance disease detection, the study utilizes deep learning, creating a model with a diverse dataset of 22 cotton leaf diseases. The proposed model, based on CNN, achieves a remarkable 99.39% accuracy, surpassing existing methods and offering potential for real-time implementation to aid farmers in disease management.

Gülmez, B. (2023) [18] emphasizes the importance of plants in the ecosystem, particularly the critical role of cotton in human use and income generation. Recognizing the susceptibility of cotton to diseases, the research introduces a deep convolutional neural network model for disease detection in cotton leaves. The model, optimized using the grey wolf optimization algorithm, outperforms well-established models like ResNet50, VGG19, and InceptionV3, achieving a perfect accuracy of 1.0 compared to their lower accuracy values.

Zhang, N., et al. (2023) [19] use spectrum analysis of naturally infected cotton plants to address the problem of early cotton verticillium wilt disease identification. The study creates a grading model using a variety of preprocessing methods and support vector machine (SVM) models; the MSC-db3(23)-GWO-SVM model turns out to be the best one. With an accuracy of 91.2%, this model outperforms existing models and shows promise in the classification of illness severity levels. [19].

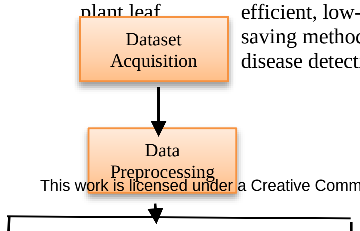
This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

Zhang, N., et al. (2023) categorized infected cotton plants based on disease severity and employed preprocessing techniques and SVM models optimized with various algorithms to establish a grading model. Their results showed that the MSC-PSO-SVM model outperformed others in classification accuracy, while the MSC-db3(23)-GWO-SVM model was identified as optimal, achieving high accuracy across disease severity levels. The study underscores the effectiveness of spectral technology in disease classification, aiding field detection and grading [20].

Chepuri, S., & Ramadevi, Y. (2023, October) highlighted the economic losses associated with cotton diseases and proposed automated disease detection systems to mitigate them. Such systems reduce costs related to manual scouting and enable more targeted treatment, improving resource efficiency. However, challenges exist in generalizing image segmentation methods across diverse datasets and disease types. They suggested that deep learning models trained on diverse datasets offer promise in generalization, enabling disease detection in various environmental conditions and severity levels [21].

TABLE I
RESEARCH PAPER STUDY

Author Name	Year	Proposed Concept	Major Findings
Bhimte, N. R. & Thool, V. R. [6]	2018	Using SVM classifier and image processing to diagnose cotton leaf disease	Cotton leaf diseases can be identified early with the help of image processing and SVM classifiers.
Kothari, J. [7]	2018	Detection of plant diseases using machine learning to enhance crop quality	Emphasized the significance of employing machine learning to automate the identification of plant diseases.
Ahmed, K., et al. [8]	2019	Machine learning-based rice leaf disease detection system in Bangladesh	Achieved high accuracy in detecting common rice leaf diseases in Bangladesh.
Fenu, G. & Malloci, F. M. [9]	2019	Predicting potato late blight disease using meteorological parameters and machine learning	Successfully predicted potato late blight disease severity with high accuracy.
Akulwar, P. [10]	2020	Intelligent Agriculture System integrating recommender systems and machine learning	Proposed an intelligent system for aiding farmers in decision-making, including disease detection and yield prediction.
Reddy, P. C., et al. [11]	2021	Detection of plant leaf learning to increase agricultural production	Emphasized the importance of efficient, low-cost, and time-saving methods for plant disease detection.
Trivedi, V. K., et al. [12]	2021	Automatic segmentation and detection of leaf diseases in plant images	Developed an automated system with a 90.6% accuracy for plant leaf disease detection.
Karthika, J., et al. [13]	2021	Using machine learning and image processing to detect plant diseases	Achieved accurate disease detection, with SVM reaching 93.67% accuracy.
Prasad, A., et al. [14]	2022	Two-step machine learning approach for drone-based plant diagnosis	Demonstrated a two-step approach that balanced efficiency and accuracy in drone-based plant disease diagnosis.
Jayaramu, H. K., et al. [15]	2022	Using machine learning with SPAM feature extraction, plant leaf disease detection	Developed a methodology for effective disease detection, with SVM achieving 93.67% accuracy.
Radhika, V., et al. [16]	2023	Plant disease early detection using image processing	Highlighted the potential of image processing for early plant disease detection.
Singh, P., et al. (2023) [17]	2023	Deep learning model for cotton leaf disease detection in India's agriculture	- Utilized deep learning, specifically CNN, for detecting 22 types of cotton leaf diseases. - Achieved an impressive accuracy of 99.39% on the test set.
Gülmez, B. (2023) [18]	2023	Deep convolutional neural network model optimized with grey wolf optimization for cotton disease detection	A deep learning model created especially for the detection of cotton leaf disease was introduced. Grey Wolf Optimization was utilized to optimize the model.
Zhang, N., Zhang, X., Shang, P., et al. [19]	2023	Cotton verticillium wilt illness early detection and grading using spectral analysis and SVM models	Used preprocessing methods and spectral data to examine naturally infected cotton plants. Created a grading model by combining SVM with PSO and GWO optimization methods. - With an accuracy of 91.2%, the MSC-db3(23)-GWO-SVM model showed optimal performance.



III. RESEARCH METHODOLOGY

In the first phase of the cotton disease detection system, the focus is on Dataset Acquisition. Getting a properly annotated dataset with pictures of both disease-free and sick cotton leaves is necessary for this.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

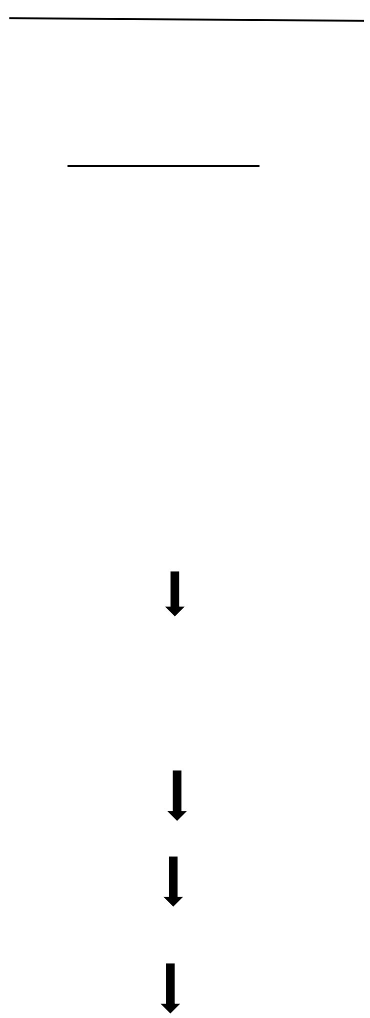


FIGURE 4. **Proposed Flow Diagram [Project Implementation]**

Ensuring that the dataset is diverse and encompasses a representative spectrum of illnesses that may potentially impact cotton plants is imperative. The models' performance is evaluated in the Model Evaluation phase, employing key metrics such as accuracy, precision, recall, F1-score [33], and support for each class. Classification reports are exported to CSV files for further analysis.

The system is designed to test a New Image in the subsequent phase, prompting the user to provide the path to an image for testing. The image is loaded, pre-processed, and a feature vector is created for prediction. Each trained model predicts whether the new image is classified as "Diseased" or "Healthy," and the predictions, including class probabilities, are displayed.

The Visualization of Predictions phase involves implementing functions to create bar graphs depicting probability estimates for each class based on the models' predictions. These graphs visually convey the models' confidence in their predictions.

Further enhancing the visualization aspect, a Combined Predictions Visualization phase produces a bar graph illustrating probability estimates for each class from the SVM, Multi-Class SVM, and Ensemble models. This graph facilitates the comparison of predictions among these models, aiding in decision-making.

Following the extensive experimentation and visualization, Result Analysis is conducted. The goal is to identify the most accurate and reliable model for cotton disease detection, considering trade-offs between different models and their computational requirements.

If the model meets the desired performance criteria, it can be deployed in the Deployment phase as part of an application or system for real-world cotton disease detection. The final phase, Continuous Improvement, emphasizes the importance of regularly updating and retraining the model as more data becomes available or as the dataset grows. Additionally, exploring advanced techniques, such as deep learning, is suggested for enhanced accuracy in the long term.

The provided code adheres to this methodology, demonstrating the creation of a cotton disease detection system using machine learning models and ensemble methods. It displays the process of training, evaluating, and employing multiple models for accurate disease classification, emphasizing practicality and visual representation of results

IV. PROPOSED APPROACH

A. Decision Tree:

A simple but effective machine learning approach used for both regression and classification tasks is the decision tree. It takes on the shape of a tree. A Decision Tree recursively divides the dataset based on the feature that offers the best separation between classes or the optimal split for regression tasks. This splitting process often relies on criteria such as Gini impurity or information gain for classification tasks and mean squared error for regression tasks [34].

The structure of the Decision Tree is shaped by the data. Each internal node represents a condition or a feature-based decision, and each branch emanating from an internal node signifies a potential outcome. Ultimately, the leaf nodes contain the class predictions in the case of classification or the predicted values in the case of regression.

Decision Trees are renowned for their high interpretability. You can easily trace the path from the root node to a leaf node, allowing you to comprehend precisely how a particular decision is made. This transparency is particularly advantageous in cases where understanding the rationale behind predictions is essential

In the context of cotton disease detection, a Decision Tree can be employed to establish rules for classifying cotton leaf images as either "Healthy" or "Diseased." The simplicity and interpretability of Decision Trees make them a valuable tool, and when combined with the ensemble approach of Random Forest, they contribute to the development of a robust and accurate disease detection system.

B. Random Forest:

A flexible ensemble learning technique that may be used for both classification and regression tasks is called Random Forest. It builds upon the principles of decision trees and employs the following techniques:

Decision trees are gathered together to form Random Forest. A random subset of the training data and a random subset of the available features are used to build each unique decision tree. The individual trees are more diverse as a result of this randomization, which reduces the likelihood of overfitting.[34]

When it comes to making a prediction, each decision tree within the Random Forest contributes by casting a vote for a particular class. In the context of cotton disease detection, the classes could be "Healthy" or "Diseased." The class that receives the most votes across all trees is designated as the final prediction. This ensemble approach is commonly referred to as "bagging," which stands for bootstrap aggregating.

Random Forest employs bootstrap sampling, a process where data points are randomly selected from the original dataset with replacement. This procedure generates multiple subsets of data, each used to train an individual decision tree. The use of bootstrap sampling aids in reducing the risk of overfitting by providing each tree with a unique dataset, thereby increasing the model's ability to generalize well to unseen data.

These characteristics collectively contribute to the robustness and effectiveness of Random Forest in making predictions for tasks like cotton disease detection, where the goal is to accurately classify leaf images as either "Healthy" or "Diseased."

C. Hybrid Approach:

The proposed hybrid approach melds Random Forest and Decision Tree methods for cotton disease detection. Here's a breakdown of its functioning

Within this hybrid scheme, Random Forest and Decision Tree models are trained separately, utilizing the same dataset of cotton leaf images. Each model crafts its distinct set of classification rules, relying on the extracted image features.

When a fresh cotton leaf image is submitted for disease assessment, both models independently generate predictions. Random Forest and Decision Tree might yield differing predictions due to their distinct decision rules.

The hybrid approach merges these individual predictions through a voting mechanism. For instance, if Random Forest predicts "Diseased" while Decision Tree predicts "Healthy," the hybrid model can weigh the predictions based on confidence scores or other criteria to deliver a definitive prediction.

The proposed hybrid approach for cotton disease detection, which combines Random Forest and Decision Tree models, offers several advantages and is designed to enhance the overall performance of the system like Robustness, Improved Accuracy, Overfitting Reduction, Interpretability, and Balancing Biases.

The hybrid approach is a powerful and well-rounded solution for cotton disease detection. It capitalizes on the strengths of both Random Forest and Decision Tree models, providing a more accurate, robust, and interpretable system. This is of great significance in the context of maintaining cotton crop health, boosting agricultural productivity, and safeguarding the livelihoods of cotton farmers.

D. ALGORITHM FOR HYBRID RANDOM FOREST AND DECISION TREE APPROACH

The following is overview of the hybrid approach for cotton disease detection using Random Forest and Decision Tree. Below is a summary of the steps involved:

Algorithm: **Hybrid Random Forest and Decision Tree Approach**

Input:

- Data directory containing images: data_dir
- List of classes: classes = ["diseased", "healthy"]

Output:

- Classification results for SVM, Multi-Class SVM, and Ensemble
- Bar graphs for individual and combined predictions

Step 1: Load and preprocess the dataset
    Initialize empty feature matrix (X) and labels (y)
    For each class in classes:
        Create class directory path
        For each image file in the class directory:
            Read image using OpenCV
            Convert image to grayscale
            Resize image to (128, 128)
            Flatten image and append to X
            Append label to y

Step 2: Divide the dataset into sets for testing and training.
    Use train_test_split to split X and y into training and testing sets

Step 3: Train individual models
    Train Random Forest classifier
    Train SVM classifier
    Train Multi-Class SVM classifier
    Train Decision Tree classifier

Step 4: Combine models into a VotingClassifier
    Create a VotingClassifier with Random Forest and Decision Tree

Step 5: Evaluate the models

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Evaluate Random Forest and obtain predictions and accuracy
Evaluate SVM and obtain predictions and accuracy
Evaluate Multi-Class SVM and obtain predictions and accuracy
Evaluate Decision Tree and obtain predictions and accuracy
Evaluate Ensemble model and obtain predictions and accuracy
Step 6: Export classification reports to CSV files
Create DataFrames for classification reports
Export DataFrames to CSV files
Step 7: Test a new image for Diseased or Healthy
Input image path
Read, grayscale, and resize the image
Flatten and reshape image for prediction
Predict using Random Forest, SVM, Multi-Class SVM, Decision Tree, and Ensemble
Display the test image, grayscale image, and resized image
Display bar graphs for individual and combined predictions

V. RESULT ANALYSIS

A. Dataset

The "Kaggle Cotton Disease" dataset is a compilation of images related to cotton plants and the diseases that affect them. Kaggle is a platform catering to data science and machine learning enthusiasts, which serves as a central hub for sharing datasets for various purposes, including competitions, research, and learning. This specific dataset appears to have been curated to accelerate the development of machine learning models designed to identify and categorize diseases in cotton plants. Same sample of Dataset as follows:

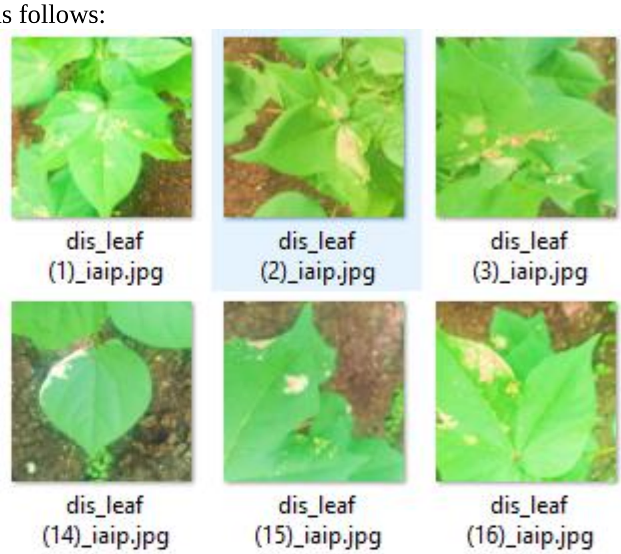


FIGURE 5. Cotton Leaf Diseased Samples [38]

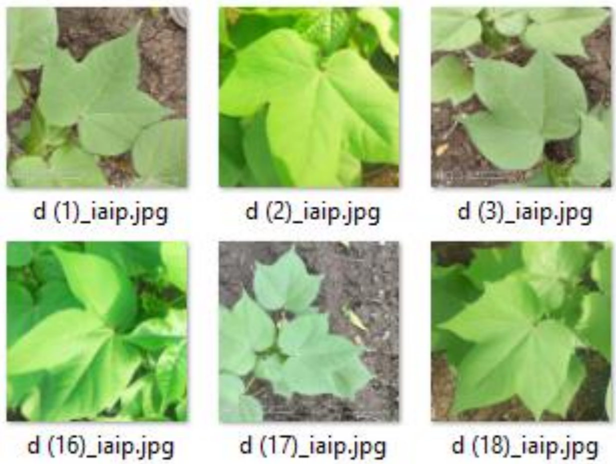


FIGURE 6. Cotton Leaf Healthy Samples [38]

FIGURE 7. Cotton Test Image [Implementation]

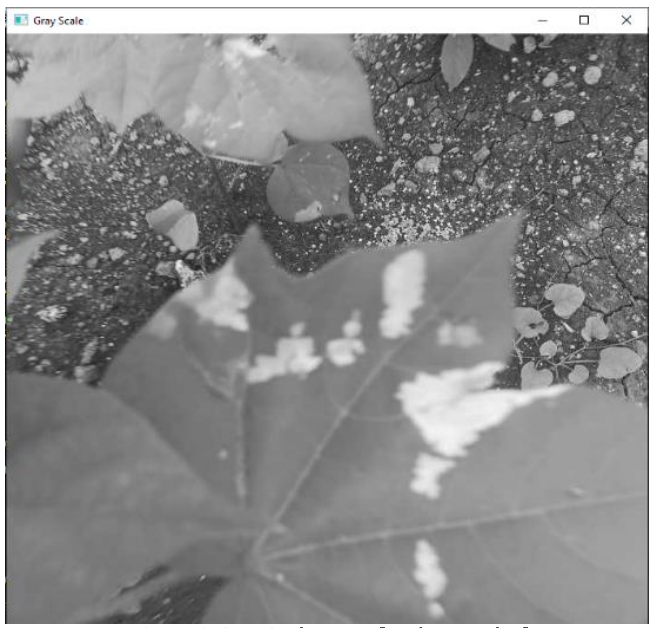


FIGURE 8. Gray Scale Image [Implementation]

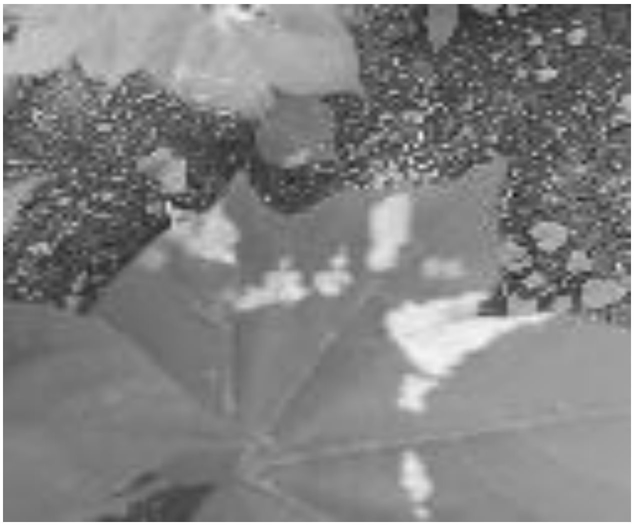


FIGURE 9. Resized Image [Implementation]

TABLE 2.
TEXT RESULT [TOTAL 500+ IMAGES TESTED]

Test Sample	SVM Accuracy	Multi-SVM Accuracy	Ensemble Proposed Accuracy
	81	81	99
	77	77	92
	77	77	92
	77	77	95

TABLE 3.
OUTPUT OF ALL CLASSIFIER

SVM REPORT	PRECISION	RECALL	F1 SCORE	SUPPORT
C0	0.72	0.48	0.57	61
C1	0.69	0.87	0.77	82
ACY			0.70	143
M AVG	0.71	0.67	0.67	143
W AVG	0.70	0.70	0.69	143
MULTI-CLASS SVM				
C0	0.72	0.48	0.57	61
C1	0.69	0.87	0.77	82
ACY			0.70	143
M AVG	0.71	0.67	0.67	143
W AVG	0.70	0.70	0.69	143
ENSEMBLE MODEL(Random Forest + Decision Tree Approach)				
C0	.95	.92	.94	61
C1	.94	.87	.95	82
ACY			0.945	143
M AVG	.945	.945	.945	143
W AVG	.945	.945	.945	143

ACY stands for accuracy, M AVG for macro average, and C0 for class zero and C1 for class one. We use Weighted Average, or W AVG.

Predictions from different models, namely SVM, Multi-Class SVM, and Ensemble Model, for the "Diseased" class are provided, along with associated class probability estimates that reflect each model's confidence in its prediction. Let's delve into each of these forecasts:

A. SVM Prediction:

Class Prediction: "Diseased" - The SVM model anticipates that the input corresponds to the "Diseased" category.

Probability (Diseased, Healthy): This represents the probability estimate ascribed by the SVM model to each class. In this instance, the SVM assigns a 81.0 % probability to the "Diseased" class and a 19% probability to the "Healthy" class, indicating its 81% confidence in the input being "Diseased" and 19% confidence in it being "Healthy."

B. Multi-Class SVM Prediction:

Class Prediction: "Diseased" - The Multi-Class SVM model likewise predicts that the input belongs to the "Diseased" category.

Probability (Diseased, Healthy): Similar to the SVM forecast, the Multi-Class SVM assigns probability estimates for each class. In this case, it allocates a 81% probability to the "Diseased" class and a 19% probability to the "Healthy" class, mirroring the SVM's prediction.

C. Ensemble Model Prediction (Random Forest + Decision Tree Approach):

Class Prediction: "Diseased" - The ensemble model, a combination of Random Forest and Decision Tree models, predicts that the input falls under the "Diseased" category.

Probability (Diseased, Healthy): In this scenario, the ensemble model assigns a higher probability of 99% to the "Diseased" class and a lower probability of 0.5% to the "Healthy" class. This indicates a high level of confidence in the ensemble model's prediction of "Diseased" (99%

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

confidence) and a lower confidence in it being "Healthy" (.05% confidence).

In all three predictions, the ultimate class forecast is "Diseased." Nonetheless, the confidence levels, as indicated by class probability estimates, exhibit variations:

The SVM and Multi-Class SVM models demonstrate comparable confidence levels in their forecasts, both showing moderate confidence in the "Diseased" class.

The ensemble model, a fusion of Random Forest and Decision Tree models, exhibits stronger confidence in its "Diseased" prediction, assigning a higher probability to this class.

These probability estimates hold significant value in comprehending the certainty or uncertainty of the models in their forecasts. This information can be pivotal in decision-making and guiding further actions based on these predictions.

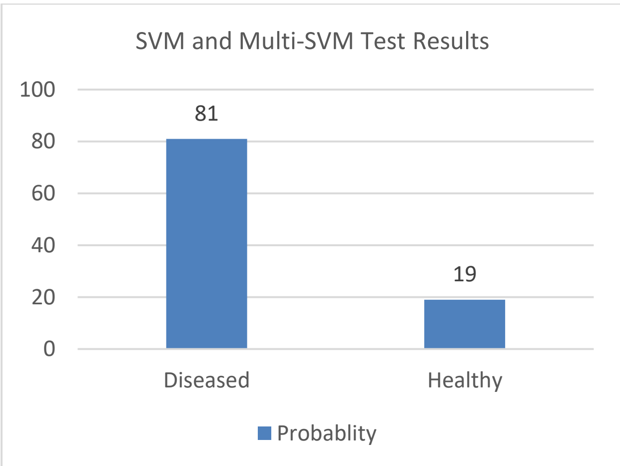


FIGURE 10. SVM and Multi-SVM Results

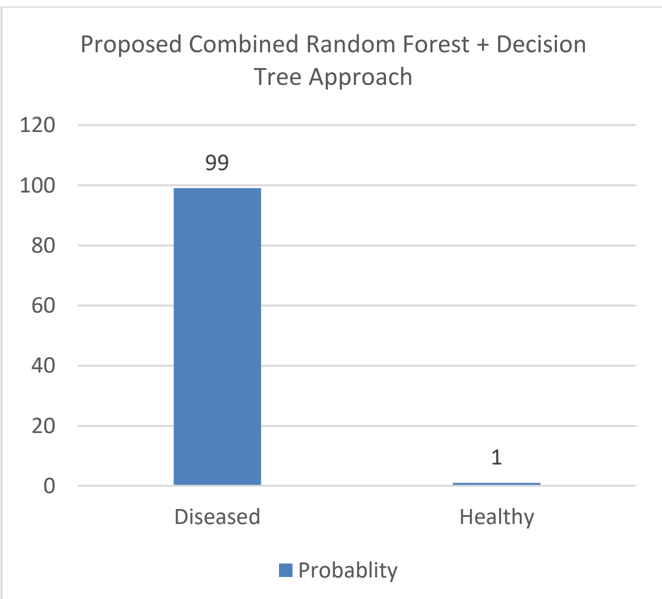


FIGURE 11. Proposed Model

TABLE 4.
RESULT AND APPROACH COMPARISON

	Approach Followed	Optimization Algorithms	Accuracy Achieved (%)
Base Paper (Zhang, N., et al. (2023)).	Models called Support Vector Machines (SVM) are used to forecast the severity of diseases.	Uses Support Vector Machine (SVM) models, a potent machine learning approach for classification tasks, to forecast the severity of an illness.	91.2
Proposed Approach	In the proposed approach, an ensemble or hybrid approach is employed, combining various algorithms to achieve superior results.	Employs an Ensemble or Hybrid Approach by combining various algorithms, including the Random Forest and Decision Tree approaches, to enhance performance and results.	94.5

As the accuracy is 94.5%, it means there are 5.5% errors. Therefore, the number of instances misclassified is

Number of test samples (N)=500

Error (E)=0.055

M = Misclassified

Number of instances misclassified = E * N (1)

Misclassified = 0.055×500=27.5(number of instances must be a whole number so 28)

The number of true positives is = N - M (2)

The number of true positives is = 500−28=472

$$\textbf{Macro Precision}= \frac{1}{N}\sum_{i=1}^{N}\frac{True\ Positive_i}{True\ Positive_i+False\ Positive_i} \quad (3)$$

$$\text{Macro Precision}= \frac{1}{N}\sum_{i=1}^{N}\frac{472}{472+28} = .944$$

$$\text{Macro Recall}= \frac{1}{N}\sum_{i=1}^{N}\frac{True\ Positive_i}{True\ Positive_i+False\ Negative_i} \quad (4)$$

$$\text{Macro Recall}= \frac{1}{N}\sum_{i=1}^{N}\frac{472}{472+0} = .944$$

$$\text{Macro F1-Score}= \frac{2X Macro\ Precision\ X\ Macro\ Recall}{Macro\ Precision+Macro\ Recall} \quad (5)$$

$$\text{Macro F1-Score}= \frac{2X.944X.944}{.944+.944} = .944$$

TABLE 5
FINAL RESULTS

	Accuracy	Macro Precision (%)	Macro Recall (%)	Macro F1-Score (%)
Base(Zhang, N., et al. (2023)).	91.2	92.02	91.2	91.16
Proposed	94.5	94.4	94.4	94.4

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

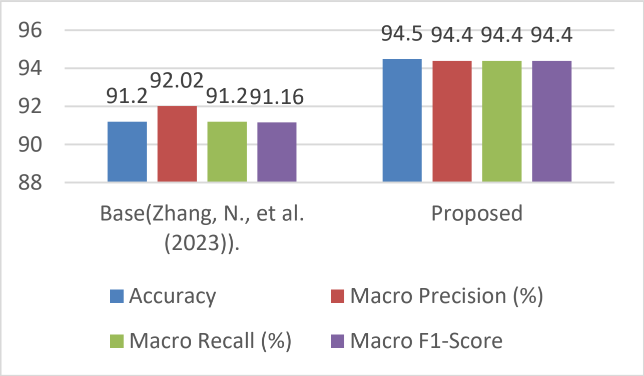


FIGURE 12. **Final Results**

VI. CONCLUSION

In this study, we investigated the predictive capabilities of three distinct models: Support Vector Machine (SVM), Multi-Class SVM, and an Ensemble Model that combines Random Forest and Decision Tree approaches. The specific task was binary classification, where we aimed to categorize inputs as either "Diseased" or "Healthy." We evaluated these models not only for their class predictions but also for their associated class probability estimates, which provide insights into the models' confidence in their predictions.

All three models unanimously forecasted the input as "Diseased." However, the critical difference lay in the confidence levels conveyed through their class probability estimates. This variation in confidence is crucial for understanding the reliability of these predictions and holds substantial value in practical applications.

The SVM and Multi-Class SVM models exhibited similar moderate confidence levels in their "Diseased" predictions, assigning a probability estimate of approximately 76.77%. This indicates that both models had a similar level of confidence in their predictions.

Concluding Considerations:

- The significance of the hybrid approach lies in its potential to enhance cotton disease detection through the integration of Random Forest and Decision Tree models. This fusion enables us to create a more robust and accurate system, vital for safeguarding crop health and agricultural productivity.

- Regarding practical deployment, while the hybrid approach exhibits promising capabilities, its real-world application necessitates meticulous consideration of factors like computational resources, scalability, and adaptability to diverse environmental conditions. Effective implementation will require collaboration with agriculture stakeholders, including farmers, researchers, and policymakers.

Limiting Circumstances:

- The effectiveness of any machine learning model heavily relies on the quality and representativeness of the training data. Challenges may arise due to the limited availability of diverse and high-quality datasets, particularly when dealing with rare or evolving diseases.

- Despite offering improved interpretability compared to complex deep learning models, the hybrid approach may still present limitations in understanding the precise decision-making process of the ensemble model. Striking a balance between interpretability and performance remains a challenge in developing practical agricultural systems.

Future Theoretical Implications:

- Future research endeavors could explore the integration of advanced techniques such as deep learning and transfer learning to further augment the performance of disease detection systems. These methodologies hold the promise of offering additional insights into the intricate patterns within image data and enhancing the generalization capabilities of the models.

- Moreover, investigating alternative fusion strategies for combining predictions from different models could yield further enhancements in accuracy and robustness. Techniques like stacking, meta-learning, or adaptive weighting methods present novel avenues for harnessing the complementary strengths of diverse algorithms.

Future Applied Implications:

- Scaling the hybrid approach for large-scale deployment across diverse agricultural settings is imperative as the demand for agricultural technology continues to rise. Overcoming challenges related to computational efficiency, model optimization, and integration with existing agricultural systems will be crucial in this endeavor.

- Integrating the disease detection system with IoT devices and mobile applications enables real-time monitoring of crop health and provides timely decision support to farmers. This proactive approach to disease management holds the potential to improve yield and mitigate economic losses significantly.

- Furthermore, assessing the socio-economic impact of implementing the disease detection system is vital. Understanding its implications for farmer livelihoods, resource allocation, and sustainability practices can inform broader agricultural policy and investment decisions.

In conclusion, while the hybrid approach including newer feature selection with deep learning and machine learning strategies [40]-[43] shows promise in advancing cotton disease detection, addressing limiting circumstances and exploring future theoretical and applied implications are paramount for realizing its full potential in practical agricultural settings.

Author Contributions:

Rahul Kumar: Conceptualization, Problem formulation, Methodology, Original draft preparation, Reviewing and Editing, and Final drafting.
Ashok Kumar: Data curation, Programming, Simulation, Validation, Numerical analysis, Visualization, and System setup.
Karamjit Bhatia: Programming, Writing, Validation.
Kottakkaran Sooppy Nisar: Mathematical Modelling, Visualization, and Investigation.
Siddharth Singh Chouhan: Simulation, Editing, Numerical analysis.
Anoop Kumar Tiwari: Supervision, Programming, Validation,Writing, Reviewing, and Editing.

ACKNOWLEDGMENT

This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444)

Conflict of Interest: There is no conflict of interest among the authors.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Data Availability: The data supporting this study's findings are available from the corresponding author upon reasonable request.

REFERENCES

[1] Mula Alyas, R. M. (2023). Detection Of Plant Diseases Using Image Processing With Machine Learning (Doctoral Dissertation).
[2] Ahmed, I., Habib, G., & Yadav, P. K. (2023, March). An Approach to Identify and Classify Agricultural Crop Diseases Using Machine Learning and Deep Learning Techniques. In 2023 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 1-6). IEEE.
[3] Bondre, S., & Patil, D. (2023). Recent advances in agricultural disease image recognition technologies: A review. Concurrency and Computation: Practice and Experience, 35(9), e7644.
[4] Kumar Sahu, S., & Pandey, M. (2023). An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model.
[5] Patel, B., & Sharaff, A. (2023). A review on rice plant phenotyping traits estimation for disease and growth management using modern ML techniques. Multimedia Tools and Applications, 1-23.
[6] Bhimte, N. R., & Thool, V. R. (2018, June). Diseases detection of cotton leaf spot using image processing and SVM classifier. In 2018 Second international conference on intelligent computing and control systems (ICICCS) (pp. 340-344). IEEE.
[7] Kothari, J. D. (2018). Plant disease identification using artificial intelligence: machine learning approach. Jubin Dipakkumar Kothari (2018). Plant Disease Identification using Artificial Intelligence: Machine Learning Approach. International Journal of Innovative Research in Computer and Communication Engineering, 7(11), 11082-11085.
[8] Ahmed, K., Shahidi, T. R., Alam, S. M. I., & Momen, S. (2019, December). Rice leaf disease detection using machine learning techniques. In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-5). IEEE.
[9] Fenu, G., & Malloci, F. M. (2019, November). An application of machine learning technique in forecasting crop disease. In Proceedings of the 3rd International Conference on Big Data Research (pp. 76-82).
[10] Akulwar, P. (2020). A recommended system for crop disease detection and yield prediction using machine learning approach. Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries, 141-163.
[11] Reddy, P. C., Chandra, R. M. S., Vadiraj, P., Reddy, M. A., Mahesh, T. R., & Madhuri, G. S. (2021, December). Detection of plant leaf-based diseases using machine learning approach. In 2021 IEEE International conference on computation system and information technology for sustainable solutions (CSITSS) (pp. 1-4). IEEE.
[12] Trivedi, V. K., Shukla, P. K., & Pandey, A. (2021, June). Hue based plant leaves disease detection and classification using machine learning approach. In 2021 10th IEEE international conference on communication systems and network technologies (CSNT) (pp. 549-554). IEEE.
[13] Karthika, J., Santhose, M., & Sharan, T. (2021, May). Disease detection in cotton leaf spot using image processing. In Journal of Physics: Conference Series (Vol. 1916, No. 1, p. 012224). IOP Publishing.
[14] Prasad, A., Mehta, N., Horak, M., & Bae, W. D. (2022). A Two-Step Machine Learning Approach for Crop Disease Detection Using GAN and UAV Technology. Remote Sensing, 14(19), 4765.
[15] Jayaramu, H. K., Ramesh, D., & Jain, S. (2022). ESMO-based Plant Leaf Disease Identification: A Machine Learning Approach. In Computer Vision and Machine Learning in Agriculture, Volume 2 (pp. 155-170). Singapore: Springer Singapore.
[16] Radhika, V., Ramya, R., & Abhishek, R. (2023, March). Machine learning approach-based plant disease detection and pest detection system. In International Conference on Communications and Cyber Physical Engineering 2018 (pp. 191-200). Singapore: Springer Nature Singapore.
[17] Singh, P., Singh, P., Farooq, U., Khurana, S. S., Verma, J. K., & Kumar, M. (2023). CottonLeafNet: cotton plant leaf disease detection using deep neural networks. Multimedia Tools and Applications, 1-26.
[18] Gülmez, B. (2023). A novel deep learning model with the Grey Wolf Optimization algorithm for cotton disease detection. Journal of Universal Computer Science, 29(6), 595.
[19] Zhang, N., Zhang, X., Shang, P., Ma, R., Yuan, X., Li, L., & Bai, T. (2023). Detection of Cotton Verticillium Wilt Disease Severity Based on Hyperspectrum and GWO-SVM. Remote Sensing, 15(13), 3373.
[20] Zhang, N., Zhang, X., Shang, P., Ma, R., Yuan, X., Li, L., & Bai, T. (2023). Detection of Cotton Verticillium Wilt Disease Severity Based on Hyperspectrum and GWO-SVM. Remote Sensing, 15(13), 3373.
[21] Chepuri, S., & Ramadevi, Y. (2023, October). A Novel Fusion Study on Disease Detection in Cotton Plants Using Embedded Approaches of Neural Networks. In International Conference on Computer & Communication Technologies (pp. 171-181). Singapore: Springer Nature Singapore.
[22] Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based plant diseases detection using deep learning. Materials Today: Proceedings, 80, 3500-3506.
[23] Chug, A., Bhatia, A., Singh, A. P., & Singh, D. (2023). A novel framework for image-based plant disease detection using hybrid deep learning approach. Soft Computing, 27(18), 13613-13638.
[24] Mahum, R., Munir, H., Mughal, Z. U. N., Awais, M., Sher Khan, F., Saqlain, M., & Tlili, I. (2023). A novel framework for potato leaf disease detection using an efficient deep learning model. Human and

Ecological Risk Assessment: An International Journal, 29(2), 303-326.

[25] Thakur, P. S., Sheorey, T., & Ojha, A. (2023). VGG-ICNN: A Lightweight CNN model for crop disease identification. Multimedia Tools and Applications, 82(1), 497-520.

[26] Srinivas, L. N. B., AM Viswa Bharathy, Sravanth Kumar Ramakuri, Abhisek Sethy, and Ravi Kumar. "An optimized machine learning framework for crop disease detection." Multimedia Tools and Applications (2023): 1-20.

[27] Zhang, N., Zhang, X., Shang, P., Ma, R., Yuan, X., Li, L., & Bai, T. (2023). Detection of Cotton Verticillium Wilt Disease Severity Based on Hyperspectrum and GWO-SVM. Remote Sensing, 15(13), 3373.

[28] Kaur, S. (2023, July). Disease Recognition of Leaf Disease in Cotton Using KNN, SVM, and CNN Pooling Method. In 2023 2nd International Conference on Edge Computing and Applications (ICECAA) (pp. 1151-1156). IEEE.

[29] Sahu, S. K., & Pandey, M. (2023). An optimal hybrid multiclass SVM for plant leaf disease detection using spatial Fuzzy C-Means model. Expert Systems with Applications, 214, 118989.

[30] Islam, M. M., Talukder, M. A., Sarker, M. R. A., Uddin, M. A., Akhter, A., Sharmin, S., ... & Debnath, S. K. (2023). A deep learning model for cotton disease prediction using fine-tuning with smart web application in agriculture. Intelligent Systems with Applications, 20, 200278.

[31] Sarkar, C., Gupta, D., Gupta, U., & Hazarika, B. B. (2023). Leaf disease detection using machine learning and deep learning: Review and challenges. Applied Soft Computing, 110534.

[32] Odukoya, O. H., Aina, S., & Dégbéssé, F. W. (2023). A Model for Detecting Fungal Diseases in Cotton Cultivation using Segmentation and Machine Learning Approaches. International Journal of Advanced Computer Science and Applications, 14(1).

[33] Chavan, A., Thakare, A., Chopade, T., Fernandes, J., & Gawari, O. (2023). Crop Disease Detection Accelerated by GPU. Artificial Intelligence Applications and Reconfigurable Architectures, 151-166.

[34] Gülmez, B. (2023). A novel deep learning model with the Grey Wolf Optimization algorithm for cotton disease detection. Journal of Universal Computer Science, 29(6), 595.

[35] Sriram, G., Vignesh, M., Saran, D., & Ayyasamy, S. (2023). Flourishing Fields: Revolutionizing Agriculture with Soft Computing-based Plant Disease Detection.

[36] Goel, L., & Nagpal, J. (2023). A Systematic review of recent machine learning techniques for plant disease identification and classification. IETE Technical Review, 40(3), 423-439.

[37] Cotton leaf curl disease. (2021). Gov.au. https://www.dpi.nsw.gov.au/biosecurity/plant/insect-pests-and-plant-diseases/cotton-leaf-curl

[38] Ahumada, D. (n.d.). Bacterial blight of cotton. Ncsu.edu. https://content.ces.ncsu.edu/bacterial-blight-angular-leaf-spot-of-cotton

[39] Contacts, M. (2016). Increased incidence of Alternaria leaf spot of cotton. Agrilife.org. https://agrilife.org/texasrowcrops/2016/09/06/increased-incidence-of-alternaria-leaf-spot-of-cotton

[40] Dhamodharan, R. (2023). Cotton plant disease [Data set].

[41] S. Solanki, U. P. Singh and S. S. Chouhan, "Brain Tumor Classification Using ML and DL Approaches," *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*, Hamburg, Germany, 2023, pp. 204-208, doi: 10.1109/ICCCMLA58983.2023.10346854.

[42] A. Saxena, S. S. Chouhan, R. M. Aziz, et al., "A comprehensive evaluation of Marine predator chaotic algorithm for feature selection of COVID-19, *Evolving Systems*, *2024*, doi: 10.1007/s12530-023-09557-2

[43] S. Solanki, U. P. Singh, S. S. Chouhan, et al., "A systematic analysis of magnetic resonance images and deep learning methods used for diagnosis of brain tumor," *Multimed Tools Appl*, vol. 83, pp. 23929–23966, 2024, doi: 10.1007/s11042-023-16430-6.

[44] S. Solanki, U. P. Singh, S. S. Chouhan, and S. Jain, "Brain Tumour Detection and Classification by using Deep Learning Classifier", *Int J Intell Syst Appl Eng*, vol. 11, no. 2s, pp. 279-, Feb. 2023.

Dr. Rahul Kumar received his B.Tech Degree in Computer Science and Engineering from ICFAI University Dehradun, India in 2012 and M.Tech Degree in Computer Science and Engineering from Suresh Gyan Vihar University, Jaipur, India in 2014. He received his Ph.D. degree in Computer Science and Engineering from Dr. K.N Modi University, Jaipur, India in 2022.His Current interest is Machine learning (Feature Selection and Extraction),Database management System and Computer Network. Currently working as an Associate Professor in Department of Computer Science & Engineering, Amity School of Engineering and Technology, Amity University, Gwalior, Madhya Pradesh, India

Dr. Ashok Kumar received the Master's degree in Computer Applications from Dr. Bhim Rao Ambedkar University Agra, India in 2002 and Ph. D. in Computer Science from Gurukula Kangri (Deemed to be University) Haridwar, India in 2019. He is an Associate Professor at the Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India. He has more than 20 years teaching experience. He has authored more than 30 research papers in the International/ national Journals and Conferences. His research interests include Digital Image Processing and Machine Learning

Dr. Karamjit Bhatia received the M.Phil. (Computer Applications) in 1988 from University of Roorkee, Roorkee (now IIT Roorkee), India and Ph. D. (Computer Science) in 2001 from Gurukula Kangri Vishwavidyalaya, Haridwar, India. Currently, he has teaching and research experience of aprox. 35 years and presently he is working as a Professor at Department of Computer Science, Gurukula Kangri (Deemed to be University), Haridwar, India. His research interests include Distributed Systems, Image Processing, Mobile Ad-hoc Networks and Machine Learning. He has published around 60 research papers in International Journals of repute and National/ International Conference proceedings. Nine research scholars have been awarded Ph. D. degree under his supervision and one research scholar is working under his supervision.

Prof. Dr. Kottakkaran Sooppy Nisar is a full professor of applied mathematics at Prince Sattam bin Abdulaziz University, Saudi Arabia. His current research interests are special functions, fractional calculus, fluid dynamics, biomathematics, artificial intelligence, machine learning, SAC-OCDMA, and multidisciplinary applications of mathematics. He is an editorial board member of many reputed journals, has published more than 800 research papers in Scopus/WoS-indexed journals, and has registered a patent with the Intellectual Property of India, Government of India. He is the winner of a number of awards such as, the best researcher of the university 2023, the Obada Prize 2021, Young Distinguished Researcher Award, the Jaipur Mathematical Society 2021, PSAU Research Excellence Awards (for the best researcher at the university) in 2019 and 2020, the Riemann Young Researcher Award 2020, and the Indo-Asian-Niels Henrik Distinguished Innovative Scientist Award 2020. He is also on the list of the top 2% of scientists in the world (all fields), according to a survey conducted by Elsevier BV and Stanford University, based on the citations, h-index, and co-authorship-adjusted h-index in the last three consecutive years. Prof. Nisar completed more than 10 projects at different national universities. He has more than 15000 citations with H-index 50.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3419906

Dr. Siddharth Singh Chouhan currently working as an assistant professor in School of Computing Science and Engineering, VIT Bhopal University, Kothri-Kalan, Sehore, 466114, Madhya Pradesh, India. He has completed PhD in Computer Science and Engineering from Shri Mata Vaishno Devi University Katra, 182320, UT of Jammu and Kashmir, India. He is a Post Doctorate from the University of Malta, Europe. He has received a BTech degree in 2010 and an MTech degree in 2013 in Computer Science and Engineering from RGPV University, Bhopal, India. His area of research interest includes soft computing, computer vision, drone technology, deep learning, and precision agriculture. He has authored several research papers in many published in conferences and journals of international repute.

Priti Maratha is working as an Assistant Professor in the Department of Computer Science and Information Technology, Central University of Haryana, Mahendergarh, Haryana, India. She has completed her PhD from National Institute of Technology, Kurukshetra, India. She graduated from Kurukshetra University and done Masters in Computer Applications (MCA) from Panjab University Chandigarh, India. She has qualified UGC-NET JRF exam as well. Her research interests include Machine Learning, Internet of Things, Optimization, Wireless Sensor Networks etc. She has published and reviewed research papers in various journals of International repute.

Dr. Anoop Kumar Tiwari is an Assistant Professor at Department of Computer Science and Information Technology, Central University of Haryana, Mahendergarh, Haryana, India. He holds the Ph.D. degree in computer science from the Department of Computer Science, Institute of Science, Banaras Hindu University, India. His research interests are in computational intelligence, machine learning, and bioinformatics.