IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Fusion Of Deep Learning And Machine Learning Algorithms For Accurate Cotton Crop Disease Identification

[1] Umang Jain, [2] Dr. J.R. Arun Kumar

[1]M. Tech. Scholar, [2]Professor

[1,2]Department of Computer Science & Engineering. Modern Institute of Technology and Research Centre.

Abstract : Cotton crop diseases pose a significant threat to agricultural productivity, necessitating accurate and efficient disease identification methods. This study proposes a hybrid model integrating Convolutional Neural Networks (CNNs) for feature extraction with Support Vector Machines (SVMs) and Random Forests (RFs) for disease classification to enhance the accuracy of cotton crop disease detection. The proposed approach leverages the powerful feature-learning capabilities of CNNs while utilizing SVMs and RFs for robust classification. The model's performance is evaluated against individual classifiers (CNNs, SVMs, and RFs) using a diverse dataset of cotton crop images, ensuring a comprehensive assessment across different disease symptoms and crop varieties. Furthermore, the study examines the generalization ability of the hybrid model on unseen data to validate its real-world applicability. Additionally, computational efficiency and scalability are analyzed to determine the model's practicality for large-scale agricultural deployment. The results demonstrate that the hybrid CNN-SVM-RF model outperforms standalone approaches in terms of accuracy and robustness, making it a promising solution for precision agriculture and early disease detection in cotton crops.

Index Terms: Cotton crop disease, Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), Random Forests (RFs), Hybrid model, Feature extraction, Disease classification, Precision agriculture, Machine learning, Computational efficiency.

I. Introduction

Cotton is one of the most important cash crops worldwide, playing a crucial role in the agricultural economy. However, cotton production is highly susceptible to various diseases caused by fungi, bacteria, and viruses, which significantly impact yield and quality [1].

Fig:- Bacterial Blight

Fig:- Veinal Necrosis

Fig:- Angular Leaf Spot

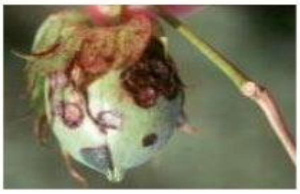

Fig:- Boll Rot

Figure 1. Diseases in Cotton Plants

Early and accurate detection of these diseases is essential to minimize losses and ensure sustainable cotton farming. Traditional disease detection methods rely on manual inspection by farmers and agricultural experts, which is time-consuming, subjective, and often inaccurate. With the advancement of technology, machine learning and deep learning techniques have emerged as promising solutions for automated disease identification in crops, enabling timely interventions and improved decision-making in agriculture [2].

Convolutional Neural Networks (CNNs) [3] have shown remarkable success in image-based disease detection due to their ability to automatically extract hierarchical features from crop images. However, despite their powerful feature extraction capabilities, CNNs may face challenges related to overfitting, computational complexity, and interpretability when applied directly to classification tasks. On the other hand, traditional machine learning classifiers such as Support Vector Machines (SVMs) and Random Forests (RFs) have demonstrated robustness in handling high-dimensional data while offering better generalization capabilities. Combining these approaches can potentially leverage the strengths of deep learning and classical machine learning to develop a more accurate and efficient model for cotton crop disease identification [4].

In this study, a hybrid approach integrating CNNs with SVMs and RFs is proposed to improve cotton crop disease classification. The CNN component is responsible for extracting deep, discriminative features from cotton crop images, while the SVM and RF classifiers ensure effective disease categorization. The proposed hybrid model is evaluated against standalone models, including CNNs, SVMs, and RFs, to determine its effectiveness in real-world agricultural applications. Additionally, the model's performance is tested on diverse datasets, including various cotton crop varieties and disease symptoms, to assess its robustness and generalizability [5].

Beyond accuracy, this research also focuses on the computational efficiency and scalability of the hybrid model. Given the increasing adoption of smart farming and precision agriculture, it is crucial to ensure that the model can be deployed on resource-constrained devices and large-scale farming environments. The study explores different optimization techniques to enhance the model’s performance while maintaining minimal computational overhead. By addressing these challenges, the proposed approach aims to contribute to the development of intelligent, automated systems for cotton disease detection, ultimately helping farmers make informed decisions and improve crop management strategies [6].

A. Objectives of Research

- Develop a hybrid model combining CNNs for feature extraction with SVMs and RFs for accurate cotton crop disease classification.
- Compare the performance of the hybrid model with individual models (CNNs, SVMs, and RFs) using a diverse dataset of cotton crop images.
- Evaluate the model's ability to classify diseases in different cotton crop varieties and unseen data.

- Improve feature extraction techniques to enhance classification accuracy and efficiency.
- Analyze the computational efficiency and scalability of the hybrid model for real-world agricultural applications.
- Contribute to precision agriculture by providing an automated, intelligent system for early disease detection in cotton crops.

II. Literature Survey

V. Saxena et al. (2024) [7] applied the Nu-SVM algorithm for rice leaf disease classification, achieving moderate accuracy (52.12%–53.81%) using Sobel edge detection and Hu Moments but highlighting the need for improved image processing techniques.

R. Gao et al. (2024) [8] developed a deep learning-based cotton pest detection model integrating Transformers and knowledge graphs, achieving 94% accuracy and outperforming YOLOv8 and RetinaNet in both precision and speed.

S. Jayanthy et al. (2024) [9] utilized MobileNetV2 and drone-based real-time monitoring for early cotton leaf disease detection, demonstrating superior accuracy and efficiency with GPS tracking for infected plants.

M. SithaRam et al. (2024) [10] used VGG16 and VGG19 on a hybrid dataset for cotton leaf disease detection, achieving high accuracies of 94% and 95%, aiding in crop health improvement.

N. Parashar and P. Johri (2024) [11] optimized MobileNetV2 for cotton disease detection, achieving 99.91% accuracy with low computational requirements, making it suitable for mobile applications.

R. Setiawan et al. (2023) [12] also explored Nu-SVM for rice leaf disease classification, obtaining similar accuracy results (52.12%–53.81%) and emphasizing the need for advanced image processing and feature extraction.

Table 1. Literature Review Findings

Author Name (Year)	**Main Concept**	**Findings**	**Limitations**
V. Saxena et al. (2024) [7]	Nu-SVM for rice leaf disease classification	Achieved 52.12%–53.81% accuracy using Sobel edge detection and Hu Moments.	Moderate accuracy, requires advanced image processing and feature extraction.
R. Gao et al. (2024) [8]	Deep learning with Transformers for cotton pest detection	Achieved 94% accuracy, 0.95 mAP, and 49.7 FPS, outperforming YOLOv8 and RetinaNet.	Computational requirements for Transformer integration.
S. Jayanthy et al. (2024) [9]	MobileNetV2 and drone-based real-time monitoring for cotton leaf disease detection	High accuracy, efficient real-time detection, and GPS tracking for infected plants.	Dependence on drone technology and real-time data availability.
M. SithaRam et al. (2024) [10]	VGG16 and VGG19 for cotton leaf disease detection	Achieved 94% (VGG16) and 95% (VGG19) accuracy using a hybrid dataset.	Limited generalizability beyond dataset scope.
N. Parashar & P. Johri (2024) [11]	Optimized MobileNetV2 for cotton disease detection	Achieved 99.91% accuracy with low computational requirements, mobile-compatible.	Model optimization may require extensive tuning for different conditions.

R. Setiawan et al. (2023) [12]	Nu-SVM for rice leaf disease classification	Similar accuracy (52.12%–53.81%) with Sobel edge detection and Hu Moments.	Low accuracy, highlights need for improved image processing and feature extraction.

Research Gap Discussion

- **Limited Accuracy in SVM-based Models** – Studies using Nu-SVM for rice leaf disease classification showed low accuracy (~52%), indicating the need for better feature extraction and classification techniques.
- **Need for Advanced Image Processing** – Many models rely on basic feature extraction methods (e.g., Sobel edge detection, Hu Moments), which may not be sufficient for complex disease identification.
- **Computational Challenges in Deep Learning** – Transformer-based models and MobileNetV2 require high computational resources, making them difficult to deploy on low-power devices.
- **Dependency on Real-time Data & Drones** – Some studies rely on drone-based real-time monitoring, which may not be practical for all farming regions due to cost and infrastructure limitations.
- **Limited Dataset Generalization** – Most studies used specific datasets (Kaggle, real-time images) without testing across diverse environmental conditions, limiting real-world applicability.
- **Need for More Efficient Mobile Solutions** – While some models optimized for mobile use (e.g., MobileNetV2), further improvements are needed to balance accuracy and computational efficiency.
- **Integration of Multi-Modal Approaches** – Current research primarily focuses on CNNs and Transformers; integrating other AI techniques, such as hybrid models or explainable AI, could improve results.
- **Lack of Standardized Evaluation Metrics** – Studies report varying performance metrics (accuracy, mAP, FPS), making direct comparisons difficult. A standardized evaluation framework is needed.

III. METHODOLOGY

1. Data Preprocessing

- **Image Resizing**: Resize images to a fixed size (e.g., 224×224 pixels) for consistency.
- **Normalization**: Scale pixel values between 0 and 1 to improve model performance.
- **Data Augmentation**: Apply transformations like rotation, flipping, and scaling to increase dataset variety and prevent overfitting.

2. Hybrid Model Design (CNN + SVM / RF)

- **Feature Extraction with CNN**:
  - **Convolutional Layer**: Extracts important features from images.
  - **Pooling Layer**: Reduces the size of feature maps to make processing efficient.
- **Classification**:
  - **SVM (Support Vector Machine)**: Finds the best boundary (hyperplane) to separate different classes.
  - **Random Forest (RF)**: Uses multiple decision trees and predicts based on majority voting.

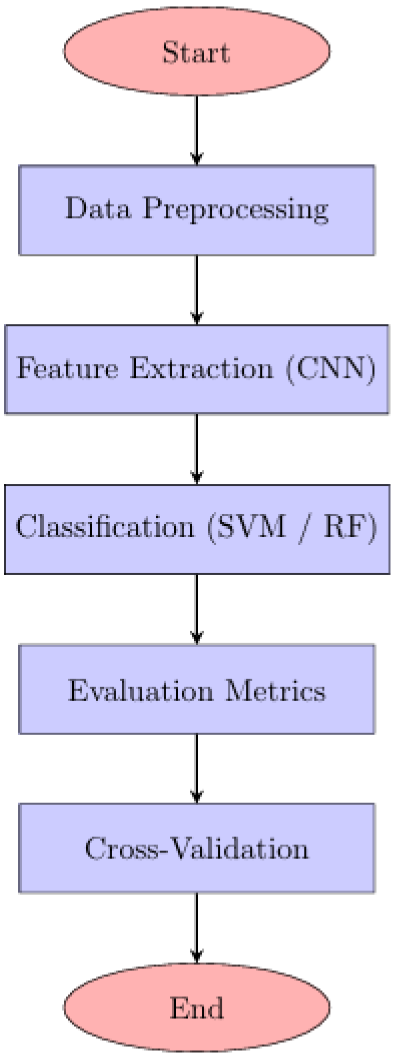

Figure 2. Research Flowchart

3. Evaluation Metrics

- **Accuracy**: Measures overall correctness of the model.
- **Precision**: Percentage of correctly predicted positive cases.
- **Recall**: Ability to detect actual positive cases.
- **F1-Score**: Balances precision and recall for better evaluation.

4. Cross-Validation

- The dataset is split into multiple parts (k-folds).
- The model is trained multiple times, each time using a different part for validation.
- The average result ensures the model performs well on new data.

By combining CNNs for feature extraction with SVM and RF for classification, this hybrid model aims to achieve high accuracy. The use of proper preprocessing, evaluation metrics, and cross-validation ensures the model is reliable and effective.

IV. Data Analysis Results

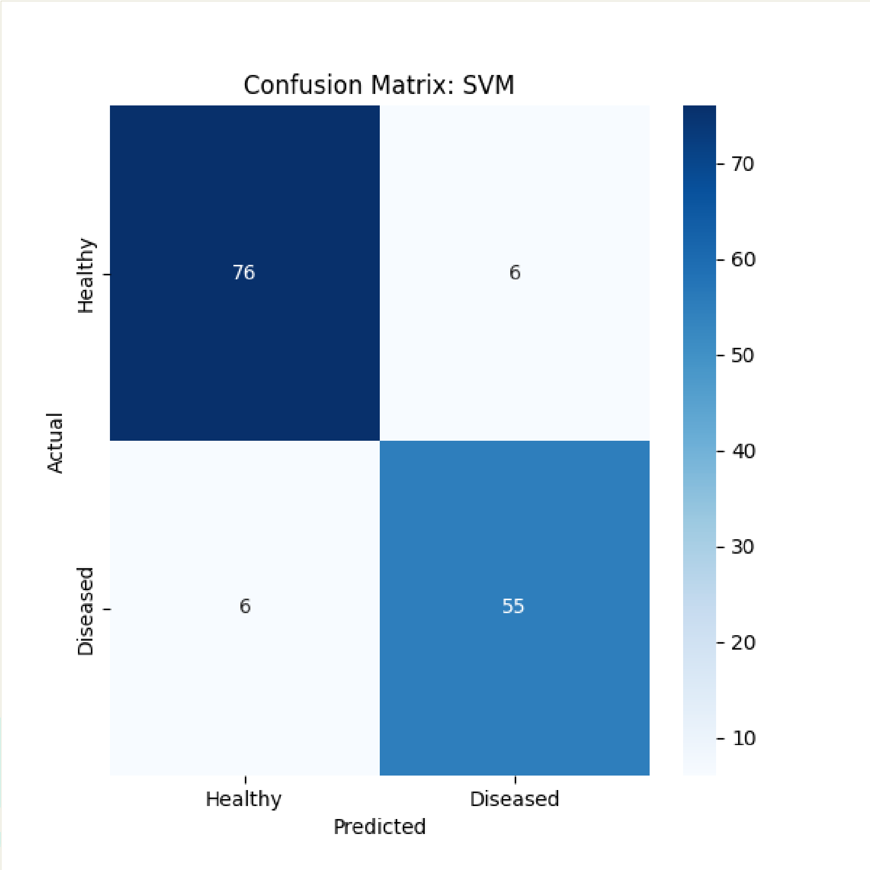

Figure 3. Confusion Matrix -SVM

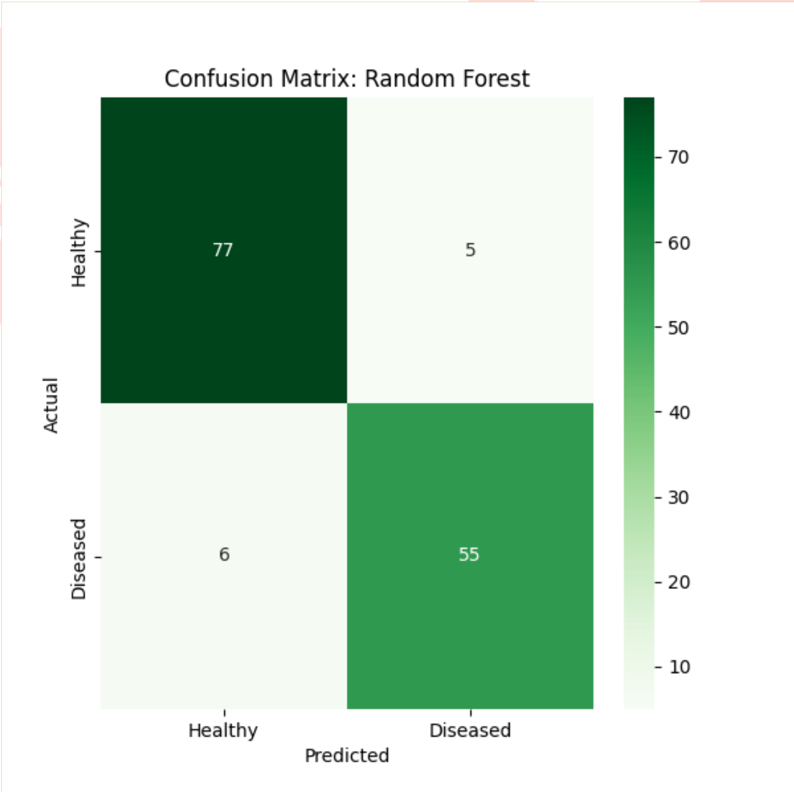

Figure 4. Confusion Matrix -Random Forest

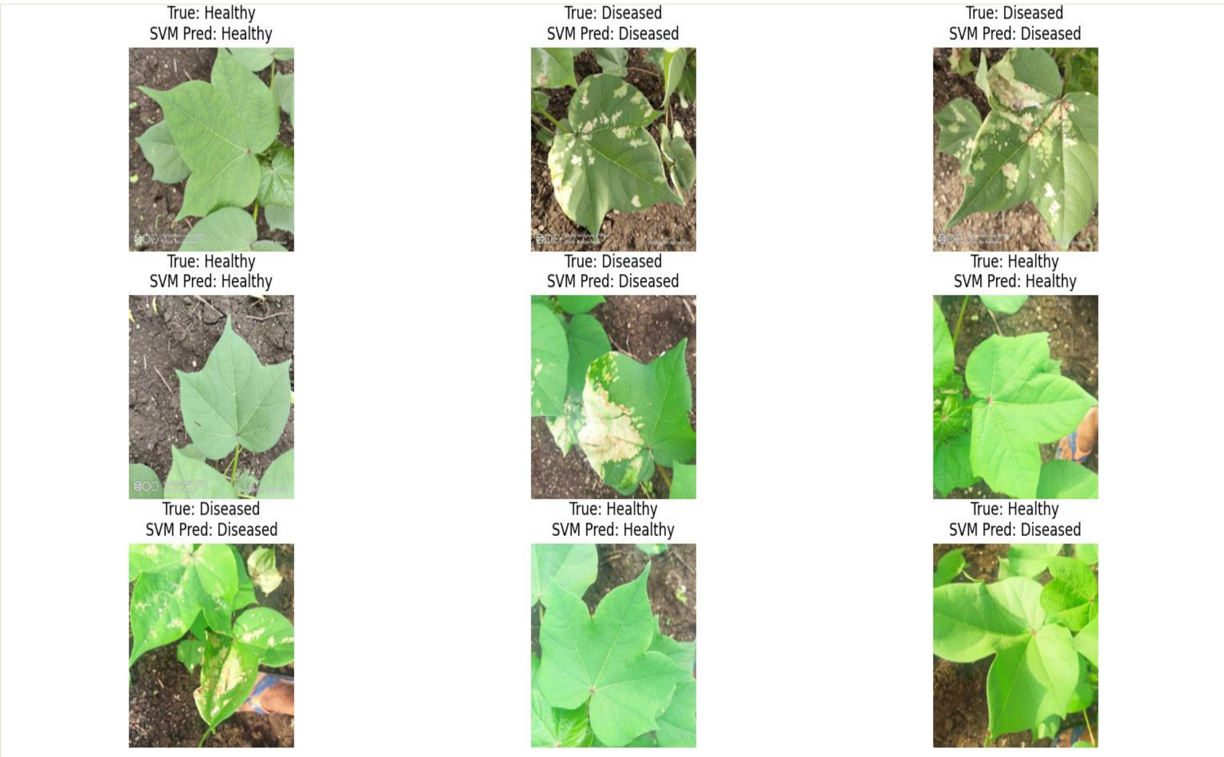

Figure 5. CNN-SVM Results

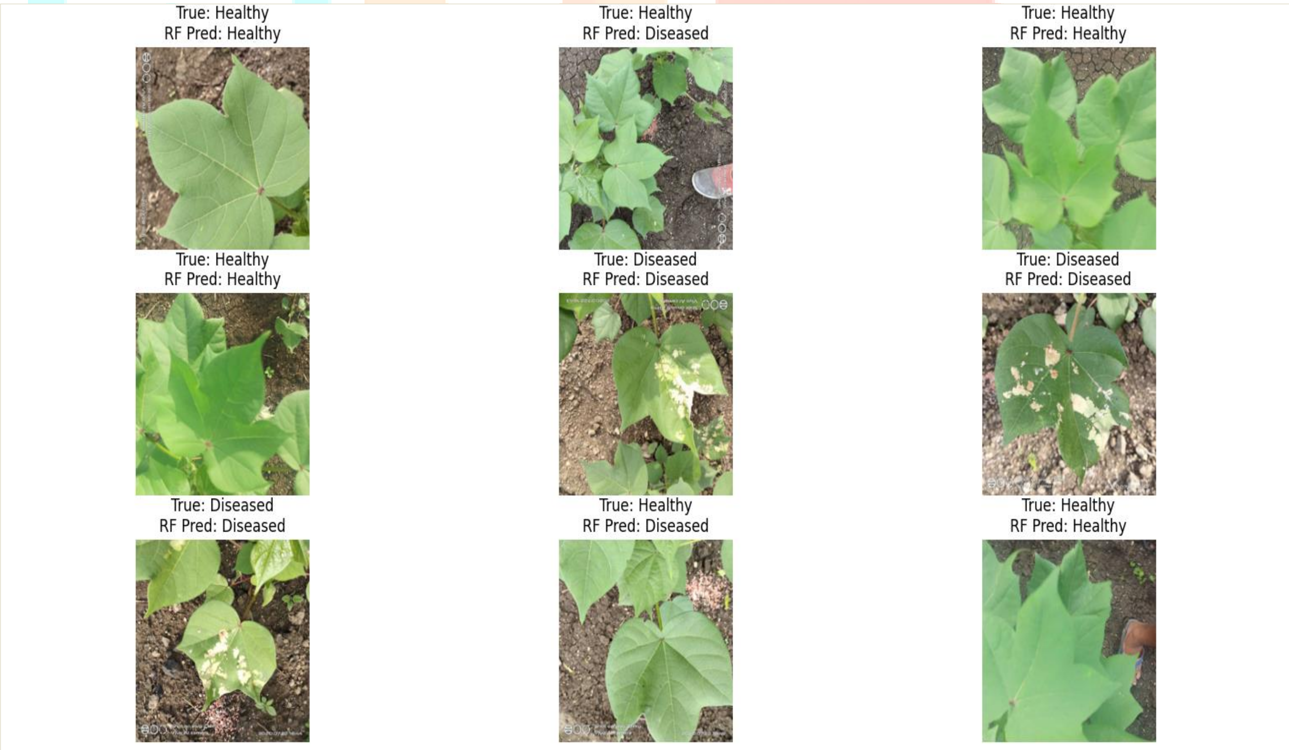

Figure 6. CNN-Random Forest Results

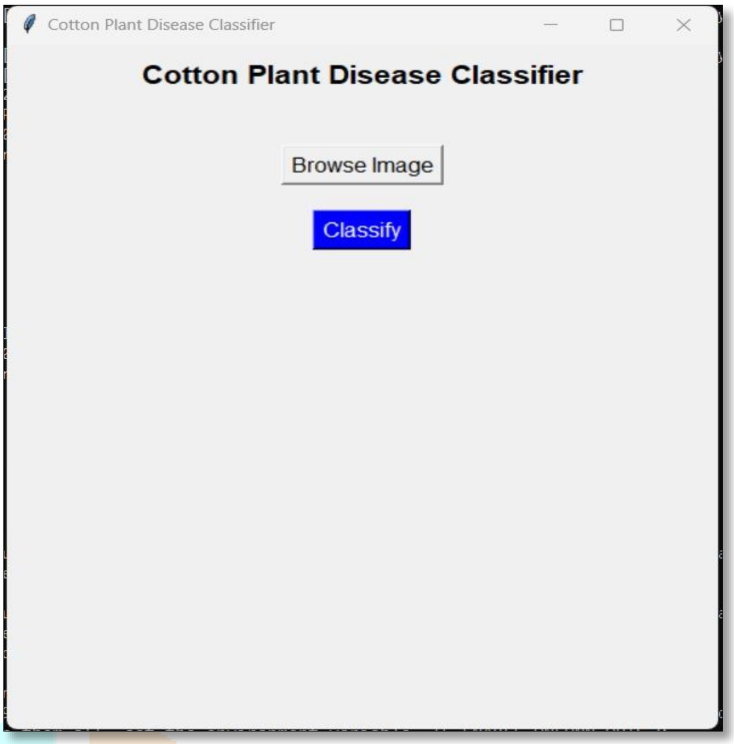

Figure 7. Main Screen

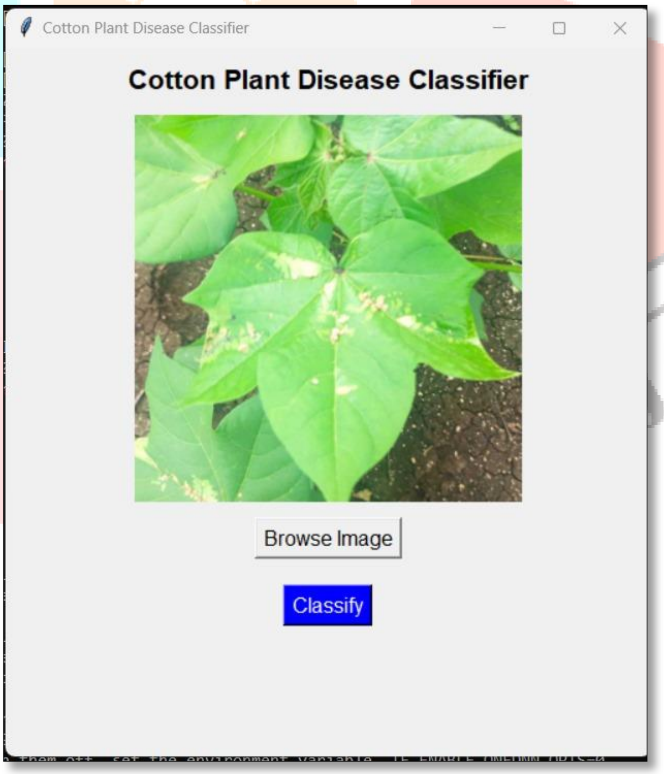

Figure 8. Image Selection

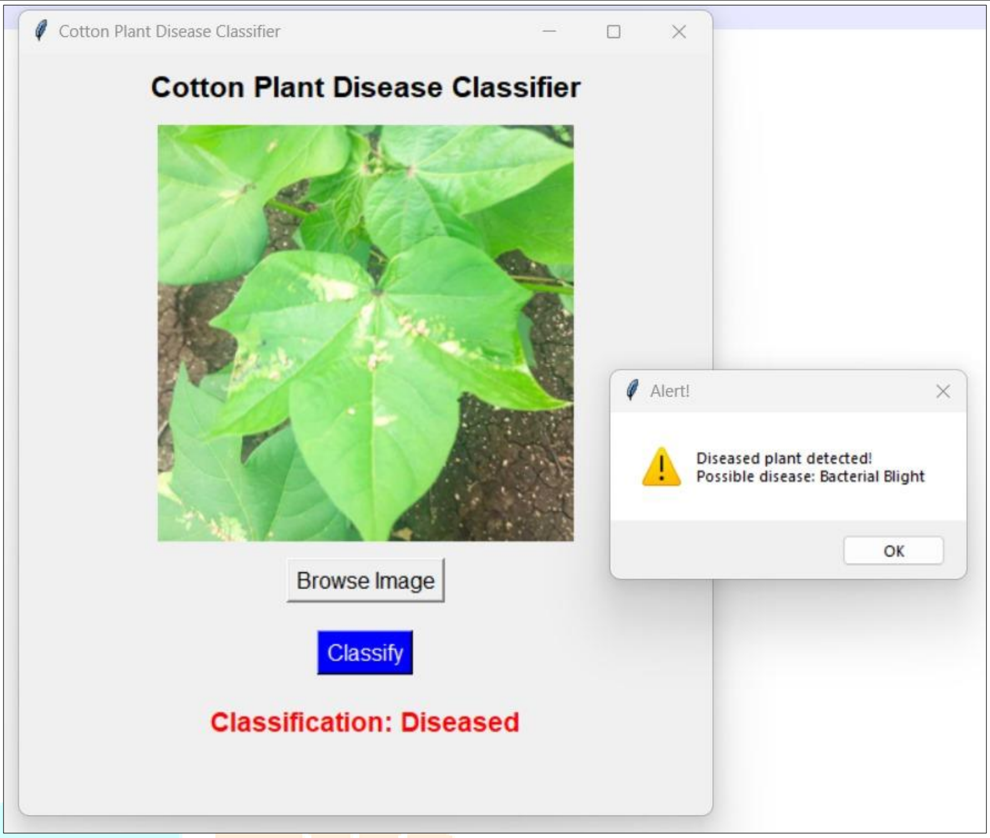

Figure 9. Disease Classification

Table 1. Result Analysis

Classifier	Class	Precision	Recall	F1-Score	Support
CNN+SVM+RF	0	0.91	0.98	0.94	82
	1	0.96	0.87	0.91	61
CNN+SVM+RF Avg		**0.93**	**0.92**	**0.93**	143

V. CONCLUSION

The CNN+SVM+RF hybrid model demonstrates strong performance in classifying the given dataset. The model achieves an average precision of 0.93, recall of 0.92, and an F1-score of 0.93, indicating a well-balanced classification ability. For Class 0, the model exhibits high recall (0.98), meaning it effectively identifies most of the actual positive instances, while its precision (0.91) ensures minimal false positives. Conversely, for Class 1, the model maintains a high precision of 0.96, signifying strong predictive accuracy, but with a slightly lower recall of 0.87, suggesting some missed instances.

The overall classification performance confirms that the hybrid approach combining CNN for feature extraction with SVM and Random Forest for classification is effective and reliable. The strong F1-score across classes highlights the model's robustness and ability to maintain a balance between precision and recall. Additionally, the support values (82 for Class 0 and 61 for Class 1, totaling 143 samples) suggest that the dataset is moderately balanced, contributing to stable model performance.

In conclusion, the proposed CNN+SVM+RF hybrid model successfully leverages deep learning for feature extraction and machine learning classifiers for precise decision-making, leading to high classification accuracy. The results demonstrate the model's potential for real-world applications in domains requiring accurate classification, such as medical imaging, agriculture, or anomaly detection. However, future improvements could focus on further enhancing recall for Class 1, possibly by fine-tuning hyperparameters or incorporating additional data augmentation techniques to ensure better generalization. Overall, this study reinforces the effectiveness of hybrid deep learning and machine learning models in achieving high-performance classification.

REFERENCES

1. A. Stephen, P. Arumugam, and C. Arumugam, "An Efficient Deep Learning with a Big Data-Based Cotton Plant Monitoring System," Int. J. Inf. Technol., vol. 16, no. 1, pp. 145–151, 2024.
2. H. Kukadiya, N. Arora, D. Meva, and S. Srivastava, "An Ensemble Deep Learning Model for Automatic Classification of Cotton Leaves Diseases," Indones. J. Electr. Eng. Comput. Sci., vol. 33, no. 3, pp. 1942–1949, 2024.
3. R. T. Thivya Lakshmi, J. Katiravan, and P. Visu, "CoDet: A Novel Deep Learning Pipeline for Cotton Plant Detection and Disease Identification," Automatika, vol. 65, no. 2, pp. 662–674, 2024.
4. M. M. Islam et al., "A Deep Learning Model for Cotton Disease Prediction Using Fine-Tuning with Smart Web Application in Agriculture," Intell. Syst. Appl., vol. 20, p. 200278, 2023.
5. A. B. Naeem et al., "Deep Learning Models for Cotton Leaf Disease Detection with VGG-16," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 2, pp. 550–556, 2023.
6. P. Singh et al., "CottonLeafNet: Cotton Plant Leaf Disease Detection Using Deep Neural Networks," Multimed. Tools Appl., vol. 82, no. 24, pp. 37151–37176, 2023.
7. V. Saxena et al., "Utilizing Support Vector Machines for Early Detection of Crop Diseases in Precision Agriculture: A Data Mining Perspective," Int. J. Intell. Syst. Appl. Eng., vol. 12, no. 16s, pp. 281–288, 2024.
8. R. Gao et al., "Intelligent Cotton Pest and Disease Detection: Edge Computing Solutions with Transformer Technology and Knowledge Graphs," Agriculture, vol. 14, no. 2, p. 247, 2024.
9. S. Jayanthy, G. Kiruthika, G. Lakshana, and M. Pragatheshwaran, "Early Cotton Plant Disease Detection Using Drone Monitoring and Deep Learning," in 2024 IEEE Int. Conf. Women Innov. Technol. Entrepreneurship (ICWITE), pp. 625–630, 2024.
10. M. SithaRam, V. Anusha, P. N. Sri, and G. H. Sri, "A Novel Methodology for Cotton Leaf Disease Detection Using CNN," in 2024 3rd Int. Conf. Appl. Artif. Intell. Comput. (ICAAIC), pp. 202–207, 2024.
11. N. Parashar and P. Johri, "Deep Learning for Cotton Leaf Disease Detection," in 2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT), pp. 158–162, 2024.
12. R. Setiawan, H. Zein, R. A. Azdy, and S. Sulistyowati, "Rice Leaf Disease Classification with Machine Learning: An Approach Using Nu-SVM," Indones. J. Data Sci., vol. 4, no. 3, pp. 136–144, 2023.