IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

SMART CROP DISEASE DETECTION USING IOT AND MACHINE LEARNING TECHNIQUES

[1] Jayagopi G, [2] Sumathi C B,
[1]Associate Professor, [2]Assistant Professor,
[1]Computer Science and Engineering,
[1]Mother Theresa Institute of Engineering and Technology, Palamanar, India

Abstract: The rapid advancement of Internet of Things (IoT) technology and the growing need for sustainable agricultural practices have driven the development of intelligent systems for crop disease detection and management. This study presents an IoT-based framework integrated with machine learning techniques for the early identification and classification of crop diseases. The system deploys a network of IoT devices in agricultural fields, equipped with environmental sensors to monitor parameters such as temperature, humidity, and soil moisture, along with camera modules to capture images of crop leaves. The collected data is transmitted to a centralized server for processing and analysis. Machine learning algorithms, including Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), are employed to analyse the image and sensor data for accurate disease detection. Both supervised and unsupervised learning methods are explored to enhance classification performance. The proposed system aims to provide farmers with timely insights, enabling proactive intervention and improving overall crop health and productivity.

IndexTerms - Internet of Things (IoT), Crop Disease Detection, Machine Learning, Agricultural Monitoring, Precision Farming, Sensor Networks, Data Analytics

I. INTRODUCTION

The agricultural sector has experienced significant transformation in recent years due to the integration of advanced technologies such as machine learning (ML) and the Internet of Things (IoT). These innovations have introduced more efficient and precise methods for crop monitoring, early disease detection, and improved management practices, replacing many traditional approaches. Crop diseases continue to pose a serious threat to global food security, often resulting in reduced yields and economic losses for farmers [2]. Conventional methods of disease diagnosis, which typically rely on visual inspection, are inherently limited—being time-consuming, subjective, and prone to human error. Moreover, by the time symptoms become apparent, the disease may have already spread extensively, making effective control more difficult and costly. To address these limitations, IoT-based disease detection systems have emerged as promising tools in modern agriculture [3]. These systems utilize interconnected sensors to monitor key environmental parameters such as temperature, humidity, soil moisture, and leaf wetness in real time [4]. The continuous data flow allows for timely insights into crop health and facilitates the early identification of abnormal conditions. When combined with machine learning techniques trained on historical datasets, these systems can detect subtle, early-stage symptoms of plant disease—often before they are visible to the human eye [6]. Furthermore, ML algorithms can support decision-making by recommending optimized interventions, such as the best times for pesticide application or irrigation scheduling. These capabilities enhance farmers' ability to manage crop diseases effectively and sustainably.

II. RELATED WORKS

Crop diseases pose a major challenge to the agricultural industry, significantly impacting both the yield and quality of produce worldwide. Traditionally, diagnosing such diseases requires specialized knowledge, often necessitating the involvement of agricultural experts. This process can be costly and time-intensive, especially for farmers in remote or underserved regions. Consequently, there is a growing need for innovative solutions that enable early disease detection while minimizing the dependency on expert intervention [7].

This research aims to address these challenges by utilizing the capabilities of machine learning (ML) and the Internet of Things (IoT) to predict and diagnose crop diseases at an early stage. By combining IoT-enabled sensors—which collect real-time environmental data—with advanced ML algorithms, it becomes possible to monitor crop health continuously and identify potential issues before they become severe. This integration offers a practical and cost-effective approach to managing disease outbreaks. Various machine learning models, such as regression algorithms and random forest classifiers, can be applied to analyze data from IoT sensors, uncovering patterns that may indicate early signs of plant diseases [8]. These models, trained on historical data, are capable of recognizing subtle symptoms that may not be immediately visible, allowing for early intervention. With timely alerts and actionable insights, farmers can take proactive measures to protect their crops and reduce potential yield losses.

The consequences of crop diseases extend beyond production loss—they can cause serious financial strain and, in some cases, discourage farmers from continuing their operations. This project seeks to empower farmers with accessible, reliable tools for disease diagnosis and management, promoting both environmental sustainability and economic stability. By leveraging the combined strengths of IoT and ML, this approach has the potential to transform how agricultural diseases are managed and ensure food security and livelihood sustainability for farmers around the world.

III. RESEARCH METHODOLOGY

This study utilized two distinct data sets: one comprising experimental data and the other consisting of real-time data collected through Internet of Things (IoT) devices. For soil moisture monitoring, we employed an Arduino Uno board in conjunction with DHT-11 sensors. The DHT-11 sensor was used to capture temperature and humidity data, while a dedicated soil moisture sensor was used to gather readings on soil humidity. A user-friendly web interface was developed to allow users to input specifications. The system is integrated with a Flask API, which processes the data through the machine learning model and presents the results in a structured JSON format, making the predictions easily accessible to users via the web interface.

The backend model of the system leverages machine learning algorithms, specifically the Random Forest Classifier (RFC), to accurately classify the data [10]. The RFC technique works by generating a decision tree for each random subset of the training data and then combining these decision trees to produce a final classification for the test data. This ensemble approach enhances the precision of the predictions. The system was applied to detect diseases affecting Areca nuts, such as Koleroga and Mahali (fruit rot), which are known to cause significant damage to tree production. By leveraging this predictive model, early detection and effective disease management are made possible, helping mitigate the impact of these infections.

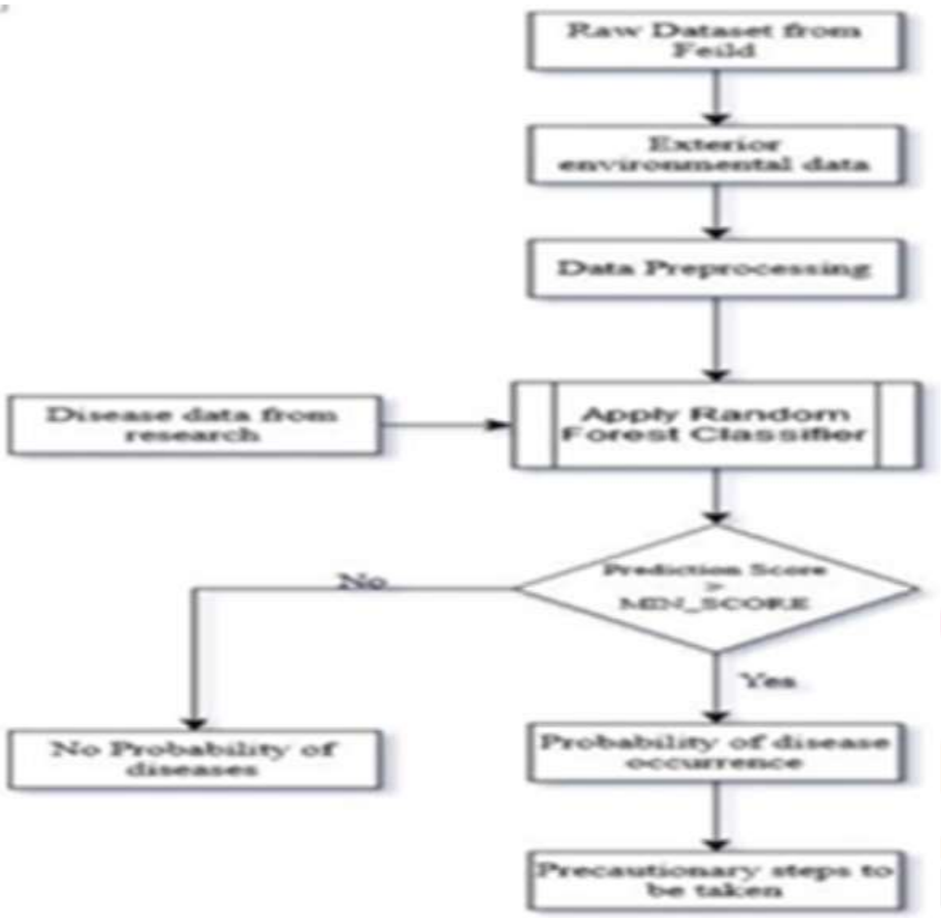

Fig.1: Depicts the flow chart for crop disease prediction using machine learning and IoT

This article explores the key environmental factors that influence the health of Areca nut trees and contribute to disease development. Environmental conditions are primary drivers of these diseases, and this article examines several critical factors in detail. Table 1 presents a straightforward dataset used to train the predictive model for Areca nut diseases. The dataset includes key environmental variables that directly affect the tree's well-being. To generate accurate predictions, users must provide information about the regional conditions, which will be used to determine the final classification of the test data. Table 2, provided by CPCRI, Vittal, summarizes the regional conditions over the past two years.

The Random Forest Classifier (RFC) processes the input data in a random manner, generating a decision tree for each subset of the training data, which may consist of hundreds or thousands of records. The final classification is based on the majority vote of the decision trees, ensuring the most reliable outcome.

Overall, the research methodology for the agricultural disease detection and management system incorporates several key elements, including the use of machine learning algorithms and structured environmental data collection, to build a robust and effective system for disease prediction and management.

Problem Definition: Crop Diseases in Agriculture

Crop diseases are a major challenge to global agriculture, leading to significant losses in both the quantity and quality of crops. These diseases not only reduce food production but also impose substantial economic burdens on farmers, particularly in regions with limited resources for disease management. Early detection of crop diseases is critical, but traditional diagnostic methods, which rely on visual inspection, are often slow, subjective, and inaccurate. This delay allows diseases to spread unchecked, further threatening food security and agricultural livelihoods.

Research Methodology for IoT-based Crop Disease Detection and Management System

1. **Data Collection**

   Collect diverse datasets that include images of both healthy and diseased crops across various types and environmental conditions. Additionally, gather environmental data such as temperature, humidity, and soil moisture through IoT sensors. The data should cover multiple regions and seasons to ensure comprehensiveness and relevance.

2. **Data Preprocessing**

   Clean the data by eliminating noise and inconsistencies. For image data, apply augmentation techniques (e.g., rotations, cropping) to enhance dataset diversity. Standardize or normalize environmental data to ensure uniformity and improve the performance of machine learning models.

3. **Model Selection**

   Choose appropriate machine learning models based on the nature of the data. For image classification tasks, **Convolutional Neural Networks (CNNs)** are ideal due to their ability to analyze image data effectively. For environmental data analysis, algorithms like **Random Forests** or **Decision Trees** are well-suited for classifying disease-related patterns based on environmental conditions.

4. **Integration with IoT Devices**

   Develop or integrate IoT devices capable of collecting environmental data from the field. Ensure seamless communication between the IoT devices and machine learning models, allowing for real-time data transfer and timely disease predictions.

5. **Prototype Development**

   Build a prototype that combines IoT devices with machine learning models to create a functional crop disease detection and management system. Test the prototype in a controlled environment to validate its operation and accuracy in disease detection.

6. **Field Testing and Validation**

   Deploy the prototype in real-world agricultural settings. Gather feedback from farmers and agricultural experts on its usability and performance. Evaluate the system's accuracy in detecting and managing crop diseases under actual field conditions.

7. **Performance Evaluation**

   Assess the system's performance using key metrics such as accuracy, precision, recall, and F1-score. Compare the performance of different machine learning models and fine-tune them to enhance disease detection effectiveness.

8. **Iterative Improvement**

   Based on feedback and performance evaluations, refine the system's design. Make necessary adjustments to improve model accuracy, system reliability, and overall usability for farmers.

9. **Documentation and Dissemination**

   Document the entire research process, including the datasets used, models developed, system architecture, and results from testing and evaluation. Share findings through publications or presentations to promote the adoption of IoT-based crop disease management solutions.

IV. RESULT AND DISCUSSION

In this section, machine learning techniques, specifically the Random Forest Classifier (RFC), were used to predict crop diseases. The system's output includes a final category of the test data, which is generated after training. Once the training phase is complete, the RFC model processes inputs randomly. To enhance prediction accuracy, we utilized 300 decision trees.

Figure 2 illustrates the input data collection process, while Figure 3 demonstrates how the system interfaces with the user. The RFC method showed a notably higher prediction accuracy compared to other models like the Decision Tree and Support Vector Machine (SVM). This improvement is primarily due to RFC's ability to mitigate the overfitting problem typically seen in decision trees, leading to more reliable and generalized predictions.

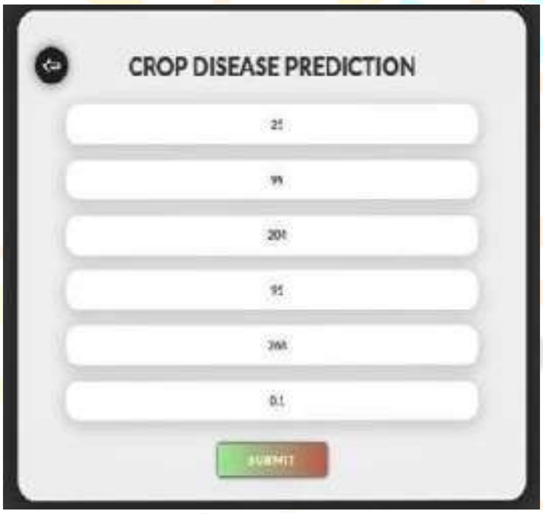

Figure 2: User interface for inputting environmental conditions.

During the training phase, data is processed to predict the test class, as outlined in Table 1. The model generates a prediction score that classifies the test data into two categories: conditions conducive to disease development and conditions unfavorable for disease. The MIN_SCORE represents the threshold for classification. If the score exceeds this threshold, it indicates a high likelihood of disease occurrence. If the score is below the threshold, the risk of disease is minimal.

The graphical representation in the system showcases the correlation between environmental conditions and disease risk, highlighting both favorable and unfavorable conditions. This visualization helps users understand how specific environmental factors contribute to potential disease outbreaks in crops.

Month	Temperature		Humidity	
	Max	Min	FN	AN
Jan	33.7	16.3	94.5	39.6
Feb	35.3	17.5	92.8	34.0
Mar	35.1	20.1	92.8	48.1
Apr	35.3	22.0	92.5	57.2
May	33.6	22.7	93.9	62.6
Jun	28.7	22.3	97.7	87.8
Jul	28.3	22.0	97.7	86.0
Aug	28.3	20.9	97.7	84.5
Sep	32.0	21.5	95.1	63.7
Oct	33.4	21.9	93.8	59.6
Nov	34.3	21.4	92.0	49.5
Dec	33.4	20.4	94.2	50.2

Table 1: Depicts the Historic data of environmental conditions of the fields

To determine the most accurate forecasting approach, three different algorithms were tested on the same dataset. The goal was to identify which method provided the most reliable predictions. Both the Support Vector Machine (SVM) and Decision Tree algorithms performed well in classifying the test data. However, the Random Forest Classifier (RFC) showed slightly better performance during the testing phase, outperforming the other approaches with the trained dataset. Additionally, data was collected through IoT-connected devices, allowing for real-time information gathering.

Although some instances involved nonlinear data, RFC demonstrated its ability to effectively handle such cases and maintain stable performance. One of the key strengths of RFC is its ability to remain unaffected by minor changes in the dataset. Since the algorithm relies on multiple decision trees, new data has a limited impact on the final outcome. This ensures stability, as new data influences only a small subset of trees at a time, making it challenging for the data to significantly alter the prediction. Furthermore, RFC is adept at managing missing values, enabling the model to operate effectively even when some data points are absent.

V. CONCLUSION AND FUTURE ENHANCEMENT

In conclusion, this study showcases the successful development and implementation of an IoT-based Crop Disease Detection and Management System powered by machine learning algorithms. By integrating IoT sensors, image processing techniques, and advanced machine learning models, the system effectively identifies crop conditions with high accuracy and provides timely management recommendations to farmers. Evaluation results demonstrate the system's ability to detect various crop diseases early, allowing for proactive interventions to minimize potential yield losses. The user-friendly interface further enhances the system's practicality, making it easy for farmers to access and interpret disease detection results.

Looking ahead, there are several opportunities for future research and development. Enhancing the system's scalability and adaptability to diverse agricultural environments and crop types could expand its applicability. Additionally, incorporating real-time weather data, such as rainfall and soil moisture, would offer more comprehensive insights into disease dynamics and crop health management. Exploring advanced machine learning techniques, like deep learning, could improve the precision and reliability of disease detection models. Remote monitoring technologies, such as drones or satellite imagery, could also be integrated to extend coverage and enable early detection over larger agricultural areas.

Moreover, collaboration with agricultural experts and stakeholders will be essential for validating the system's performance in real-world farming conditions. Incorporating feedback from farmers will enable continuous optimization, ensuring that the system evolves to meet practical needs and achieves broad adoption, ultimately enhancing the effectiveness of agricultural practices.

V. REFERENCES

[1] Y. R. Jin and S. Ji, "Mapping hotspots and emerging trends of business model innovation under networking in the Internet of Things," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, 2018.

[2] M. T. *et al.*, "Internet of Things enabled energy-efficient flying robots for agricultural field monitoring using smart sensors," in *Intelligent Technologies for Sensors*, 1st ed., Apple Academic Press, 2023. ISBN: 9781003314851.

[3] Y. X. Wei, "Study on the application of Internet of Things-based intelligent microscope in blood cell analysis," *Journal of Computational and Theoretical Nanoscience*, vol. 14, no.2 , pp. 1199–1203, 2017.

[4] P. S. *et al.*, "An experimental analysis of crop yield prediction using modified deep learning strategy," in *Proc. Int. Conf. Advances in Computing, Communication and Artificial Intelligence (ACCAI)*, Chennai, India, 2022, pp. 1–6, doi: 10.1109/ ACCAI53970. 2022.9752492.

[5] N. Gururaj, V. Vinod, and K. Vijayakumar, "Deep grading of mangoes using Convolutional Neural Network and Computer Vision," *Multimedia Tools and Applications*, 2022. [Online]. Available: https://doi.org/10.1007/s11042-021-11616-2

[6] Z. Ahmed, S. Zeeshan, D. Mendhe, and X. Dong, "Human gene and disease associations for clinical genomics and precision medicine research," *Clinical and Translational Medicine*, vol. 10, pp. 297–318, 2020. doi: 10.1002/ctm2.28

[7] S. Ramadasan, M. Ramadasan, K. Vijayakumar, and G. Balaji, "Classification of Rice- Plant Images into Healthy/Disease Class using ResNetV2 Variants," in *Proc. 2023 Int. Conf. Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2023, pp. 1–4, doi: 10.1109/ICSES60034.2023.10465409.

[8] Z. Lv, "Security of internet of things edge devices," *Software: Practice and Experience*, vol. 51, no. 12, pp. 2446–2456, 2021.

[9] M. D. Alshehri, F. K. Hussain, and O. K. Hussain, "Clustering-driven intelligent trust management methodology for the Internet of Things (CITM-IoT)," *Mobile Networks and Applications*, vol. 23, no. 3, pp. 419–431, 2018.

[10] D. Hebri, R. Nuthakki, A. K. Digal, K. G. S. Venkatesan, S. Chawla, and C. Raghavendra Reddy, "Effective facial expression recognition system using machine learning," *EAI Endorsed Trans. IoT*, vol. 10, Mar. 20, 2020. doi: 10.4108/eai.20-3-2020.1634