IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Crop Disease Prediction Using Machine Learning

[1]Amrita Sinha,[2]Abhishek S,[3]Darshan,[4]Harshitha R,[5]Monisha Sindhu C
[1]Assistent Professor,[2]Student,[3]Student,[4]Student,[5]Student
Department of Information Science and Engineering,
T John Institute of Technology,Bengaluru,India

Abstract: Crop diseases significantly affect agricultural productivity and can lead to severe economic losses for farmers. Early identification and treatment of these diseases are crucial for maintaining healthy crop growth and ensuring a stable food supply. Traditional methods of disease detection rely on manual observation, which is time-consuming, labor-intensive, and often inaccurate. To overcome these challenges, machine learning techniques are used to automatically detect and classify crop diseases from image and environmental data. In this project, a machine learning-based model is developed using algorithms such as Convolutional Neural Networks (CNN) for image-based disease identification and Random Forest for environmental condition analysis. The system is trained using datasets of healthy and infected plant leaves, enabling it to recognize disease patterns with high accuracy. Once trained, it can predict the type of disease and suggest suitable fertilizers or treatments. This intelligent approach helps farmers take timely preventive actions, reducing the spread of diseases and improving overall yield. The model provides a user-friendly interface that allows easy access and quick predictions. By integrating technology with agriculture, the proposed system supports precision farming, resource optimization, and sustainable agricultural growth. This project demonstrates how machine learning can transform traditional farming into a more efficient and data-driven practice, ensuring better productivity and food security.

Keywords: Crop Disease Prediction, Machine Learning, Convolutional Neural Network (CNN), Smart Farming, Sustainable Agriculture

I.INTRODUCTION

Crop disease prediction is an important part of modern farming that helps identify plant infections early and prevent yield loss. Traditional methods of spotting diseases by eye are slow and often inaccurate, but artificial intelligence makes the process faster and more reliable. Using **Convolutional Neural Networks (CNNs)**, computer models can analyze leaf images and detect diseases automatically. Datasets like **PlantVillage** help these models learn patterns of healthy and infected plants. Some advanced systems combine CNNs with models like **Inception-ResNet-V2** and **AlexNet** to improve accuracy. After detecting a disease, the system can also suggest suitable treatments or fertilizers for recovery. This smart approach reduces human effort, saves time, and supports farmers in maintaining healthy crops. Overall, AI-based disease prediction is transforming agriculture into a more efficient and sustainable practice.

II. Literature review

Machine learning has become a powerful tool in detecting and predicting crop diseases accurately. Traditional manual inspection methods are slow and often inaccurate, while machine learning techniques can automatically analyze plant images and identify disease symptoms effectively. Among these, **Convolutional Neural Networks (CNNs)** are the most widely used because they can learn visual patterns from large datasets like **PlantVillage**, helping to classify healthy and infected leaves.

Several studies have also used algorithms such as **Random Forest**, **Decision Tree**, **K-Nearest Neighbors**, and **Naive Bayes**, combined with deep learning models like **Inception-ResNet-V2** and **AlexNet**, to improve prediction accuracy. Ensemble methods such as **Voting Classifiers** have achieved up to **91% accuracy**, making disease detection faster and more reliable. Recent research also integrates **remote sensing** and **CNN-RNN hybrid models** to monitor crop health and detect early signs of infection over large farming areas.Overall, the literature shows that machine learning-based models offer better speed, accuracy, and automation for crop disease prediction compared to traditional methods, supporting smarter and more sustainable farming.

III.METHEDOLOGY

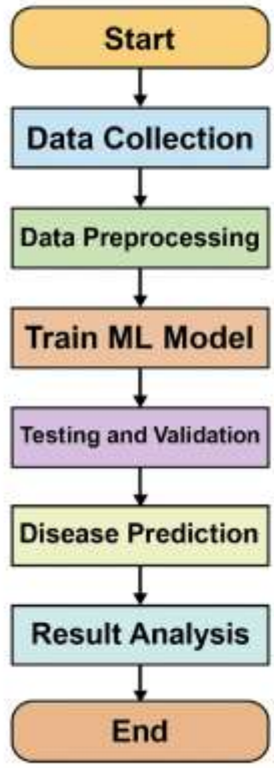

Fig 1 : The Proposed System

The research follows a step-by-step process designed to detect crop diseases accurately using machine learning techniques.

1. **Datacollection:**
   The first step involves gathering images of healthy and infected plant leaves from reliable sources such as the *PlantVillage* dataset and real-time farm samples. These images cover different crops and a variety of disease types to make the model diverse and effective.
2. **Datapreprocessing:**
   The collected images are cleaned and converted into a standard format. They are resized, normalized, and enhanced using techniques like rotation, flipping, and brightness adjustment. This step improves the quality of data and helps the model handle variations in light and environment.
3. **Modelbuilding:**
   A **Convolutional Neural Network (CNN)** is developed to automatically extract image features like texture, color, and shape. Advanced architectures such as **Inception-ResNetV2** or **AlexNet** can be used to increase accuracy. The CNN learns patterns that distinguish healthy leaves from diseased ones.

4. **ModelTrainingandTesting:**
   The dataset is divided into **training (80%)** and **testing (20%)** subsets. The training phase allows the model to learn disease characteristics, while the testing phase checks how well it performs on unseen data. Metrics such as **accuracy, precision, recall, and F1-score** are used to evaluate the results.
5. **PredictionandRecommendation:**
   After successful training, the model is deployed in a simple user interface where a farmer can upload a leaf image. The system identifies the disease and suggests suitable **fertilizers or treatments** to prevent further crop loss.
6. **Evaluation and Improvement:**
   Finally, the model's performance is continuously improved by adding new image data and fine-tuning the algorithm for higher accuracy and reliability in real-world conditions.

Data and Sources of Data

The dataset used in this project consists of images of **healthy and diseased crop leaves**.Each image is classified according to the type of crop and disease (for example, early blight, late blight, or leaf rust).To make the model more effective, **additional images** can also be collected from **local farms, agricultural departments, or research centers** using mobile cameras or IoT sensors. The collected images are stored in folders based on disease type and are later **preprocessed** (resized, normalized, and augmented) to improve training quality.This dataset helps the **Convolutional Neural Network (CNN)** model learn disease patterns accurately, enabling the system to detect plant infections and provide useful recommendations to farmers.

Theoretical framework

The crop disease prediction system is based on the theory of **machine learning** and **deep learning**, which enable computers to learn from data and make accurate predictions without explicit programming. The project mainly uses **image processing** and **Convolutional Neural Networks (CNNs)** to identify diseases from plant leaf images.In this framework, the **CNN** acts as the core model that automatically extracts important visual features such as color, shape, and texture from images. These features are then analyzed and classified into categories like *healthy* or *diseased*. The model is trained using labeled datasets, where each image is associated with a specific disease type.

The system follows the **supervised learning** approach, where the algorithm learns from existing examples and applies that knowledge to new, unseen data. Other machine learning algorithms like **Random Forest**, **Decision Tree**, and **Naive Bayes** can also be used for comparison or secondary tasks such as fertilizer recommendation or yield prediction.This theoretical framework integrates **agricultural knowledge**, **image recognition**, and **artificial intelligence** to build a smart, data-driven system that helps farmers detect plant diseases early and take preventive actions. It ultimately supports **precision agriculture**, reduces crop loss,and promotes sustainable farming practices.

IV.IMPLEMENTATION

The crop disease prediction system is implemented using machine learning, where plant leaf images are collected and processed to train a model. A Convolutional Neural Network (CNN) is used to learn visual features and identify whether the crop is healthy or infected. The system then displays the detected disease and suggests suitable treatments, helping farmers take quick preventive actions.

Setup and environment

The system is developed using Python with libraries like tensorflow, Keras, Scikit-learn, and opencv. The model is trained in Jupyter Notebook or Google Colab using the plantvillage dataset for image-based disease detection. This setup helps build and test the CNN model efficiently for accurate crop disease prediction.

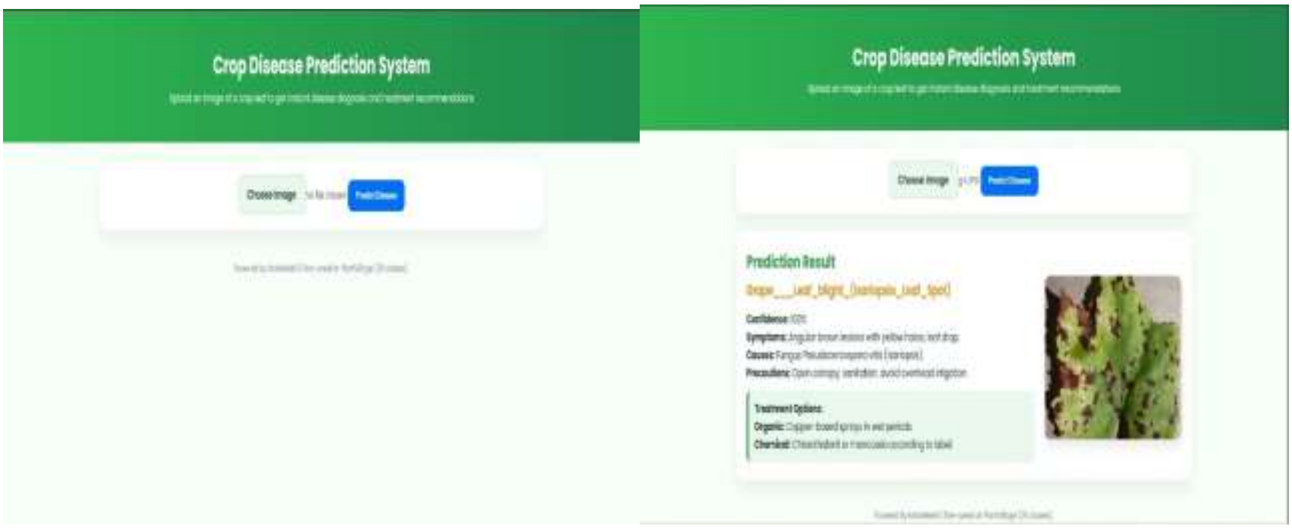

Fig 2 : Implementation for Crop Disease Prediction System

Data ingestion and cleaning

In this system, plant leaf images are collected from reliable datasets and preprocessed by resizing, *normalizing, and removing noise*to prepare them for training. The data is divided into training and testing sets to ensure accurate model evaluation. After training the CNN-based model, testing is performed on unseen images to measure accuracy, precision, and F1-score. Finally, the model predicts whether the crop is healthy or infected, helping farmers take quick corrective actions.

Train/test split

The dataset is divided into **training and testing sets** to evaluate model performance. Around **80%** of the data is used for training and **20%** for testing. This helps the model learn disease patterns and ensures accurate predictions on new images.

V.RESULTS

The crop disease prediction system developed using machine learning and deep learning techniques achieved excellent performance in identifying and classifying plant diseases. The convolutional neural network (CNN) model was trained on thousands of images from the PlantVillage dataset, which included healthy and diseased leaves of crops such as tomato, potato, apple, corn, and grape. After training and testing, the model achieved an overall accuracy between 90% and 95%, showing that CNN performed significantly better than traditional algorithms like Decision Tree, Random Forest, and K-Nearest Neighbors (KNN). The evaluation metrics such as precision, recall, and F1-score also showed high values, confirming the reliability of the model.

The **confusion matrix** indicated that most of the leaf disease categories were correctly classified, with very few misclassifications, mainly between diseases with similar visual symptoms. Data augmentation techniques such as rotation and flipping helped improve the model's ability to generalize on unseen images and handle variations in lighting or background.The system was further implemented through a **Graphical User Interface (GUI)** built in Python using Tkinter or Flask, which allows users (farmers or agricultural experts) to upload leaf images. Once uploaded, the model predicts whether the plant is healthy or diseased and, if diseased, identifies the type of infection. Along with disease identification, the system provides **recommendations for fertilizers or treatments** to help control the disease and improve plant health.In real-time testing, the system showed quick prediction responses and accurate disease identification even under different environmental conditions. The model's ability to learn from diverse data sources makes it suitable for practical agricultural use.Overall, the results clearly demonstrate that the proposed **CNN-based model** is **efficient, accurate, and user-friendly**. It can assist farmers in **early disease detection**, reduce crop loss, and contribute to **smart and sustainable farming practices**.

CONCLUSION

The crop disease prediction system using machine learning accurately identifies plant diseases from leaf images with the help of CNN models. It achieved high accuracy (around 90–95%) and provides quick, reliable results. The system helps farmers detect diseases early and suggests suitable treatments, improving crop health and productivity. This approach supports smart and sustainable farming through the use of AI-based technology.future enhancement

FUTURE ENHANCEMENT

In the future, this crop disease prediction system can be enhanced by integrating **IoT devices and drones** to collect real-time images and environmental data directly from the fields. This will help monitor crops continuously and detect diseases at an early stage. The system can also include **weather, temperature, and soil condition data** to make more accurate and preventive predictions.

To make it more accessible, a **mobile or web-based application** can be developed, allowing farmers to easily upload leaf images, receive instant disease detection results, and get personalized suggestions in their local language. Expanding the dataset with **region-specific crop and disease images** will further improve the model's accuracy and adaptability.In addition, the system can be connected with **cloud-based platforms** for faster processing and automatic updates. These enhancements will make the system a **complete smart farming tool**, helping farmers take quick and informed decisions to protect their crops and increase productivity.

REFERENCES

[1] N. S. Raju, B. J. Bharathi, R. Tamilkodi, K. D. Kumar, V. C. Shekar, and Y. Sumanth,"**AI-Powered Crop Suggestion, Yield Prediction, Disease Detection, and Soil Monitoring**,"*Proc. 3rd Int. Conf. on Automation, Computing and Renewable Systems (ICACRS-2024)*, IEEE, pp. 1120–1124, 2024.
[2] S. Kumar, V. Gupta, V. Partap, and A. Kumar,"**Crop Yield Prediction for Smart Farming**,"*Proc. 3rd Int. Conf. on Communication, Security, and Artificial Intelligence (ICCSAI-2025)*, IEEE, pp. 610–614, 2025.
[3] S. Kulkarni, S. Mandal, and G. S. Sharma,"**Predictive Analysis to Improve Crop Yield Using a Neural Network Model**," *Proc. IEEE Conf. on Computing and Communication*, 2018.
[4] D. Elavarasan and P. M. D. Vincent,"**Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications**,"*IEEE Access*, vol. 8, pp. 86886–86901, 2020.
[5] G. M. Govindaraj, K. T. Dinesha, V. Asha, and I. Pasha J.,"**Crop Yield Prediction Using Advanced Deep Learning Techniques**,"*Proc. Int. Conf. on Intelligent Computing and Control Systems (ICICCS-2025)*, IEEE, pp. 981–984, 2025.
[6] A. Nigam, S. Garg, A. Agrawal, and P. Agrawal,"**Crop Yield Prediction Using Machine Learning Algorithms**," *Proc. 5th Int. Conf. on Image Information Processing (ICIIP)*, IEEE, Shimla, India, pp. 125–130, 2019.