International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Plant Disease Detection and Classification

Dr. Sajja Suneel
Assistant Professor
Department of Data Science
Institute of Aeronautical Engineering
Hyderabad, India
sajja.suneel@gmail.com

1st Lambu Vyshnavi
dept of Data Science
Institute of Aeronautical Engineering
Hyderabad, India
vyshnavireddy488@gmail.com

2nd Neelala Vamshi
dept of Data Science
Institute of Aeronautical Engineering
Hyderabad, India
vamshineelala200@gmail.com

3rd Panuganti Shivani
dept of Data Science
Institute of Aeronautical Engineering
Hyderabad, India
shivanipanuganti0712@gmail.com

***Abstract*—Everyone's daily existence in the twenty-first century involves a large amount of machine learning.These days, a wide range of applications, including object identification, object categorisation, and medicinal uses, use it.This seeks to identify illnesses in plant leaves using deep convolutional neural networks. Farmers typically use manual disease detection techniques since they are ignorant of illnesses on plant leaves. They frequently become less abundant as the virus multiplies.However, in many parts of the world, rapid identification has to be enhanced due to a lack of necessary infrastructure.The efficient handling of plant diseases is essential for both food security and agricultural production.The proposed application uses CNNs for plant disease detection and classification. It attempts to identify illnesses from photos and offers disease descriptions and preventive steps and suggests suitable pesticides along with product link for further action and other features of the application are crop recommendation and fertilizer prediction based on some values(nitrogen,phosphorous,pottasium).The CNN model architecture is optimized for accurate disease classification,taking into account various types of plant diseases and leaf variations.This aims to enhance agricultural practices by providing a scalable and efficient tool for early disease detection,classification, and treatment recommendation,crop recommendation,thereby contributing to improved crop yield and sustainable farming practices.**

***Index Terms*—Plant Disease Detection, CNN (Convolutional Neural Network), Automated Disease Description,Pesticide Suggestion,Precision Farming.**

I. Introduction

Agriculture is a critical sector that sustains the global economy and meets the food demands of the ever-growing population.However,one of the key challenges farmers face is the prevalence of plant diseases, which can drastically impact both crop productivity and standard.The timely identification and precise diagnosis is vital for reducing the negative effects on productivity[9].When diseases are not identified and treated promptly,they can result in considerable economic losses and threaten food security,especially in regions that heavily depend on agriculture.Traditionally,farmers have relied on visual inspections of crops to identify signs of infection,such as changes in leaf color, drooping.Although this approach is simple and cost-effective,it is often subjective and prone to inaccuracies,as symptoms can vary between different diseases,crop types, or environmental conditions.Furthermore,expertise is required to correctly identify these diseases,which may not be readily available,particularly in rural or less-developed areas.The inefficiency of manual inspection methods often results in delayed responses to infections,allowing diseases to spread and affect overall agricultural productivity.

In many parts of the world,farmers still face challenges due to limited access to modern diagnostic tools and infrastructure,making timely disease detection and control difficult.The rapid spread of infections,especially in the case of viral or bacterial pathogens,complicates disease management further,as they often go unnoticed until the later stages.This highlights the urgent need for scalable, automated,and affordable systems that can effectively detect plant diseases.Recent advancements in ML and DL offer promising solutions for addressing agricultural challenges[5]. Image-based analysis through deep learning has gained prominence for its ability to detect,classify, and monitor diseases with a high degree of accuracy[1]. Machine learning models that are trained on extensive image datasets have the capability to recognize fine differences in leaf texture, color, and shape details that may not be easily discernible to the human eye.[4].CNNs are specialized dl models designed specifically for handling and analyzing visual information. They have demonstrated remarkable effectiveness in tasks involving image analysis, including object detection and classification. This capability makes them particularly well-suited for detecting plant diseases. By training these models on labeled datasets of plant leaf images, CNNs can learn to identify specific diseases and differentiate between healthy and infected plants. These models not only automate disease detection but also deliver more accurate and consistent results compared to traditional methods.

II. Literature Review

[1] Presents a comprehensive work on plant disease detection using image processing techniques. The authors reviewed

various datasets used for plant disease classification, including public and proprietary ones. They discussed traditional algorithms such as thresholding, edge detection, and color-based segmentation, while highlighting their limitations under varying environmental conditions. The paper emphasized a shift toward ml and dl techniques like CNNs, which showed improved performance on large, complex datasets. The survey demonstrated the growing potential of these methods for scalable, automated disease detection in agriculture.

[2] The authors presented an Improved YOLOv5 model for plant disease detection. They utilized the Plant Village dataset for training and validation, containing various plant leaf images with labeled diseases. The lightweight version of YOLOv5 was optimized to reduce computational load while maintaining high detection accuracy. This optimization improved processing speed, making the model suitable for real-time applications in resource-constrained agricultural environments. Their research demonstrated that the model effectively detects and classifies plant diseases, enabling scalable, efficient agricultural solutions.

[5]"The authors trained and assessed machine learning models for illness detection and classification using the Plant Village dataset, which consists of a sizable collection of labelled photos of plant leaves. To examine illness trends, they used Random Forest methods and SVM. Utilising feature extraction methods like colour and texture analysis improved the performance of the models.The results of the study demonstrated the use of varied datasets, such as Plant Village, in reaching high classification accuracy. This illustrated ml's scalability and dependability for automated plant health monitoring.

[7]suggested categorising and identifying tomato plant illnesses using dl algorithms. The study concentrated on using CNNs to examine tomato leaf photos. The scientists showed that CNNs can diagnose a wide range of illnesses that harm tomato crops with great accuracy, outperforming conventional approaches in terms of automation and precision. Their study demonstrated how dl may improve agricultural disease control and provide farmers with real-time remedies to increase crop health and productivity.

[10]The authors looked at the efficacy of deep learning methods and conventional machine learning techniques for diagnosing illnesses of plant leaves. The study contrasted several algorithms, such as CNNs, decision trees, and support vector machines, pointing out the advantages and disadvantages of each in terms of efficiency and accuracy. The results showed that in terms of accuracy and speed of illness diagnosis, dl models—in particular, CNNs—perform noticeably better than traditional techniques. This study offers insights for creating reliable disease detection systems suited for agricultural applications and emphasises the need of utilising cutting edge technology to improve plant health monitoring.

III. EXISTING SYSTEM

Existing systems for plant disease detection have traditionally relied on manual inspections carried out by farmers or agricultural experts, who visually examined plants for signs like discoloration, wilting, or unusual growth. While this approach was straightforward and widely accessible, it was often subjective, prone to errors, and could lead to delays in identifying diseases, increasing the risk of crop loss. Early technological solutions involved basic image processing methods, such as color thresholding and edge detection, to analyze plant leaves. However, these techniques struggled to cope with the variability of disease symptoms under different environmental conditions, limiting their effectiveness.In addition to image processing, cnn models such as SVM and decision trees were used to classify plant diseases by extracting handcrafted features from images. Although these models improved accuracy compared to basic image analysis, they required substantial domain expertise to identify relevant features. Moreover, they had difficulty handling complex datasets with significant variability. Rule-based expert systems were another attempt to replicate human diagnosis but were limited in adaptability, often becoming outdated as new diseases or variations emerged.Overall, while these systems marked important steps forward, they were constrained by limitations in accuracy, scalability, and the ability to handle diverse datasets. These challenges underscored the need for more advanced methods, paving the way for the adoption of Convolutional Neural Networks (CNNs), which offer greater precision and automation in plant disease detection.

IV. METHODOLOGY

CNN Architecture

Convolutional Neural Networks (CNNs) are a type of deep learning models specifically designed to handle structured grid data, such as images. Several critical layers make up a CNN's architecture, which functions as a unit to automatically extract hierarchical information from input pictures. The input layer, which is the initial layer, is where the image's raw pixel values are received. Then come several convolutional layers that use filters (or kernels) to execute convolution operations and extract features like textures, edges, and patterns. To provide non-linearity to the model, an activation function, such as ReLU (Rectified Linear Unit), is typically used for each convolutional layer.

After the convolutional layers, the feature maps are down-sampled using pooling layers, which reduces dimensionality and computational complexity without losing crucial information. Two common pooling techniques are max pooling and average pooling. In order to incorporate the high-level characteristics that the preceding layers have learnt and provide final predictions, one or more fully linked layers are added after these levels. The output layer commonly utilises a softmax activation function for multi-class classification tasks, yielding probabilities for each class.Because of its hierarchical nature, CNNs are especially good at tasks like object identification and picture classification because they can recognise complex patterns and spatial hierarchies in images. CNNs are useful in plant disease detection and classification because they can automatically identify and extract key features from images, such as textures, patterns, spots, and color differences, which

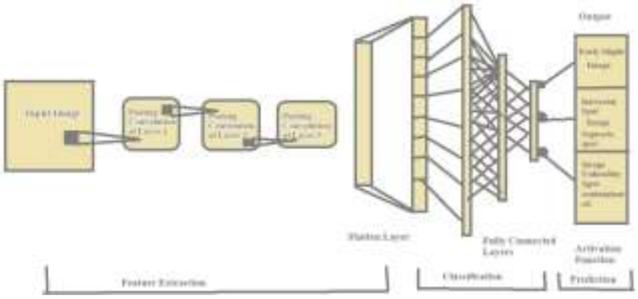

Fig. 1. CNN Architecture

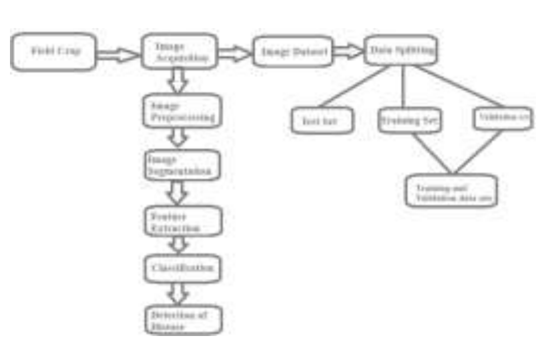

Fig. 2. Flow of working

are essential for recognizing different diseases. The primary image processing and manual observation methods used in earlier approaches were less scalable and more prone to human error. With the introduction of CNNs, automatic detection has improved in accuracy, enabling systems to recognize a variety of plant diseases from pictures with little to no preprocessing. To train models and get high classification accuracy, researchers have looked at a variety of datasets, including the Plant Village dataset. By lowering pesticide usage and crop loss, these models support sustainable agricultural practices in addition to helping farmers detect diseases early. Random Forest, an ensemble learning method, is commonly used for both classification and regression tasks. It constructs multiple decision trees during the training phase and, for classification, outputs the most frequent class (mode) among the trees. In the case of regression, it averages the predictions from all the trees to produce the final result.Overfitting is reduced and the model's robustness is raised since each tree in the Random Forest is built using a random selection of features and training data.This aids in the analysis of soil and environmental characteristics to forecast appropriate fertilisers and offer crop suggestions.

V. Implementation

Implementing Plant disease detection and classification using convolutional neural networks involves several steps, including data preparation, defining the model architecture, training the model, and evaluating its performance.

1. *Importing Dependencies:*

- This step imports necessary Python libraries.
- Numpy and pandas are for data manipulation, matplotlib for visualization, torch for PyTorch (the deep learning framework), and other modules for dataset handling, image transformation, and neural network functionality.

2. *Dataset Collection:*

- The Plant Village Dataset comprises 61,486 images, including both healthy and diseased plant leaves, utilized for training and evaluating models. It contains images across 39 distinct categories of plant leaves, along with background images.
- A dataset which has nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall is used for crop and fertilizer prediction.
- The transform object defines image pre-processing steps such as resizing images to 255x255, cropping them to 224x224, and converting them to tensors (PyTorch's data format).

3. *Train-Test-Validation Split:*

- The data-set containing 61,486 images.
- The dataset is divided into three subsets: training, validation, and testing. A total of 36,584 images are allocated for training, 15,679 images for validation, and the remaining images are designated for testing purposes.
- Trainsampler, validationsampler, and testsampler are created to randomly sample data for training, validation, and testing, respectively.The data loaders read data in batches for efficient processing, each with a batch size of 64.
- Index-to-Class Mapping,This dictionary maps the output class indices (0 to 38) to the corresponding plant diseases. For instance, index 0 corresponds to Apple Apple Scab, and index 38 corresponds to Tomato Healthy.

4. *Model Creation:*

- CNN is used for image classification, specifically to classify the images into 39 categories (different plant diseases).
- The Adam optimizer is used for optimization, and CrossEntropyLoss is used as the loss function, as this is a multi-class classification problem. Batch Gradient Descent for training and validation. It computes losses for both the training and validation sets, updating model weights at every epoch.
- Random Forest is used for crop recommendation and fertilizer prediction, which predicts the most suitable crop and fertilizer based on environmental factors like nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall.

5. *Model Evaluation:*

- The accuracy function assesses the model's performance using datasets for testing, validation, and training. Defined as the percentage of properly predicted occurrences

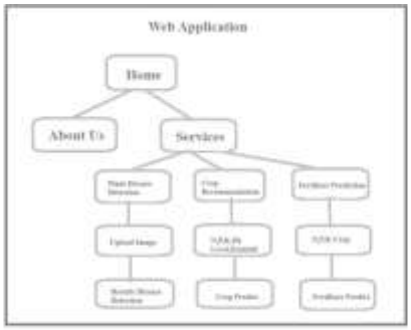

Fig. 3. Flask-based web application

compared to the total number of predictions, accuracy is a crucial performance parameter for models. On the training data, the model obtains an accuracy of 96per, on the validation data, 98per, and on the test data, 98per.

6. *Saving the Model:*

- The trained model is saved to detect and classify plant diseases.

Flask-based web application The proposed model provides three main features: plant disease detection,crop recommendation and fertilizer suggestion.

- The web application is built using Flask and serves multiple pages.
- Homepage,which directs users to different parts of the application and it consists of information about the page and its services
- Plant Disease Detection page,The user can upload a plant image, and the application uses the CNN model to predict the disease. Information about the disease (name, description, preventive steps) is retrieved from the dataset.The application also displays supplementary information,that can be used to treat the plant and provides a link for purchasing it.
- Crop Recommendation page,It accepts input parameters including nitrogen, phosphorus, potassium levels, pH, rainfall, and the city.Based on the information,it recommends the crop.
- Fertilizer Recommendation page,Users can input their crop name and the levels of nitrogen, phosphorus, and potassium. The application compares these values with ideal nutrient levels from the fertilizer dataset.Based on the differences between actual and ideal values, it suggests whether the user needs more or less of nitrogen, phosphorus, or potassium.
- The application provides a user-friendly interface for farmers and agriculturalists to predict the best crops, suggest fertilizers, and diagnose plant diseases.

VI. Results

The results of flask based web application are: Upon uploading an image of a diseased plant leaf, the application identifies the disease,provides a detailed description of the disease, possible preventive measures, and a link to purchase relevant supplements. After inputting soil and environmental parameters,the application suggests the best-suited crop. Based on the crop type and nutrient levels entered, the application advises which nutrient (N, P, or K) to increase or decrease and provides a recommended fertilizer type. The web application helps farmers by accurately detecting plant diseases through CNNs, recommending suitable crops based on soil and environmental conditions, and predicting fertilizer needs

Fig. 4. Plant Disease Detection And Classification

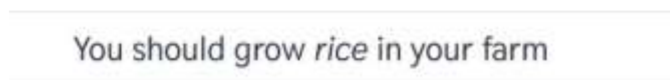

Fig. 5. Crop Recommendation

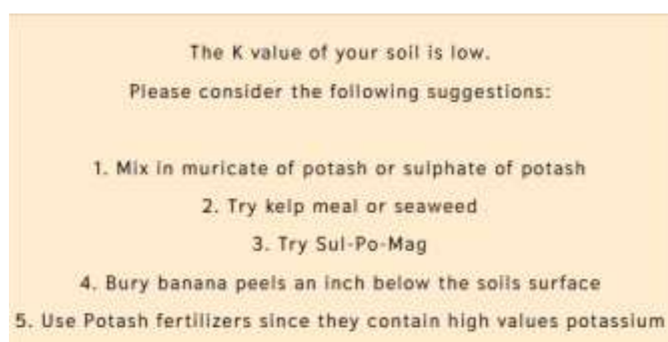

Fig. 6. Fertilizer Prediction

for optimal yields. It promotes sustainable farming practices, increases productivity, and reduces crop losses.

VII. Conclusion and future enhancements

In conclusion, proposed a web application,plant disease detection and classification provides a comprehensive approach to tackle the challenges faced by farmers in managing crop health and maximizing productivity. By leveraging advanced image-based analysis through CNNs,application enables accurate,timely detection of plant diseases directly from leaf images. Once a disease is identified, the system provides users with detailed descriptions of the disease, along with prevention steps, pesticide suggestions, and links to purchase the recommended pesticides. This approach empowers farmers with actionable insights to treat and control plant diseases effectively.The application also includes additional features such as crop recommendation and fertilizer prediction by using random forest algorithm. Based on input parameters like soil nutrients (nitrogen, phosphorus, potassium), pH value, and environmental conditions such as rainfall, the system recommends suitable crops for optimal yield. Moreover, it helps farmers manage soil health by predicting the right fertilizer combinations for their crops, thereby promoting sustainable farming practices.Overall, the web application equips farmers with modern, automated tools for improving crop management, enhancing yields, and making informed decisions to adopt sustainable agricultural practices. Future enhancements that can further improve the effectiveness and usability of this web application, A crucial area for enhancement involves broadening the dataset to encompass a greater range of crops and diseases. This expansion will enable the system to support a wider network of farmers from various agricultural regions. Additionally, implementing a multi-language interface would help cater to farmers from different linguistic backgrounds, increasing accessibility and user adoption.

References

[1] M. B. Ahmad Supian, H. Madzin and E. Albahari, ”Plant Disease Detection and Classification Using Image Processing Techniques: a review,” 2019 2nd International Conference on Applied Engineering (ICAE), Batam, Indonesia, 2019, pp.

[2] Haiqing Wang,Shuqi Shang,Dongwei Wang,Xiaoning He,Kai Feng and-Hao Zhu,”Plant Disease Detection and Classification Method Based on the Optimized Lightweight YOLOv5 Model”,2022

[3] Roy, A.M.; Bhaduri, J. A Deep Learning-Enabled Multi-Class Plant Disease Detection Model Based on Computer Vision. AI 2021

[4] L. S. Puspha Annabel, T. Annapoorani and P. Deepalakshmi, ”Machine Learning for Plant Leaf Disease Detection and Classification – A Review,” 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019.

[5] P. Panchal, V. C. Raman and S. Mantri, ”Plant Diseases Detection and Classification using Machine Learning Models,” 2019

[6] S. Y. Yadhav, T. Senthilkumar, S. Jayanthy and J. J. A. Kovilpillai, ”Plant Disease Detection and Classification using CNN Model with Optimized Activation Function,” 2020

[7] H. Hong, J. Lin and F. Huang, ”Tomato Disease Detection and Classification by Deep Learning,” 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Fuzhou, China, 2020

[8] B. Rajesh, M. V. Sai Vardhan and L. Sujihelen, ”Leaf Disease Detection and Classification by Decision Tree,” 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, 2020

[9] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis.

[10] Hardikkumar S.Jayswal and Jitendra P. Chaudhari,”Plant Leaf Disease Detection and Classification Using Conventional Machine Learning and Deep Learning”

[11] Real-Time Plant Health Detection Using Deep Convolutional Neural Networks by Mahnoor Khalid 1,Muhammad Shahzad Sarfraz 1,Uzair Iqbal 2ORCID,Muhammad Umar Aftab 1ORCID,Gniewko Niedbała 3,ORCID andHafiz Tayyab Rauf 4.

[12] Deng, L.; Deng, Q. The basic roles of indoor plants in human health and comfort. Environ. Sci. Pollut. Res. 2018, 25, 36087–36101

[13] Balasundram, S.K.; Golhani, K.; Shamshiri, R.R.; Vadamalai, G. Precision agriculture technologies for management of plant diseases. In Plant Disease Management Strategies for Sustainable Agriculture through Traditional and Modern Approaches; Springer:Berlin/Heidelberg, Germany, 2020; pp. 259–278

[14] Hasan, R.I.; Yusuf, S.M.; Alzubaidi, L. Review of the state of the art of deep learning for plant diseases: A broad analysis and discussion. Plants 2020, 9, 1302.

[15] Lin, K.; Gong, L.; Huang, Y.; Liu, C.; Pan, J. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front. Plant Sci. 2019, 10, 155