Early Stage Plant Disease Detection Using Machine Learning

*Abstract* - Plant diseases generate significant productivity and economic losses, as well as a decline in both the quality and quantity of agricultural goods. Plant disease identification is now receiving more attention in wide fields of agricultural monitoring. Farmers face significant challenges when transitioning from one disease management regime to another. The typical technique used in practise for detecting and identifying plant diseases is skilled naked eye inspection. In this research, we examine the necessity for a simple plant leaf disease detection system that would aid agricultural innovations. Early information on crop health and disease identification can aid in disease control through correct management measures. This strategy will increase agricultural productivity. This report also examines the advantages and disadvantages of these.

I. Introduction

The reason that agriculture is the foundation of any country is because it not only produces food but also many other essential resources like cotton, textiles, and many others. India is the second-largest producer of wheat and grains in the world, contributing to both global food security and economic growth through exports. India has the sixth-largest economy in the world, with approximately 15% of its GDP coming from the agricultural industry. Knowing how essential our agriculture industry is makes it imperative to safeguard it and implement policies that will maximise its potential. According to accounts, pests, weeds, and other diseases damage 15–25% of the crop output each year. This brings us to our project, "Early Stage Crop Disease Detection Using Machine Learning Algorithm."

Our project is based on a model where we first use a drone camera to survey the entire field, taking detailed pictures for our data and then scrutinising the pictures for every tiny detail. Once we have the data, our model (using CNN, Open cv software’s) works to find the precise disease and at precise positions on the field. Once we have identified the illness and the affected area, we search for the ideal treatment before using a drone to spread pesticides and medications over the affected area. It lessens the farmer's workload and, in our trials, had a precision of up to 95%, which, if put into practice, could be very effective.

Given the significance of agriculture, it is our moral obligation to make improvements to this industry that will have a beneficial effect on the economy and development of our nation. Therefore, we are developing a model that will identify the illness at an early stage and tend to eliminate it, preventing it from spreading and harming our prosporous produce. We are equipped with information that will enable us to combat various illnesses and find the best treatment for each one. We attempted to provide a way that will assist our farmers and get the best out of their produce because they work extremely hard for their crop and if it gets infected it causes them a lot of trouble.

II. Literature Survey

[1]. The use of image processing in agriculture is becoming increasingly important, as it can help detect defects and diseases in crops that may not be visible to the naked eye. This can help farmers prevent significant crop losses due to pests or other factors. Various techniques are used for disease recognition, including image acquisition, pre-processing, segmentation, feature extraction, and categorization. Some of the methods which are used are : K-means Clustering, GLCM & SVM, Otsu’s Detection, CNN-ANN-KNN, Histogram Technique.

Firstly the images of healthy & unhealthy leaf are stored for experimentation & then they are sent for image pre-processing. Then K-means clustering is used for segmentation and feature extraction using GLCM. Then SVM is used for classification. Otsu’s detection is used to convert the RGB image of leaf into HSV (Hue, Saturation Value). The classification approach is then carried out by KNN, ANN & CNN. The KNN method classifies samples by finding the nearest distance between trained and testing subjects. ANN method is used as a classifier.

One of the earliest systems on early stage crop disease detection was developed by Singh et al. (2016) [2]. The system used image processing techniques to analyze images of plants infected with different types of diseases. The system achieved an accuracy of 85% in detecting different types of diseases.

In 2017, Khan et al. proposed a system that used machine learning algorithms to detect the presence of wheat rust disease. The system achieved an accuracy of 98.5% in detecting the disease. The authors also reported that the system could be used to detect other types of crop diseases with some modifications.

In 2018, Wei et al. [3] proposed a system that used deep learning algorithms to detect tomato diseases. The system used a convolutional neural network (CNN) to classify images of tomato leaves into different categories based on the type of disease present.

In 2019, Wang et al. [4] proposed a system that used hyperspectral imaging to detect apple diseases. The system used a support vector machine (SVM) algorithm to classify the hyperspectral images of apple leaves into healthy and diseased categories. The system achieved an accuracy of 94.1% in detecting apple diseases.

In 2020, Wang et al. [5] proposed a system that used machine learning algorithms to detect maize diseases. The system used a combination of image processing and machine learning techniques to analyze images of maize leaves and detect the presence of diseases. The system achieved an accuracy of 95.1% in detecting different types of maize diseases.

In 2021, Zhang et al. [6] proposed a system that used deep learning algorithms to detect wheat diseases. The system used a CNN to analyze images of wheat leaves and classify them into different categories based on the type of disease present. The system achieved an accuracy of 96.2% in detecting different types of wheat diseases.

III. Proposed Model

Here, we are utilizing our model to determine which leaf in a crop is good or ill. Therefore, in order to accomplish this, we go through a number of processes, including picture resizing, data preprocessing, and feature extraction. We use OpenCV to resize images. In OpenCV, picture resizing is accomplished using the following methods.

1. Keep Aspect Ratio in Mind (height to width ratio of image is preserved)

i) Downscale (Reduce the image's area);

ii) Upscale (Increase the size of the image)

2. Don't maintain the aspect ratio

i) Only adjust the breadth (Increase or decrease the width of the image keeping height unchanged)

ii) Only adjust the height (Increase or decrease the height of the image keeping width unchanged)

3. Specify a new breadth and height for

And after that, we resize our picture in OpenCV using the methods below to make it

256 x 256 pixels. Afterward, we divide our picture collection into training and validation groups. After that, we insert our collections into the CNN architecture model, where various levels allow us to separate healthy and unhealthy images. Here, we employ three different layer types—32 px, 64 px, and lastly, 128 px—to analyze our picture and determine whether it is healthy or not.

IV. Algorithm Description

Amass a collection of crop picture data that contains both healthy and disease-affected plants. The pictures should be preprocessed by being uniformly sized and having their pixel values normalized.

Create training and validation groups from the information. Describe the structure of the CNN model, including the quantity and dimensions of the convolutional, pooling, and completely linked layers. Compile the model while indicating the loss function, algorithm, and assessment measures.

Using the fit() method, train the model on the training set while specifying the group size and number of epochs. Utilize the evaluate() function to gauge the model's success on the validation collection.

Retrain the model on the training set after fine-tuning the hyperparameters (such as learning rate and number of layers).

Use the algorithm to forecast the health state of new crop images after it has been optimized.

Analyze the model's F1 score, precision, memory, and accuracy using the new data.

Automatic or semi-automatic disease diagnosis procedure.

CNN: A CNN is a machine learning algorithm that receives an image input, prioritises distinct features and items, and distinguishes them from one another. It operates by removing characteristics from pictures.

Decision Trees: Decision trees are used to represent choices and conclusions in an intuitive and visual manner. It is an effective categorization and regression instrument. As the name implies, it is a tree structure where each leaf node has a class title, each branch represents a test outcome, and each internal node represents a test on an attribute.a Random forest:

Healthyy

i. The model is given more unpredictability thanks to Random Forest. It searches for the finest characteristics within a randomly selected subset of features rather than the most crucial ones.

ii. Random forest adds extra randomness in image.

Open CV: To resize the picture up or down to the required dimensions in OpenCV. You can't just enter one or the other for 1000 and let adjust () figure out the other because it needs either the target size (in both dimensions) or the scaling (in both dimensions).

Input Plant Images

In CNN Architecture Image processed in different layers

Resize Image into 256 x 256 Px

Unhealthy

Fig 1. Flowchart on training Model

V. Conclusion

Considering the aforementioned information, we can draw the conclusion that this article provides a general overview of the methodology and provides a concise summary of various techniques helpful for the early detection of plant diseases.Our model can add significant value to a farmer's crop because it was created by attempting to merge modern technology with traditional agricultural methods. The result is a model with a high degree of accuracy and efficiency, which immediately increases the total yield from the

land. We have made an effort to meet the demand for cost-effective, efficient methods that provide farmers with solid solutions for the improvement of their farming fields. This project will assist farmers in addressing a major issue that can quickly undo all of their hard work and had developed a less complicated method with a high effectiveness (roughly 95%) that will have a significant impact and led to the improvement of farmers as well as the world. In the future, our goal is to create a system for early automatic tracking and fixing that is more effective and automated and can be expanded to detect all potential illnesses.

References

[1]. Singh er al.Using Deep Learning for Image-Based Plant Disease Detection https://www.frontiersin.org/articles/10.3389/fpls.2016.01419/full

[2] Wei et al Applying Big Data Based Deep Learning System https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9142151

[3] Wang et al Detection of Early Rottenness on Apples by Using Hyperspectral Imaging Combined with Spectral Analysis and Image Processing

<https://www.researchgate.net/publication/271646924_Detection_of_Early_Rottenness_on_Apples_by_Using_Hyperspectral_Imaging_Combined_with_Spectral_Analysis_and_Image_Processing>

[4] Identification of Corn Leaf Disease Based on Image Processing <https://www.researchgate.net/publication/340896870_Identification_of_Corn_Leaf_Disease_Based_on_Image_Processing>

[5] Zhang et al Wheat leaf disease identification based on deep learning algorithm

https://www.sciencedirect.com/science/article/pii/S0885576522001552