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

Potato Leaf Disease Detection Using Machine Learning

Dr. Anjali Singhal

1. Pranjal Singh, 2. Arpit Jaiswal, 3. Pranav Goyal, 4. Satyam Singh

Dept. of Computer science & Engineering
Inderprastha Engineering College
Ghaziabad U.P, BHARAT

Abstract - Analysing plant phenotype stands as a pivotal component in understanding plant development. This study introduces a streamlined methodology aimed at discerning between healthy and diseased or infected leaves through the integration of image processing and machine learning methodologies. Diverse ailments impair leaf chlorophyll, manifesting as discoloured spots on the leaf surface, typically brown or black in hue. The proposed approach employs a sequence of techniques including image preprocessing, segmentation, feature extraction, and machine learning-based classification. Notably, the incorporation of a convolutional neural network (CNN) notably enhances detection accuracy.

***Key Words*: Convolutional Neural Network (CNN), Confusion matrix, Machine Learning**

1.INTRODUCTION

In India, agriculture plays a pivotal role in economic development, serving as a crucial resource. Continuous efforts are made by agricultural industries to enhance food production by seeking efficient methods to safeguard crops from damage. This prompts researchers to explore new, precise technologies aimed at maximizing productivity. Crop diseases significantly impact production levels, causing economic losses to farmers and agricultural sectors. Identifying diseases is vital for a thriving farming system. Typically, farmers rely on visual inspections to detect symptoms, necessitating constant monitoring. However, pinpointing diseases, especially those affecting leaves, poses challenges. Image processing methods utilizing plant leaf images offer a promising solution for disease detection. Despite the benefits of continuous health monitoring and disease detection, the process is financially burdensome. Machine learning algorithms are increasingly being explored for their superior accuracy. However, selecting appropriate classification algorithms remains a challenge due to varying accuracy across different input data. The objectives of this study are to detect leaf disease from images, extract features from diseased leaf segments, and recognize the condition using Support Vector Machines (SVM). Additionally, the accuracy of Convolutional Neural Networks (AlexNet) is evaluated and compared. The paper is organized into five sections: introduction, literature review, methodologies including image feature extraction, SVM, and CNN, classification results, and conclusion and future scope. The sections are structured as follows: Section 2 covers the literature review, Section 3 discusses the system architecture, Section 4 outlines research methodologies including dataset, collection, preprocessing, and classification techniques. Theoretical and mathematical aspects of feature selection and classification algorithms are detailed. Finally, Section 5 addresses the conclusion and future directions, while the last section acknowledges contributions and lists references.

2. LITERATURE REVIEW

In a prior investigation conducted by Aditi Singh and Harjeet Kaur [1], they employed a device to input data and conducted segmentation on individual plant leaves post context removal. Subsequently, they analyzed the segmented diseased components of the leaf using a high-pass filter. The outcomes revealed a motivation for the proposed model, aimed at detecting and classifying affected and unaffected potato leaves. Remarkably, the proposed framework achieved an impressive accuracy of 95.99%.

Another scholarly inquiry by Tan Soo Xian and Ruzelita Ngadiran [2] delved into the realm of plant diseases classification utilizing machine learning algorithms. Their study centered on the efficacy of the Extreme Learning Machine (ELM) classifier, aiming to match or surpass the performance of established classification models like Decision Trees (DT) and Support Vector Machines (SVM). The findings indicated that ELM exhibited promising results, showcasing comparable performance with decision tree classifiers using proposed image features, while simplifying complexity.

Furthermore, a study on plant diseases and pests detection based on deep learning, as outlined in [3], analyzed various methods within the deep learning framework. This analysis encompassed network structures for classification, detection, and segmentation of plant diseases and pests.

In [4], the focus was on detecting leaf disease portions from images, extracting features from exposed leaf areas, and recognizing diseased leaves through Support Vector Machines (SVM) and Convolutional Neural Networks (CNN), specifically AlexNet. The study highlighted the potential of transfer learning and other CNN models to enhance the accuracy of tomato leaf disease detection.

3. ARCHITECTURE

Dataset collection is the act of gathering information containing potato leaf images. we get this dataset from kaggle dataset. we use some data cleaning and preprocessing process. As we are using some deep learning, we are using tensor flow library accurate results. We use tf dataset which help us to represents a sequence of elements, in which each element consists of one or more components. Data augmentation is used to get more data from the existing type of data and helps us to zoom out and zoom in for it. And after that we are using convolution neural network model for model building which extract features from it and get accurate results.to create a website where we can click photos and upload it.

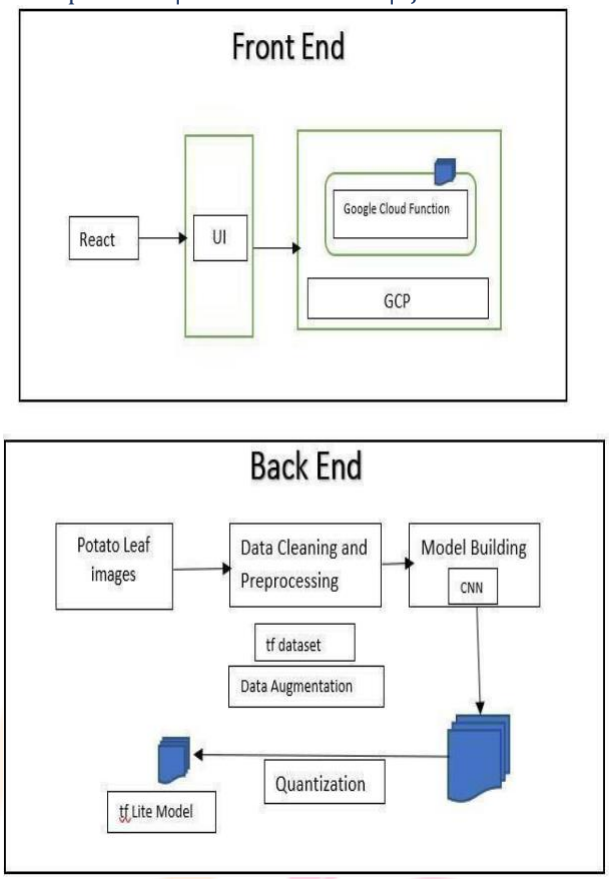

Fig 2.1 System Architecture

4. RESEARCH METHODS

A. Dataset:

The potato leaf disease detection is utilized in the study. It has been downloaded online from Kaggle. The dataset contains various input features, which are broken down into the following categories:

1. Potato healthy leaf
2. Potato early blind
3. Potato late blind

B. Data Collection and Preprocessing:

A collection of methods known as data preparation are used on data to enhance its quality. These methods include addressing missing values, changing the type of feature, and many more. There are two types of process.

1.TF Datasets: A collection of ready-to-use datasets. A Tensor Flow dataset is a group of ready-to-use datasets using Tensor Flow or other Python ML frameworks such as Jax. All datasets are provided as tf. data. Datasets

enabling easyto-use and influential input pipelines. Check out our guide and list of recordings to get started. The tf dataset can be used to transform blurred images.

2.Data augmentation: It is a process of artificially increasing the amount of data by generating new data points from existing data. We can generate new points from the existing points and can generate new type of data from existing data.it helps to predict accurate data. From the existing data and generating new type of data from it we can get zoom in zoom out details from different angles with the help of data.

C. Model building:

1.CNN (Convolution neural network)

To classify more complex images, we need to use more sophisticated artificial neural networks called convolutional neural networks (CNNs). In convolutional neural networks, each neuron in the next layer connects only to groups of neighbouring neurons in the previous layer, usually called local approachable fields or patches. The next neuron in the hidden layer connects to the patch gained by shifting the patch by one pixel. Therefore, each occult neuron learns to analyze its specific local approachable field or patch. instead of all input neurons. Convolutional neural networks are more complex than fully connected networks. Convolutional neural networks differ from other neural networks because of their greater performance on speech, image, or audio signal inputs.

Conclusion

In conclusion, the detection and identification of plant diseases leverage image processing and machine learning algorithms. Data sourced from the Potato Plant Diseases Dataset on Kaggle, comprising over 12,949 images of healthy and unhealthy crop leaves, undergoes segmentation to isolate diseased areas. Utilizing the grayscale co-occurrence matrix (GLCM), key features are extracted, and SVM is employed for classification, yielding an 80% accuracy rate. To enhance precision, convolutional neural networks (CNNs) are employed, achieving an impressive 97.71% accuracy, surpassing traditional hard-coding approaches. This advancement paves the way for automated plant disease identification, fostering early detection and thereby boosting agricultural productivity. Future efforts may explore transfer learning and alternate CNN architectures to further refine potato leaf disease detection accuracy.

REFERENCES

[1] Aditi Singh, Harjeet Kaur"Potato Plant Leaves Disease Detection and Classification using Machine Learning Methodologies [1]" ICCRDA 2020 [2] Tan Soo Xian, Ruzelita Ngadiran, "Plant Diseases Classification using Machine Learning", Hindawi 2021 [3] Jun Liu and Xuewei Wang "Plant diseases and pests detection based on deep learning", (ICoEngTech) 2021 [4] Utkarsha N. Fulari , Rajveer K. Shastri , Anuj N. Fulari, Leaf Disease Detection Using Machine Learning",ResearchGate 2019 [5] Gurleen Kaur Sandhu, Rajbir Kaur, "Plant Disease Detection Techniques: A Review," ICACTM, IEEE, 2019. [6]Md. Selim Hossain, Rokeya Mumtahana Mou, "Recognition and Detection of Tea Leaf's Diseas Using Support Vector Machine," International Colloquium on Signal Processing & its Applications, IEEE, 2018.