

Machine Learning Based Plant Leaf Disease Detection Using CNN

Prof.Dr.Sanjay Nalbalwar¹, Priyanka Nagarale²

^{1,2}Department of electronics and Telecommunication Engineering, Dr. Babasaheb Ambedkar Technical University, Lonere

Abstract - Leaf diseases are a major factor affecting agricultural productivity and crop quality worldwide. Early detection of plant diseases is essential to prevent large-scale losses and improve yield. Traditional disease detection methods rely on manual inspection, which is time-consuming, labour-intensive, and less accurate. This paper presents a machine learning-based plant leaf disease detection system using Convolutional Neural Networks (CNN). The proposed system captures leaf images and processes them using image processing techniques. A trained CNN model is used to classify diseases based on features such as color, texture, and shape. The system is capable of identifying diseases like bacterial blight, rust, and leaf spot with high accuracy. The proposed approach reduces human effort, enables real-time monitoring, and supports precision agriculture.

Key Words: Leaf disease detection, CNN, Machine learning, Image processing, Precision agriculture

1. INTRODUCTION

Agriculture is a big part of the economy, especially in countries like India where many people rely on farming for their livelihood. One of the biggest problems farmers face is plant diseases, which can lower crop production and affect the quality of the harvest. Catching these diseases early is really important to stop them from spreading and to reduce the loss of crops. In the past, identifying plant diseases was done by experts through manual checks, but this method is slow and costly. Now, with the help of artificial intelligence and machine learning, there are better ways to detect plant diseases automatically. Convolutional Neural Networks, or CNNs, are especially good at image classification because they can find detailed patterns in images. In this project, a CNN model is used to accurately and quickly identify plant leaf diseases. This helps farmers act quickly and makes farming more productive overall.

2. LITERATURE REVIEW

Plant Leaf Disease Classification and Detection System Using Machine Learning” This paper is presenting Agricultural intervention in the livelihood of rural India

indulges by about 58%. Among the agricultural products, tomato is one of the most used crops. The application of Artificial Neural Networks (ANN) for plant disease classification has been extensively explored by S. S. Sannakki et al. [1], where a Back Propagation Neural Network (BPNN) is employed. The methodology involves multilayer feed forward architecture consisting of input, hidden, and output layers. Feature vectors derived from infected leaf regions are used to train the network using gradient descent optimization. The error between predicted and actual output is minimized using back propagation, enabling the system to achieve robust classification performance across multiple disease categories.

P. R. Rothe et al. [2] utilized pattern recognition techniques for cotton leaf disease identification. Their system emphasizes statistical feature extraction, including mean, variance, and standard deviation of pixel intensities. Segmentation is achieved using thresholding techniques, and classification is performed using supervised learning models. Their work highlights that feature selection significantly impacts classification accuracy and computational efficiency.

A hybrid model combining computer vision and fuzzy inference systems (FIS) has been proposed by Aakanksha Rastogi et al. [3]. In this approach, uncertainty in disease severity is handled using membership functions and fuzzy rule-based systems. Features extracted from images are mapped to linguistic variables such as “low infection,” “moderate infection,” and “severe infection.” This approach improves interpretability and provides a human-like decision-making mechanism.

Godliver Owomugisha et al. [4] introduced a mobile-based vision system for real-time disease detection in banana plants. The methodology includes image normalization, region segmentation using clustering algorithms, and feature extraction, followed by classification using machine learning techniques. The system is optimized for low computational resources,

making it suitable for deployment on smartphones in rural agricultural environments.

Yuan Tian et al. [5] proposed a Support Vector Machine (SVM)-based multiclass classification framework. SVM works by constructing an optimal hyperplane in a high-dimensional feature space. Kernel functions such as Radial Basis Function (RBF) are used to handle non-linear separability. The study demonstrates that SVM achieves high generalization accuracy and is less prone to overfitting, especially in cases with limited training samples.

A Probabilistic Neural Network (PNN) model has been developed by S. Yun et al. [6], incorporating both image-based features and meteorological parameters such as temperature, humidity, and rainfall. PNN is based on Bayesian classification and uses a radial basis function for probability density estimation. The integration of environmental data significantly enhances prediction accuracy, especially for diseases influenced by climatic conditions.

A comprehensive survey of digital image processing techniques has been conducted by J. G. A. Barbedo [7], focusing on key stages such as preprocessing, segmentation, feature extraction, and classification. The study discusses techniques like median filtering for noise removal, histogram equalization for contrast enhancement, and clustering algorithms for segmentation. It also highlights challenges such as illumination variability, occlusion, and complex backgrounds.

Caglayan et al. [8] proposed a leaf recognition system based on shape descriptors and color histograms. Edge detection methods such as Canny and Sobel operators are used to extract leaf boundaries. Shape features like aspect ratio, perimeter, and area are computed and used for classification. These techniques are crucial for identifying structural abnormalities caused by plant diseases.

Although primarily focused on medical imaging, Zhen et al. [9] introduced regression forest-based models, which are ensemble learning methods combining multiple decision trees. These models are capable of handling large datasets and complex feature interactions. Their adaptability makes them suitable for plant disease severity estimation and prediction tasks.

3. METHODOLOGY

The proposed system uses image processing and CNN-based classification for disease detection. The overall process consists of the following steps:

3.1 Image Acquisition

Leaf images are captured using a digital camera or mobile device.

3.2 Pre-processing

The captured images are enhanced by removing noise and improving contrast.

3.3 Segmentation

The infected regions of the leaf are separated using image segmentation techniques.

3.4 Feature Extraction

Features such as color, texture, and shape are extracted from the segmented image.

3.5 Classification using CNN

The extracted features are fed into a CNN model which classifies the leaf as healthy or diseased.

3.6 Block Diagram

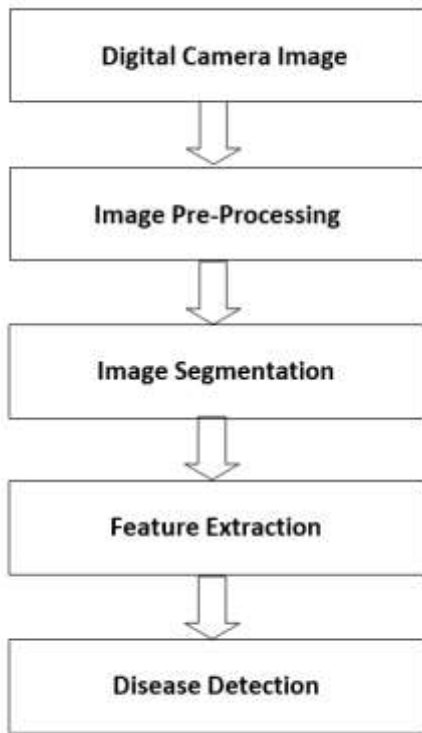


Fig-1: Block Diagram

3.7 CNN (Convolutional Neural Network)

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other

A convolutional neural network is a feed-forward neural network that is generally used to analyze visual images by processing data with grid-like topology. It's also known as a ConvNet. A convolutional neural network is used to detect and classify objects in an image.

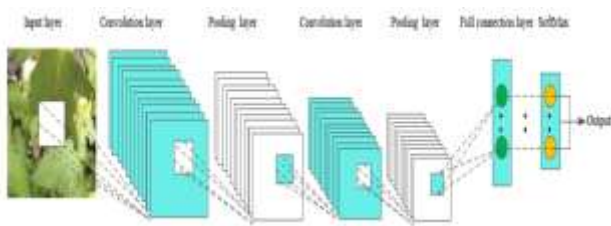


Fig-2: Layers in CNN

3.8 DATASET

The system is trained using the Plant Village dataset, which contains a large number of images of healthy and diseased plant leaves. The dataset includes multiple plant species and disease categories, which helps in improving model accuracy.

The dataset is divided into training and testing sets to evaluate the performance of the model. Data augmentation techniques such as rotation and flipping are applied to improve model generalization.

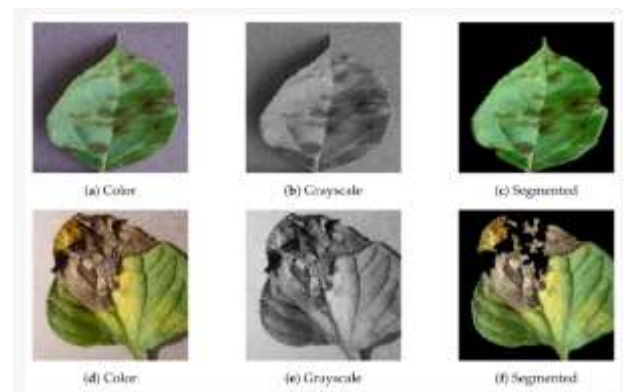


Fig-3: Dataset of Leaf

Table: Detailed description of Plant Village dataset with relative information.

Class	Plant Name	Disease Name	Causes Virus Name	Type of Disease	No of Images
C1	Tomato	Healthy			1591
C2	Tomato	Bacterial spot	Xanthomonas perforans	Bacterial	2127
C3	Tomato	Early Blight	Alternaria sp	Fungal	3000
C4	Tomato	Late blight	Phytophthora infestans	Fungal	1909
C5	Tomato	Leaf Mold	Lycopersicon	Fungal	953
C6	Tomato	Septoria leaf spot	Septoria Lycopersici	Fungal	1773
C7	Tomato	Spider mites	Tetranychus spp	pest	1678
C8	Tomato	Target Spot	Cornespora Cassicola	Fungal	1404
C9	Tomato	Tomato mosaic virus	Tomato mosaic	viral	373
C10	Tomato	Tomato Yellow Leaf	Begomovirus	viral	5357

Table-1: Plant Village Dataset

4. RESULTS AND DISCUSSION

The proposed system is tested using leaf images, and the CNN model successfully identifies plant diseases with good accuracy.

The developed GUI-based system takes an image as input and predicts the plant type and disease name.



Fig-4: Cotton



Fig-5: Tomato

5. ADVANTAGES

- Early detection of plant diseases
- High accuracy using CNN
- Reduces manual effort
- Suitable for large-scale farming
- Supports real-time monitoring

6. FUTURE SCOPE

The system can be further improved by integrating IoT sensors for real-time environmental monitoring. Mobile applications can be developed to make the system easily accessible to farmers. Advanced deep learning models can be used to improve accuracy further.

7. CONCLUSIONS

The system developed in this work focuses on detecting plant leaf diseases using a CNN-based approach. Compared to traditional methods, where farmers depend on manual inspection, this system makes the process faster and more practical. It helps in identifying diseases

at an early stage, which is important to avoid major crop losses and maintain good yield. The results show that the model is able to classify leaf diseases with satisfactory accuracy. It also reduces the need for continuous expert supervision, which can be difficult in large farming areas. This makes the system useful for real-world agricultural applications. IN the future, this work can be improved by connecting it with real-time monitoring systems or mobile applications so that farmers can easily use it in their daily practices. With such improvements, the system can become a part of modern smart farming solutions.

REFERENCES

1. S. S. Sannakki and V. S. Rajpurohit, "Classification of Pomegranate Diseases Based on Back Propagation Neural Network," International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015
2. P. R. Rothe and R. V. Kshirsagar, "Cotton Leaf Disease Identification using Pattern Recognition Techniques", International Conference on Pervasive Computing (ICPC),2015.
3. Aakanksha Rastogi, Ritika Arora and Shanu Sharma, "Leaf Disease Detection and Grading using Computer Vision Technology &Fuzzy Logic" 2nd International Conference on Signal Processing and Integrated Networks (SPIN)2015.
4. Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, "Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease ", Preceding of the 1'st international conference on the use of mobile ICT in Africa ,2014.
5. uan Tian, Chunjiang Zhao, Shenglian Lu and Xinyu Guo, " SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases," Proceedings of 2010 Conference on Dependable Computing (CDC'2010), November 20-22, 2010.
6. S. Yun, W. Xianfeng, Z. Shanwen, and Z. Chuanlei, "Pnn based crop disease recognition with leaf image features and meteorological data," International Journal of Agricultural and Biological Engineering, vol. 8, no. 4, p. 60, 2015.
7. J. G. A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying

plant diseases,” Springer Plus, vol. 2, no.660, pp. 1–12, 2013.

8. Caglayan, A., Guclu, O., & Can, A. B. (2013, September). “A plant recognition approach using shape and color features in leaf images.” In International Conference on Image Analysis and Processing (pp. 161-170). Springer, Berlin, Heidelberg.

9. Zhen, X., Wang, Z., Islam, A., Chan, I., Li, S., 2014d. “Direct estimation of cardiac bi-ventricular volumes with regression forests.” In: Accepted by Medical Image Computing and Computer-Assisted Intervention– MICCAI 2014.

[10] Wang P., Chen K., Yao L., Hu B., Wu X., Zhang J., et al. (2016).” Multimodal classification of mild cognitive impairment based on partial least squares