Plant Disease Prediction System Using Deep Learning

Suyash Dhake[1] , Jeet Trivedi[2] , Manav Sachdeva[3] , Asst. Prof. Ms. Monali Parikh [4]

[123]*Research Scholar, Institute of Information Technology, Krishna School Of Emerging Technology & Applied Research,* KPGU University , *Varnama, Vadodara, Gujarat, India*

[4] *Assistant Professor, Department of Information Technology and Engineering, Krishna School Of Emerging Technology & Applied Research,* KPGU University. *Varnama, Vadodara, Gujarat, India*

---***---

Abstract - Plant diseases significantly threaten global food security, resulting in substantial agricultural losses annually. Early detection and accurate classification of plant diseases can help farmers take timely preventive measures, thereby mitigating crop damage and improving yield. This research aims to develop a machine learning-based model on top of EfficientNetB3 for detecting and classifying plant diseases from leaf images. The proposed model has achieved 99.4% testing accuracy. This study is focused on a deep learning model to detect disease in plant leaves. But, in the future model can be integrated with a drone or any other system to detect diseases from plants and report the diseased plant's location to people so that they can be cured accordingly.

***Key Words*:** Plant Disease, Convolution Neural Network (CNN), Deep Learning, Agriculture, and PlantVillage

1.INTRODUCTION

Agriculture is the backbone of the global economy, providing food, raw materials, and employment for a significant portion of the world's population. However, one of the most pressing challenges in agriculture today is crop loss due to plant diseases, which can devastate yields and lead to severe economic and food supply issues. Effective disease management relies on early and accurate diagnosis, allowing farmers to take timely action to protect their crops.

Traditionally, disease identification has relied on expert knowledge, which can be time-consuming, costly, and inaccessible for farmers in remote or underserved areas. Thus, there is a growing demand for technology-driven solutions that can detect plant diseases with precision and speed.

Recent advancements in machine learning and image processing present promising avenues for addressing this challenge. Machine learning algorithms, particularly Convolutional Neural Networks (CNNs), excel in recognizing image patterns, making them well-suited for identifying disease symptoms on plant leaves. By developing an automated plant disease detection system, this research aims to leverage these technologies to create an accessible, cost-effective solution for early disease diagnosis.

Some of the common symptoms of disease in plants disease are Leaf rust (common leaf rust in corn), Stem rust (wheat stem rust), Sclerotinia (white mold), Powdery mildew, Birds-eye spot on berries (anthracnose), Damping off of seedlings (phytophthora), Leaf spot (septoria brown spot), Chlorosis (yellowing of leaves). These diseases can be identified by the physical condition of plant leaves. The experts can identify whether the plant is defective or not by looking at leaves stems or fruit. This approach requires having lots of human resources for this particular job.

In this era of technology and automation, it is not a very efficient approach, it would be much better if we had an automated system that detects disease in plants automatically. Many researches are already done to fill this purpose most of them utilize traditional machine learning approaches [].

This research investigates the development of a machine learning-based plant disease detection model trained on an extensive dataset of plant leaf images named PlantVillage.

2 Literature Summary

Sr No	Paper Author/ Title/Year Of Publication	Method	Data set	Limitations	Future Scope
1	Garima Shrestha , Deepsikha, Majolica Das & Naivrita Dey Plant Disease Detection Using CNN [July 2020]	CNN, Image Net	a dataset of labeled plant leaf images.	Lack of real time processing	Predict the disease more accurately
2	Yan Guo, Jin Zhang, Chengxin Yin, Xiaonan Hu, Yu Zou, Zhipeng Xue, Wei Wang **Plant Disease Identification Based on Deep Learning Algorithm in Smart Farming** 18 August 2020	**Region Proposal Network (RPN)**	Data sets of labelled plant leaf images	Simple backgrounds and transfer learning limit effectiveness and adaptability.	Dataset expansion, advanced algorithms, real-time detection, novel diseases, field tests, apps.
3	Naresh Trivedi, Vinay Gautam, Abhineet Anand, Hani Moaiteq Aljahdali, Santos Gracia Villar, Divya Anand, Nitin Goyal, Seifedine Kadry Erly Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network 30 November 2021	Convolutional Neural Network (CNN)	Data sets of labelled plant leaf images	More time was required to process the images because of its complex computations	Accuracy can be increased with help of more subsequent technology
4	LILI LI, Shujuan Zhang, and Bin Wang Plant Disease Detection and Classification by Deep Learning	Unsupervised classification, hybrid models (2D-CNN-BidLSTM/GRU), supervised 3D-CNN, and ELM	P1antVillage, Plant Pathology Challenge, author-collected images,	Resource-intensive, small samples, adjustable parameters, metric learning, data generation	Deep learning, data augmentation, transfer learning, hyperspectral
5	MH Saleem, J Potgieter, KM Arif Plant disease detection and classification by	Deep Learning, Convolutional Neural Network (CNN)	The Plant Village dataset was used	Lacks recent visualization techniques and modified deep learning models	Enhance deep learning, improve disease detection,

	deep learning **31 October 2019**			for plant disease identification.	integrate advanced imaging technologies.
6	F Mohameth, C Bingcai, KA Sada Plant disease detection with deep learning and feature extraction using plant village June 2020	Deep Learning, Convolutional Neural Network (CNN)	Plant Village Dataset was used	Limited architecture comparisons, database reliance, and real-time processing challenges.	Collect diverse data, improve detection accuracy with new datasets, and explore classifiers and architectures.
7	Sammy V. Militante1 , Bobby D. Gerardoij, Nanette V. DionisioĴ Plant Leaf Detection and Disease Recognition using Deep Learning2019	Image acquisition, pre-processing, and classification using CNN to detect plant diseases.	Images from plant village dataset	Limited to 32 varieties and diseases;	Expand dataset, explore CNN architectures, optimize performance.
8	Sharada p Mohanty, David p Hughes, Marcel Salathe Using Deep Learning for Image-Based Plant Disease Detection 22 September 2016	Convolutional Neural Networks (CNNs)	AlexNet and GoogleNet were used to create ImageNet dataset	Limited data access, no real-time updates, cannot browse the internet	Enhanced AI, improved data integration, real-time processing
9	Maummer Turkoglu, Davut Hanbay Plant disease and pest detection using deep learning-based features 2019	Support vector machine(svm), Extreme learning machine(elm), K-nearest neighbour(knn)	AlexNet, VGG16, and VGG19	Data biases, small sample size, limited generalizability.	Explore methodologies, expand diversity, apply findings.
10	M Chohan, A Khan, R Chohan, SH Katpar **Plant disease detection using deep learning** May 2020	Convolutional Neural Networks (CNNs), Artificial Neural Network (ANN)	PlantVillage dataset	Data dependence, high resource needs, overfitting, low interpretability.	Accuracy improvement, real-time detection, disease expansion, IoT integration, collaboration.

3. Materials

The following materials were used in the development of the Plant Disease Detection:

- **Programming Language**: Python 3. x
- **Development Environment**: Anaconda or Jupyter Notebook (for experimentation and initial development)
- **Libraries**:
- **OpenCV (cv2)**: For image processing tasks.
- **Pandas**: For data manipulation and analysis.
- **Seaborn**: A data visualization library based on Matplotlib.
- **Matplotlib**: For data plotting.
- **Scikit-learn (sklearn)**: For data preprocessing, model training, and evaluation
- **TensorFlow**: For building and deploying deep learning models.
- **Keras**: For experimentation with deep learning model.
- **Numpy**: For handling numerical data and arrays during image processing.

Deep learning is a powerful machine learning approach that has mitigated the traditional machine learning headache of feature engineering. It doesn't need any domain expertise now and all credit goes to deep learning. The core of deep learning is the artificial neural network (ANN). Artificial neural networks are mathematical models that replicate with their neurons and synapses interconnecting them with the general principles of brain function.

A. Convolutional Neural Network:

Convolutional Neural Networks (CNNs) have become the preferred machine learning architecture for image-based plant disease detection due to their ability to automatically extract and learn complex patterns from images. A CNN is a type of deep learning model specifically designed to handle image data, making it ideal for analyzing visual symptoms of plant diseases. Convolution is used to detect edges of patterns in an image and pooling is used to reduce the size of an image. We have developed a model on top of the EfficientNetb3 (a convolutional neural network built upon a concept called compound scaling). Moreover, training of this model is done using Jupyter Notebook and Keras API of TensorFlow.

B. Dataset Discussion

We have used a PlantVillage dataset with over 50,000 expertly curated images of healthy and infected leaves of crop plants. PlantVillage consists of a total of 38 classes of which there are 26 classes are diseased plants and 12 are healthy plants. It is an open-source dataset. A description of these classes and dataset is given in the following table:

Plant Name	Healthy/Diseased	Disease Name	Number of Images
Apple	Diseased	Apple Scab	630
Apple	Diseased	Black Rot	621
Apple	Healthy	-	1645
Blueberry	Diseased	Powdery Mildew	1502
Cherry	Diseased	Powdery Mildew	1052
Cherry	Diseased	Leaf Spot	854
Cherry	Healthy	-	234
Corn (Maize)	Diseased	Cercospora Leaf Spot, Gray Leaf Spot	513
Corn (Maize)	Diseased	Common Rust	1192
Corn (Maize)	Diseased	Northern Leaf Blight	985
Corn (Maize)	Healthy	-	1162
Grape	Diseased	Black Rot	1180
Grape	Diseased	Esca (Black Measles)	1383
Grape	Diseased	Leaf Blight	1076
Grape	Healthy	-	423
Orange	Diseased	Huanglongbing (Citrus Greening)	5507
Peach	Diseased	Bacterial Spot	2297
Peach	Healthy	-	380
Pepper (Bell)	Diseased	Bacterial Spot	997
Pepper (Bell)	Healthy	-	1470
Potato	Diseased	Early Blight	1000
Potato	Diseased	Late Blight	1909
Potato	Healthy	-	152
Raspberry	Diseased	Leaf Spot	371
Soybean	Healthy	-	5090
Squash	Diseased	Powdery Mildew	1835
Strawberry	Diseased	Leaf Scorch	1109
Strawberry	Healthy	-	456
Tomato	Diseased	Bacterial Spot	2127
Tomato	Diseased	Early Blight	1000

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Tomato	Diseased	Late Blight	1909
Tomato	Diseased	Leaf Mold	952
Tomato	Diseased	Septoria Leaf Spot	1771
Tomato	Diseased	Spider Mites	1676
Tomato	Diseased	Target Spot	1404
Tomato	Diseased	Tomato Mosaic Virus	373
Tomato	Diseased	Tomato Yellow Leaf Curl Virus	5357
Tomato	Healthy	-	1591

Sample Images from PlantVillage Dataset

The dataset is divided into two parts one for training and the other for Testing. Splitting of the dataset is an 80/20 ratio randomly. 80% for the training dataset and the rest 20% for the testing dataset.

C. Model Description

We have developed a CNN which has 5 layers.

EfficientNetB3 (1st layer) – a pre-trained model used for efficient training

Batch Normalization (2nd layer) - Batch Normalization is used to scale data on a particular scale.

Dense (256 units) (3rd layer) – Reduces/ compresses the vector output from the previous layers to 256.

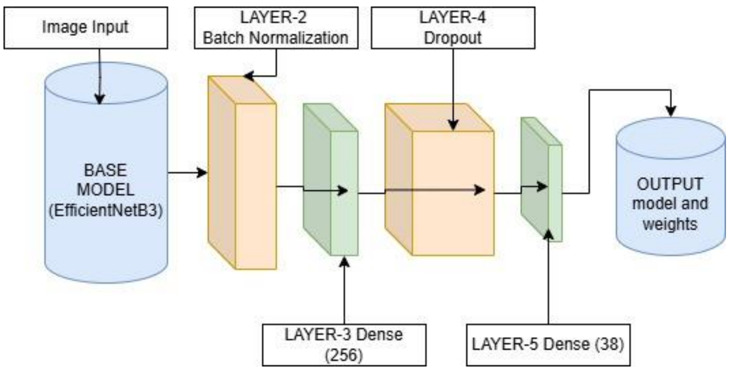
Dropout: (4th layer) - added to prevent any overfitting problem with the training.

Dense_1 (38 units): - (5th layer) - Outputs a vector of length 38 (using the input from the previous layer - to perform classification as the remaining 38 neurons will determine which input belongs to which class of the image provided (example - Apple_Healthy)).

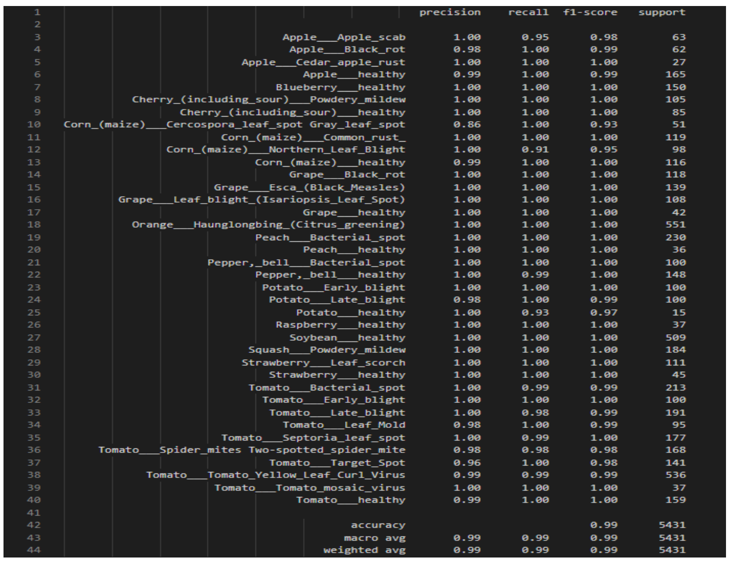
CNN training Parameters:

Parameter	Value
Epochs	40
Batch Size	40
Learning Rate	1e-3
Activation in the middle Layer	Relu
Activation in Final Layer	Softmax

Applied Methodology



Confusion Matrix



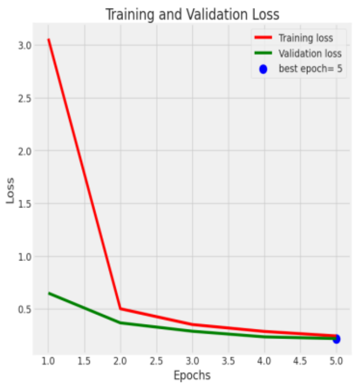
	precision	recall	f1-score	support
Apple___Apple_scab	1.00	0.95	0.98	63
Apple___Black_rot	0.98	1.00	0.99	62
Apple___Cedar_apple_rust	1.00	1.00	1.00	27
Apple___healthy	0.99	1.00	0.99	165
Blueberry___healthy	1.00	1.00	1.00	150
Cherry_(including_sour)___Powdery_mildew	1.00	1.00	1.00	105
Cherry_(including_sour)___healthy	1.00	1.00	1.00	85
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot	0.86	1.00	0.93	51
Corn_(maize)___Common_rust_	1.00	1.00	1.00	119
Corn_(maize)___Northern_Leaf_Blight	1.00	0.91	0.95	98
Corn_(maize)___healthy	0.99	1.00	1.00	116
Grape___Black_rot	1.00	1.00	1.00	118
Grape___Esca_(Black_Measles)	1.00	1.00	1.00	139
Grape___Leaf_blight_(Isariopsis_Leaf_Spot)	1.00	1.00	1.00	108
Grape___healthy	1.00	1.00	1.00	42
Orange___Haunglongbing_(Citrus_greening)	1.00	1.00	1.00	551
Peach___Bacterial_spot	1.00	1.00	1.00	230
Peach___healthy	1.00	1.00	1.00	36
Pepper,_bell___Bacterial_spot	1.00	1.00	1.00	100
Pepper,_bell___healthy	1.00	0.99	1.00	148
Potato___Early_blight	1.00	1.00	1.00	100
Potato___Late_blight	0.98	1.00	0.99	100
Potato___healthy	1.00	0.93	0.97	15
Raspberry___healthy	1.00	1.00	1.00	37
Soybean___healthy	1.00	1.00	1.00	509
Squash___Powdery_mildew	1.00	1.00	1.00	184
Strawberry___Leaf_scorch	1.00	1.00	1.00	111
Strawberry___healthy	1.00	1.00	1.00	45
Tomato___Bacterial_spot	1.00	0.99	0.99	213
Tomato___Early_blight	1.00	1.00	1.00	100
Tomato___Late_blight	1.00	0.98	0.99	191
Tomato___Leaf_Mold	0.98	1.00	0.99	95
Tomato___Septoria_leaf_spot	1.00	0.99	1.00	177
Tomato___Spider_mites Two-spotted_spider_mite	0.98	0.98	0.98	168
Tomato___Target_Spot	0.96	1.00	0.98	141
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.99	0.99	0.99	536
Tomato___Tomato_mosaic_virus	1.00	1.00	1.00	37
Tomato___healthy	0.99	1.00	1.00	159
accuracy			0.99	5431
macro avg	0.99	0.99	0.99	5431
weighted avg	0.99	0.99	0.99	5431

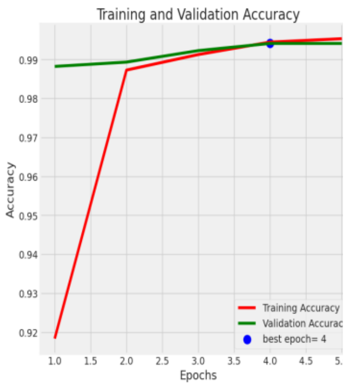
Performance Metrics:

Accuracy: 99.44%
Precision: 99.25%
Recall: 99.25%
F1-Score: 99.3%

D. Results

The results of the plant disease detection model demonstrate high accuracy in classifying plant diseases across a variety of crop types. Using a deep learning-based Convolutional Neural Network (CNN), the model achieved an overall accuracy of 99.44% in identifying both healthy and diseased plant leaves. Precision, recall, and F1-scores for most classes are near perfect, reflecting the model's ability to correctly distinguish between diseases with minimal false positives and false negatives.





5. CONCLUSION & Future Work

In conclusion, this study demonstrated the effectiveness of Convolutional Neural Networks (CNNs) in accurately detecting and classifying plant diseases. The system has achieved an overall 99.4% testing accuracy on publically accessible dataset. This system can be integrated into mini-drones to live detection of diseases from plants in cultivated areas. In the future, this system can also adopt 3 layer approach where the first layer detects if there's any plant in an image or not, the second layer tells the plant type and disease and the third layer tells the possible solution or cure for that particular plant disease.

REFERENCES

1 Garima Shreshtha, Deepsikha, Majolica Das & Naivrita Dey
Plant Disease Detection Using CNN
[July 2020]

2 YanGuo, Jin Zhang, Chengxin Yin, Xiaonan Hu, Yu Zou, Zhipeng Xue, Wei Wang
Plant Disease Identification Based on Deep Learning Algorithm in Smart Farming
[2020]

3 Naresh Trivedi, Vinay Gautam, Abhineet Anand, Hani Moaiteq, Aljahdali, Santos Garcia Villar, Divya Anand, Nitin Goyal, Seifedine Kadry
Early Detection and Classification of Tomato Leaf Disease Using High-Performance Deep Neural Network[2021]

4 SB Jadhav, VR Udupi, SB Patil
Identification of plant diseases using convolutional neural networks [2020]

5 MH Saleem, J Potgieter, KM Arif
Plant disease detection and classification by deep learning [May 2019]

6 F Mohameth, C Bingcai, KA Sada
Plant disease detection with deep learning and feature extraction using Plant Village [May 2020]

7 Sharada P. Mohanty, David P. Hughes, Marcel Salathe
Using Deep Learning for Image-Based Plant Disease Detection [2016]

8 Konstantinos P. Ferentinos
Deep learning models for plant disease detection and diagnosis [2018]

9 Maummer Turkoglu, Davut Hanbay
Plant disease and pest detection using deep learning-based features [2019]

10 M Chohan, A Khan, R Chohan, SH Katpar
Plant disease detection using deep learning [2020]