INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 09 | SEPT - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Plant Leaf Disease Prediction Using Transfer Learning

HarisAhmed I Nidagundi
Department of Computer Science and Engineering
PES University
Bengaluru, India
harrisnidagundi@gmail.com

Mohammed Ismail I
Department of Computer Science and Engineering
PES University
Bengaluru, India
m.d.ismailirshad008@gmail.com

Zafeer Ahmed
Department of Computer Science and Engineering
PES University
Bengaluru, India
emailzaff@protonmail.ch

Sneha Sujit Saha
Department of Computer Science and Engineering
PES University
Bengaluru, India

Prof. Sreenath MV
Department of Computer Science and Engineering
PES University
Bengaluru, India

***Abstract*—According to a poll conducted in 2018, agriculture employed around 80% of the labour force in India and generated between 16 and 18% of the GDP. It is regarded as the foundation of the Indian economy. A lot of work and research is being done to automate agricultural tasks, and with the development of technologies like Deep Learning, one of those tasks is the prediction of leaf disease. Identification of leaf disease is crucial for preventing the loss in the yield and quantity of the agricultural product. Machine learning and Image Processing techniques are used for detection of plant diseases. The strategies that already exist have limitations when it comes to the ability to adapt and reuse the results of learning. As a result, we employ the *transfer learning* approach to enhance performance and shorten the training period. We collected PlantVillage dataset from Kaggle. Images of 15 different plant leaf classes from the three different plants potato, pepper, and tomato are shown. The dataset was split into three separate datasets. The model provides a maximum accuracy of 99%. The final training accuracy was 87% and the final validation accuracy was 99%, with low training and validation loss. The experimental results showed that our model achieved a good accuracy rate for plant leaf disease prediction.**

***Keywords—Plant Disease, Machine Learning, Transfer Learning, Kaggle, Python*.**

I. INTRODUCTION

The importance of agriculture to a nation's economy depends on the quality of the produce, which in turn depends on the quality of the plants. The mainstay of India's economy is agriculture, which produces the majority of the nation's exports. For ages, farmers and experts have used their unaided eyes to identify illnesses on plant leaves. Some Indian farmers lack the necessary field expertise and information needed for this strategy. Additionally, due to a lack of money, they are unable to recruit such specialists to determine the type of illness, which results in subpar disease catering and decreased agricultural output. Early detection and management of plant diseases are essential for a healthy crop yield and increased agricultural profitability. The key to reducing losses in the yield and quantity of the agricultural output is the identification of the plant disease. The research on plant disease focuses on the patterns on the plant that can be observed with the naked eye. A vital component of sustainable agriculture is the monitoring of the plant's health and the detection of disease. The monitoring of plant diseases is extremely challenging and takes a lot of time. As a result, methods for machine learning and image processing are utilized to identify plant diseases. Image acquisition, picture pre-processing, image segmentation, feature extraction, and classification are processes in the disease detection process. Climate change and sustainable agriculture are both issues that are directly tied to the issue of effective plant disease protection. A plant pathologist should develop superior observational skills so that they can recognize distinctive characteristics in order to diagnose plant diseases accurately. Artificial neural networks and Support Vector Machines are currently the most widely used methods for identifying plant disease. The study attempts to correctly identify the disease by examining the pattern on the plant's leaves using deep learning algorithms like convolutional neural networks. The primary goal of the project is to identify leaf diseases in plants like pepper, potato, and tomato by using a leaf disease image as an input and an output that includes the illness's name and recommended treatments. The project's objective is to increase the prediction model's level of accuracy. In this paper, we considered three different plant species. The Kaggle website hosted the dataset. It has 10 different tomato leaf varieties, 3 different potato leaf varieties, and 2 different pepper leaf varieties. The sample images in the dataset (for Tomato, Potato, Pepper Bell leaves) are shown in Fig.1, Fig.2, Fig.3.

VOLUME: 08 ISSUE: 09 | SEPT - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Fig. 1. Tomato Spot, Leaf Mold, Yellow leaf Curl Virus, Mosaic Virus, Bacterial Spot, Leaf Spot, Early Blight, Late Blight, Spider Mite.

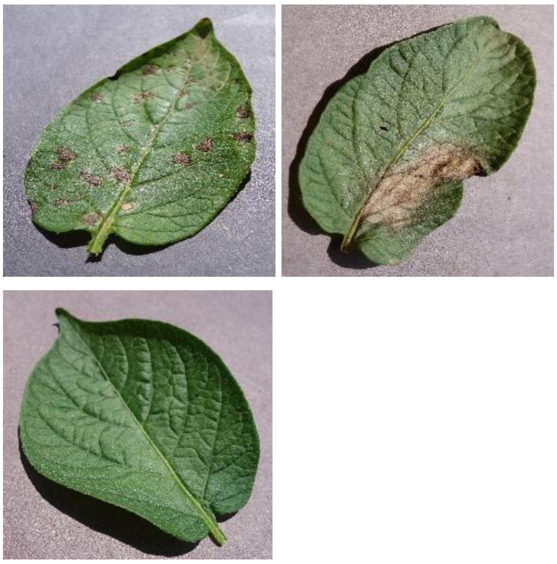

Fig. 2. Potato Early Blight, Late Blight, Healthy

Fig. 3. Bacterial Spot, Pepper Bell Healthy

II. REALATED WORK

The use of ML/DL in agriculture is crucial for increasing national output. ML and DL algorithms were used by several studies to detect plant leaf disease. P. Rubini .Et Al [1] has worked on Tomato Plant Disease and developed a machine learning model that identifies whether a plant is diseased or healthy. The author worked with CNN VGG16 and DenseNet using transfer learning and their respective accuracy was calculated. The result described that CNN showed accuracy of around 97%. A. Lakshmanrao .Et Al [2] has used some traditional methods for identifying plant disease. The author worked on Tomato and Potato Plants and have applied ConvNets on the respective datasets. The author worked with CNN and Random Forest and their respective accuracy was calculated. The result described that Random Forest showed accuracy of around 97% and CNN showed accuracy of around 98%. Jingyao Zhang .Et Al [3] developed a novel identification approach of cucumber leaf diseases based on small sample size and deep convolutional neural network. Experimental results showed that the proposed approach achieved the average identification accuracy of 96.11% and 90.67% when implemented on the datasets of lesion and raw field diseased leaf images with three different diseases of anthracnose, downy mildew and powdery mildew. J. Liu .Et Al [4] has studied powdery mildew, bacterial leaf spot, black spot and downy mildew of sunflower leaves. The author worked with K-means and Random Forest algorithms. According to the test results, the overall accuracy rate of four common sunflower diseases is 95%, which has a practical application level. P. Sharma .Et Al [5] have introduced the artificial intelligence based automatic plant leaf disease detection and classification for quick and easy detection of disease and then classifying it and performing required remedies to cure that disease. The author used several techniques which includes KNN, SVM, CNN and Logistic Regression. The result described that KNN, SVM, CNN and Logistic Regression showed accuracy of around 54%, 53%, 97% and 66% respectively.

III. PROPOSED METHODOLOGY

First, we downloaded a dataset from Kaggle called Plant Village. The dataset includes photos of plant leaves from three different plants—tomato, pepper, and potato—in 15 different categories. Table 1 displays the dataset's specifics. The dataset was then split into three separate datasets for the Potato, Pepper, and Tomato datasets. After that, we applied Convnets (Convolutional Neural Network) for plant leaf disease detection. The proposed framework is shown in Figure-4. The purpose of the proposed system is to detect the plant leaf diseases using the methods of feature extraction (features=shape, color and texture), CNN a machine learning algorithm is used for classification of the leaf which is followed by transfer learning to decrease the training time and improve the efficiency and performance.

TABLE 1. DETAILS OF DATASET

Class	Disease Name	Count of Images
Potato	Healthy	152
	Late-Blight	1,000
	Early-Blight	1,000
Bell Pepper	Healthy	1,478
	Bacterial Spot	997
Tomato	Healthy	1,591
	Septorial-Leaf-Spot	1,771
	Mosaic-virus	373
	Leaf-Mold	952
	Late-Blight	1,909
	Bacterial-Spot	2,127
	Spider-Mite	1,676
	Target-Spot	1,404
	Yellow Leaf-Curl-Virus	3,209
	Early-Blight	1,000

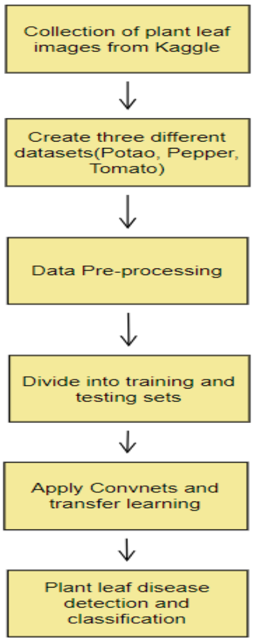

Fig. 4. Proposed Framework for Plant Disease Detection

A. Data Pre-processing

The dataset exhibits homogeneity if all of the photographs are represented by files of the same size. So, before applying CNNs, we did some pre-processing which includes Image Augmentation, Image Segmentation.

Image Augmentation: It is a method of modifying the current data to produce new data for the model training procedure. In other words, it is the process of enhancing the dataset that is made available for deep learning model training.

Image Segmentation: It is a technique that divides a digital image into a number of smaller groupings known as image segments, which aids in bringing down the complexity of the image to make subsequent processing or analysis of the image easier. In plain English, segmentation is the process of giving pixels labels. Each pixel or piece of a picture assigned to the same category has a unique label.

B. ConvNets

Convolutional neural networks, often known as Convnets, are the most practical deep learning model and have gained popularity in recent years as the primary tool for image categorization issues. They resemble the forward and backward propagating feedforward neural network model. Each neuron receives information, carries out some mathematical operations, adds nonliteral plays out a spot item, and alternately follows it with a non-linearity. One distinct single function is communicated throughout the entire network. It uses the pixels from the image as input and outputs class label scores.

The following steps are used in ConvNets:

Step 1: Convolutional Operation: This stage involves finding the Feature Map using a Feature Detector.

Step 2: Activation Function: Activation function is used following Convolution operations.

Step 3: Pooling: Images are down sampled and shrunk in size in this process.

Step 4: Fully Connected Layer: This is similar to multilayer perceptron.

C. Transfer Learning

In general, the term "transfer learning" refers to a process where a model developed for one problem is used in some capacity to another similar problem. Transfer learning offers the advantage of cutting down on neural network model training time and potentially lowering generalization error. We have used ResNet9 transfer learning model in our paper. A state-of-the-art picture categorization model that can be used is Resnet, a convolutional neural network. The ImageNet dataset, a sizable classification dataset, served as pre-training data for the Resnet models that we will utilise in this lesson.

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 09 | SEPT - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

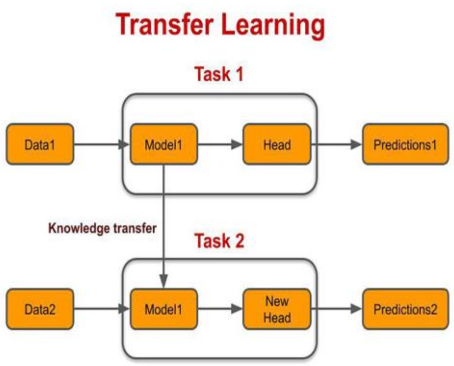

Fig. 5. Transfer Learning

The Residual Blocks idea was created by this design to address the issue of the vanishing/exploding gradient. We apply a method known as skip connections in this network.

The skip connection bypasses some levels in between to link layer activations to subsequent layers. This creates a leftover block. These leftover blocks are stacked to create resnets.

IV. EXPERIMENTATION

The dataset contains images of Tomato, Potato and Pepper plant leaves.

Splitting into Training and Testing Sets

The dataset is divided into training and testing sets based on 80:20 ratio. The details of the division is shown in the Table-2.

TABLE 2. DETAILS OF DIVISION OF DATASET

Dataset	No. of Images
Total Images	20,639
Training Set	16,511
Testing Set	4,128

After dividing into Training and Testing sets, we build a CNN model. As there is no rule of thumb for selecting parameters for deep learning model. We conducted several experiments with the dataset. The hidden layer activation function is RELU. Relu is well suited for hidden layers. The accuracy achieved for the dataset is 98.3%. The output function used is SoftMax. The final training accuracy was 87% and the final validation accuracy was 99%, with low training and validation loss.

V. RESULTS

The model used Convolution Neural Network and Transfer Learning Model ResNet9, which is trained using PlantVillage Dataset. Fig. 7. Shows the Training and Validation Loss of the model.

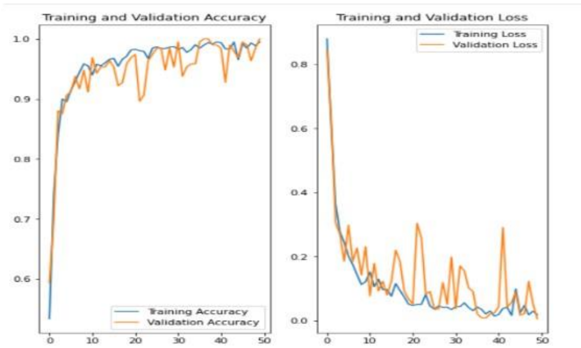

Fig. 7. Accuracy, Training and Validation Loss of the model

COMPARISON OF ACCURACIES WITH OTHER MODELS

Paper	Dataset	Model	Accuracy
P. Rubini	Tomato	CNN VGG16	97%
A. Lakshmanrao	Tomato, Potato	CNN, Random Forest	98%, 97% respectively
Jingyao Zhang	Cucumber	Deep CNN	96.11%
J. Liu	Sunflower	k-means, Random Forest	95%
This Paper	Tomato, Pepper, Potato	Resnet9	99%

VI. CONCLUDSION

The project represents a simple and efficient way of predicting leaf disease with increased accuracy and performance in comparison to other algorithms. The model provides a maximum accuracy of 99%. The final training accuracy was 87% and the final validation accuracy was 99%, with low training and validation loss. In the future, we will focus on estimating the severity of the defected disease and we will test several datasets with our model.

REFERENCES

[1] Sethy, P.K.; Barpanda, N.K.; Rath, A.K.; Behera, S.K. Deep feature based rice leaf disease identification using support vector machine. Comput. Electron. Agric. 2020, 175, 105527.

[2] Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y.A. Using deep transfer learning for image-based plant disease identification. Comput. Electron. Agric. 2020, 173, 105393

[3] Bai, X.; Cao, Z.; Zhao, L.; Zhang, J.; Lv, C.; Li, C.; Xie, J. Rice heading stage automatic observation by multi-classifier cascade based rice spike detection method. Agric. For. Meteorol. 2018, 259, 260–270.

[4] Lee, S.H.; Goëau, H.; Bonnet, P.; Joly, A. New perspectives on plant disease characterization based on deep learning. Comput. Electron. Agric. 2020, 170, 105220.

[5] Too, E.C.; Yujian, L.; Njuki, S.; Yingchun, L. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. 2021, 161, 272–279.

[6] Babu, M.P.; Rao, B.S. Leaves Recognition Using Back Propagation Neural Network-Advice for Pest and Disease Control on Crops.

[7] Shima Ramesh Maniyath, Shima Ramesh Maniyath," Plant Disease Detection Using Machine Learning", 2019 International Conference on Design Innovations for 3Cs Compute Communicate Control.

[8] Pushkara Sharma, Pankaj Hans, Subhash Chand Gupta, “CLASSIFICATION OF PLANT LEAF DISEASES USING MACHINE LEARNING AND IMAGE PREPROCESSING TECHNIQUES”,2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)

[9] Jun Liu, Fang Lv, Penghui Di, Identification of sunflower leaf diseases based on random forest algorithm, 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS)

[10] Hiroki TANI, Ryunosuke KOTANI, Satoshi KAGIWADA, Hiroyuki UGA, Hitoshi IYATOMI, “Diagnosis of Multiple Cucumber Infections with Convolution Neural Networks”, 2021 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)

[11] Herlambang Dwi Prasetyo, Hendi Triatmoko, Nurdiansyah, Ika Nurlaili Isnainiyah, “The Implementation of CNN on Website-based Rice Plant Disease Detection”, 2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)