Vegetable Leaves Disease Prediction using Deep Learning

ASHISH KUMAR[1], DIVIJ BHATIA[2], SAURAV KHAN[3], VIKASH KUMAR[4]
[1, 2, 3] *CSE, Maharaja Agrasen Institute of Technology, Delhi*
[4] *CSE, Jain University – FET, Bangalore*

Abstract- ***The population across the world are increasing rapidly which leads to the high demand of food but the supply chain of food is not enough to meet the requirements. This issue leads to high production of food but one of the major problem arises in agricultural sector is plant disease that is the major threat to farmers. By early detection of these plants disease can help farmers to prevent the plant wastage and also prevent their financial losses. In this project, we train a Deep Learning CNN model to detect different plant diseases that are affecting farmers adversely. Farmers will be able to detect these plant diseases at early stage of agricultural production and can take various measures to prevent these diseases. We have used dataset of 17,055 images of 5 different plant categories in 15 classes including healthy and unhealthy plant leaves. Our model is achieves an accuracy of 93.22% upon training these datasets. We are also proposing the web application which will detect the disease of these 5 different plant categories for 15 classes.***

I. INTRODUCTION

With the increasing population growth over recent decades across the world leads to high demand of agricultural worldly goods, which results in large expansion of farming. In order to meet these high demands for food, biomass and several other products, the production of crop must be increasing amid these rising population by 2050to its double speed. This can only be achieved if crop production enhance by 2.4% each year but it is hardly able to reach up to 1.3% per year.

In India, Agriculture is plays a vital role in contribution of economic development. Production of suitable crops are decided by the type of soil, Fertilizers, location, weather condition and economic value. But changing nature of soil, weather condition force the researchers to start looking for new technologies to overcome these challenges so that high productivity of crop production can be achieved. Several other reasons for the crops destruction are Pests and unavailability of water for irrigation at proper time that results in diseased crop. These reasons resulting the less productivity and diminish of quality crop yield.

In recent times, several countries came up with various modern technologies that reduce the harvesting time and also maximize the production of crop with proper care of disease, irrigation time, soil type etc. Various tools for these technologies are temperature and moisture sensors, Precision agriculture, resolution camera, disease recognition, modern greenhouse, livestock monitoring, climate monitoring, nutrition management, hydroponics etc. These tools and techniques helps and modern agriculture but they are not very much cost effective and also time consuming. In modern days Machine Learning and Deep Learning algorithm rising rapidly in helping farmers to resolve their issue sat shorter period of time by recognizing various leaf disease with high accuracy. A diseased leaf might varies in several characteristics like size of leaf, shape of leaf and color of leaf but some of leaf might vary in shape but similar in color while some vary in color but similar in shape. All these patterns can be easily recognized by CNN model by capturing numbers of images of different plant categories. Later these trained model can be automated to some device that can easily recognize the diseased leaves. These trained models will benefit a lot to farmers in order to get rid of these plant disease and to yield higher productivity. In fact the result we will get is more accurate at a shorter period of time. The growth of these technologies i.e. CNN or Deep Learning

increasing day by day and proving its importance in every sector.

In this study, we will learn about how Neural Network or CNN (Convolutional Neural Network) able to identify the healthy plant leaves and diseased plant leaves. Deep learning is a class of Machine Learning algorithms used to model complex patterns in datasets using multiple hidden layers and activation function. This study three major stages to follow: Collection of dataset, preprocessing of dataset and model building. The dataset is collected from multiple source that it contains 5 different categories of 15 classes of healthy and unhealthy leaves i.e. bell pepper, potato, corn, cassava and tomato. Then several preprocessing are done over the dataset like image processing, data cleaning, data augmentation etc. Model building is done basically in 3 layers input layer, hidden layer and output layer. Input layer holds the data on which the model will train in which each neuron represents a unique attribute of datasets. Hidden layer are typically full connected layers, each neuron receives input from previous layer's neuron where an activation function is applied. Output layer receives an input from previous hidden layers and return an output representing the model's prediction.

II. OBJECTIVE AND METHODOLOGY

- OBJECTIVE

Our aim is to prevent the wastage of crop as it needs lot of effort to grow a crop and wastage never worth it. It will also help the farmers to prevent from financial losses. As the accessibility to reach the correct department to analyze the disease in the crop is very limited, so we will help our farmers to find the solution by their own. It will definitely deal with the problem of food insecurity in the country. The better accuracy of identification of disease is our primary aim with our best working model.

- METHODOLOGY

1. Dataset Discussion

We will collect the different healthy and unhealthy pictures of 5 plant crop species from the online platform and classify the dataset. After that we will use train the dataset with CNN (Convolutional Neural Network) model. We will use Fast API framework to work with the real time image of the potato, tomato, corn, cassava, bell pepper crop disease to identify, it contains any disease or not. Dataset consists of 15 classes with healthy and unhealthy plant leaves with a total number numbers of 17,055 images. It is also openly available on internet.

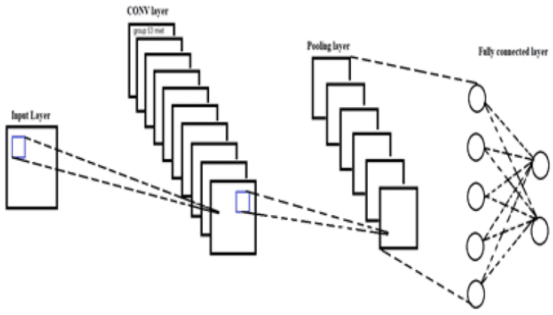
class	Vegetable Name	Infected or Not	Infection Name	Datasets(Number)
0	Bell pepper	Disease	Bell Pepper bacterial spot	1496
1	Bell Pepper	Healthy	-	1137
2	Cassava	Disease	Cassava Bacterial Blight	781
3	Cassava	Disease	Cassava Green Mottle	1507
4	Cassava	Disease	Cassava Mosaic Disease	1476
5	Cassava	Healthy	-	370
6	Corn	Disease	Corn Common rust	1528
7	Corn	Disease	Corn Gray leaf spot	1491
8	Corn	Healthy	-	1104
9	potato	Disease	Potato Early blight	1000
10	Potato	Healthy	-	152
11	Potato	Disease	Potato early blight	1000

12	Tomato	Disease	Tomato early blight	1542
13	Tomato	Healthy	-	971
14	Tomato	Disease	Tomato Late blight	1500
Total			17,055	

2. CONVOLUTION NEURAL NETWORK

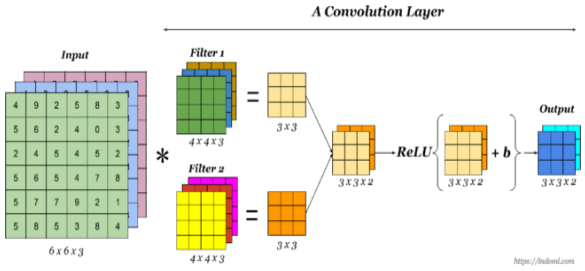
Deep Learning is a part of Machine Learning that also uses Artificial Neural Network to model complex patterns in datasets using multiple hidden layers. There is huge demand if deep learning in almost every field from banking for fraud detection to mobile phones, AI proctored examination, speech and video analysis, etc.

CNN (Convolutional Neural Network) is another type of Neural Network that differs from ANN(Artificial Neural Network) in the fact that it makes use of convolution operation to be applied on dataset. CNN is highly effective in learning the features from data like – images and infer weights and biases which in future are used for prediction. CNN model are trained iteratively using optimization techniques like gradient descent.



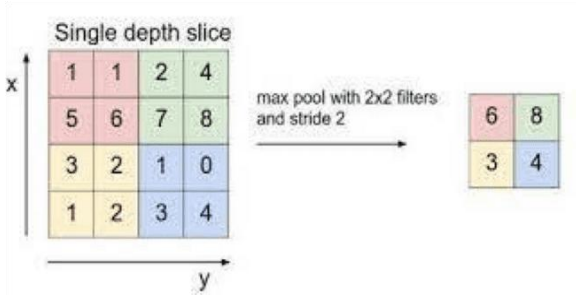
a) CONVOLUTIONAL LAYER

Convolutional Neural Network image classification takes an input image, process it and classify it under certain categories by considering an input image as an array of pixels. When CNN model's are applied to train and test, each input image will pass through a series of convolutional layers with filters, pooling, flatten layer, fully connected layers and apply activation function to classify an object with probabilistic value between zero and one. Convolutional operation is applied using a filter. In convolutional layer, the image pixels are multiplied with filter matrix to generate a feature matrix. After few repeated convolutional operation, the feature map is flatten and then a dense layer is applied with softmax activation function.



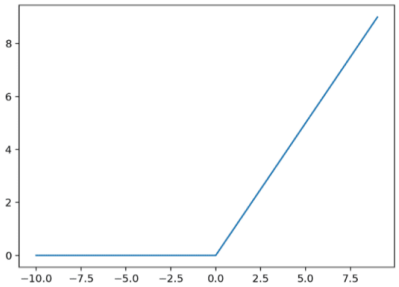
b) POOLING LAYER

Pooling layer is used to reduce the dimension on the feature map obtained after convolutional layer. Max pooling layer is applied to select the maximum value from that window of given stride. Max pooling along with convolutional helps to detect position invariant feature detection. Pooling layer also prevents over fitting as there are less parameters.



c) RELU FUNCTION

The RELU stand for the Rectified Linear Unit which is commonly used activation function in Neural Network. This function helps to prevent the exponential growth. This function used in Hidden Layers that ranges from zero to infinity. It also returns zero if it receives any negative input. It also suffers from Dying Relu Problem. Relu function is important as it does not activate all the neurons at the same time.



Relu Function Graph

d) ACCURACY CALCULATION

Accuracy is a term that describes how well our models performs during the training of the Datasets. It is also important when all class keep equal importance. The Accuracy score is calculated by dividing the number of correct predictions to total number of prediction.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

For the Binary Classification, Accuracy of the model can be calculated as a Positive and Negative as follows:

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}}$$

III. METHODOLGY STEP BY STEP

The steps are as follows:

1. Dataset acquisition or collection

The image dataset for training the model was obtained from a variety of open source sites, including the plant village dataset. Images were manually downloaded and organized into folders based on their appropriate classifications. Here we have taken 5 different classes (with some example):

a) Bell pepper(Bacterial Spot, Bacterial Healthy)

Bacterial spot Healthy

b) Cassava (Bacterial Blight, Green Mottle, Mosaic Disease, Healthy)

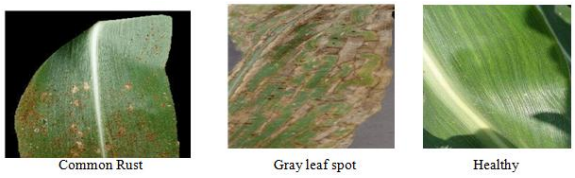
Bacterial

Green

Mosaic disease

Healthy

c) Corn (Common Rust, Gray Leaf Spot, Healthy)



Common Rust | Gray leaf spot | Healthy

d) Potato (Early Blight, Late Blight

Early blight | Healthy | Late blight

e) Tomato (Early blight, healthy, late Blight

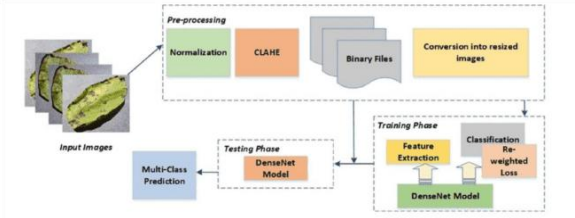
Early Blight | Healthy | Late

2. Image Pre-processing

Photos were pre-processed to minimize their size to match input layer criteria. Colored 256 X 256 resolution images are used in this study.

3. Model Building

A typical CNN model was built in to train the data. CNN Model design also plays an important role to determine the final accuracy and other results.



4. Training

The model was trained using the pre-processed dataset for about 40 epochs.

Parameter	Value
epochs	40
Batch size	32
Middle Layer	relu
Final Layer	softmax

5. Classification

First features are extracted from input images using CONV and pooling layers and then classification is done using fully connected layers.

IV. RESULT AND CONCLUSION

- RESULT

During the training of deep learning model, 40 epochs were used to train the model which attained an accuracy rate of 93.22%. While testing on random image samples, the model successfully reached a maximum accuracy rate of 89.75%. Our project will help the farmers to easily identify the disease of leaf. It will be very convenient to use as they don't have to worry about moving to another area for the search of people to successfully analyze the disease of the leaf. It will prevent the financial loss of the farmers. More importantly, there will be the proper defense for the threat of food insecurity in the society or in country. There may be imbalance of supply chain in the market which may cause the hike of the product cost drastically.

- CONCLUSION

Our proposed solution will add values to the society and help the farmers to know the accurate disease of the potato, tomato, corn, cassava, bell pepper and hence can cure the disease more accurately and prevent from the financial losses and also prevent wastage of crop. Our application is made with user friendly and predict the disease more accurately.

REFERENCES

[1] Divyansh Tiwari,Mritunjay Ashish, Nitish Gangwar -"Potato Leaf Diseases Detection Using Deep Learning" . May 2020.

[2] ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network Mohit Agarwal, Abhishek Singh, Siddhartha Arjaria, Amit Sinha, Suneet Gupta . 16 April 2020.

[3] Tomato Leaf Disease Detection Using Deep Learning Techniques. Publisher: IEEE Surampalli Ashok; Gemini Kishore; Velpula Rajesh.

[4] Fuentes, S. Yoon, S. C. Kim and D. S. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition", *Sensors 2017. Proceedings of TENCON 2018 — 2018 IEEE Region 10 Conference*, 28–31 October 2018.

[5] "Cassava Leaf Disease Detection Using Convolutional Neural Networks". Rafi Surya, Elliana Gautama.

[6] "An Overview of the Research on Plant Leaves Disease detection using Image Processing Techniques". Ujwalla Gawande. January 2014.

[7] "Bell Pepper Crop Disease Detection Using Convolution Neural Network". K P Shreyas, Gurugubelli V S Narayana. 07 November 2022.

[8] "Detection and Classification of Cassava Diseases Using Machine Learning". August 2019.

[9] "A deep learning approach for RGB image-based powdery mildew disease detection on strawberry leaves". Jaemyung Shin Young K. Chang, Brandon Heung, Tri Nguyen-Quang, Gordon W. Price, Ahmad Al-Mallahi.27 February 2021.

[10] "Leaf Disease Detection Using Machine Learning".Utkarsha N. Fulari , Rajveer K. Shastri , Anuj N. Fulari.

[11] "Pumpkin Leaf Disease Detection: Convenience of CNN Over Traditional Machine Learning in Terms of Image Classification". Mosaddek Ali Mithu, Shampa Islam Momo, Md. Mehedi Hasan, Kazi Motiour Rahman & Abdus Sattar. 04 September 2021.

[12] Oo, Y.M., Htun, N.C.: Plant leaf disease detection and classification using image processing. Int. J. Res. Eng. (IJRE) (2018).

[13] Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. Frontiers in Plant Science (2016).

[14] Bhangea, M., Hingoliwalab, H.A.: Smart farming: pomegranate disease detection using image processing. Procedia Comput. Sci. 58, 280–288 (2015).

[15] "Plant Disease Detection and Classification by Deep Learning". Muhammad Hammad Saleem, Johan Potgieter and Khalid Mahmood Arif.Published: 31 October 2019.

[16] "Blended Features Classification of Leaf-Based Cucumber Disease Using ImageProcessing Techniques". Jaweria Kainat,Syed Sajid Ullah.Published:30 Dec 2021.

[17] "Detection using convolution neural network for plant disease identification in tomato, potato and bell pepper."Rheshwan Raj Ravichandran, Norzihani Yusof.

[18] "Bell pepper leaf disease classification using neural network". Faisal Mahamud, Mithun Das, Nazmun Nahar.

[19] **"**Pepper bell leaf disease detection and classification using optimized convolutional neural network**".** Hassan Mustafa, Muhammad Umer**.**Published: 19 September 2022.

[20] "Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease". Muhammad Shoaib, Tariq Hussain, Mudassar Shah. 07 October 2022.