2021 International Conference on Artificial Intelligence and Machine Vision (AIMV) | 978-1-6654-4211-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIMV53313.2021.9670918

Plant Disease Prediction and classification using Deep Learning ConvNets

A. Lakshmanarao
Department of Information Technology
Aditya Engineering College
Surampalem, India
laxman1216@gmail.com

M.Raja Babu
Department of Information Technology
Aditya Engineering College
Surampalem, India
raajababu.makineedi@gmail.com

T.Srinivasa Ravi Kiran
Department of Computer Science
P.B. Siddhartha College of Arts & Science
Vijayawada, A.P, India
kirantsr1@gmail.com

***Abstract*—A country's inventive growth is dependent on the agricultural sector. Agriculture, the foundation of all nations, offers food and raw resources. Agriculture is hugely important to humans as a food source. As a result, plant diseases detection has become a major concern. Traditional methods for identifying plant disease are available. However, agriculture professionals or plant pathologists have traditionally employed empty eye inspection to detect leaf disease. This approach of detecting plant leaf disease traditionally can be subjective, time-consuming, as well as expensive, and requires a lot of people and a lot of information about plant diseases. It is also possible to detect plant leaf diseases using an experimentally evaluated software solution. Currently, machine learning and deep learning are using in recent years. The agriculture sector is also not a exception for machine learning. In this paper, we applied "Convnets" for plant disease detection and classification. We collected a PlantViallge dataset from Kaggle. It contains images of 15 different classes of plant leaves of three different plants potato, pepper, tomato. We divided the dataset into three datasets and applied Convnets on three datasets. We achieved an accuracy of 98.3%,98.5%,95% for potato plant disease detection, pepper plant disease detection, tomato plant disease detection. Experimental results have shown that our model achieved a good accuracy rate for plant leaf disease detection and classification.**

Keywords*—*Plant disease,Machine Learning,Deep Learning, Kaggle,Convnetssion,UCI,Python.

I. INTRODUCTION

India is a country where agriculture plays major role in the enhancement of human civilization. Crops were exclusively utilized to feed people and animals. Farming has become far more significant in recent years. People are researching production methods. Increase product activity, use fewer pesticides, and reduce environmental impact are important. The goal is to improve flat land for farming, enhance food production, and create lucrative systems. Possibilities for work are also provided. Agriculture is the primary source of income and provides the food sector with raw ingredients. The agricultural areas are confronted with challenges, including significant crop losses. Agricultural discretion will have an impact on the entire economy. Plant diseases have become a conundrum since they have raised concerns mostly about the production of agricultural outputs. There are numerous diseases in this climate. For these reasons, it is beneficial to identify these diseases effectively and on time to recognize the losses they create. Plantation, nurturing, and preventing plant diseases are key for a country's or region's overall governance. Plant diseases can be checked using a variety of approaches, including man-based and technology-based procedures. Some of the issues in plants can be seen with the naked eye. Some diseases are discovered later in the life cycle of the leaves and have already caused significant harm to the leaves and plants. Plant illnesses such as pathogens, live microorganisms, bacteria issues, fungi-infected plants, microbes, and virials cause problems in plants. It is important to identify the problem in the early stage. We considered three different plant types in this paper. The dataset was collected from Kaggle. It contains 3 types of potato leaves,2 types of pepper leaves, and ten types of tomato leaves. The sample images in the dataset (for Tomato, Potato, Pepper Bell leaves) are shown in Fig.1,Fig.2,Fig.3.

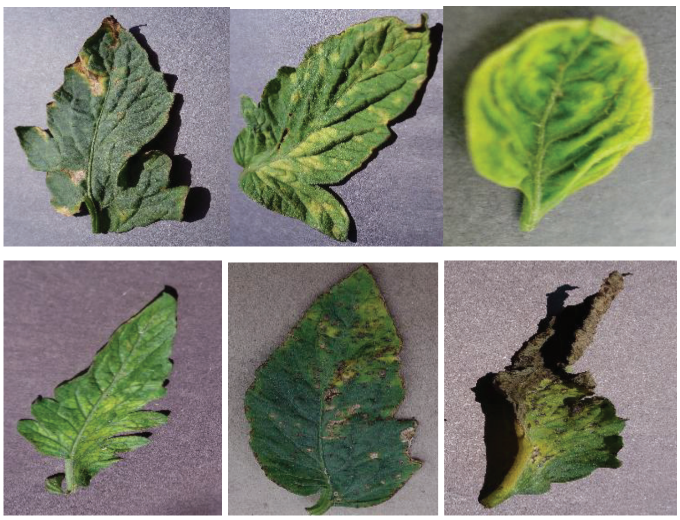

Fig. 1. Tomato Spot,Leaf_Mold,Yellow Leaf Curl Virus,Mosaic Virus,Bactorial Spot,Leaf_Spot,Early Blight,Late Blight, spider_mite

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

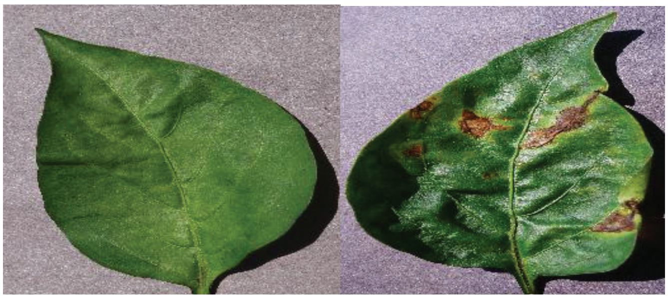

Fig. 3. Pepper Bell Healthy,Bactorial_Spot

II. RELATED WORK

Applying ML/DL in horticulture is an important role to enhance the productivity of the country. Several researchers applied ML and DL algorithms for plant disease detection. Monalisa Saha[1] et.al used convolutional neural networks for extraction of suitable features from image datasets. Later, clustering used for classifying the images as healthy and unhealthy.H. Durmuş[2] et.al applied AlexNet and then Squeeze Net. architectures on plant leaf images and achieved good results.Vishnoi[3] et.al discussed various types of diseases of plants and also discussed how they are classified so far.Qiaokang Liang[4]et.al applied the Modified ResNet50 framework for plant disease detection and achieved good results. Mohit Agarwal[5] et.al applied CNN for tomato plant leaf disease identification.Andre S. Abade[6] et.al reviewed various approaches used for plant disease Detection. They presented 121 papers with a full analysis of plant disease detection research work. A. Lakshmanrao[7] et.al applied a machine learning technique for rice plant disease detection. They applied several ML algorithms and achieved good results with random forest. Badage [8] et.al applied several Machine Learning procedures for Crop Disease Detection. They developed a Plant disease detection model in two steps. In the first step, they prepared a clean dataset. In the second step, they applied ML models and achieved good results. Iqbal[9] utilized image segmentation procedures for accurate plant image data preparation and later applied shallow learning models for potato plant disease detection T.Prajwala[10] applied a neural network model for tomato leaf disease detection and achieved an accuracy of 94%.

III. PROPOSED METHODOLOGY

First, we collected a Plant Village dataset from Kaggle [11]. The dataset comprises of 15 categories plant leaf images with three plants namely Tomato, Pepper, Potato. The details of the dataset are shown in Table-1. Next, we divide the dataset into three different datasets for Potato, Pepper, Tomato datasets. Later, we converted all the images into 64x64 numeric vectors. After that, we applied Convnets (Convolutional Neural Networks) for plant disease detection with three different plants typed. The proposed framework is shown in figure-1.

TABLE I. DETAILS OF DATASET

class	Plant Name	Images
Potato	Healthy	152
	Late-blight	1,000
	Early blight	1,000
Pepper Bell	Healthy	1,478
	Bacterial-spot	997
Tomato	Healthy	1,591
	Early-blight	1,000
	Yellow Leaf-Curl-Virus	3,209
	Target_Spot	1,404
	Two_spotted-spider-mite	1,676
	Bacterial-spot	2,127
	Late-blight	1,909
	Leaf-Mold	952
	Mosaic-virus	373
	Septoria-leaf-spot	1,771

A. *Dataset Preparation*

The original dataset is divided into three datasets. The first dataset contains 2152 images of Potato. The second dataset contains 2475 images of Pepper Bell with two different classes. The third dataset contains 16,011 images of tomato plant leaves with 10 different classes.Table-2 shows details of the prepared three datasets.

TABLE II. DETAILS OF PREPARED DATASETS

Dataset	No.of images
Dataset-1(Potato)	2152
Dataset-2(Pepper Bell)	2475
Daatset-3(Tomato)	16011

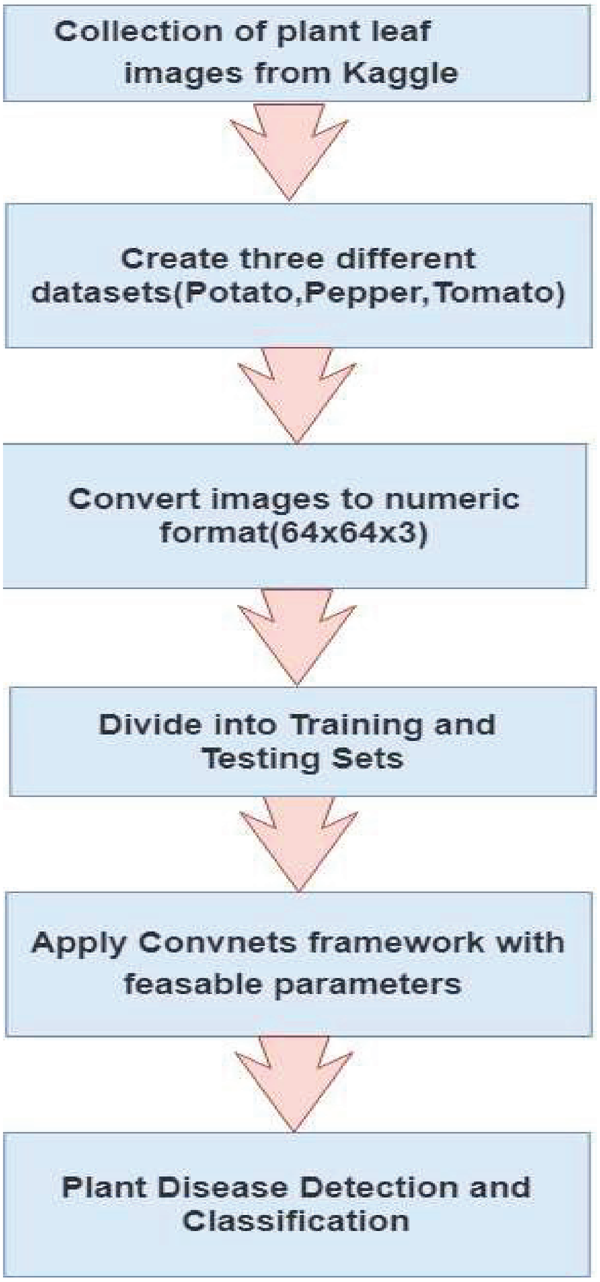

Fig. 4. Proposed Framework for Plant Disease Detection

B. Convnets

Convnets or Convolutional neural networks are at the most useful model of deep learning, popularized in recent years as the major tool for image classification problems. They are similar to Feedforward neural network model with forward propagation and backward propagation. Every neuron gets inputs, performs some mathematical operations, and adds nonliterary plays out a spot item, and alternatively follows it with a non-linearity. The entire network communicates a solitary differentiable single function. It has taken the image's pixels as input and produces class label scores as output. The difference between regular Network and ConvNet is that it can consider three dimensions namely width, height, depth of a image (Example:64x64x3).

The following steps are used in ConvNets:

Step1: Convolutional Operation: In this step, a Feature Detector is used to find Feature Map.

Step2: Applying activation function: After convolution operations activation function is applied.

Step3: Pooling: In this step, images are down sampled and reduced to small sizes.

Step4: Fully Connected Layer: This is similar to multilayer perceptron.

After the above steps, it behaves like a normal Neural Network model.

C. Data Preprocessing

If all the images are represented in the same size, then the dataset has uniformity. ConvNet also expects dimensions of input in three dimensions. So, before applying CNNs, we converted all the images in the three datasets into 64x64 numeric vectors. The final dimensions of every image before applying CNN is 64x64x3(here 3 indicates color images).

We also created a class label vector for all three datasets. This label vector is useful for checking the performance of the built CNN model.

IV. EXPERIMENTATION

A. Experimentation with Dataset1(Potato Dataset)

The dataset1 contains images of potato plant leaves. This dataset contains different leaves of two different potato diseases namely Early Blight, Late Blight. It also contains images of healthy potato leaves. So, the dataset has three labels.

Splitting into Training and Testing sets:

The Potato dataset is divided into training and testing sets. The details of division is shown in Table-3.

TABLE III. DETAILS OF POTATO DATASET

Potato Dataset	No.of Images
Total Images	2152
Training set	1721
Testing set	431

Applying ConvNets:

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 Potato dataset. The number of epochs is 80.The architecture of final CNN model has three CNN layers. The number of neurons used in the three layers is 150,80,50 respectively. The hidden layer activation function is RELU. Relu is well suited for hidden layers. For the max-pooling layer, we used a stride size of 2x2. The accuracy achieved for the Potato dataset is 98.3%. The output function used is Softmax. The optimizer used in the model is 'adam'.We also tried sgd(Stochastic Gradient Descent). But most of the times 'adam' given accurate results. The accuracy, loss chart of training and testing sets are shown in Fig.5.From Figure-5, it is observed that we achieved good accuracy for training and testing sets.So, there is no overfitting in the model. Also, by observing the loss function, it is concluded that ConvNets are able detect plant diseases effectively.

Fig. 5. Accuracy ,loss for Potato disease Detection

B. Experimentation with Dataset2(Pepper Bell Dataset)

The dataset2 contains images of Pepper Bell plant leaves. This dataset contains different leaves of one disease namely Bactorial_spot.It also contains healthy pepper bell leave images.So the dataset has two labels.

Train-test set split:

The details of the Pepper Bell dataset division is shown in Table-4.The total number of images in the dataset are 2,475.We divide the total images into train, test sets with 80%,20% split ratio.Training set contains 1980 images and Testing set contains 495 images.

TABLE IV. DETAILS OF PEPPER BELL DATASET

Potato Dataset	No.of Images
Total Images	2475
Training set	1980
Testing set	495

Applying ConvNets:

After dividing dataset into Training and Testing sets, we build a CNN model. First,we applied a CNN architecture which we applied for the Potato dataset.After experimentation, we achieved good results.So,we used the same CNN architecture,which we used for dataset-1 and achieved 98.5 % accuracy for Pepper Bell disease detection.The accuracy,loss details of training and testing sets for the model was shown in Figure-6.

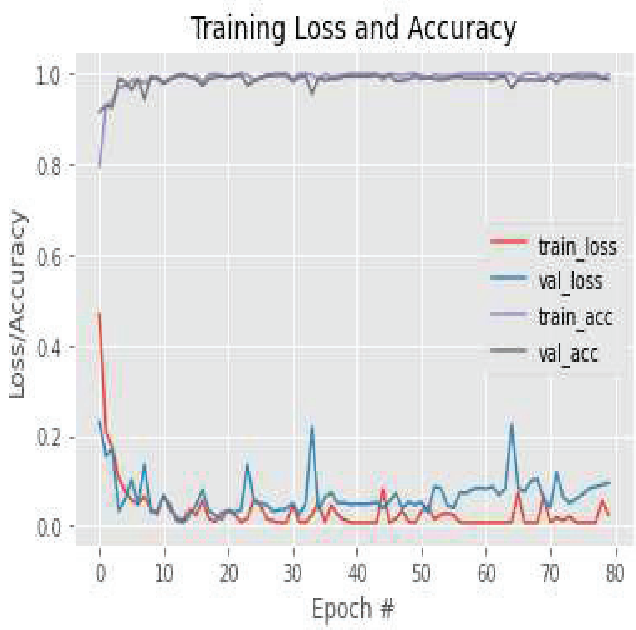

Fig. 6. Accuracy ,loss for Pepper Bell disease Detection

C. Experimentation with Dataset3(Tomato Dataset)

The dataset3(Tomato Dataset) contains images of Tomato plant leaves. This dataset contains different leaves of nine tomato diseases. It also contains tomato healthy leaf images.The total number of images in this dataset are 16011.

Train-test set split:

The details of Tomato dataset division are shown in Table-5. The total no. of images are 16011.We used 80%,20% split ratio. The no. of training images is 12808 and number of testing images is 3203.

TABLE V. DETAILS OF TOMATO DATASET

Tomato Dataset	No. of Images
Total Images	16011
Training set	12808
Testing set	3203

Applying ConvNets:

After dividing into Training and Testing sets, we build a CNN model. We tried different parameters(number of epochs, activation functions, optimizers).The number of CNN layers used for detection of tomato plant diseases is 3.The number of neurons are 250,120,60 respectively.The activation function used in hidden layers is RELU.The function to produce output types is Softmax. The number of epochs are 120.With this architecture, we achieved an accuracy of 95%.

D. Results

After conducting experiments for three datasets (potato, pepper bell, tomato), the results are tabulated.They are shown in Figure-7 and Table-6.

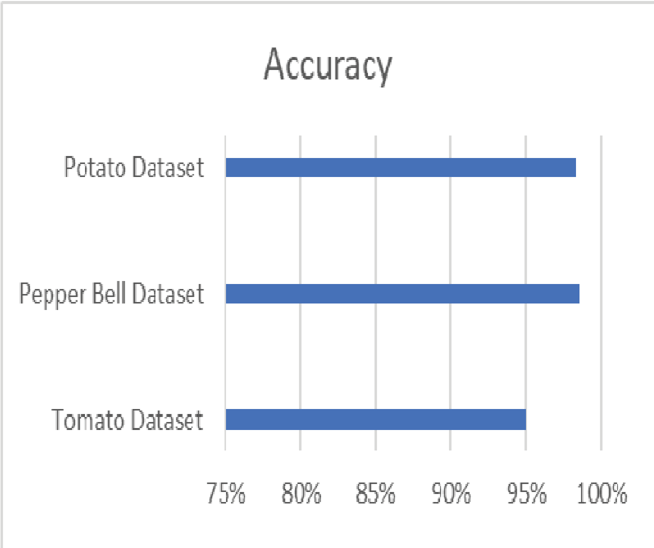

Fig. 7. Accuracy of three datasets

TABLE VI. ACCURACY

Dataset	Accuracy
Tomato Dataset	95%
Pepper Bell Dataset	98.5%
Potato Dataset	98.3%

E. Comparison with Existing Work

In [9], an accuracy of 97% achieved with random forest. In [10] ANN applied and 0.94% score obtained. Here, the accuracy of 98.3% (Table-7, Fig.8) achieved.

TABLE VII. ACCURACY COMPARISON

Model	Accuracy
Random Forest [9] (Potato disease detection)	97%
Artificial Neural Network [10] (Tomato disease detection)	94%
Proposed model (Potato disease detection)	98.3%
Proposed model (Tomato disease detection)	95%

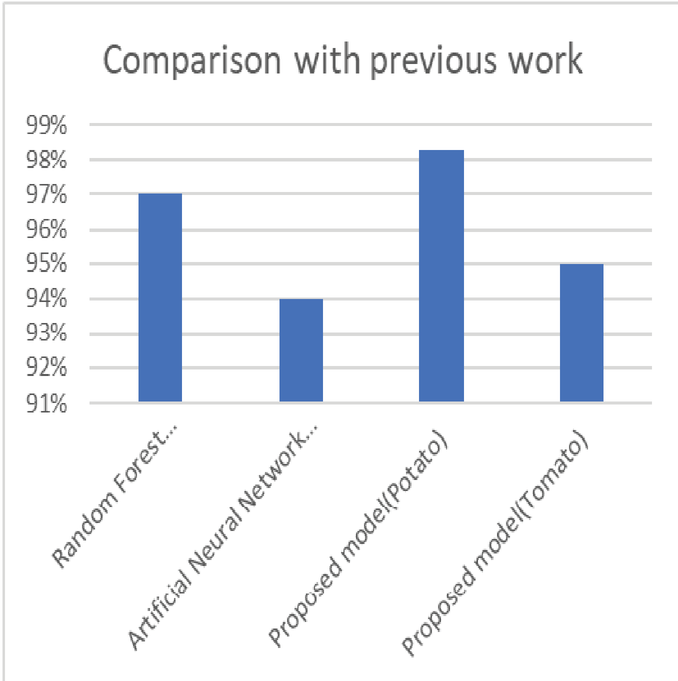

Fig. 8. Comparison with Previous Work

V. CONCLUSION

In this paper, the authors applied ConvNets for plant disease detection. First, a plant village dataset was downloaded from Kaggle Repository. Next, the authors divide the dataset into three different datasets namely the Potato dataset, Tomato dataset, Pepper Bell dataset. Later ConvNets are applied on three different datasets.The authors achieved an accuracy of 98.3%,98.5%,94% for potato disease detection,pepper bell disease detection,tomato disease detection respectively.In the future, we will test several datasets with our model.

REFERENCES

[1] M Saha,Sasikala, "Identification of Plants leaf Diseases using Machine Learning Algorithms",International Journal of Advanced Science and Technology, 29(9s),2900-2910,2020.

[2] Durmuş.H, “Disease Detection on leaves of the Tomato plants by using Deep Learning Techniques” , International Conference on Agro_Geo_Informatics, 2017, pp. 1-5.

[3] V.K.Vishnoi,K.Kumar,B.Kumar,“Plant disease detection using computational intelligence and image processing” Journal of Plant Diseases and Protection,Springer,19–53 (2021).

[4] Q.Liang, Xiang,Y Hu, G.Coppola, “Computer-assisted plant disease diagnosis and severity estimation network ”, Computers and Electronics in Agriculture,Vol-157,Pages 518-529,ISSN 0168-1699,2019.

[5] M Agarwal, A.Singh, S.Arjaria, “Tomato Leaf Disease Detection using Convolution Neural Network”, Procedia Computer Science,Volume 167,2020,Pages 293-301,ISSN 1877-0509.

[6] A S. Abade, P.A.Ferreira, F. de B.Vidal, “Plant Diseases Recognition in Images with Convolutional Neural Networks-A Review,” arXiv:2009.04365v1 [cs.CV] 9 Sep 2020.

[7] A.Lakshmanarao, G.Prasanna, S.Ramsai,”Machine Learning Methods for Plant-Disease-Detection”, Journal of Information and Computational Science,Vol- 10 Iss- 6,2020.

[8] A.Badage, “Crop Disease Detection using Machine Learning : Indian Agriculture”,IRJET, Volume: 05 Issue: 09 | Sep 2018.

[9] Asif Iqbal,K H Talukder,"Detection of Potato Disease Using Image Segmentation and Machine Learning,"International Conf. on Wireless-Communications Processing & Networking,2020, pages:43-47.

[10] T.Prajwala, “Tomato Leaf Disease Detection Using Convolutional Neural Networks”,International Conference on Contemporary Computing-IC3, 2018, pp. 1-5.

[11] https://www.kaggle.com/emmarex/plantdisease