

Identification of Plant disease in leaves, using Deep Neural Networks

Hemalatha N, Anusha T A, Asha Nair

Abstract— Plant diseases have become a concern as they can lead to a significant reduction in both the quality and quantity of agricultural products. Immediate identification of plant diseases is a key research topic as it can prove useful in the monitoring of large crop fields and thus automatically identify the signs of pathogens as soon as they appear on plant leaves. The proposed efficient algorithm could successfully identify and recognize the diseases under investigation and model could achieve an accuracy of 95.18.

Keywords: Tomato leaf diseases, Potato leaf diseases, Pepper leaf Diseases.

I. INTRODUCTION

In Biological Science information and data retrieved from images plays an important role. Based on advancements in systems engineering and devices, digital image processing and image analysis technology has many uses in biology and it avoids the issues associated with conventional photography. Plant diseases cause periodic disease outbreaks that leads to death and crop failure on a large scale. Since the effects of plant diseases have been devastating, some crop cultivation has been abandoned. The bare eye observation of plants by the experts is the main approach followed for the detection and identification of plant diseases. However, this requires continuous monitoring of specialists who may be incredibly expensive in large farms. Furthermore, the farmers may have to go in search of consulting experts which is too expensive and time-consuming.

Automatic detection of plant diseases is indeed a significant research subject since it can prove helpful in the monitoring of huge agricultural fields and thus instantly identify diseases from the symptoms that appear on the leaves of the plant. Besides, the current solutions to fight against various diseases including the heavy use of plant safety products that are harmful to the environment and to the consumer. This enables the machine to identify image-based auto-inspection of diseased leaves and provide machine guidance. Comparatively, visual identification is less accurate and it can only be done in small areas.

This paper has been divided into five sections where related works are discussed in the next section. Section three discusses the materials and methodology used for creating the model. This is followed by Result section and later by Discussion section which explains the result obtained in the experiment. Discussion compares the present work with a similar work. Paper is concluded with the Conclusion section.

Revised Manuscript Received on January 31, 2020.

N Hemalatha, Associate Professor, Department of IT and Bioinformatics, AIMIT, St Aloysius College, Mangalore, India.

Anusha T A, Guest faculty, Department of Bioinformatics, St Aloysius College, Mangalore, India.

Asha Nair, Pursuing, Msc In Big Data Analytics, St. Aloysius Of Science & Technology Mangalore, India.

II. LITERATURE REVIEW

Rangarajan *et al.*, in his paper has used the pre-trained deep learning algorithm [1]. In the analysis, the photos of tomato leaves which were taken from the dataset of PlantVillage are given as inputs to two deep learning architectures, namely AlexNet and VGG16 net. The function of the number of images and the significance of the hyperparameters, i.e. minibatch length, weight and bias learning speed, in classification accuracy and execution time, were analyzed. The accuracy of the classification using 13,262 images was 97.29 percent for VGG16 net and 97.49 percent for AlexNet. Model performance was evaluated by changing the number of pictures, by setting different minibatch sizes and varying the weight and bias learning rate. In terms of computation load, AlexNet gives good accuracy at minimal execution time compared to the deep VGG16 net.

Vijay Singh and A.K. Misra, in their paper did a survey on different disease classification techniques used for plant leaf disease detection and an algorithm for image segmentation technique that can be used for automatic detection as well as classification of plant leaf diseases later [2]. The proposed algorithm is applied on some of the species such as Banana, beans, jackfruit, lemon, mango, potato, tomato, and sapota. The advantage of using this approach is that plant diseases can be detected at the earliest stage or at an initial stage.

Arivazhagan *et al.*, in their paper explained, the use of a texture analysis to diagnose and identify plant leaf diseases [3]. The proposed algorithm was therefore checked on ten species of plants, including bananas, peas, jackfruit, lemon, mango, potatoes, tomatoes and sapota. The experimental results imply that the applied method can recognize and categorize leaf diseases with some computational effort. By this approach, plant pathogens can be detected at an earlier stage and pest control techniques can be used to solve pest problems while reducing risks to humans and the environment. In order to enhance the level of disease identification at different stages, training specimens can be improved and the shape and colour characteristics, along with the optimal characteristics, can be provided as an input condition for the identification of the disease.

Amara J *et al.*, in their paper have discussed an approach based on convolution neural networks to identify and classify banana diseases [4]. The proposed model will serve as a decision-making support tool to help farmers to recognize the disease in a banana plant. As a result, the farmer can take a picture of the leaf with the symptoms and then the system will identify the type of disease. Their main contribution was to apply deep neural networks in order to detect two famous banana diseases, under challenging circumstances such as illumination, complex environment, specific resolution, volume, pose and orientation.

Identification of Plant disease in leaves, using Deep Neural Networks

After several trials, their program was able to find good classification tests. It has shown that the proposed method can greatly help the accurate detection of leaf diseases with little computational effort.

Alvaro Fuentes *et al.*, in their paper proposed a robust deep-learning-based detector for real-time tomato diseases and pest's recognition [5]. This process provides a realistic and functional approach for the identification of the category and position of diseases in tomato plants, which is, in addition, a major difference compared to other approaches for the classification of plant diseases. Their goal was to find the more appropriate deep-learning architecture for their mission. As a result, experimental results and correlations between different deep-meta-architectures and feature extractors has shown how the deep-learning detector is able to successfully identify nine different categories of diseases and pests, including complex intra-and inter-class variants.

III. MATERIALS AND METHODS

A. Datasets: - The dataset is collected from www.kaggle.com. Images of 12 different diseases and 3 healthy samples of tomatoes, potatoes and pepper crops are collected from the PlantVillage dataset. Segmented dataset images were used, where all three-channel reference pixels (Red Green Blue) are 0 except for the leaf-associated pixels used in this analysis (as shown in Fig. 1). A total of 17,364 segmented images are available for identified diseases and healthy ones from the datasets.

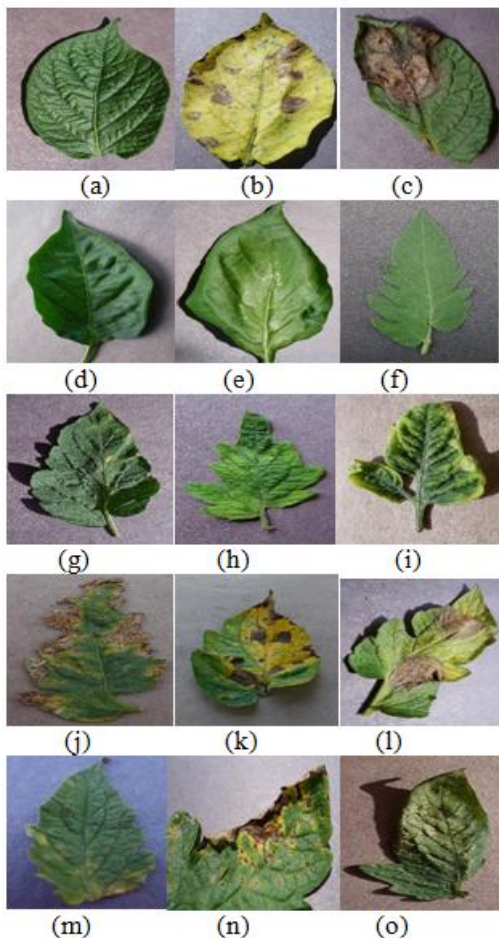


Fig. 1 (a)Potato Healthy (b)Potato_Early_blight (c)Potato_Late_blight (d)Pepper Healthy

- (e)Pepper__bell__Bacterial_spot(f) Tomato_healthy(g) Tomato_Target_Spot
- (h) Tomato_Tomato_mosaic_virus (i) Tomato__Tomato_YellowLeaf__Curl_Virus
- (j) Tomato_Bacterial_spot(k)Tomato_Early_blight(l)Tomato_Late_blight
- (m)Tomato_Leaf_Mold
- (n)Tomato_Septoria_leaf_spot
- (o)Tomato_Spider_mites_Two_spotted_spider_mite

B. Algorithm: - Artificial Intelligence has undergone a significant increase in the connection between the capacities of humans and machines. Advances in Computer Vision with Deep Learning have been developed and improved over time, mainly through a single algorithm the Convolution 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. The pre-processing required for ConvNet is much lower than for other classification algorithms. In basic methods filters are hand-engineered whereas ConvNets can learn these filters/characteristics with enough preparation.

C. Methodology: -Algorithm used for developing this prediction tool is given below.

Step 1: Start

Step 2: Packages are imported. (The package we used here is keras and scikit learn)

Step 3: Initialization of Epochs, images, height, width etc., is done

Step 4: The dataset is extracted into folder and the images are converted into array

Step 6: The data is split into train and set datasets in the ratio 2:8, where 20% test and 80% trainset are given.

Step 7: Sequential Model of keras is fitted (We have changed the convolutional layers and the size of epochs and checked).

Step 8: The datasets is checked with the epochs accordingly and the desired accuracy and loss at each stage is calculated.

Step 9: The final accuracy and the loss of the result is printed.

Step 10: The result of test accuracy and the loss in epochs is plotted and the final CNN model is saved.

Step 11: Stop.

IV. RESULT

All the experiments are performed in python (Spyder). For input data, samples of plant leaves such as Tomato leaf with the bacterial disease, Pepper leaf with the bacterial disease, Potato leaf with the bacterial disease are included. Fig.1, shows the images taken for the analysis. The software features are calculated by mapping the R, G, B components of the objects, extracting the features of the sheets and comparing them with the corresponding values of the features that are recorded in the library.

The amount of leaf disease specimens that has been categorized into 13 categories as diseased leaves and 3 categories as healthy using the proposed algorithm. Here we mainly check with the accuracy of the model; the accuracy of the model always depends on the number of epochs that is taking and also the number of Convolutional Layers. and the MaxPooling Layers of CNN (Convolutional Neural Networks)

Table- I: Algorithm accuracy according to the Epochs, CNN and MaxPooling layers Given

Epochs	CNN Layers	MaxPooling Layers	Accuracy	Loss
8	Conv2D(3 2, (3, 3), Conv2D(6 4, (3, 3), Conv2D(6 4, (3, 3), Conv2D(1 28, (3, 3), Conv2D(1 28, (3, 3)	MaxPooling2D(pool_size=(3, 3)), (MaxPooling2D(pool_size)=(2, 2)), MaxPooling2D(pool_size)=(2, 2))	91.08	8.92
10	Conv2D(3 2, (3, 3), Conv2D(6 4, (3, 3), Conv2D(6 4, (3, 3), Conv2D(1 28, (3, 3), Conv2D(1 28, (3, 3)	MaxPooling2D(pool_size=(3, 3)), (MaxPooling2D(pool_size)=(2, 2)), MaxPooling2D(pool_size)=(2, 2))	93.20	6.79
12	Conv2D(3 2, (3, 3), Conv2D(6 4, (3, 3), Conv2D(6 4, (3, 3), Conv2D(1 28, (3, 3), Conv2D(1 28, (3, 3)	MaxPooling2D(pool_size=(3, 3)), (MaxPooling2D(pool_size)=(2, 2)), MaxPooling2D(pool_size)=(2, 2))	91.19	8.80
8	Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(6 4, (4, 4), Conv2D(6 4, (4, 4)	MaxPooling2D(pool_size)=(4, 4)), (MaxPooling2D(pool_size)=(3, 3)), MaxPooling2D(pool_size)=(3, 3))	88.78	11.2126294
10	Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(6 4, (4, 4), Conv2D(6 4, (4, 4)	MaxPooling2D(pool_size)=(4, 4)), (MaxPooling2D(pool_size)=(3, 3)), MaxPooling2D(pool_size)=(3, 3))	95.18	4.81

12	Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(2 0, (4, 4), Conv2D(6 4, (4, 4), Conv2D(6 4, (4, 4)	MaxPooling2D(pool_size)=(4, 4)), (MaxPooling2D(pool_size)=(3, 3)), MaxPooling2D(pool_size)=(3, 3))	93.84	6.15
----	--	--	-------	------

From the table 1, we have that algorithm which run with 10 epochs, 4 CNN layer and 3 max pooling layers have maximum accuracy of 95.18 and loss of only 4.8 compared algorithm, when run in different epochs and parameter for CNN and max pooling layers, graphical representation of training and validation accuracy is shown in the fig.,1.1 and fig.,1.2.

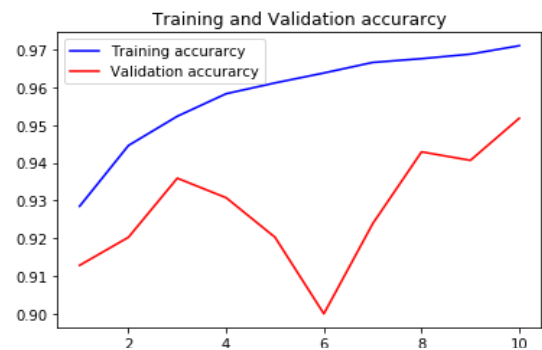


Fig., 2.1 Training and Validation accuracy

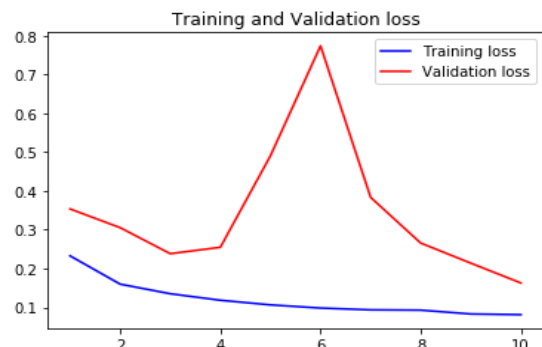


Fig., 2.2 Training and Validation loss

Table.,1 shows the accuracy of the model in each stage accordingly with the numbers of epochs, MaxPooling layers and the CNN layers. The accuracy of the model increases with the number of epochs, CNN layers and MaxPooling layers.

V. DISCUSSION

Rangarajan et al., in their paper, they used the pre- trained deep learning algorithm [1]. They used AlexNet and VG16 algorithm in MATLAB and got accuracy 97.29% for VGG16 net and 97.49% for AlexNet. They used the data from PlantVillage dataset and used only the Tomato disease leaf images. In this paper we have used the Sequential model of keras in python and the accuracy was changed according to the epochs, CNN layers and MaxPooling layers. As in [1] we choosed PlantVillage Datasets but in addition we used Tomato, Pepper, Potato leaf disease images for the model.



VI. CONCLUSION

In this paper, Tomato crop disease classification has been performed with the images from PlantVillage dataset using pretrained deep learning architecture of Sequential model of keras. The accuracy of the model was going on increasing with the number of epochs and the convolutional layers. From the experiment conducted with a dataset, we could conclude that when the algorithm was run with 10 epochs, 4 CNN layers and 3 max pooling layers with different parameters we decided an accuracy of 95.18 compared to other epochs.

REFERENCES

1. A.K Rangarajan, R. Purushothaman, and A. Ramesh. "Tomato crop disease classification using pre-trained deep learning algorithm". Procedia computer science., vol 133, 2018, pp.1040-1047.
2. V. Singh, and A.K. Misra. "Detection of plant leaf diseases using image segmentation and soft computing techniques". Information processing in Agriculture, vol. 4(1), 2017, pp.41-49.
3. S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S.V. Varthini, 2013. "Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features.", Agricultural Engineering International: CIGR Journal, vol 15(1), pp.211-217.
4. J. Amara, B. Bouaziz, and A. Algergawy, 2017. "A Deep Learning-based Approach for Banana Leaf Diseases Classification." In BTW (Workshops) (pp. 79-88).
5. A. Fuentes, S. Yoon, S. Kim and D. Park, 2017. "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition". Sensors, 17(9), p.2022.
6. S. Sladojevic, M. Arsenovic, A. Anderla., D. Culibrk, and D. Stefanovic. "Deep neural networks-based recognition of plant diseases by leaf image classification". Computational intelligence and neuroscience, 2016.

AUTHORS PROFILE



N Hemalatha, PhD is an Associate Professor in the Department of IT and Bioinformatics, AIMT, St Aloysius College, Mangalore, India. She earned her Ph.D. in the area of Computational Biology using Machine Learning. Her research interest includes Bioinformatics, Pattern Recognition, Biomedical image processing, Machine learning and Deep learning.



Anusha T A, completed her masters in Bioinformatics from St Aloysius College, Mangalore. She worked as a BIITP trainee in Research and Development Department in Jigsaw BioSolutions Pvt Ltd, Bangalore. Currently She is working as a Guest faculty, department of Bioinformatics, St Aloysius College, Mangalore. Her areas of interest are machine learning, Drug design, Biological data analysis, System Biology.



Asha Nair, is currently pursuing her MSc in Big Data Analytics from St. Aloysius of science & Technology Mangalore, India. Her research interests are Big data, machine learning, deep learning and Natural Language Processing.