Plant Disease Classification using Deep Learning Google Net Model

Satwinder Kaur, Garima Joshi, Renu Vig

Abstract*: Plant diseases have been a major crisis that is disturbing the food production. So there is a need to provide proper procedures for plant disease detection at its growing age and also during harvesting stage. Timely disease detection can help the user to respond instantly and sketch for some defensive actions. This detection can be carried out without human intervention by using plant leaf images. Deep learning is progressively best for image detection and classification. In this effort, a deep learning based GoogleNet architecture is used for plant diseases detection. The model is trained using public database of 54,306 images of 14 crop varieties and their respective diseases. It achieves 97.82% accuracy for 14 crop types making it capable of further deployment in a crop detection and protection application.

***Index Terms*: Deep learning, GoogleNet, Plant disease detection**

I. INTRODUCTION

Safety measure in food production is a major concern because the total population in world is expected to grow over more than 9.7 billion by 2050. There are many cultivation losses due to pathogens such as microbes, virus and fungus. So to minimize these losses induced in crops during growth and harvest, advanced disease detection and prevention techniques for crops are exigent. Diagnoses of plant disease by human specialist require high degree of proficiency because they must have experience in detecting large variety of disease symptoms [1]. Even experienced plant pathologists may fail to accurately detect specific diseases and diagnose them. Therefore, with progress in digital cameras and information technology, expert systems in farming have been used for crop management. Disease diagnostics and treatment shall help in improving production capacity of plants [2]. Hence, a strategy that automates the classification of diseases using images is necessary. The advantage of computerized system for the analysis of plant diseases is that it helps the agronomist to perform such diagnosis through optical observation [3]. The chief goal of these approaches is to automatically identify the diseases in order to provide suitable treatment in time [4].

In recent times, tools based on cell phones have grown rapidly, smart phones offer efficient portable solution to identify diseases due to their computing power and high degree displays.

Revised Manuscript Received on July 02, 2019.

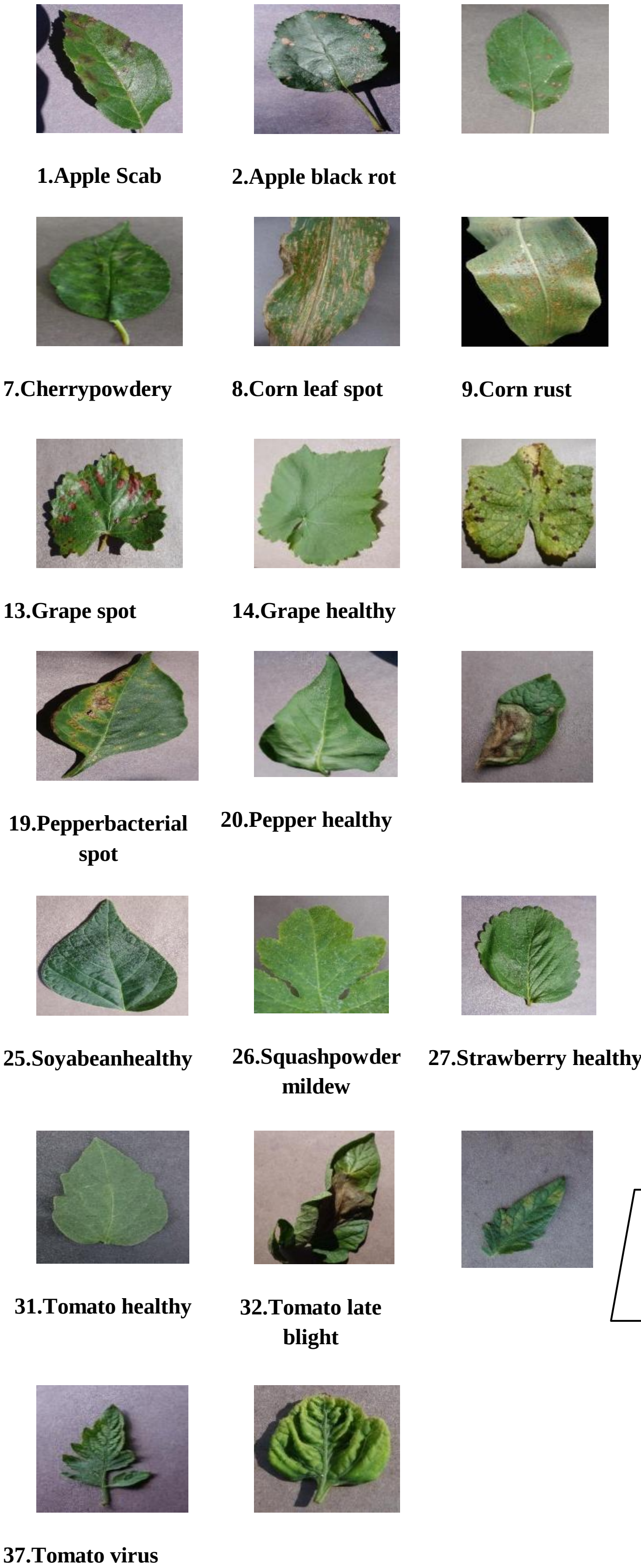
Satwinder Kaur, Garima Joshi, Renu Vig Panjab University, Chandigarh, India

The combined capabilities of mobile devices along with appropriate algorithm can lead to automated disease diagnoses. Moreover, these approaches are based on machine learning and computer vision to classify diseases using only images of plants [5].

II. RELATED WORKS

Various methods are deployed for plant disease identification applying computer vision. Koushik *et al.* reviewed the claim for developing a fast, cost efficient, and consistent system for farming. They described presently the technologies that include spectroscopic and image capturing based plant disease detection methods for monitoring strength and disease identification in plants under field conditions [7]. In [8-10], image processing disease recognition approach was used for plant disease diagnostics. Chaudhary *et al*. extracted color features, disease spots were identified by extracting some attributes such as shape feature method. Patil and Bodhe used segmentation to resolve leaf area and triangle thresholding for affected area for disease identification in sugarcane leaves getting the average accuracy of 98.60% [12]. Furthermore, extracting texture feature is also used in plant disease detection. Patil and Kumar projected model for plant disease detection using all three color, shape and texture features. This collective color extraction is used on detecting diseases of Soyabean leaves. Arrangement of all these features leads to a robust system for plant image classification [13]. Recently in 2018, a large, deep convolution neural network (CNN) was developed for plant disease detection using (LVQ) algorithm [14]. Too *et al.* detected 26 diseases with a fully connected CNN approach. The presentation of model was 99.75%. However, the capability to estimate the exact plant disease pair was limited [15].

With the help of literature review it can be concluded that to improve the accuracy, researchers have studied many methods for plant diseases diagnoses based on conventional machine learning such as k-nearest neighbor, and support vector machine (SVM). In the following years, deep CNN decreased the error rate and enhanced the accuracy. While training of neural networks take much time but the trained models can classify images very simply and rapidly. CNN contains one of the most powerful techniques for performance of pattern detection in an application with huge quantity of data. In this paper, GoogleNet based CNN architecture is trained and assessed to perform an automatic plant disease classification system.



III. MATERIALS AND METHODS

In the work presented in this paper, a deep learning approach utilizing a large open dataset, available on PlantVillage website is used. 54,306 images of 14 crops with 26 diseased and 12 healthy leaves are downloaded from PlantVillage [16]. The crop-disease pair is shown in Fig 1. A total number of 38 classes are assigned to them. Each class label is a plant disease pair, and an effort has been made to predict the plant disease pair using just an image of plant leaf. The algorithms are implemented on the GPU of DELL PRECISION Tower 7810 parallel programming platform.

IV. PROPOSED WORK

1. GoogleNet based deep learning model

Szegedy *et al.* described GoogleNet model [17] as a pre-trained network that has been trained on the basis of large parameters to classify images into 1000 classes. Block diagram of GoogleNet architecture is shown in Fig 2. GoogleNet with 22 layers deep has been applied. The initial layers (C1, C2, C3, C4) are simple convolutional layers with pooling followed by C1 and C4. After that there are numerous blocks of inception modules uses equivalent 1×1, 3×3 and 5×5 convolutions so that it can capture large amount of features. The max layers affect the spatial dimensions. GoogleNet is time consuming but accuracy is very high. For an input image, the trained model predicts the label and gives the probability at the output.

2. Training GoogleNet

The entire database is divided into training and test set. The splitting ratio of (training/testing) most commonly used in CNN application is (80/20). Transfer learning is a method used by GoogleNet to learn a new assignment. A network with transfer learning is hasty and simple than training a network from scrape with aimlessly initialized weights. Training using transfer learning is possible even using a less amount of data. For re-training the GoogleNet to classify latest images, the layers in the end of the network are replaced. These layers include information regarding combining the extracted features into probability and class labels. The last fully connected is of same size as the number of classes.

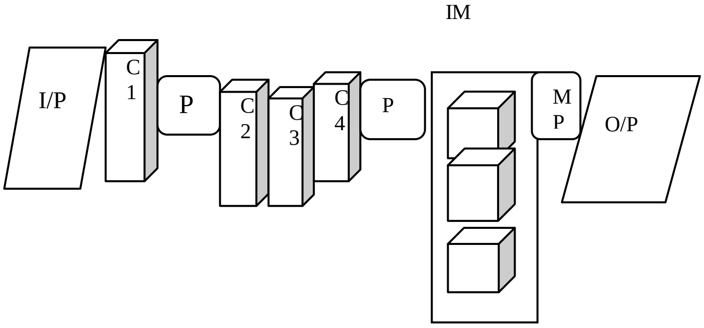


Fig 2: GoogleNet architecture (I/P-input, IM-inception module, C-convolutional, P-polling, MP-max pooling, O/P-output)

In order to ensure faster learning, the learning rate of the fully connected layer is increased. Also the learning rates in earlier layers if set is zero can congeal the weights of earlier layers in the network. Thus during re-training, it does not modify the parameters of the frozen layers. This significantly speeds up network training process. In case of smaller dataset, freezing previous network layers also avoids over fitting the new dataset.

Table 1: GoogleNet Training Validation Parameters

S. No.	Prameter	Value
1.	Input Layer size	224x224x3
2.	Epochs	6
3.	Validation Frequency	3 iterations
4.	Learning Rate	0.0001
5.	Image Augmentation	Flipping
6.	Hardware resource	Single GPU

The performance of GoogleNet on PlantVillage dataset is analyzed by training the network using transfer learning. Segmentation and feature extraction is not requirement in CNN based deep learning models as they have inbuilt layers to classify the important and unimportant features of a given class of images. The parameters used in training the GoogleNet for PlantVillage dataset is listed in Table 1 are used. The first step is to load the dataset images using augmented image data store. Load the pretrained GoogleNet network and extract the graph from trained network. Replace the final layers and freeze initial layers. Then training the network using training data by augmentation is done and classification accuracy is calculated using validation data.

V. RESULTS AND DISCUSSIONS

Table 2 presents the results of 38 diverse classes of plant disease pair. The outcomes are shown in terms of total number of images, accuracy and time. The overall accuracy obtained on PlantVillage dataset improved from 87.32% (while performing on single CPU by basic CNN architecture) to 97.82% (in case of GoogleNet performing on a single GPU). In some cases like cherry accuracy achieved is 100% due to less number of sub-classes as shown in Table 2. A snapshot of the result is shown in Fig 3. All the investigational configurations run for total of 6 epochs. The learning rate starting from 0.01 (in a case of CNN) reduced to 0.0001in GoogleNet. However, the time taken to train the network increased by 40% in case of GoogleNet.

Table 2: Plant diseases and their classification

Class No.	Plant Name	Disease Name	No. of images	Time Taken	Accuracy
1	Apple	Apple scab	2520	133m31s	99.59%
2	Apple	Apple black rot	2484		
3	Apple	Apple cedar rust	2200		
4	Apple	Apple healthy	2510		
5	Blueberry	Blueberry healthy	2270	23m35s	100%
6	Cherry	Cherry healthy	2282	95m21s	100%
7	Cherry	Cherry powdery mildew	2104		
8	Corn	Corn leaf spot	2052	188m3s	92.47%
9	Corn	Corn common rust	2317		
10	Corn	Corn healthy	2324		
11	Corn	Corn leaf blight	2385		
12	Grape	Grape black rot	2360	39m6s	98.74%
13	Grape	Grape black measles	2400		
14	Grape	Grape healthy	2115		
15	Grape	Grape leaf blight	2152		
16	Orange	Orange citrus greening	2513	18m35s	100%
17	Peach	Peach bacterial spot	2297	6m52s	98.88%
18	Peach	Peach healthy	2160		
19	Pepper	Pepper bacterial spot	2391	18m49s	99.45%
20	Pepper	Pepper healthy	2485		
21	Potato	Potato early blight	2424	28m36s	98.55%
22	Potato	Potato healthy	2280		
23	Potato	Potato late blight	2424		
24	Raspberry	Raspberry healthy	2226	23m32s	100%
25	Soyabean	Soyabean healthy	2527	29m25s	100%
26	Squash	Squash powdery mildew	2170	22m34s	100%
27	Strawberry	Strawberry healthy	2280	23m25s	100%
28	Strawberry	Strawberry leaf scorch	2218		
29	Tomato	Tomato bacterial spot	2127	81m55s	81.84%
30	Tomato	Tomato early blight	2400		
31	Tomato	Tomato healthy	2407		
32	Tomato	Tomato late blight	2314		
33	Tomato	Tomato leaf mold	2352		
34	Tomato	Tomato leaf spot	2181		
35	Tomato	Tomato spider mite	2176		
36	Tomato	Tomato target spot	2284		
37	Tomato	Tomato mosaic virus	2238		
38	Tomato	Tomato leaf curl virus	2451		

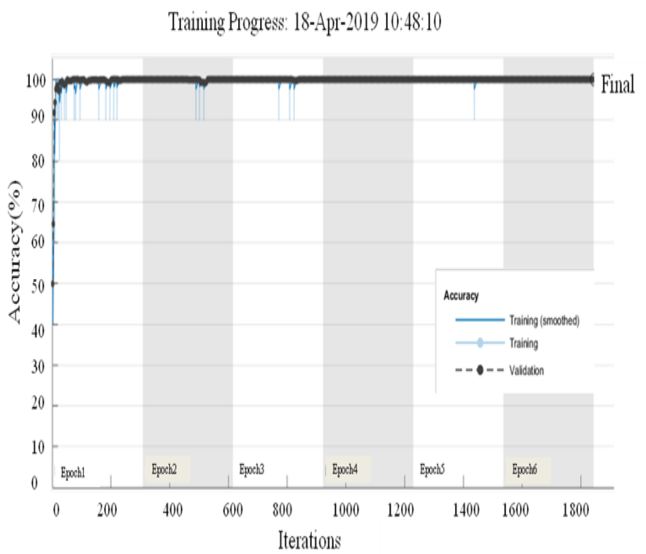


Fig 3: Snapshot of Result after training and validation for Cherry leaf images

VI. CONCLUSION

In this work, a deep learning model has been trained based on CNN architectures, for the detection and classification of plant diseases for healthy and diseased plant leaf images. Training/validation of the models is done using an openly available dataset. The most successful model architecture is GoogleNet which achieved a success rate of 97.82%. For the experiments performed, GoogleNet has proved to be the best in this case. This can be used as a practical tool for farmers to protect plants against diseases and design a viable additional method that helps to prevent food production loss.

REFERENCES

[1] Yang, Xin, and Tingwei Guo. "Machine learning in plant disease research." European Journal of Biomedical Research 3, no. 1 6-9.2017.

[2] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." Frontiers in plant science 7:1419.2017.

[3] Hughes, David, and Marcel Salathé. "An open access repository of images on plant health to enable the development of mobile disease diagnostics." arXiv preprint arXiv: 1511.08060 2015.

[4] Slodojevic, Srdjan, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep neural networks based recognition of plant diseases by leaf image classification." Computational intelligence and neuroscience 2016.

[5] Rehman, Tanzeel U., Md Sultan Mahmud, Young K. Chang, Jian Jin, and Jaemyung Shin. "Current and future applications of statistical machine learning algorithms for agricultural machine vision systems." Computers and Electronics in Agriculture 156: 585-605. 2019.

[6] Zhang, Shanwen, Haoxiang Wang, Wenzhun Huang, and Zhuhong You. "Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG." Optik 157: 866-872.2018.

[7] Nagasubramanian, Koushik, Sarah Jones, Asheesh K. Singh, Arti Singh, Baskar Ganapathysubramanian, and Soumik Sarkar. "Explaining hyperspectral imaging based plant disease identification: 3D CNN and saliency maps." arXiv preprint arXiv: 1804.08831 2018.

[8] Johannes, Alexander, Artzai Picon, Aitor Alvarez-Gila, Jone Echazarra, Sergio Rodriguez-Vaamonde, Ana Díez Navajas, and Amaia Ortiz-Barredo. "Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case." Computers and electronics in agriculture 138: 200-209.2017.

[9]Ferentinos, Konstantinos P. "Deep learning models for plant disease detection and diagnosis." Computers and Electronics in Agriculture 145 : 311-318. 2018

[10] Iqbal, Zahid, Muhammad Attique Khan, Muhammad Sharif, Jamal Hussain Shah, Muhammad Habib ur Rehman, and Kashif Javed. "An automated detection and classification of citrus plant diseases using image processing techniques: A review." Computers and electronics in agriculture 153: 12-32.2018.

[11] Chaudhary, Piyush, Anand K. Chaudhari, A. N. Cheeran, and Sharda Godara. "Color transform based approach for disease spot detection on plant leaf." International Journal of Computer Science and Telecommunications 3, no. 6: 65-70.2012.

[12]Patil, Sanjay B., and Shrikant K. Bodhe. "Leaf disease severity measurement using image processing." International Journal of Engineering and Technology 3, no. 5: 297-301.2011.

[13] Patil, Jayamala Kumar, and Raj Kumar. "Analysis of content based image retrieval for plant leaf diseases using color, shape and texture features." Engineering in agriculture, environment and food 10, no. 2: 69-78.2017.

[14] Sardogan, Melike, Adem Tuncer, and Yunus Ozen. "Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm." In 2018 3rd International Conference on Computer Science and Engineering (UBMK), pp. 382-385. IEEE, 2018.

[15] Too, Edna Chebet, Li Yujian, Sam Njuki, and Liu Yingchun. "A comparative study of fine-tuning deep learning models for plant disease identification." Computers and Electronics in Agriculture 2018.

[16] Barbedo, Jayme Garcia Arnal. "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification." Computers and electronics in agriculture 153: 46-53.2018.

[17] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9. 2015.

[18] Erhan, Dumitru, Christian Szegedy, Alexander Toshev, and Dragomir Anguelov. "Scalable object detection using deep neural networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2147-2154. 2014.

AUTHORS PROFILE

Satwinder Kaur is B.E in Electronics and Communication Engineering from Panjab University, Chandigarh, India. Here area of research includes image processing, plant disease detection. Currently, she is pursuing M.E from Panjab University, Chandigarh, India.

Garima Joshi is Ph.D. from Panjab University, Chandigarh, India. Her area of research includes image processing and VLSI design. She is working as Assistant Professor (ECE). She has more than 10 years of work experience. She has many publications in international Journals/Conference proceedings.

Renu Vig is Ph.D. from Panjab University, Chandigarh, India. Her area of research includes signal processing and soft computing techniques. Presently, she is working as Professor (ECE). She has more than 29 years of work experience. She has many publications in international Journals/Conference proceedings.