International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019

Deep Learning based Model for Plant Disease Detection

S. V. Kogilavani, S. Malliga

Abstract: ***Plant disease prediction is vital in Agriculture sector. Farmer's economic growth depends on the quality of the products that they produce, which relies on the plant's growth and the yield they get. Identifying the disease can lead to quicker interventions that can be implemented to reduce the effects of major economic loss. There may be high degree of complexity in diagnosing the various types of diseases in plants through leaves of the plants. Manual mode of plant disease detection is a tedious process. In this proposed work, Convolutional Neural Network methodologies like Sequential model and SmallerVGG model were utilized for detecting diseases in plants and diagnosis using plant leaf image. Among these two models, SmallerVGG model achieved more accuracy rate of 87% than 65% of sequential model.***

Keywords**:** ***Convolutional neural network, Deep learning, Plant disease detection, Image processing***

I. INTRODUCTION

Indian economy mainly based on Agriculture industry and this field is also known as backbone of India. Farmer's economic growth depends on the quality of products that they produce and the yield they get. Nowadays, Plants are highly prone to various diseases that certainly affect the growth of plants and leads to economic losses. Usage of manual mode of detection is a tedious job. Therefore, automatic identification of diseases in the plants is considered as an instrumental task.

Deep Learning is a sub field in artificial intelligence that emulate the mechanism of the human brain in data processing task and creating useful patterns for decision making. The learning algorithm learns from plenty of training samples available. Iteratively, the learning task can be repeated in order to obtain the correct output. Hence, in the proposed work, two different computational models were developed using Convolutional Neural Network also known as CNN. CNN is the model which utilizes some features of visual cortex. These models are designed to identify whether the plant leaf is affected or not. In this work, plant leaf dataset is collected and the collected images are preprocessed. Then, with this preprocessed images training is performed and different models are constructed.

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

S.V.Kogilavani*, Department of Computer Science and Engineering, Kongu Engineering College, Erode, Tamil Nadu, India. Email: kogilavani.sv@gmail.com

S.Malliga, Department of Computer Science and Engineering, Kongu Engineering College, Erode, Tamil Nadu, India. Email: mallisenthil@kongu.ac.in

CNN based models are used to compare and analyze the results. Initially, Sequential model is used which is a pre-built keras model with three added layers. This model is one of the majorly used model in keras. Sequential model consists of stack of layers in which several sequential layers can be designed with various convolution and pooling layers. In the second model design, SmallerVGG model was utilized which consist of five layers with fixed filter size. Finally the model is validated to predict whether the leaf is infected or not.

II. LITERATURE REVIEW

KamleshGolhani, Siva K. Balasundramet et al [1] mentioned that detecting diseases in plants is the fundamental way to prevent the losses in the yield and also improve the mass of the agricultural product. It requires massive amount of data of plant and additional execution time. Hence, hyperspectral data is to detect diseases in plants.The Neural Network with different model, classifiers and algorithms are used to process hyperspectral images. In early detection of plant diseases, two approaches are incorporated namely non imaging field spectroscopy and imaging spectroscopy. Hyperspectral data can detect diseases at various levels, like from individual plant to fields. To improve the process CNN model can be used.

G. Prem Rishi Kranth, M. HemaLalithet et al [2] presented an approach to automatically grade disease based on naked eye and disease scoring scale that leads to problems related to manual grading. Machine learning is used to predict the process by past or historical data. It is based on classification approach. The classification is based on different techniques such as decision tree, K-mean clustering, Naive Bayes, random forest and Artificial neural network. Perform the processing step of dataset gathered. Finally compare the output of different technique used. The accuracy differs from one technique to another based on the dataset taken.

Konstantinos P. Ferentinos et al [3] stated that, supervised learning is used as main characteristic to train the model. CNN based architectures are developed for image recognition. They have used 5 basic CNN architectures for detecting plant diseases which include AlexNet, GoogLeNet, VGG, AlexNetOWTBn and Overfeat. First step is to train and test dataset. The dataset gathered undergoes some process like preprocessing(resizing the image), splitting of data based on 80-20 rule. Compared to other model, VGG and AlexNetOWTBn is the most successful model. The success rate of final model is 99.53%. But there is some problematic situation in identifying the classes.

Satish Madhgoria, Marek Schikora et al [4] proposed automatic classification method which utilizes pixels to detect unhealthy or infected portions in the images of plant leaves.

The algorithms have been tested comprehensively. Linear Support Vector Machine has been used to classify every component. The results from SVM can be improved remarkably victimization the neighborhood check technique. The task is performed in three steps. Initially perform segmentation to separate the image into foreground and background portions. Next, SVM is applied to identify the class of the foreground portion. Finally, perform neighborhood-check in order to remove all falsely-classified pixels in previous step.

Heeb Al Bashish, Malik Braik et al [5] represents an image processing based work for leaf as well as stem disease detection. They concentrates on five diseases which affect the plants like Early scorch, Cottony mold, ashen mold, late scorch, tiny whiteness. In this work, initially the images are segmented using the K-Means algorithm, and after that the segmented images are passed through a pre-trained neural network.

P.Revathi, M.Hemalatha et al [6] described about the work which is based on Image edge detection segmentation techniques in which, the captured images are processed for enrichmentdisease spots are identified using RGB color image segmentation. The features of the images are extracted from the disease spots in order to recognize diseases. This work concentrates on cotton leaf spot, cotton leaf color segmentation and Edge detection.

S. Pramod landge, A Sushil Patil et al [7] developed an application for automatic detection and classification of plant diseases through image processing techniques. They developed image processing algorithms that can recognize and identify problems in crops from images, based on color, texture and shape to automatically detect diseases or different conditions which may have an effect on crops. They provide quick as well as correct solutions to the farmer with the help of messages.

S. Bhumika Prajapati, K. Vipul Dabhi et al [8] proposes an approach for detection and classification of cotton plant disease. They identified that background removal color house conversion from RGB to HSV is useful. They also identified that threshold technique produces better results compared to other background removal techniques. They also found that SVM produces smart results, based on accuracy, for disease classification task.

MalvikaRanjan, Manasi Rajiv Weginwar et al [9] represented a method that is widely used and required correct judgment to identify the diseases. In this approach, various diseased leaf images are collected and processed . Artificial neural network (ANN) is used as a model to segregate the healthy and diseased samples.

III. METHODOLOGY

A. Neural Network

A neural network will be a simplified model of the method the human brain processes data. This network consists of a colelction of “neurons” which are structured in layers. The bottom layer consists of input , and the top layer represents output. The neural network sometimes consists of additional layer called hidden layer which may have hidden neurons. The neural networks become non-linear when hidden layers are incorporated.

B. Convolutional Neural Network

Convolutional Neural Network is the mainly used neural network model for image classification problem. CNN architecture is formed by a sequential model consisting of different layer that transform the input volume into an output volume. CNN architecture has several layers consisting of convolution layer, pooling layer and fully connected layer. A convolutional layer has a number of filters to perform convolutional operation. The function of pooling layer is to reduce the spatial size of the representation of layer. Then the final layer is fully connected layer which reduces the number of connections. Activation function plays an important role to learn the mapping between the input and the response variables. Activation function checks for the neuron should be activated or not by calculating weighted sum and further adding bias with it. This activation function is used to deal non linear or not linearly separable data samples.

IV. PROPOSED SYSTEM

The proposed system utilizes Convolutional Neural Network. CNN is used to recognize the essential and unimportant features from the image dataset. The proposed work concentrates on training CNN model so that it will classify the plant leaves into diseased or not. In this process, model is trained by feeding existing data samples. CNN comprises of one or more convolutional layers which is followed by one or more fully connected layers. Dataset containing leaf images of healthy and infected plants were used for the training and testing of the CNN models. The entire dataset was initially divided into two portions, the training set, and the test set, by randomly splitting images. Out of total data samples, 80% of them formed the training set and remaining 20% formed the test set. Sequential and SmallerVGG models were implemented in this work for train and assess the dataset. Among the Sequential and SmallerVGG architectures, SmallerVGG consistently performs better than Sequential model based on the method training.

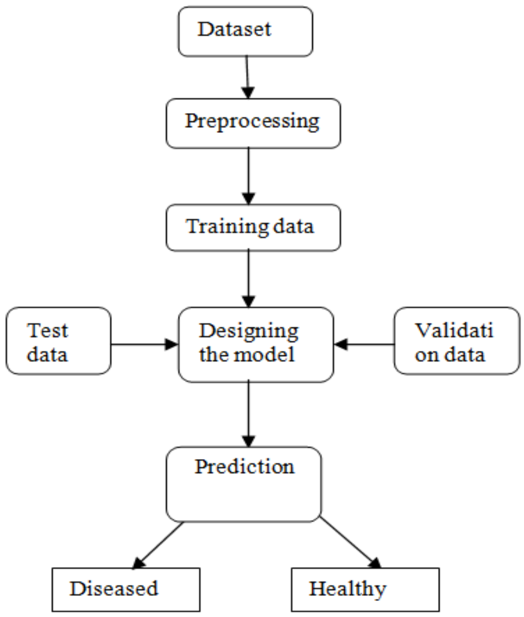

Fig.1. Proposed System Design

Initially the datasets are collected from PlantVillage which consists of plant leaf images. The dataset contains healthy and diseased leaves. A total of 2,000 images were taken into account. The dataset is split into training and test set based on 80/20 rule. The pre-processing step is carried out to re-size the labeled images in a fixed size of 96 x 96. For RGB images, we have used channel 3. Then the images are sorted. The resized image which is being processed is converted to numpy array initially and then casted to float32 to save space. Then the images are normalized by dividing it by 255.This ensures that values lie between 0 and 1. This helps us to train the model faster.

The first model utilized is the Sequential model which allows to create the model layer by layer. The first hidden layer used in convolutional neural network is convolutional layer with 32 filters for first 2 layers and last layer with filters 64 each of size 2x2.This layer produces feature map as output and is given as input to the pooling layers. The second layer used is the pooling layer, which uses max pooling to reduce dimensions. Max pooling with 2x2 window is utilized and it considers only maximum value in 2x2 window. Then, Flatten function is added to convert the matrix format output to a single list format. The last layer of this network is the fully connected layer which uses Dense function with 64 filters to check whether every node is connected with each other. ReLU is the activation function used. Dropouts are included in the fully connected layer to prevent over-fitting. Sigmoid activation function is used in final layer to determine the probability between 0 and 1.

The second model utilized is VGG model which has several layers in it but we have utilized only 5 hidden layers, so we call this model as SmallerVGG. This model has employed 5 layers of convolution with fixed filter size of 3 x 3. Padding is embedded in the convolutional layer to add zeros in input matrix and the ReLU activation function is added. With the convolutional layer output, max pooling of 3 x3 as pool size is performed. Dropouts are utilized to disconnect nodes from the current layer to the next layer. This way of random disconnects in nature helps the neural network to minimize overfitting. Likewise, the following actions are performed for the other four layers. Then, the Flatten function is added which helps in converting the matrix array into linear array. Dense function is used to connect each input node to each output node. Finally, the model is compiled and saved for future use.

Finally prediction is performed with labeled images at first which can be chosen randomly. Later, it can be done without labeling of data to check whether the leaf is detected correctly. For predicting, randomly images are chosen and is loaded and resizing is done. Then, the image is converted to numpy array and the class labels are predicted in the model to determine whether the labeled and unlabeled images are predicted as healthy or diseased.

V. EXPERIMENTAL AND RESULT ANALYSIS

The dataset consists of plant leaf images which are collected from PlantVillage. The dataset contains healthy and diseased leaves A total of 2,000 images were taken into account.

Precision: It is used to evaluate the performance of retrieved instances. Precision is given as in equation 1.

$$\Pr ecision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (1)$$

True positive denotes that the leaf has disease and it actually predict the leaf has diseases. False positive denotes that the leaf has disease and it actually predict the leaf doesn't has diseases.

Recall : Recall is a fraction of relevant instance that have been retrieved over the total amount of relevant instances. Recall is given as in equation 2.

$$\operatorname{Re} call = \frac{TruePositive}{TruePositive + FalseNegative} \quad (2)$$

False negative denotes that the leaf doesn't has disease and it actually predict the leaf has diseases.

Accuracy: It is one metric for evaluating classification models. Accuracy is given as in equation 3.

$$Accuracy = \frac{Numberofcorrectprediction}{Tota\ln umberofprediction} \quad (3)$$

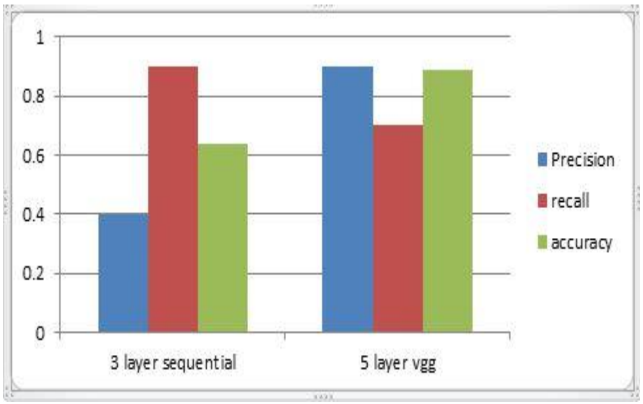

Fig.2. Performance Analysis of Sequential and SmallerVGG models

Two different CNN models were trained by applying the various parameters. After appropriate training, these values gave the best results. The learning rate followed a specific annealing schedule, starting from 1e-3 and epoch of 5 with batch size 32. The models were compared based on their performance on the test set. The SmallerVGG model reached the high accuracy of 87%. The sample output of SmallerVGG model is shown in figure 3.

Fig.3. Sample outcome of SmallerVGG model

Mapping of loss and accuracy for SmallerVGG model is depicted in figure 4.

Fig.4. Mapping of loss and Accuracy for SmallerVGG Model

VI. CONCLUSION AND FUTURE WORK

In this research work, two deep learning models like Sequential and SmallerVGG models were implemented and analyzedd, which works based on specific convolutional neural network, for the detection of plant diseases through simple leaf images of healthy or diseased plants. The training of the models was performed using a leaf images, taken in cultivation fields. The data classified into two distinct classes namely healthy and diseased. Out of two models, a SmallerVGG convolutional neural network, achieved a success rate of 87% compare with Sequential convolutional neural network which gave 65% of accuracy.

In the future work, large number of datasets can be collected to improvise the system. Further, it can be implemented to predict how the leaf gets affected and the system must provide necessary pesticides and fertilizers for the infected plant leaves.

APPENDIX

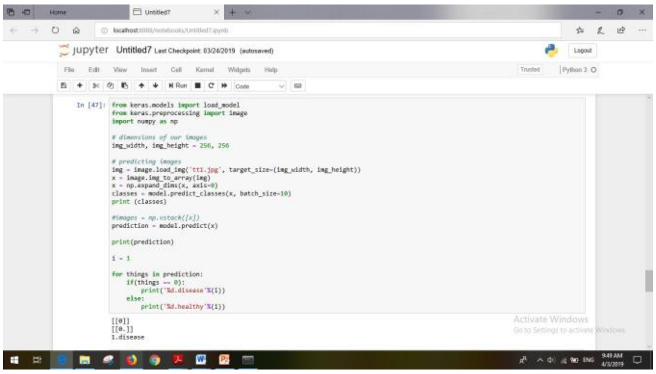

Fig. A1 Accuracy of Sequential Model

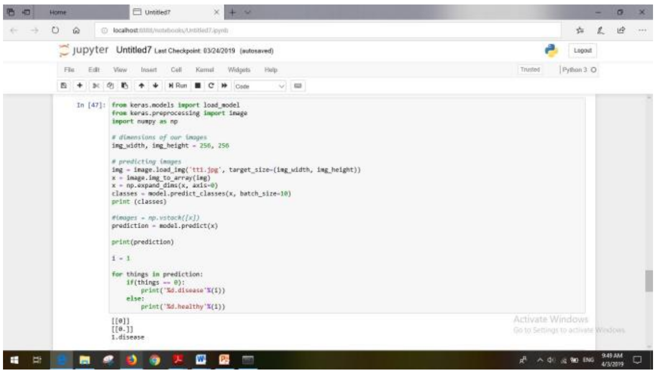

Figure A2. Validating the data with Sequential Model

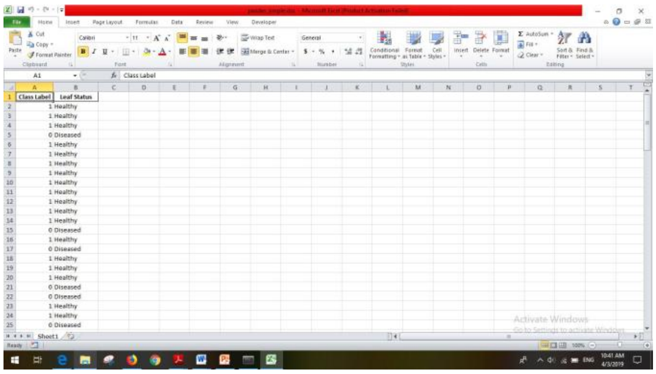

Fig. A3. Predicted labels of Test data of Sequential Model

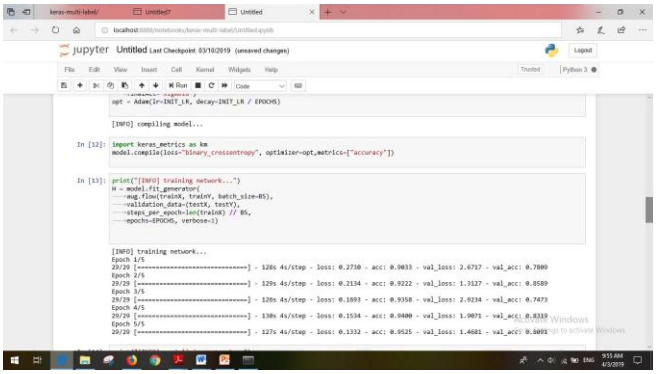

Fig. A4 Accuracy of SmallerVGGModel

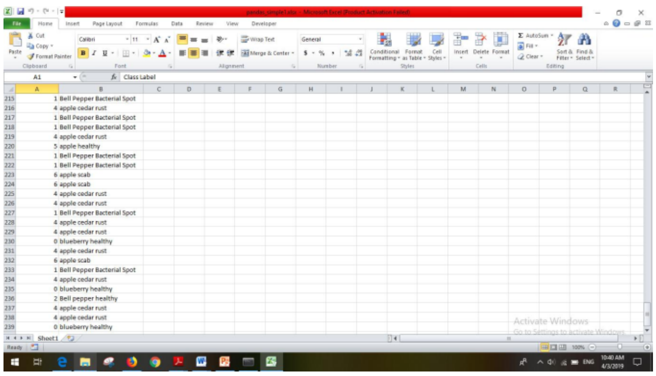

Fig. A5 Predicted Labels of SmallerVGG Model

REFERENCES

1. Kamlesh Golhani, Siva K. Balasundram, GanesanVadamalai, BiswajeetPradhan, "A review of neural networks in plant disease detection using hyperspectral data", Information Processing in Agriculture, Vol:5, Issue 3, 2018.
2. G. Prem Rishi Kranth, M. HemaLalitha, LaharikaBasava, Anjali Mathur, "Plant Disease Prediction using Machine Learning Algorithms", International Journal of Computer Applications, Vol:182(25), 2018.
3. Konstantinos P. Ferentinos. "Deep learning models for plant disease detection and diagnosis", Computers and Electronics in Agriculture, Vol:145, 2018.
4. SatishMadhgoria, MarekSchikora, and Wolfgang Koch, "Pixel-Based Classification Method for Detecting Unhealthy Regions in Leaf Images", Proceedings of 6th Workshop on Sendor Data Fusion, 2011.
5. Heeb Al Bashish, Malik Braik, and SuliemanBani-Ahmad, "A Framework for Detection and Classification of Plant Leaf and Stem Diseases", IEEE International Conference on Signal and Image Processing, 2010.
6. P.Revathi, M.Hemalatha, "Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection U sing Image Edge detection", IEEE conference on Computing, Communication and Networking, 2012.

7. Mr.Pramod S. landge, Sushil A. Patil, Dhanashree S. Khot, "Automatic Detection and Classification of Plant Disease through Image Processing", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, 2013.
8. BhumikaS.Prajapati, VipulK.DabhiHarshadkumar, B.Prajapati, "A Survey on Detection and Classification of Cotton Leaf Diseases", International Conference on Electrical, Electronics, and Optimization Techniques, 2016.
9. MalvikaRanjan, Manasi Rajiv Weginwar, NehaJoshi, Prof.A.B. Ingole, "detection and classification of leaf disease using artificial neural network", International Journal ofCamps-Valls G, Bruzzone L. Kernel-based methods for hyperspectral image classification. IEEE Transaction on Geoscience Remote Sensing. 43(6), 2005.
10. Balasundram SK, Kassim FA, Vadamalai G, Hanif AHM. "Estimation of red tip disease severity in pineapple using a non-contact sensor approach", Agriculture Science, 4(4), 2013.
11. Pen˜ uelas J, Gamon JA, Fredeen AL, Merino J, Field CB, "Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves", Remote Sense Environment 48(2), 1994.
12. Mahlein AK, RumpfT,Welke P, Dehne HW, Plu¨ mer L, Steiner U, et al, "Development of spectral indices for detecting and identifying plant diseases", Remote Sense Environment 128, 2013.
13. Ashourloo D, Mobasheri M, Huete A, "Developing two spectral disease indices for detection of wheat leaf rust (Pucciniatriticina)", Remote Sense Environment, 6(6), 2014.
14. Wang G, Sun Y, Wang J, "Automatic image-based plant disease severity estimation using deep learning", Computational Intelligence Neuroscience, 2017.
15. Haut JM, Paoletti M, Plaza J, Plaza A, "Cloud implementation of the K-means algorithm for hyperspectral image analysis", Journal of Supercomputer, 73(1), 2017.
16. Quirita VAA, da Costa GAOP, Happ PN, Feitosa RQ, d.S. Ferreira RQ, Oliveira DAB, "A new cloud computing architecture for the classification of remote sensing data", IEEE Journal of Selected Topics in Applied Earth Observations and Remote sensing, Vol:10, Issue:2, 2017.

AUTHORS PROFILE

Dr.S.V. Kogilavani is associated with the Department of Computer Science and Engineering as Associate Professor at Kongu Engineering College, Tamil Nadu, India. She has presented 30 papers in national and international conferences and published 25 papers in national and international journals. She received Summer Research Fellowship from Indian Academy of Sciences in the year 2018. Her Area of Interest is on Machine Learning and Natural Language Processing.

Dr. S.Malliga is working as a Professor in the Department of Computer Science and Engineering, Kongu Engineering College, Tamil Nadu, India. Her main research area is Network and Information Security. She has done consultancy project for BPL and offered several courses on latest technology. She has published 36 articles in international journals and presented more than 25 papers in national and international conferences in her research and other technical areas.. She is also interested in cloud and virtualization technolgies.

.