International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Enhanced Plant Leaf Disease Detection Using Deep Learning Techniques

DR.R.Gopi[1], R.Aarthy[2], Kothamasu Venkata Ratna Sai[3], Jasthi Manikanta[4], Chunduri Sai Babu[5], Bommisetti Pardha Sai[6]
[1] Associative Professor, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur
[2] Assistant Professor, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur
[3,4,5,6] Student, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur
Dhanalakshmi Srinivasan Engineering College (AUTONOMOUS),Perambalur-621212

***Abstract*: Indian economy is largely depending on the growth in the agricultural production. Nowadays, iden -tifying the plant leaf disease is one of the most challenging tasks in the agricultural sector because plant leaf diseases spread very quickly and effects the yield of the crop. In this paper, we proposed three deeplearning models i.e, VGG16, VGG19 and InceptionV3 for detecting various plant leaf diseases because manual detection sometimes leads to inaccurate results.The dataset consists of several variety of plants like tomato, apple, potato, sugarcane, grapes and corn of both healthy and diseased leaves and all the images arecollected manually and also from various freelyavailable resources. Using the three models we achieved an average accuracy as 91.8% accuracy which demonstrates the feasibility of the profound deep neural network approach**

Keywords:
VGG16, VGG19, InceptionV3, deep learning,Convolution Neural Network (CNN).

I. Introduction

Agriculture is one of the major sectors that influences a country's economic growth. Food and health care industries are mainly dependent on plants. About 40%of medicines are derived from plant leaves, roots, flowers and stems In countries like India, the majorityof population depends on agriculture for their livelihood. There are many factors which effects the yield of the crop some of them are crop harvesting, weather forecasting, disease detection [1]. Between
26 and 40 percent of the world's potential crop production is lost annually because of weeds, pests and diseases, and these losses could double without the use of crop protection practices Some of the above factors cannot be controlled but the factors like disease detection can be controlled and farmers face many turbulences in the identification of the plant diseases. Plants diseases are mainly caused by pathogens and they comprise of fungi, viruses, nematode, parasitic bacteria, and bacteria [2]. These plant pathogens cause diseases for root, fruits, vascular system, stem and leaf [3]. Some of the diseases will spread rapidly which means the diseaseswill spread from one plant to other very rapidly thus identification of such diseases in early stage is very crucial to avoid the losses[4]. So in order to detect the plant leaf disease which directly increases the yield of the crop we proposed a model which uses CNN deeplearning techniques to identify the type of disease. The dataset used in the proposed work contains around 87,000 images of 38 different classes and the data set is split into test and train datasets [5]

The convolutional-neural-network (CNN) is a subclass or just one kind of neural-networks or more precisely Artificial neural networks which have Pooling layer, Activation layer, Fully Connected layer and at least one convolution layer [6]. Neural networks do the linear combination between the previous layers output and current layers weights andthen passes the data to the next layer by passing through the activation function. CNN performs the convolution between previous layers outputs and current layers kernel and pass the data to the next layer by passing through the activation function. Deep learning does not require any feature engineering unlike traditional machine learning. Here our model utilizes the pictures of plant leaves to detect the disease present in plants just like previous research [7].

The concept of VGG16, VGG19 and InceptionV3 are used for the detection of plant leaf disease. The proposed three models are the specific convolution network methods specially designed for the classification and localization whereas CNN is the main concept of a neural network [8]. CNN model is used to extract some features like horizontal edges, vertical edges, RGB values etc., from images of the leaves [9]. In the previous research only a single type or few types of plant leaves are studied in one go but, in this research, we studied nearly 38 classes of plantleaves in one go using the proposed architecture models [10]. The functionality differences between VGG16 and VGG19 model is because of the three convolution layers present in VGG19 and they will certainly works great for good neural networks but they may not function well to extract complex features because they contain a simple stack of

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

convolution and also max- pooling layers followed by finally fully connected layers where as in InceptionV3 there are some inception modules that consists of filters ranging from 1*1 known as pointwise convolution subsequently by convolution layers with different filter sizes applied all together. In the previous research only a single or two architectures are used and also performed using only a specific typeof plant leaves but in this research, we developed a model using three architectures which uses 38 different classes of plant leaves and obtained the accuracy of nearly 94 percent.

This paper is divided into multiple sections where section I is the introduction, we discussed about the importance of detecting plant diseases at early stage and also described briefly about the implementation of the model. In section 1 we described about project implementation and also the process of collecting of data set, dividing the dataset, preprocessing the dataset and finally retrieving the results from the model [13], section II proposes methodology of the model and the steps involved during the implementation of the model is discussed., section III presents the results in the form of graphs plotted between accuracy, validation accuracy and loss, validation loss, section IV presents the conclusion by discussing the difference between vgg16, vgg19 and InceptionV3 and by comparing the results with other research papers by taking the references mentioned insection V [14].

II. Dataset Description

The dataset was taken from an open-source repository Kaggle and also collected manually. The dataset contains nearly 87,000 images of healthy and diseasedleaves of various variety of plants like tomato, apple, potato, sugarcane, grapes and corn. Every plant contains more than one type of leaf disease and eachtype is contemplated as a separate type of plant leaf disease, The number of images in every class are not same.

III. Proposed Model

The dataset, which comprises of 38 classes, is used to detect plant leaf diseases. CNN models are used to predict the type of disease afflicted out of the 38 classes. Vgg16, Vgg19 and InceptionV3 are the sophisticated CNN architectures with per-trained layers and a thorough grasp of how an image isdefined in terms of shape, color, and structure. The only pre-processing done is subtracting the mean RGB value from each pixel. The three deep neural networks have been trained on millions of photos withcomplex classification problems. As a result, these Architecture can accurately forecast plant disease, by allowing us to load a pre-trained version of the network and trained on more than a million images from the Image Net database. The pre-trained network is capable of classifying images into 1000 object categories. The network has therefore been trained to recognize rich features representations for a wide range of images, including computer keyboards,mice, pencils. An many more animals. The network'simage input size is 224 by 224[5].

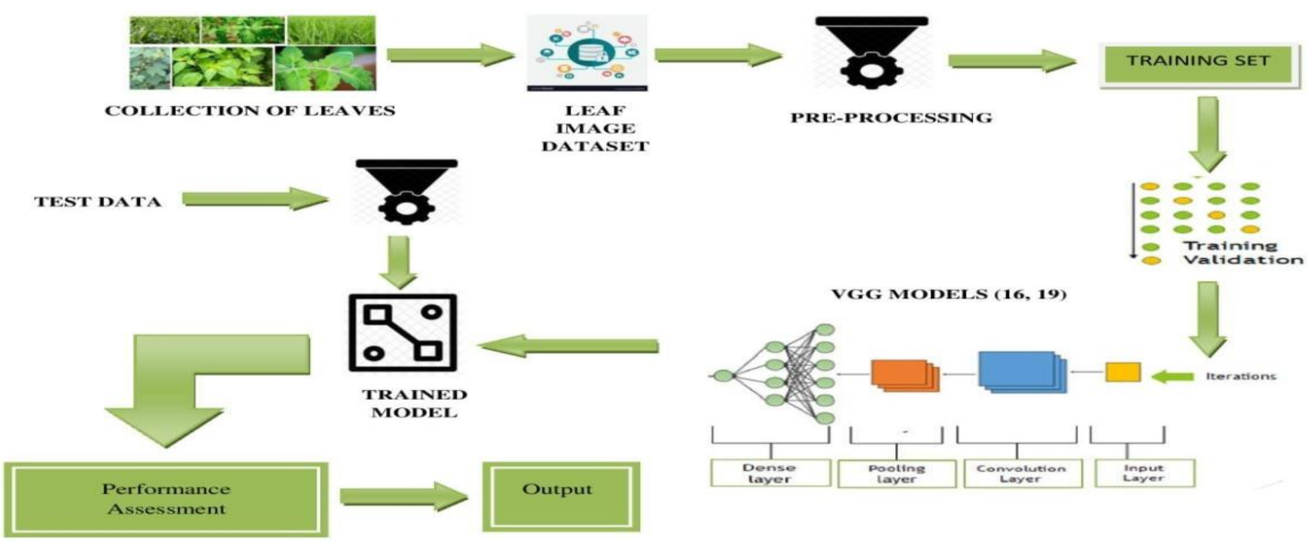


Fig 1: Proposed Methodology

IV. Proposed Methodology

Above block diagram represent proposed methodology and few blocks are discussed below:

I. Leaf Image Dataset

Different images are obtained from several sources, like Kaggle, surrounding, and from google images bycollecting all, we created a dataset with 87,000 images

II. Pre-processing

The very first step is to reduce the amount of noise inthe image by clipping out the parts of the image that aren't relevant. The picture will not be utilized if it hastoo much noise [16]. to standardize input photos in the dataset, images obtained from many sources of various sizes must be shrunk to 224x224 pixels.

III. Data Set

The dataset collected from various sources is divided into two parts in 80/20 ratio, one for training and the other for testing. 80 percent is for training the model and remaining 20 percent is used for testing the model. Training dataset consists of 56,236 images andtesting dataset consists of 14,057 images.

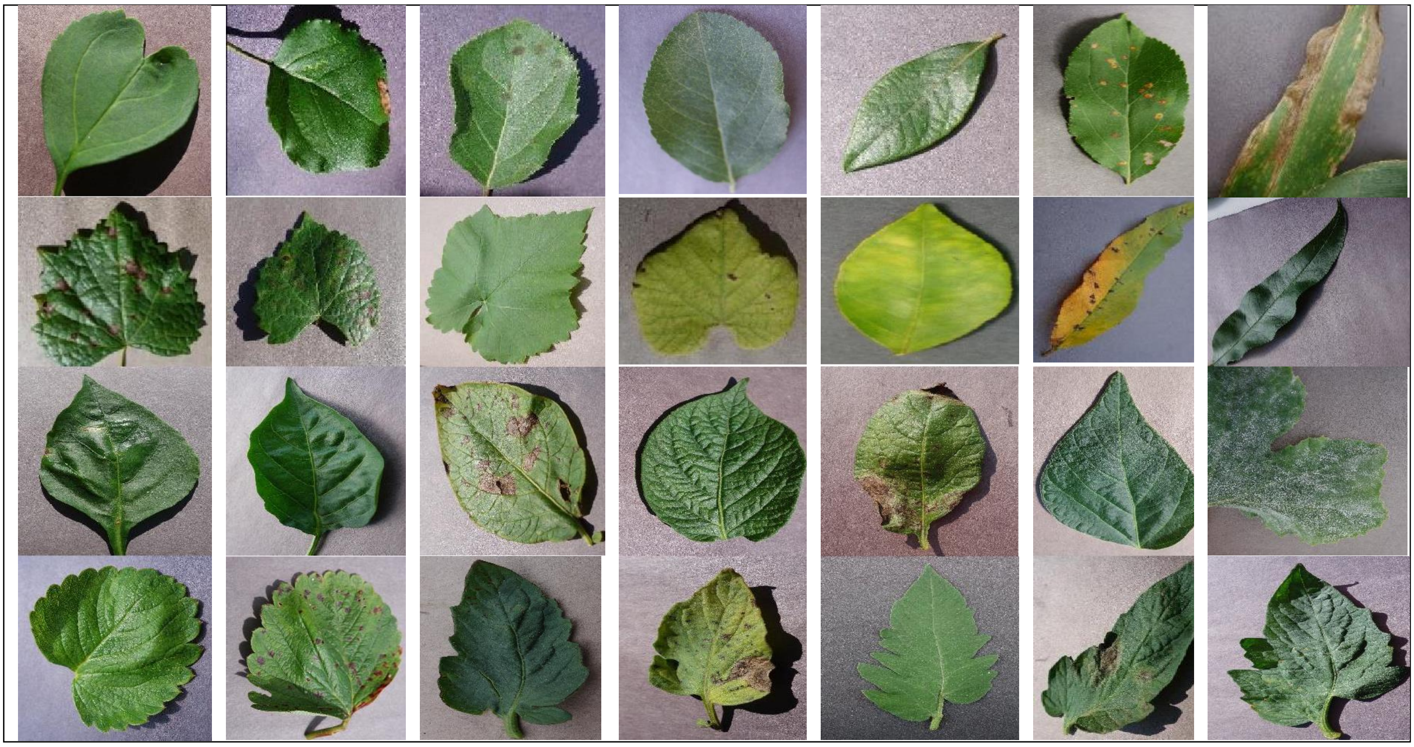


Fig:2 Collected Leaves

(a)

(b)

(c)

Fig.3 (a) Healthy leaf (b) Unhealthy leaf –I (c) Unhealthy leaf-II

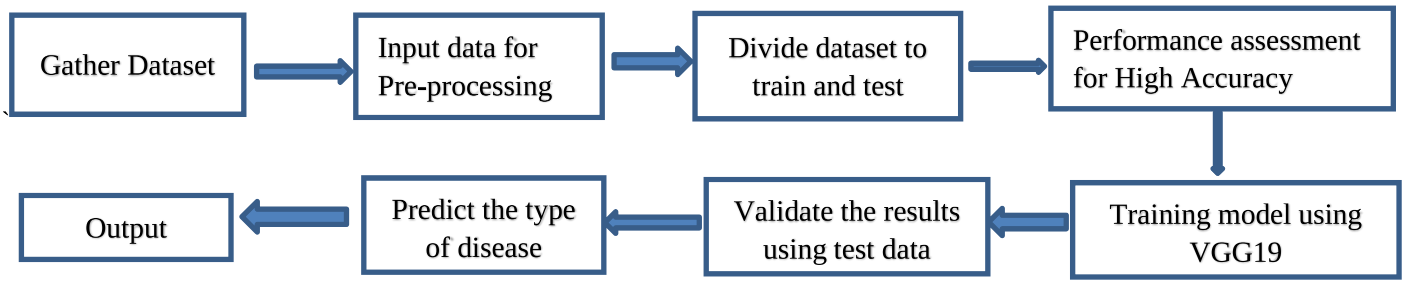


Fig 4. Workflow of proposed model

The Algorithm above represents the workflow or steps contained in the sequence. Flowcharts are used to analyze, design, document, or manage processes orprograms in different disciplines. First, collect the data and create the dataset. The dataset contains data from different classes of plants in different conditions Rusted or dry leaves. In addition, the data on Sheet can be collected at night or during the day. After collecting the data, extended format preprocessing is applied to the dataset to improve the size and accuracy of the dataset. After preprocessing, the dataset is splitinto two parts. One for training and one for testing. The training dataset trains the model according to your requirements and uses the test dataset to explain the actual solution to the problem. The dataset is randomly split at a ratio of 80/20. 80% of the training dataset and the remaining 20% of the test dataset. The image is hidden in front of the model so that you can see the accuracy of the model. To get an accurate solution, the training and test datasets must be trained with a convolutional neural network of the VGG19 architecture. A variant of the VGG architecture, the VGG 19 features 19 deeply connected layers, consistently providing superior performance compared to other state-of-the-art models. The performance assessment is very accurate with respect to disease detection, the time required to detect the disease, and a total of positive and false-positive results are validated against the test data. After successful validation, the developed model can accurately predict the nature of the disease.

V. Architecture: CNN

Convolution neural networks are generally used in segmentation tasks, image processing and classification etc. Convolution means the sliding of the channel or filter over the image to acquire the knowledge of few critical features of the input image generally the image is nothing but some numerical values placed in a matrix and convolved with the help of a filter with the input image and learn necessary features at multiple stages.

Hidden layers of the CNN consist of pooling layers, Convolution layers, normalization and fully connected layers and the output of fully connected layers is given to the activation function. Activation function makes the model to learn the complicated and complex form data and represents the non-linearity arbitrary functional mapping between the outputs and inputs. Hence, we can generate nonlinear mapping from inputs and outputs using a non-linear activation function and here it is ReLU which was present in SoftMax layer. Pooling is used for reducing the dimensionality of input matrix. Pooling is of three types min pooling, max pooling and average pooling. The name itself represents the functionality of the pooling, max pooling means extracting the maximum value from the matrix, min pooling means extracting the minimum value from the matrix and average pooling means taking the average of all the values from the matrix. Fully connected layers represent the input to the present layer is the weights from the previous layer. The normalization layers is used to normalize activations of the previous layer maintaining the mean activation close to zero.

`

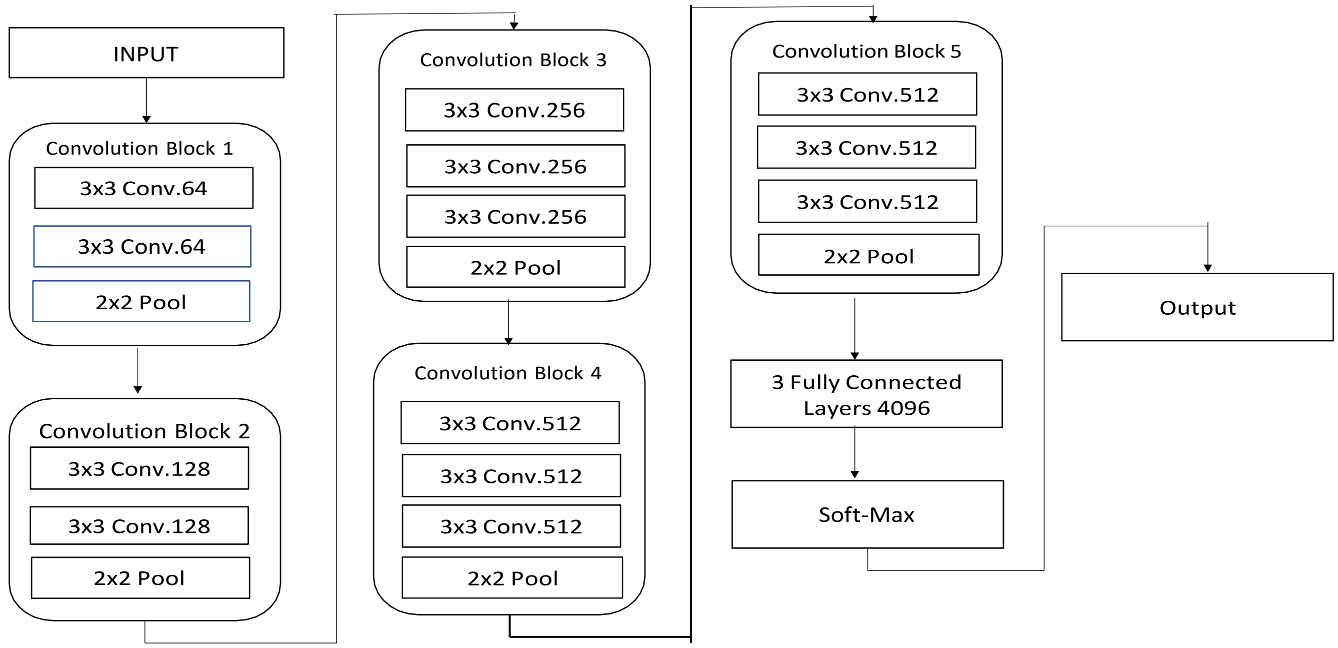


Fig 5: Vgg16 Architecture

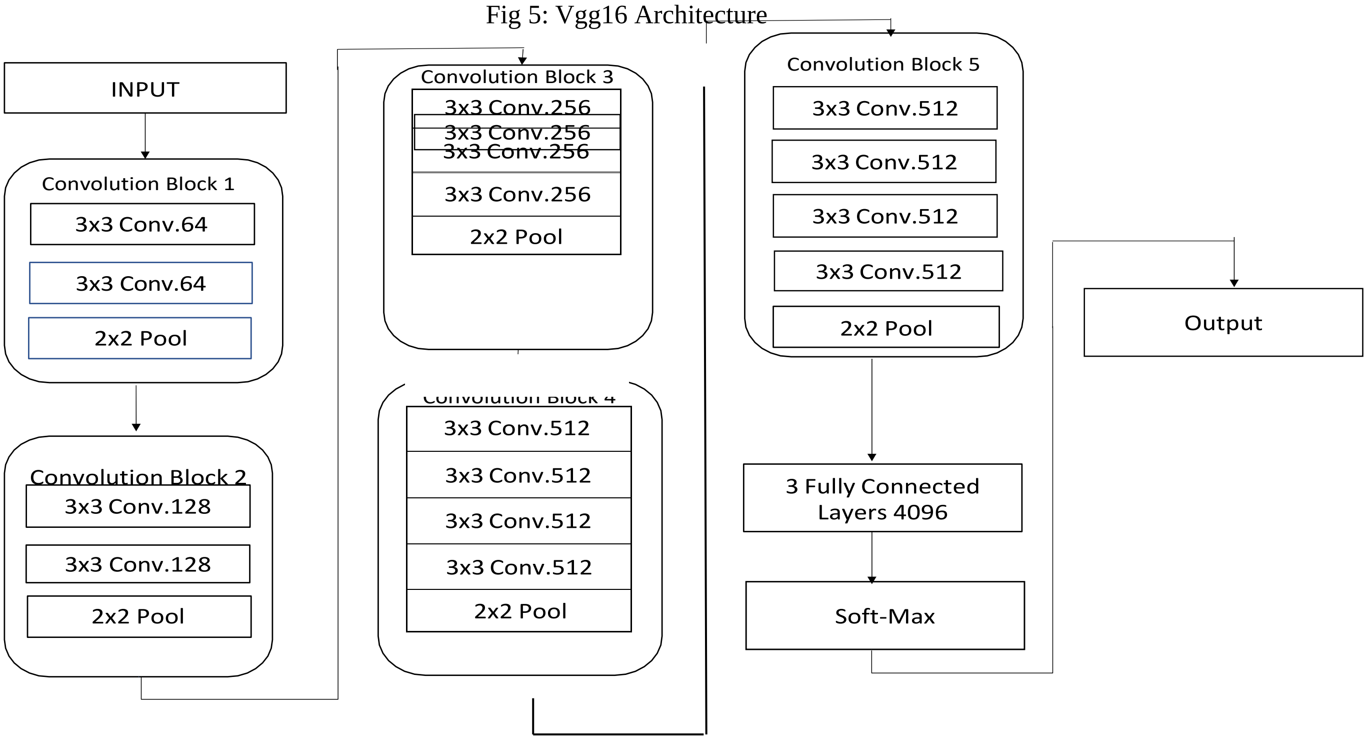


Fig 6: Vgg19 Architecture

VI. Results

In the proposed model the accuracy the accuracy of the models are as following Vgg16 obtained the accuracy of about 93.8 percent, Vgg19 obtained accuracy of about 89.8 percent and InceptionV3 obtained accuracy about 92 percent and these comparisons are explained in table I, in table IIpredicted type of leaf disease is explained and in table III various plots of graphs which includes the accuracy, validation accuracy, training loss and validation loss of all the three models are shown.

Table I: Comparison of results with other models

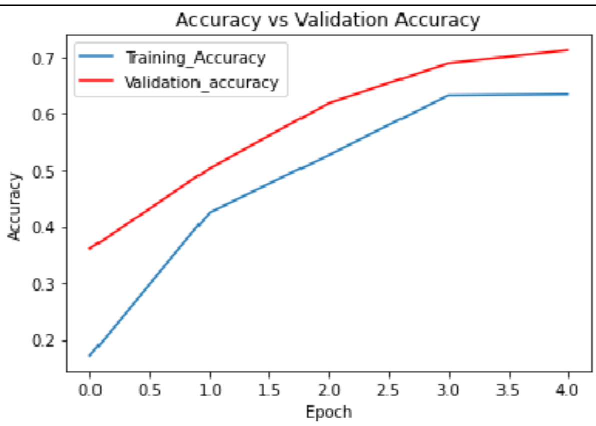
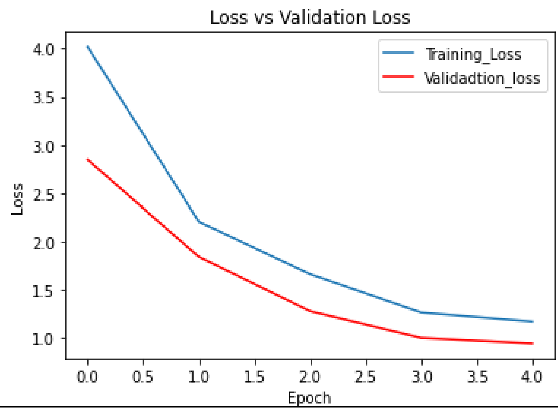
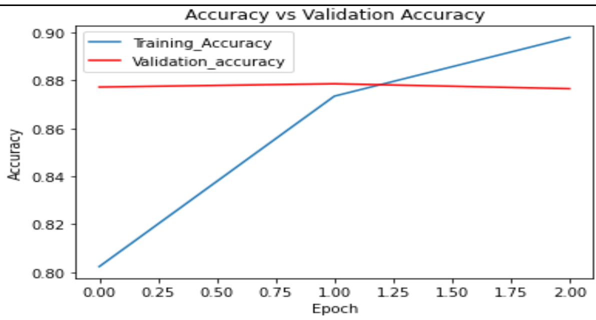
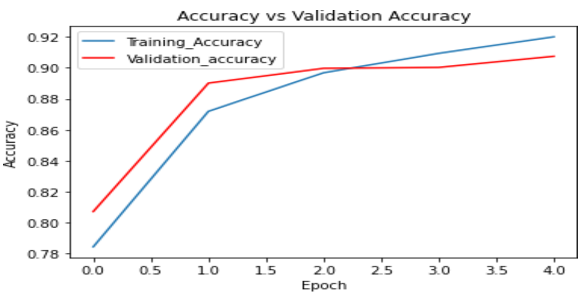
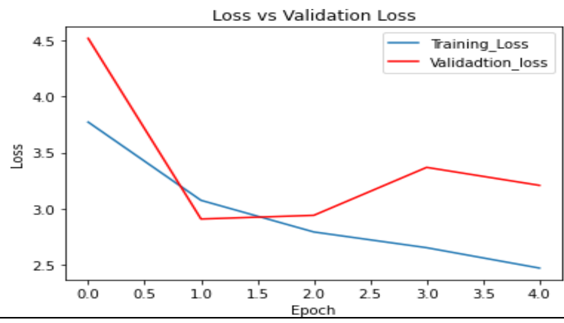
	VGG 16	VGG 19	INCEPTION V3	Average
Accuracy	93.8	89.8	92.00	91.9
Validation Accuracy	93.1	87.6	90.75	89.8
Loss	29.8	48.2	0.24	26.0
Validation Loss	38.7	68.3	0.32	35.8

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 Impact Factor: 8.176 ISSN: 2582-3930

Table II: Results of predicted type of disease

Type of plant Leaf	Type of Disease
	The image belongs to Apple_ _cedar_Apple_rust
	The image belongs to Corn_(maize)__common rust
	The image belongs to Tomato__tomato_Yellow_leaf_curl_virus

Table III: Comparison of graphs with other three models

<table>
<tr><th>Type of CNN</th><th>Plot of accuracy and validation accuracy</th><th>Plot of loss and validation loss</th></tr>
<tr><td>Vgg16</td><td>
</td><td>
</td></tr>
<tr><td>Vgg19</td><td>
</td><td>
</td></tr>
<tr><td>InceptionV3</td><td>
</td><td>
</td></tr>
</table>

VII. Conclusion

This model uses deep learning CNN technique to fulfill the automatic plant leaf disease detection system. Fully connected layers in CNN are utilized for the final prediction. This research is carried out using data sets available publicly which contains 70000 images of 38 classes. The model has achieved 91.8 percent average testing accuracy on publicly available dataset and shows the same accuracy on the images collected manually.

In future this model can be trained with robots or drones for predicting the disease easily without any involvement of humans. In addition, more environmental images can be added for more accuracy and to classify more plant types as well as disease types. Finally, we investigated a new approach to automatically classify and detect plant diseases from leaf images using deep learning techniques.

REFERENCES

[1] Beals KA, "Potatoes, nutrition and health", American Journal of Potato Research, vol.96, no.2, pp-102-110, 2019.

[2] R. K. Arora and S. Sharma, "Pre and Post Harvest Diseases of Potato and Their Management," in Future Challenges in Crop Protection Against Fungal Pathogens (Eds. Aakash Goyal and C.Manoharachary), New York, Springer, 2014, pp. 149-183.

[3] P. TM, P. Alla, K. S. Ashrita, N. B. Chittaragi and S. G. Koolagudi, "Tomato Leaf Disease Detection using Convolutional Neural Network," International Conference on Contemporary Computing (IC3), 2018

[4]C. Xie, Y. Shao, X. Li, and Y. He, ‘‘Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging,’’ 2015.

[5] Shruthi, U., Nagaveni, V., & Raghavendra, B. K. (2019, March). A Review on Machine Learning Classification Techniques for Plant Disease Detection. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 281-284). IEEE

[6] Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala. "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition" , TENCON 2018 - 2018 IEEE Region 10 Conference, 2018. [18] 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 (Jeju, Korea, 28-31 October 2018).

[7] C. Valenzuela, R. B. Baldovino, A. A. Bandala and E. P. Dadios, "Optimization of Photosynthetic Rate Parameters using Adaptive NeuroFuzzy Inference System (ANFIS)," in 2017 International Conference on Computer and Applications (ICCA), Doha, 2017.

[8] J. Shijie, J. Peiyi, H. Siping and s. Haibo, "Automatic detection oftomato diseases and pests based on leaf images," in 2017 Chinese Automation Congress (CAC), Jinan, 2017.

[9] Fuentes, S. Yoon, S. C. Kim and D. S. Park, " A Robust Deep-LearningBased Detector for Real-Time Tomato Plant Diseases and Pests Recognition," Sensors, 2017. Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018)

[10] Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 2012.

[11] Y. Tian, P. Zheng and R. Shi, "The Detection System for Greenhouse Tomato Disease Degree Based on Android Platform," in 2016 3rd International Conference on Information Science and Control Engineering (ICISCE), Beijing, 2016

[12] A. K. Hase, P. S. Aher and S. K. Hase, "Detection, categorization and suggestion to cure infected plants of tomato and grapes by using Open CV framework for andriod environment," in 2 nd International Conference for Convergence in Technology (I2CT),Mumbai, 2017.

[13] 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.

[14] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh and S. Ma, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, 2015.

[15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and F.-F. Li, "ImageNet: a Large-Scale Hierarchical Image Database," in IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[16] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard and W. Hubbard, "Backpropagation Applied to Handwritten zip code Recognition," Neural Computation, vol. 1, no. 4, 1989. https://www.kaggle.com/c/cassava-leaf-disease-classification.