IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

Disease Detection in Plants using Deep Learning

Saji Kumar T. V.
Assistant Professor, Department of Electronics Engineering
College of Engineering, Chengannur, Kerala, India
sajikumartv@gmail.com

Abstract: *Crop growth and yield are essential aspects that influence the field of agriculture as well as farmer economically, socially, and in every possible way. So, it is necessary to have close monitoring at various stages of crop growth to identify the diseases at right time. But, humans naked may not be sufficient and sometimes it would be misleading scenarios arise. In this aspect, automatic recognition and classification of various diseases of a specific crop are necessary for accurate identification. This thought gave inspiration for the present proposed framework. The proposed framework mainly concentrated on the transfer learning phenomenon based on three different pre-trained models such as VGG-16, ResNet-50, and ResNet-50 v2, and then compared the three models based on transfer learning models based on various standard evaluation metrics. The dataset is considered for the implementation is the "PlantVillage" dataset which includes the various diseased and healthy leaves of Pepper, Potato, and Tomato.*

Keywords: Convolutional Neural Network, Visual Geometry Group, Support Vector Machine.

I. INTRODUCTION

Agriculture is an essential sector in countries like India as those countries' economy directly or indirectly dependent on agriculture. It indicates the necessity of taking care of plants from seedling until the expected crop obtains. Through this process, the crop needs to cross a lot of phases to obtain the expected crop such as weather conditions, the survival of the crop from various diseases, and the survival of the crop from various animals. Of these major phases, the crops can be protected from the various animals by providing proper protection for the field and this issue can be solvable. The next major issue is weather conditions which will not be in the control of humans, humans can only pray for better weather conditions to obtain a better crop. Finally, The major issue is very crucial to protect the crop from various diseases as these diseases can impact the complete growth and yield of the crop. If one can able to identify these diseases in time, then the crop can be protected using appropriate fertilizers. If this process of identification and classification of diseases able to digitalize which would be helpful for the agriculturists. It will decrease the time for the identification of diseases and precision in classifying the diseases. Leaves with disease part are shown in fig.1.

Fig. 1 Leaves with disease part

With the development of science, plants were identified as living-thing. It could also respire, reproduce, and even get prone to various diseases. These are different types of diseases by various microorganisms may it be bacteria, viruses, or fungi. Plant diseases can damage crops to a great extent. It can even be fatal to human beings. One such situation emerged in 1840 when a large amount of potato crop was destroyed due to a disease called the Late blight of potato. This is also known as the Irish famine and this was a darker phase of European history, where people were dying of hunger. We all know that plants are very important in our lives so we need to protect them not only from deforestation

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

but also from various plant diseases. In India still, a large population is engaged in agriculture. It ranks 2nd in the world in terms of agriculture production This is again a major challenge. Crops are damaged in India because of various factors one of the major causes is natural calamities and other is the microbial diseases. According to United Nations data 96-billion-dollar loss was in agriculture in a decade (starting from 2005). As humans do not have any control over natural disasters but we can control microbial infections in plants.

Once a crop is infected by some disease it is difficult for farmers to find out the real cause of the disease. pathogens and pests are affecting the crops badly. The crop produces for 5 main food crops by 10 -40%, This data is about a study report which is published by UC Agriculture and Natural Resources.

II. LITERATURE REVIEW

Wallelign, S., et al. in 2018 discussed the viability of convolutional neural network architecture for the classification of various plant diseases with the aid of leaf images [1]. The mentioned framework is implemented by utilizing the LeNet, one of the popular CNN architecture, for disease classification in the aspect of soybean plants. The soybean plant leaf images of 12,763 samples are obtained from the standard database called PlantVillage. The mentioned framework able to achieve an accuracy of 99.32% indicating the viability of CNN with plant disease classification utilizing the leaf images. Sladojevic, S., et al. in 2016 concerned the generation of the new-age model for the identification of various diseases of 13 plant diseases out of the healthier plant leaf images. The deep learning architecture called Caffe was utilized for training the data. The results were obtained from the mentioned framework with a precision of 91percent to 98percent.

Fuentes, A., et al. in 2017 proposed a framework and can be applied in two stages. At first, the meta architectures of Faster R-CNN, R-FCN, and SSD will be combined to form a single meta-architecture. Lastly, certain methodologies such as VGG-16, VGG-19 [3] , and ResNet-50 will be attached to extract the features from more depth and these models' efficiency was estimated. When compared to many other models, the proposed framework efficiency is better. Arivazhagan, S. and Ligi, S. V. in 2018 proposed a framework based on automated deep learning for the recognition and classification of various diseases in mango plants. The dataset utilized for this framework consists of 1200 images which include both diseased and healthy leaves of mango. The accuracy obtained from the proposed framework is 96.67%. Oppenheim.

Brahimi, M., et al. in 2017 proposed a framework based on a convolutional neural network for the detection and classification of various diseases in the tomato crop. The dataset utilized for this framework consists of 14,828 tomato leaf images with almost nine diseases from the plant village image database. The proposed framework able to achieve an accuracy of 99.18%. Shrivastava, V. K., et al. in 2019 [9]focused on the detection and classification of various diseases in the rice plants using a framework with the aid of CNN architecture along with SVM. The framework was implemented on the dataset consists of 619 rice plant leaf images with all four categories of diseases. The accuracies are evaluated for various proportions of training and testing datasets and the maximum accuracy achieved is 91.37%.

Vijai Singh et.al [2017] An advancement of genetic algorithm is proposed by the author named minimum distance algorithm to find the infected plant part of the plant that is to perform image segmentation. After the image segmentation step the author has checked the accuracy of the algorithm with other classification algorithms like k mean clustering and SVM [17].

Konstantinos 2018 In this paper author has used a convolution neural network technique to identify various plant diseases. A detailed study has been done by the author. Images of various plant leaves are taken which includes both the infected leaves images and healthy leaves images and then the author has classified it in various classes and all CNN architectures gave more than97% accuracy. The CNN architectures include AlexNet, AlexNetOWTBn, GoogLeNet, Overfeat, VGG.

KamleshGolhani et.al, 2018 In this paper author has done a detailed review of various deep learning algorithms along with their advantages and disadvantages also their optimization techniques. A comparison has also been made for these techniques about the related workHarshalWaghmare and RadhaKokare 2016[25] the author in this paper proposed a framework where one could utilize leaf textures for identification of diseases in plants. A coloured image is provided as input to system which is segmented to find the infected region in the image and special section of leaf is obtained. On the basis of features a model based on texture is derieved. The texture of leaf is very much unique of every new

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

category of the leaf disease. This is the data which SVM classifier process to identify the disease. The author in this work used multiclass SVM classifier to categorize and identify the disease in grapes plant leaf images. The image pattern is then classified as an SVM designation for multiclass in groups who are healthy or ill classes respectively. The proposed study mainly concentrates commonly and worst affecting disease which are downy mildew disease & black red. The recommendation system proposed in this work quickly provides tenants with expert guidance with 96.6 percent of accuracy.

III. EXISTING METHODOLOGY

General block diagram of feature based approach is shown in fig.2. After pre-processing the captured image, segmentation will take place and sufficient features will be extracted. Finally classification will take place.

Image Acquisition: image acquisition in the first load the image in digital picture process and that consist capturing the image through digital camera and stores it in digital media for additional MATLAB operations [5].

Image Preprocessing: The main aim of image pre-processing is to enhance the image information contained unwanted distortions or to reinforce some image features for any processing [3]. Preprocessing technique uses various techniques like dynamic image size and form, filtering of noise, image conversion, enhancing image and morphological operations[2].

Image Segmentation: In image segmentation is used K-means cluster technique for partitioning of pictures into clusters during which a minimum of one part of cluster contain image with major space of unhealthy part [7]. The k means cluster algorithmic rule is applied to classify the objects into K variety of categories per set of features [3].

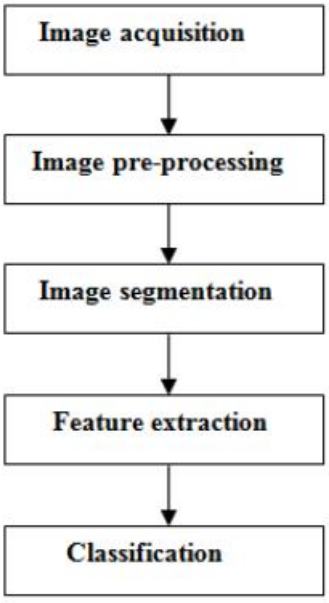

Fig.2 Feature-Based Approach

Feature extraction: After clusters are formed texture features are extracted using GLCM [13]. (Gray-Level Co-occurrence Matrix).

Classification: In classification is used for testing the leaf disease. The Random forest classifier is used for classification.[3]

IV. PROPOSED METHODOLOGY

Convolutional neural networks (CNN) can be used for the computational model creation that works on the unstructured image inputs and converts to output labels of corresponding. classification[3]. They belong to the category of multi-layer neural networks which can be trained to learn the required features for classification purposes [3].Less pre-processing is required in comparison to traditional approaches and automatic feature extraction is performed for better performance. For the purpose of leaf disease detection, the best results could be seen with the use of a variation of the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

LeNet architecture [13]. LeNet consists of convolutional, activation, max-pooling and fully connected layer also LeNet is simple CNN model. This architecture used for the classification of the leaf diseases in LeNet model [4]. It consists of an additional block of convolution, activation and pooling layers in comparison to the original LeNet architecture. The model used in this paper been shown in Fig. 3..Each block consists of a convolution, activation and a maxpooling layer. Three such blocks followed by fully connected layers and soft-max activation are used in this architecture. Convolution and pooling layers are used for feature extraction whereas the fully connected layers are used for classification. Activation layers are used for introducing non-linearity into the network [3].

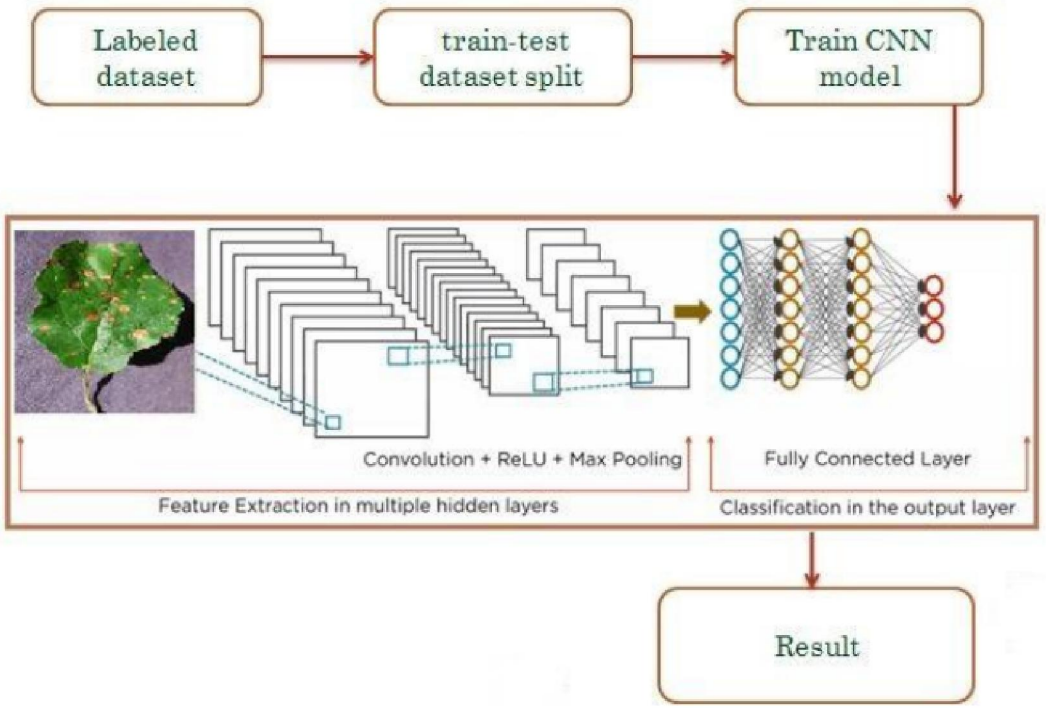

Fig 3: Proposed system

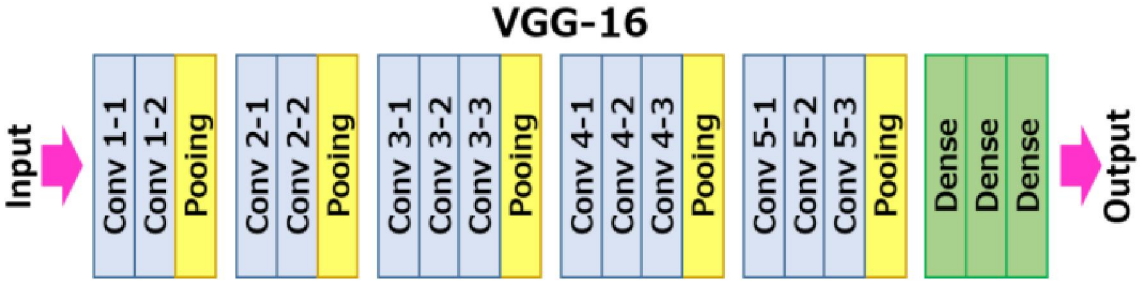

Fig .4 VGG16 layered architecture

Transfer learning is a knowledge- sharing method that reduces the size of the training data, contains 224*224 image fix size. To transfer the learning of a pre-trained model to a new model Transfer learning is useful. Transfer learning has been used in various applications, such as plant classification, software defect prediction, activity recognition and sentiment classification [3]. In this, the performance of the proposed Deep CNN model has been compared with popular transfer learning approach VGG16.

Convolutional neural networks (CNN) can be used for the computational model creation that works on the unstructured image inputs and converts to output labels of corresponding

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

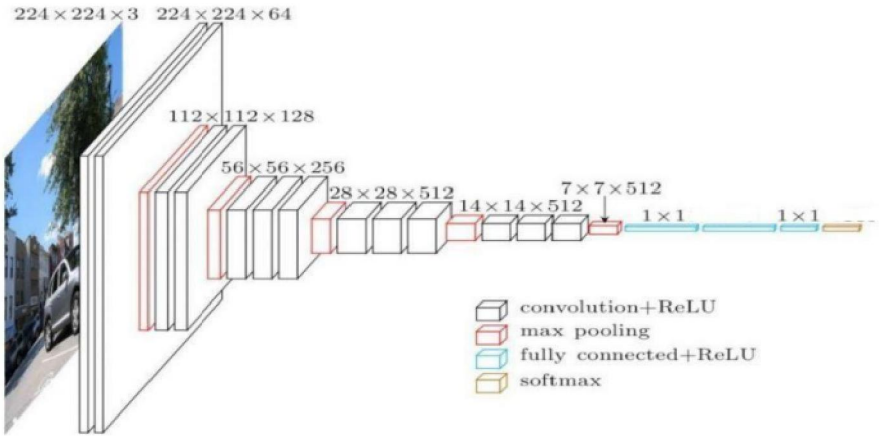

Fig 5. VGG16 architecture

4.1 Dataset

The Plant leaf diseases dataset with augmentation data-set, 39 different classes of plant leaf and background images are available. The data-set containing 61,486 images. We used six different augmentation techniques for increasing the data-set size. These techniques are 1)image flipping,2) Gamma correction, 3) noise injection, 4) PCA color augmentation, 5) rotation, and 6) Scaling[4].We use The Plant leaf diseases dataset with augmentation dataset only 30,052 images with 24labels. The apple label namely: Apple scab, Black rot, apple rust, and healthy. Corn label namely: Corn Cercospora spot Gray spot, Corn rust, Corn healthy, Corn Northern Blight[3].Grape label namely: Black rot, Esca, healthy, and Leaf blight. Potato label namely: Early blight, healthy, and Late blight. Tomato label namely: bacterial spot, early blight, healthy, late blight, leaf mold, septoria leaf spot, spider mite, target sport, mosaic virus, yellow leaf curlvirus[4].

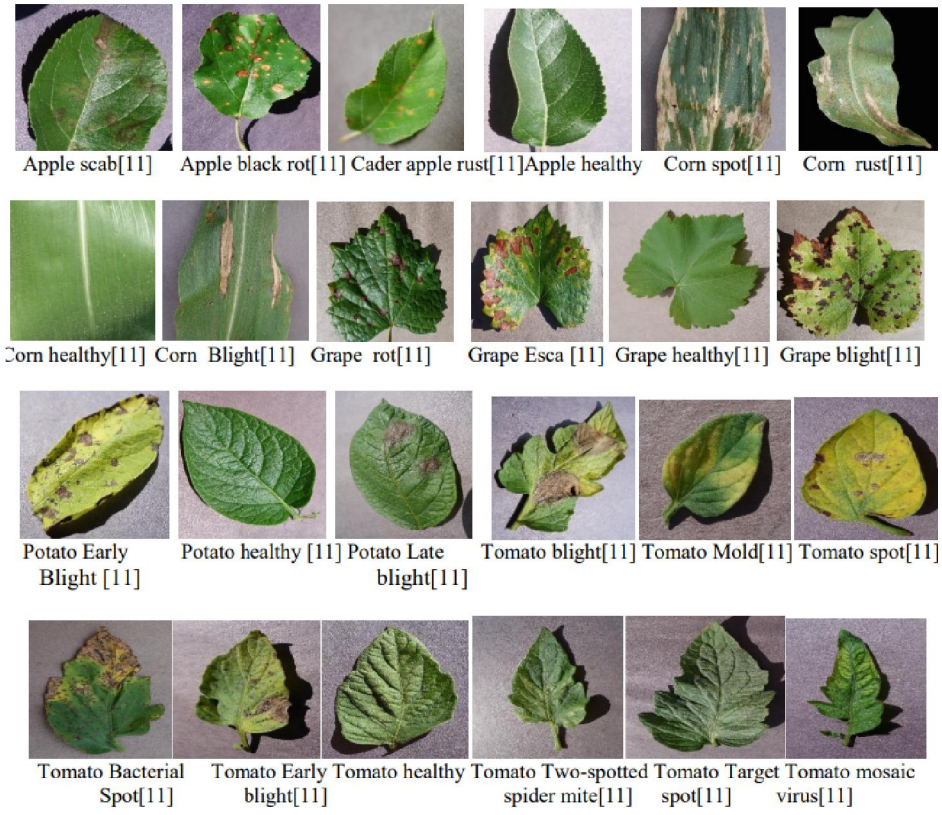

Fig.6 Vegetable and fruits leaves with diseases

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

4.2 Implementation

Tools and technologies used for implementation are:

JUPITER NOTEBOOK:-The Jupyter Notebook is an open-source web application that enables you to make and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more [23].

PYTHON:-Python as a language has a vast community behind it. Any problem which may be faced is simply resolved with a visit to Stack Overflow. Python is among the foremost standard language on the positioning that makes it very likely there will be straight answer to any question Python has an abundance of powerful tools prepared for scientific computing Packages like NumPy, Pandas and SciPy area unit freely available and well documented. Packages lik ethese will dramatically scale back, and change the code required to write a given program. This makes iteration fast. Python as a language is forgiving and permits for program that appear as if pseudo code. This can be helpful once pseudo code given in tutorial papers must be enforced and tested. Using python this step is sometimes fairly trivial. However, Python is not without its errors. The language is dynamically written and packages are area unit infamous for Duck writing. This may be frustrating once a package technique returns one thing that, for instance, looks like an array instead of being an actual array. Plus the actual fact that standard Python documentation does not clearly state the return type of a method, this can lead to a lot of trials and error testing that will not otherwise happen in a powerfully written language. This is a problem that produces learning to use a replacement Python package or library more difficult than it otherwise may be.

NUMPY:-Numpy is python package which provide scientific and higher level mathematical abstractions wrapped in python. It is the core library for scientific computing, that contains a strong n dimensional array object, provide tools for integrating C, C++ etc. It is additionally useful in linear algebra, random number capability etc. Numpy's array type augments the Python language with an efficient data structure used for numerical work. Numpy additionally provides basic numerical routines, like tools for locating Eigenvectors.

SCIKIT LEARN:-Scikit-learn could be a free machine learning library for Python. It features numerous classification, regression and clustering algorithms like support vector machine, random forests, and k-neighbors', and it additionally supports Python numerical and scientific libraries like NumPy and SciPy. Scikit-learn is especially written in Python, with some core algorithms written in Python to get performance. Support vector machines are enforced by a python wrapper around LIBSVM .i.e., logistic regression and linear support vector machines by a similar wrapper around LIBLINEAR.

TENSORFLOW:-TensorFlow is an open source software library for numerical computation using data flow graphs. Nodes inside the graph represent mathematical formula, whereas the graph edges represent the multidimensional knowledge arrays (tensors) communicated between them. The versatile architecture permits you to deploy computation to at least one or more CPUs or GPUs in a desktop, server, or mobile device with a single API. Tensor Flow was originally developed by researchers and engineers acting on the Google Brain Team at intervals Google's Machine Intelligence analysis organization for the needs of conducting machine learning and deep neural networks research, however, the system are general enough to be applicable in a wide range of alternative domains as well.

KERAS:-Keras is[20] a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with attention on enabling quick experimentation. Having the ability to travel from plan to result with the smallest amount doable delay is key to doing great research. Keras permits for straightforward and quick prototyping (through user-friendliness, modularity, and extensibility).Supports each convolutional networks and recurrent networks, furthermore as combinations of the two Runsseamlessly on CPU and GPU. The library contains numerous implementations of usually used neural network building blocks like layers, objectives, activation functions, optimizers, and a number of tools to create operating with image and text data easier. The code is hosted on GitHub, and community support forums embody the GitHub issues page, a Glitter channel and a Slack channel.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

V. RESULTS

Images are resized and assigned label name to all images during the training process.For validation, split dataset images into a training set and testing set. There are 1500 dataset images and are divided into two sets 1200 are in the training set and 300 are in the testing set.

TABLE 1. Comparison table of CNN vs. VGG16

Epochs	CNN Accuracy	VGG16 Accuracy
150	90.229166 %	51.166334 %
120	86. 666667%	50.140005 %
90	86.133333 %	47.157776 %
60	85.085556 %	46.872223 %
30	84.166664 %	45.708333 %

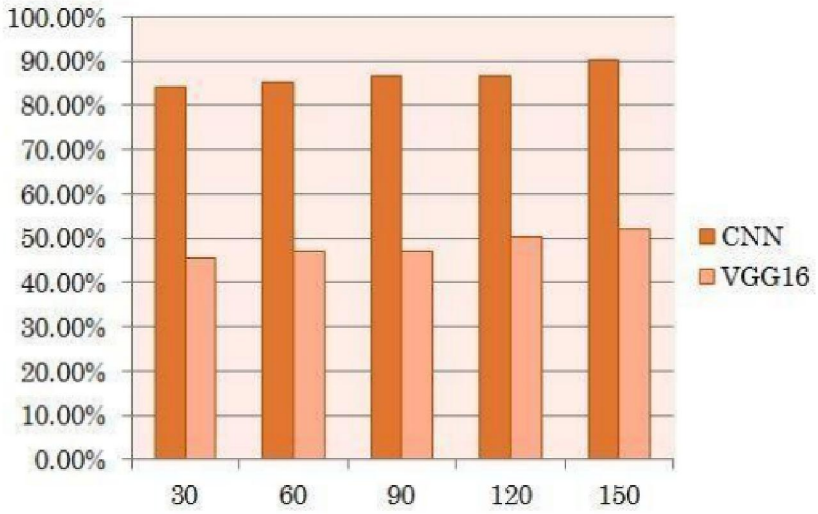

Fig 7. Chart of CNN vs. VGG16

VI. CONCLUSION

The Plant leaf diseases dataset with augmentation data set is used to predict plant diseases. The dataset implementation is the "PlantVillage" dataset which includes the various diseased and healthy leaves of Pepper, tato, and Tomato. Transfer learning method that reduces the size of the training data, contains 224*224 image fix size. In his analysis, the performance of the proposed Deep CNN model has been compared with the popular transfer learning approach VGG16. The result shows the CNN implementation has a better accuracy of 90.23%compared to the VGG16 model to detect leaf diseases.

REFERENCES

[1]. Abirami Devaraj, Karunya Rathan, Sarvepalli Jaahnavi and K Indira, "Identification of Plant Disease using Image Processing Technique," International Conference on Communication and Signal Processing, IEEE 2019.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 19, May 2023

[2]. Velamakanni Sahithya, Brahmadevara Saivihari, Vellanki Krishna Vamsi, Parvathreddy Sandeep Reddy and Karthigha Balamurugan, "GUI based Detection of Unhealthy Leaves using Image Processing Techniques," International Conference on Communication andSignal Processing 2019.

[3]. Balakrishna K Mahesh Rao, "Tomato Plant Leaves Disease Classification Using KNN and PNN," International Journal of Computer Vision and Image Processing 2019.

[4]. Masum Aliyu Muhammad Abdu, Musa Mohd Mokji, Usman Ullah Sheikh, Kamal Khalil,"Automatic Disease Symptoms Segmentation Optimized for Dissimilarity Feature extraction in Digital Photographs of Plant Leaves," IEEE 15th International Colloquium on Signal Processing & its Applications 2019.

[5]. Suja Radha, "Leaf Disease Detection using Image Processing," Article in Journal ofChemical and Pharmaceutical Sciences, March 2017.

[6]. Sneha Patel, U.K Jaliya, Pranay Patel, "A Survey on Plant Leaf Disease Detection,"International Journal for Modern Trends in Science and Technology, April 2020.

[7]. Priyanka Soni, Rekha Chahar, "A Segmentation Improved Robust PNN Model for DiseaseIdentification in Different Leaf Images," 1st IEEE International Conference on PowerElectronics, Intelligent Control and Energy Systems (ICPEICES-2016).

[8]. S. Arivazhagan, R. Newlin Shebiah S. Ananthi, S. Vishnu Varthini, "Detection ofunhealthy region of plant leaves and classification of plant leaf diseases using texturefeatures," CIGR Journal, March 2013.

[9]. AdedamolaAdedoja& Pius AdewaleOwolawi&TemitopeMapayi, "Deep LearningBased on NAS Net for Plant Disease Recognition Using Leave Images," 2018.

[10]. Prajwala TM, AllaPranathi, Kandiraju Sai Ashritha, Nagaratna B. Chittaragi, ShashidharG. Koolagudi, "Tomato Leaf Disease Detection using Convolutional Neural Networks,"Proceedings of 2018 Eleventh International Conference on Contemporary Computing(IC3), 2018