A Project Report
on

Crop Disease Detection and Diagnosis

by

1. Aayush Singh
2. Deepanshu Sonparote
3. Nitish Talekar

under the guidance of

Prof. Dilip S. Kale

MCT
MANJARA CHARITABLE TRUST
RAJIV GANDHI INSTITUTE OF TECHNOLOGY, MUMBAI

Department of Computer Engineering

University of Mumbai

April - 2020

MANJARA CHARITABLE TRUST
RAJIV GANDHI INSTITUTE OF TECHNOLOGY, MUMBAI

Juhu-Versova Link Road Versova, Andheri(W), Mumbai-53.

Certificate

Department of Computer Engineering

This is to certify that

1. Aayush Singh
2. Deepanshu Sonparote
3. Nitish Talekar

Have satisfactory completed this project entitled

Crop Disease Detection and Diagnosis

Towards the partial fulfillment of the

BACHELOR OF ENGINEERING
IN
(COMPUTER ENGINEERING)
as laid by University of Mumbai.

Guide | **Head of Department**

Prof. Dilip S. Kale | **Dr.Satish Y. Ket**

Principal

Dr.Sanjay Bokade

Project Report Approval for B. E.

This project report entitled **Crop Disease Detection and Diagnosis** by **Aayush Singh, Deepanshu Sonparote and Nitish Talekar** is approved for the degree of Computer Engineering.

Examiners

1. ______________________

2. ______________________

Date: Place:

Declaration

We wish to state that the work embodied in this project titled "Crop Disease Detection and Diagnosis" forms our own contribution to the work carried out under the guidance of "Prof. Dilip S. Kale" at the Rajiv Gandhi Institute of Technology.

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. we also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. we understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Students Signatures)
Aayush Singh (858)
Deepanshu Sonparote (859)
Nitish Talekar (862)

Abstract

The crop agriculture industry faces economic losses due to pest infections, bacterial or viral contagions. Farmers lose nearly 10-20% of the total profit on an average, annually in India. This project proposes a solution to the aforementioned agricultural problem by applying crop disease recognition using deep learning techniques. In this project, the system sets out to classify crop leaf images of different diseases, and also whether the crop is infected or not. Also, we endeavour applications that give the farmers readily available means to identify the diseases on their crop and take appropriate damage control actions. This project uses a dataset of 14 crops and their 26 underlying diseases to help train the model for precise disease detection. The trained models employing Neural Network techniques have provided a performance reaching a 99% success rate in identifying the corresponding crop diseases. The project also provides relative precautionary information to tackle the diseases that are recognized by the model. The application of this project is a useful advisory warning tool for the farmers and agricultural workers for the identification of diseases in the early stage so that immediate action can be taken. This project can cut down loss of crops on a large scale thereby helping the agricultural industry and decreasing economic loss from the same.

Contents

List of Figures

List of Tables

List of Algorithms

This page contains list of algorithms

List of Acronyms

This page contains list of acronyms

List of Symbols

This page contains list of symbols

Chapter 1

Introduction

1.1 Introduction Description

Agriculture is the primary occupation in India. Crops are part and parcel of the agricultural industry. Nowadays, a tremendous loss in the quality and quantity of crop yield is observed due to various diseases in the plants. Crop disease classification is a critical step, which can be useful in early detection of pest, insects and diseases and help in disease control and productivity boost, among other examples. In today's times, farmers try to recognize diseases manually with foregoing symptoms of plants, but the various diseases are hard to distinguish with the naked eye, and it is time-consuming to predict whether the crop is healthy or not. Symptoms of diseases in cotton predominantly come out on leaves of plants. Diagnosing the crop disease symptoms on plant leaves incorporates a high degree of complexity. Due to this complexity, even experienced agronomists and plant pathologists often fail to successfully diagnose specific diseases and are consequently led to mistaken conclusions and treatments.

The existing techniques for disease detection have utilized various image processing methods followed by a wide variety of classification techniques. Crop Disease Detection has been an area of interest for producers, agricultural-related organizations. Economic and agricultural losses can be minimized by genuine disease detection.

This project proposes to help predict crop diseases in plants by processing the images of the crop leaf. For this, Image Processing techniques are used for the quick, accurate and appropriate classification of diseases. The proposed system in this project uses an artificial neural network to classify the health of a crop and predict the disease it is infected by. The existence of an automated system for the detection and diagnosis of plant diseases would offer a support system to the agronomists who are performing experiments based on diagnosis through observation and farmers who need to prevent infections and diseases in their crop yield.

1.2 Organization of Report

- **Ch.1 Introduction:** An introduction to motivation for the project, proposed plan and system and a brief description of existing system.

- **Ch.2 Review of Literature:** The survey of existing systems and current methods of dealing with the problem of crop diseases.

- **Ch.3 Proposed System:** The detailed description of proposed system using Classification techniques with help of neural networks to determine disease in crops.
- **Ch.4 Implementation Plan:** The proposed time allocation and plan of action for the development of project.
- **Ch.5 Simulation, Results & Analysis:** The resulting output system and analysis of results obtained during development.
- **Ch.6 Conclusion:** The final results and conclusion of results obtained during the development and testing of project.

Chapter 2

Literature Review

Crop Yield Forecasting has been an area of interest for producers, agricultural-related organizations. Timely and accurate crop yield forecasts are essential for crop production. The existing techniques for disease detection have utilized various image processing methods followed by various classification techniques. However, some unconventional approaches have led to classification of diseases using unconventional factors. Chopda et al. [1] propose a system which can predict the cotton crop diseases using decision tree with the help of the parameters like temperature, soil moisture, etc. based on the previous year data and through sensors. However, these data might not be fully dependable to predict or classify diseases.

Image classification and regression techniques play a very important role because it allows identifying, group, and properly of organisms from a standardized system. We apply an algorithm for image segmentation technique on data for automatic detection and classification of plant leaf diseases. In [2], Kamble defines the application of texture analysis for detecting plant diseases with the help of different image processing technique. Further with the use of Decision-Making Module, the disease is classified. Singh and Misra [8] suggest different diseases classification techniques that can be used for plant leaf disease detection and an algorithm for image segmentation, the advantage of using this method is that the plant diseases can be identified at an early stage or the initial stage.

Deep learning is a set of learning methods attempting to model data with complex architectures combining multiple non-linear transformations. The element of deep learning is the neural networks that are combined to form the deep neural networks. These techniques have enabled significant progress in the fields of image processing and image classification. Kulkarni [3] formulates an application of Deep Convolutional Neural Network to identify and classify crop disease on images, testing it on five classes of crops and three types of diseases for each class. Mique Jr, Eusebio L [4] proposed an application that will help farmers in detecting cotton insect pests and diseases using Convolutional Neural Network (CNN) and image processing. The searching and comparison of captured images to a stack of cotton pest images was implemented using a model based on CNN. Collected images were pre-processed and were used in training the model.

There exist several types of architectures for neural networks:

1. The multilayer perceptions, that are the oldest and simplest ones,

2. The Convolutional Neural Networks (CNN), particularly adapted for image processing

3. The recurrent neural networks and dense neural network used for sequential data such as text or times series.

A different approach is taken by Petrellis [6] where mobile phone application for plant disease diagnosis is presented which is based on the detection of the disease signature that is expressed as a number of rules that concern the color, the shape of the spots, historical weather data among other factors..

2.1 Survey Existing system

2.1.1 Plant Disease Detection using CNN

CNN architectures vary with the type of the problem at hand. The proposed model consisted of three convolutional layers each followed by a max-pooling layer. The final layer was fully connected MLP. ReLu activation function was applied to the output of every convolutional layer and fully connected layer. The first convolutional layer filters the input image with 32 kernels of size 3x3. After max-pooling is applied, the output is given as an input for the second convolutional layer with 64 kernels of size 4x4. The last convolutional layer has 128 kernels of size 1x1 followed by a fully connected layer of 512 neurons. The output of this layer is given to softmax function which produces a probability distribution of the four output classes. The model is trained using adaptive moment estimation (Adam) with batch size of 100 for 1000 epochs.

2.1.2 Cotton Blast Disease Recognition Using a CNN and SVM

The SVM is a powerful classifier that works well on a wide range of complex classification problems. SVM with different kernel functions can transform a nonlinear separable problem into a linear separable problem by projecting data into a higher dimensional space to maximize the classification distance and achieve the desired classification. In this study, the radial basis function (RBF), a popular kernel function of SVM, is chosen as the kernel function. The LIBSVM, as an efficient open source tool, is chosen to build SVMs in experiments. CNN works as a trainable feature extractor and SVM performs as a recognizer. This hybrid model automatically extracts features from the raw images and generates the predictions.

2.2 Limitation Existing system or research gap

- No public platform to help farmers
- Does not use actual real life plant dataset

The occurrence of cotton disease is regular, and the type and the probability of the cotton disease vary with the stages of cotton growth. Therefore, different cotton disease identification systems should and can be established using the method presented by this study, and then the automated cotton disease diagnosis can be realized by combining identification models and domain knowledge of cotton disease. Although our method of automatic identification of cotton blast has achieved satisfactory results, substantial further work is needed to improve its accuracy and reliability in cotton disease diagnosis systems. In particular, we plan to address the following two issues in future studies:

1. Expand the dataset of cotton disease, and establish a comprehensive tool for cotton disease diagnosis system. Te data augmentation method will be employed for building a good classifier when the number of samples is insufficient.
2. Study other deep neural network architectures and take full advantage of the deep learning algorithms to improve the classification accuracy, and enhance the reliability and robustness of the cotton disease diagnosis systems.

2.3 Problem Statement and objectives

A tremendous loss in crop yield has been observed globally every year, subject to various diseases which are infested upon them. Detection of plant diseases is an essential topic as it may prove beneficial in monitoring large fields of crops, and thus automatically detect the diseases and provide appropriate diagnosis as soon as they appear on plant leaves. The proposed system is a web-based solution for automatic detection and classification of plant leaf diseases. Once analyzed, the system provides methods and techniques to control and limit the spread of disease.

2.3.1 Objectives

- To compare different models of classification on basis on accuracy and efficiency and select the best fitting model for disease detection in crops.
- To create a web-platform for detection of disease in crops by use of classification.
- To provide treatment and precautionary information for the related disease.

2.4 Scope

The next stage of this project could be to develop an application compatible with smart mobile devices with features such as displaying recognized diseases in plants based on leaf images captured by the mobile phone camera. A helpline and contact details could be added to procure said precautionary measures. The application could offer a discussion forum for farmers and agronomists to discuss the treatment and early precautions of diseases that they have encountered. This could be achieved by clustering the users on the basis of their activity on the application.

Chapter 3

Proposed System

The proposed system is a web-based solution for automatic detection and classification of crop diseases based on leaf images uploaded to the web application. Once analyzed, the system provides methods and techniques to control and limit the spread of disease.

Frontend: The system includes a front-end that displays an option to upload an image of the crop leaf. It will display the image as well as the detected disease of the plant. Along with the results, the symptoms, triggers and precautionary measures for the detected disease are displayed on the screen.

Backend: On the server-side, pre-trained Convolutional Neural Network models are to be used. These models are trained with sufficient accuracy for the detection of diseases. The uploaded image is pre-processed using Image Processing and given as input to the model which returns the predicted disease for the crop. Simultaneously, the disease database is accessed to fetch information and precautionary measures for the detected disease.

Classifier: Different classifiers are tested the proposed system and most accurate is chosen to be used as the classification model. Neural Network classifiers are trained on the obtained dataset.

3.1 Framework

The proposed system looks forward to use one of the following frameworks: CNN, DNN, SVM, Random Forest, or ResNet

3.1.1 Tensorflow

TensorFlow is an open-source library for numerical calculation and provides an enormous scope in Machine Learning. It has a huge number of machine learning, deep learning models and algorithms. It makes it easier to collect data, prepare models, give predictions, and refine future results.TensorFlow can train and run deep neural networks for image classification and other applications. TensorFlow creates data flow graphs i.e structures that describe how data moves through a graph or a series of processing nodes present in a neural network. Each node in the graph represents a mathematical operation, and each connection, link or edge between nodes is a multidimensional data array or tensor. The data flows through this link adjusting the values of the nodes until the requirements are met.TensorFlow offers additional facilities to debug and gain insight into TensorFlow objects. The execution mode can evaluate and modify each graph operation separately and transparently, instead of constructing the entire graph as a single object and evaluating it all at once. The TensorBoard visualization suite inspects and profiles the way graphs run by an interactive, web-based dashboard.

3.1.2 Flask

Flask is a python based lightweight framework used for developing web UI with python backend. It is quick, easy to use and scalable for complex applications. The Flask framework is used in this project along with HTML and CSS to create a user-friendly UI that makes the use of the classification model easy and simple. Flask creates a backend where the model is loaded, and the uploaded image is processed. The output of the classifier is then displayed.

3.2 Details of Hardware and Software

3.2.1 Software Requirements

- Windows, MacOS or Linux
- Python - Tensorflow, Scipy, Pillow, Flask, Scikit-Learn

3.2.2 Hardware Requirements

- Processor - Intel Core i5+
- RAM - 8GB+
- ROM - 1GB+
- GPU - 2GB+

3.3 Design Details

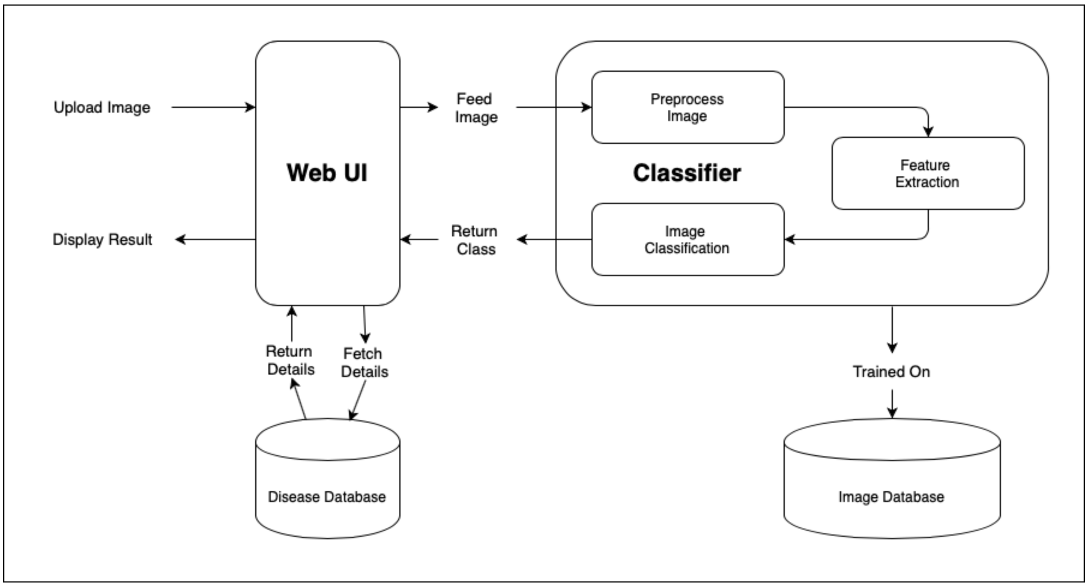


Figure 3.1: System Architecture

3.3.1 Detailed Design/UML Diagrams

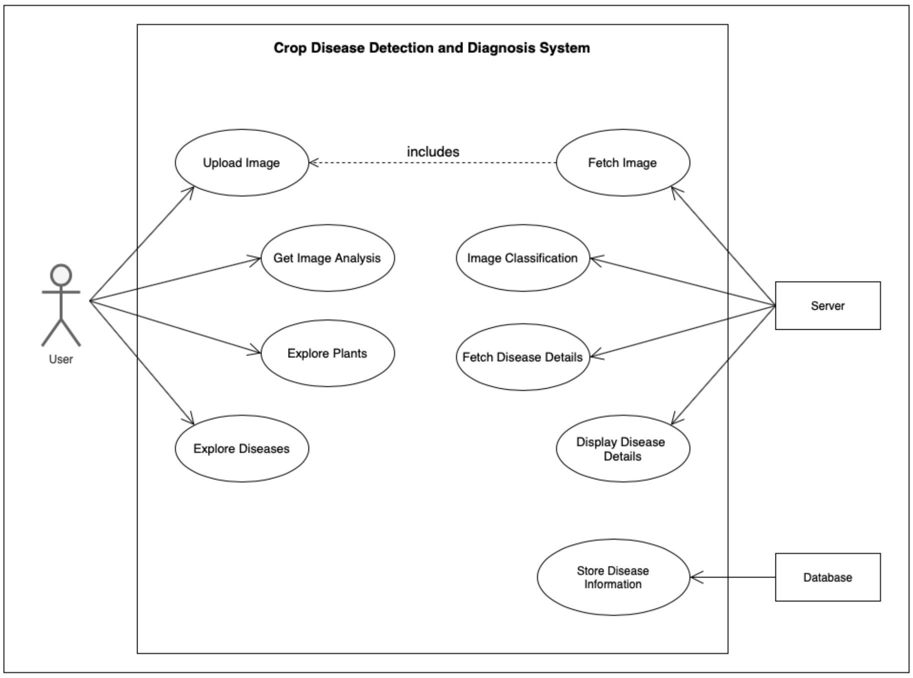


Figure 3.2: Use Case Diagram

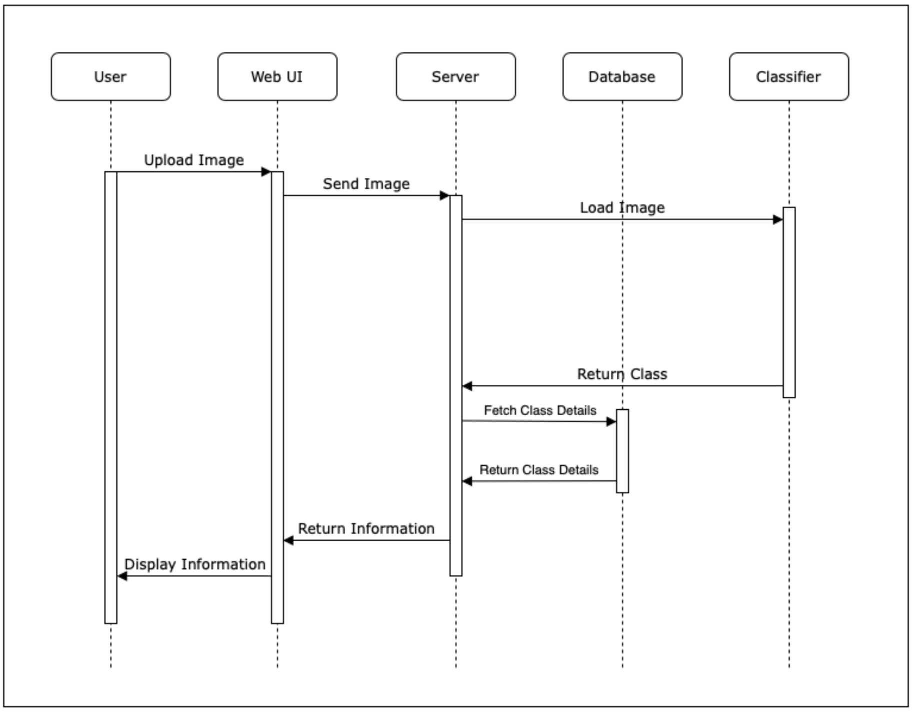


Figure 3.3: Sequence Diagram

3.4 Methodology

Input Images: For the initial process, images [19] of various image sizes are taken from datasets as input by setting IMGW, IMGH to 224 ppi with 3 channels (RGB), for better visibility, preferred with dimensions ¿ 180. The features that are used for the classification of the images.The feature normalization used in the study is the min-max normalization. It is the ratio of the difference between the instance's feature value and the minimum value of a feature in the instance to the difference between the maximum and minimum values of features in the instance.

3.4.1 Dataset

Dataset and their Labels: The dataset contains a total of 53705 images of 38 different classes of 14 crops.

Table 3.1: Dataset Counts

Crop	Sub-Classes	Images
Tomato	10	18160
Apple	4	3171
Corn	4	3852
Grape	4	4062
Potato	3	2152
Bell Pepper	2	2475
Cherry	2	1906
Peach	2	2657
Strawberry	2	1565
Blueberry	1	1502
Orange	1	5507
Raspberry	1	371
Soybean	1	5090
Squash	1	1835
Total	38	54305

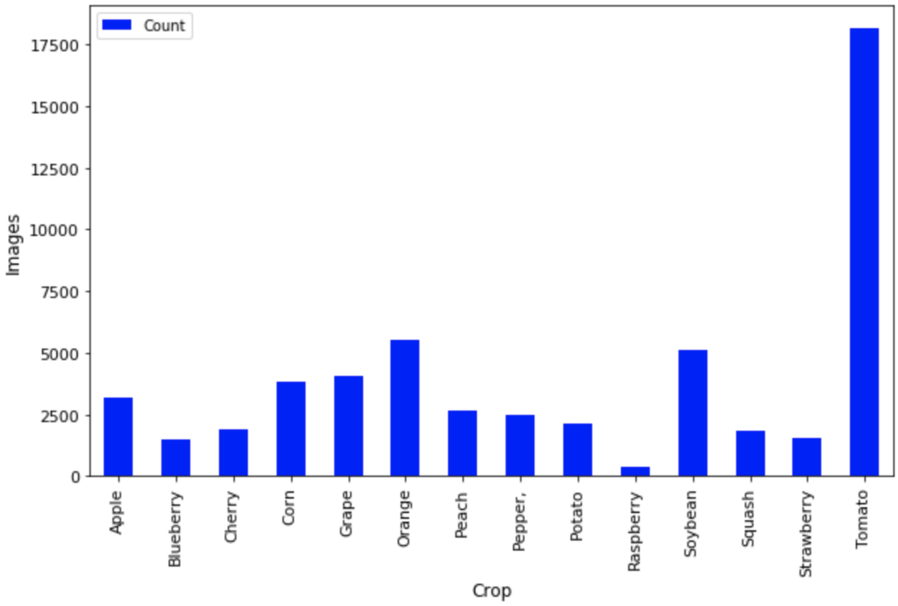


Figure 3.4: Dataset

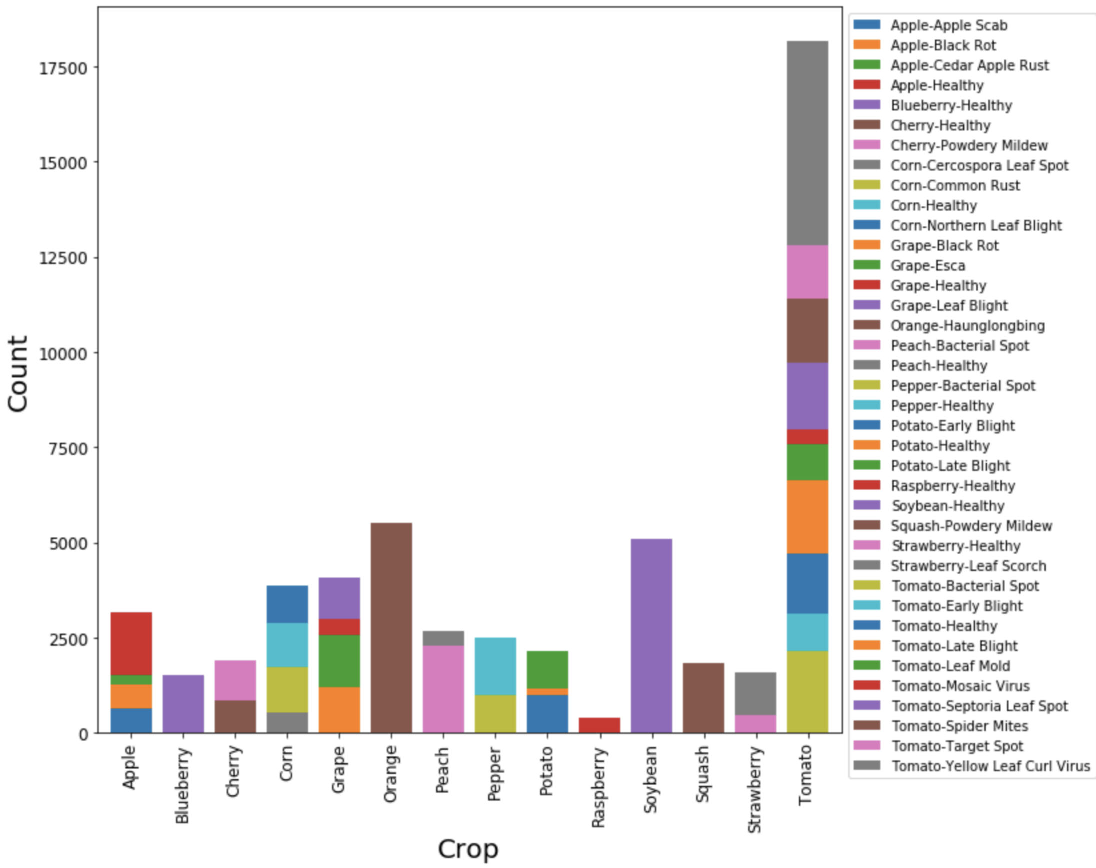


Figure 3.5: Dataset with their Classes

3.4.2 Procedures

Procedure 1:

- Input images are read using the Python Imaging Library (PIL).
- The input image is then resized to IMG_H and IMG_W variable.
- Image Pixel data is loaded into the X list and Image Category to Y list.
- X list and Y list are then converted to NumPy arrays and are normalized.
- The images are then fed to different classification models to get training and validation accuracy.

Procedure 2:

- Input images are processed using Keras ImageDataGenerator Library.
- ImageDataGenerator is given the input data directory
- Validation Spilt size, Normalization Factor, Resize size, Zoom Factor and other parameters are set.
- Training and Validation trains are fed to different classification models to get training and validation accuracy.

3.4.3 CNN Architectures

Alexnet

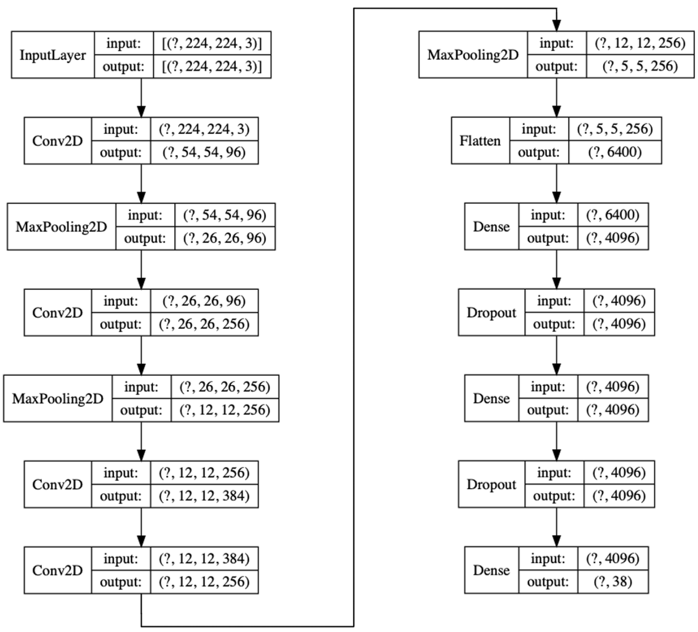


Figure 3.6: Alexnet Architecture

Deep Convolutional Neural Network

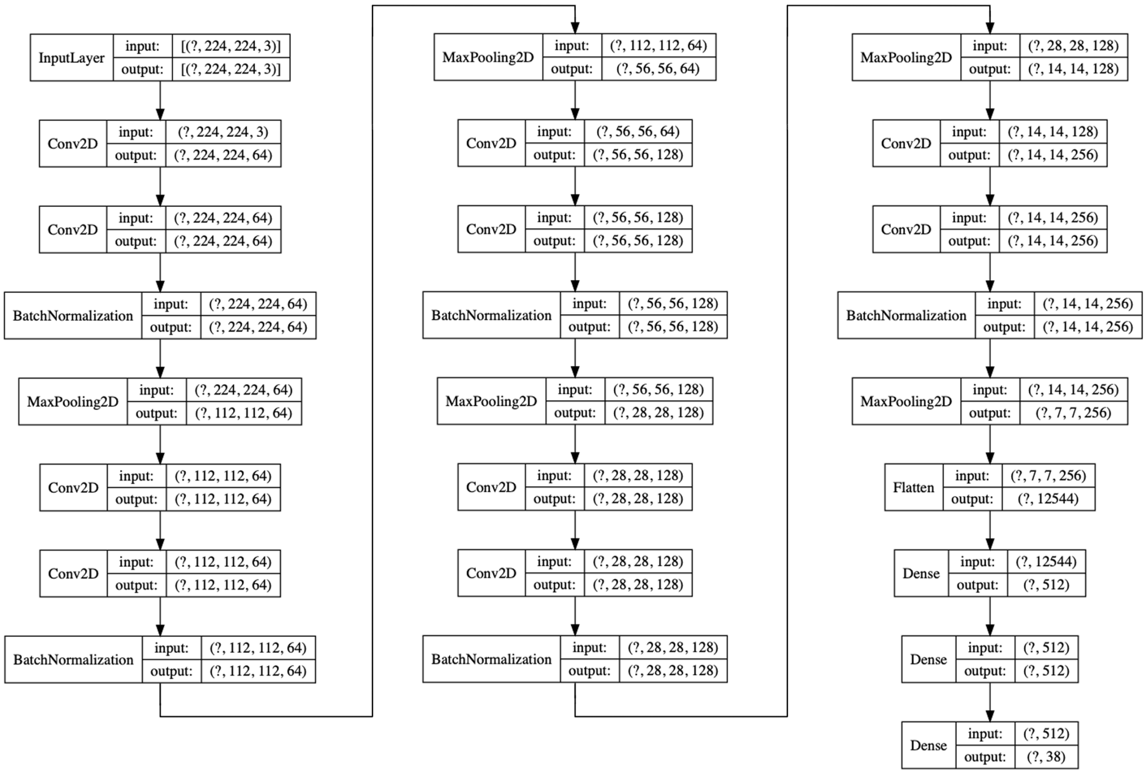


Figure 3.7: Deep CNN Architecture

Residual Network

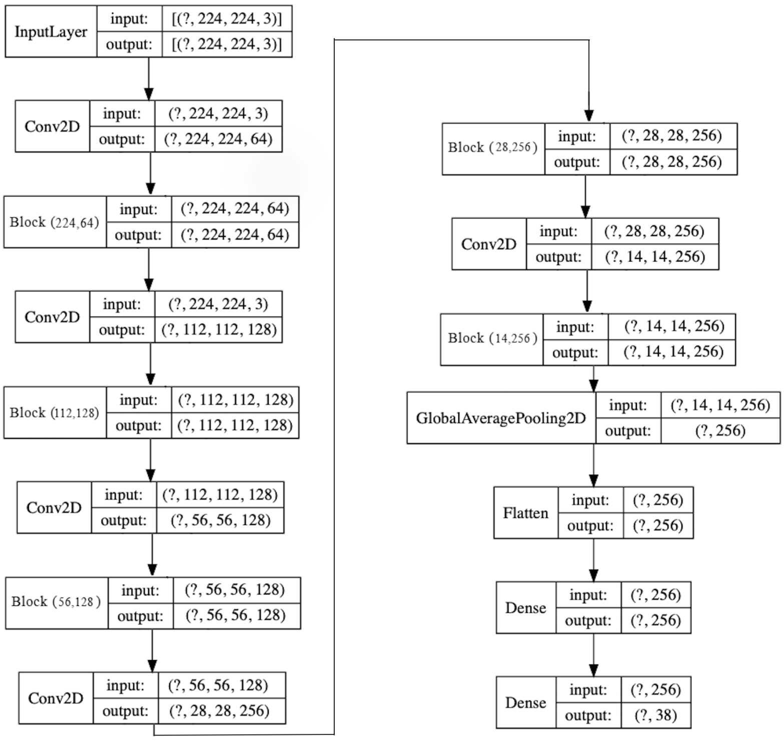


Figure 3.8: Residual Network - Architecture

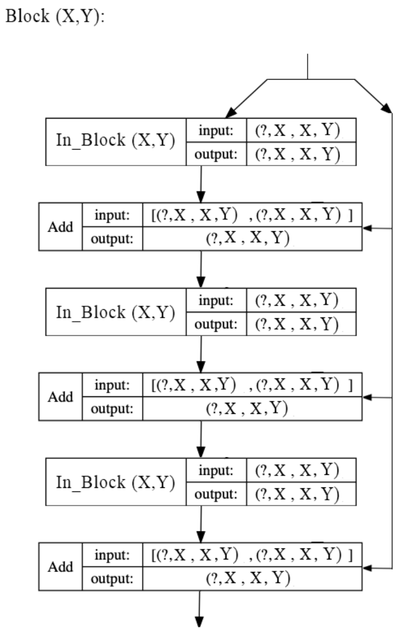


Figure 3.9: Residual Network - Block

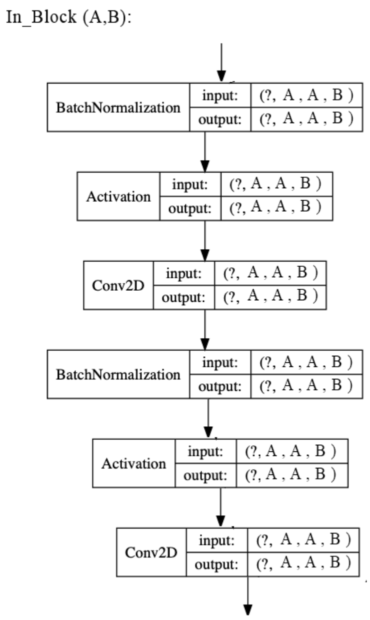
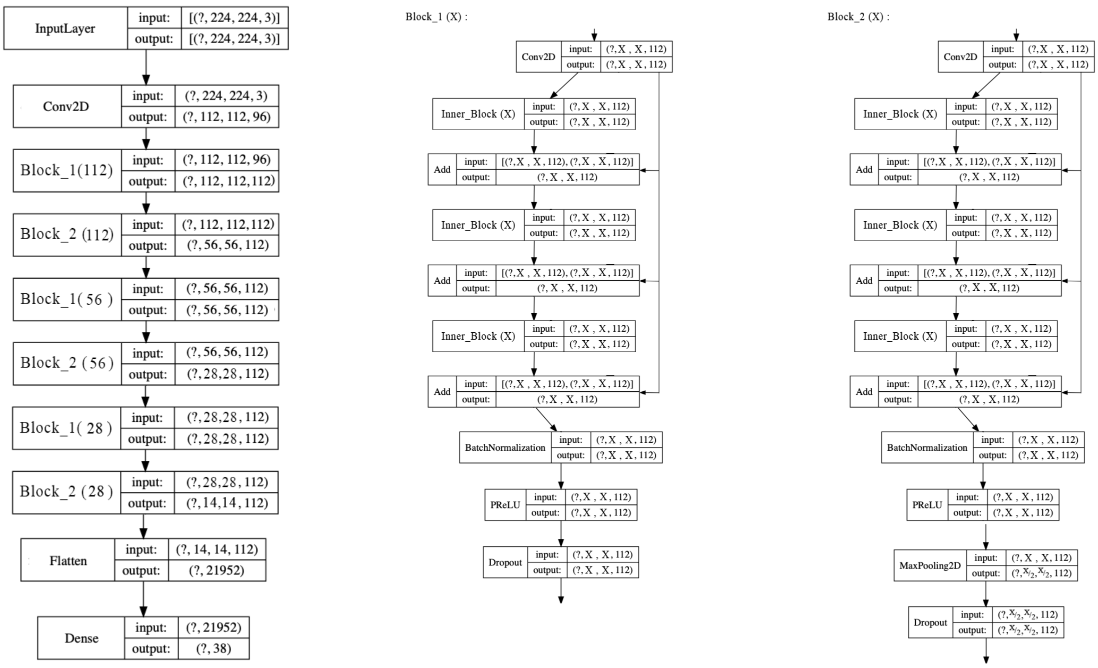


Figure 3.10: Residual Network - Inner Block

Recurrent Convolutional Neural Network



(a) Architecture (b) Block 1 (c) Block 2

Figure 3.11: Recurrent CNN

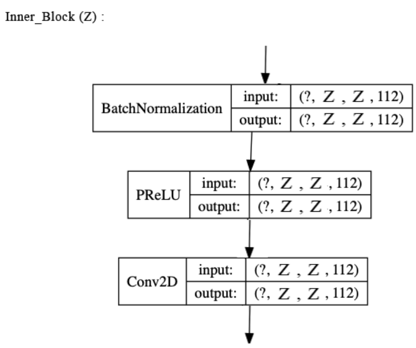


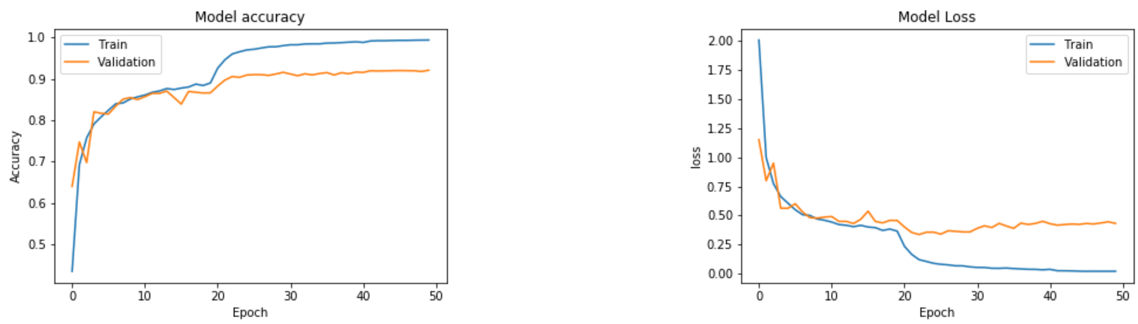
Figure 3.12: Recurrent CNN - Inner Block

Chapter 4

Results & Discussions

4.1 Implemented Algorithm

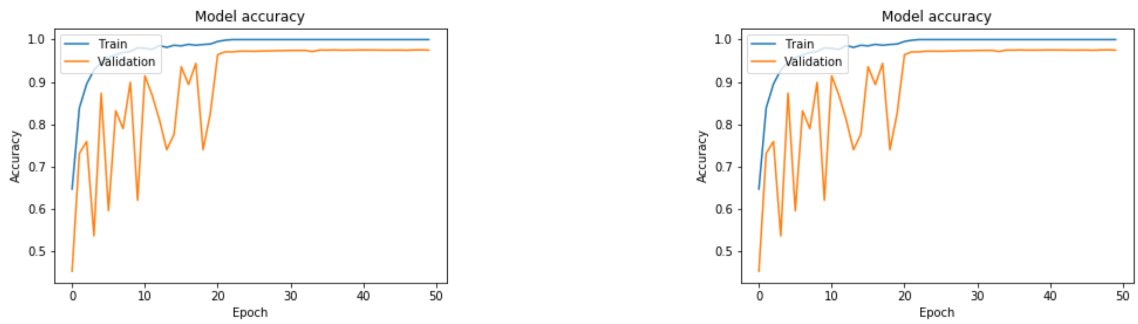
4.2 Results & Analysis



(a) Accuracy Graph

(b) Loss Graph

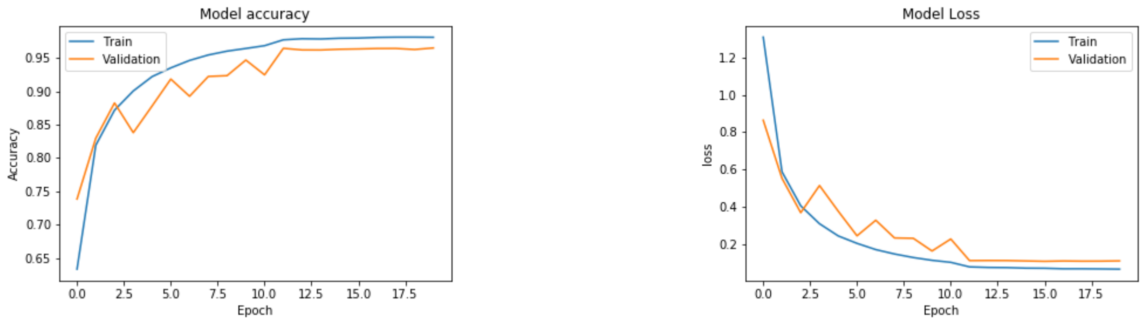
Figure 4.1: Alexnet



(a) Accuracy Graph

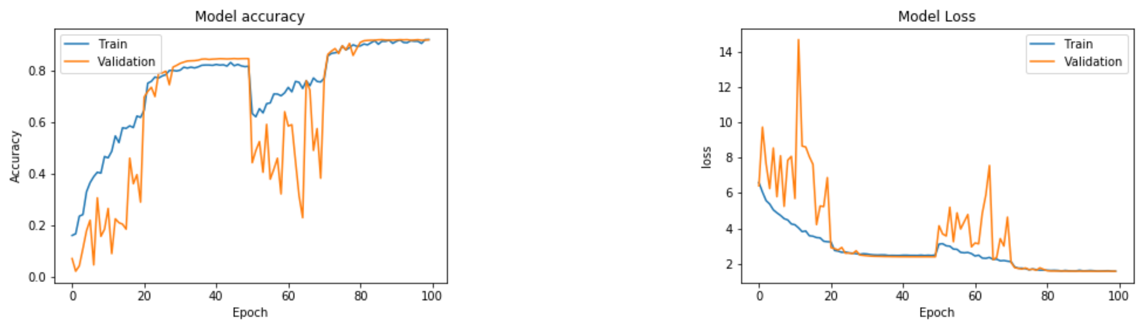
(b) Loss Graph

Figure 4.2: Deep Convolutional Neural Network



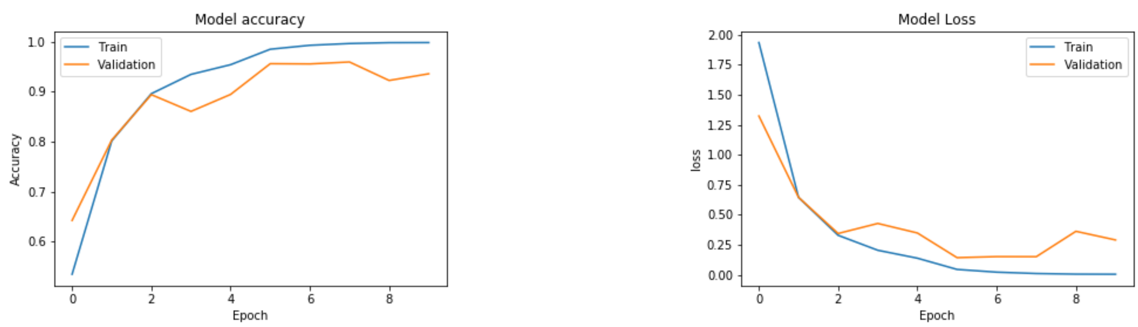
(a) Accuracy Graph (b) Loss Graph

Figure 4.3: Deep Convolutional Neural Network with Data Augmentation



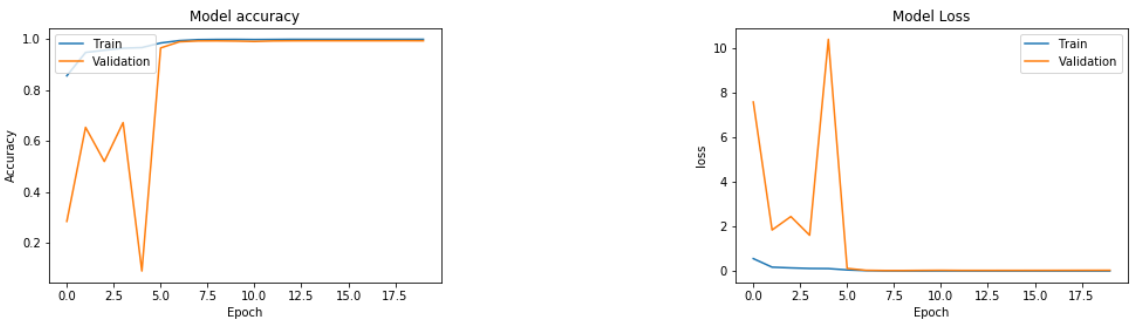
(a) Accuracy Graph (b) Loss Graph

Figure 4.4: Residual Network



(a) Accuracy Graph (b) Loss Graph

Figure 4.5: Recurrent Convolutional Neural Network



(a) Accuracy Graph

(b) Loss Graph

Figure 4.6: Google's MobileNet

4.3 Simulation

4.3.1 Web Application

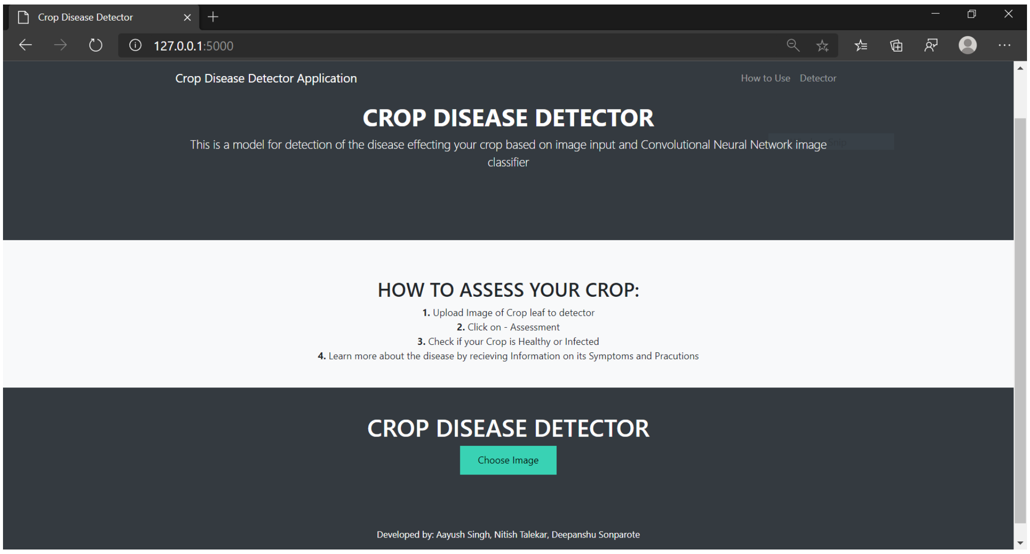


Figure 4.7: Web Application: Home Page (Steps)

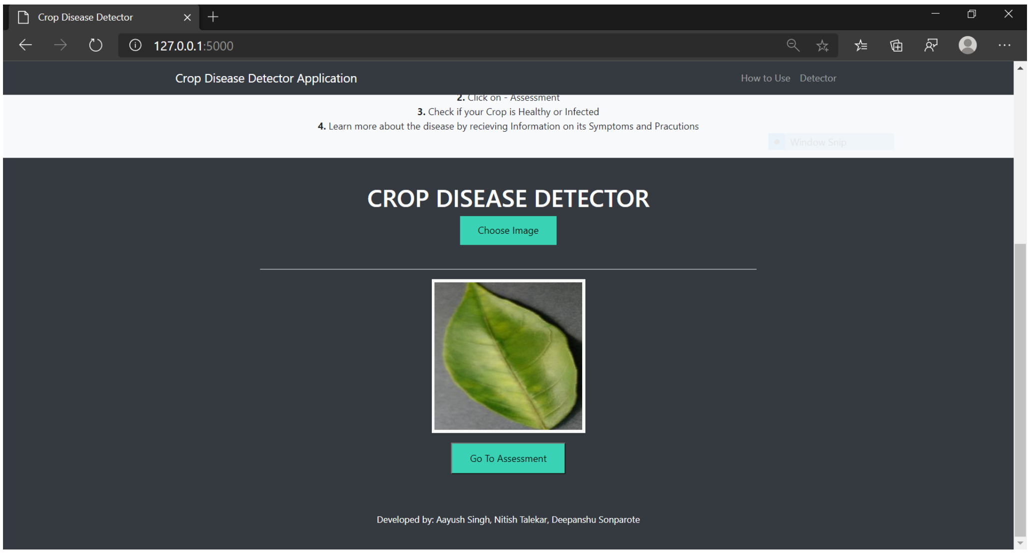
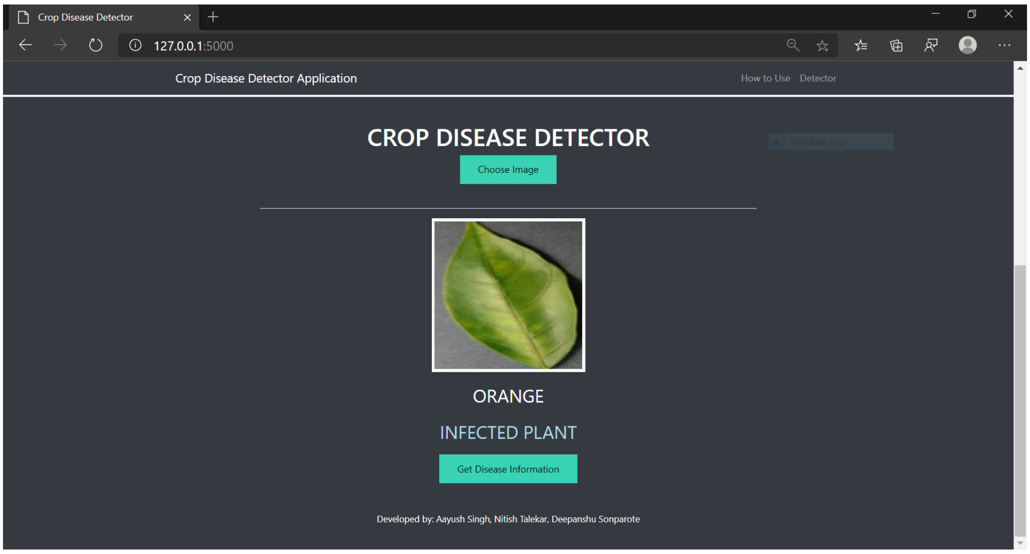


Figure 4.8: Web Application: Uploading Image



s

Figure 4.9: Web Application: Crop Recognition

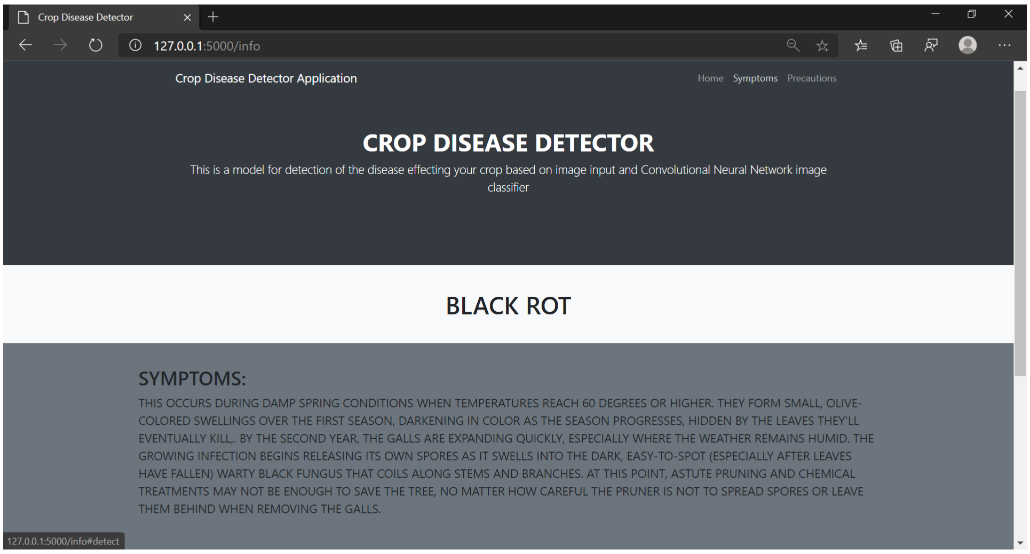


Figure 4.10: Web Application: Crop Disease Diagnosis

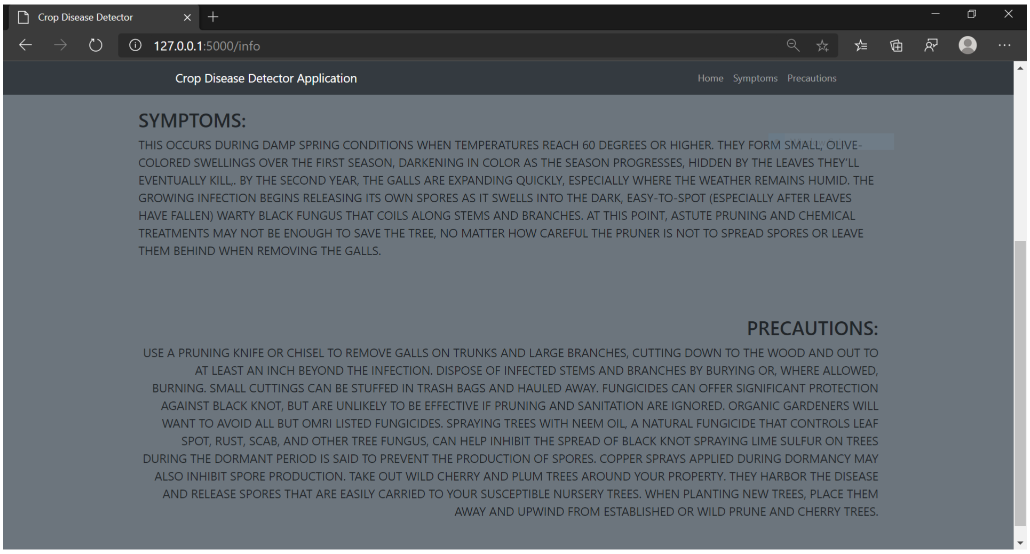


Figure 4.11: Web Application: Crop Disease Information

Figure 4.12: Home Page

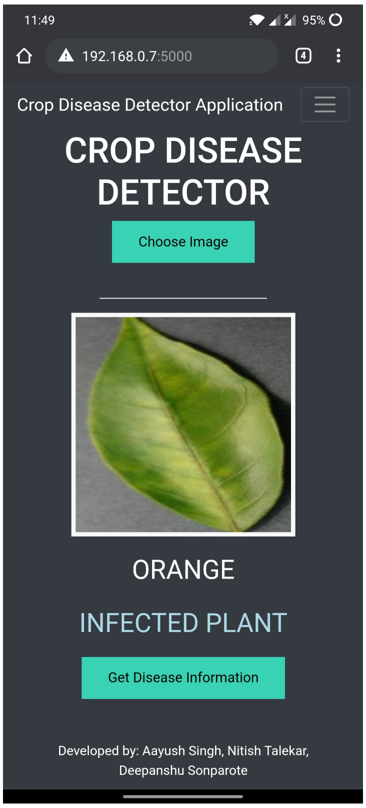


Figure 4.13: Upload Image

Figure 4.14: Get Results

4.3.2 Database

Chapter 5

Conclusion

This project is a crop disease detection system that uses neural networks and image processing techniques to recognize underlying diseases in crops based on crop leaf images. The model is trained over a Convolution Neural Network (CNN) on a dataset of over 54000 images labeled in 38 different classes. The maximum accuracy achieved from the trained model is 99%.

A web application is deployed with a complete system consisting of server-side components containing a trained model with features such as displaying recognized diseases in crops. A database of disease symptoms and precautions is maintained to fetch relevant data to display after disease is recognized.

This project is useful for farmers and agronomists in aiding them to identify underlying diseases in their crops so that immediate actions can be taken to prevent further losses.

Appendix

Detailed information, lengthy derivations, raw experimental observations etc. are to be presented in the separate appendices, which shall be numbered in Roman Capitals (e.g. "Appendix I"). Since reference can be drawn to published/unpublished literature in the appendices these should precede the "Literature Cited" section.

Bibliography

[1] Chopda, J., Raveshiya, H., Nakum, S. and Nakrani, V., Cotton Crop Disease Detection using Decision Tree Classifier, *International Conference on Smart City and Emerging Technology (ICSCET)*, IEEE, pp. 1-5, January 2018.

[2] Kamble, J.K., Plant Disease Detector, *International Conference on Advances in Communication and Computing Technology (ICACCT)*, pp. 97-101, February 2012.

[3] Kulkarni, O., Crop Disease Detection Using Deep Learning, *Proceedings of Journal of Global Research in Computer Science* , Volume 4, No.4, April 2013.

[4] Mique Jr, E.L. and Palaoag, T.D., Rice Pest and Disease Detection Using Convolutional Neural Network, *International Conference on Information Science and System*, ACM, pp. 147-151, April 2018.

[5] Padol, P.B. and Yadav, A.A., SVM classifier based grape leaf disease detection, *Conference on advances in signal processing (CASP)*, IEEE, pp. 175-179, June 2016.

[6] Petrellis, N., Mobile Application for Plant Disease Classification Based on Symptom Signatures, *21st Pan-Hellenic Conference on Informatics*, ACM, p. 1, September 2017.

[7] Zeng, W., Li, M., Zhang, J., Chen, L., Fang, S. and Wang, J., High-Order Residual Convolutional Neural Network for Robust Crop Disease Recognition,*2nd International Conference on Computer Science and Application Engineering*, ACM, p. 101, October 2018.

[8] Singh, V. and Misra, A.K., Detection of unhealthy region of plant leaves using image processing and genetic algorithm, *International Conference on Advances in Computer Engineering and Applications*, IEEE, pp. 1028-1032, March 2015.

[9] Sardogan, M., Tuncer, A. and Ozen, Y., September. Plant leaf disease detection and classification based on CNN with LVQ algorithm, *3rd International Conference on Computer Science and Engineering (UBMK)*, IEEE, pp. 382-385, September 2018.

[10] Kluepfel M. , Blake J. and Keinath A., Tomato Disease & Disorders, Dec 2018 [online] Available at: https://hgic.clemson.edu/factsheet/tomato-+diseases-disorders/ [Accessed 24 Jan. 2020].

[11] Common Diseases Of Apple Trees - Gardening Channel, 2020, [online] Available at: https://www.gardeningchannel.com/common-diseases-of-apple-trees [Accessed 24 Jan. 2020].

[12] Corn Leaf Diseases, Aug 2018 [online] Available at: https://www.krugerseed.com/en-us/agronomy-library/corn-leaf-diseases.html [Accessed 24 Jan. 2020].

[13] Doman, E., 7 Most Common Grapevine Diseases, Oct 2015[online] Available at: https://learn.winecoolerdirect.com/common-grapevine-diseases [Accessed 24 Jan. 2020].

[14] Scheufele, S., Potato, Identifying Diseases, Nov 2016[online] Available at: https://ag.umass.edu/vegetable/fact-sheets/potato-identifying-diseases [Accessed 24 Jan. 2020].

[15] Bloomfield, C., Common Pepper Plant Problems – Pepper Plant Diseases And Pests., Jan 2020 [online] Available at: https://www.gardeningknowhow.com/edible/vegetables/pepper/common-pepper-plant-problems.htm [Accessed 24 Jan. 2020].

[16] Cherry (Prunus Spp.)-Powdery Mildew., [online] Available at: https://pnwhandbooks.org/plantdisease/host-disease/cherry-prunus-spp-powdery-mildew [Accessed 24 Jan. 2020].

[17] Doubrava, N., Miller R., J., Blake, J., Peach Diseases., Nov 2019 [online] Available at: https://hgic.clemson.edu/factsheet/peach-diseases/ [Accessed 24 Jan. 2020].

[18] Ellis, M., Strawberry Leaf Diseases., Apr 2016 [online] Available at: https://ohioline.osu.edu/factsheet/plpath-fru-35 [Accessed 24 Jan. 2020].

[19] Mohanty, S.P., Hughes, D.P. and Salathé, M., Using deep learning for image-based plant disease detection, *Frontiers in plant science*, 7, p.1419, September 2016.

[20] Krizhevsky, A., Sutskever, I. and Hinton, G.E., Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[21] He, K., Zhang, X., Ren, S. and Sun, J., Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, IEEE, pp. 770-778, 2016.

[22] Liang, M. and Hu, X., Recurrent convolutional neural network for object recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, IEEE, pp. 3367-3375, 2015.

[23] He, K., Zhang, X., Ren, S. and Sun, J., Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*, IEEE, pp. 1026-1034, 2015.

[24] Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S., A survey of the recent architectures of deep convolutional neural networks. *arXiv preprint*, arXiv:1901.06032, 2019.

Publication by Students

Paper entitled "An Intelligent Traveling Companion For Visually Impaired Pedestrian" is presented at *International Conference on Circuits, Systems, Information and Communication Technology Applications (CSCITA - 2014)*, In collaboration with IEEE Bombay Section,Mumbai with the approval of IEEE, USA.

Annexure

=210mm =297mm

Cotton Crop Disease Detection

Dilip S. Kale
Asst. Professor
Department of Computer Engineering
Rajiv Gandhi Institute of Technology

Aayush Singh
Department of Computer Engineering
Rajiv Gandhi Institute of Technology
aayushipsingh@gmail.com

Deepanshu Sonparote
Department of Computer Engineering
Rajiv Gandhi Institute of Technology
deepanshusonparote@gmail.com

Nitish Talekar
Department of Computer Engineering
Rajiv Gandhi Institute of Technology
nitishtalekar.nt503@gmail.com

Abstract - **The crop agriculture industry faces the economic losses due to the pest infections, bacterial or viral contagions, the farmers lose nearly 18-20% of the total profit on an average annually in India. This paper proposes a solution to the agricultural problem, which involves crop disease recognition by using machine learning and image processing techniques. In this paper, the study sets out to classify cotton crop images into classes, whether the crop is infected by a disease or not. Also, we endeavour applications that give the farmer readily available means to identify the diseases on their crop and take appropriate damage control actions. The dataset used to train the model was user created (mobile capture images with high-resolution camera) from various crop farms. Cotton crops, of different varieties, containing three types of diseases, namely Rust, Mosaic Virus, Woolyaphids and Healthy plants are taken as classification ideals. The trained models have provided a performance reaching an 87.5% success rate in identifying the corresponding cotton plant disease. The model used in the study delivers significant accuracies of classification on the dataset used by employing Convolutional Neural Networks. The model is a very useful advisory or early warning tool for the farmers for identification of diseases in the early stage so that immediate action can be taken.**

Keywords – Crop Disease Detection, Cotton plant leaf, Convolutional Neural Network, Image processing.

I. INTRODUCTION

Agriculture is the primary occupation in India. Crops are part and parcel of the agricultural industry. Nowadays, a tremendous loss in the quality and quantity of crop yield is observed, subject to various diseases in the plant. Crop Plant disease classification is a critical step, which can be useful in early detection of pest, insects, disease control, increase in productivity, among other examples. Farmers recognize disease manually with foregoing symptoms of plants, and with experts, whereas the actual diseases are hard to distinguish with naked eye, and it is time-consuming to predict whether the crop is healthy or not. Cotton is one of the major agricultural crops in India and it has a dominant impact on the overall Indian agriculture sector. Cotton plant leaf disease diagnosis is very difficult through simple observation. Due to the complexity, even experienced agronomists and plant pathologists often fail to successfully diagnose specific diseases and are consequently led to mistaken conclusions and treatments.

Our study helps to predict crop diseases in cotton plants by processing the images of the crop. For this, Image Processing techniques are used for fast, accurate and appropriate classification of diseases. Symptoms of diseases in cotton predominantly come out on leaves of plants. The existence of an automated system for the detection and diagnosis of plant diseases would offer a support system to the agronomist who is performing such diagnosis through observation of the leaves of infected plants.

The existing techniques for disease detection have utilized various image processing methods followed by varied classification techniques. The proposed system uses a convolutional neural network to classify the health of a cotton leaf plant. The flow diagram of the proposed system given in Fig. 1. consists of steps used to acquire the desired output.

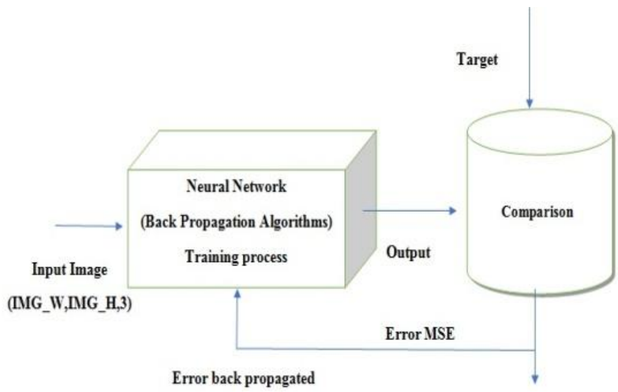


Fig 1. Flow Diagram of Proposed System

II. LITERATURE REVIEW

Crop Yield Forecasting has been an area of interest for producers, agricultural-related organizations. Timely and accurate crop yield forecasts are essential for crop production. The existing techniques for disease detection have utilized various image processing methods followed by various classification techniques. However, some unconventional approaches have led to classification of diseases using unconventional factors. Chopda et al. [1] proposes a system which can predict the cotton crop diseases using decision tree with the help of the parameters like temperature, soil moisture, etc. based on the previous year data and through sensors. However, these data might not be fully dependable to predict or classify diseases.

Image classification and regression techniques play a very important role because they allow identifying, grouping, and organizing from a standardized system. We apply an algorithm for image segmentation technique on data for automatic detection and classification of plant leaf diseases. In [2], Kamble defines the application of texture analysis for detecting plant diseases with the help of different image processing technique. Further with the use of Decision-Making Module, the disease is classified. Singh and Misra [8] suggest different disease classification techniques that can be used for plant leaf disease detection and an algorithm for image segmentation, the advantage of using this method is that the plant diseases can be identified at an early stage or the initial stage.

Deep learning is a set of learning methods attempting to model data with complex architectures combining multiple non-linear transformations. The element of deep learning is the neural networks that are combined to form the deep neural networks. These techniques have enabled significant progress in the fields of image processing and image classification. Kulkarni [3] formulates an application of Deep Convolutional Neural Network to identify and classify crop disease on images, testing it on five classes of crops and three types of diseases for each class. Mique Jr, Eusebio L [4] proposed an application that will help farmers in detecting rice insect pests and diseases using Convolutional Neural Network (CNN) and image processing. The searching and comparison of captured images to a stack of rice pest images was implemented using a model based on CNN. Collected images were pre-processed and were used in training the model.

There exist several types of architectures for neural networks:

1. The multilayer perceptions, that are the oldest and simplest ones,
2. The Convolutional Neural Networks (CNN), particularly adapted for image processing
3. The recurrent neural networks and dense neural network (DNN) used for sequential data such as text or times series.

A different approach is taken by Petrellis [6] where mobile phone application for plant disease diagnosis is presented which is based on the detection of the disease signature that is expressed as a number of rules that concern the colour, the shape of the spots and historical weather data among other factors.

III. Diseases in Cotton Plant

A. Rust

Rust is a disease usually found in older leaves, where a yellow, orange-ish lesions are found on the upper layer of the leaves. These lesions grow with time and with time burst to release spores, which often coalesce and give rise to irregular dark brown spots.

Fig 2. Leaf infected with Rust disease

B. Mosaic Virus

Mosaic Virus is a viral infection in plants which can be transmitted via multiple carriers and does not get cured. The inability to cure this has made it necessary to conserve and defend against the contraction of virus. The virus affects the plant in multiple ways at multiple stages of a plant's lifetime. It can vary among Yellow/White spots, wrinkled leaves, stunted growth and development of 'warty' areas.

Fig 3. Leaf infected with Mosaic Virus

C. Woolly Aphids

Woolly Aphids are small parasite-like insects which utilize hosts such as cotton plants to lay eggs or feeding purposes. These generally feed on seeds, foliage or sometimes even roots of plants. The aphids are generally blue or green in colour, but appear white or cloudy due to the waxy material surrounding their body. There are rare cases of severe attack by woolly aphids, therefore only a slight care of plants can ensure protection from these insects.

Fig 3. Woolly Aphids' attack on Stem of Plant

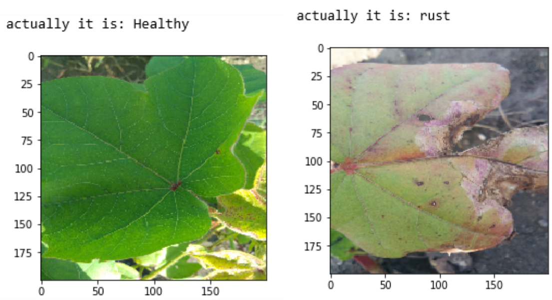
IV. METHODOLOGY

A. Dataset

The dataset contains a total of 8380 images. This is constituted by 2055 healthy plants images, 1800 having the Mosaic virus, 2155 plants being infected by Rust and 2370 samples of plants having Woolly aphids. Table 2 shows the composition of dataset of leaf images with class name and the number of plant leaf samples.

Image Class	Number of Samples
Healthy	2055
Mosaic virus	1800
Rust	2155
Woolly aphids	2370
Total	8380

Table 1. Amount of data samples per class



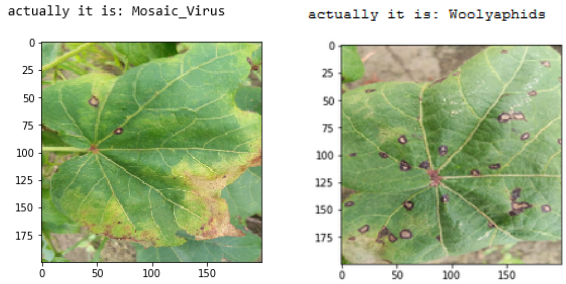


Fig 4. Dataset Examples and their labels

B. Pre-processing Data

For this initial process images of high resolution (4160 x 3120) are taken from datasets as input by setting IMG_W, IMG_H with 3 channels (RGB), for better visibility. The images are foremost pre-processed into a 4160 x 3120, RGB format with pixel values ranging from 0 to 255. Once normalized, the image is converted to a 224x224x3 input tensor and appended to the input array. The feature normalization used in the study is the min-max normalization. It is the ratio of the difference between the instance's feature value and the minimum value of a feature in the instance to the difference between the maximum and minimum values of features in the instance.

$$Norm(x) = \frac{x - xmin}{xmax - xmin}$$

C. Classification Model

The convolution network model used for classification utilizes the architecture of a VGG16 Network. VGG16 model comprises of 13 Layers of convolution with increasing number on kernels after certain intervals, and 3 Dense layers for classification. The VGG16 networks therefore has around 14M training parameters. To quicken up the training process and lower the time of execution, we have kept the number of kernels constant throughout the feature extraction process - 128. We have also added Batch Normalization layer after convolutions to drop unwanted nodes during the training process. These tweaks have proved to reduce the training parameter count to 9M and reduce training period significantly.

Fig 5a. Layers of standard VGG16 w.r.t project

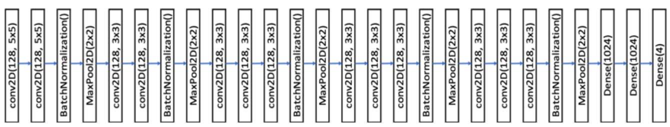


Fig 5b. Modified Layers of VGG16 for project

The activation function, ReLU is used for each of the hidden layers and the SoftMax function used in the last layer. Adaptive Moment optimizer is used for optimization. The dataset is divided into a train test split of 70%-30%.

V. Discussions and Results

The study was set out to classify if the crop is infected by a disease or is healthy. The training of model provided a significant quick learning process due to selected convolution layer structure and tweaks of layer addition.

Adagrad is used as an optimization function to converge to the optimized classes quicker. The number of true values predicted correctly are represented as T, several false values predicted correctly are represented as F. The error was calculated using the error function mentioned.

$$Error = \left(\frac{F}{F+T}\right) * 100$$

The following graphs are mapping of training accuracy and loss over 50 epochs of training.

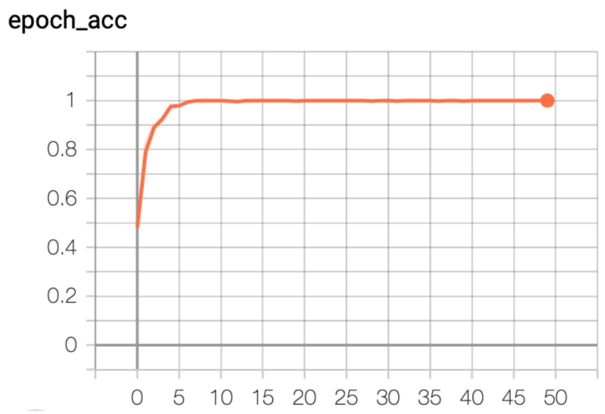


Fig 6a. Training Accuracy

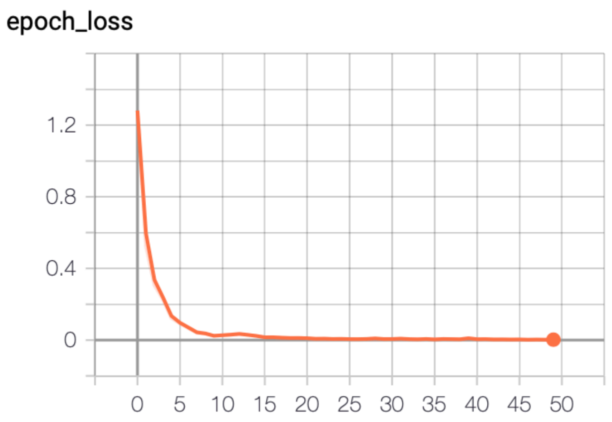


Fig 6b. Training Loss

The training accuracy chart shows a significant consistency of training accuracy after 5 epochs of training, thus continuing to reduce loss over next 50 epochs. Consequently, the training loss can be observed to diminish exponentially over epoch 1 to 10, further decreasing as epoch progressing. This denotes the finer understanding of the pattern and regions in images, by the Network.

VI. Conclusions and Future scope

In this paper, we have considered the cotton crop as it is the most important cash crop in India. Normally Rust, Mosaic virus and Woolly aphids are the hazardous diseases that the cotton crop suffers from in our country. Here, we consider a Convolutional Neural Network for crop disease recognition using leaf images for classification. There are several methods in computer vision for plant disease detection and classification process, but still, this research field is lacking. At present, there are no commercial solutions available in the market dealing with plant disease recognition based on the leaf images. We use a new approach of deep learning which automatically classifies and detect crop diseases from leaf images.

The data set of size 8380 images is collected from local farm with different diseases. The crop leaf images are considered for the training and testing purpose of the network. Initially, with the use of Gradient Descent and Back-Propagation algorithm classification are performed and it gives the prediction of diseases with 87.5% efficiency.

Furthermore, Convolutional Neural Network can be used for better classification accuracy. The main aim is to detect the crop leaf diseases from the database and train the images in such a way that the trained model gives the solution to farmers. The proposed model can recognize 11 different types of plant diseases. Here we consider plant stream and affected area by the disease boundaries, colour variation, size and shape of plant leaves.

The future work of this project is to develop a complete system consisting of server-side components containing a trained model and an application for smart mobile devices with features such as displaying recognized diseases in plants, based on leaf images captured by the mobile phone camera. This application will serve as an aid to farmers, enabling fast and efficient recognition of plant diseases and facilitating the decision-making process when it comes to the use of chemical pesticides.

References

[1] Chopda, J., Raveshiya, H., Nakum, S. and Nakrani, V., 2018, January. Cotton Crop Disease Detection using Decision Tree Classifier. In 2018 International Conference on Smart City and Emerging Technology (ICSCET) (pp. 1-5). IEEE.

[2] Kamble, J.K., 2018, February. Plant Disease Detector. In 2018 International Conference on Advances in Communication and Computing Technology (ICACCT) (pp. 97-101). IEEE.

[3] Kulkarni, O., 2018, August. Crop Disease Detection Using Deep Learning. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-4). IEEE.

[4] Mique Jr, E.L. and Palaoag, T.D., 2018, April. Rice Pest and Disease Detection Using Convolutional Neural Network. In Proceedings of the 2018 International

Conference on Information Science and System (pp. 147-151). ACM.
[5] Padol, P.B. and Yadav, A.A., 2016, June. SVM classifier based grape leaf disease detection. In 2016 Conference on advances in signal processing (CASP) (pp. 175-179). IEEE.
[6] Petrellis, N., 2017, September. Mobile Application for Plant Disease Classification Based on Symptom Signatures. In Proceedings of the 21st Pan-Hellenic Conference on Informatics (p. 1). ACM.
[7] Zeng, W., Li, M., Zhang, J., Chen, L., Fang, S. and Wang, J., 2018, October. High-Order Residual Convolutional Neural Network for Robust Crop Disease Recognition. In Proceedings of the 2nd International Conference on Computer Science and Application Engineering (p. 101). ACM.
[8] Singh, V. and Misra, A.K., 2015, March. Detection of unhealthy region of plant leaves using image processing and genetic algorithm. In 2015 International Conference on Advances in Computer Engineering and Applications (pp. 1028-1032). IEEE.
[9] Sardogan, M., Tuncer, A. and Ozen, Y., 2018, September. 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.

Acknowledgement

We wish to express our sincere gratitude to **Dr. Sanjay U. Bokade, Principal** and **Dr. Satish. Y. Ket , H.O.D.** of Department Computer Engineering of Rajiv Gandhi Institute of Technology for providing us an opportunity to do our project work on **"(Project Title) "**.

This project bears on imprint of many peoples. We sincerely thank our project guide **Mr. Dilip S. Kale** for his/her guidance and encouragement in carrying out this synopsis work.

Finally, we would like to thank our colleagues and friends who helped us in completing project work successfully

1. ABC(name of the student)
2. PQR(name of the student)
3. XYZ(name of the student)