Image Based Plant Disease Detection- A deep Learning Approach

Thesis Submitted By:

Sikandar Hayat & 01-243172-049

Supervised By:

Dr. Sumaira Kausar

Session (2017-2019)

Bahria University
Discovering Knowledge

MS-13

Thesis Completion Certificate

Scholar's Name: Sikandar Hayat Registration No. 29703

Programme of Study: MS (CS)

Thesis Title: Image based plant disease detection – A deep Learning Approach

It is to certify that the above student's thesis has been completed to my satisfaction and, to my belief, its standard is appropriate for submission for Evaluation. I have also conducted plagiarism test of this thesis using HEC prescribed software and found similarity index at 15% that is within the permissible limit set by the HEC for the MS/MPhil degree thesis.
I have also found the thesis in a format recognized by the BU for the MS/MPhil thesis.

Principal Supervisor's Signature: ____________

Date: ____________ **Name:** Dr. Sumaira Kausar

Bahria University
Discovering Knowledge

MS-14A

Author's Declaration

I, Sikandar Hayat hereby state that my MS thesis titled "Image based plant disease detection- A deep Learning Approach"

is my own work and has not been submitted previously by me for taking any degree from this university

(Name of University)

or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduate the university has the right to withdraw/cancel my PhD degree.

Name of scholar: Sikandar Hayat
Date: 25-Jun-2019

Bahria University
Discovering Knowledge

MS-14B

Plagiarism Undertaking

I, solemnly declare that research work presented in the thesis titled "Image based plant disease detection - Ad deep learning Approach" is solely my research work with no significant contribution from any other person. Small contribution / help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Bahria University towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred / cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the university reserves the right to withdraw / revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC / University website on which names of students are placed who submitted plagiarized thesis.

Student / Author's Sign:

Name of the Student: Sikander Hayat

BAHRIA UNIVERSITY

Islamabad Campus

(Department of Computer Science)

Student's project dues paid for final presentation

Signature of Account Officer

Thesis Submission Form

(To be filled by Supervisor)

Date:

1. Thesis Title: Image Based Plant disease detection- A deep learning Approach
2. Thesis Start Date: Thesis Ending Date:
3. Information of Students:

S. #	Students Name	Enrolment #	Email Address	Contact #
1	Sikander Hayat	01-243172-049	Sikanderhayat194@gmail.com	0314-5073093

4. Do you recommend this thesis for external evaluation: Yes ✓ No
5. If the answer to 4, is yes then complete the following:
   a. Have you *thoroughly* read the degree thesis report: Yes ✓ No
   b. Have you taken the demonstration of the thesis: Yes ✓ No
   e. Have you checked the *plagiarism* report: Yes ✓ No

Supervisor's Signatures

Abstract

Identification of plant disease is beneficial not to protect plants from diseases but also to increase the budgetary advancement of a country. The study of plant disease identification means the study of visually pattern seen on the plant. The traditional approach adopted for plant disease detection was naked eye observation of expert but it is very complex to find the plant disease manually because it requires high processing time and expertise in plant's diseases. So, it was necessary to develop a system that detect the plant disease in less time and cost effective manner. Hence, we have applied convolution neural networks (CNNs) architectures VGG19 and ResNet50 on both laboratory and natural images of plants for the identification of plant disease. We have trained and tested our CNN models on 10,066 images which contains the steps likes image capturing, image pre-processing, data augmentation, CNN training and testing. Furthermore, we have trained and tested our models on mixture of laboratory and natural images but the amount of natural images plus classes was greater than laboratory images and classes of plants which contain diseases. Results in that ResNet50 outperformed in detection of plant disease.

Acknowledgments

I would like to thanks my parents who supported me financially to complete my degree. Moreover, my special thanks to Department of plant pathology, PMIS University and National Agriculture and Research Center (NARC) for helping me to gather the dataset to conduct my research as well as Prof. Dr. Gwanggil Jeon who supported me in the understanding of Convolution Neural Networks. At long last, I am obliged to thank my supervisor, Dr. Sumaira Kausar, who came up with the idea of exploring the field of Plant disease detection and i am glad that she put confidence in me to undertake this research topic. Her support, guidance and advice throughout the thesis is highly appreciated.

SIKANDAR HAYAT
Bahria University Islamabad, Pakistan

June 2019

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Plants play an important role not only in the source of energy but also the primary source of current issues of global warming. The damages caused by the plant diseases in plant system significantly affect the economic growth of a country. An approximately 26% of Pakistan's economy depends on agriculture growth. Furthermore, plant diseases impact to the human's being disease and environmental damages. As these plant's diseases have been dispersed worldwide result in damaging the normal functioning of the plant as well as financial cost by minimizing the quantity of plant grown. More consequently, it also impacts on the quality of plant as well and these diseases are sometimes not visible with the naked eye.

Plants are also a source of food to more the 7.5 billion people living in this earth but it is threatened by several factors such as climate change [59] and plant disease [57]. Plant diseases are not the only risk to food security but also cause the financial growth of a poor farmer who livelihood only depends only on the cultivation of plants. In many countries, 70% of the plant production is produced by poor farmers [22] and approximately 50% of the plant production are damaged due to diseases calculated in [17]. In many places, worldwide farmers are still judging the type of disease by their own observation but this is not a proper way.

In plants, most of the diseases are initiated due to fungal and bacterial infections. The United Nations Food and Agriculture (FAO) have proposed in their research that agriculture production needs to be increased by 70% by 2050 to overcome the world's food needs [1]. Moreover rapid use of synthetic such as fungicide and bactericides to capture plant diseases has been causing effects in the agro-ecosystem. So, we need rapidly and effectively early disease detection technique to control plant diseases and to sustain the agro-ecosystem.

There are various types of plant leaf diseases which are bacterial, viral and fungal. Their symptoms are described below

1.1.1 Bacterial disease symptoms

This symptom is characterized by tiny pale green spots which soon come into view as water soaked. The lesions enlarge and then appear as dry dead spot as shown in figure 1.1

Figure 1.1: Coffee leaf with bacterial spot

1.1.2 Fungal disease symptoms

Plant disease which is caused by fungus e.g. powdery mildew. At the first stage, it looks like water soaked and spot of grey-green color which further affects other leaves and the older spot become darker. An example Figure is shown in 1.2

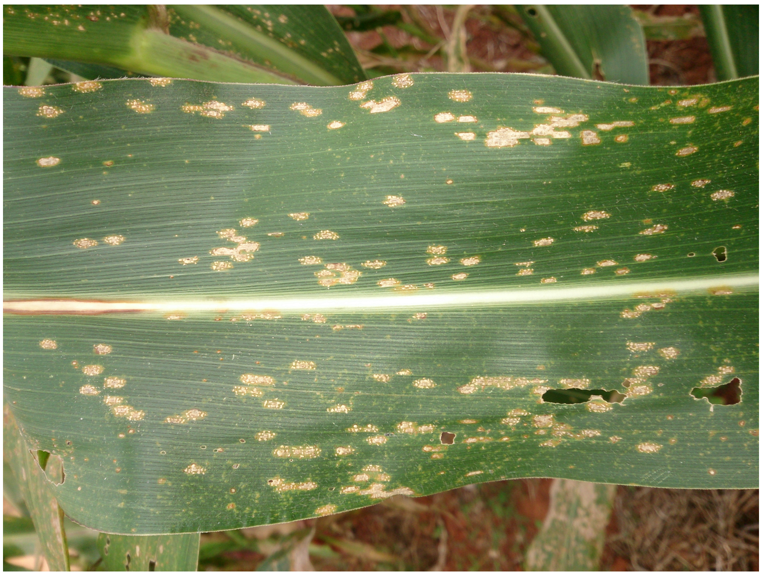

Figure 1.2: Corn leaf with fungal tropical rust disease

1.1.3 Viral disease symptoms

These symptoms are difficult to address. Viruses produce no sign to observe and detect and sometimes confused with nutrient deficiencies and herbicides injury. Aphides, leafhopper, whiteflies, and cucumber beetles are the cause of these type of diseases. An illustration is given in the Figure 1.3

Figure 1.3: Passion tree leaf with viral disease

The core approach adapted for the identification and detection of plant disease in many places is naked eyes observation of experts. This approach is also affected by several factors such as the physical condition of expert, eyesight and working condition such as climate. That is the reason that the above-mentioned approach is not proper to identify the disease of plant and it is also a time-consuming approach by monitoring the large covered area of land by an expert.

Many laboratory-based approaches have also been developed to investigate plant's disease such as polymerase chain reaction but these solutions to identify the disease of the plant is not cost effective and also time-consuming. So, in order to get rid of this threating problem of plant disease we need an effective system which not only identifies the disease of the plant in cost effective manner but gives the response of identification of plant disease rapidly.

The machine learning methods such as an Decision Trees, K-mean and support vector machine (SVM) have been applied in plant disease domain [49]. The traditional approach for image classification is on hand-engineered featured such as SIFT [32] and HoG [12] etc and then these features have been given to the learning algorithms for disease detection but sometimes the main disadvantage of these hand engineered features is that it may give accuracy less than the expected accuracy. However, a recent trend of using the convolution neural network has showed that learned representations are more effective and efficient. The core benefit of this approach is that algorithms automatically learn features and analyze images from the huge collection with minimum error [8].

In this research thesis, we are going to automate the system which overcomes the problems to detect plant disease by using different deep learning techniques.

1.2 Research Problem

The proposed research aims to develop a system that identifies the diseases of plants. While significant researches have been carried out to identify the plant disease, a major proportion of these efforts target identification of plant's leave disease by using laboratory setup images. We plan to investigate both on laboratory and real condition images for the detection of plant's disease.

1.3 Contribution in Research Thesis

The research contribution of our document is described below

- We detect the plant disease on both laboratory and natural images captured in real cultivation field while most of the work has been performed on laboratory images.
- The detection of plant's diseases has been carried out by using different models of convolution neural networks (CNN). The training has been performed on pre-trained networks of convolution neural network.
- The major contribution of the research was to collect the dataset of plants with diseases.
- Furthermore, other major contribution was to select the plant diseases on which the work has not been done before.

1.4 Organization of Research Thesis

The related work of our research proposal is written under section chapter 2. Chapter 3 represents the detail description data preparation and CNN models that we have used in our research thesis. Chapter 4 represents the experimental results and discussion. In the last chapter, we have concluded our entire document.

Chapter 2

Related Work

This chapter presents an past researches for plant disease detection. Moreover, Methods and techniques for design of plant's disease detection systems are also discussed.

2.1 Related Work

The existence of deep learning models has allowed the researchers to design, train and test the system from beginning to the end instead of using the traditional handcrafted approaches. Due to the tremendous performance of the convolution neural network as a feature extractor in image processing problems, this concept has also been extended to solve other problems like digit recognition, agriculture, and robotics. We have divided our literature review into four phases which are discussed below.

2.1.1 Existing Methodologies used for Plant disease detection

Many types of research on plant disease identification have been investigated in the past with innovative techniques and a survey on these techniques is presented in [35, 50]. Most widely used techniques were deoxyribose nucleic acid (DNA), serological and molecular. In [65] the author developed a system for citrus canker plant diseases with an overall accuracy of 87.99% and 86.87% for both image processing techniques and through experts. A. Camargo et al in [8] have used banana and plantain images for a different type of plant's diseases. At the first stage, they have converted the RGB images into H, l3a and l3b transformation and segment the image by analyzing the intensity distributed. After that, they have removed the unwanted pixel from the important part of the image for disease identification.

The authors in [20] proposed a system to identify the wheat disease. The high and low appropriate wavelengths of plant's disease were identified by using a RELIEF-F algorithm and they only considered three wheat plant disease i.e powdery, mildew and yellow rust with the classification accuracy rate of 86.5% for powdery, 85.2% for mildew and 91.6% for yellow rust disease. Zulkifli et al in [21] has proposed a plant's disease system to identify the health of chili plant. They have proposed that the chemicals for disease recovery can only apply to chili plant. They have used a software tool named MATLAB for feature extraction and image recognition. The preprocessing is performed by using Fourier filter, edge detection, and morphological methods. And at the last computer vision is used for object classification. In [53] Helmi et al have made the system to operate on oil palm leaves to detect the Ganoderma basal stem rot disease. They have used Airborne hyperspectral for extracting the data in continuous spectral bands for plant's disease detection.

Many other work has also been done for plant disease detection by applying computer vision approaches. The author in [9] has identified plant disease by using three color feature model i.e YcbCr, CIELB and HIS concluding that the disease spot is correctly identified without the disturbance of unwanted pixels. The author in [43] a system for sugarcane plant leave detection has been proposed in which they have used threshold segmentation for plant's leaf area and triangle thresholding for unwanted area concluding with the classification success rate of 98.60%. The author in [16] has detected the disease on citrus leaves. They have targeted three diseases of citrus plants which are Citrus ulcer, Overwatering, and Citrus greening. In this paper image enhancement techniques involves color space conversion and discrete cosine transform. At last, they have used GLCM to see feature extraction measurements like energy, contrast, and homogeneity.

2.1.2 Plant disease spot detection by applying thresholding

Different images processing method were also used to quantifying the severity of plant's disease. In [46] the author compared the image processing techniques with visual techniques to estimate the extremity of coffee leaf rust. They have captured images with two type camera the first one was black and white and the other was colored camera.The segmentation was performed by applying thresholding into the images. In light of the author argues, the image processing method performed better than visual techniques. They also said that the image processing techniques were better in case of the disease images which have more severe disease. They also proposed that the image processing based system were also suitable in discriminating the more severe image disease with less severe image disease. Tucker et al in [60] has developed the system for the identification of the diseases in sunflower and oat leaves. They have considered two disease blight and rust to detect. The first stage they have performed image segmentation by applying thresholding to extract the affected part of the plant leaf. After that, they have classified the disease by judging the characteristics of the lesion. The results were good but give some error due to the illumination while capturing the images.

In [55], the author estimates the damage caused by the spider mites on leaves. They have proposed the algorithm which was firstly separate the image from the background and then extracted the damage part from the healthy portion of the image.The final calculation of damage caused by the spider mites on leaves was estimated by dividing the number of pixel in contains the damage region with the total number of pixel in the leaf. The author has also compared his method with other techniques like leaf damage index and chlorophyll fluorescence. Results concluded that their method outperformed than the other two methods. A less similar approach was also adopted by Weizheng et al. In [62] to quantify the lesion in soybean plant leaves. They have used two methods of threshold to separate the leaf from the background after that the separated leaf region was converted into HCI color space. The Sobel operator was applied to find the edge of the infected portion. The second thresholding was applied to the image discover from the Sobel operator to remove the small objects from the binary image and fill the hole in the white portion of the object. This helped the author to extract the damage portion of the leaf.

Many other thresholding base techniques to identify the damage region of the plant leaves were also discussed. In [29], the author captured the images from an analog video camera in the presence of red light illumination to find the necrotic area. The images then stored in the computer. The method was performed from the plant leaves of tomato, sycamore and bracken fern. The necrotic region was extracted from simple thresholding. They have also proposed an algorithm that removes the healthy region pixel from the disease region pixel that was misclassified as a part of the damage region. Similarly in [34], calculates the damage caused by the frost in oat crops. The dataset was collected from real cultivation field. In the first step, the author converted the RGB images into L*a*b channels representation. They have applied three different thresholding technique to identify the severe portion of the plant which were otsu's, Isodata, and Fuzzy thresholding. Each thresholding emerges a unique value for each channel. They have repeated this process until some stopping criteria are met. The final unique values represent the number of class after applying labelling illuminates the damage caused by frost in oat crops.

Lloret et al. in [31] have developed a system to observe the health of the vineyard. The images were collected and monitored through simple webcam. Their main objective was to monitor the health and to quickly detect and quantify the diseased plant. Their system were consisted of five stages which are (distance between leaf and the camera) leaf size estimation, thresholding which they have used to separate the unhealthy leaves from disease leaves, morphological operation which they have used to remove the noise from the images, detection of disease leaf and at last the alarm was generated that your vineyard wants attention. In [43], a method to estimate the extremity of disease in sugar cane leaves. They have segmented out the images from two methods. The first method was to separate the leaf from the background they have used simple thresholding for the separation of leaf from the background. The second method they have used was to convert the RGB image into HIS color space. After that binarization method was used to separate the diseased portion of the leaf. They have used triangle thresholding for finding the threshold to binarize the image. Finally, the binary image is to estimate the fungi related diseases in sugar cane leaves.

Many Fuzzy based approaches have also been used in the past which detect the disease region of the plants. In [52], detect the symptoms of disease in cucumber and pumpkins leaves. They have captured the images from the scanner. Firstly the leaves were detached from the entire plant and then image taken from the scanner. Then the RGB images are converted into HSV model and the V component is discarded. They have used MATLAB functions for the implementation of their ideas. The image of the leaf was separated with the help of thresholding. Furthermore, they have applied C-mean fuzzy algorithm to make the two clusters first one was of disease and the second one was of healthy leaves. Results in that their algorithm performed better than using the third party packages.

2.1.3 Literture Review for Plant Disease by using Deep Learning Methods

In recent past years, the problem of plant disease identification and classification has also been done by using a neural network to enhance the classification significantly [30]. The authors in [47] identify the diseases occur in the tomato plant. They have considered 6 tomato plant diseases and identification has been done through two deep convolution neural network architecture AlexNet and VGG16 net but the major drawback of this paper is that they have used pre-trained classifier for tomato plant's disease with a classification success rate of 97.43% for AlexNet and 97.23% in case of VGG16 net.

The author in [38] uses two different architecture of convolution neural network for the detection of plant disease but the vital drawback of this system is that they have only used the laboratory-based images for plant's disease and haven't focused on real condition images. A similar approach has been introduced in Pawara et al in [44] compared the result of convolution neural network models with some traditional pattern recognition techniques for plant disease using three different databases of full plant, fruits, and plants leaves, result in, that the traditional techniques for the identification of plant disease is not suitable than CNN techniques. The author in [56] has also used the CNN model to detect the plant's disease with an amount of 13 diseases and 5 plant leaves.

In [39] has proposed a system for tomato plant leaf disease detection system.In this paper, noise removal, image smoothing, background removal and image resizing have applied as an image pre-processing techniques. The author has used a Gabor wavelet change technique for feature extraction and they have used SVM for the classification of tomato plant leaf disease. The author in [37] has detected the diseases on sugarcane leaf. In the pre-processing stage, firstly they have converted the color images into the grayscale images then they have applied image segmentation to separate the unwanted area from the infected one. At last, they have used Non-linear SVM, linear SVM and multiclass SVM for the identification of sugarcane plant disease.

Menukaewjinda et al [36] have applied BPNN for color extraction in grape plant disease and the detection have performed by using the combination of self-organizing feature map and Genetic algorithm which provide automatic adjustment of parameters to detect the grape leaf disease color extraction. Song kai et al [24] have recognized maize disease of corn leaf by using the BPNN. Image segmentation has been performed by using the YCbCr color space to separate the useless area from the infected region of corn leaf and Co-Occurrence and gray level layer has been used to extract the disease spot texture features. Song Kai et al have used the K-mean algorithm for clustering and BPNN for the detection of corn leaf disease.

Another approach adopted by Shoby Sunny et al [58] to detect the citrus canker leaf disease. They have used CLAHE as an image pre-processing technique.At the first stage, they have segmented out the region of interest using K-mean clustering and texture feature extraction using GLCM. After they have extracted the features then they have applied Support vector machine (SVM) for the detection of citrus canker disease. In [64], Shi Yun et. al.used feature of color,

Table 2.1: Literature papers of plant disease detection along with their Accuracies-1

Ref#	Author Name	Species,Disease	Database	CNN Architectures, Other Algorithms	Accuracy
[18]	Sharada P.Mohanty et al	14,26	PlantVillage (lab)	AlexNet, GoogleNet	99.36%
[19]	Pawara et al (plant type recognition)	10 types	AgrilPlant (n), leafSnap (lab), Folio (lab)	AlexNet, GoogleNet	98.33%, 89.62%
[20]	Srdjan Sladojevic et al	5,13	Custom (lab)	CNN	96.3%
[23]	Usama Mokhtar et al	1,2	Custom (200 images, lab)	SVM and GWC for FE	99.5%
[24]	Prajakta Mitkal et al	1,6	Custom	SVM	—
[25]	Meunkaewjinda, A et al	1,3	Custom (1472 images)	BPNN and GA for DD	86.03%
[26]	Song kai et al	1,3	Custom (lab)	BPNN	98%
[28]	Shoby Sunny et al	1,1	Custom	SVM (without CLAHE, with CLAHE)	95%, 94%
[29]	Shi Yun et. al et al	1,3	Custom (300 images, natural)	BPNN	90%

texture, and shape for feature extraction and they have applied probabilistic based neural network (PNN) for the detection of plant disease. Furthermore, a similar approach adopted by the author in [54] but the difference was that they have used the combination of color and shape feature. They have also developed new features mean and peak indices of the histogram of shape for the detection of plant disease. Total 163 shape and 260 color features were extracted based on different aspects like the probability of feature error, targeted class relevancy, and correlation. Results in that the combination of feature gives better accuracy then finding the accuracy from the individual feature.

The work in [41] has been done by identifying the disease on brinjal plants. In this paper, the study of interest was to detect the brinjal's plant disease by using only the leaf image of brinjal plant instead of choosing the entire plant of brinjal.The diseases of brinjal have been investigated by using used K-mean and neural network. In [63], the author used 32 kinds of plants for detection. They have used PCA for feature extraction. The PCA has extracted Twelve feature and they have been given to the probabilistic neural network (PNN) for identification of 32 kinds of plants. The accuracy achieved by PNN was greater than 90%.

The author in [48] has proposed system for the detection of lesion region of disease. They have used edge detection to find the lesion region of plant disease then they have applied HPCCDD algorithm for the classification. This paper proposed an RGB feature based methodology to segment the portion of the infected area of plant leaf and edge feature were extracted by using canny and Sobel filters. D. Oppenheim et al in [42] used potato plant leaf images for tuber disease detection. They have used five classes of tuber disease of the potato plant in which four consisted of tuber disease and one was from a healthy plant. The images that they have used were from different size and shape and labeled by an expert. They have applied the VGG deep convolution network for the detection of tuber disease on potato plants. Images were cropped and resized into the dimension of 224 X 224 and feed into the VGG convolution neural network for classification. The Table 2.2 is illustrating the summary of the past papers which had used deep learning for the detection of plant's disease.

In [15], proposed a system for which contains seven types of cucumber plant disease. They have used image processing techniques and four layer convolution neural network along with 4-fold cross validation methodology for the detection of cucumber plant disease results in that the average accuracy rate was 82.3%. A similar approach in [26] has used three layers convolution neural network for training purpose to detect the disease in cucumber leaf. The amount of disease were two in this paper. The author in [40] has used method color, FAST oriented and BRIEF rotation to extract features from the plant's leaves and has used K-nearest neighbors and linear support vector machine to identify four diseases in cassava plant. They have also made a smartphone application for real-time detection of cassava plant leaf which helps to protect the farmer's garden from cassava plant disease.

Many other deep learning approaches have been developed for plant disease detection by applying different models of a convolution neural network. In [61], the author has used deep convolution network VGG to identify 14 crops diseases by using an open source public dataset PlantVillage and they have achieved an overall accuracy of 90.4% by applying transfer learning

Table 2.2: Literature papers of plant disease detection along with their Accuracies-2

Ref#	Author Name	Species,Disease	Database	CNN Architectures, Other Algorithms	Accuracy
[17]	AK Rangarajan et al	1,7	PlantVillage (lab)	AlexNet, VGG16	97.43%, 97.23%
[31]	Smita Naikwadi et al	n/a,6	Custom	Neural Network	94%
[32]	Stephen Gang Wu et al	32 kind	Flavia (lab)	Probabilistic NN	90%
[33]	P. Revathi, et al	1,1	custom	HPCCDD algorithm	98.1%
[34]	D. Oppenheim et al	1,5	Custom (400 images)	VGG	95%
[35]	Erika Fujita et al	1,8	7,520 from Saitama Agr. Ins., JAPAN (n)	CNN	82.3%
[36]	Yusuke Kawasaki et al	1,3	800 from Saitama Agr. Ins., JAPAN	CNN	94.9%
[37]	Godliver Owomugisha et al	n/a,5	7,386 images (n)	Linear SVM	80%
[38]	Guan Wang et al	1,1	PlantVillage (lab)	VGG	90.4%
[39]	Ilke Cugu et al	57 kind Flavia (lab)	CNN	98.44%	
[40]	Alvaro Fuentes et all	n/a,9	Custom (n) with 10 diseases	see table 2.3	see table 2.3
[41]	Jihen Amara et al	1,3	(PlantVillage) (lab)	LeNet CNN	see table 2.4
[42]	Peng Jiang et al	1,5	26,377 (both lab and natural)	GoogleNet and Inception	78.80%

with top-1 accuracy of 96.3%. Another approach adopted to classify the plant leaves from 57 different plants species by using a convolution neural network. They have extracted features from handcrafted image processing techniques and SVM for classification [11].

Alvaro Fuentes et al [14] has proposed meta-architecture for the detection of nine tomato plant leaf diseases. They have used three different convolution neural network architectures as a detector of tomato plant disease thus known as meta-architecture and was based on the Faster R-CNN, R-FCN and SSD. For feature extraction, they have used the VGG net and ResNet. Image segmentation was also applied on complex images of tomato plant leaf so that better feature and accuracy can be accomplished by the author. The overall accuracy achieved by the author presented in table 2.3

Table 2.3: different combinations of CNN (Detector, Feature Extractor) along with their accuracies

Combination	Accuracy
Faster R-CNN with VGG-16	83%
ResNet-50 with VGG-16	75.37%
ResNetXt-50 with VGG-16	71.1%
SSD with ResNet-50	82.53%
R-FCN with ResNet-50	85.98%

The author in [4] has used LeNet CNN architecture model for the identification of banana plant disease. These model were trained and tested on 3700 disease leaf images. The performance matrix consisted of overall accuracy, loss, and F1 score. They have targeted two diseases of banana plant which is banana Sigatoka and banana speckle. Different results of training and testing sets on different splits are illustrated table 2.4

Table 2.4: Accuracies on different sizes of training and test sets

size of Train set	size of Test set	Accuracy
20%	80%	98.61%
40%	60%	98.61%
80%	20%	92.61%

Furthermore, they have used hyper-parameters for training and testing purpose. Peng Jiang et al [23] has worked on to detected five apple plant disease. They have used complex dataset which was based on both laboratory and natural images of plants. The dataset was constructed by applying different data augmentation and image annotation techniques. The overall size of the dataset after data augmentation and annotation techniques was 26,377. They have applied the improved convolution neural network with the combination of GoogleNet inception structure along with Rainbow concatenation and have achieved the overall accuracy of 78.80% under Caffe framework. Summary of the papers are shown in table 2.5

Other methods without using the deep convolution network are also discussed in the past. In [19] the author has used k-nearest neighbors' classifier to detect diseases from different plants. They have extracted texture features from the disease image of various plants. They achieved an

Table 2.5: Literature papers of plant disease detection along with their Accuracies-3

Ref#	Author Name	Species,Disease	Database	CNN Architectures, Other Algorithms	Accuracy
[43]	Eftekhar Hossain et al	n/a,5	Arkansas and Reddit-plant (237 images)	KNN	96.7%
[44]	Konstantinos P. Ferentinos	25,58	PlantVillage (62.7%)	VGG, Overfeat, AlexNet, GoogleNet, AlexNetOWT	99.5%
[56]	H. Al-Hiary et al	n/a,5	Custom (lab)	NN and CCM	94%
[57]	Saraansh Baranwal et al	1,4	PlantVillage	CNN	98.54%
[58]	Anakha Krishnakumar et al	n/a,6	Custom (n)	SVM	86%
[59]	Juncheng Ma et al	1,4	Custom (n)	CNN	96.3%
[60]	Ardi Hidayat et al	1,3	PlantVillage (3,854)	CNN	99%
[61]	Ronnel R. Atole	1,3	Custom (600 images, natural)	AlexNet	91.23%
[62]	Ahmad Arib Alfarisy et al	n/a,13	Custom (4,511 images, natural)	CaffeNet	87%
[64]	Artzai Picona et al	1,3	Custom (8,178)	RNN	87%
[64]	Solemane Coulibaly et al	1,1	Custom (711 images)	VGG	95%

accuracy of 96.7%. The main disadvantage of this paper was they have used laboratory-based images instead of using the natural images.

Konstantinos P. Ferentinos in [13] has used deep learning architecture VGG and AlexNet for the detection of plant disease. The author has used 87,484 total images out of which 37.3% images were gathered from real cultivation field and 62.7% images were collected from the public dataset PlantVillage. This 62.7% of images were laboratory-based images. He has used two mechanisms for training and testing the dataset. The first mechanism was to train the convolution neural network on laboratory base images and test on images captured from the real environment. The second method was to train the network on natural images and test on laboratory images but the major drawback of this paper was that they have not trained and tested the CNN on images captured from the real environment. They concluded with a result that the overall accuracy of the system decreases if trained on laboratory images and test on natural images. They achieved an accuracy of 99.5%.

In [2], the author has developed a system to identify the diseases of agriculture plant leaves. They have detected five different diseases of plants. Due to the time consuming and expensive method of visually detect the plant disease the author uses the machine learning method for instant investigation of diseases of plant leaves. In the first step, the author has applied different images processing tasks such as image pre-processing, image segmentation and feature extraction. In the second step the author has used a k-mean clustering algorithm to detect the infected portion of the plant, features have been extracted by applying the color co-occurrence method and classification has been performed by using the artificial neural network (ANN). They achieved an overall accuracy of 94.67%. Moreover, in [6], Saraansh Baranwal et al has detected the apple leaves disease by applying convolution neural network (CNN). They have collected the images of apple leaves from an online public dataset PlantVillage which contain more than 50,000 laboratory images on different plant's species. They have only collected the images of apple leaves with a total amount of 2561 label images. Image compression, image generation, and image filtering techniques were used to increase the size and quality of the dataset. They have used pre-trained convolution neural networks and achieve an overall accuracy of 98.54%.

Anakha Krishnakumar et al [27] have estimated the plant disease severity by calculating the affected area divided by the total area of the plant's leaf. They have also classified the cucumber malady spot disease. The images were collected in the real condition field. They have used support vector machine to classify the malady spot on the cucumber plant. The overall accuracy achieved by the author was 86%. MATLAB was used for the implementation of this work. A similar approach in [33] has also been developed by Juncheng Ma et al but the difference was that they have used four cucumber leaf diseases. The images were captured in the real environment field. They have applied data augmentation techniques to increase the size of the data. The total amount of data after data augmentation was 14,208 images. They have used deep convolution neural network for detection of plant disease. The paper's limitation was that they have applied data segmentation which separates the leaf's disease spot from the entire image. The author has also compared his model's result with other models like random forest and support vector machine concluded with remarks that their deep convolution network was best suitable to recognize the cucumber leaf disease.

Ardi Hidayat et al in [18] have used deep CNN for the detection of corn plant leaf disease. The diseases were Common Rust, Gray Leaf Spot, and Northern Leaf Blight. They have collected the images from public dataset namely PlantVillage. The total amount of images were 3,854. They achieved an accuracy of 99%. The author in [5] has used AlexNet for the detection of rice plant leaf disease. They have made three-class classifier of normal healthy and snail infected plant via transfer learning. The overall accuracy achieved by the author was 91.23% with a learning rate of 0.0001. The total amount of images were 600. Ahmad Arib Alfarisy et al in [3] have classified 13 classes of paddy pests and paddy disease. They have used CaffeNet using Caffe framework for the identification of plant disease. Some of the diseases were from the English language some from Indonesian and some from Japanese. They achieved an overall accuracy of 87%. The total amount of images were 4,511. The author in [28] has made a deep learning IoT based system that monitors the water level of different plants. They have considered only 10-20 classes to adjust the soil moisture level.

Artzai Picona et al in [45] have used deep residual NN for the detection of wheat plant diseases Septoria, Tan Spot and Rust. Images were collected in Spain and Germany in the year 2014 to 2016 and the size of the images was 8178. The overall accuracy achieved by the author was 87%. The author in [10] has used VGG-16 for the detection of the millet crop of mildew disease. They have used transfer learning for the identification of millet crop disease. The images were collected from different web resources. Initially, they have downloaded 124 images of mildew disease. Then they have applied data augmentation techniques which enhance their dataset up to 711 images. They achieved an overall accuracy of 95%.

Literature also proposed several survey papers on plant disease detection which includes the comparison of different plant disease detection system. In [25], the author has written a survey paper on 40 different research paper of plant disease identification which has adopted the deep learning techniques applied to various agriculture and food-related production problem. They have written the agriculture problems under study, the specific models, and the overall accuracy achieved by each model in each research paper. Furthermore, they have compared the study to detect plant disease with other traditional approaches in term of classification and feature extraction. Concluded that deep learning approaches are more effective and suitable for the detection of plant disease. A similar approach was also adopted by Jayme Garcia et al in their survey paper [7] but they have only included the research articles which explore visible symptoms and stem in leaves. In [51], the author conduct a survey which discussed various techniques of image pre-processing like background removel, object detection, image disease portion extraction, image enhancement, noise removal and this paper has also discussed some methodologies for feature extraction and disease classification in various research papers for the identification of plant disease. Andreas Kamilaris et al in [25] have written a survey paper from 40 research items. They have examined the particular agriculture problem, the dataset used, pre-processing, data augmentation techniques and a tool they have used for implementation and overall accuracies achieved by models. They have also compared the deep models used for plant disease detection with other existing popular techniques. Concluded with remarks that the deep networks are more convenient than any other technique for the detection of plant disease.

2.2 Summary of the Chapter

In this Chapter of research thesis, we have discussed the past work that has used deep learning and other methods for the detection of plant diseases. The majority of the literature review is to identify the plant disease using deep learning methods. We have read different research papers which detects the plant disease by using different deep learning models.The major step involves for the detection of plant disease were image pre-processing, image labeling, data augmentation, image segmentation and classifiers for the classification of crop disease. Some paper used the CNN model for both as a feature extractor and at the same time classifier as well. Some papers have used the CNN model only as a classifier. We have also learned from the literature that CNN models can be trained through the scratch and fine-tuning. In majority papers, the pre-trained network gives better accuracies than the network trained from scratch the reason was that the pre-trained networks are trained on thousands of classes. Majority of the paper have used PlantVillage dataset which contains more than 50,000 laboratory-based images. Some work has been done by using their own dataset but the number of plant diseases was less. Many papers have only extracted the disease spot at the leaf by applying different thresholding methods from the entire plant. Some papers have measured the severity of plant disease by extracting the spot of plant disease from entire plant and they have calculated the severity of plant disease by dividing the area of plant leaf disease from the healthy area of the plant leaf. We have also inserted some of the survey paper on plant disease. These papers were based on the comparison of different research work by employ the metric which involves the overall accuracy, the nature of the images. Very few papers have used mixture of natural and labortory images to train and test the deep learning model such as in [23]. So, we will also train and test our deep learning models with the mixture of both type of images to achieve better accuracy then the research conducted [23]. Concluding remarks that the deep learning methods are best suitable for the complex problem and the detection of plant disease.

Chapter 3

Materials and Methods

This chapter introduces an overview of the proposed system. We then describe in detail the steps involves in our proposed study.

3.1 Materials and Methods

We start developing the system by choosing images of plants with different diseases to their disease detection. The system overview diagram of the entire research is shown in figure 3.1 where the data is annotated and converted into label data.

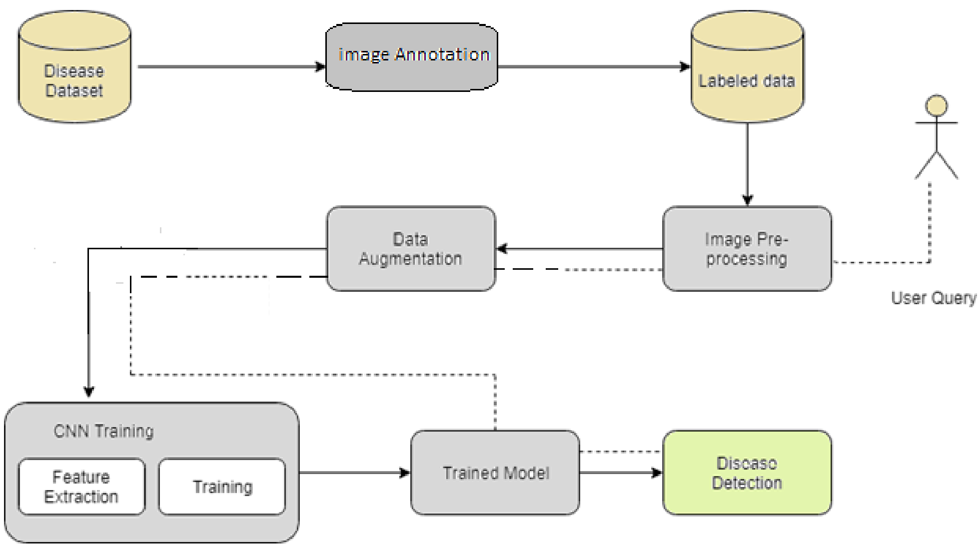

Figure 3.1: Proposed research system overview

After that, the label data is pre-processed and augmented. After Augmentation, the augmented data is given to the convolution neural network for feature extraction and training. Once the CNN model has been trained we test our CNN model by providing the user query which will be the image of the plant contains the disease. The dotted line shows that the user input will firstly be pre-processed to convert the image into the proper input size of the CNN model then the input image is given to the trained model. After that, the trained model predict the disease of the user's input plant disease image.

3.2 Design Stages of the proposed Research

The overall procedure of developing the entire system using deep learning is described in different stages as shown in figure 3.2 which involves image collection, image pre-processing, data augmentation, CNN training and CNN testing.

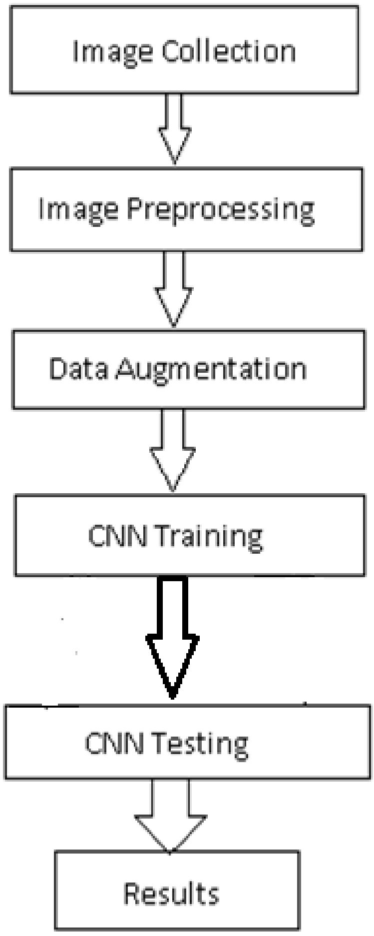

Figure 3.2: Design stages of the proposed research

3.2.1 Dataset

We have used the dataset collected from different sources for plant disease detection research, started from the training stage up until the validation stage to observe the performance of the learning algorithm. We have evaluated our learning algorithm on thirty three different classes of nine plant species with real condition and laboratory images described in table 3.1. The amount of Natural images classes is 27 and the laboratory images classes is 6. Furthermore, we have not added the class of healthy plants because our focus is to find the disease of the plant and healthy plants can be examined easily with naked eye without any cost.

3.2.2 Image Pre-processing

The images that we have collected from different resources are in different resolution. In order to get good results, we need to preprocess the images so that our deep learning algorithm gives good accuracy. In the image pre-processing, we have resized our dataset images to 224x224 so that we can give to our CNN model.

3.3 Data Augmentation

The purpose of applying the data augmentation is to enhance the dataset and to reduce the over-fitting (to reduce the training error). We have used keras (a popular python library) for images augmentation. We have used ImageDataGenerator class of keras library. Initially, we have 1500 natural plant and laboratory images. we split the dataset into train and test sets. we have augmented the train set. After Augmentation our dataset is converted into 10066 images.

Some of the examples of augmented images are shown in figures 3.3, 3.4 , 3.5 , 3.6 and 3.7

Figure 3.3: Augmented Dataset Examples

3.4 Convolution Neural Network (CNN)

CNN is a newer version of Artificial NN that helps in the process of image recognition and that is designed to processed the pixel data. CNN is the most powerful image processing technique based on deep learning to perform descriptive and generative tasks. Traditional Artificial neural network was not suitable for image recognition tasks because it takes images with low-quality resolution.

Table 3.1: Plants along with disease classes

Plant Name	**Disease**
Coffee	Cercospora Leaf Spot (N), Ferrugem(Rust) (N), Bacterial Blight (N), Brown leaf spot (L),
Cashew Tree	Anthracnose (N), Black Mould (N), Gummosis (N),
Sugarcane	Ferrugem(Rust) (N), Ring Spot (N), Red Stripe (N),
Coconut Tree	Aceria guerreronis (L), Lixa Grande(L), Lixa grande and lixa pequena (N), Lixa pequena (N)
Dry Bean	Anthracnose (N), Cowpea Mild Mottle Virus (N), Common bacterial blight (N), Phytotoxicity (N)
Dry Bean	Hedylepta indicata (N), Target leaf spot (L), Web blight, Powdery mildew (N)
Papaya	Smallpox (N)
Passion Tree	Scab (N), Dione juno juno (L), Bacterial spot (N), Mosaic virus (N)
Passion Tree	Senescence (N), Septoria spot (N), Woodiness virus (N)
Corn	Tropical rust (N)
Black Pepper	Phytophthora Foot Rot (N), Bacterial canker (L)

Figure 3.4: Rotation at 40 of corn plant with Diplodia leaf streak disease

This problem overcomes by CNN arranging the neuron just like the frontal lobe “Area responsible for visual stimuli for human and animal”.

Convolution neural network is a multilayer architecture designed to reduce the processing requirement. CNN consists of many layers which are fully connected layer, Convolution layer, pooling layer, ReLu layer, and normalization layers. A basic CNN architecture is shown in figure 3.8

3.4.1 Convolution Layer

A convolution layer of CNN is a building block of CNN. It contains the majority of the CNN processing load. This layer performs a dot product between the kernel also known as filter and the restricted portion of the receptive field. Conventionally, the size of the kernel is small then the image but is in more depth. This means that if the image is in RGB channel, the filter size will be small in both width and height but the depth is extended up to all the three channel.

During the forward pass, the filter slides across the receptive portion of the image and produce a new image representation of the receptive filed also known as feature maps. The forward pass size of the filter is known as a stride. An exemplary diagram of the forward pass of filter on the receptive region of an image is shown in figure 3.9

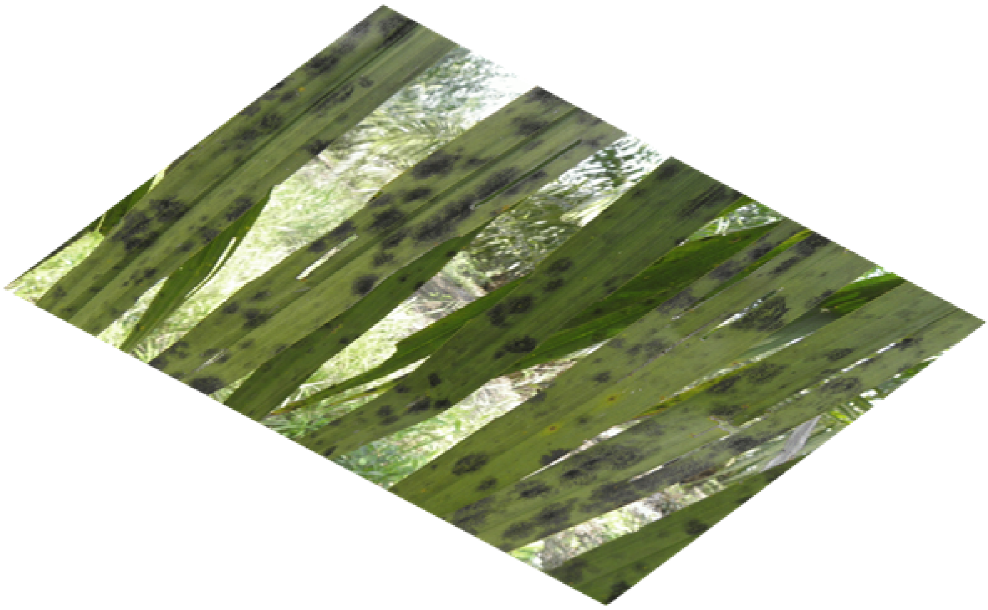

Figure 3.5: rotation of 40 degree of oil palm plant with Black sooty mold disease

3.4.2 Pooling Layer

The purpose of pooling layer is to decrease the parameters that help in reducing the computational power of the system and also helps in avoiding the overfitting. It applies to every slice of representation separately. There are many functions of pooling which is average of rectangle region, weighted average but the most popular is the max pooling which gives the maximum output from the neighborhood. An example of pooling is given in figure 3.10

3.4.3 Fully connected Layer

In this layer, every neuron is connected to all the neurons in the preceding and successive layer as you can see in figure 3.8. It helps to map the representation between the input and output.

3.4.4 Softmax Layer

It is implemented just before the output layer of CNN. Its neurons are equal to the neurons of the output layer. It is created when we have multiple class problem. It assigns decimal probabilities to each class. Those decimal probabilities are between 0 and 1. This helps to converge the training more quickly.

3.5 CNN Models

Are mathematical model that mimics the behavior of human brain function with their neurons interconnected with each other. The main characteristics of CNN are to train the network through supervised learning. In this process, CNN is trained to model some system by using existing data.

Figure 3.6: Horizontal flip of passion tree with Scab disease

CNN are the replacement of traditional artificial neural network their main advantage is that they reduced the number of structuring elements as compared to ANNs. We have used two different architecture of Convolution neural network which are

- VGG19
- ResNet50

These models along with their training and testing has been implemented by writing the python script on Google Colab.

3.5.1 VGG19 CNN Model

Visual Geometry Group (VGG) is developed by Oxford. It takes advantage over AlexNet by replacing large filter size 11 in the first convolution layer and 5 in the second convolution layer with 3x3 multiple filters one after another in increasing depth. The volume size in VGG is handled by a max pooling layer. With a given receptive filed (an input image on which output depends), It consists of multiple small kernel size which takes an advantage of extracting more complex features from the receptive field in a lower cost than the larger kernel size and in this way increase the depth of the network. An architectural diagram is shown in figure 3.11

it consists of 19 convolution layers and it is very attractive because of its uniform architecture. The width of VGG starts with value 64 and increase by the factor of 2 after every max pooling and decreased the size of the receptive filed by factor of 2. In the last layer, it consists of two fully connected plus ReLu layers with 4096 neuron along with softmax classifier. It has 138 million parameter which can be difficult to handle. It is trained on ImageNet dataset which contains images of more then 1000 classes.

Figure 3.7: Horizontal flip of dry bean plant with Phytotoxicity disease

3.5.2 ResNet50 CNN Architecture

ResNet50 is a type of convolution neural network which is trained on more 25.6 million images from the ImageNet database. It is 50 layers deep architecture and can classify 1000 of objects like pencil, animal, mouse, keyboard, etc. As a result, the network has the capability to learn advance rich feature from a different variety of images. Its input size is 224x224.

It is similar to GoogleNet but it uses average pooling followed by the classification layer. It achieves better accuracy then GoogleNet and VGG and computationally more efficient then VGG. It achieves a top-5 accuracy of 93%. ResNet50 architecture is more similar then the VGG network consists mostly of 3x3 kernel filters. An architectural diagram of ResNet50 is shown in figure 3.12

3.6 CNN Training

The preparation of training the CNN models has been described in the dataset, image pre-processing and data augmentation steps. We have used the pre-trained networks VGG16 and ResNet50 for our CNN training or fine-tuned the pre-trained network. The VGG16 and ResNet50 models both are pre-trained on ImageNet dataset or it takes weights of ImageNet dataset for their training. We have frozen all the layers of VGG10 except the last four layer and add dense layer with softmax activation function at the end of the model. In case of ResNet50, we have froze all the layer except the last five layer and add dense layer with softmax activation function at the end of the model. We have used the average pooling layer for the dimensionality reduction in both the CNN architectures. The classification function in both the VGG16 and ResNet50 is the softmax classifier which computes the probability of 1,000 classes of the ImageNet dataset. We have identified that fine-tuning or using the pre-trained method for training the model was efficient than the training from scratch because it requires less time to change the existing implemented parameters as compared to the

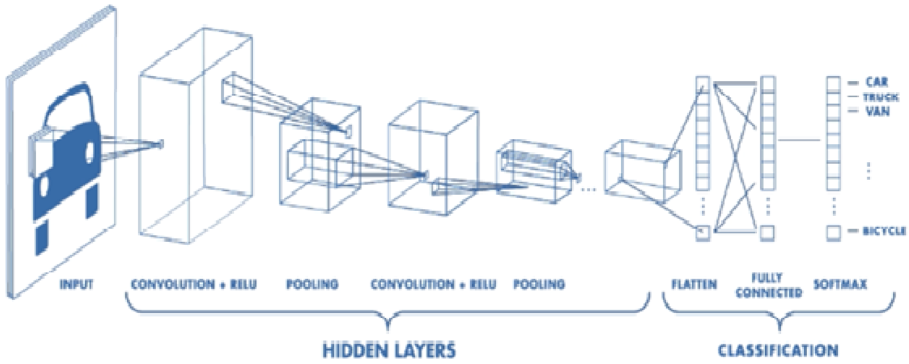

Figure 3.8: A basic CNN architecture

scratch. The final Convolution neural network training parameters are given in table 3.2 which includes

- Initial Number of Epocs
- Learning Rate
- Batch size Training
- Batch size Validation
- Early Stop Patience
- Steps PER EPOCH Training
- Steps PER EPOCH Validation

Table 3.2: parameters along with their values for CNN model training

Parameter Name	Value
Initial Number of Epocs	150
Learning Rate	0.004
Early Stop Patience	3
Batch Size Training	100
Batch Size Validation	50
Early Stop Patience	3
Momemtum	0.9

3.7 Summary of the Chapter

In this chapter, we have discussed the methods to design our proposed research which includes data-set information which described the number of classes (plant, disease) to detect the plant

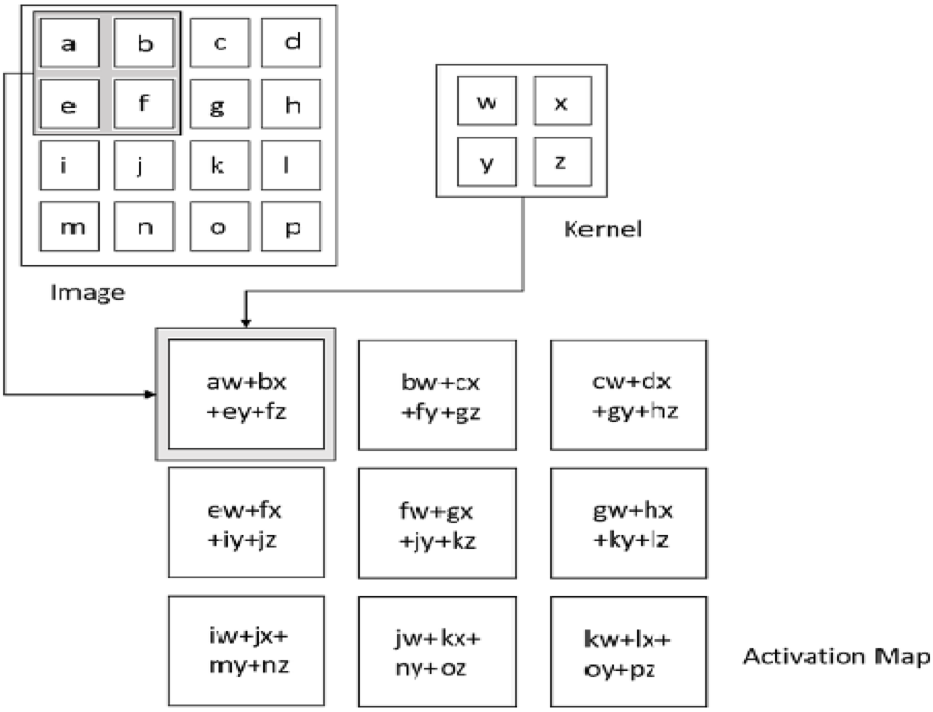

Figure 3.9: A example of forward pass of filter on the receptive region of an image

disease. We have mentioned in that chapter that we will detect 33 plant disease from nine plant species, image pre-processing to re-size the image into the proper input format of the model, data augmentation to enhance the number of images for better model accuracy and CNN models (VGG16 and ResNet50) that we have trained and validated for the detection of plant disease. This chapter has also presented a detailed about the architecture we have trained for plant disease detection.

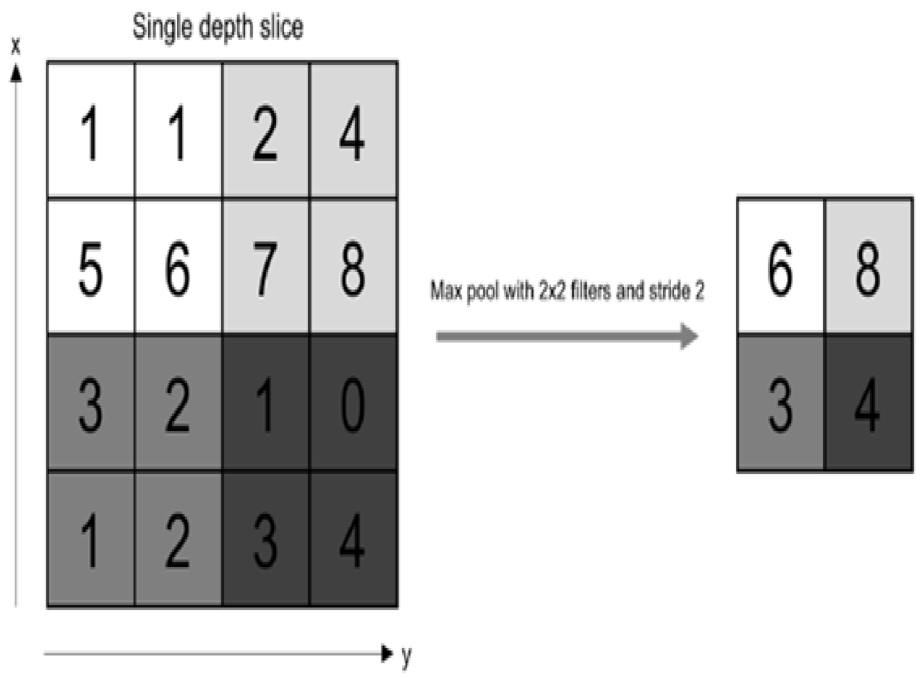

Figure 3.10: Max Pooling operation with filter size 2x2

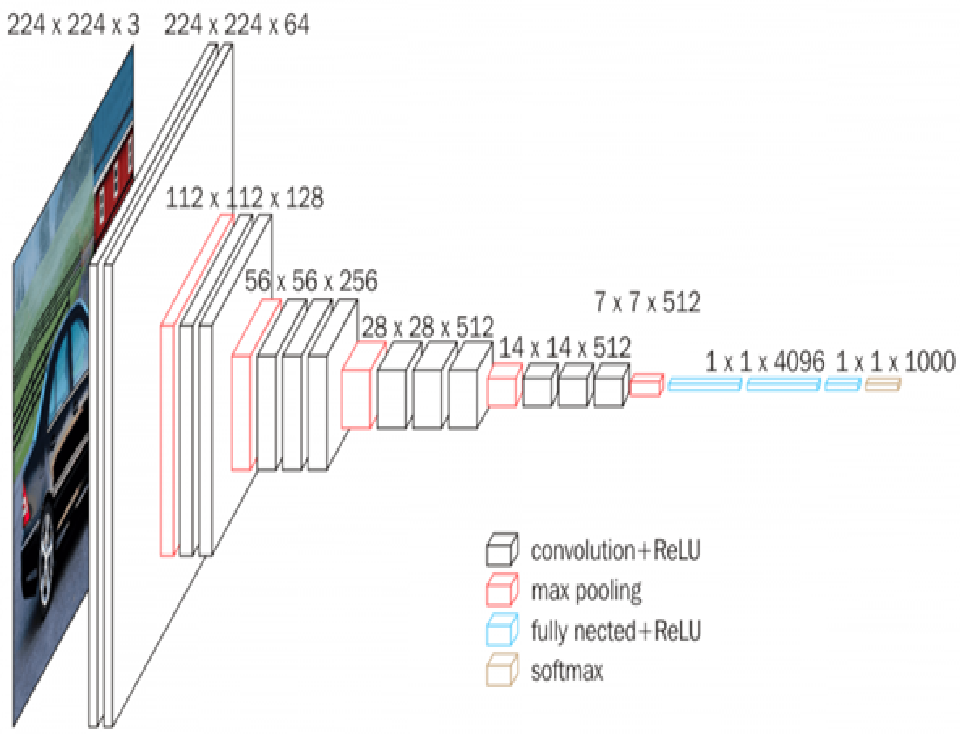

Figure 3.11: VGG19 Architectural diagram

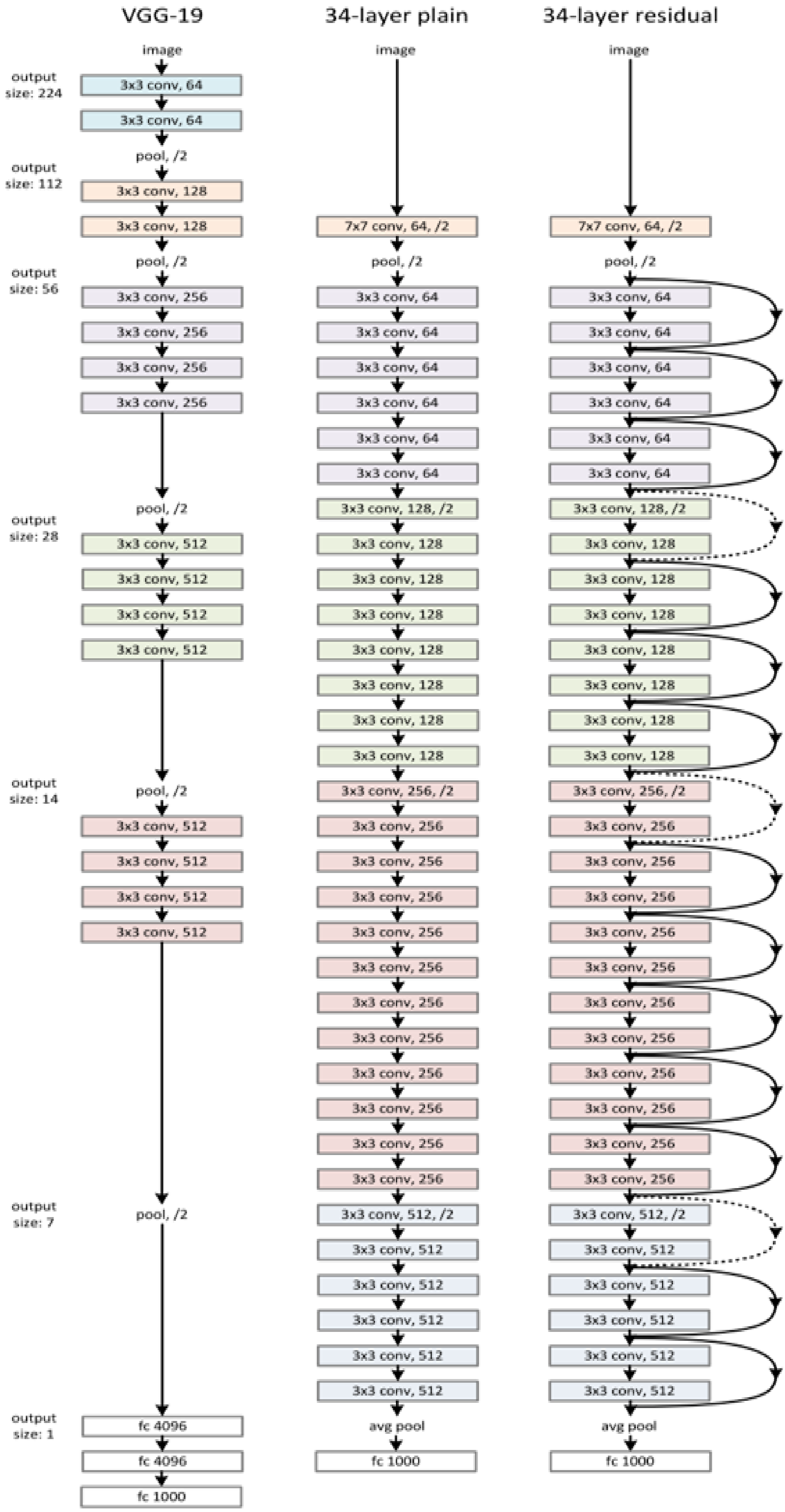

Figure 3.12: ResNet50 Architectural diagram

Chapter 4

Results and Discussion

4.1 Results and Discussion

The deep learning models presented in chapter 3 were trained by using the parameter defined in table 3.2. Comparison of these model was based on their testing set. We have extracted 500 images for testing out of 1500 natural and laboratory images from our entire dataset. The CNN models have been trained and validated on 10066. Initially, we set the number of epochs to 150 so that our model can best train and validate. We have also used early stopping with patience 3 in our model's training and validation. The purpose of early stopping is to stop the training if the training and validation loss of model not increases nor decreases.

4.1.1 VGG19 Training and Validation Graph

The graphs of training and validation loss and accuracy and model summary of VGG19 is shown in figure 4.1, 4.2 and 4.3

it can be seen from figure 4.2 the early stopping stops the training of the model after 25 epochs when the validation loss starts to increase.

4.1.2 ResNet50 Training and Validation Graphs

The graph of training and validation loss and model summary of ResNet50 architecture is shown in figures 4.4, 4.5 and 4.6

4.2 Results

Table 4.1 represents the successful percentage during the testing set of 33 different classes. The table includes the following parameters

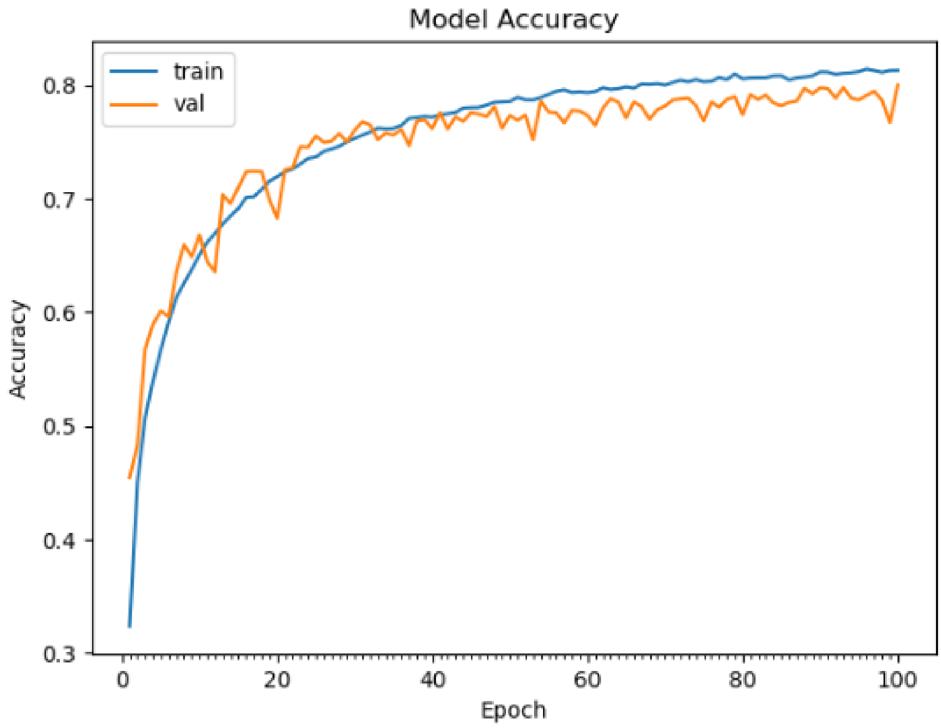

Figure 4.1: VGG19 Training and validation Accuracy

- Successful detection rate when providing the test set in the same directions in which the models were trained.
- Training epochs to achieve the performance of the model.
- Training time

Table 4.1: Performance of CNN models to detect plant disease by using the testing set of original images (pre-trained model)

Model Name	**Training Time (hrs)**	**Same Directions**	**No. of Epochs**
VGG19	3.5	84%	100
ResNet50	5	86%	50

It can be seen from table 4.1 that ResNet50 gives better accuracy than VGG19. Compared to the results of our models to the results written in literature, our model performances are better than the researches conducted in the past because such as in [23] we have choose large amount of disease classes plus number of species which increase the complexity of our models. Furthermore, in literature most of work has been done by choosing very small amount of disease classes and also the number of species 2.1 2.2 2.5. Moreover, our research focus was to work on those plant disease on which the work has not been done before.

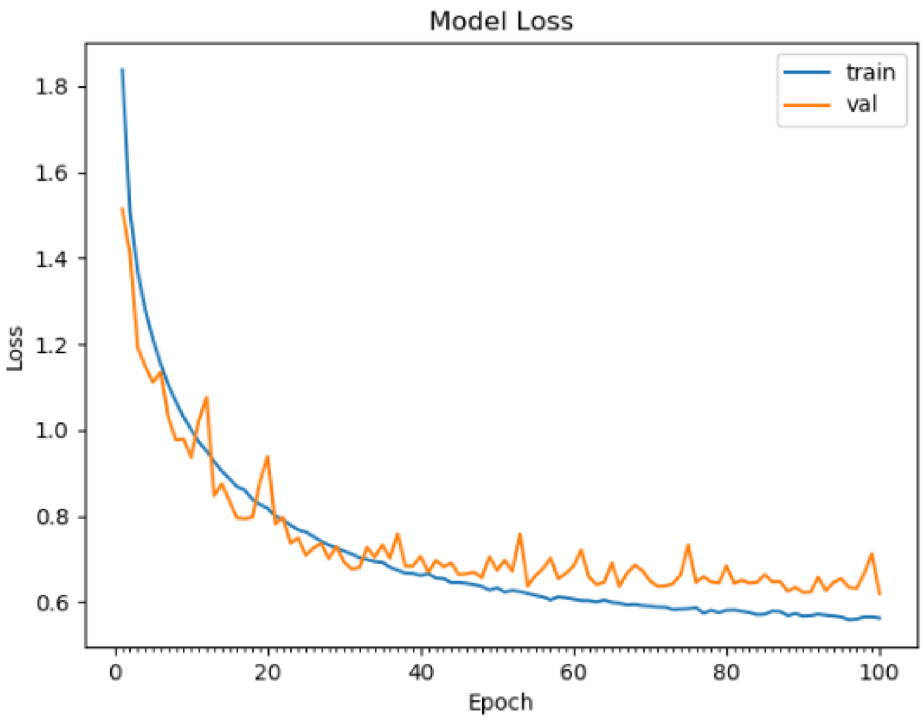

Figure 4.2: VGG19 Training and validation Loss

4.3 Conclusion of the Chapter

In this chapter, we have conducted different experiments to achieve the detection accuracy of plant disease. The accuracy has been achieved from the models trained on both laboratory and natural images but the number of natural images classes was greater than the number of laboratory images classes. We have tested VGG19 and ResNet50 for model accuracy results in that ResNet50 outperformed in detecting the disease of the plants. We have concluded this chapter with remarks that our models give better accuracies when providing the data in the same directions in which the models were trained.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg19 (Model)                (None, 512)               20024384
_________________________________________________________________
dense (Dense)                (None, 33)                16929
=================================================================
Total params: 20,041,313
Trainable params: 16,929
Non-trainable params: 20,024,384
_________________________________________________________________
```

Figure 4.3: VGG19 model summary

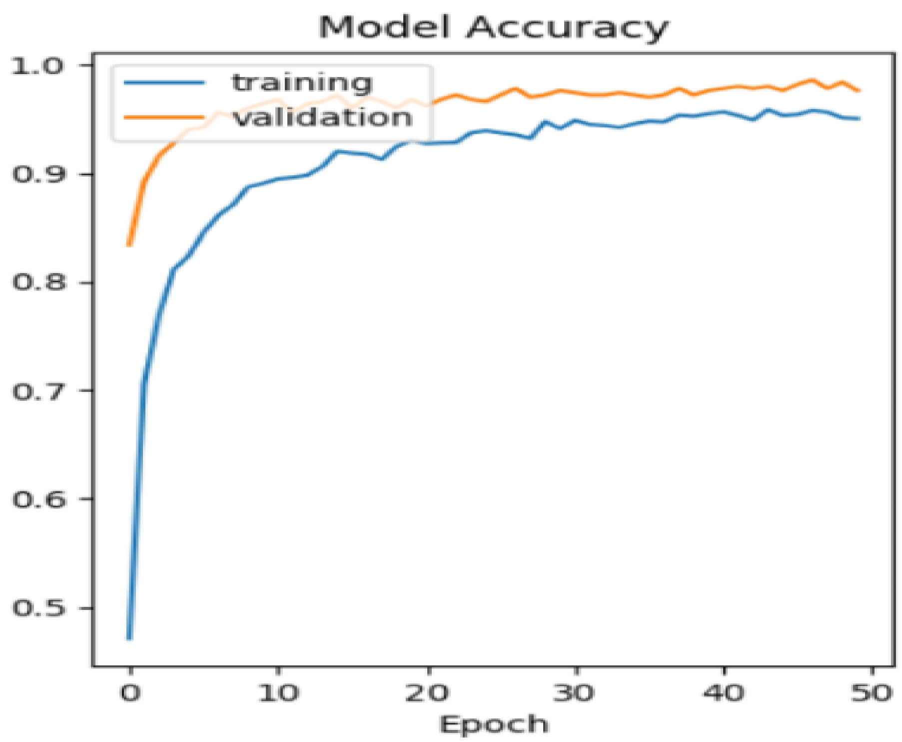

Figure 4.4: ResNet50 Training and validation Accuracy

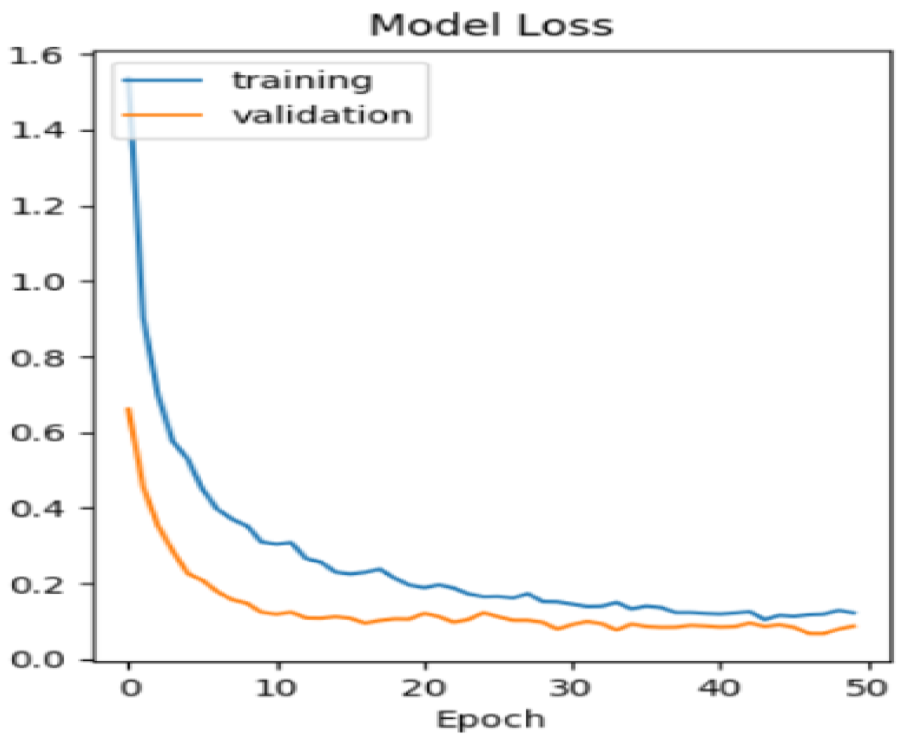

Figure 4.5: ResNet50 Training and validation Loss

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
resnet50 (Model)             (None, 2048)              23587712

_________________________________________________________________
dense (Dense)                (None, 33)                67617
=================================================================
Total params: 23,655,329
Trainable params: 23,602,209
Non-trainable params: 53,120

_________________________________________________________________
```

Figure 4.6: ResNet50 model summary

Chapter 5

Conclusion and Perspectives

5.1 Conclusion and Perspectives

In this research, deep learning models were implemented to identify the plant diseases through images of plants. The primary steps of conducting the research consisted of different steps which involve image gathering, image pre-processing, data augmentation, CNN training, and testing. At the first step of our concerned study, we have separated the training and test set from the large dataset. We have tested our model on 500 images augmented during runtime when given to the trained network but the most successful method was when providing the plant's disease images in which the model was trained and validated. We have used CNN models VGG19 and ResNet50 both as a feature extractor and a detector as well at the same time.

The training and validation of the model were conducted using 10,066 images of diseased plants taken from the real condition field. The data consisted of 33 different classes of nine plants species. The most successful model was ResNet50 with a detection rate of 86% with original images. Based on this detection rate, it is evident that convolution neural network is suitable for the identification of plant disease detection.

Furthermore, we have used pre-trained CNN network VGG19 and ResNet50 for the detection of plant disease. We have fined tuned these networks by adding a dense layer with softmax classifier for our focused study. In this study, we have trained and validated our models on both laboratory and natural images of plant diseases but the number of natural images classes was greater than the laboratory images classes. This leads to our majority concern to train the model. Our major research contribution was to collect the images of plant disease in the real condition field.

Moreover, for small power utilization of training the single image indicated that this mechanism of identification of plant disease could also be developed by making mobile applications that can be used by the farmer for detection of plant disease with low cost. For the future point of view, the focus of detecting the plant disease can be on pure natural images with a huge amount of data on the diseases that have not covered in the literature.

Chapter 6

Appendix

- Our paper has been accepted in 25th international conference on automation and Computing, Lancaster University, UK.

Bibliography

[1] Fao, http://www.fao.org retrieved from food and agriculture organization of the unites states: http://www.fao.org/news/story/en/item/35571/icode. 2009, September 23. Cited on p. 1.

[2] Heba Al-Hiary, Sulieman Bani-Ahmad, M Reyalat, Malik Braik, and Zainab ALRahamneh. Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications*, 17(1):31–38, 2011. Cited on p. 13.

[3] Ahmad Arib Alfarisy, Quan Chen, and Minyi Guo. Deep learning based classification for paddy pests & diseases recognition. In *Proceedings of 2018 International Conference on Mathematics and Artificial Intelligence*, pages 21–25. ACM, 2018. Cited on p. 15.

[4] Jihen Amara, Bassem Bouaziz, Alsayed Algergawy, et al. A deep learning-based approach for banana leaf diseases classification. In *BTW (Workshops)*, pages 79–88, 2017. Cited on p. 12.

[5] Ronnel R Atole and Daechul Park. A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies. *INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS*, 9(1):67–70, 2018. Cited on p. 15.

[6] Saraansh Baranwal, Siddhant Khandelwal, and Anuja Arora. Deep learning convolutional neural network for apple leaves disease detection. *Available at SSRN 3351641*, 2019. Cited on p. 13.

[7] Jayme Garcia Arnal Barbedo. Digital image processing techniques for detecting, quantifying and classifying plant diseases. *SpringerPlus*, 2(1):660, 2013. Cited on p. 15.

[8] A Camargo and JS Smith. An image-processing based algorithm to automatically identify plant disease visual symptoms. *Biosystems engineering*, 102(1):9–21, 2009. Cited on pp. 3 and 5.

[9] Piyush Chaudhary, Anand K Chaudhari, AN Cheeran, and Sharda Godara. Color transform based approach for disease spot detection on plant leaf. *International Journal of Computer Science and Telecommunications*, 3(6):65–70, 2012. Cited on p. 6.

[10] Solemane Coulibaly, Bernard Kamsu-Foguem, Dantouma Kamissoko, and Daouda Traore. Deep neural networks with transfer learning in millet crop images. *Computers in Industry*, 108:115–120, 2019. Cited on p. 15.

[11] İlke Çuğu, Eren Şener, Çağrı Erciyes, Burak Balcı, Emre Akın, Itır Önal, and Ahmet Oğuz Akyüz. Treelogy: A novel tree classifier utilizing deep and hand-crafted representations. *arXiv preprint arXiv:1701.08291*, 2017. Cited on p. 12.

[12] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *international Conference on computer vision & Pattern Recognition (CVPR'05)*, volume 1, pages 886–893. IEEE Computer Society, 2005. Cited on p. 3.

[13] Konstantinos P Ferentinos. Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, 145:311–318, 2018. Cited on p. 13.

[14] Alvaro Fuentes, Sook Yoon, Sang Kim, and Dong Park. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9):2022, 2017. Cited on p. 12.

[15] Erika Fujita, Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. Basic investigation on a robust and practical plant diagnostic system. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 989–992. IEEE, 2016. Cited on p. 10.

[16] Kiran R Gavhale, Ujwalla Gawande, and Kamal O Hajari. Unhealthy region of citrus leaf detection using image processing techniques. In *International Conference for Convergence for Technology-2014*, pages 1–6. IEEE, 2014. Cited on p. 6.

[17] Celia A Harvey, Zo Lalaina Rakotobe, Nalini S Rao, Radhika Dave, Hery Razafimahatratra, Rivo Hasinandrianina Rabarijohn, Haingo Rajaofara, and James L MacKinnon. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in madagascar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1639):20130089, 2014. Cited on p. 1.

[18] Ardi Hidayat, Ucuk Darusalam, and Irmawati Irmawati. Detection of disease on corn plants using convolutional neural network methods. *Jurnal Ilmu Komputer dan Informasi*, 12(1):51–56, 2019. Cited on p. 15.

[19] Eftekhar Hossain, Md Farhad Hossain, and Mohammad Anisur Rahaman. A color and texture based approach for the detection and classification of plant leaf disease using knn classifier. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pages 1–6. IEEE, 2019. Cited on p. 13.

[20] Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Dongyan Zhang. New optimized spectral indices for identifying and monitoring winter wheat diseases. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):2516–2524, 2014. Cited on p. 5.

[21] Zulkifli Bin Husin, Ali Yeon Bin Md Shakaff, Abdul Hallis Bin Abdul Aziz, and Rohani Binti S Mohamed Farook. Feasibility study on plant chili disease detection using image processing techniques. In *2012 Third International Conference on Intelligent Systems Modelling and Simulation*, pages 291–296. IEEE, 2012. Cited on p. 6.

[22] UNEP IFAD. Smallholders, food security and the environment. *Rome: International Fund for Agricultural Development*, 2013. Cited on p. 1.

[23] Peng Jiang, Yuehan Chen, Bin Liu, Dongjian He, and Chunquan Liang. Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. *IEEE Access*, 7:59069–59080, 2019. Cited on pp. 12, 16, and 30.

[24] Song Kai, Liu Zhikun, Su Hang, and Guo Chunhong. A research of maize disease image recognition of corn based on bp networks. In *2011 third international conference on measuring technology and mechatronics automation*, volume 1, pages 246–249. IEEE, 2011. Cited on p. 8.

[25] Andreas Kamilaris and Francesc X Prenafeta-Boldu. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147:70–90, 2018. Cited on pp. 15 and 16.

[26] Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks. In *International Symposium on Visual Computing*, pages 638–645. Springer, 2015. Cited on p. 10.

[27] Anakha Krishnakumar and Athi Narayanan. A system for plant disease classification and severity estimation using machine learning techniques. In *International Conference on ISMAC in Computational Vision and Bio-Engineering*, pages 447–457. Springer, 2018. Cited on p. 13.

[28] Jessica Kwok and Yu Sun. A smart iot-based irrigation system with automated plant recognition using deep learning. In *Proceedings of the 10th International Conference on Computer Modeling and Simulation*, pages 87–91. ACM, 2018. Cited on p. 15.

[29] SE Lindow and RR Webb. Quantification of foliar plant disease symptoms by microcomputer-digitized video image analysis. *Phytopathology*, 73(4):520–524, 1983. Cited on p. 7.

[30] Bin Liu, Yun Zhang, DongJian He, and Yuxiang Li. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, 10(1):11, 2017. Cited on p. 8.

[31] Jaime Lloret, Ignacio Bosch, Sandra Sendra, and Arturo Serrano. A wireless sensor network for vineyard monitoring that uses image processing. *Sensors*, 11(6):6165–6196, 2011. Cited on p. 7.

[32] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. Cited on p. 3.

[33] Juncheng Ma, Keming Du, Feixiang Zheng, Lingxian Zhang, Zhihong Gong, and Zhongfu Sun. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Computers and Electronics in Agriculture*, 154:18–24, 2018. Cited on p. 13.

[34] Antonia Macedo-Cruz, Gonzalo Pajares, Matilde Santos, and Isidro Villegas-Romero. Digital image sensor-based assessment of the status of oat (avena sativa l.) crops after frost damage. *Sensors*, 11(6):6015–6036, 2011. Cited on p. 7.

[35] Federico Martinelli, Riccardo Scalenghe, Salvatore Davino, Stefano Panno, Giuseppe Scuderi, Paolo Ruisi, Paolo Villa, Daniela Stroppiana, Mirco Boschetti, Luiz R Goulart, et al. Advanced methods of plant disease detection. a review. *Agronomy for Sustainable Development*, 35(1):1–25, 2015. Cited on p. 5.

[36] A Meunkaewjinda, P Kumsawat, K Attakitmongcol, and A Srikaew. Grape leaf disease detection from color imagery using hybrid intelligent system. In *2008 5th international conference on electrical engineering/electronics, computer, telecommunications and information technology*, volume 1, pages 513–516. IEEE, 2008. Cited on p. 8.

[37] Nagane M. Bhosale P. Padwal M. Mitkal P., Pawar P. and Nagane P. Leaf disease detection and prevention using image processing using matlab. *International Journal of Recent Trends in Engineering and Research (IJRTER)*, Volume, 2:2455–1457, 2016. Cited on p. 8.

[38] Sharada P Mohanty, David P Hughes, and Marcel Salathé. Using deep learning for image-based plant disease detection. *Frontiers in plant science*, 7:1419, 2016. Cited on p. 8.

[39] Usama Mokhtar, Mona AS Ali, Aboul Ella Hassenian, and Hesham Hefny. Tomato leaves diseases detection approach based on support vector machines. In *2015 11th International Computer Engineering Conference (ICENCO)*, pages 246–250. IEEE, 2015. Cited on p. 8.

[40] Ernest Mwebaze and Godliver Owomugisha. Machine learning for plant disease incidence and severity measurements from leaf images. In *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 158–163. IEEE, 2016. Cited on p. 10.

[41] Smita Naikwadi and Niket Amoda. Advances in image processing for detection of plant diseases. *International journal of application or innovation in engineering & management (IJAIEM)*, 2(11), 2013. Cited on p. 10.

[42] D Oppenheim and G Shani. Potato disease classification using convolution neural networks. *Advances in Animal Biosciences*, 8(2):244–249, 2017. Cited on p. 10.

[43] Sanjay B Patil and Shrikant K Bodhe. Leaf disease severity measurement using image processing. *International Journal of Engineering and Technology*, 3(5):297–301, 2011. Cited on pp. 6 and 7.

[44] Pornntiwa Pawara, Emmanuel Okafor, Olarik Surinta, Lambert Schomaker, and Marco Wiering. Comparing local descriptors and bags of visual words to deep convolutional neural networks for plant recognition. In *ICPRAM*, pages 479–486, 2017. Cited on p. 8.

[45] Artzai Picon, Aitor Alvarez-Gila, Maximiliam Seitz, Amaia Ortiz-Barredo, Jone Echazarra, and Alexander Johannes. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, 161:280–290, 2019. Cited on p. 15.

[46] TV Price, R Gross, Wey J Ho, and CF Osborne. A comparison of visual and digital image-processing methods in quantifying the severity of coffee leaf rust (hemileia vastatrix). *Australian Journal of Experimental Agriculture*, 33(1):97–101, 1993. Cited on p. 6.

[47] Aravind Krishnaswamy Rangarajan, Raja Purushothaman, and Aniirudh Ramesh. Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia computer science*, 133:1040–1047, 2018. Cited on p. 8.

[48] P Revathi and M Hemalatha. Classification of cotton leaf spot diseases using image processing edge detection techniques. In *2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)*, pages 169–173. IEEE, 2012. Cited on p. 10.

[49] T Rumpf, A-K Mahlein, U Steiner, E-C Oerke, H-W Dehne, and L Plümer. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and electronics in agriculture*, 74(1):91–99, 2010. Cited on p. 3.

[50] Sindhuja Sankaran, Ashish Mishra, Reza Ehsani, and Cristina Davis. A review of advanced techniques for detecting plant diseases. *Computers and Electronics in Agriculture*, 72(1):1–13, 2010. Cited on p. 5.

[51] Lalit Saxena and Leisa Armstrong. A survey of image processing techniques for agriculture. 2014. Cited on p. 15.

[52] Joanna Sekulska-Nalewajko and Jaroslaw Goclawski. A semi-automatic method for the discrimination of diseased regions in detached leaf images using fuzzy c-means clustering. In *Perspective Technologies and Methods in MEMS Design*, pages 172–175. IEEE, 2011. Cited on p. 7.

[53] Helmi ZM Shafri and Nasrulhapiza Hamdan. Hyperspectral imagery for mapping disease infection in oil palm plantationusing vegetation indices and red edge techniques. *American Journal of Applied Sciences*, 6(6):1031, 2009. Cited on p. 6.

[54] Punnarai Siricharoen, Bryan Scotney, Philip Morrow, and Gerard Parr. Texture and shape attribute selection for plant disease monitoring in a mobile cloud-based environment. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 489–493. IEEE, 2016. Cited on p. 10.

[55] Barbora Škaloudová, Vlastimil Křivan, and Rostislav Zemek. Computer-assisted estimation of leaf damage caused by spider mites. *Computers and electronics in agriculture*, 53(2):81–91, 2006. Cited on p. 6.

[56] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, 2016, 2016. Cited on p. 8.

[57] Richard N Strange and Peter R Scott. Plant disease: a threat to global food security. *Annu. Rev. Phytopathol.*, 43:83–116, 2005. Cited on p. 1.

[58] Shoby Sunny and MP Indra Gandhi. An efficient citrus canker detection method based on contrast limited adaptive histogram equalization enhancement. *International Journal of Applied Engineering Research*, 13(1):809–815, 2018. Cited on p. 8.

[59] Amos PK Tai, Maria Val Martin, and Colette L Heald. Threat to future global food security from climate change and ozone air pollution. *Nature Climate Change*, 4(9):817, 2014. Cited on p. 1.

[60] CC Tucker and S Chakraborty. Quantitative assessment of lesion characteristics and disease severity using digital image processing. *Journal of Phytopathology*, 145(7):273–278, 1997. Cited on p. 6.

[61] Guan Wang, Yu Sun, and Jianxin Wang. Automatic image-based plant disease severity estimation using deep learning. *Computational intelligence and neuroscience*, 2017, 2017. Cited on p. 12.

[62] Shen Weizheng, Wu Yachun, Chen Zhanliang, and Wei Hongda. Grading method of leaf spot disease based on image processing. In *2008 international conference on computer science and software engineering*, volume 6, pages 491–494. IEEE, 2008. Cited on p. 7.

[63] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang, and Qiao-Liang Xiang. A leaf recognition algorithm for plant classification using probabilistic neural network. In *2007 IEEE international symposium on signal processing and information technology*, pages 11–16. IEEE, 2007. Cited on p. 10.

[64] Shi Yun, Wang Xianfeng, Zhang Shanwen, and Zhang Chuanlei. Pnn based crop disease recognition with leaf image features and meteorological data. *International Journal of Agricultural and Biological Engineering*, 8(4):60–68, 2015. Cited on p. 10.

[65] Min Zhang and Qinggang Meng. Automatic citrus canker detection from leaf images captured in field. *Pattern Recognition Letters*, 32(15):2036–2046, 2011. Cited on p. 5.