

ROSE LEAF DISEASES DETECTION USING CNN

BY

MD. NAYAN HOSSAIN

ID: 172-15-10239

MD. SAHADAT HOSSAIN KHAN

ID: 163-15-8520

And

SITHI MOZUMDER

ID: 172-15-10245

This Report Presented in Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science in Computer Science and Engineering

Supervised By

Ms. Nazmun Nessa Moon

Assistant Professor

Department of CSE

Daffodil International University

Co-Supervised By

Ms. Nusrat Jahan

Sr. Lecturer

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

APPROVAL

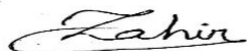
This Project/internship titled “**Rose Leaf Diseases Detection**”, submitted by “**Md. Nayan Hossain**”; “**Md. Sahadat Hossain Khan**” and “**Sithi Mozumder**”, ID No: **172-15-10239, 163-15-8520 and 172-15-10245** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **09-09-2021**.

BOARD OF EXAMINERS



Dr. Touhid Bhuiyan
Professor and Head

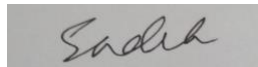
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Chairman

Mr. Gazi Zahirul Islam
Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Mr. Md. Sadekur Rahman
Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Shad Md Imran
Industry Promotion Expert

LICT Project, ICT Division, Bangladesh

External Examiner

DECLARATION

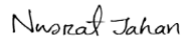
We hereby declare that, this project has been done by us under the supervision of **Ms. Nazmun Nessa Moon, Assistant Professor** and **co-supervision** of **Ms. Nusrat Jahan, Sr. Lecturer**, **Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:



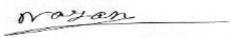
Ms. Nazmun Nessa Moon
Assistant Professor
Department of CSE
Daffodil International University

Co-Supervised by:

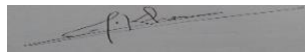


Ms. Nusrat Jahan
Sr. Lecturer
Department of CSE
Daffodil International University

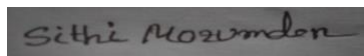
Submitted by:



Md. Nayan Hossain
ID: 172-15-10239
Department of CSE
Daffodil International University



Md. Shadat Hossain Khan
ID: 163-15-8520
Department of CSE
Daffodil International University



Sithi Mozumder
ID: 172-15-10245
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Supervisor Ms. Nazmun Nessa Moon and Co-Supervisor of Ms. Nusrat Jahan, Sr. Lecturer**, Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “Machine Learning and Data Mining” to carry out this project. Her endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

As Bangladesh is fully dependent on agriculture so identifying the diseases of the plants will aid the farmers for better productivity. This paper elucidates the recognition of rose leaf disease which is based on deep learning algorithms. By using deep learning algorithms, the defected images of leaves can be identified precisely which may further help the farmer to identify the leaf disease accurately and then they will immediately take precaution according the disease. In this paper, first we have collected the images from the nursery and labelled each disease after preprocess them in such a way that it is fitted to the particular model that we used. Later, we have modified and implemented CNN model compatible to our dataset to get better outcome. By doing so, we can identify the diseased leaves 96% accurately.

TABLE OF CONTENTS

CONTENTS	PAGE NO.
Approval	ii
Declaration	ii
Acknowledgements	iii
Abstract	iv
 CHAPTER	
CHAPTER 1: INTRODUCTION	1-3
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	2
1.4 Research Questions	3
1.5 Expected Output	3
1.6 Report Layout	4
 CHAPTER 2: BACKGROUND	5-9
2.1 Introduction	5
2.2 Related Works	5
2.3 Research Summary	8
2.4 Scope of the Problem	8
2.5 Challenges	8

CHAPTER 3: RESEARCH METHODOLOGY	10-21
3.1 Introduction	10
3.2 Research Subject and Instrumentation	10
3.3 Data Collection Procedure	11
3.3.1 Attributes	11
3.3.2 Data Pre-processing	12
3.3.3 Data Organizing	13
3.3.4 Data Storing	14
3.3.5 Deep Learning Algorithm	14
3.4 Statistical Analysis	19
3.5 Implementation Requirements	21
 CHAPTER 4: EXPERIMENTAL RESULT AND DISCUSSION	 22-29
4.1 Experiment Setup	22
4.2 Model Summary	23
4.3 Experimental Result and Analysis	25
4.4 Discussion	29

CHAPTER 5: SUMMARY, CONCLUSION, RECOMMENDATION, IMPLICATION FOR FUTURE RESEARCH	30-31
5.1 Summary of the Study	30
5.2 Conclusions	30
5.3 Recommendations	30
for 5.4 Implication Further Study	31
 REFERENCES	 32-33
APPENDIX	34
PLAGIARISM REPORT	35

LIST OF FIGURES

FIGURES	PAGE NO.
Figure 3.1: Abiotic	12
Figure 3.2: Balckspot	12
Figure 3.3: Hole	12
Figure 3.4: Salt	12
Figure 3.5: Dataset Labelling	13
Figure 3.6: Leaf Disease Classes	14
Figure 3.7: Maxpooling From Matrix	16
Figure 3.8: Kernel Size	17
Figure 3.9: Rectified Linear Unit	17
Figure 3.10: Sigmoid Function	18
Figure 3.11: Proposed Model Structure	20
Figure 4.1: Model Visualization	23-24
Figure 4.2: Sample Output	25-26
Figure 4.3: Loss Graph	27
Figure 4.4: Accuracy Graph	27
Figure 4.5: Confusion Matrix	28

TABLE NAME	PAGE NO.
Table 1.1: Type of Rose Leaf Diseases	1
Table 2.1: Summary of Previous Researches	7
Table 3.1: Dataset Overview	19
Table 4.1: Experimental Result	25

CHAPTER 1

INTRODUCTION

1.1 Introduction

In, Bangladesh, a lot of people are depending upon agriculture. Flowers are providing a huge income in it. Among flowers, roses are the most famous because individuals send roses to their loving or beloved person to show love, integrity. Whenever there's an occasion roses are used at high measure. Some worldwide occasions such as valentine, mothers day, friendship day and in all of them every people give roses to their loved person. Rose has always been representing love. Not only in occasion or ceremony ,even when someone you loved passed away we remember them by giving them flowers to ease our heart and to show love to our loved ones.

These flowers are cultivated by our hardworking farmers. Each year, a tremendous number of roses are wasted because of various diseases. Many farmers are not aware of the rose diseases and are not able to take precautions. If they knew why there's a sudden defect occurred in rose leaves then they would solve it immediately. For these reasons, identification of diseases of rose leaves in Bangladesh is essential for the farmers because it aid them to identify the causes of it, which in terms, will help them to find the solutions for each disease. There are several types of rose leaf diseases and among them four types of disease will be discussed.

The four classifications of the image that we have used to identify the rose leaf disease detection are as follows:

Table 1.1: Type of Rose Leaf Diseases

Name of Genre	Description
Abiotic	This can be caused by nonliving factors. For instance, it can be caused by freezing injury or high temperature, drought, injury occurred by chemicals etc. Moreover, it can also be occurred by overwatering conditions. Overall, it took place by unorganized conditions of plants to their lifestyle.
Blackspot	Blackspot usually appear in leaves when the fungus <i>diplocarpon rosae</i> infect the trees. It happens usually in rose plants. It is a highly contagious disease and it spreads rapidly.
Hole	There is a larvae or as known as rose slug which is rose sawfly cause this hole in the rose leaves. As the larvae grows the hole became larger. This is also called rose slug damage.
Salt burn	It occurs when the minerals of the water stocked around the root of the plant. It stops the water absorption of the plants which in terms causes salt burn to the leaves.

1.2 Motivation

There are variations of diseases in rose leaves and each of them occurred for some reasons. For instance, as mentioned above, Hole in the leaf diseases happen because of a larvae named as rose sawfly. As each of the disease happen for plethora of issues in the trees so if they can be identified then farmers can easily find solution for each disease . Moreover, sometimes the farmer doesn't even realize that some of the disease of leaves are contagious and it can happen in his entire farm which then will be a reason for his great loss. Sometimes, the leaves are affected by insects and if the farmer can be aware of what kind of problem occurred this then they can easily resolve that issue.

- For reducing the difficulties to identify the leaf disease.
- To ameliorate the productivity, profit and mitigate the losses.

1.3 Rationale of the Study

There are significant amount of work has been done to rose leaf disease. One individuals have used pretrained CNN model to identify the disease of rose leaves which will be discussed later, however, the accuracy and the output is still undefined.

In this paper, we want to apply our own proposed CNN model to identify the images which have a better outcome and it can successfully identify the defects of each leaf disease. As we mentioned earlier, there are less significant work have been done previously with various types of datasets. Some people have used pretrained models to identify leaf disease from the images, however, in our case we want to propose our own variation of CNN model where we will fine tuning the model and ameliorate the accuracy than others.

1.4 Research Questions

This work is concerned with the problem of automated inspection of rose leaf disease through the use of computer vision. The issues are as follows:

- Is there any way we could fine tuning the convolutional neural network to get better results?
- Can we use popular deep learning techniques such as CNN and what are the parameters and hyper parameters that we can change to improve our work?

1.5 Expected Outcome

- Deep learning approach that successfully recognize the disease of images at a higher rate.
- Can detect various types of disease effectively.
- Publish a research paper based on findings.

1.6 Report Layout

Chapter 1 In this section, we discussed the inspiration for our assignment, the goals we set for ourselves, and the typical outcome of our endeavor.

Chapter 2 In this section, we have discussed the foundation of our study and examined other connected works, similar investigations, the scope of the issues and difficulties, and the extent of the issues and difficulties.

Chapter 3 We're talking about our research subject and the instrument we're using, as well as our data collection technique, statistical analysis, and eventual application of our findings.

Chapter 4 Here, we are taking out our research experimental result and descriptive analysis and find out summary.

Chapter 5 In this chapter we are talking about summary of prediction and conclusion and we have also added further study process.

CHAPTER 2

BACKGROUND

2.1 Introduction

To begin with, we will talk about what this research is all about and what it's all about. The background study includes a review of the area under investigation as well as current information surrounding the subject of the research, previous studies on the subject, and relevant historical information on the subject. In this section, the study should, in the best case scenario, adequately present the history and background information on this topic. In accordance with our suggestion, we will examine and determine the obstacles to be addressed, and we will summarize the research and find solutions to the issues and challenges. The benefits and automated solution must be discovered after conducting extensive investigation. The proposal is a slew of leaf diseases that are identified through the proposal's flaws, which is in line with the proposal. Furthermore, the problem of individuals trying to figure out the disease will be addressed in this research, and the benefits of the research will be evaluated for the benefit of humanity as a whole.

We will discuss our related efforts, a project description, and any challenges that we have experienced while performing this research in this section of the website. Other research articles that are linked to our topic will be discussed, and we will give links to such publications in the related works section. Following that, we will discuss the highlights of our project in the summary section and the challenges we found during our research in the challenge section.

2.2 Related Works

Identifying the leaf disease is variedly known and many kinds of leaf disease has been identified till now. However, Rose leaf detection is still not widely known by the usage of deep learning.

Nishkala & Bhavya [1] in their paper they identified Rose leaf disease and for this process first, they have segmented the images. Then they have used k-means clustering method to

get a pixel distribution. Moreover, after distributed the pixels they have implied canny edge detection by measuring the difference between pixel color and the weights of the images. Finally, they applied CNN for classification of the disease. In this paper, the accuracy and the classes of the datasets are unidentified but process seems extremely good to continue our research.

The paper [2] clarified that they have identified rose diseases from the images of rose. However, the rose disease can't be a proper solution as if the tree is already affected with certain kind of disease then the outcome will be disastrous. They have implemented transfer learning by usage of MobileNet in their research. They have even achieved an optimistic output of 95.63% accuracy in their research.

Thus, many people currently working on variation of leaf disease and among them Surampalli Ashok, Gemini Kishore has worked on tomato leaf disease [3]. As mentioned in paper [1] they have used similar techniques to identify the leaf disease detection. Similarly, they have segmented the images and then clustered them to get the weights of the images and finally they have used open source algorithms for recognition of diseased leaves.

To begin with with paper [4], R. Mounika, Dr. p. Shayamala Bharathi has worked with plant leaf disease. In this paper they have used healthy and disease images of leaves. In order to identify the images, they have segmented the images and clustered them to the beginning of the experiment by k means algorithm. After clustering them they extract the features of them by using SVM algorithm. If the image output is a diseased image then it will illustrate disease detected otherwise the output will be healthy. However, in this paper it doesn't conclude how much accurately they can identify the images.

Kawcher Ahmed; Tasmia Rahman Shahidi [5] worked with three types of rice plant disease in their paper. These three types of disease are used in their dataset. Before applying algorithms, they first used white background for each of the images and after that they have applied multiple machine learning algorithms to identify the disease by the completion of necessary algorithms. J48, KNN, Naïve Bayes, Logistic Regression and Decision Tree algorithms has been used in this process. However, among all of them decision tree

algorithms gave satisfactory result with 10fold cross validation. It predicted with 97% accuracy from the test dataset.

In our paper, we will be using CNN to identify the diseased images because among all of the research CNN provide better result if it can be compatible with the dataset.

Table 2.1: Summary of Previous Researches

SL	Author	Methodology	Description	Outcome
1.	Varsha J. Sawarkar, Dr. Seema Kawathekar	They have used deep learning and image processing methods.	The goal of this paper is rose leaf disease identification by using the concepts of deep learning.	In this paper, they successfully identified the rose leaf disease by using CNN. (The accuracy and the classes of the datasets are unidentified)
2.	Aditya Rajbangshi, Toma sharkar	Transfer learning has been used here with mobileNet algorithm	Rose disease has been identified from the images by using pretrained model.	A satisfactory outcome has been achieved with this algorithm which is 95.63%.
3.	Surampalli Ashok, Gemini Kishore	Open source algorithms has been used in this paper.	This paper identifies leaf tomato leaf disease accurately.	This paper safely identifies the tomato leaf disease by using some open source algorithms and these algorithms has been applied after segmenting and clustered the images.
4.	R. Mounika, Dr.p.	They have implemented	After segmenting and cluster the images matlab	If the image is defected then the

	Shayamala Bharathi	image segmentation, K-means clustering, SVM.	has been used to identify the defects.	output will be diseased otherwise the output will be healthy.
5.	Kawcher Ahmed, Tasmia Rahman Shahidi	Machine learning algorithms has been used in this process.	Three types of diseased leaf has been used in this paper where the background has been white as default. Multiple machine learning algorithms has been used for this research.	Decision tree with 10 fold cross validation achieved 97% accuracy on the dataset.

2.3 Research Summary

Our examination work plans to see what type of disease each image holds. Moreover, each types of disease will provide how much accurate it is. Since our output index depends on four types of diseases so the output index will be four and among them the highest accuracy of each indexes will be selected as diseased images. For instance, if an image has hole in it then then in the output the hole will have higher accuracy than all other classifications of images.

We have used CNN algorithms to find out the accuracy of the diseased leaves and we have also made a confusion matrix to see the output comparison with each types of disease. Higher the accuracy the better will be the system.

2.4 Scope of the Problem

According to the research there have a big impact of rose leaf disease to solve out the issue of identifying disease for the sake of human kind to make life easy. Analyzing the previous research, we figure out that deep learning and automated leaf disease classification using the deep learning method will be an ideal way to solve out the difficulties to identifying the disease of the images.

2.5 Challenges

In order to improve our predictive accuracy, the most difficult task is data collection. From the standpoint of Bangladesh, data collection is a time-consuming and complicated operation. We have to collect the data from nursery and it was quite difficult for us to get the permission because nobody gives permission to get the data because of its privacy. After collecting the dataset with much hardship we have to preprocess the dataset and that was another tedious work. There was some imbalanced between the dataset so that was another problem for this reason we have balanced the dataset manually. Applying CNN is another challenge because we have to preprocess the dataset according to our model. If the preprocessing doesn't go according to the model then the output will not be good. Moreover, to ameliorate the accuracy the hyperparameter tuning is quite a big challenge as well.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The amelioration of deep learning algorithms has fascinated the entire era which has been aiding us to identify images and separating them in various ways. As we have used a convolutional neural network to identify our leaf disease so it can predict the images accurately.

Before going on to elucidate our work, the emphasizing of CNN is essential. A CNN is a deep learning algorithm and it is applied to identify images, time-series data, objects, or various aspects. It can also be used to classify sounds from various types. As mentioned that Convolutional neural network can be used in identifying sound, data, or objects and even in the video so there is some classification in them as well. In terms of times, series data conv1D is used which two dimensional and when it works with images then we need to use conv2D which is for three-dimensional data. Finally, the last one is conv3D which is used for multi-frameworks where are frames in each time span. Thus, we will be using conv2D in our work which is used for a 2-dimensional array. To exemplify, when we will try to find an image we will use a 2-dimensional array which will later be used as individual data as an image array. Furthermore, conv2D will be used in convolutional neural networks.

In this section the description of algorithm and how the dataset has been prepared will be discussed in a sequence manner. First, the description of research subject will be discussed properly. A more thorough explanation of each of these steps will be provided, in order to better carry out this research. In the data collection and preprocess the elucidation of how the data has been collected and their processing will be described. Then in the proposed methodology we will describe our algorithm and methodology that has been used to identify genre for classification. Consequently, in the statistical analysis part the flow chart and our way of work along with some visualization will be provided. Finally, the chapter will be closed about how we have accomplished our project.

3.2 Research Subject and Instrumentation

We can see that the information is the most important component of the examination. It is an extremely important aspect of our investigation work for a specialist to locate great data as well as an excellent technique or model. We also need to look at previous examination papers that are comparable. At that moment, we'll have to decide between a few options:

- Which types of data should be gathered?
- How can we assure that the data gathered is accurate?
- Is the data to be structured in the same manner for each piece?
- How do you suggest labeling each piece of data?

3.3 Data Collection Procedure

The most difficult part of our thesis is to collect dataset. For security and health precaution purposes no one was allowing us to collect the dataset. However, somehow, we have managed to get the permission from the nursery which is situated in savar golapbagan. In this process, there were variation of disease but many of them can't be included because there weren't enough diseased images. We have to roam around a lot to collect all of the images. In this process, first we have captured the images with digital camera and then we stored it. The images weren't that perfect because of the lighting, for this reason we have to preprocess it in such a way that we could use it in our algorithm. The preprocessing methods will be elucidated later to the upcoming section.

3.3.1 Attributes

The obtained dataset contains 4 attributes. The main attributes are (abiotic,blackspot, hole, saltburn). Training dataset and testing dataset have been separated into two separate portions, with each containing the dataset. One of the several datasets used for training is also set aside for validation.

Some sample images of the dataset are illustrated below:



Fig 3.1: Abiotic



Fig 3.2: Blackspot



Fig 3.3: Hole



Fig 3.4: Salt

These four types of diseased images are used in our dataset.

3.3.2 Data Pre-processing

In this process, first we manually crop only the defected images and put them in folder to classify them. Then we have divided them into two folder train and test.

```

In [7]:
random.shuffle(train_data)
random.shuffle(test_data)

for lbl in train_data[:10]:
    print(lbl[1])

X_train=[]
y_train=[]

for features,label in train_data:
    X_train.append(features)
    y_train.append(label)

```

```

1
2
2
0
0
0
3
0
2
2

```

Figure 3.5: Dataset labelling

Then we have labelled each of the classes. Abiotic has been labelled as ‘1’ and blackspot has been labelled as ‘2’ Hole has been labelled as ‘3’ and saltburn has been labelled as ‘4’. Some random labelling has been illustrated in the figure 3.6.

In our dataset the images are (64, 64, 3) pixels, where 3 resembles that this image is an RGB image and not greyscale. Moreover, if this image converts into an array then there will be three 64*64 matrices. As we know when we work with greyscale images there remains only one matrices, however in this this case each matrices will have a huge value to be accumulated later on in our algorithms to identify the variation of each images been converted to 64*64 pixels.

3.3.3 Dataset Organizing

We have trained and tested the data to organize the data. The data are then saved in two folders. Furthermore, we have used a test folder to test validation of train data. Then we have created sub-folders in the test and train folders like abiotic, blackspot hole, and salt.

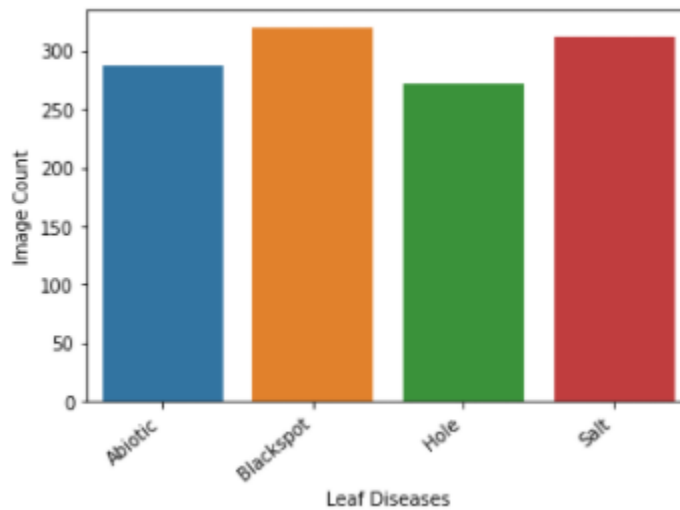


Figure 3.6: Leaf Disease Classes

3.3.4 Data Storing

Here, we keep all of our data in Kaggle, as we use the Kaggle notebook. Using saved data from Kaggle, we can incorporate the stored data in our current project. This means that their output will be saved as a .h5 file. This can be utilized in online applications, as well as in anything else. In the future, these data could be useful in our project work by providing clear step-by-step instructions.

3.3.5 Deep Learning Algorithms

We have modified CNN model in our thesis for better outcome. The hyperparameter tuning and modification process will be elucidated later on. All the applied algorithms on python and libraries are NumPy, Pandas, Scikit-learn and Matplotlib.

Modified Convolutional Neural Network Known as a CNN (ConvNet/CNN), a Deep Learning algorithm, a CNN (ConvNet/CNN) receives images as input and uses them to comprehend its weights and biases for distinct parts of the image, as well as to distinguish them from one another. The amount of data pre-processing required by ConvNet is less and more straightforward than the amount of data pre-processing required by alternative classification techniques. Unlike traditional neural networks which learn to discriminate

between characteristics over time, ConvNet is capable of learning the images of attributes and distinguishing them from one another in the future.

As convolutional neural network has the ability to identify the images by extracting its features and separating them so we have used this in our image dataset. In convolutional neural network there are several features such as kernel, strides, pooling, dropout, filter activation function etc. All has their own importance in building a full neural network.

Firstly, the image identification process is done via detecting the edges. The edges could be vertical or horizontal.

When we get an output by involving over the image by the usage of filter then we attach bias to those results and then we use an activation function to compute the number of activations. Then the one layer of CNN equation are as follows:

$$\mathbf{z}^{[1]} = \mathbf{w}^{[1]} * \mathbf{a}^{[0]} + \mathbf{b}^{[1]}$$

$$\mathbf{a}^{[1]} = g(\mathbf{z}^{[1]})$$

To this scenario, the input image is 64*64*3 and filters are (4*4*4) which is weights $\mathbf{w}^{[1]}$. This is the first layer in our CNN the in the second layer the activation $\mathbf{a}^{[1]}$ from layer 1 will act as the activation for layer 2. With this in mind, the number of parameters are dependent on the size of the images. A layer structure in our CNN is comprised of three kinds.

- Convolution layer
- Pooling layer
- Fully connected layer

The modified CNN has 4 layers of 2-dimensional Convolution layer has been used and in each layer there are maxpooling layers available. These layers are usually used to extract the features of the images. It reduces the input size of the images as well as increase the computation speed. For instance, if we imagine a 4*4 pixel image then the application of max pooling will deteriorate the matrix result in the following way:

1	3	2	1		
2	9	1	1		
1	3	2	3	9	2
5	6	7	8	6	8

Figure 3.7: Maxpooling from matrix

In fig 3.5.2 the illustration is done for 2*2 pooling where for each 4 pixels we deduces the amount and return 1 pixels. Similarly in our model 2*2 maxpooling has been used which will help to identify the features of the images. The maxpooling actually works by taking the maximum value in 2*2 pixels. Additionally, there are other pooling methods are available as well such as average pooling, where the average value of 2*2 strides will be taken.

There are a lot of hyper parameters and parameters available in convolutional neural network. Among them kernel size is really important. Deep neural networks, or CNNs as they are more specifically known, are essentially a stack of layers formed by the action of a number of filters on an input signal. Those filters are often referred to as kernels. In our proposed model kernel size of 3*3 has been used which means that in an image the stack of layers will be built upon 3*3 kernel and after each completion of calculation of a kernel in a convolutional neural network it will traverse to the next kernel and so on. This is how a fully stack of layers will be created.

1	3	2	1	5	6
2	9	1	1	2	9
1	3	2	3	4	8
5	6	7	8	1	5
7	8	7	1	4	4
4	2	5	6	1	3

Figure 3.8: Kernel Size

In each model there are activation function are added. Each activation function are necessary for the enumeration of the next layer. In my proposed CNN relu or rectified linear unit has been used as an activation function. For its faster and better performance relu has been used as a default neural network activation function. We didn't change the activation function because of its advanced performance.

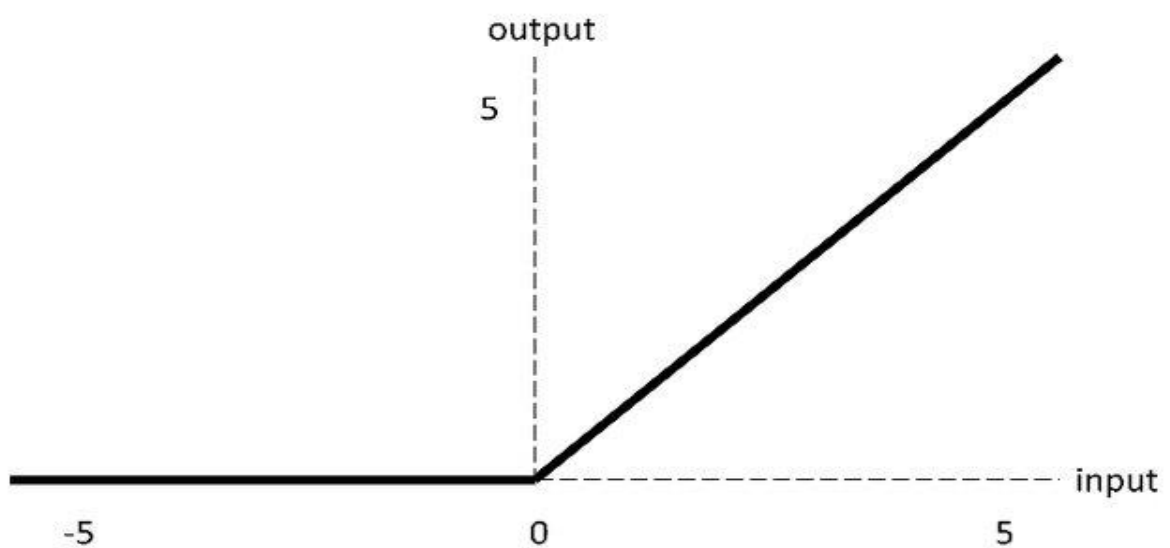


Figure 3.9: Rectified Linear Unit

Currently, just half of the ReLU has been revised (from bottom). In this case, the function $f(z)$ is equal to zero, and when z is greater than or equal to zero, the function $f(z)$ is equal to the value of the variable.

$$R(z) = \max(0, z)$$

Specifically, the difficulty in this scenario is that all of the negative values are immediately turned to zero, reducing the model's ability to appropriately fit or train from the data set. In other words, any negative input given to the ReLU activation function is immediately converted to zero in the graph, which has an adverse effect on the resultant graph because the negative values are not adequately mapped to the positive values. That is why in the last layer we have used the sigmoid function so that the negative values can be measured as well in order to get the expected outcome, as shown in the image below. In terms of training in each layer relu has been used so the negative value doesn't affect much.

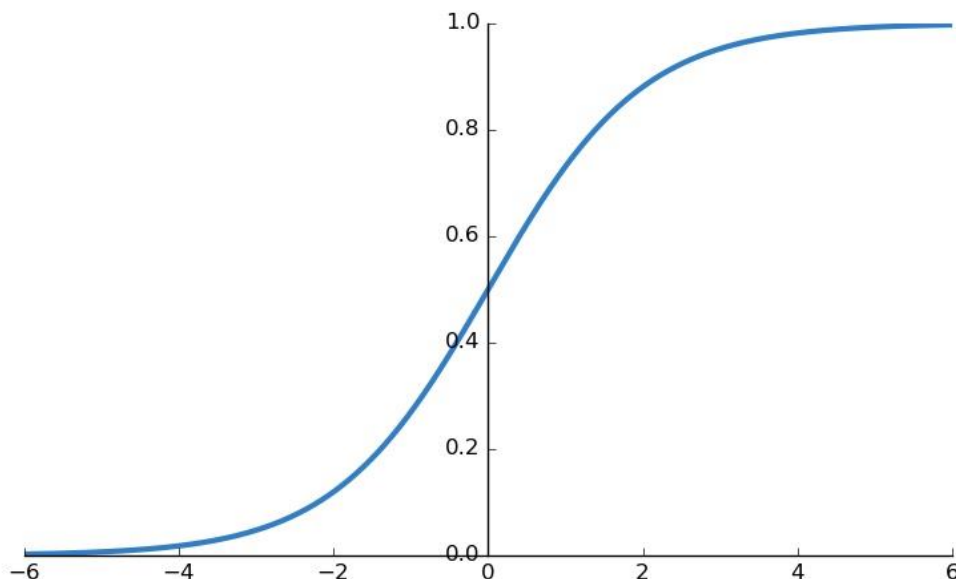


Figure 3.10: Sigmoid function

However, as sigmoid function provide values from zero to positive 1 to in the output layer we have used sigmoid function so that output remains accurate. The main reason of using sigmoid function in the output layer is because sigmoid function always returns value (0,1). The equation of sigmoid function:

$$S(x) = \frac{1}{1 + e^{-x}}$$

The value of the sigmoid function will be either zero or one, depending on the equation, which makes it an excellent choice for obtaining the probability as an output.

Finally, the total structure of CNN is compile and trained on the training dataset with a learning rate of .001. Furthermore, there has been a dropout of 20% has been added in my proposed CNN. Usually, dropout is used to reduce the error of the model by removing the garbage value and focusing more on the main features.

3.4 Statistical Analysis

We have used 1430 images in our datasets and among them 1192 images were used in our train images and 238 images were used in test images. We have accumulated 20% of the train images in our test dataset. In figure 3.5 the amount of each classes of images are illustrated.

Table 3.1: Dataset Overview

Overall Data	Train	Test	Resolution
1430	1192	238	64*64*3

To apply deep learning algorithms, we have preprocessed the images by classifying them in different folder, not only that we have chosen lower pixels because the defects on the

images are hardly be shown. So, we have cropped them first then we have to use lower pixels. Otherwise, if we have picked full resolution images it could provide us lower accurate result at times of applying algorithm.

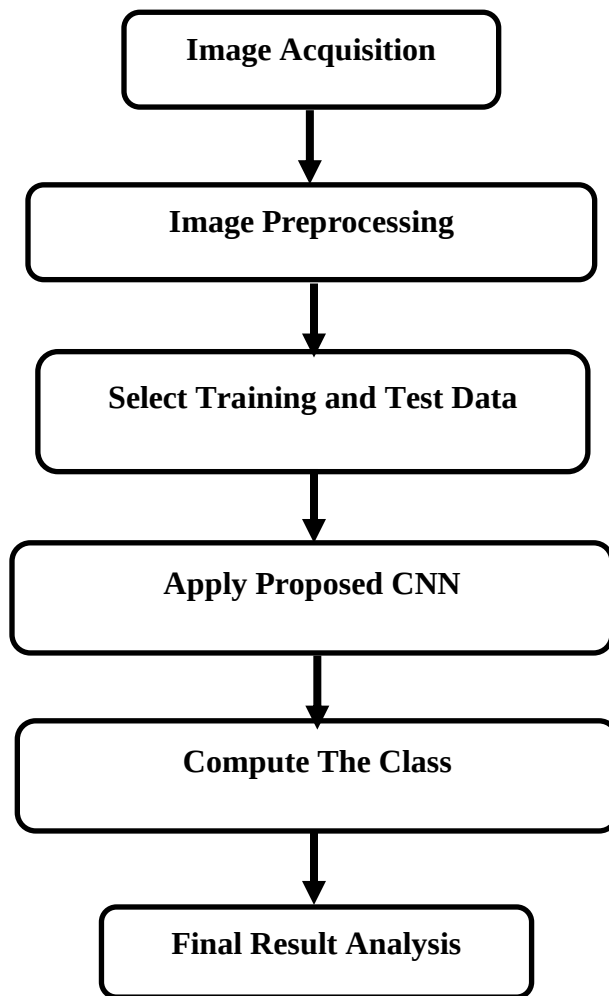


Figure 3.11: Proposed Model Structure

In this figure, we have shown that how we have done our research shortly details. In this figure (3.6), we can know how to go ahead for our target step by step

3.6 Implementation Requirements

The implementation of our work requires a lot of machine power as there are plenty of images. Deep learning algorithms are intensive to use in machine but as they require high power hardware requirement so the output is also fascinating compared to other machine learning algorithms. To ameliorate the proficiency of the model we needed to use a lot of parameters to train the datasets and the surge of these parameters are the only reason for this high end hardware requirements. When we are running the models, in the time of training the datasets, first the datasets are saved into the volatile memory (RAM) to be used later for training and then when the matrix computation is occurred then the GPU or CPU is needed to compute them according to our proposed CNN. The surge of dataset will require more RAM and high end GPU, moreover, the increase of variety of images as well as improved quality of images all are relatable with the hardware requirements.

Our team developed a list of the hardware, software, and development tools that we would require to forecast the leaf disease after thoroughly studying all relevant statistical or theoretical concepts and approaches. The following items are most likely to be required:

Hardware/Software Requirements

- Operating System (Windows 10 or Ubuntu 20.04 or compatible operating system)
- Ram (more than 16 GB)
- Web Browser (preferably chrome)
- GPU Nvidia gtx1660 super.

Developing Tools

- Python 3.6 or more
- Anaconda
- Jupyter notebook

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental Setup

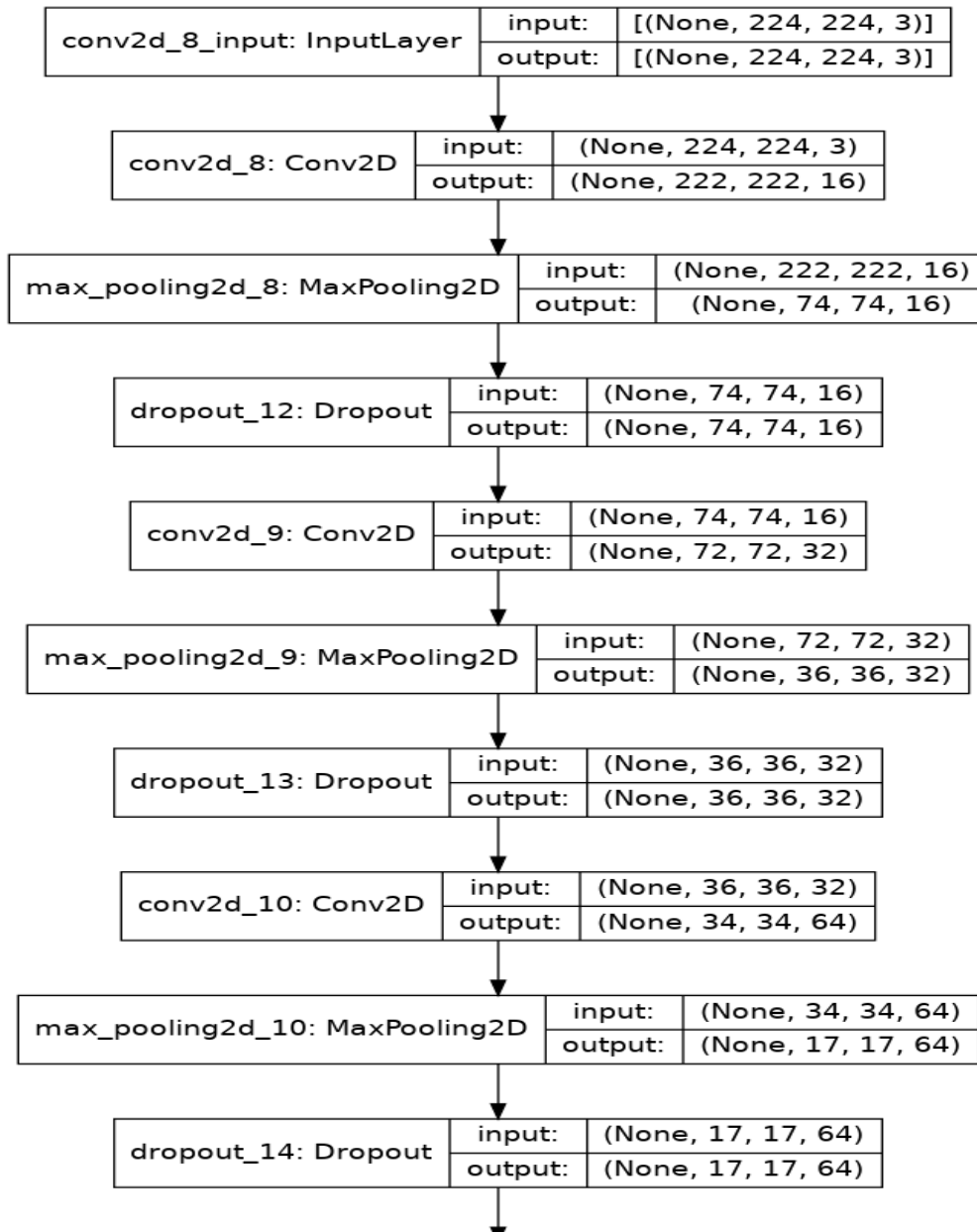
To the next part the explanation of my CNN model and how it works as well as the dataset description has been provided. In this section, the output of our experiment will be provided. We will also elucidate the analysis that we have gathered so far by doing this project.

The final model is created for a specific image type which is later trained in the notebook to get the best outcome. In each images as the features are with lots of variation so it is difficult to separate them in different classes. We were successive to do this by increasing the layers of neurons and lowering the strides per pixels.

- As we have worked with the prediction of leaf diseases we have collected diseased leaf dataset.
- Most of our time expenditure on collecting data as it was the most tedious work
- We have likewise gathered data from a nursery as described before.
- After we have preprocessed the data, they become suited for our model..
- At that point we have finished and preprocessed the data so we can applying our algorithms.

4.2 Model Summary

To identify the disease various types of algorithm could be used. Previously, many researchers have implemented pre-trained model of transfer learning technique. However, in our thesis work we will be applying our own proposed CNN by fine tuning the CNN model to get the best outcome. In this approach, we have used convolutional neural network where there are 470,629 parameters. The whole model is shown in the figure below:



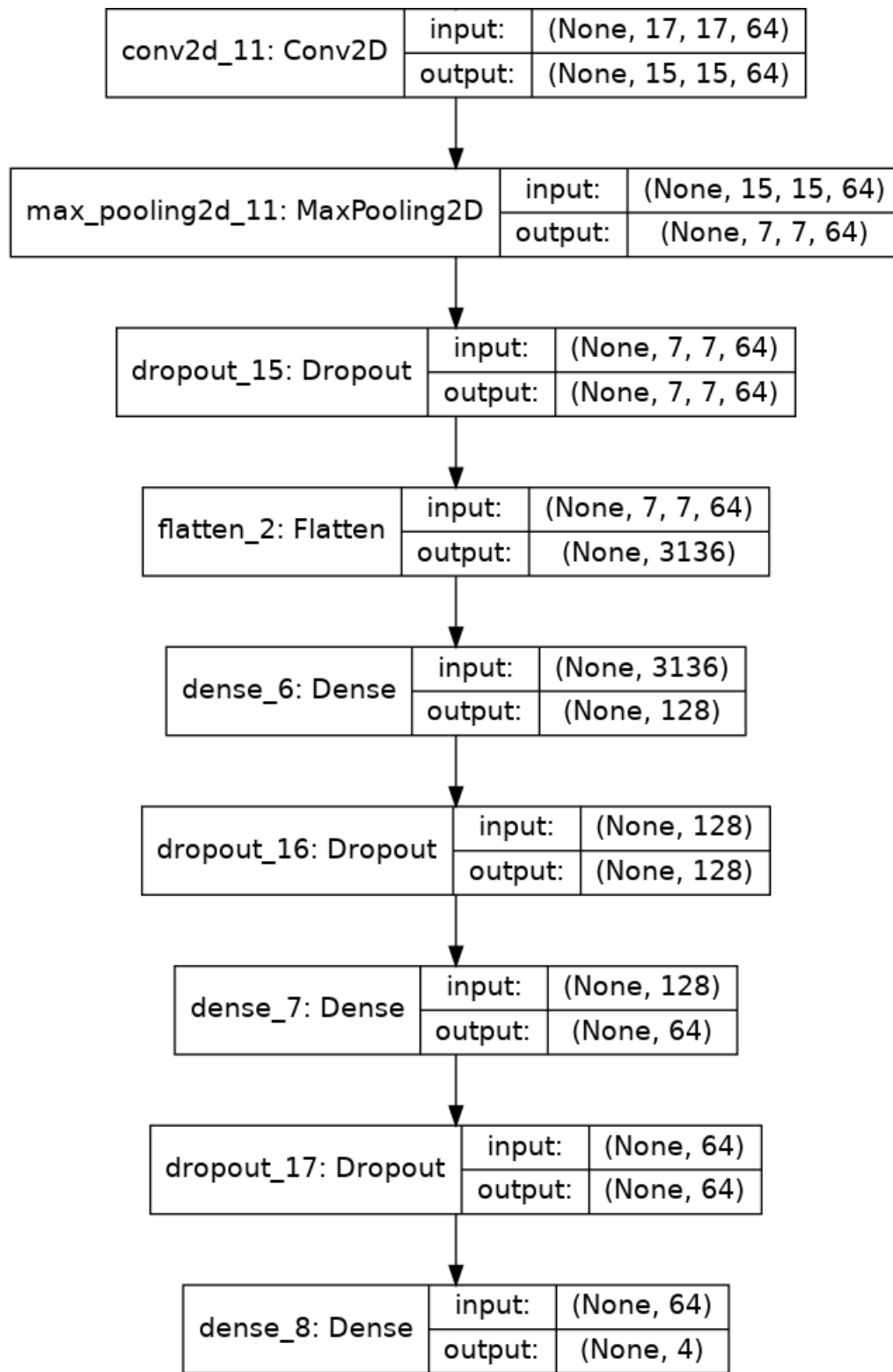


Figure 4.1: Model visualization

In this proposed CNN model we have used 4 layer of convolutional neural network where maxpooling have been used in each layer with 20% of dropout. To the next section CNN will be briefly explained with proper equations and examples and how the CNN has been used according to my dataset will also be explained.

4.3 Experimental Result and Analysis

To predict the leaf disease we have used Modified CNN model to get the best outcome. We got an accuracy of 96% and got this result after 100 epochs. So, our prediction was satisfactory for my modified CNN

Table 4.1: Experimental Result

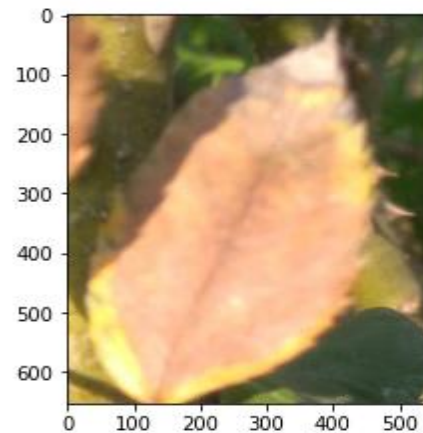
Model	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy
CNN	0.071	96%	.24	88%

The output of the images according to the accuracy are as follows:

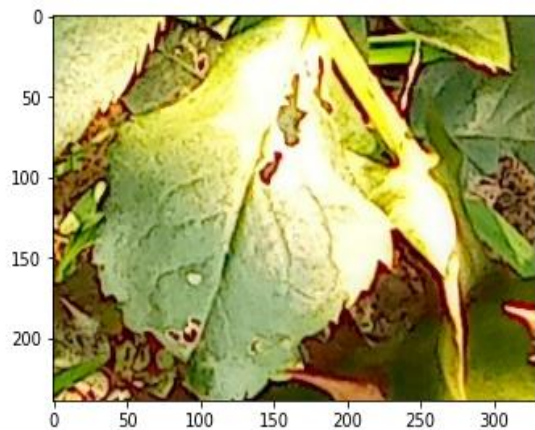
```
Abiotic---->99.91601
Blackspot--->0.00000
Hole----->0.08356
Salt----->0.00044
```



```
Abiotic---->0.02724
Blackspot--->93.48217
Hole----->4.12866
Salt----->2.36192
```



```
Abiotic---->0.00000  
Blackspot--->0.00000  
Hole---->99.99982  
Salt---->0.00018
```



```
Abiotic---->0.00000  
Blackspot--->0.00002  
Hole---->0.00001  
Salt---->99.99998
```

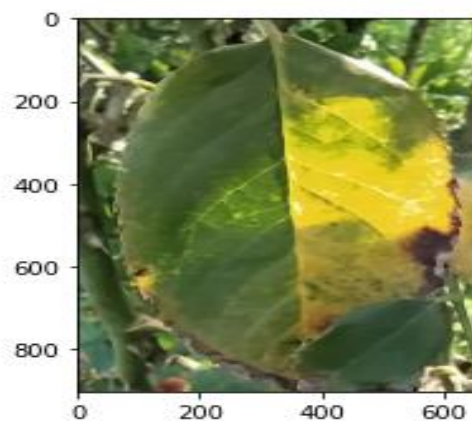


Figure 4.2: Sample Output

In the above images it illustrated the output of each disease. The disease is identified via the highest of its accuracy.

The accuracy graph is illustrated below:

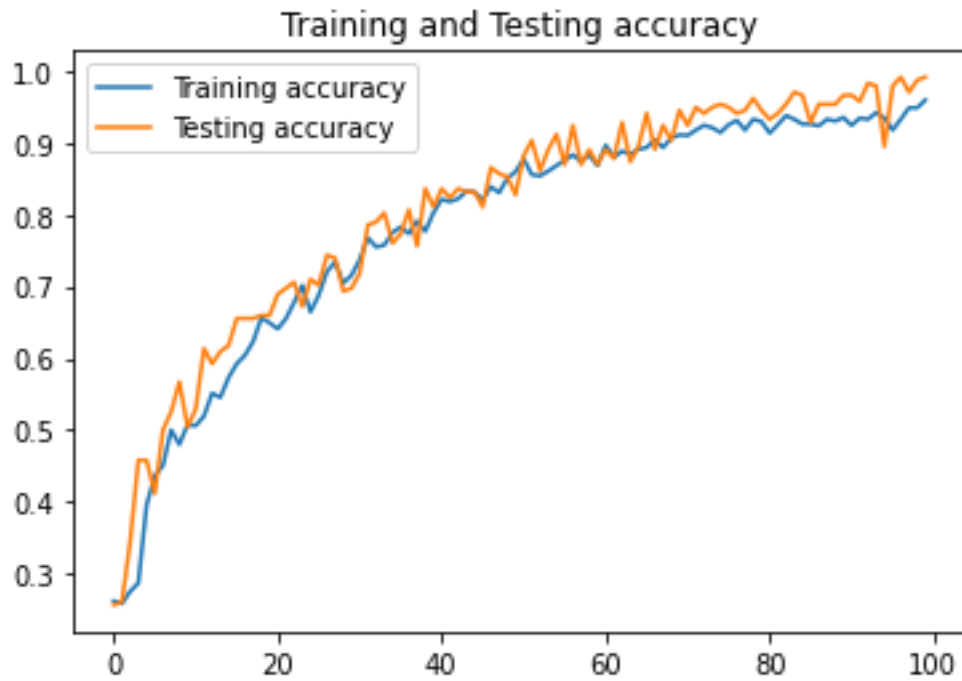


Figure 4.3: Loss graph

Figure 4.4: Accuracy graph

The confusion matrix is given below:

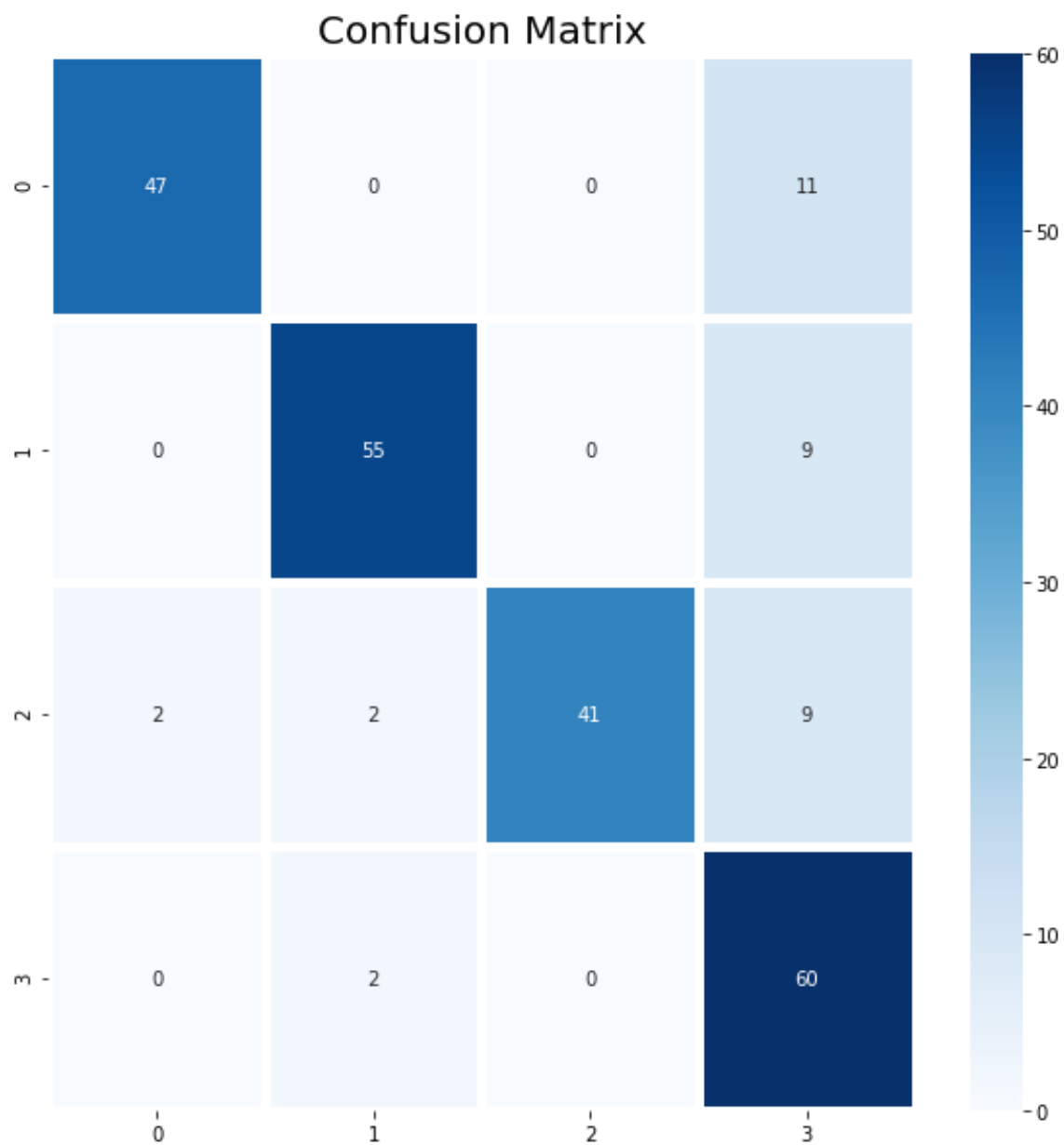


Figure 4.5: Confusion Matrix

4.4 Discussion

The examples of four classes of images are the preparation datasets for CNN.

We have pre-processed and augmented the dataset so that our CNN function can work admirably. For each categories we have added different parameters and took necessary steps so that our model could easily differentiate the images from these four classes. In each image of the dataset has plethora of features which is difficult to identify. Kernel size and padding and all setups for neural network creates a clear perception of dataset for machine to learn and classify the images.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION, IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

We have built up this research undertaking to predict the rose leaf diseases which will further aid the farmers for better production and profit. The research has demonstrated that our modified CNN provides the best outcome and it has successfully predicted the images of the defected leaves. We have gathered four types of disease images in our research and preprocess them. Thereafter, we painstakingly applied data pre-handling decides with the goal that we could make them ideal for the algorithm to be applied. Data collections should be prepared so they can meet the data dealing with reason. In our research works, a portion of the data collections are prepared while different ones are tried.

5.2 Conclusion

Our primary goal was to develop a model that, when compared to other models, can aid in the identification of diseased leaves. To anticipate the images, we used our own modified of a CNN model as a deep learning approach, which we developed myself. This research found that my testing datasets had an accuracy of 98 percent, which suggests that almost all of the images in the test data set could be correctly identified. It appears that my model with an effective marking solution has a promising future and deserves to be investigated in greater depth.

The results of this study will assist farmers in more properly identifying the leaf disease. They can then use the research to examine the leaf disease and come up with treatments for it even if they are not aware of any specific rose leaf diseases at the time of the first research.

5.3 Recommendations

There are some recommendations for this disease identification such are:

- Amelioration of image resolution can also provide better outcome

- Try to clean the dataset and augment them properly.
- Bigger datasets can provide higher accuracy.

5.4 Implication for Further Study

Few Implications that are possible in further study using this Leaf disease images are:

1. This thesis can be further used in web development or in a mobile application for users.
2. More types of disease can be added as addition to this thesis.
3. An addition of datasets can be added into this thesis to improve the output.

REFERENCES

- [1] Varsha J. Sawarkar, Dr. Seema Kawathekar, "Rose Leaf Disease Detection using Digital Image Processing & Deep Learning", International Journal of Emerging Trends & Technology in Computer Science, June,2021.
- [2] Aditya Rajbongshi,Toma Sharkar, "Rose Diseases Recognition using MobileNet", 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT),October,2020.
- [3] Surampalli Ashok; Gemini Kishore; Velpula Rajesh; S. Suchitra; S.G. Gino Sophia; B. Pavithra, Tomato Leaf Disease Detection Using Deep Learning Techniques. 2020 5th International Conference on Communication and Electronics Systems (ICCES), IEEE, (June 2020).
- [4] R. Mounika, Dr. p. Shayamala Bharathi, "Detection of Plant Diseases Using Image Processing", Journal Critical Reviews, 2020.
- [5] Kawcher Ahmed; Tasmia Rahman Shahidi; Syed Md. Irfanul Alam; Sifat Momen, "Rice Leaf Disease Detection Using Machine Learning Techniques", IEEE Xplore,2019.
- [6] MD. Kamran Hosain, Harun-ur- Rashid, Tasnova Binte Tahar, Mohamamd Masudur Rahman, "Genre Recognition of artworks using Convolution Neural Network", IEEE Xplore,2020.
- [7] Zhao-Min chen, Xiu-Shen-Wei, Peng Wang, Yanwen Guo, "Multi- Lable image recognition with grape convolutional networks" CVF Computer vision foundtion , CVPR-2019, IEEE-2019.
- [8] Karim and Zacccone (March 2018). Deep Learning with TensorFlow. Packt Publishing. pp. Chapter 4. ISBN 9781788831109.
- [9] C. Boukouvalas, J. Kittler, R. Marik, M. Mirmehdiand, M. Petrou, "Ceramic Tile Inspection for Color and Structural Defects", under BRITE-EURAM, project no. BE5638, pp 6, University of Surrey, 2006.
- [10] Marina Ivasic-Kos; Miran Pobar; Luka Mikec, "Movie posters classification into genres based on low-level features", 37th Proceedings of the International Convention MIPRO conference 2014. Publisher-IEEE.
- [11] Very Deep Convolutional Networks for Large-Scale Image Recognition by Karen Simonyan.
- [12] Masci, J., Meier, U., Ciresan, D., Schmidhuber, J., Fricout, G.: Steel defect classification with max-pooling convolutional neural networks. In: Proc. of International Joint Conference on Neural Networks (IJCNN), pp. 1–6 (June 2012).

- [13] Eran Ozsarfati, Egmen Sahin, can jozef saul, Alper Yilmaz, "Book genre Classification Based on titles with comparative Machine Learning algorithms", IEEE Xplore,2019
- [14] Nishkala G, Bhavya P R "Novel Genre Classification Using deep Learning", International Research journal of engineering and technology (IRJET) Sept 2020.
- [15] Shikha Gupta, Mohit Agarwal, satbir jain, "Automated Genre Classification of books Using Machine Learning and natural Language Processing", IEEE Xplore,2019.

APPENDIX RESEARCH REFLECTION

Several issues have arisen during the course of our research efforts. However, there were three key issues that needed to be addressed. The first step is to determine the optimal algorithm to use, as well as how to apply that method. The following step is to locate, gather, and preprocess the dataset in accordance with our dataset viewpoint. Last but not least, tuning the model is the most difficult step. When we attempted to make changes to the dataset, we ran into some difficulties. We had a terrible time gathering data because it was so difficult. Because, in the event of a Covid-19 epidemic, data collecting on the ground is time-consuming and too difficult for students to complete. We have, however, obtained authorization to visit the nursery and have acquired the necessary data. And, after a lengthy period of time, numerous efforts, and considerable effort, we were successful.

PLAGIARISM REPORT

Rose Leaf

ORIGINALITY REPORT

4%

SIMILARITY INDEX

4%

INTERNET SOURCES

1%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1

dspace.daffodilvarsity.edu.bd:8080

Internet Source

4%

2

M Anandkumar. "Texton Features and Deep Belief Network for Leaf Disease Classification", Multimedia Research, 2020

Publication

<1%

3

H. Ness. "Influence of the tip-induced electric field on the STM contrast of chemisorbed C₂H₄ on the Si(001) surface", Physical Review B, 04/1997

Publication

<1%

4

Shekoofeh Azizi, Sharareh Bayat, Pingkun Yan, Amir Tahmasebi et al. "Deep Recurrent Neural Networks for Prostate Cancer Detection: Analysis of Temporal Enhanced Ultrasound", IEEE Transactions on Medical Imaging, 2018

Publication

<1%

5

Shupe, S.M.. "Cover- and density-based vegetation classifications of the Sonoran Desert using Landsat TM and ERS-1 SAR

<1%

imagery", Remote Sensing of Environment,
20041030

Publication

-
- | | | |
|--|--|------|
| <div style="background-color: #8B4513; color: white; display: inline-block; width: 20px; height: 20px; text-align: center; line-height: 20px;">6</div> | V. M. Babich. "Hadamard's method and the asymptotics of the spectral function of a second-order differential operator",
Mathematical Notes of the Academy of Sciences of the USSR, 1980 | <1 % |
|--|--|------|
- Publication
-

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off

