APPLE LEAF DISEASE DETECTION USING CONVOLUTIONAL NEURAL

NETWORK

BY

Moriom Akhter
ID: 213-25-073

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

Supervised By

M. Ismail Jabiullah
Professor Dr.
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

SEPTEMBER 2022

APPROVAL

This Thesis titled **"Apple Leaf Disease Detection Using Convolutional Neural Network",** submitted by **Moriom Akhter** ID No: **213-25-073** 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 M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **11-09-2022**.

BOARD OF EXAMINERS

Chairman

Dr. S M Aminul Haque, PhD
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Ms. Most. Hasna Hena
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Md. Abbas Ali Khan
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Mohammad Shorif Uddin, PhD
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Dr. Md. Ismail Jabiullah**, Professor, 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:

M. Ismail Jabiullah
Professor Dr.
Department of CSE
Daffodil International University

Submitted by:

Moriom

Moriom Akhter
ID: 213-25-073
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.

I really grateful and wish our profound our indebtedness to **Supervisor Dr. Md. Ismail Jabiullah**, 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.

I would like to express our heartiest gratitude to **Prof. Dr. Touhid Bhuiyan** and Head**,** Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

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

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

Being totally dependent on agriculture, Bangladesh, detecting plant diseases can help farmers stimulate economic growth. This study explains how to recognize Apple leaf disease using deep learning algorithms. Deep learning algorithms can properly detect the defective leaf photos, which may enable the farmers, diagnose the leaf disease accurately and take urgent precautions in accordance with the disease. In order to classify each illness in our article, we first gathered the photos from the nursery and preprocessed them so that they fit the specific model that we employed. Later, in order to achieve a better result, we changed and applied a CNN model that was consistent with our dataset. By doing this, we can testify the diseased leaves 99.05% accurately.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In Bangladesh, many people are dependent on agriculture. Fruit brings great income to it. Among fruits, apples are best known as they are a good source of dietary fiber, vitamin C and various antioxidants. Apples are widely used for sick visits, morning breakfasts, and jams. China produces the majority of the world's apple supply, averaging over 60 million tons per year. More than half of America's crops are typically consumed as fresh fruit. One-fifth is used for apple butter, juice, jelly and vinegar. One sixth is canned as applesauce or pie stock. In Europe, the production of cider, wine and brandy accounts for the majority of the harvest. A quarter of the world's production is used for cider. These apples are grown by hardworking farmers. Every year, huge amounts of apples are wasted due to multiple diseases. Many farmer are unaware of apple disease and are unable to take preventative measures. If I knew why Apple's leaves suddenly malfunctioned, I'd fix it quickly. For these reasons, identifying apple leaf diseases in Bangladesh is very important as it helps farmers to identify their causes and find solution for each disease. There are different type of apple leaf disease, three of which are described. Here are his four classifications of images that he used to identify apple leaf disease detection:

Table1.1: Type of Apple Leaf Disease

Name of Genre	Description
Black rot	The fungus Diplodia seriata is the perpetrator behind black rot (syn Botryosphaeria obtusa). The fungus may infect living trunks, branches, leaves, and fruits in addition to dead tissue.

Scab	The fungus Venturia inaequalis, typically transmits by oocysts and spends the winter on dead leaves, is the source of the illness known as apple scab.
Apple Cedar Rust	The fungus Gymnosporangium juniperi-virginianae is the pathogen that causes cedar apple rust. Red cedars and apple trees are also impacted by this strange fungal disease.

1.2 Motivation

On our dataset, we applied six different models (ANN, RF, DT, GNB, LR, and SVM) to evaluate model performance. Among them, KNN gives the highest accuracy on the data set, GNB, DT model gives 83° accuracy. The second lowest accuracy is given by the SVM model. The RF model works well with an accuracy of 91.0 and LR shows an accuracy of 87.0.

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

1.3 Rationale of the Study

Considerable research has been done on apple leaf blight. A person who used a trained CNN model to identify apple leaf disease is described below, but the accuracy and results have yet to be determined. In this article, I would like to implement my own proposed CNN model to determine which image give better results and can better identify disease defect in each leaf. As mentioned earlier, different types of datasets have not been used so extensively before. Some have used pre-trained model to identify leaf diseases from images. However, in my case, I would like to develop my own type of the CNN model to improve it and give it better accuracy than other.

1.4 Research Questions

This work solves the problem of automatic apple leaf disease screening through the implementation of computer vision. The problems are:

• Is there a way to tweak the complex neural network for better result?

• Can I use popular deep learning technique like CNN and what parameter and hyperparameters can I change to improve my work?

1.5 Expected Outcome

• Deep learning method successfully recognizes image disease with higher rate.

• Can effectively detect diseases.

• Publish a research paper based on the results.

1.6 Report Layout

Chapter 1 In this section, we discussed the inspiration for our mission, the goal we set for ourselve, and the typical result of our efforts. Chapter 2 In this section, we discussed the basis of our research and looked at other related work, similar investigation, the extent of problems and difficulties, and the extent of problems and difficulty. Chapter 3 We discuss the research topic and the tools we used, and the data collection techniques, statistical analysis, and applicability of our results. Chapter 4 Here, we take the experimental and descriptive analysis results and find a summary. Chapter 5 In this chapter, we summarize the predictions and conclusions and also add more in-depth research.

CHAPTER 2

BACKGROUND

2.1 Introduction

Fundamental research includes an assessment of the research field, current information on the research topic, previous research on the topic, and relevant information on the topic. In this section, research should present as much background and background information about the topic as possible. My proposal is to review and identify obstacles to overcome, summarize research findings, and find solution to problems and challenges. Automated benefits and solutions should be discovered after thorough research. It has been proposed that different leaf diseases be identified according to the proposed defects, which is consistent with the proposal.

We will discuss our related effort, a description of the project, and any challenges we faced in conducting this research in this section of the website. Other research article related to our topic will be discussed and we will provide links to these publication in the related works section. We will then discuss the highlight of our project in the abstract and the challenges we encountered during our research in the challenges section.

2.2 Related Works

The identification of foliar diseases is well known and many foliar diseases have been identified to date. However, the discovery of apple leaves is still not well known through the use of deep learning. In their study, Kawcher Ahmed and Tasmia Rahman Shahidi [1] studied three different type of rice diseases. Their data set includes three different disease categories. To detect the disease by completing the necessary algorithm, they first used a white background for each photo. Then they apply a variety of machine learning algorithms. In this procedure, algorithm J48, KNN, Naive Bayes, logistic regression and decision tree were used. However, the decision tree algorithms outperform all of them with 10x cross-validation. From the test data set, it makes predictions that are accurate to 97%.

Nishkala & Bhavya [2] They recognized apple leaf disease and initially subdivided the images for this procedure. Then, to get the pixel distribution, they used k-means clustering algorithm. Furthermore, after distributing the pixels, they infer intelligent edge detection by comparing the change in pixel color and image weight. Finally, CNN was used to classify the disease. Although the accuracy and classification of the datasets were not clearly stated in this work, this approach appeared to be quite effective for furthering our investigation.

Beginning with article [3], R. Dr. Pauline Mounika Leaf disease is an area of expertise for Shayamala Bharathi. They used images of healthy and sick people in this essay. They divided the photos into several categories and grouped them at the start of the experiment using k-means to identify the images. By using SVM technique, their features are extracted after clustering. Disease detected will be displayed if the image output is a diseased image; otherwise the output will be a sane image. The accuracy with which they were able to recognize the photos, however, was not specifically addressed in this article. Therefore, many people are currently working on different types of leaf diseases, and Surampalli Ashok and Gemini Kishore have been working on tomato leaf diseases [4]. Similar methods have been used to determine the detection of foliar diseases, as reported in [1]. In a similar way, they segmented the photos, grouped them to determine their weight, and then used open-source algorithms to identify damaged panels. Article [5] They recognized apple leaf disease and they initially subdivided the images for this procedure. Then, to get the pixel distribution, they used k-means clustering algorithm. Furthermore, after distributing the pixels, they infer intelligent edge detection by comparing the change in pixel color and image weight. Finally, CNN was used to classify the disease. Although the accuracy and classification of the dataset are not clearly stated herein, this method seems to be quite effective for furthering our investigation. In our article, we will use CNN to identify sick image because out of all searches, CNN provides better results if it can be compatible with the dataset.

Table 2.1:Summary of Previous Researches

SL	Author	Methodology	Description	Outcome
1.	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 algorithm has been used for this research.	Decision tree with 10 fold cross validation achieved 97% accuracy on the dataset.
2.	Varsha J. Sawarkar, Dr. Seema Kawathekar	They have used deep learning and image processing methods.	The goal of this paper is Apple leaf disease identification by using the concepts of deep learning.	In this paper, they successfully identified the Apple leaf disease by using CNN.
3.	R. Mounika, Dr.p. Shayamala Bharathi	They have implemented image segmentation, K-means clustering , SVM.	After segmenting and cluster the images matlab has been used to identify the defects.	If the image is defected then the output will be diseased otherwise the output will be healthy.
4.	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.

5.	Aditya Rajbangshi, Toma sharkar	Transfer learning has been used here with mobileNet algorithm	Apple disease has been identified from the images by using pretrained model.	A satisfactory outcome has been achieved with this algorithm which is 95.63%.

2.3 Research Summary

Our research work provides the type of disease contained in each image. In addition, each disease type indicates its accuracy. Since the output metrics depend on 4 types of disease, there will be 4 output metrics, of which the one with the highest accuracy for each metric is selected as the disease image. For example, if the image contains holes, the output holes will be more accurate than other image classifications.

Using the CNN algorithm, we examined the accuracy of diseased leaves and created a confusion matrix to see how the results compare with each type of diseases. The higher the accuracy, the better the system.

2.4 Scope of the Problem

According to research, apple leaf disease has a huge impact on solving the problem of disease identification for the benefit of mankind to make life easier. By analyzing from previous research, we found that deep learning and automatic classification of leaf diseases by deep learning would be an proper way to solve the difficulty of foliar disease identification.

2.5 Challenges

The most difficult task for improving prediction accuracy is data collection. From Bangladeshi perspective, data collection is a long and complex activity. I need to collect data from an incubator, but it's very difficult to get permission because no one has given permission to take the data for privacy reasons. After collecting a very difficult dataset, the dataset needs to be preprocessed, which is also a tedious task. There was an imbalance between records, but that was another issue. So I manually balanced the record. Applying a CNN is another challenge, as the dataset has to be preprocessed according to the model. If the preprocessor doesn't match the model, the result will be bad. Improving the accuracy of hyperparameter tuning is also a major issue.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Improved deep learning algorithm have fascinated an entire era, helping us identify and separate images in different ways. Using complex neural networks to identify leaf disease can help predict image accurately. Before further elucidating my work, it is important to emphasize CNN. CNN is a deep learning algorithm, applied to identify image, time series data, objects or various aspects. It can also be used to classify different types of sounds. As mentioned earlier, Agglomeration Neural Networks can also be used to identify sounds, data, object, and even videos, so there is also some classification. Temporarily, the serial data transform is used in 2D and the 3D data transform should be used when working with images. Finally, an object used by many frameworks that contains frames for each interval. Therefore, I use this transformation in my work with 2D arrays. When trying to find an image, I explain using a 2D array. This will later be used as separate data as an array of images. Additionally, convolution transforms are used in complex neural networks. This section provides a step-by-step description of the algorithm and preparation of the dataset. First, the explanation of the research theme is discussed logically. In order to better conduct this research, we will elaborate on each of these steps. Data Collection and Preprocessing describes how data collection and data processing are performed. Then, in the proposed method, I describe my algorithm and the method used to determine the genus for taxonomy. Therefore, working charts and methods as well as some visualizations are provided in the statistical analysis section. Finally the chapter on how I implemented my project is closed.

3.2 Research Subject and Instrumentation

You will find that information is the most important part of the exam. Finding good data and good methods or models for experts is a very important aspect of our research work. You should also consider comparable past exam questions. At this point, you have to decide between several option.

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

3.3 Data Collection Procedure

The hardest part of our thesis was collecting the dataset. For security and health reasons, no one has given us permission to collect the data set. In this procedure there is a variation of the disease, but many diseases cannot be included because there is not enough disease picture. We had to move a lot to collect all the pictures. In this process, we first take pictures with a digital camera and then store them. The images are not so perfect due to the lighting, so we need to preprocess them to be able to use them in our algorithm. The preprocessing method will be explained later in the next section.

3.3.1 Attributes

The resulting dataset contain 4 attributes. The main attribute are (black rot, apple scales and cedar apple rust). The training dataset and the test dataset have been split into two separate parts, each containing the data set. One of the many dataset used for training is also devoted to validation.

Some sample image of the dataset are illustrated below:

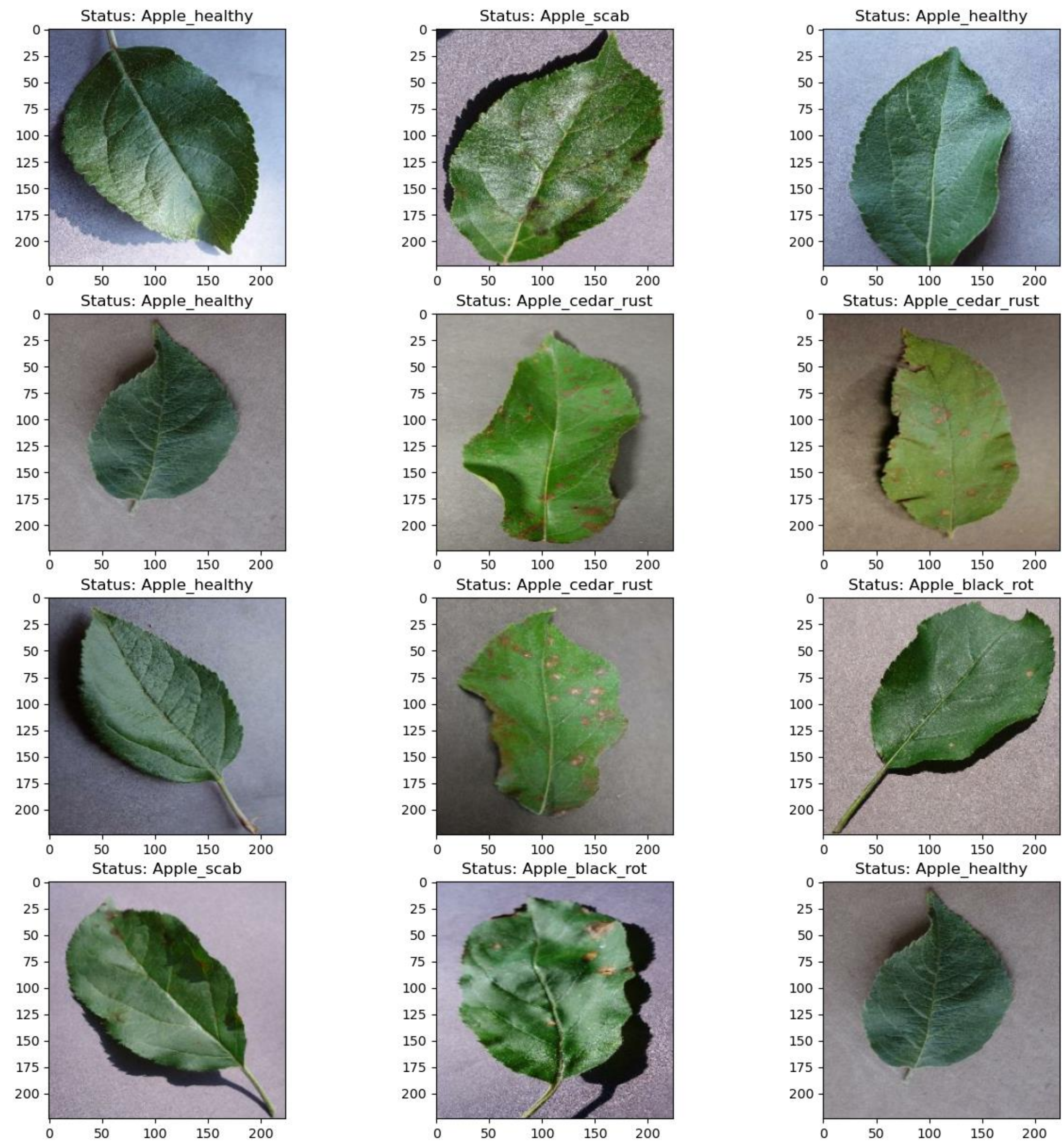


Figure 3.1 Samples of our dataset

These three type of diseased images are used in our dataset.

3.3.2 Data Pre-processing

In this process, firstly we manually crop only the defected image 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 3.2: Dataset labeling

Then we have labelled each classes. Black Rot has been labeled as ‘1’ and scab has been labeled as ‘2’ ceder rust has been labeled as ‘3’ and healthy leaf has been labeled as ‘4’. Some random labelling has been illustrated in the figure 3.3.

In this dataset, the image is (224,224.3) pixels. 3 indicates that this image is an RGB image and not a grayscale image. Also, converting this image to an array results in three 224*224 matrices. We know that when dealing with grayscale images, we are left with only one matrix, but in this case each matrix contains a number of algorithms that will later be used to identify changes in each image that will be converted to 224*224 pixels. contains a huge value that needs to be accumulated in .

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-folder in the test and train folder like Black Rot, Apple Scab, and Apple Cedar Rust.

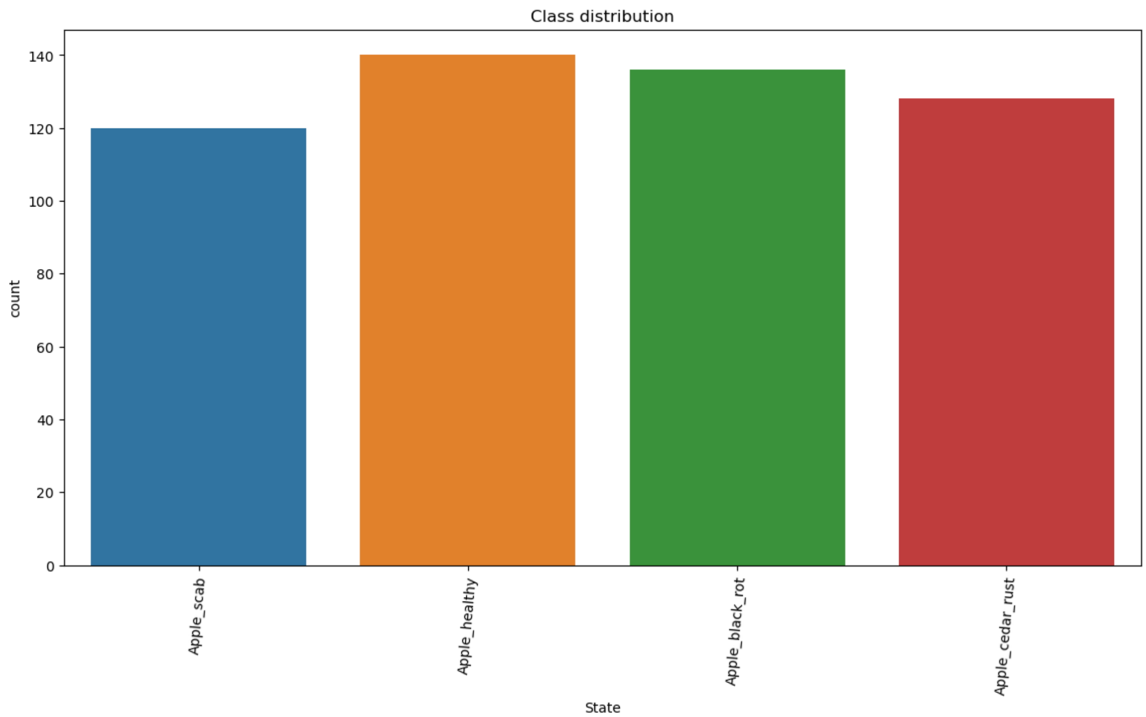


Figure 3.3: Apple Leaf Disease Class

3.3.4 Data Storing

This is where you store all your Kaggle data when using Kaggle Notebook. Saved data from Kaggle allows you to integrate saved data into your current project. This means the output will be saved as a .h5 file. This can be used in other applications as well as online applications. In the future, this data may help project work by providing clear step-by-step instructions.

3.3.5 Deep Learning Algorithms

We modified the CNN model in the paper to get better results. The process of customizing and modifying hyperparameters is described in the next section. All algorithms are applied to Python and libraries are NumPy, Panda, Scikit-learning and Matplotlib.

A CNN (ConvNet/CNN) called Modified Convolutional Neural Network CNN (ConvNet/CNN), a deep learning algorithm, takes images as input, uses them to understand the weights, and computes them for individual parts of the image. or increase them. distinguish from each other. The amount of data preprocessing required by ConvNets is increasingly less straightforward than that required by alternative classification techniques. Unlike traditional neural networks that learn to distinguish features over time, ConvNets can learn images from attributes and distinguish them from each other in the future.

A complex neural network was used on the image dataset because it can identify images by extracting and separating features. Complex neural networks have multiple functions such as cores, feeds, pooling, tasks, filters, and activation functions. Everything has its own significance in building a perfect neural network.

First, image recognition processing is performed by edge detection. Edges can be vertical or horizontal.

Once I get the image implied output through the use of a filter, I attach bias to those result and then use the activation function to calculate the number of activation. Then the single layer of the CNN equation is as follows:

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

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

In this scenario, the input image is 224*224*3 and the filter is (4*4*4), corresponding to the weight w^([1]). This is the first layer of CNN. In the second layer, the layer 1 activation a^([1]) serves as the layer 2 activation. In this sense the number of parameter depends on the size of the image. My CNN hierarchy consists of three types.

- Convolution layer

- Pooling layer
- Fully connected layer

The modified CNN has 4 layer of the bidirectional convolutional layer used and in each layer the maximum layers are available. These layers are commonly used to extract features from images. It reduce the input size of the image and increases the calculation speed. For example, if we imagine a 4*4 pixel image, applying maximum aggregation will worsen the result of the matrix as follows:

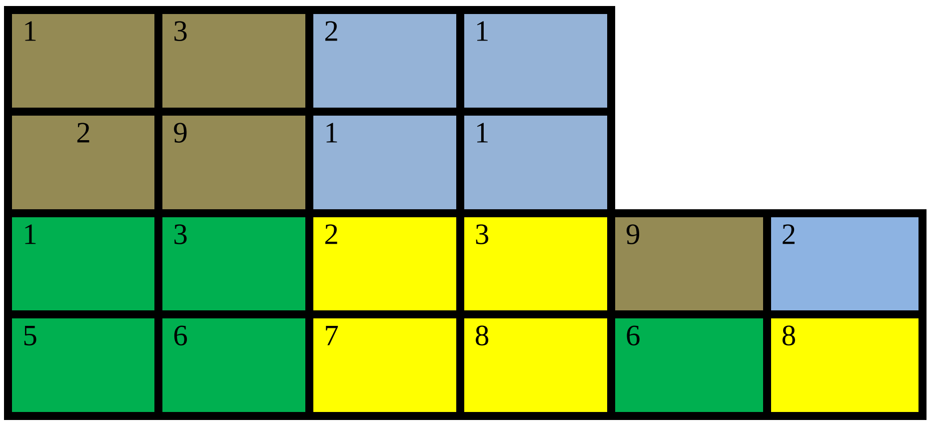


Figure 3.4: Maxpooling from matrix

In Figure 3.5.2, the illustration is made for 2 * 2 pooling where for every 4 pixels I subtract the amount and return 1 pixel. In addition, in the 2*2 model, our maximum sharing was used to help determine the characteristics of the image. Maxpooling actually works by taking the largest value in 2 * 2 pixels. In addition, other pooling methods are also available, such as average pooling, where the mean of 2 * 2 strides will be taken.

There are many hyperparameters and parametrics available in accumulative neural networks. Among them, core size really matters. Deep Neural Networks, or CNN as it is more specifically called, is basically a stack of layers formed by the action of several filters on the input signal. These filters are often referred to as kernels. In our proposed model, a kernel size of 3*3 was used, which means that in an image the layer stack will be built on top of a 3*3 kernel and after each computation is completed kernel in the complex neural network, it will pass to the next kernel and so on. This is how a full stack of classes 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 3.5: Kernel Size

Each model has additional activation features. Enumeration of the next layer requires each activation function. In my proposed CNN, relu or rectified linear units were used as activation functions. relu is used as the default neural network activation function because it is fast and performs well. The activation function is high performance, so I haven't changed it.

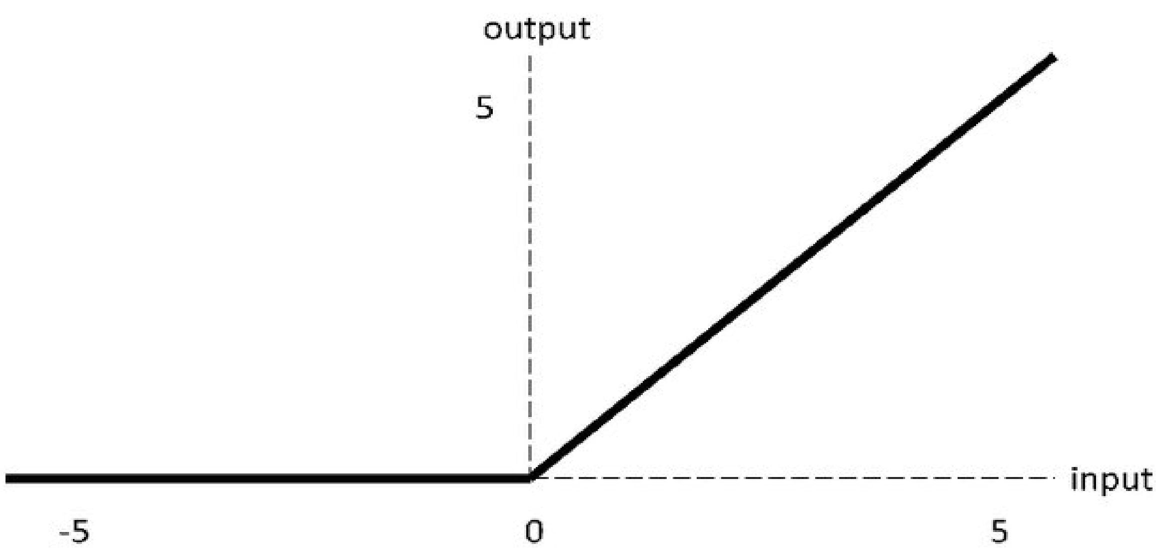


Figure 3.6: Rectified Linear Unit

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

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

In particular, the difficulty in this scenario is that all negative value are immediately reset to zero, which reduces the likelihood of an appropriate fit or training of the model from the dataset. In other words, any negative input given to the ReLU trigger function will immediately be converted to 0 in the graph, which has a negative effect on the resulting graph as negative values are not detected. correctly maps to a positive value. That's why in the last layer we use the sigmoid function so that negative values can also be measured to get the expected result as shown in the image below. Regarding training in each class, relu was used, so the negative value does not affect much.

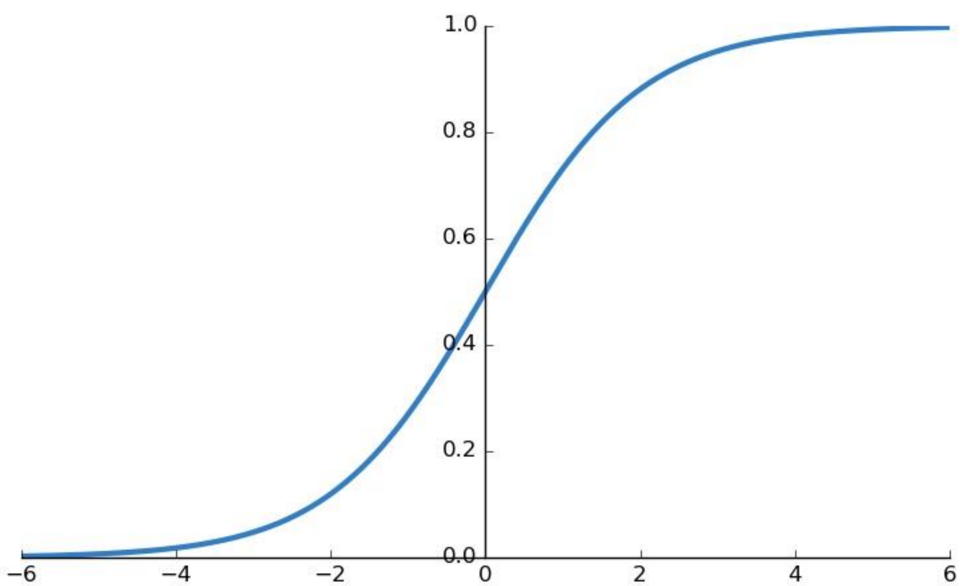


Figure 3.7: Sigmoid function

However, since the sigmoid function returns values between zero and positive 1 in the output layer, we used the sigmoid function to keep the output accurate. The main reason for using the sigmoid function in the output layer is that the sigmoid function always returns the value (0,1). The formula for the 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 CNN structure is compiled and trained on the training dataset with a learning rate of 0.001. Also, 20% dropouts were added to my CNN recommendation. Typically, omitting is used to reduce model errors by eliminating spurious values and focusing more on key features.

3.4 Statistical Analysis

We used 1430 image in our dataset and of those 1192 were used in the train image and 238 were used in the test image. We have accumulated 20% of train image in our test dataset. In Figure 3.5, the number of each image layer is shown.

Table 3.1: Dataset Overview

Overall Data	Train	Test	Resolution
701	524	177	224*224*3

To apply deep learning algorithm, we have preprocessed the images by classifying them into different folders, not only that we chose lower pixels because the defects on the image are almost nonexistent. can be seen. So we cropped them first, then we had to use the lower pixel. On the other hand, if we choose full resolution images, it may give us less accurate results when applying the algorithm.

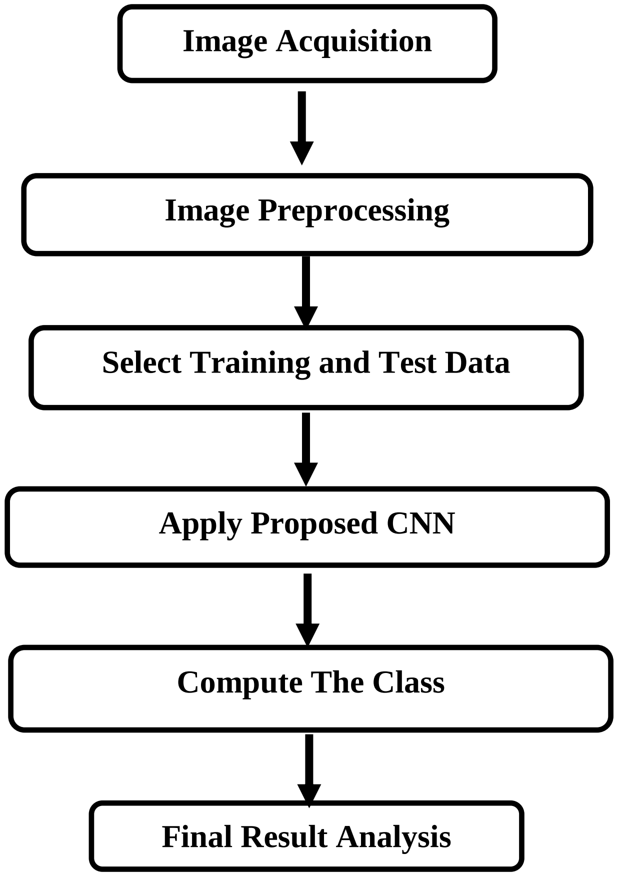


Figure 3.8: Proposed Model Structure

This diagram shows how we briefly investigated. In this figure (3.6) you can see how to reach the goal step by step.

3.5 Implementation Requirements

Realizing my work require a lot of machine power because there are a lot of image. Deep learning algorithm are widely used in machines, but because they require high hardware power, the outputs are also attractive compared to other machine learning algorithm. To improve the skill of the model, I need to use many parameter to train the dataset and the scaling of these parameters is the only reason for such high end hardware requirement. When I run the model, at the time of training the datasets, the dataset are first stored in volatile memory (RAM) to be used later for training, then when computing matrix takes place, GPU or CPU is needed for the calculation in my opinion. CNN suggested. Increasing

the dataset will need more RAM and high-end GPUs. Furthermore, increasing image diversity as well as improving image quality are related to hardware requirements.
After a thorough examination of relevant statistical or theoretical concepts and approaches, our team developed a list of hardware, software and development tools needed to predict leaf disease. In most cases you will need:

Hardware/Software Requirements

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

Developing Tool

- Python 3.6 or more
- Anaconda

CHAPTER 4

EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experimental Setup

In the following, an explanation of my CNN model and how it work along with a description of the dataset has been provided. In this section, our experimental results will be provided. I'll also shed light on the analysis we've gathered so far while working on thisproject.

The final model is created for a specific type of image then trained in a notebook to achieve the best results. In each image, the features are very variable, so it is difficult to divide them into different layers. I take turns doing this by increasing the neural layers and decreasing the number of steps per pixel.

- As we have worked with the prediction of leaf disease we have collected diseased leaf dataset.

- Most of our time expenditure on collecting data as it was the most tedious work

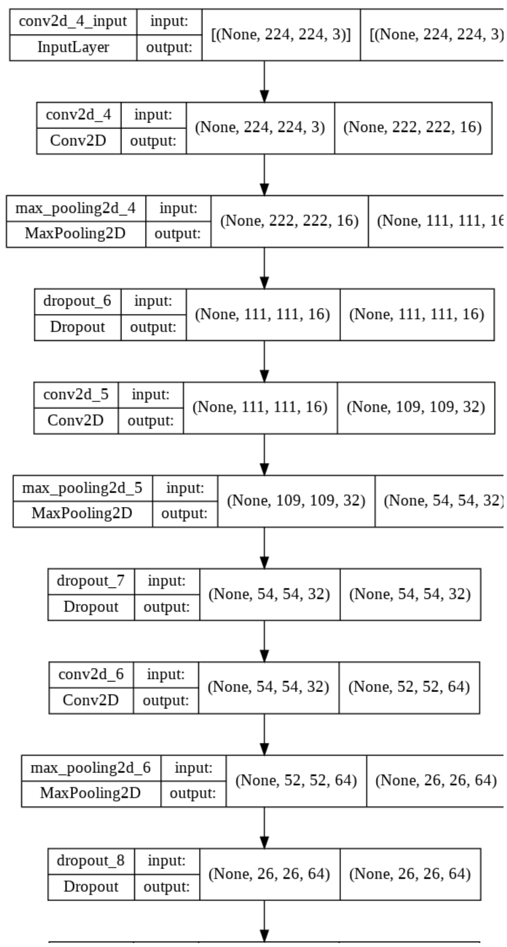
- We have likewise gathered data from a nursery as described before.

- 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 algorithm.

4.2 Model Summary

For disease identification, different types of algorithms can be used. In the past, many researcher have implemented a pre-training model on transfer learning techniques. However, in my thesis work, I will apply my own CNN proposal by refining the CNN model to get the best results. In this approach, I used an accumulative neural network with 51,68,676 parameters. In this proposed CNN model, I used 4 layers of a complex neural network where maximization is used in each layer with 20% skips. In the next section, CNN will be briefly explained with appropriate equation and example and how CNN has been used according to my data set will also be explained.



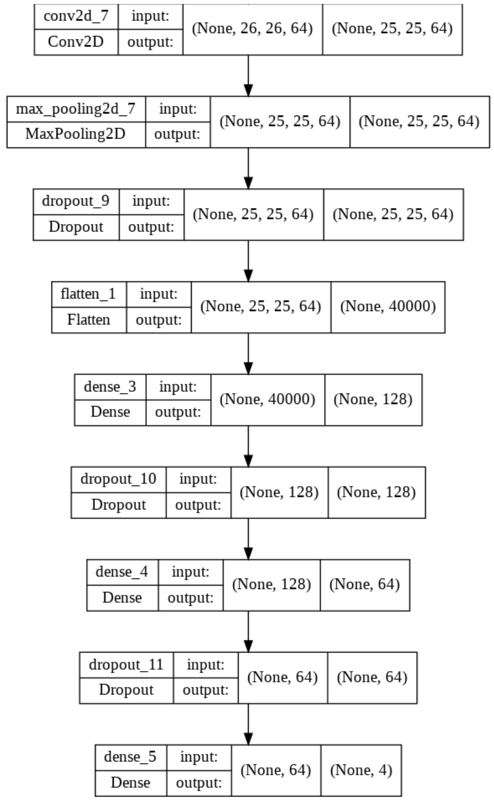


Figure 4.1 Model Visualization

4.3 Experimental Result and Analysis

A modified CNN model was used to predict leaf disease with the best results. We got 99.05% accuracy and got this result after 100 epochs. Therefore my prediction was satisfactory for the modified CNN

Table 4.1:Experimental Result

Model	Train Accuracy	Test Accuracy	Precision Score	Recall core	Train Loss	Test Loss
CNN	99.05%	87.01%	87.01%	87.01%	0.013%	21%

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

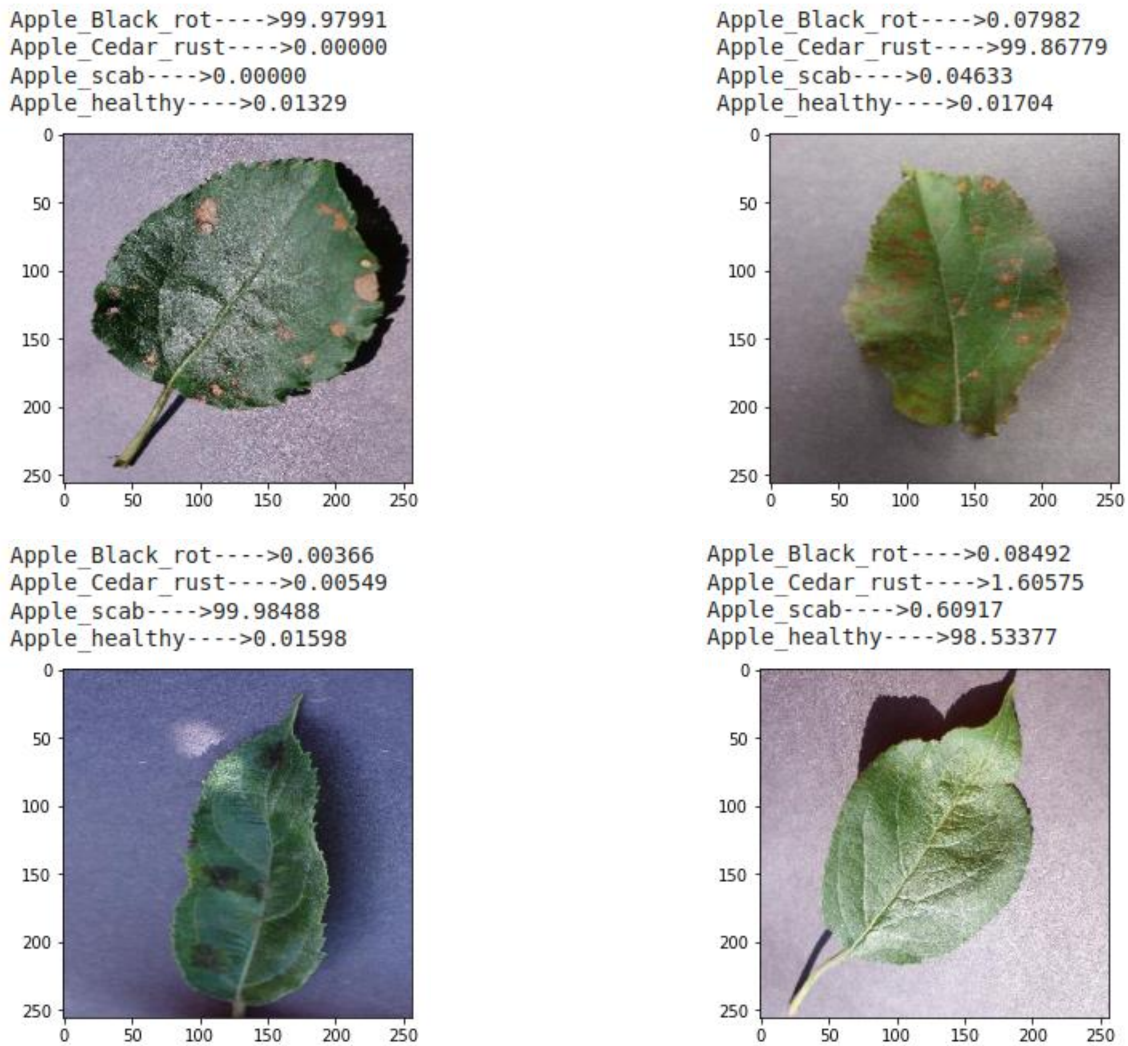


Figure 4.2: Sample Output

In the given images it showed 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

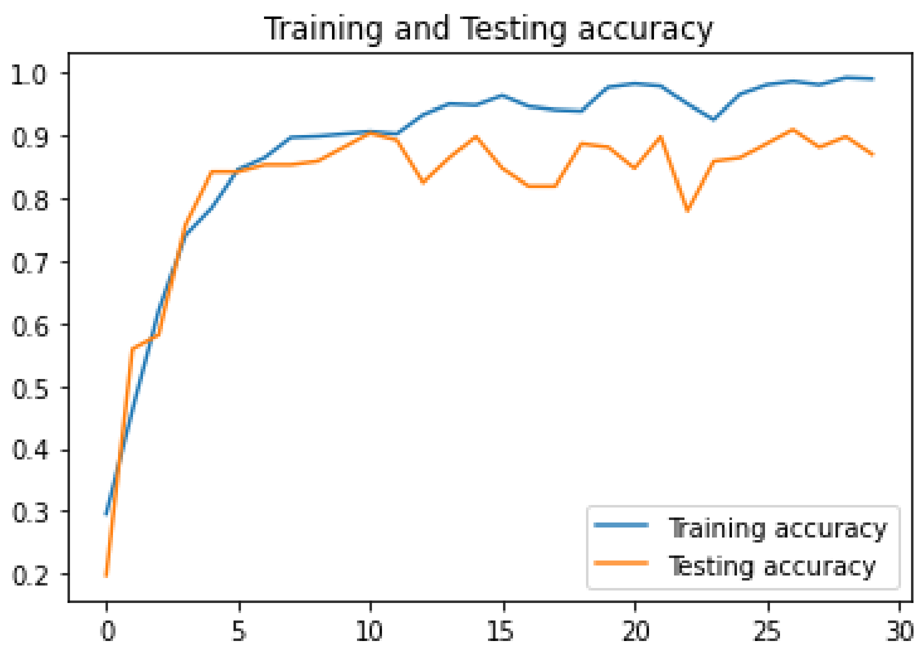


Figure 4.4: Accuracy graph

The confusion matrix is given below:

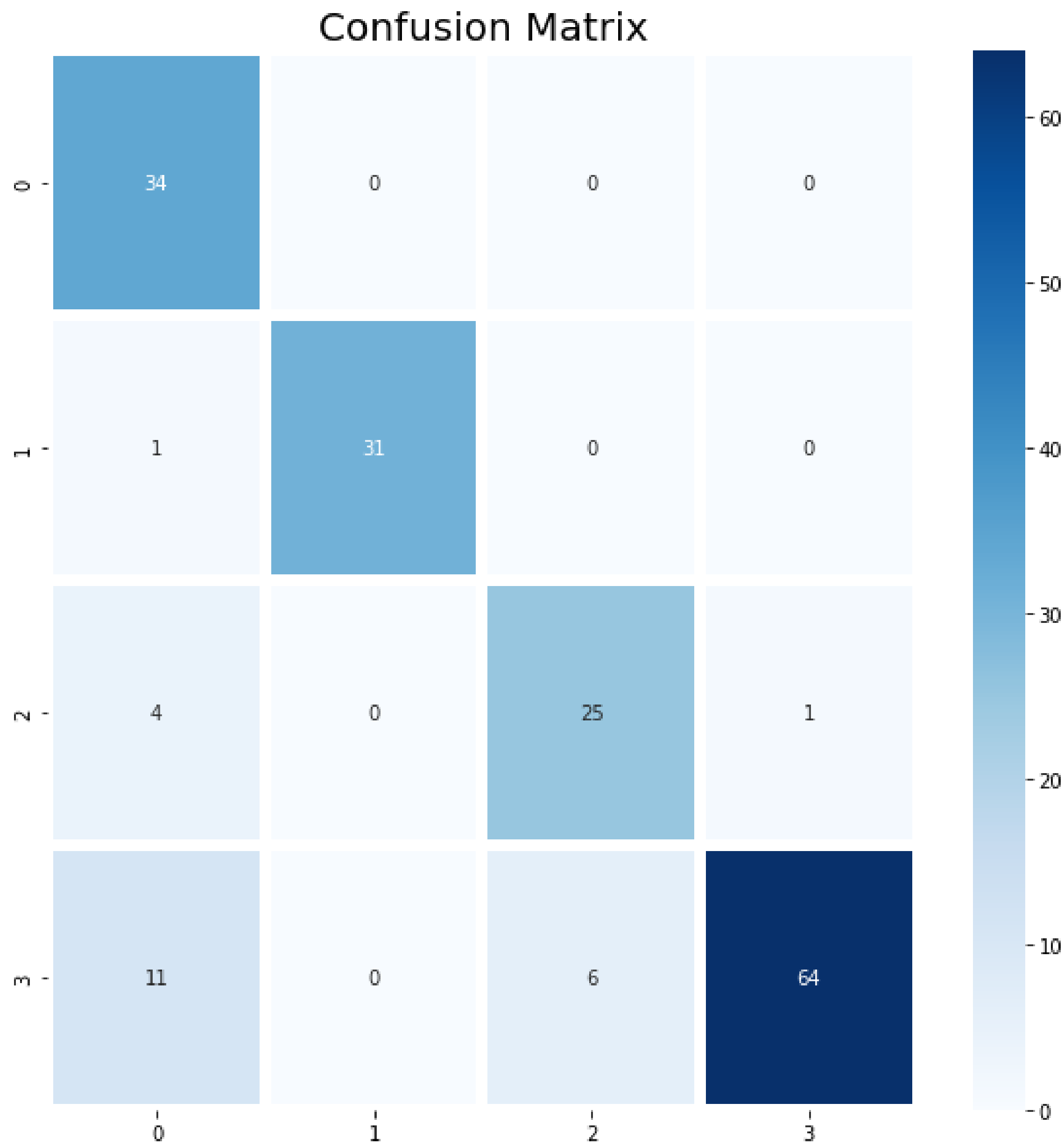


Figure 4.5: Confusion Matrix

4.4 Discussion

An example of a 4-level image is a dataset prepared for CNN. We preprocessed and enriched the data set so that the CNN function works well. For each category, we added different parameters and took the necessary steps to allow the model to easily distinguish between images at these four levels. Each image in the dataset has many feature that are difficult to identify. Kernel size and padding, as well as all neural network configurations, provide a well-defined data set for machine learning and image classification.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION, IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study

We built this research business to predict apple leaf diseases, which will help farmers improve yields and profits even further. The study proved that our modified CNN provided the best results and it successfully predicted the faulty leaf images. We collected four type of disease image in our study and preprocessed them. We then carefully applied the data preprocessing decisions with the aim of making them ideal for the applied algorithm. Data collections must be prepared in such a way that they can reasonably respond to data. In our studies, one part of the data collection is prepared while the other is tested.

5.2 Conclusion

My main goal is to develop a model that, compared to other models, can help identify diseased leaves. To predict the images, I used my own CNN model modification as a deep learning method I developed myself. This search revealed that my test dataset had an accuracy of 99.05%, indicating that most of the images in the test dataset can be correctly identified. It looks like my model with an effective tagging solution has a promising future and deserves further study. The results of this study will help farmers better identify leaf diseases. They can then use the research to look at leaf diseases and recommend treatments even if they weren't aware of any specific apple leaf disease at the time of the initial study.

5.3 Recommendations

There are certain recommendations for identifying this disease, such as:

- Improved images resolution can also yield better results
- Try to clean the data sets and enhance them properly.
- Larger datasets can provide more accuracy.

5.4 Implication for Further Study

Few Implication 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 user.

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, "Apple Leaf Disease Detection using Digital Image Processing &Deep Learning”,International Journal of Emerging Trends & Technology in Computer Science, June,2021.

[2] Aditya Rajbongshi,TomaSharkar, "Apple 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, "DetectionofPlant Diseases Using Image Processing”,Journal Critical Reviews, 2020.

[5] Kawcher Ahmed; Tasmia Rahman Shahidi; Syed Md. IrfanulAlam; SifatMomen, "Rice Leaf Disease Detection Using Machine Learning Techniques", IEEE Xplore,2019.

[6] MD. Kamran Hosain, Harun-ur- Rashid, TasnovaBinteeTahar, MohamamdMasudur 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 recognation with grape convolutional networks" CVF Computer vision foundtion , CVPR-2019, IEEE-2019.

[8] Karim and Zaccone (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; MiranPobar; 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, EgmenSahin, canjozefsaul, 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, satbirjain, "Automated Genre Classification of books Using Machine Learning and natural Language Processing", IEEE Xplore,2019.

Apple leaf

ORIGINALITY REPORT

9%	9%	1%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	9%
2	M Anandkumar. "Texton Features and Deep Belief Network for Leaf Disease Classification", Multimedia Research, 2020 Publication	<1%
3	e-journal.uajy.ac.id Internet Source	<1%
4	Shupe, S.M.. "Cover- and density-based vegetation classifications of the Sonoran Desert using Landsat TM and ERS-1 SAR imagery", Remote Sensing of Environment, 20041030 Publication	<1%
5	www.vorsers.com Internet Source	<1%