LEMON DISEASE CLASSIFICATION USING CNN-BASED ARCHITECTURES

BY

Mst. Fatematuz Zohura
ID: 201-15-3195

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

Supervised By

Sharmin Akter

Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Md. Sabab Zulfiker

Senior Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

23 JANUARY 2024

APPROVAL

This Project titled "**lemon disease classification using cnn-based architectures**", submitted by Mst.Fatematuz Zohura, ID:201-15-3195 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 23/01/2024.

BOARD OF EXAMINERS

____

Dr. Sheak Rashed Haider Noori (SRH) **Chairman**
Professor & Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Afjal Hossan Sarower (AHS) **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Amir Sohel (ARS) **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

____

Dr. Md. Arshad Ali (DAA) **External Examiner**
Professor
Department of Computer Science and Engineering
Hajee Mohammad Danesh Science and Technology University

Declaration

We hereby declare that, this project has been done by us under the supervision of **Sharmin Akter, 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.

Supervised by:

____

Sharmin Akter
Lecturer (Senior Scale)
Department of CSE
Daffodil International University

Co-Supervised by:

____

Md. Sabab Zulfiker
Sr. Lecturer
Department of CSE
Daffodil International University

Submitted by:

Fatema
____

Mst.Fatematuz Zohura
ID: 201-15-3195
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

At first, I want to express my heartiest thanks and gratefulness to Almighty Allah for His divine blessing makes us possible to complete the final year project successfully.

I am really grateful and wish our profound our indebtedness to **Sharmin Akter**, **Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Machine Learning interest of our supervisor in the field of AI 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 my heartiest gratitude to the Almighty Allah and thanks to 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, we must acknowledge with due respect the constant support and patients of our parents.

Abstract

The focus of my study, "Lemon disease classification using CNN-based architecture," is image-based disease detection. Significant progress has been achieved in image processing, and artificial intelligence and its uses are being applied in many design fields. Humanity has already stepped into the digital age. We use sophisticated cameras to take pictures, and the better, more useful, and more productive the outputs and consequences are, the clearer the pictures are. I utilized fresh, totally infected, and infested lemons for this study. In addition to employing neural network (CNN) tools for AI to analyze the findings, I am creating my groupings of features using the VGG16, MobileNetV2, NASNetMobile, ResNet152, EfficientNetV2, and DenseNet201 models. Applications primarily utilizing growth modeling in agricultural production and lemon agronomic research can benefit from the knowledge of lemon features. The direct measuring techniques used in the past were often labor and time-intensive, basic, and not very dependable. Effective techniques for seeing and recognizing exterior illnesses and other aspects provide the foundation of recommended vision. As of right now, image processing algorithms are evolving quickly to identify and recognize certain color frames of afflicted areas in order to diagnose illnesses. Ultimately, the afflicted region is isolated from the picture. The diseases that impact fruit production, such as citrus canker disease, citrus rust mite disease, and delicious lemons, were then recognized from pictures of lemons. The results of the VGG16, MobileNetV2, NASNetMobile, ResNet152, EfficientNetV2, and DenseNet201 models in my example demonstrate that this autonomous vision-based system performs admirably, with the highest results in VGG16 coming in at 91%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Citrus crops, with their economic and nutritional significance, play a vital role in global agriculture. Among citrus fruits, lemons (Citrus limon) stand out not only for their culinary applications but also for their versatility in various industries. However, the health of lemon crops is frequently threatened by the emergence of diseases, posing a significant challenge to agricultural productivity and sustainability.

The impact of diseases on citrus crops is multifaceted, affecting both yield and fruit quality. Recognizing the signs of disease early on is crucial for implementing effective disease management strategies. Traditional methods of disease identification often rely on visual inspection by experts, making the process time-consuming and subjective. In light of these challenges, there is a pressing need for automated systems that can rapidly and accurately classify citrus diseases.

This research focuses on leveraging Convolutional Neural Network (CNN)-based architectures for the automated classification of lemon diseases. CNNs, a subset of deep learning models, have demonstrated exceptional capabilities in image recognition tasks. By enabling the automatic extraction and interpretation of intricate features within images, CNNs offer a promising avenue for the development of robust and efficient disease classification systems.

The primary objective of this study is to contribute to the advancement of automated agricultural disease detection, specifically targeting lemon diseases. By harnessing the power of CNNs, we aim to develop a model capable of distinguishing between healthy lemons and lemons affected by various diseases. This research addresses the need for accurate and timely disease identification, which is paramount for implementing targeted intervention strategies to safeguard citrus crops.

In this report, we detail our methodology, including data collection, pre-processing, and the design of the CNN architecture. The study's findings, presented in the results and discussion sections, provide insights into the performance of the classification model and its potential implications for the field of citrus crop management. Through this research, we contribute to the ongoing efforts to enhance agricultural practices and mitigate the impact of diseases on lemon crops.

1.2 Motivation

The global agricultural landscape faces an increasing challenge to ensure food security and sustainability. Citrus crops, including lemons, contribute significantly to nutrition and economies worldwide. The devastating impact of diseases on citrus crops poses a threat to food production, livelihoods, and economies. Addressing this challenge requires innovative solutions that can efficiently detect and manage diseases, ensuring the resilience of citrus cultivation on a global scale.

Automated disease classification using advanced technologies, such as CNN-based architectures, holds promise for revolutionizing agricultural practices. By providing timely and accurate identification of diseases, these technologies can empower farmers and policymakers with actionable insights.

In the context of Bangladesh, agriculture is a cornerstone of the economy, providing livelihoods for a significant portion of the population. Citrus cultivation, including lemons, contributes substantially to both domestic consumption and export markets. However, the agricultural sector faces various challenges, including the impact of climate change and the emergence of plant diseases.

In Bangladesh, where agriculture is a vital component of the national economy, the motivation for developing advanced disease classification models is particularly poignant. Lemon diseases not only jeopardize the income of local farmers but also influence the availability of nutritious produce for the population. Implementing automated and accurate disease detection systems can significantly benefit Bangladesh by enhancing crop management practices, optimizing resource utilization, and contributing to sustainable agricultural development.

Furthermore, advancements in agricultural technology align with Bangladesh's commitment to the United Nations Sustainable Development Goals (SDGs). Leveraging innovative solutions for crop protection reflects the country's dedication to achieving food security, promoting economic growth, and ensuring environmental sustainability.The motivation for this research extends beyond individual countries to address global challenges in agriculture. For the world, it represents a step towards securing the future of food production, while for Bangladesh, it offers a targeted and impactful approach to fortify a crucial sector of its economy and contribute to broader sustainable development objectives.

1.3 Problem Definition

I am trying to build an accurate model for lemon disease from a different angle of view. In my model, the main goal is to detect lemon diseases and increase production. Prepared a model, which is lemon To detect citrus canker disease and citrus rust mite disease, we selected the image classification technique using the CNN deep learning architecture.

1.4 Research Questions

The following are the primary questions addressed in this paper:

RQ1.Why do we require image classification to identify lemon disease?

RQ2.Which learning approach is most suited to picture classification? How do we recognize it?

RQ3.Why is CNN superior to other image categorization algorithms?

RQ1: Through this, we can easily find the diseased lemons and take appropriate measures so that the production of lemons is good. RQ2: Convolutional Neural Networks are distinguished by their architecture, which depends on big labeled datasets and makes them efficient for tasks involving images. RQ3: In my case, the VGG-16 algorithm is the most advanced among the various algorithms in CNN.

1.5 Expected Outcome

This research-based initiative is expected to produce the following results:

1.With 91% accuracy, detect illness and fresh lemons.

2.Learn and train the set to its full potential.

3.A graph displaying the affected level will be shown.

4.The projected disease kind or healthy plant species will be output here.

1.6 Report Layout

Chapter 1:In this section I can look at the introduction, motivation and goals and talk about what results are expected. In a nutshell Part 1 is a description of the development of this project.

Chapter 2:In this section I have discussed the various works of past days related to the issue and discussed the position of Bangladesh.

Chapter 3:This section discusses the research topic and materials discussion, information spectrum techniques, practical investigation, and implementation requirements.

Chapter 4:Discussion Results Empirical and descriptive results are presented.

Chapter 5:In this part the results are analyzed.

Chapter 6:This part discusses the future planning of the problem.

CHAPTER 2

REASERCH BACKGROUND

2.1 Introduction

Everywhere we want to touch everything, technology makes our life so much faster and easier. We will explore relevant research or initiatives on deep learning and data categorization in this chapter. The full technique for developing the model for plant recognition using deep CNN is given in detail. The complete procedure is broken into a few critical stages in the following sub-sections, beginning with images of social events for categorization using deep neural networks. We will review prior relevant work in the first section, then display the findings or a summary of my research of the related work in the second section, and then analyze the project's advantages and obstacles.

2.2 Related Works

Citrus Plant Disease Detection Using Genetic Algorithms, initial pictures of sample leaves are obtained and separated into numerous portions. The genetic algorithm is then utilized to choose the optimal solution for diagnosing and determining the color and textural features of the leaves[1].Citrus canker is a bacterial disease that causes significant damage to citrus trees [2].Citrus, a plant with several medical and therapeutic properties, has long been used in herbal medicine due to its low side effects and versatility. This article [3]provides information on the medicinal and health benefits of lemon. In a research done in Sylhet, Bangladesh, citrus diseases such as greening, die-back, scab, and canker were discovered to be frequent among 17 varieties. Rough lemon was the most sensitive type, while die-back and scab were the most common illnesses. China-lemon had the lowest overall sickness incidence. However, the environment in Bangladesh promotes disease growth while also promoting year-round citrus production. Citrus canker, scab, gummosis, dieback, Citrus Tristeza Virus, and greening can all injure citrus trees. The purpose of this study is to investigate the disease prevalence and disease index in citrus orchards in the Sylhet district[4]. By enhancing images of citrus canker-infected leaves with RSWHE and a median filter, the suggested approach improves visual imaging and disease diagnosis. Future advances may involve segmentation of sickness detection[5].Zahid Iqbal estimates that 50% of citrus peel produced each year is damaged or lost. Every year, a significant quantity of waste is created in the citrus industry. They have image pre-processing, fragmentation, feature extraction, and selection techniques, as well as stratification. Photographs of their study findings are used to isolate plant components for further processing. Following the preceding phase, segmentation entails classifying the provided data into distinct categories. Depending on whose capability is selected[6]. It has the ability to reduce the requirement for transfer learning. A neural variation of this is the large number of training samples that Networks often require, which will assist such models in overcoming the technical limitations caused by utilizing a limited data set[7]. This study proposes a CNN-based approach for identifying plant diseases, with an emphasis on food and cash crops. The model is applied to 15 distinct cases, including 12 ill leaves and three healthy leaves. The model's accuracy is 88.80%, and several performance matrices are generated to identify illnesses. This technique has the potential to increase agricultural productivity while also ensuring the Indian economy's environmental sustainability[8]. Deep Learning advancements have revealed the promise of Automatic Image Recognition systems based on Convolutional Neural Networks. Transfer Learning was used to construct a small dataset, and the accuracy was 92.46%[9]. The study paper's work focuses on the many methods utilized for lemon leaf disease detection and categorization. The authors contrasted the methods of illness diagnosis using convolutional neural networks (CNN) and support vector machines (SVM). The illnesses were accurately classified using the SVM method. These findings show that CNN fared better in properly identifying and categorizing lemon leaf illnesses than SVM[10]. A article titled "Detection of plant leaf diseases using image segmentation and soft computing techniques" was published by Vijai Singh and A.K. Misra. In order to automate the detection and treatment of plant leaf diseases, this research presents a computation for the image division process[11]. Agricultural productivity depends on the identification of plant diseases, yet the available techniques are expensive and time-consuming. In order to detect plant illnesses, this research suggests a Disease Recognition Model that combines convolutional neural networks (CNNs) with picture categorization. Large-scale agricultural losses can be avoided by detecting plant illnesses early on. Deep learning and neural network-based automated disease detection techniques show promise in this regard. The use of artificial neural networks and machine learning algorithms for plant disease identification has been investigated in a number of research. The suggested CNN model seeks to distinguish between diseased and healthy leaves as well as identify plant ailments[12]. Using deep learning, Prasanna Mohanty and associates created a deep convolutional neural network that can identify 26 diseases and 14 distinct crops. The training set model achieved an accuracy of 99.35 percent on a held-out test set, demonstrating the usefulness of this approach.

When evaluated on a set of photos obtained from reliable websites, taken in conditions different from those used for training, the model still achieves 31.4 percent accuracy. Even though this accuracy is significantly higher than the 2.6% based on random selection, more training data must be collected in order to improve accuracy overall[13]. Crop infections can lead to crop mortality and food loss, posing a serious danger to crop health and productivity. The development of deep learning, computer vision, and smartphone usage has made detection possible. The effectiveness of deep learning-based methods for identifying plant diseases in practical settings is investigated in this study. In a collection of crop leaf photos, 38 illnesses affecting 14 different crops were categorized using a sequential architecture. With a 95% accuracy rate, a basic convolutional neural network helped novice researchers with deep learning applications[14]. Despite advances in technology and genetics, famines, and plant diseases, many suffer. Artificial intelligence and deep learning can help farmers detect and classify plant diseases, improving yield and quality. This study uses the Plant Village database and transfer learning technology[15]. In order to detect lemon plant diseases, this research proposes a deep learning (DL) method that makes use of traditional neural networks such as Google Net, Res-Net, and squeeze Net. Six09 photos in the dataset—which was gathered from Mendeley data are grouped according to prevalent illnesses that affect lemon plants[16]. This study suggests a method for identifying cotton leaf illnesses using visual cues to mitigate the financial losses incurred by India's cotton farming sector. The expert system classifies plant leaves and detects diseased and healthy leaves using overlapping pooling, multi-layered perceptron, graph-based MLP network models, k-nearest neighbor, and support vector machine algorithms[17].

Table 2.1: Comparison with previous related work

Reference	Main Task	Used Algorithm	Best Model	Accuracy
[18]	Plant disease identification	CNN	CNN	75.00%
[19]	Detector for Real-Time Tomato Plant Diseases and Pests Recognition	CNN(VGG-16, ResNet-50)	ResNet-50	83.00%
[20]	Disease detection scheme for Citrus fruits	CNN	CNN	89.1%
[21]	Citrus Diseases Recognition	CNN	CNN	88%
[22]	Examination of Lemon Bruising	CNN(ResNet, DenseNet, ShuffleNet, MobileNet)	ResNet	90.47%

2.3 Bangladesh Situation

We know that lemon is planted in Bangladesh, and its cultivation is expanding because it is profitable. The Bangladesh Export Development Bureau reports that during the most recent fiscal year 2021–2022, the country exported citrus fruits valued at 17.35 million dollars and citrus peels valued at twenty thousand dollars.Most farmers in our nation are uneducated or illiterate, making it difficult for them to identify lemon illness. If they do not recognize lemon illnesses at an early stage, it will have a significant influence on lemon cultivation and cause them to lose enthusiasm. To address this issue, we created a model that allows farmers to detect lemon infections and take appropriate action. We collect many sorts of lemons. Any data may be used to test our model. CNN is used in our model.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This part will look at and detail how to identify, measure, and investigate data utilizing the research approach of my examination technique. A talk of the research project's apparatuses, information gathering, the research subject, pre-handling and preparation, factual examination, and execution will also take place. We were using CNN with the feature extraction method, and the output accuracy was 91 percent. We are using only 3 classes of lemon.

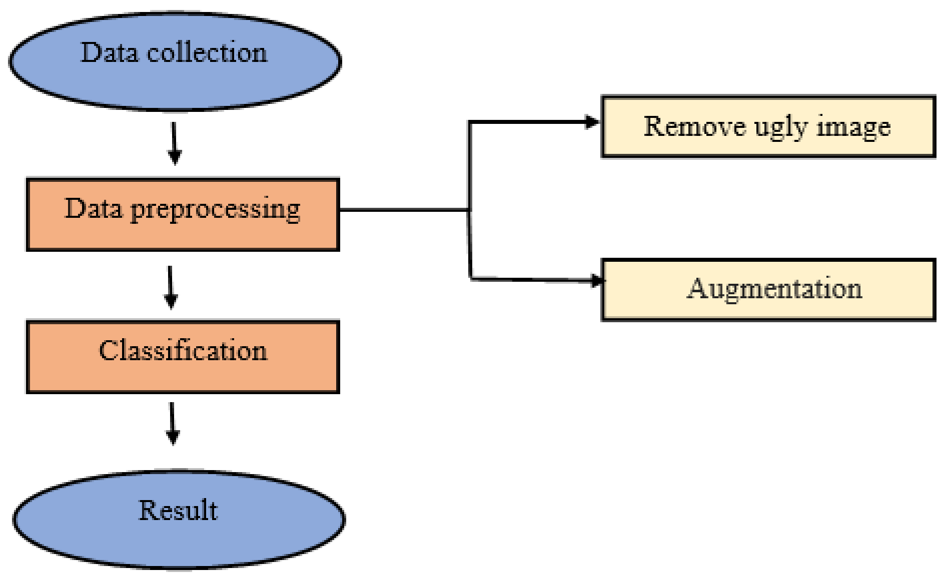

Figure-3.1.1:Methodology

3.2 Implement CNN

CNN dominance is not an easy feat to achieve. This is the first stage of my journey, so brush up on the fundamentals before proceeding. As it were, an introduction to Convolutional Neural Networks.

3.2. 1 Convolution Operation

Convolution activity is the primary building component in our attack strategy. In this phase, we will look at include indicators, which act as conduits in the brain organization. We will also look into highlight maps, learning about the boundaries of such guides, how instances are identified, the layers of identification, and how the discoveries are depicted in Figure 2.[23]

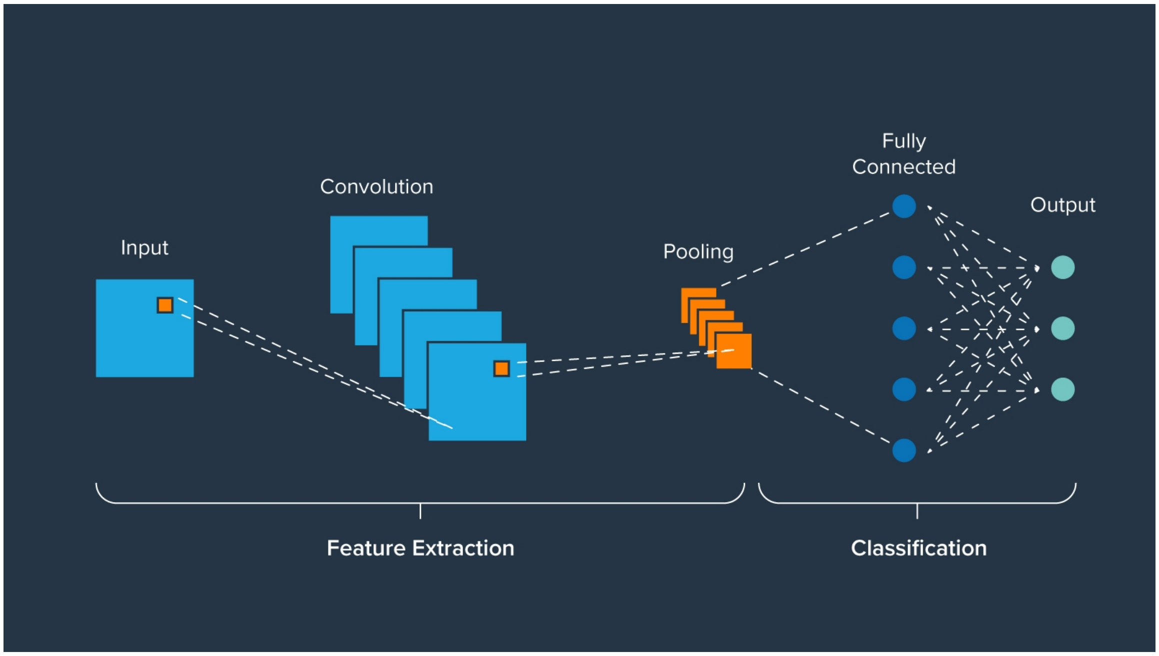

Figure -3.2.1: Image Classification steps with CNN

3.2. 2 Relu _layer

The Rectified Linear Unit, or ReLU, will be the second step in this development. I will discuss ReLU layers and study linearity capabilities in convolutional neural networks. A quick workout to improve my skills, on the other hand, isn't intrinsically hazardous.

3.2. 3 Pooling_layer

This part is about pooling, and we'll look at how it works in general. However, we will concentrate on a specific type of pooling: maximal pooling. There will be a discussion of several ways to compute the mean (or total). Following that, I'll see a visual intelligence technology that will deduce the entire notion for us.

3.2. 4 Flattening

The straightening cycle and how we transition from pooling to leveled layers will be briefly covered while dealing with Convolutional Neural Networks.

3.2. 5 Complete connection

This section will recap everything we've discussed thus far. With this knowledge, you'll have a better understanding of what Convolutional Neural Networks perform and how the "neurons" that are produced gradually become familiar with visual organization.

After a while, I'll come to a finish and offer a short overview of the episode's concept. I will most likely encounter these ideas when working with Convolutional Neural Networks, and being familiar with them will be quite beneficial. SoftMax and Cross-Entropy facts to help you comprehend Convolutional Neural Networks. Let's get this part started.

3.3 Research Subject and Instrumentation

The essential component of the exploration effort is information. Finding spectacular knowledge and outstanding calculations or models for my assessment task is an extremely essential component for an expert. I must also learn about related exam papers. Here, my research topic is good lemons and finding diseased lemons my research tools are Google, Coolab, Python,different types of research papers, etc.

3.4 Data Collection Procedure

Data Collection: I gathered photos of lemons that were both healthy and diseased.Good-lemon:380, Citrus-canker-disease:111, Citrus-rust-mite-disease:58 As a result, I've developed three classes that can distinguish between healthy and infected lemons.

After augmentation Good-lemon:1140, Citrus-canker-disease:1110, Citrus-rust-mite-disease:1160, however, I employed a CNN-based architecture trained and evaluated on the Lemon disease dataset.

Figure-3.4.1: (a)-(c) From Left to right (a) Good lemon (b) citrus canker disease (c) citrus rust mite disease

Data pre-processing: Data pre-processing is an important step in ensuring that your CNN-based model can learn successfully from the dataset and classify lemon diseases accurately. Depending on the dataset and the needs of your study, the particular pre-processing processes may differ. Data that has been properly preprocessed is the foundation for developing strong and accurate machine learning models. Due to the short size of our data collection, we start by removing unpleasant photographs and then do image augmentation.Through data augmentation additional pictures are formed from a single photograph in this case.

Import Image Data:I have some code in Co-Lab that imports our data from Google Drive in this section. All picture data must be uploaded to Google Drive as soon as possible. In Google Co-Lab, I import images with Python. Figure-3.4.2 depicts some of the steps.

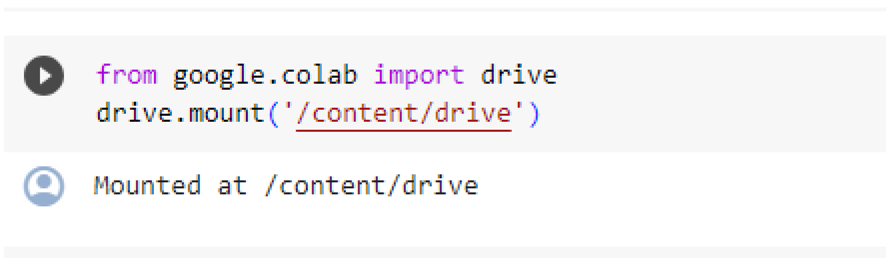

Figure-3.4.2: Data importing by Colab

3.5 Implimentation requirements

• 3.8 Python

One version of Python is Python 3.8. This programming language is sophisticated. For their inspection, the majority of analysts employ it. Because it is so easy to learn and understand, it is highly recommended as a programming language for AI-based work and is well-known among modern software professionals.

• CoLab at Google

The open source Python programming language merchant Google CoLab is permitted to use. Both Jupiter Notepad and our software allow us to work online here. Nevertheless, the main benefit of Google CoLab is that it allows us to use virtual GPUs online for free.

CHAPTER 4

IMPLEMENTATION STEPS AND EXPERIMENTAL OUTCOMES

4.1 Introduction

Image classification depends on the basis of the pixel. Identifying the disease is the most difficult. My model can detect diseases in the lemon and classify them properly. My model plays a vital role in detecting diseases where the lemon leaves are affected.

4.2 Arrangement by CNN

Image size trialing: The goal of stage one is to investigate how picture size affects model execution. Five estimated images total, ranging in size from 1280 x 720 to 256 x 256, are attempted. First, pre-prepared loads for Resnet34 are downloaded. Except for the final two levels, all layers are frozen by default in move learning.

These are specifically related to the lemon disease order task and contain novel loads. Freezing makes it possible to construct these layers independently of illness without having the angles spread backward. In exactly this manner.

The final layers are prepared using the 1-cycle technique. The remaining layers are provided with this total. A plot that contrasts learning rate with misfortune is made and broken down to aid in the adjusting cycle. The model is then run after suitable learning is selected from this. After recording the findings, the model is recreated in the additional four image sizes. All parameters, including the learning rate, remain constant in the preliminary Additionally, Figure -5 convolutional operation processing [24].

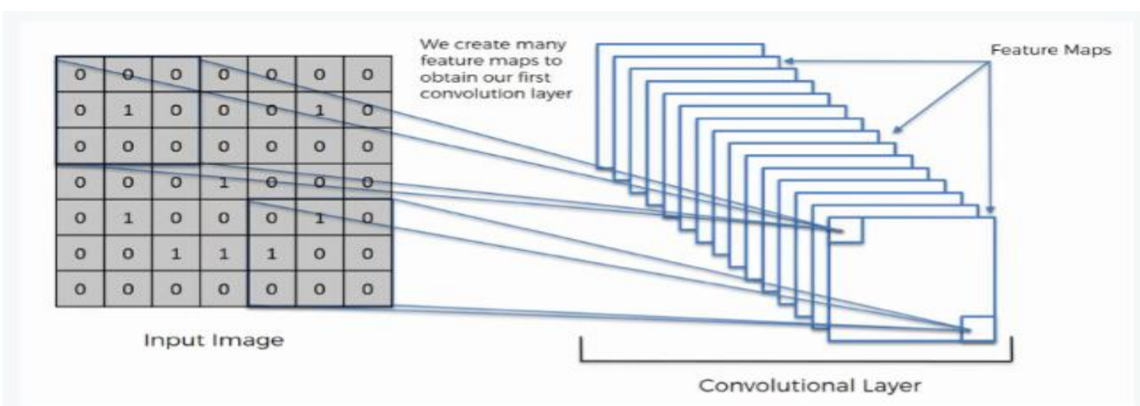

Figure -4.2.1: Input image process by convolutional layer

The model is improved by making use of the best possible image size. More growth options are included to further enhance the model's presentation (Figure 4.3.1.1). Activities include twisting (0.5) and splendor changes (0.4, 0.7). The final two layers are then produced at the default learning rate and disengaged. Using this sum, adjustments are made, and many preliminary tests are undertaken to evaluate a range of learning rates and age groups.

4.3 The Convolutional Operation

CNN is widely used and considered the best for jobs linked to images. Each image has three dimensions: weight, height, and RGB value, or the quantity of channels. RGB stands for red, green, and blue, and height and weight for picture control and color intensity, respectively. In order to tackle underfitting concerns, a neural network modifies the size of the picture and speeds up processing. Assuming that a picture has dimensions of 224 by 224 by 3, we may get our input vector as 1555, which is one dimension.

- The convolution layer
- Activation layer
- Pooling layer
- Batch norm layer
- Dropout layer
- Fully connected layer.

4.3.1 Convolution Layer

The most crucial layer in CNN is the convolution layer. A convolution filter is used for creating a feature map. The input convolution layer is represented by the left side 5*5 matrix in the picture, while the filter matrix, or kernel, is represented by the right side 3*3.

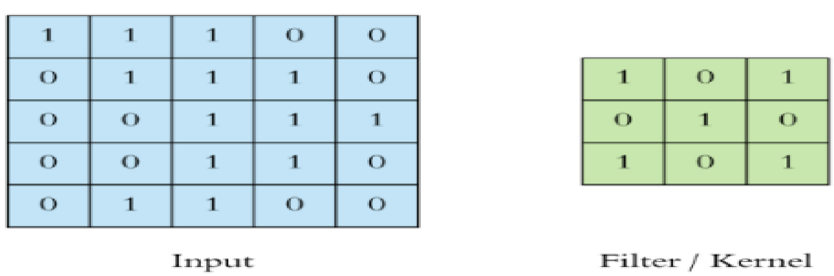

Figure-4.3.1.1: Image matrix multiplies kernel [medium]

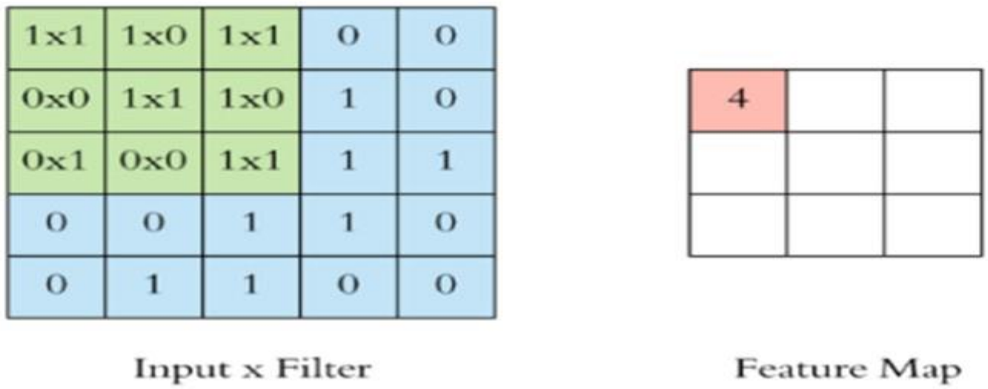

Figure-4.3.1.2: 3*3 output matrix [medium]

We are counting the convolution while swiping the filter matrix. We do matrix multiplication at each position and then add the results.

We then plot the total on the feature map after adding Since we computed the top left matrix of the image*filter, we placed our result in the top left position on the feature matrix. The filter matrix will then be shifted to the right of the cell by an interaction, which will also compute as before and set the sum in the next right cell of the feature matrix. Till the final exchange, this task will be done.

4.3.2 Activation Layer

In CNNs, an activation function is applied to the output of each neuron to introduce non-linearity. This non-linearity allows the neural network to learn complex patterns and relationships in the data. The most common activation function used in CNNs is the Rectified Linear Unit (ReLU), although other functions like Sigmoid or Hyperbolic Tangent (tanh) can also be used.

4.3.3Pooling Layer

We employ pooling in order to reduce dimensionality. For this reason, the issue of overfitting and training time lessens. Diverse forms of pooling exist. The most widely used of these is max pooling. Max pooling is used to extract the maximum value from the window. Pooling is parameter-free. However, we must define the stride value and window size.

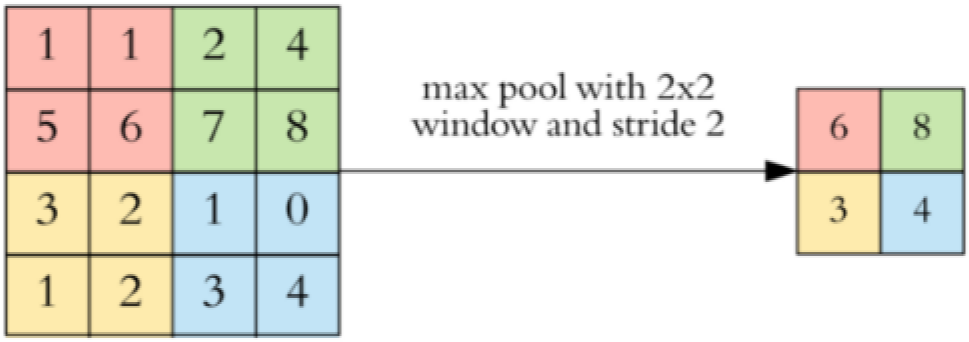

Figure-4.3.3.1: Max Pooling vs. Convolutions [mc.ai]

The graphic shows that our window is partitioned into four 2*2 matrices, with our window size being 4*4, max-pooling being 2,2, and our stride value being 2. We create a new 2*2 matrix by taking the maximum value from 4 matrices.

4.3.4 Dropout Layer

Accuracy is powerfully increased by the dropout layer. An further justification for employing the Dropout layer is to avoid overfitting issues. Every repetition of the training causes a neuron to decrease. At every level, the drop-out neurons undergo redesigns in order to advance to the next. Dropout can be used in input and hidden layers. We can only employ dropout during training period since the neurons we are losing will become dysfunctional.

4.3.5 Fully Connected Layer

A completely linked layer operates in the same way as a fed forward network. Every node in the next layer is linked to every other node. The input of the fully linked layer will be the result of final pooling and convolution. It is pressed flat. Using the flatten means figure. A 3*3 matrix is transformed into a vector by it.

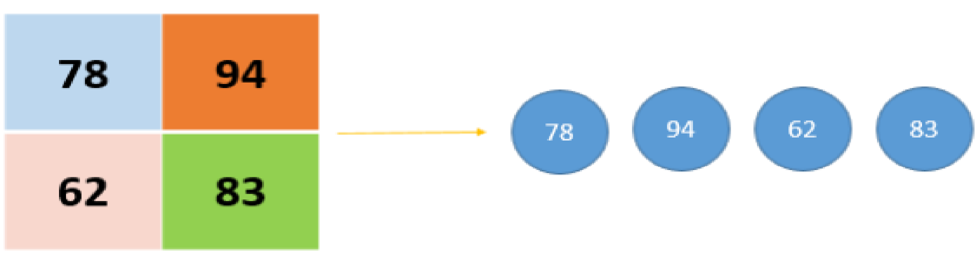

Figure-4.3.5.1: Fully connected Layer [Super data science]

4.4 Tools used

Any model or project development requires a few tools. In a similar vein, I needed a few tools to develop our model. The tools are listed below :

4.4.1 Google Colab

The Google Co Laboratory is a free and open-source Python library. Google Co. Labs offers GPU services. The procedure to enter Google CoLab is rather straightforward; all you need is an email account. Its internal memory is accessible to anyone. After mounting your disk, you may access and utilize your Google drive. This is an internet platform where you have to run each cell again if you close it.

4.4.2 Tensorflow python library

A framework that supports both low-level and high-level API is called TensorFlow. It was created by the Brain Team of Google. A Python library for deep learning is called TensorFlow. It wasn't free for everyone when it was initially built, but now that it's a well-known library, Google has made it free. C++ was used in its development to enable rapid running. The scientific calculation was applied to the numerical data. TensorFlow is utilized in the development of several applications, such as sequence to sequence modeling, handwritten digit recognition, and picture segmentation. It is able to train and DNN.

4.4.3 Keras python library

Keras is a free and open-source Python toolkit for high-level neural networks that supports many competition engines for backend neural networks. with the use of the user-friendly machine learning tool Keras. We are aware that Keras supports low-level APIs. Use of TensorFlow is required if we want to offer low-level APIs. Thus, keras may be considered a component of TensorFlow.

4.4.4 Technology used

In this modern day, we have all the modern technology we need to do any particular activity. The issue must be taken into consideration when selecting technology. Since our model was primarily created for picture categorization, a deep learning approach is being used. As is well known, CNN works best for classifying images.

4.5 Steps of Implementation

Now, I will describe the implementation process. I will try to provide a clear concept of our implementation process.

4.5.1 Library Importing

The model made use of some of the required libraries.

```
[ ] import numpy as np
    import pandas as pd
    import matplotlib.pyplot as plt
    from matplotlib.image import imread
    %matplotlib inline
    import seaborn as sns
    import cv2
    import os
    import re

    from sklearn.model_selection import train_test_split

    from tensorflow.keras.models import Sequential
    from tensorflow.keras.layers import Activation, Dropout, Flatten, Dense, Conv2D, MaxPooling2D, BatchNormalization
    from tensorflow.keras.callbacks import EarlyStopping
    from tensorflow.keras.optimizers import Adam
    from tensorflow.keras.preprocessing.image import ImageDataGenerator

    import warnings
    warnings.filterwarnings("ignore")
    warnings.warn("this will not show")

[ ] import tensorflow as tf
    import os
```

Figure-4.5.1.1: Library Importing

For picture datasets, Keras performs incredibly well. Keras is essentially used for data augmentation. There are many different kinds of data; the question is which kind we should preprocess. As we can see from the above, Keras is a component of TensorFlow. TensorFlow is required for Keras to function. Thus, the TensorFlow library is being used. In this case, the Matplotlib library is being used. A Python package called Matplotlib is used to plot datasets. A component of a matplot is a pyplot. It is essentially a matplot sub-library. It is employed in the numerical plot. Another scientific library is NumPy. It is employed in matrix computation. Here, we are utilizing an image's array operation. Keras contains all of the deep neural network's components.

4.5.2 Image visualization

```
[ ]  fig, axes = plt.subplots(2, 5, figsize=(15, 7))

     for ax in axes.flat:
         i=np.random.randint(0, data.shape[0])
         ax.imshow(Image.open(data.loc[i, "Imagen"]))
         ax.set_title(data.Calidad[i])
     #plt.tight_layout()
     plt.show()
```

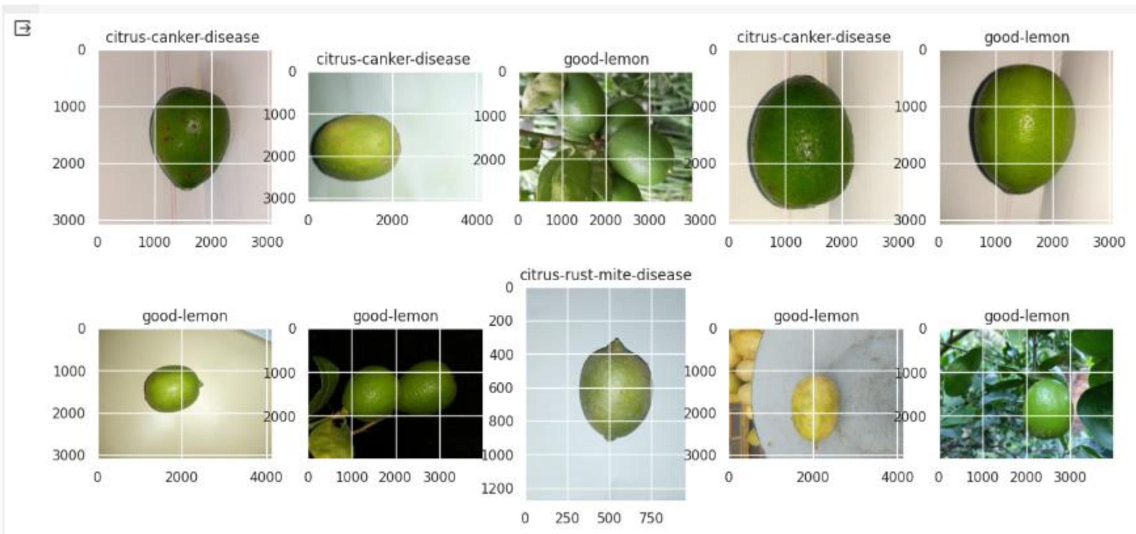

Figure-4.5.2.1: Image visualization

The method of displaying visual information in a picture so that people may readily grasp it is known as image visualization. Color mapping, 2D/3D visualization, histograms, and interactive tools are some of the techniques used in it. Image visualization helps us make better decisions, improve diagnoses, and get a deeper understanding of the complex data found in imaging, research, and engineering.

4.5.3 Preprocessing Images

Since my own collected data set is small, I use data augmentation techniques.

```
[ ]  data_augmentation = keras.Sequential([RandomFlip("horizontal"),keras.layers.RandomRotation(factor=0.2),
                                          RandomTranslation(height_factor=0.1, width_factor=0.1),
                                          RandomZoom(height_factor=(0.2, 0.3), width_factor=(0.2, 0.3)),
                                          RandomBrightness(factor=0.1),
                                          RandomContrast(factor=0.1)])
```

Figure-4.5.3.1 Preprocessing Images

4.5.4 Model building and compile

Among the several models in Keras Sequential is one. The way we constructed it was supposed to be sequential. Keras keeps track of the parameters' order and which comes first.

There are two layers in the feature extractor layer. Pool, Conv2D. Layer after layer, Keras passes pictures in a sequential manner. I used VGG16, MobileNetV2, NASNetMobile, ResNet152, EfficientNetV2S, and DenseNet201 models here. I mention the VGG16 model here because I have had more success with the VGG16 model.

```
[ ] VGG16_model_results = VGG16_model.fit(X_train, y_train, batch_size=32,
                                          epochs=10,
                                          validation_data=(X_test, y_test),
                                          callbacks=[early_stop, checkpoint, reduce_lr])
```

```
Epoch 1/10
14/14 [==============================] - ETA: 0s - loss: 1.5051 - accuracy: 0.4579
Epoch 1: val_loss improved from inf to 5.90519, saving model to VGG16_otosave.h5
14/14 [==============================] - 360s 26s/step - loss: 1.5051 - accuracy: 0.4579 - val_loss
Epoch 2/10
14/14 [==============================] - ETA: 0s - loss: 0.7567 - accuracy: 0.7380
Epoch 2: val_loss improved from 5.90519 to 3.12504, saving model to VGG16_otosave.h5
```

Figure-4.5.4.1: Model Compile

The model must first be trained, and then it must be assembled. When we train the model in the backend, gradient descent will function. In order to create a consistent gradient descent speed, we employ an optimizer. Every class is binary classified using binary cross-entropy.

4.5.5 Model Accuracy

When our model has completed training. Following that, creating the model's output and charting it is our first task. The training loss and validation loss portions of the code were displayed initially. Additionally, the training accuracy and validation accuracy of the second portion of the coding were displayed.

4.5.6 Graphing the Outcome

We are analyzing the model performance by visualizing accuracy and loss for each epoch using a graph.

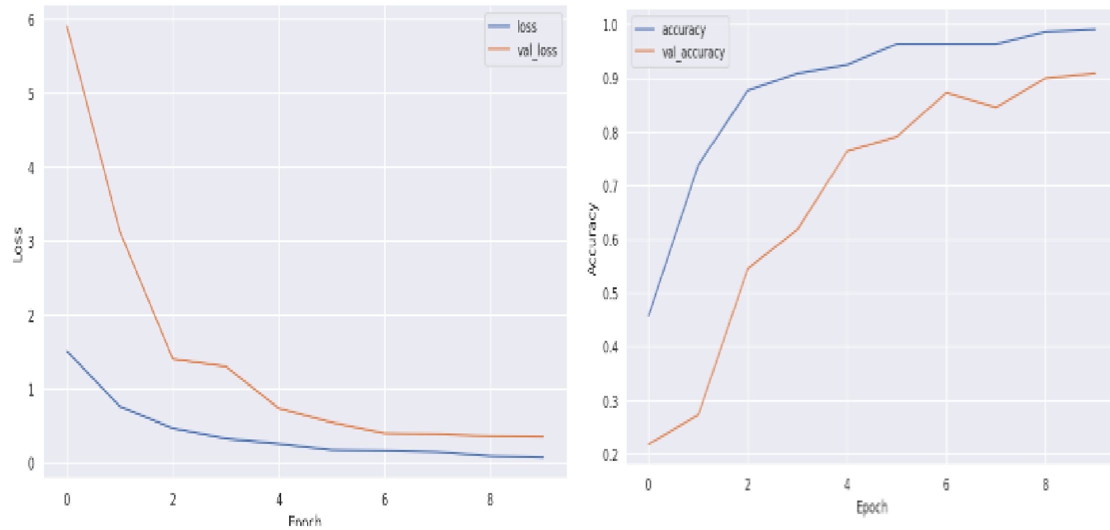

Figure-4.5.6.1: Graph of VGG16 Model

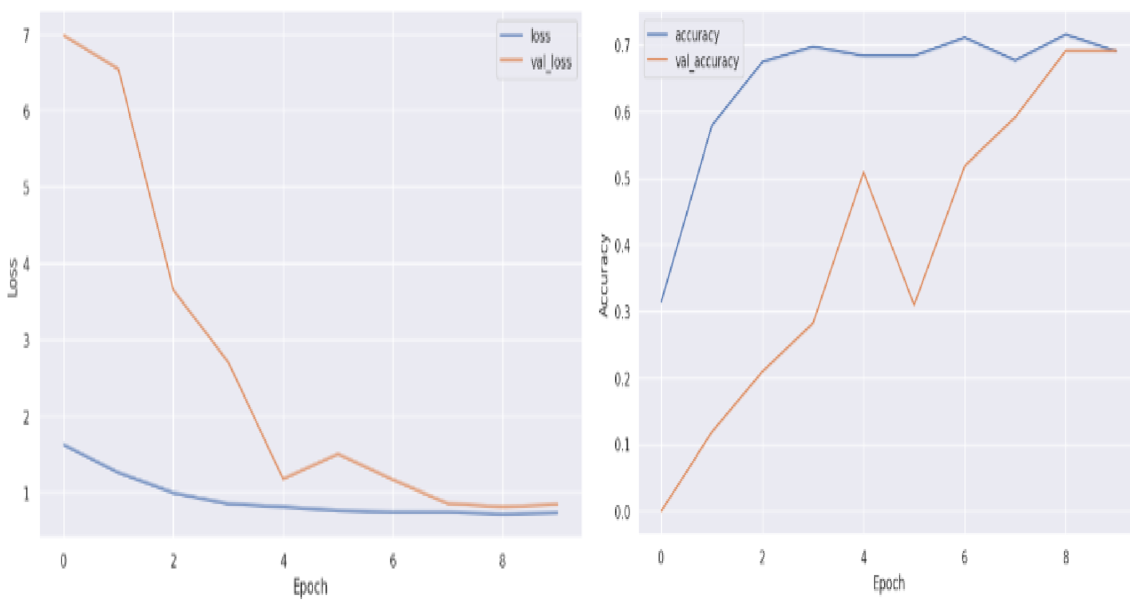

Figure-4.5.6.2: Graph of ResNet152 Model

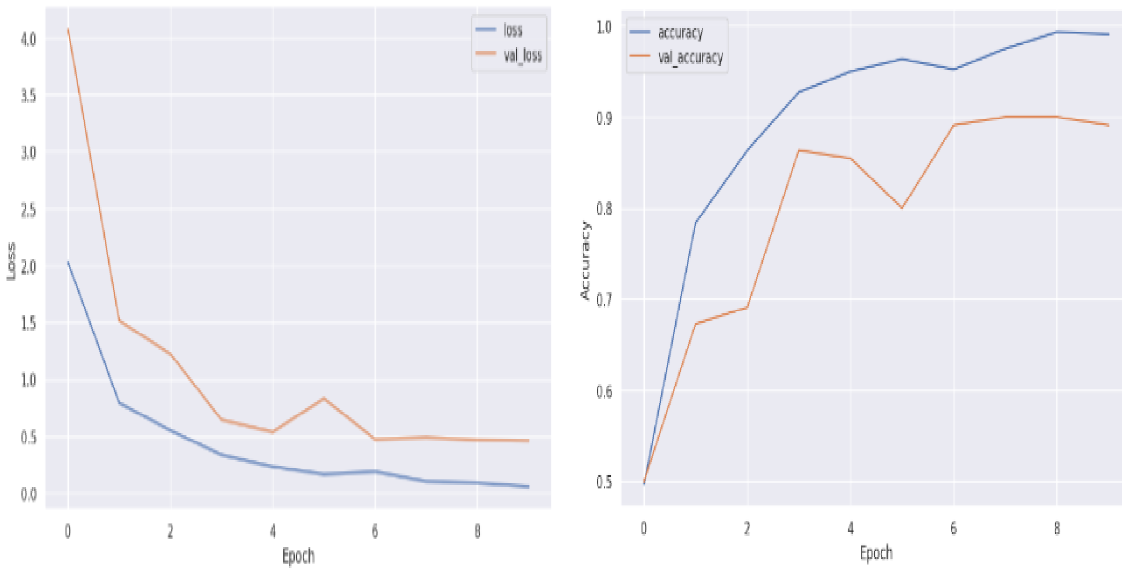

Figure-4.5.6.3: Graph of EfficientNetV2 Model

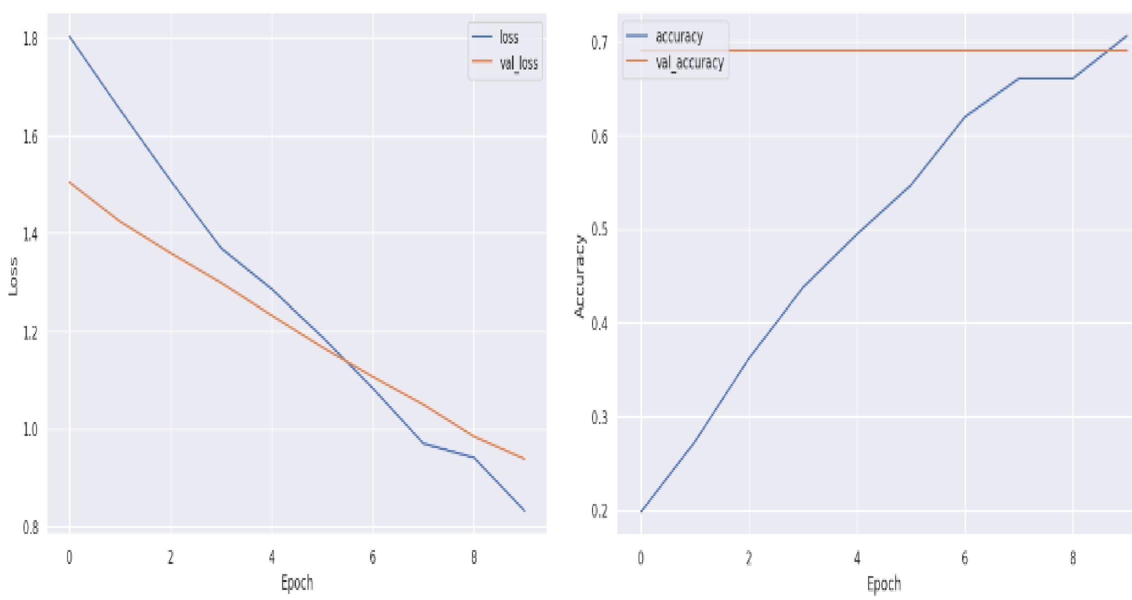

Figure-4.5.6.4: Graph of MobileNetV2 Model

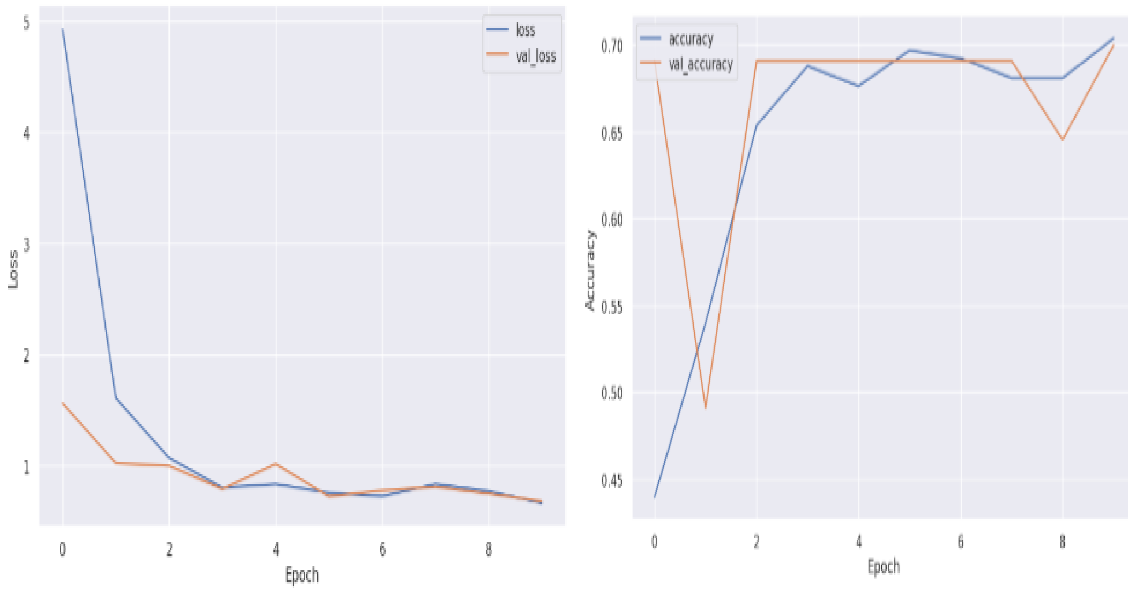

Figure-4.5.6.5: Graph of DenseNet201 Model

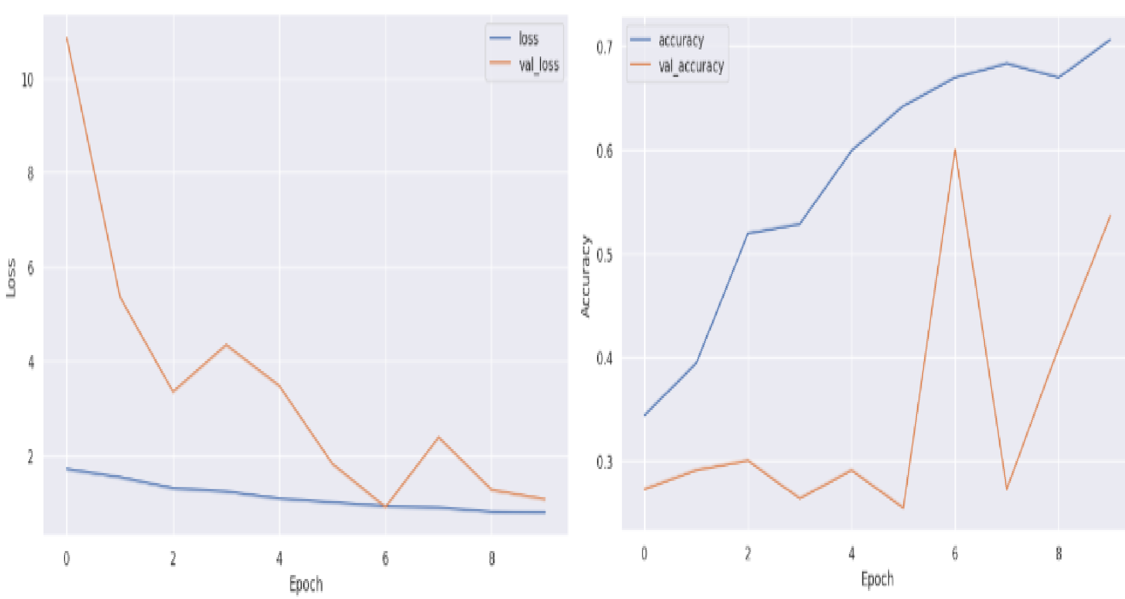

Figure-4.5.6.6: Graph of NASNetMobile Model

4.5.7 Classificaion report

Here I am showing the classification report for 10 epochs for each model.

Table 4.1: Classificaion report of my work

		precision	recall	f1-score	support
VGG16	0	0.90	0.97	0.94	76
	1	0.89	0.73	0.80	22
	2	1.00	0.83	0.91	12
	Accuracy			0.91	110
	macro avg	0.93	0.84	0.88	110
	weighted avg	0.91	0.91	0.91	110
ResNet152	0	0.69	1.00	0.82	76
	1	0.00	0.00	0.00	22
	2	0.00	0.00	0.00	12
	Accuracy			0.69	110
	macro avg	0.23	0.33	0.27	110
	weighted avg	0.48	0.69	0.56	110
EfficientNetV2	0	0.94	0.89	0.92	76
	1	0.79	0.86	0.83	22
	2	0.79	0.92	0.85	12
	Accuracy			0.89	110
	macro avg	0.84	0.89	0.86	110
	weighted avg	0.90	0.89	0.89	110
MobileNetV2	0	0.69	1.00	0.82	76
	1	0.00	0.00	0.00	22
	2	0.00	0.00	0.00	12
	Accuracy			0.69	110
	macro avg	0.23	0.33	0.27	110
	weighted avg	0.48	0.69	0.56	110

DenseNet201	0	0.70	1.00	0.82	76
	1	0.00	0.00	0.00	22
	2	1.00	0.08	0.15	12
	Accuracy			0.70	110
	macro avg	0.57	0.36	0.33	110
	weighted avg	0.59	0.70	0.58	110
NASNetMobile	0	0.93	0.37	0.53	76
	1	0.39	0.91	0.55	22
	2	0.38	0.92	0.54	12
	Accuracy			0.54	110
	macro avg	0.57	0.73	0.54	110
	weighted avg	0.76	0.54	0.53	110

4.5.8 Accuracy Score

The results of the tests conducted based on the above model are given below

Table 4.2: Accuracy Score of my work

	Model	Accuracy Score
0	VGG16	0.91
1	EfficientNetV2	0.89
2	DenseNet	0.70
3	ResNet152	0.69
4	MobilNet	0.69
5	NasNet	0.54

CHAPTER 5

RESULT ANALYSIS

5.1 Impact on Society

One of the main goals of this research is to help lemon producers financially. Bangladeshi farmers, like those in other South Asian countries, are not used to modern farming methods. As a result, they usually suffer financial losses. Because lemons are always susceptible to many illnesses, farming them is a challenging endeavor. Farmers have to rely on luck in order to make money. Our findings will aid farmers in the early disease detection of lemons. After this detection, you may take action for your benefit and reduce sickness. My technology has benefited farmers monetarily in this way. This disease detection technique can be applied to other plant diseases. CNNs are up-to-date with technology and easy to use on any platform. So, my work will be useful to farmers and the general public, who are not familiar with plant diseases. The country will grow economically. The foreign income of the country will increase. So, this research has a big social impact.

5.2 Impact on Environment

My research will employ a method of picture categorization. I've given this approach a lot of thought. Modern technologies are being used to overcome this issue. I get certain data from the field and use the internet to gather other necessary data. After gathering all of the necessary data, I constructed a useful model that produced quality results. With this approach, we have an effective outcome of ninety-one percent. Determining if lemons are impacted by disease is my primary objective.

5.3 Sustainability Plan

My study attempts to help farmers identify crop diseases early on. The Department of Agriculture and other agricultural organizations may operate more swiftly with the help of this format. Farmers stand to gain financially from my study, which will also have an impact on the national economy as a result of the growing demand for lemons in both domestic and foreign markets.

CHAPTER 6
CONCLUSION AND FUTURE WORK

6.1 Conclusion

CNN is only one of several techniques used to categorize images. Although we already know that CNN is the best, we are still researching its benefits, drawbacks, model complexity, and other factors before deciding whether or not to use it for our own model. We covered how each layer operates in Chapter 3.

Our data set is first divided into three folders: train, test, and validation. It is then stored in a root folder called lemon. There are several photos for each class, and I have handled this task by hand.

Next, I resize these pictures to fit inside 224 by 224 pixels. After rescaling, I must upload the mobile net in order to extract features. Max-pooling and convolution are two crucial stages in feature extraction. In essence, the convolution layer reduces the image's dimension by creating a feature map matrix and Max pool. We are correctly constructing this model after creating it in a sequential manner. Finally, it forecasts the picture.

6.2 Recommendation

To meet the requirements of models and algorithms, users need to adjust the connection of the modern internet. An ethical justification is undoubtedly needed for each project, which makes it safer.

Explain the decision-making process involved in the project and provide a basic explanation of the algorithm it uses. Crop production is very important for human. Because without food, human cannot survive. To complete the project, both teamwork and research are required. Those who produce lemons should know the results that this projector produces and how it can be used. Through the implementation of my proposal, I want to improve the production of current agriculture.

6.3 Scope of Future

- I will increase the dataset in the future, and I will work on other citrus fruits and other crops.
- In the future I will bring better results through more research.

- I will work more on how this project can help farmers better in the future.
- I'll put more effort into figuring out how this initiative might benefit farmers in the long run.
- I have worked on three types of lemons and in future I will work on more types of lemon diseases so that farmers can also benefit from it.
- I can use a computer-based application or a mobile-based application. I can work with all diseases by adding all the lemon data.
- I can use drones for real-time monitoring in a large field.

REFERENCE

[1] K. Thangadurai and K. Padmavathi, "Citrus Canker Disease Detection Using Genetic Algorithm in Citrus Plants." [Online]. Available: www.ijtrd.com

[2] K. Lalitha, K. Muthulakshmi, and A. Vinothini, "Proficient acquaintance based system for citrus leaf disease recognition and categorization." [Online]. Available: www.ijcsit.com

[3] "Medicinal and Health Benefits of Lemon," *Journal of Science and Technology*, no. Volume 5, pp. 16–20, Nov. 2020, doi: 10.46243/jst.2020.v6.i1.pp16-20.

[4] T. Ahmed, A. Muqit, J. Datta, M. Haque, and M. K. Haque, "Prevalence and severity of different citrus diseases in Sylhet region," *Journal of Bioscience and Agriculture Research*, vol. 23, no. 2, pp. 1957–1968, 2020, doi: 10.18801/jbar.230220.240.

[5] K. Padmavathi and K. Thangadurai Assistant Professor, "The Role of Image Enhancement in Citrus Canker Disease Detection," 2016. [Online]. Available: www.ijacsa.thesai.org

[6] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. ur Rehman, and K. Javed, "An automated detection and classification of citrus plant diseases using image processing techniques: A review," *Computers and Electronics in Agriculture*, vol. 153. Elsevier B.V., pp. 12–32, Oct. 01, 2018. doi: 10.1016/j.compag.2018.07.032.

[7] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Comput Electron Agric*, vol. 153, pp. 46–53, Oct. 2018, doi: 10.1016/j.compag.2018.08.013.

[8] G. Shrestha, Deepsikha, M. Das, and N. Dey, "Plant Disease Detection Using CNN," in *Proceedings of 2020 IEEE Applied Signal Processing Conference, ASPCON 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 109–113. doi: 10.1109/ASPCON49795.2020.9276722.

[9] S. Ray, A. Kundu, T. Na, Institute of Electrical and Electronics Engineers. Kolkata Section, and Institute of Electrical and Electronics Engineers, *CALCON 2020 : 2020 IEEE Calcutta Conference : proceedings : 28-29 February, 2020, Kolkata, India*.

[10] B. Subramanian, Jayashree. S, Kiruthika. S, and Miruthula. S, "Lemon Leaf Disease Detection and Classification using SVM and CNN," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 4, pp. 11485–11488, Nov. 2019, doi: 10.35940/ijrte.D4245.118419.

[11] R. Banni, "CITRUS LEAF DISEASE DETECTION USING IMAGE PROCESSING APPROACHES," *International Journal of Pure and Applied Mathematics*, vol. 120, no. 6, pp. 727–735, 2018.

[12] N. Shelar, S. Shinde, S. Sawant, S. Dhumal, and K. Fakir, "Plant Disease Detection Using Cnn," *ITM Web of Conferences*, vol. 44, p. 03049, 2022, doi: 10.1051/itmconf/20224403049.

[13] A. Dhakal and S. Shakya, "Image-Based Plant Disease Detection with Deep Learning," 2018. [Online]. Available: http://www.ijcttjournal.org

[14] A. Pawar, M. Singh, S. Jadhav, V. Kumbhar, T. P. Singh, and S. K. Shah, "Different Crop Leaf Disease Detection Using Convolutional Neural Network," 2023, pp. 966–979. doi: 10.2991/978-94-6463-136-4_85.

[15] D. R. Hammou and M. Boubaker, "Tomato Plant Disease Detection and Classification Using Convolutional Neural Network Architectures Technologies," in *Smart Innovation, Systems and Technologies*, Springer Science and Business Media Deutschland GmbH, 2022, pp. 33–44. doi: 10.1007/978-981-16-3637-0_3.

[16] D. Singh Solanki and R. Bhandari, "Lemon Leaf Disease Detection and Classification at Early Stage using Deep Learning Models", doi: 10.48047/nq.2022.20.22.NQ10345.

[17] Amity University, Institute of Electrical and Electronics Engineers. United Kingdom and Republic of Ireland Section, and Institute of Electrical and Electronics Engineers, *Abstract proceedings of International Conference on Automation, Computational and Technology Management (ICACTM-2019) : 24th - 25th April 2019*.

[18] J. G. Arnal Barbedo, "Plant disease identification from individual lesions and spots using deep learning," *Biosyst Eng*, vol. 180, pp. 96–107, Apr. 2019, doi: 10.1016/j.biosystemseng.2019.02.002.

[19] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors (Switzerland)*, vol. 17, no. 9, Sep. 2017, doi: 10.3390/s17092022.

[20] Kongunadu College of Engineering & Technology and Institute of Electrical and Electronics Engineers, *Proceedings, International Conference on Smart Electronics and Communication (ICOSEC 2020) : 10-12, September 2020*.

[21] W. N. Jasim, S. A. S. Almola, M. H. Alabiech, and E. J. Harfash, "Citrus Diseases Recognition by Using CNN Model," *Informatica (Slovenia)*, vol. 46, no. 7, pp. 85–94, 2022, doi: 10.31449/inf.v46i7.4284.

[22] R. Pourdarbani, S. Sabzi, M. Dehghankar, M. H. Rohban, and J. I. Arribas, "Examination of Lemon Bruising Using Different CNN-Based Classifiers and Local Spectral-Spatial Hyperspectral Imaging," *Algorithms*, vol. 16, no. 2, Feb. 2023, doi: 10.3390/a16020113.

[23] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front Plant Sci*, vol. 7, no. September, Sep. 2016, doi: 10.3389/fpls.2016.01419.

[24] SuperDataScience <<https://www.superdatascience.com/blogs/the-ultimate-guide-to-convolutional-neural-networks-cnn>>29.12.2023

APPENDIX

Appendix: Studies and Reflections

Throughout the course of my project activities, I encountered various challenges. However, three problems stood out as particularly significant. In this document, I acknowledge some of the major issues and shortcomings of works that utilized Convolutional Neural Networks (CNNs) to automatically detect crop diseases. Additionally, we have provided guidelines and strategies to further enhance the capabilities of CNNs in real-world applications. My supervisor is a kind person who is ever ready to help. She was quite helpful, and she immediately gave me some very insightful advice. While performing the study, she learned a great deal, including how to create CNN models and better datasets.

PLAGIARISM REPORT

Lemon Disease Classification Using CNN-based Architectures

ORIGINALITY REPORT

23%	20%	8%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	9%
2	Submitted to Daffodil International University Student Paper	7%
3	www.researchgate.net Internet Source	1%
4	weldonj.github.io Internet Source	1%
5	Submitted to University of Wales Institute, Cardiff Student Paper	<1%
6	"Exploiting Ensemble Learning for Rainfall Prediction using Meta Regressors and Meta Classifiers", International Journal of Engineering and Advanced Technology, 2020 Publication	<1%
7	Submitted to Grambling State University Student Paper	<1%
8	Submitted to University of Bristol Student Paper	<1%