# TABLE OF CONTENTS

[Chapter 1 **3**](#_bookmark0)

[Introduction **3**](#_bookmark1)

* 1. [Motivation 3](#_bookmark2)
  2. [Problem Statement 3](#_bookmark3)
  3. [Objectives 4](#_bookmark4)

[Chapter 2 **5**](#_bookmark5)

[Literature Survey **5**](#_bookmark6)

[Chapter 3 **8**](#_bookmark7)

[Theory And Fundamentals **8**](#_bookmark8)

* 1. [Agriculture 8](#_bookmark9)
  2. [Python 9](#_bookmark10)
  3. [Digital Image Processing 9](#_bookmark11)
  4. [Deep Learning 13](#_bookmark12)
  5. [Convolutional Neural Network 15](#_bookmark13)

[3.5.1 Introduction: **15**](#_bookmark14)

* 1. [Flutter 19](#_bookmark15)
  2. [Visual Studio Code 19](#_bookmark16)
  3. [Google Colab 20](#_bookmark17)

[Chapter 4 **21**](#_bookmark18)

[Implementation **21**](#_bookmark19)

* 1. [Classification Of Fruits And Vegetables: 21](#_bookmark20)
  2. [Disease Detection: 24](#_bookmark21)
     1. [Diseases In Potato: 24](#_bookmark22)
     2. [Diseases In Apple 30](#_bookmark23)

[Chapter 5 **41**](#_bookmark24)

[Result 41](#_bookmark25)

* 1. [Result –Classification Of Fruits And Vegetables 41](#_bookmark26)
  2. [Result – Potato Disease Prediction 41](#_bookmark27)
  3. [Result – Apple Disease Prediction 42](#_bookmark28)

[Chapter 6 **45**](#_bookmark29)

[Conclusion & Future Scope **45**](#_bookmark30)

* 1. [Conclusion 45](#_bookmark31)
  2. [Future Scope 46](#_bookmark32)

# CHAPTER 1

# INTRODUCTION

The Introduction chapter motivates the need for a mobile application that leverages deep learning for plant disease identification and management. It also states the problem statement and objectives of the project.

## MOTIVATION

Agriculture is critical for global food security, yet it faces challenges such as climate change, limited resources, and plant diseases. Plant diseases can lead to crop losses and food insecurity, particularly in developing countries, and traditional methods of disease identification can be inaccessible and time-consuming. To overcome these challenges, we propose a mobile application that leverages deep learning and CNN technology for efficient and accessible plant disease identification and management. By quickly and accurately identifying and managing diseases, farmers can improve yields and food security. The app will provide easy access to expert knowledge, empowering farmers to make informed decisions and optimize their farming practices. Ultimately, this project has the potential to significantly impact the lives and livelihoods of millions of farmers worldwide, contributing to the development of a sustainable and resilient agricultural sector.

## PROBLEM STATEMENT

Farmers struggle with managing plant diseases, often spending large sums without adequate technical support, leading to poor disease control, pollution, and harm to natural ecosystems. This project aims to address these challenges by utilizing Digital Image-Processing and machine/deep learning algorithms to develop an effective and non-invasive method for examining agronomic variables. By providing farmers with accurate and accessible information, this project can help identify efficiencies that lead to higher productivity, profitability, and lower input costs, while also reducing the negative impact of plant diseases on the environment.

## OBJECTIVES

The objective of this project is to develop a mobile application that utilizes Digital Image-Processing and machine/deep learning algorithms to provide efficient and accessible tools for plant disease identification and management. By leveraging these technologies, the app aims to accurately and quickly identify plant diseases, provide relevant information on their symptoms and treatments, and offer expert knowledge and advice to farmers. The ultimate goal is to empower farmers to make informed decisions and optimize their farming practices, leading to improved yields, lower input costs, and a more sustainable and resilient agricultural sector.

* + - This project aims to predict important factors such as Sorting of fruits/vegetables, Crop Health, and Plant species Identification/Classification for various crops.
    - Deep learning will be used to improve prediction accuracy.
    - The project aims to provide farmers with effective solutions for diseased crops.
    - The model will be designed to be feasible and user-friendly for farmers.
    - The project will provide real-time updates for farmers.

# CHAPTER 2

# LITERATURE SURVEY

Literature Survey, is a review of relevant research and literature on plant disease identification and management using digital image processing and machine/deep learning techniques.

A method to Classification of Carrots based on Shape Analysis using Machine Learning Technique. This method includes extraction of several features followed by Convolutional Neural Network and K-Nearest Neighbours. The training accuracy of 96% was achieved and a validation accuracy of 77%. [1]

External properties of fruits like colour, shape, size and the defects are considered as important features and the database comprises good quality mango images. The fruits are sorted and graded depending on these properties using t-Artificial vision systems and Image processing systems. The output product with maximum efficiency is obtained. [2]

This study aims to design a robust AI-based controller for tomato sorting processes as well as to provide a design for a budget tomato sorting machine. The performance of the most well-known convolutional and deep learning neural network algorithms was investigated in order to improve the classification accuracy of tomato sorting controllers. . The CNN algorithm yielded this accuracy using either RGB or grayscale features. Numerous AI-based algorithms were trained with the objective to assess their abilities to differentiate between those classes given that the highest detection accuracy values obtained using ANN, SVM, SOM and LVQ networks are 100%, 90%, 68% and 62.2%, respectively. [3]

In this paper, soft computing techniques have been successfully applied to modelling orange mass regression estimation. Adaptive neuro-fuzzy inference system (ANFIS) is adopted to predict mass of blood orange. Best ANFIS, linear and nonlinear regression models, yielded values of R2 , SSE, MSE of 0.99, 21.50, 1.65 (ANFIS), 0.91, 1156.69, 12.05 (linear) and 0.88, 1538.10, 15.86 (non-linear), respectively.[4]

In this study automatic apple grading by size, color, bruise and blemish using digital cameras and computerized image processing techniques.Systems accuracy rate is around 80%. [5]

This paper presents hardware for fruit grading where all the mangoes were graded using computer vision techniques. The accuracy that is obtained through classification is 96% . Thus this hardware is very cheap so that a farmer himself can analyse the quality of mangoes in the field itself. In this only 300 mangoes /hr. is accepted for processing. [6]

Different methods can be used to classify size and colour of fruit with machine vision. These methods are based on detecting and processing colour,shape features of fruit. In experimental works a certain amount of fruit was classified by workers and then the same fruits were classified by a grading system. According to these work results, accuracy rate reaches 69-75% depending on the kind of fruit. Accuracy rate can be increased to 94-98 % by applying NNA (Neural network algorithm) approach.[7]

This survey helps researchers to collect knowledge about methodologies used for fruit recognition such as pre-processing, segmentation, feature extraction, feature selection and classification techniques. Various issues related to fruit recognition were discussed and a common acceptable framework suggested. [8]

The aim of this study is to develop a system for detecting weeds and distinguishing them from agricultural crops based on machine learning algorithms and neural networks using its own data set. To achieve this goal to explore different geographical areas to identify common types of weeds and analyze the main types of grasses that are often found in the fields.with accuracy Knn - 83.3 %, Random Forest - 87.5 %, Decision Tree - 80 %[9]

To distinguish between weed and crop the machine vision-based method employs form texture, colour and location-based characteristics separately or combination. We must partition the dataset into a 70:30 ratio in the training and testing model, which is based on the scaling procedure. [10]

This paper explores the potential of machine learning algorithms for weed and crop classification from UAV images. The UAV images were collected from an Australian chilli farm and these images were pre-processed using image processing techniques.

Three different classifiers were tested using those properties: RF, SVM, and KNN. The . RF and SVM offered 96% and 94% accuracy in weed detection from RGB images, respectively whereas KNN offered only 63% accuracy. [11]

In this paper the authors build an algorithm to detect weeds based on a threshold on the color green and size base feature extraction. This approach includes classical computer vision techniques that focus on color matching and geometry of leaves. The results for the baseline model and the deep learning model for yield prediction. Model Mean Absolute Error of Linear Regression (Baseline) 0.185664 and Deep Neural Network 0.1585. [12]

The attributes such as shape, texture and color of plant leaf images are utilized to detect plant infections.Various kinds of machine and deep learning techniques were used for plant disease recognition and classification and achieved Overall accuracy 90%. [13]

This paper presented a framework for recognizing and classifying diseases in guava plants. For evaluation high-resolution guava leaf and fruit dataset was used. Segmentation was used to obtain colour histogram RGB, HSV, and textural LBP descriptors. Advanced classifiers such as Fine KNN, Cubic SVM, Complex tree, Boosted tree, and Bagged tree ensemble for image-level and disease-level classification were used. Overall the classification accuracy is 99%. [14]

# CHAPTER 3

# THEORY AND FUNDAMENTALS

The chapter "Theory and Fundamentals" covers the background theory and concepts related to deep learning and computer vision, including convolutional neural networks (CNNs), transfer learning, data pre-processing, and model evaluation techniques.

## AGRICULTURE

Agriculture, which comes from the Latin words for soil and cultivation, is the science and art of growing crops, raising livestock, and preparing plant and animal products for human use. It is a vital industry that provides most of the world's food and materials, with nearly two-thirds of the global population relying on agriculture for their livelihood. In India, agriculture plays a particularly significant role, employing nearly 60% of the country's workforce and occupying 43% of its geographical area. Despite a decline in its share of India's GDP, agriculture remains the largest contributor to the country's economy.

Technology continues to transform the agriculture industry, with new innovations and advancements constantly emerging. Precision agriculture, for example, uses technology such as GPS, sensors, and drones to optimize farming practices by providing detailed information on soil conditions, crop growth, and weather patterns. This data can be used to improve crop yields, reduce waste, and minimize environmental impact.

Modern agriculture has become increasingly dependent on technology and scientific innovation, with agricultural engineers playing a crucial role in fields such as irrigation, drainage, conservation, and sanitation. Digital Image Processing, Machine Vision, and Computer Vision are among the technologies that have revolutionized the agricultural sector. Automated systems that utilize these technologies can serve a variety of purposes, including crop monitoring, disease identification, and quality control. These advancements have the potential to improve productivity, reduce costs, and make agriculture more sustainable and efficient.



**Fig. 3.1:** Agriculture Application

## PYTHON

Python is a high-level, interpreted programming language that has gained popularity among developers due to its simplicity, readability, and extensive library support. One of the reasons Python is preferred for deep learning and CNN is the availability of libraries such as Scikit-learn, Pandas, Keras, and TensorFlow, which simplify the coding process and allow developers to focus on model creation and analysis. Scikit- learn is a popular library for basic ML algorithms such as clustering, regression, and classification, while Pandas Grayscale is used for advanced data analysis, including data merging and filtering.

Keras is another library widely used for deep learning, which allows developers to create and train artificial neural networks using both the computer's CPU and GPU, enabling faster calculations and prototyping. Tensor Flow is another popular library used for manipulating deep learning, which allows the building, training, and deployment of artificial neural networks on large datasets. The extensive support for libraries and pre-written code segments in Python makes it an ideal programming language for implementing complex algorithms such as deep learning and CNN in an efficient and straightforward manner.

## DIGITAL IMAGE PROCESSING

### IMAGE

To understand image processing, it's important to first understand the basic components of an image. An image is composed of pixels, each representing a point on the image with a specific shade, opacity, or color. The size of an image is determined by its dimensions, which are typically measured in terms of height and width, represented by the number of pixels. For instance, an image with dimensions of 500 x 400 (width x height) will have a total of 200,000 pixels.

The value of each pixel is typically represented in one of three ways. The first is as an integer between 0 to 255, where 0 represents complete blackness and 255 represents complete whiteness. Alternatively, a pixel can be represented using the RGB format, which is made up of three integers between 0 to 255, with each integer representing the intensity of the red, green, and blue color components respectively.

Image processing involves a series of fixed operations that are applied to each pixel of an image. These operations are performed sequentially, with the image processor moving through each pixel of the image and applying each operation in turn. The output of these operations can be calculated at any pixel of the image, and the process is repeated until all the operations have been applied to every pixel. Overall, image processing is a powerful tool that can be used in a wide range of applications, from computer vision to medical imaging and beyond.

### IMAGE PROCESSING

Image processing is a method of converting an image into a digital format and then using specific operations to extract useful information from it. Generally, image processing treats all images as two-dimensional signals and applies predetermined signal processing techniques.

### TYPES OF IMAGE PROCESSING

There are five main types of image processing: visualization, recognition, sharpening and restoration, pattern recognition, and retrieval.

* + - * Visualization is used to find objects that may not be visible in the image.
      * Recognition involves distinguishing or detecting objects in the image.
      * Sharpening and restoration techniques are used to create an enhanced image from the original image.
      * Pattern recognition is used to measure the various patterns around objects in the image.
      * Retrieval techniques are used to browse and search images from a large database of digital images that are similar to the original image.

### FUNDAMENTAL IMAGE PROCESSING STEPS

**Image Acquisition:** Image acquisition is the first step in image processing. This step is also known as pre-processing in image processing. It involves retrieving the image from a source, usually a hardware-based source.

**Image Enhancement:** Image enhancement is the process of bringing out and highlighting certain features of interest in an image that has been obscured. This can involve changing the brightness, contrast, etc.

**Image Restoration:** Image restoration is the process of improving the appearance of an image. However, unlike image enhancement, image restoration is done using certain mathematical or probabilistic models.

**Colour Image Processing:** Colour image processing includes a number of Colour modelling techniques in a digital domain. This step has gained prominence due to the significant use of digital images over the internet.

**Wavelets and Multi resolution Processing:** Wavelets are used to represent images in various degrees of resolution. The images are subdivided into wavelets or smaller regions for data compression and for pyramidal representation.

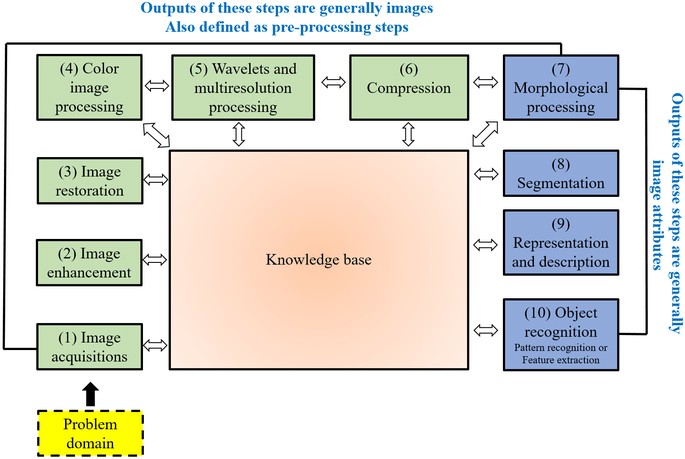
**Compression:** Compression is a process used to reduce the storage required to save an image or the bandwidth required to transmit it. This is done particularly when the image is for use on the Internet.

**Morphological Processing:** Morphological processing is a set of processing operations for morphing images based on their shapes.

**Segmentation:** Segmentation is one of the most difficult steps of image processing. It involves partitioning an image into its constituent parts or objects.

**Representation and Description:** After an image is segmented into regions in the segmentation process, each region is represented and described in a form suitable for further computer processing. Representation deals with the image’s characteristics and regional properties. Description deals with extracting quantitative information that helps differentiate one class of objects from the other.

**Recognition:** Recognition assigns a label to an object based on its description.



**Fig. 3.2** Fundamental Image Processing Steps

### APPLICATIONS OF IMAGE PROCESSING

Image processing is a field of study that involves manipulating digital images to extract useful information or transform them into a more desirable form. It has numerous applications across different industries, including medicine, traffic management, photography, and security. In medicine, image processing is widely used in research to enable more efficient and accurate treatment plans. However, since these applications require highly trained image processors, they require significant implementation and evaluation before they can be accepted for use.

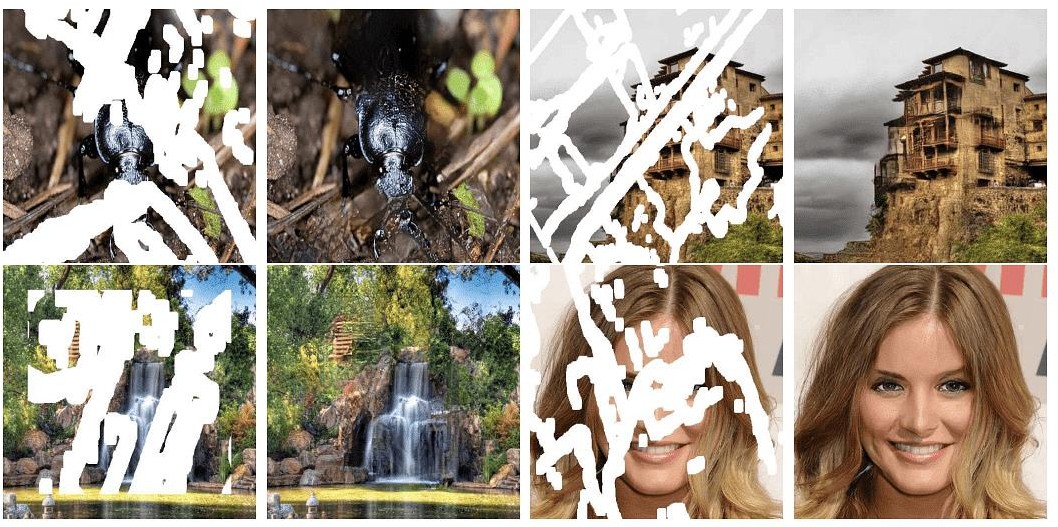
Another application of image processing is in traffic sensing technologies. The video image processing system (VIPS) is used in traffic sensors, and it consists of an image capturing system, a telecommunication system, and an image processing system. The detection zones can be set up for multiple lanes and used to sense the traffic in a particular station. By using this system, it is possible to monitor the traffic flow and detect any abnormality that may cause congestion or accidents. Additionally, VIPS can also detect vehicles' speeds, and with the help of telecommunication systems, authorities can notify drivers about any upcoming hazards.

Image reconstruction is another important application of image processing, where it can be used to recover and fill in the missing or corrupt parts of an image. This involves using image processing systems that have been trained extensively with existing photo datasets to create newer versions of old and damaged photos. Image reconstruction is not limited to photo restoration alone but can be used in fields such as astronomy to enhance the quality of images captured by telescopes.

Finally, face detection is one of the most common applications of image processing used today. It uses deep learning algorithms where the machine is first trained with the specific features of human faces, such as the shape of the face, the distance between the eyes, etc. Face detection is a vital tool used in security, biometrics, and even filters available on most social media apps these days. With the advancement in machine learning, face detection has become more sophisticated and accurate, making it a valuable tool in different areas of life.



**Fig. 3.3(a):** Normal Traffic **Fig. 3.3(b):** VIPS Image with detection zones



**Fig. 3.4:** Reconstructing damaged images using image processing

## DEEP LEARNING

Deep learning is a branch of machine learning that uses artificial neural networks to learn and improve on its own by analyzing computer algorithms. Unlike machine learning, deep learning networks are designed to imitate the human brain's structure and learn from large datasets. Due to advancements in Big Data analytics, these neural networks have become more complex and sophisticated, enabling computers to recognize and react to complex situations faster than humans. Deep learning has revolutionized fields such as image classification, language translation, and speech

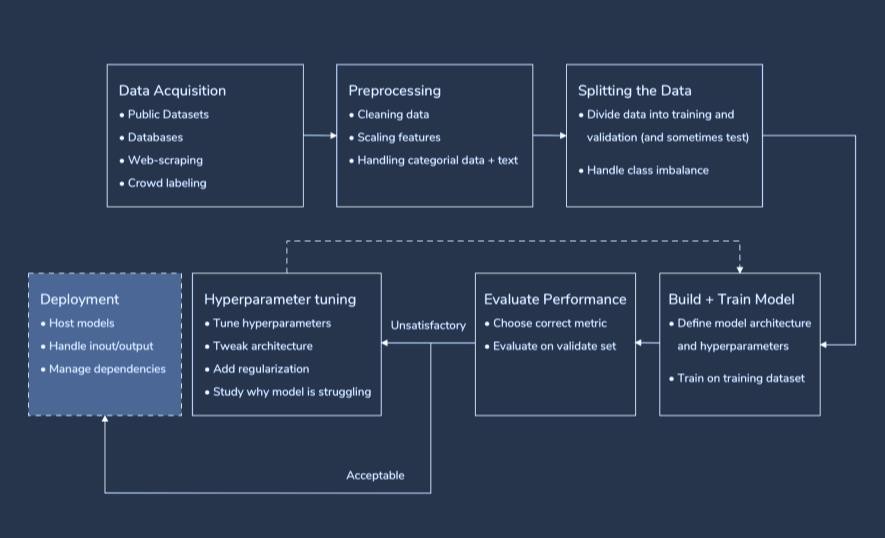
recognition, and has the potential to solve any pattern recognition problem without human intervention.

Artificial Intelligence and machine learning are critical components of the next wave of computing. They depend on the ability to recognize patterns, predict future outcomes based on past observations, and automate decision-making. Deep Neural Networks (DNNs) are a type of neural network that can perform complex operations such as abstraction and representation, making sense of images, sound, and text. As the fastest-growing field in machine learning, deep learning has become a truly disruptive technology, and businesses are increasingly using it to create new business models.

### WORKING OF DEEP LEARNING:

[Neural networks](https://www.simplilearn.com/tutorials/deep-learning-tutorial/neural-network) are layers of nodes, much like the human brain is made up of neurons. Nodes within individual layers are connected to adjacent layers. The network is said to be deeper based on the number of layers it has. A single neuron in the human brain receives thousands of signals from other neurons. In an artificial neural network, signals travel between nodes and assign corresponding weights. A heavier weighted node will exert more effect on the next layer of nodes. The final layer compiles the weighted inputs to produce an output.

* + - * Data acquisition is the process of measuring real-world physical events and converting them into digital signals that can be processed by computers.
      * Preprocessing is a necessary step to ensure that data is in a format that can be accepted by the network.
      * Data is split into training, validation, and test sets, and the validation set is used to fine-tune the model's hyperparameters after each epoch.
      * Deep learning models are built using neural networks, which process inputs through hidden layers using adjustable weights to produce predictions.
      * Evaluation metrics are used to measure the quality of a model, and there are many different types of metrics available.
      * The hyperparameters to tune in a deep learning model include the number of neurons, activation function, optimizer, learning rate, batch size, and epochs.



**Fig. 3.5:** Working Of Deep Learning

## CONVOLUTIONAL NEURAL NETWORK

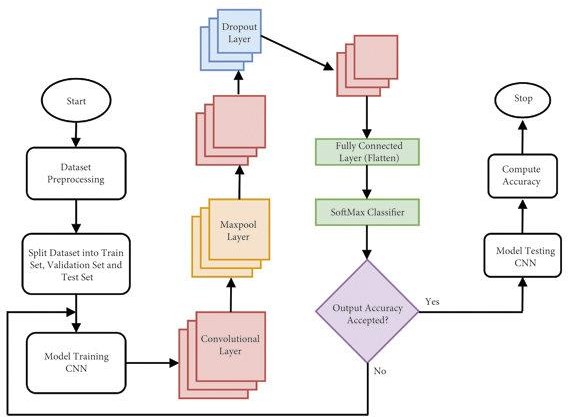
### INTRODUCTION:

Convolutional Neural Networks are a type of Deep Learning algorithm that are especially well-suited for image and video recognition tasks. They are able to automatically identify important features in images and use those features to make predictions about new images. This makes them an important tool for a variety of applications, from self-driving cars to medical diagnosis. The architecture of a ConvNet is inspired by the way neurons in the human brain are organized. In a ConvNet, each neuron in a given layer is only connected to a small region of the previous layer, called its receptive field. This allows the network to focus on local features in the image and helps to reduce the number of parameters that need to be learned.

The first layer of a ConvNet is typically a convolutional layer. In this layer, the network learns a set of filters that can be applied to the input image to extract various features. For example, the network might learn filters that detect edges or corners in the image. After the convolutional layer, the network typically includes a ReLU (Rectified Linear Unit) layer, which applies a non-linear function to the output of the convolutional layer. This helps to introduce non-linearity into the network and allows it to learn more complex features.

Finally, the network includes one or more pooling layers, which downsample the output of the convolutional and ReLU layers. This helps to reduce the number of

parameters in the network and can also help to make the network more robust to small variations in the input. Overall, Convolutional Neural Networks are a powerful tool for image and video recognition tasks. With enough training data and computational resources, they can learn to recognize a wide range of objects and scenes, making them an important tool for a variety of applications.



**Fig. 3.6:** Convolutional Neural Network

### LAYERS IN A CONVOLUTIONAL NEURAL NETWORK:

A convolution neural network has multiple hidden layers that help in extracting information from an image. The four important layers in CNN are:

1. Convolution layer
2. ReLU layer
3. Pooling layer
4. Fully connected layer

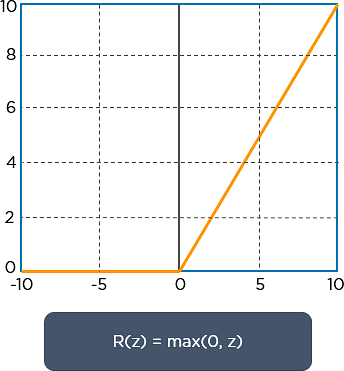
**Convolution Layer:** The Convolution Layer is the primary stage in the process of extracting relevant features from an image in a Convolutional Neural Network (CNN). It involves using several filters to perform the convolution operation.

During the convolution process, a filter matrix, also called a kernel, is applied to the input image matrix. The kernel matrix is generally of a smaller size, such as 3x3, and is slid over the entire input image. The values of the input image and the

kernel matrix are multiplied element-wise, and then the resulting products are summed up. The final result is then saved to a feature map. This operation is repeated for each position of the kernel matrix over the entire input image, generating multiple feature maps. The filters in the convolution layer can learn to detect specific features in the input image, such as edges or curves. By performing this operation multiple times using different filter matrices, the convolution layer extracts a variety of features from the input image, which can be used to classify the image later in the CNN.

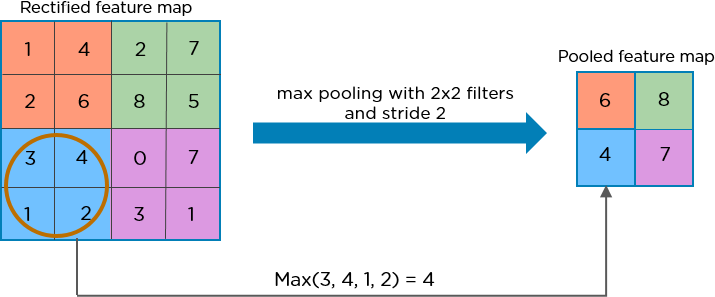
**Fig. 3.7:** Convolutional Layer

**ReLU layer:** The rectified linear unit (ReLU) is a crucial component of deep learning networks used to introduce non-linearity in the model. It operates on feature maps generated by the convolutional layer, where negative pixel values are set to 0. This element-wise operation generates a rectified feature map that helps improve the accuracy of the model.

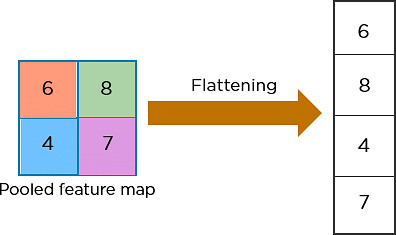
ReLU is a simple function represented by a graph that looks like an L shape. It has become a standard choice in deep learning architectures, thanks to its computational efficiency and effectiveness. The image to be analyzed is processed through a series of convolutions and ReLU layers to locate the significant features, which helps in achieving better results.

**Fig. 3.8:** RelU Layer

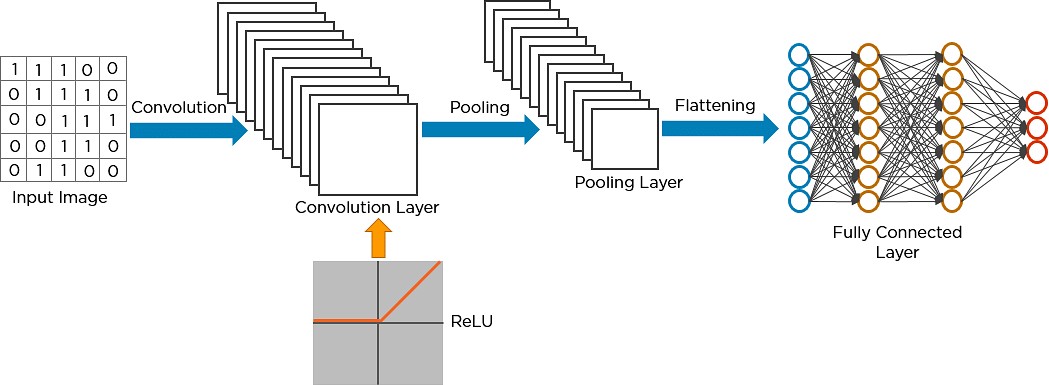
**Pooling Layer:** Pooling is a technique used to decrease the size of the feature map and to extract important information from it. After the feature map is generated by the

ReLU layer, it is passed through a pooling layer. This layer uses filters to identify various characteristics of the image, such as edges, corners, and specific features of objects like eyes or beaks. The pooled feature map contains only the most relevant information from the original feature map, which is important for reducing the dimensionality of the data and making it easier for the neural network to process.

**Fig. 3.9:** Pooling Layer

**Flattening:** After the pooled feature map is generated, the next step is to flatten the 2- Dimensional arrays into a single long continuous linear vector. This process is called flattening. The flattened matrix is then passed on as input to the fully connected layer, which is responsible for classifying the image.

**Fig. 3.10:** Flattening



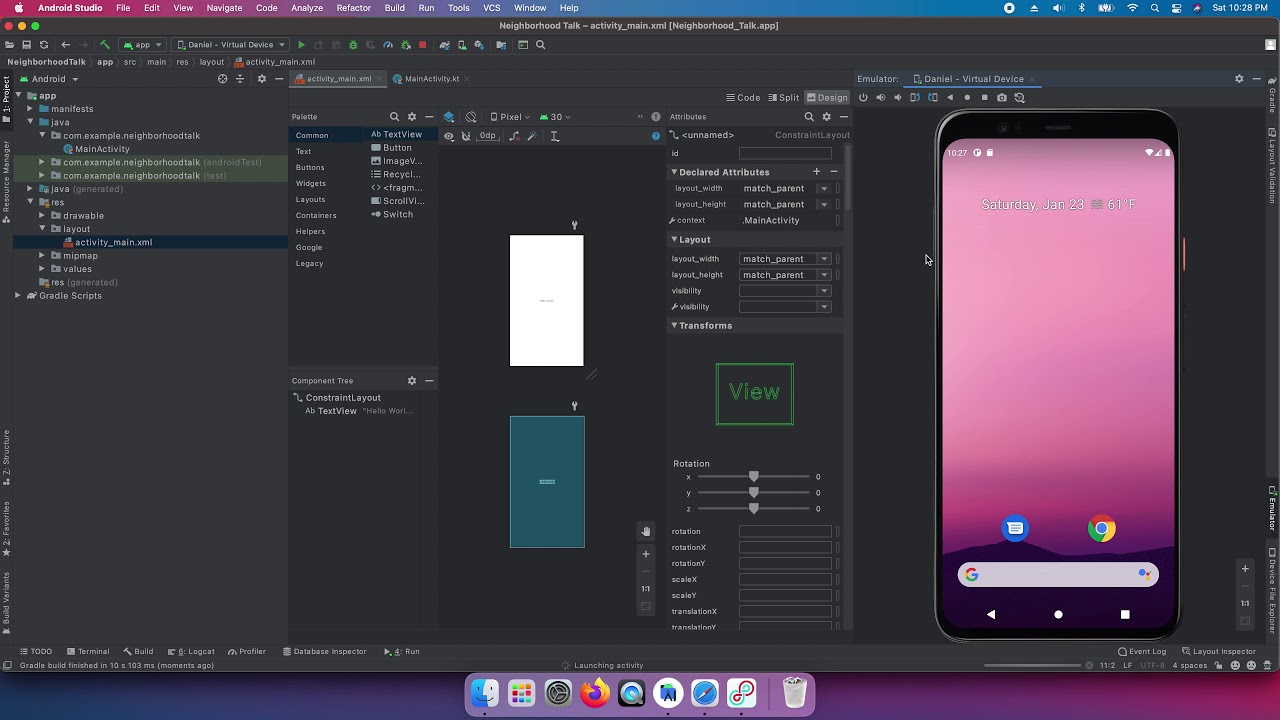
**Fig. 3.11:** Combination of layers

## Android Studio

**Android Studio** is the official integrated development environment (IDE) for Google's Android operating system, built on JetBrains' IntelliJ IDEA software and designed specifically for Android development. It is available for download on Windows, macOS and Linux-based operating systems. It is a replacement for the Eclipse Android Development Tools (E-ADT) as the primary IDE for native Android application development.

Android Studio was announced on May 16, 2013, at the Google I/O conference. It was in early access preview stage starting from version 0.1 in May 2013, then entered beta stage starting from version 0.8 which was released in June 2014. The first stable build was released in December 2014, starting from version 1.0. At the end of 2015, Google dropped support for Eclipse ADT, making Android Studio the only officially supported IDE for Android development.

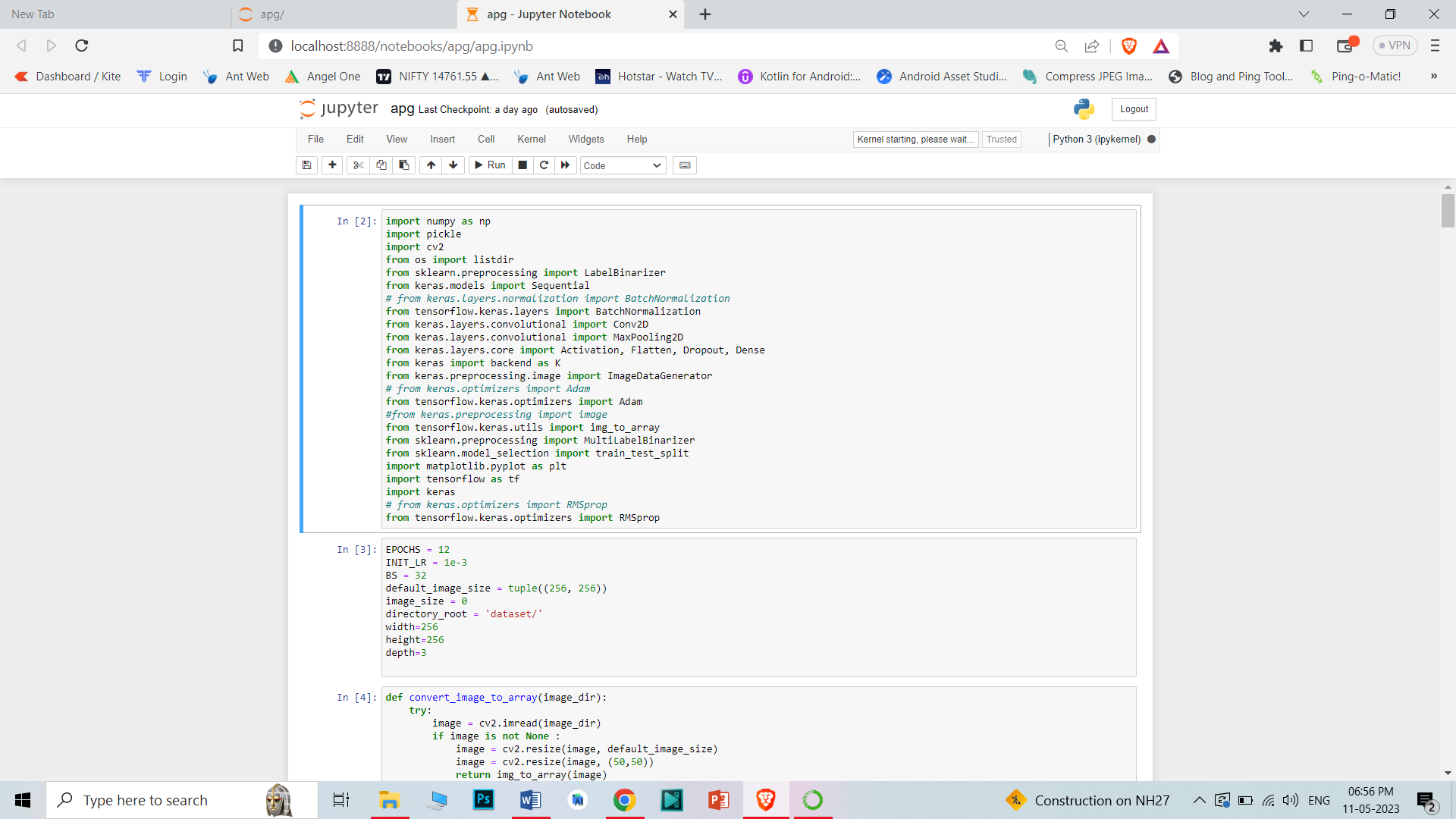
.



**Fig. 3.12:** Android Studio Outlook

## Jupyter Notebook

**Project Jupyter** is a project to develop open-source software, open standards, and services for interactive computing across multiple programming languages. It was spun off from IPython in 2014 by Fernando Pérez and Brian Granger. Project Jupyter's name is a reference to the three core programming languages supported by Jupyter, which are Julia, Python and R. Its name and logo are an homage to Galileo's discovery of the moons of Jupiter, as documented in notebooks attributed to Galileo. Project Jupyter has developed and supported the interactive computing products Jupyter Notebook, JupyterHub, and JupyterLab. Jupyter is financially sponsored by NumFOCUS..



**Fig. 3.14:** Jupyter Nootbook Outlook

# CHAPTER 4

# IMPLEMENTATION

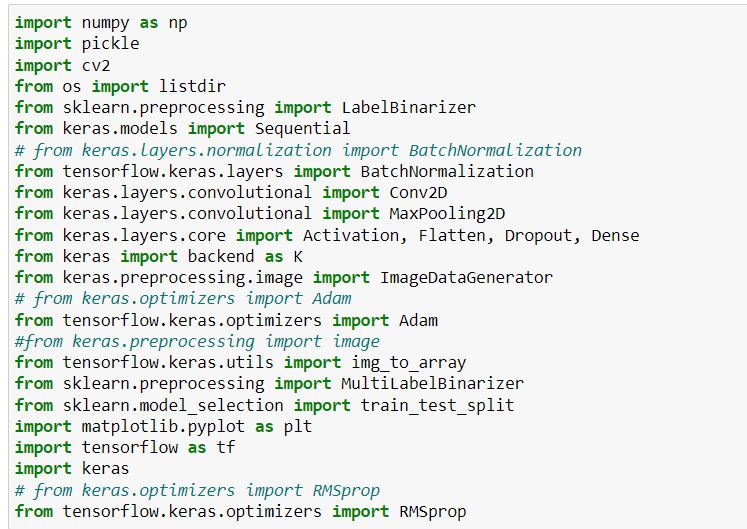
The chapter on implementation discusses the technical details of the proposed plant disease identification and management system, including the software architecture, data preprocessing, model development, and evaluation metrics.

## CLASSIFICATION OF FRUITS AND VEGETABLES:

The design flow involves importing libraries, loading dataset, compiling the model, Fit the model, predictions.

### IMPORTING LIBRARIES

* + - * Pandas: Pandas are an important library for data scientists. It is an open-source machine learning library that provides flexible high-level data structures and a variety of analysis tools. It eases data analysis, data manipulation, and cleaning of data. Pandas support operations like Sorting, Re-indexing, Iteration, Concatenation, Conversion of data, Visualizations etc,.
      * Numpy: The name “Numpy” stands for “Numerical Python”. It is the commonly used library. It is a popular machine learning library that supports large matrices and multi-dimensional data. It consists of in-built mathematical functions for easy computations.
      * Matplotlib: This library is responsible for plotting numerical data. And that‟s why it is used in data analysis. It is also an open-source library and plots high- defined figures like pie charts, histograms, scatterplots, graphs, etc.
      * TensorFlow: TensorFlow is an open-source numerical calculation library with high performance. AI organization and is now widely used for complex mathematical computations by mathematics, physics, and also machine learning researchers.
      * Pathlib: pathlib module provides an easier method to interact with the filesystem no matter what the operating system is. It allows a more intuitive, more pythonic way to interface with file paths.



**Fig. 4.1:** Libraries Imported for Classification of Fruits and Vegitable

### LOADING DATASET

To train complex models, we need to load large datasets. Hence to load data directly from Google Drive by using the mount drive method. This will import all the data from the Drive to the runtime instance. Training data is the set of the data on which the actual training takes place. Validation split helps to improve the model performance by fine-tuning the model after each epoch. The test set informs us about the final accuracy of the model after completing the training phase. A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model’s hyperparameters.



**Fig. 4.2:** Loading Dataset

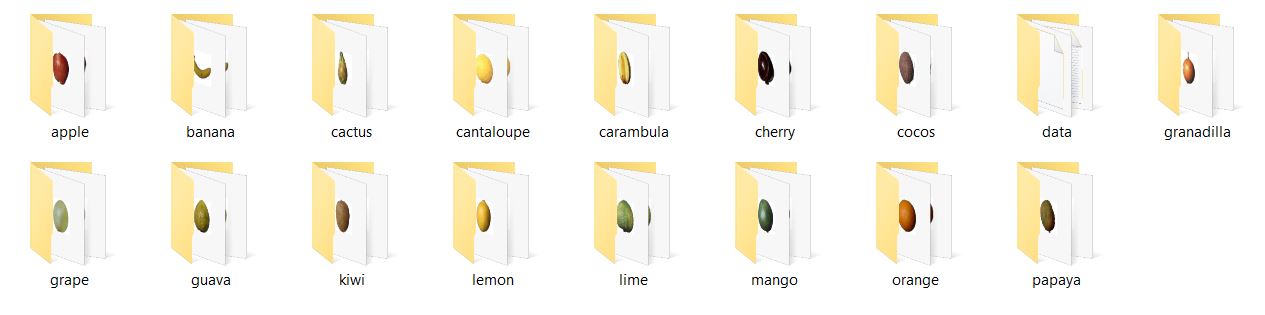
### COMPILE THE MODEL

After defining the model in terms of layers, you need to declare the loss function, the optimizer and initial weights, BIAS.

Optimizers shape and mold your model into its most accurate possible form by futzing with the weights. The loss function is the guide to the terrain, telling the optimizer when it's moving in the right or wrong direction. Bias is a phenomenon that skews the result of an algorithm in favor or against an idea. Bias is considered a

systematic error that occurs in the model itself due to incorrect assumptions the process.

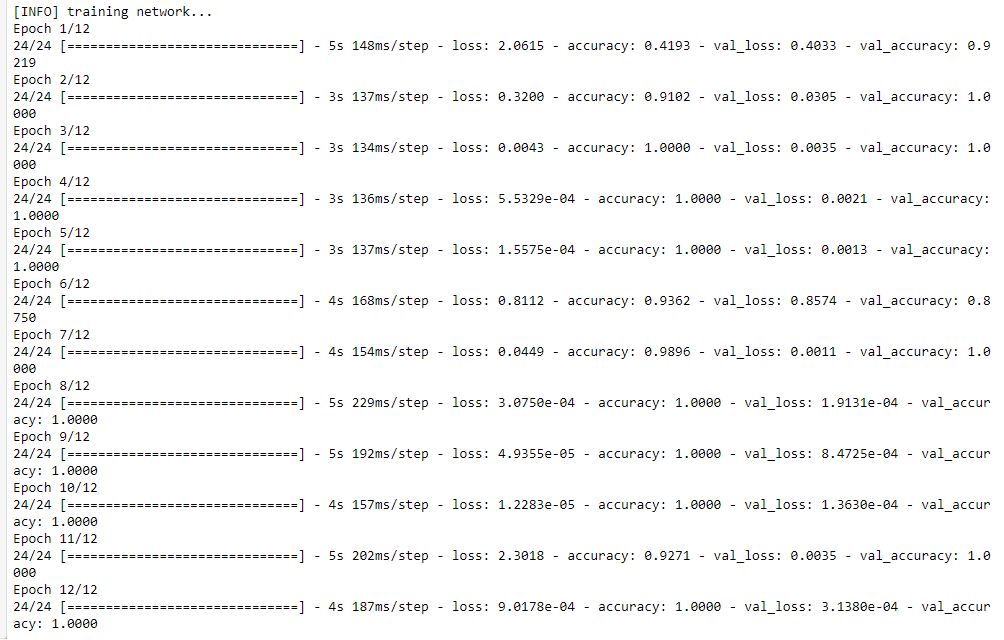
MobileNetV2 is a general architecture and can be used for multiple use cases. Depending on the use case, it can use different input layer size and different width factors. This allows different width models to reduce the number of multiply-adds and thereby reduce inference cost on mobile devices.



**Fig. 4.2:** Output of list of Fruits

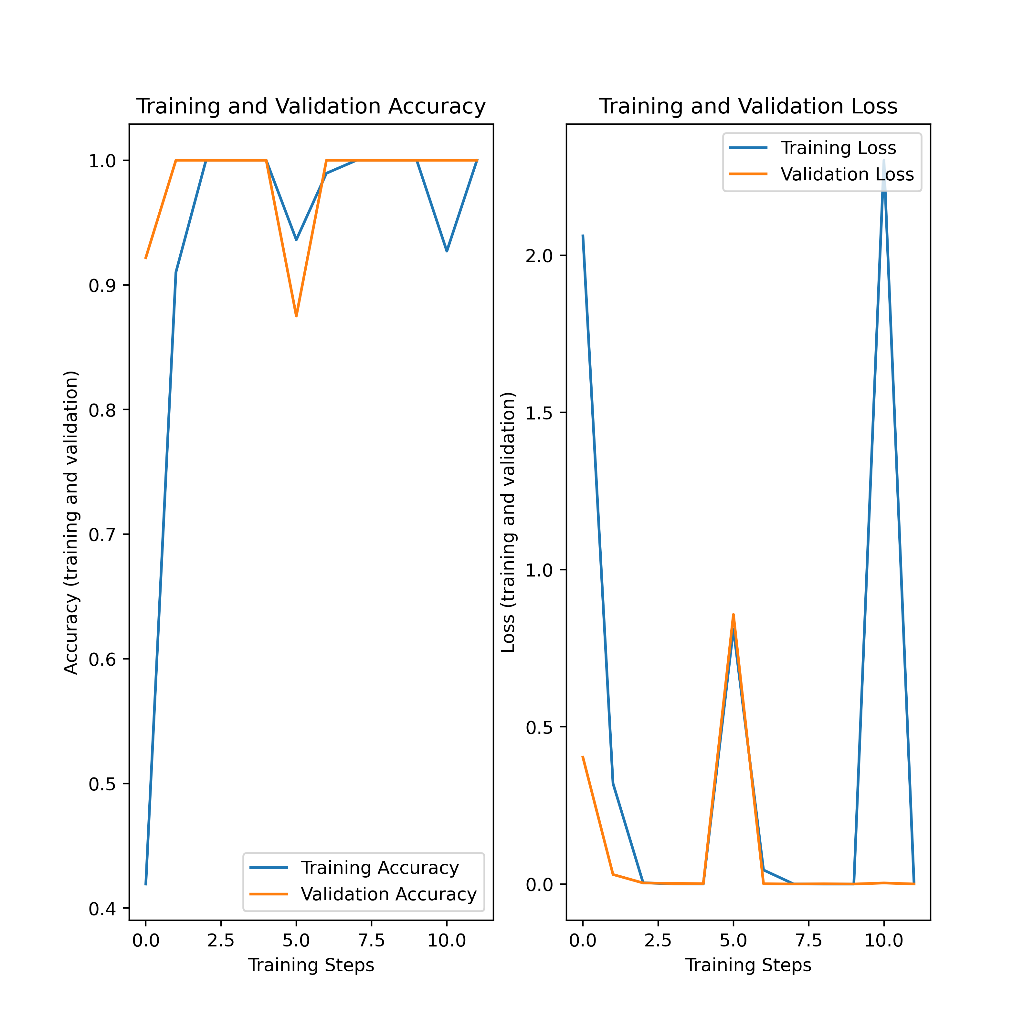
## FIT THE MODEL

Predictions are made by executing the model on the same data. Epochs are the number of iterations for the training process to run through the dataset and Batch size is a number of instances that are evaluated before the weight update.



**Fig. 4.3:** Fitting the Model

## Training and Validation Accuracy Graph

****

**Fig. 4.4:** Training and Validation Accuracy

# DISEASE DETECTION:

Identifying plant diseases is crucial for effective agricultural production as it allows for monitoring of plant growth and health, ensuring successful crop yields. The use of machine vision equipment in the field of plant disease research has become increasingly important. This technology involves capturing images of plants and using them to detect the presence of diseases or pests. In our project, we will be focusing on two crops, namely potatoes and apples.

### DISEASES IN POTATO as Example:

Potatoes contain high amounts of fibre that can assist in managing cholesterol and blood sugar levels. The fibre, known as "resistant starch", comprises both soluble and insoluble fibres that are advantageous for maintaining a healthy digestive system. Potatoes are a versatile vegetable with several health advantages. In addition to their nutritional value, potatoes have various industrial uses, including starch and alcohol production. Potato starch is utilized in textile mills for sizing yarn and in laundries. Additionally, potatoes are used in the production of glucose and dextrin.\

### Late Blight:

Late blight is a plant disease that causes damage to leaves, stems, and tubers. The affected leaves have a blistered appearance as if they have been scalded by hot water, eventually resulting in rot and drying out. As the disease progresses, the leaves turn brown or black in color, and spots covered with a flour-like substance appear on the underside of the leaves. The stems of affected plants turn black from the tips and eventually dry out. In severe cases, the disease causes all the foliage to rot, dry out, and fall off, causing the stems to dry out, leading to plant death. The affected tubers show dry, brown-colored spots on their skins and flesh. This disease acts rapidly and can cause plant death within two or three days if left untreated.

**Fig. 4.5:** Representation of early blight potato plant

The disease-causing organism can live on and in the soil among dead plant matter. It can also spread via contaminated seed tubers. The pathogen tends to thrive in conditions that are humid with low temperatures and wet leaves.

### Early Blight:

Early Blight is a common disease that affects potatoes at any stage of growth and results in leaf spots and blight. The symptoms of the disease usually become visible during the tuber bulking stage and progressively worsen until harvest. Small, black lesions are the first visible signs of the disease on potato plants, mostly appearing on the older foliage As time passes, these spots grow in size, forming a bull's eye pattern with concentric rings in the center of the affected area.The tissue surrounding the spots may turn yellow. In cases of high temperature and humidity, a considerable portion of the foliage may die.Stem lesions are similar to leaf lesions and can girdle the plant if they occur near the soil line.



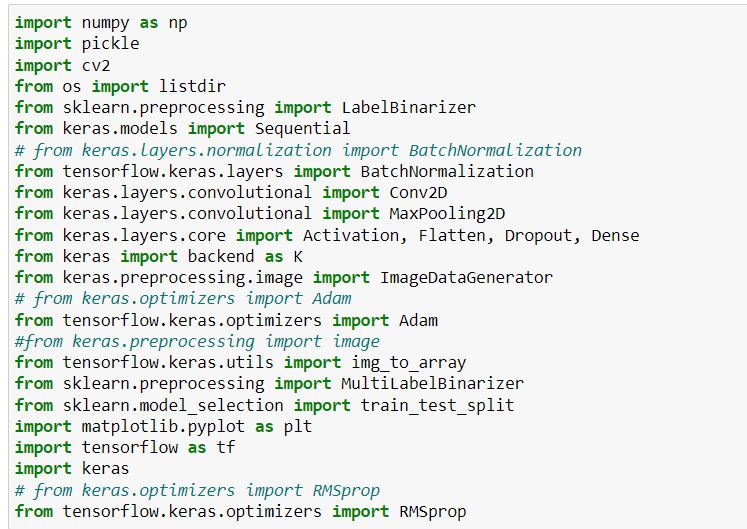
**Fig. 4.6:** Representation of early blight potato plant

The pathogen persists in infected plant debris in the soil and can survive for several years. It can also be present in infected seeds. The spores of the pathogen can be dispersed through various means such as water, wind, insects, animals (including humans) and machinery. The disease thrives in warm, rainy and wet conditions.

### IMPORTING LIBRARIES

The following libraries are used for the disease detection in potato:

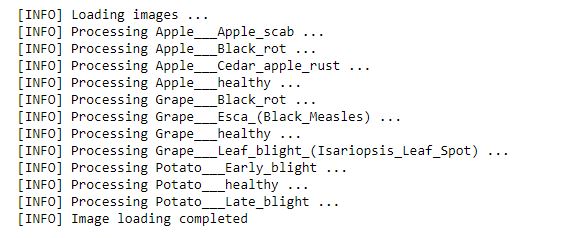
* + - * + Matplotlib: This library is responsible for plotting numerical data. And that‟s why it is used in data analysis. It is also an open-source library and plots high- defined figures like pie charts, histograms, scatterplots, graphs, etc.
        + TensorFlow: TensorFlow is an open-source numerical calculation library with high performance. AI organization and is now widely used for complex mathematical computations by mathematics, physics, and also machine learning researchers.



**Fig. 4.7:** Importing Libraries for Potato Disease Detection

### LOADING DATASET

To train complex models, we need to load large datasets. Hence to load data directly from Google Drive by using the mount drive method. This will import all the data from the Drive to the runtime instance. Training data is the set of the data on which the actual training takes place. Validation split helps to improve the model performance by fine-tuning the model after each epoch. The test set informs us about

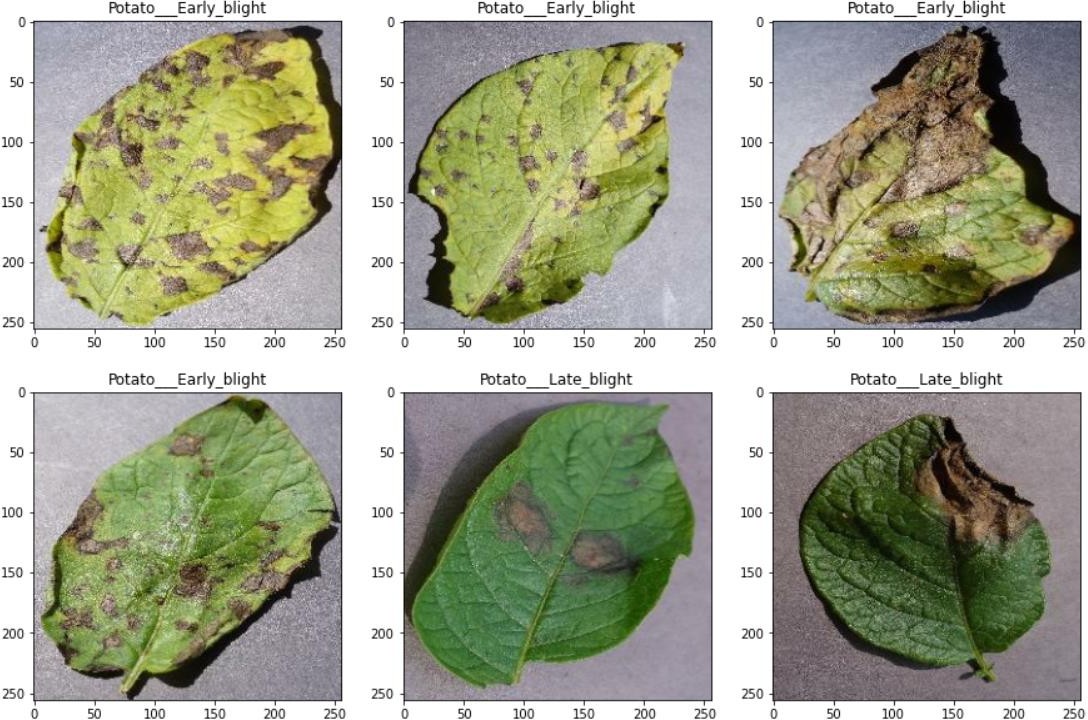


**Fig. 4.8:** Loading Dataset

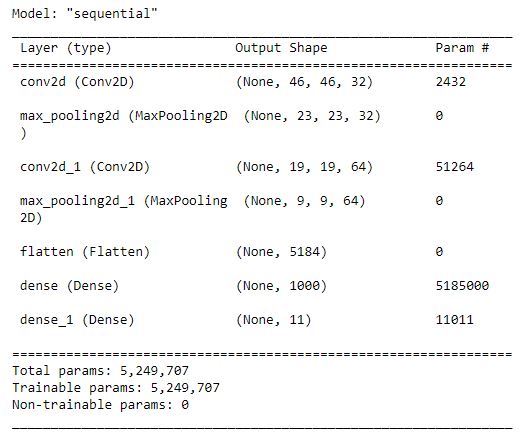
the final accuracy of the model after completing the training phase. A validation dataset is a sample of data held back from training your model that is used to give an estimate of model skill while tuning model’s hyperparameters.

### COMPILE THE MODEL

After defining the model in terms of layers, you need to declare the loss function, the optimizer and initial weights, BIAS.

Optimizers shape and mold your model into its most accurate possible form by futzing with the weights. The loss function is the guide to the terrain, telling the optimizer when it's moving in the right or wrong direction. Bias is a phenomenon that skews the result of an algorithm in favor or against an idea. Bias is considered a systematic error that occurs in the model itself due to incorrect assumptions the process.

**Fig. 4.9:** List of various Diseases in Potato Plant

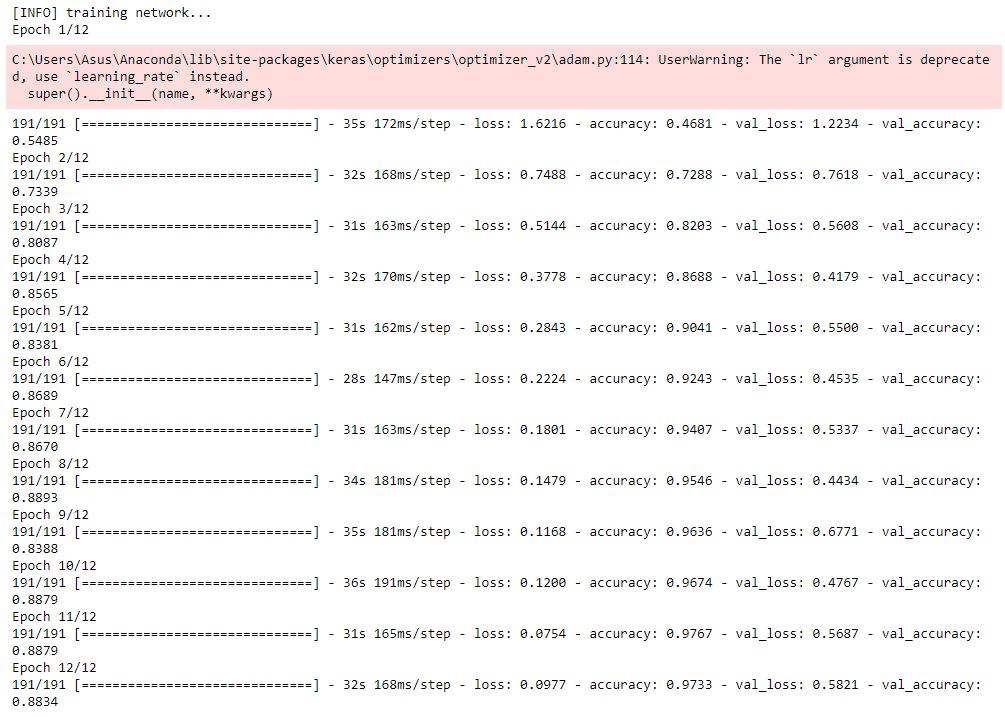


**Table 4.1:** Sequential\_2 Model

### FIT THE MODEL

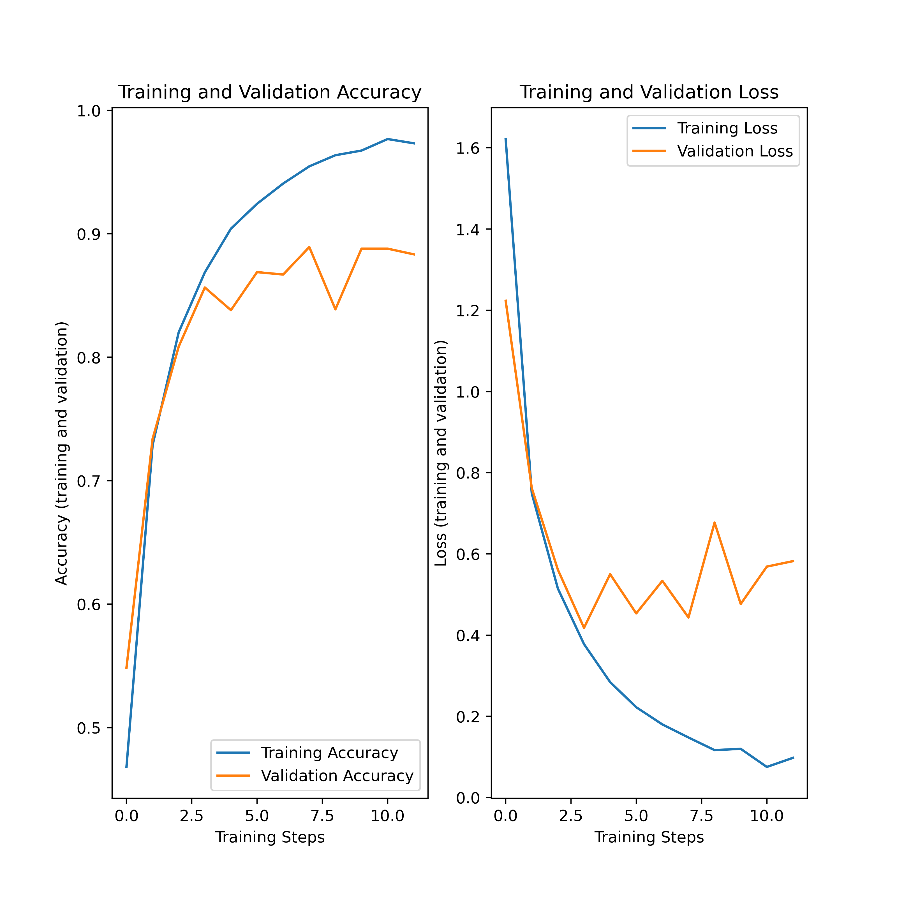
Predictions are made by executing the model on the same data. Epochs are the number of iterations for the training process to run through the dataset and Batch size is a number of instances that are evaluated before the weight update.

Verbose is the choice that how you want to see the output of your Neural Network while it's training.By default verbose = 1, which includes both progress bar and one line per epoch verbose = 0, means silent verbose = 2, one line per epoch i.e. epoch no./total no. of epochs.

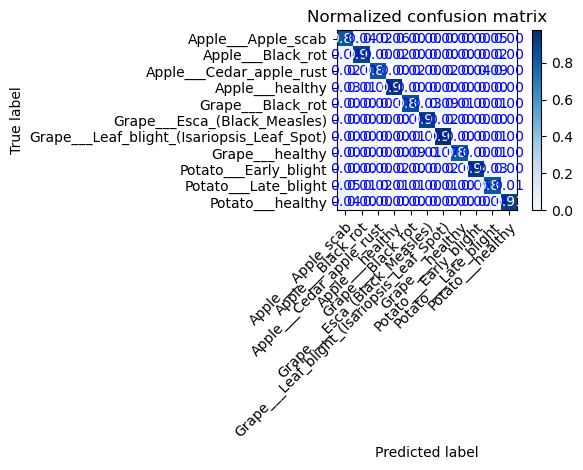


**Fig. 4.9:** Fit the Model

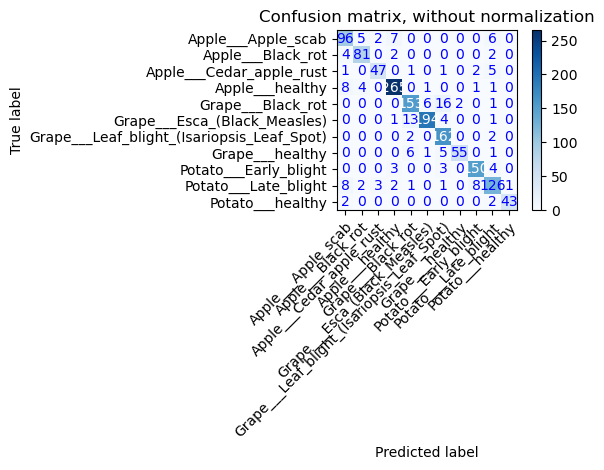
### Training and Validation Accuracy with Confusion Matrix



**Fig. 4.10:** Training and Validation Graph



**Fig. 4.10:** Confusion Matrix with Normalization

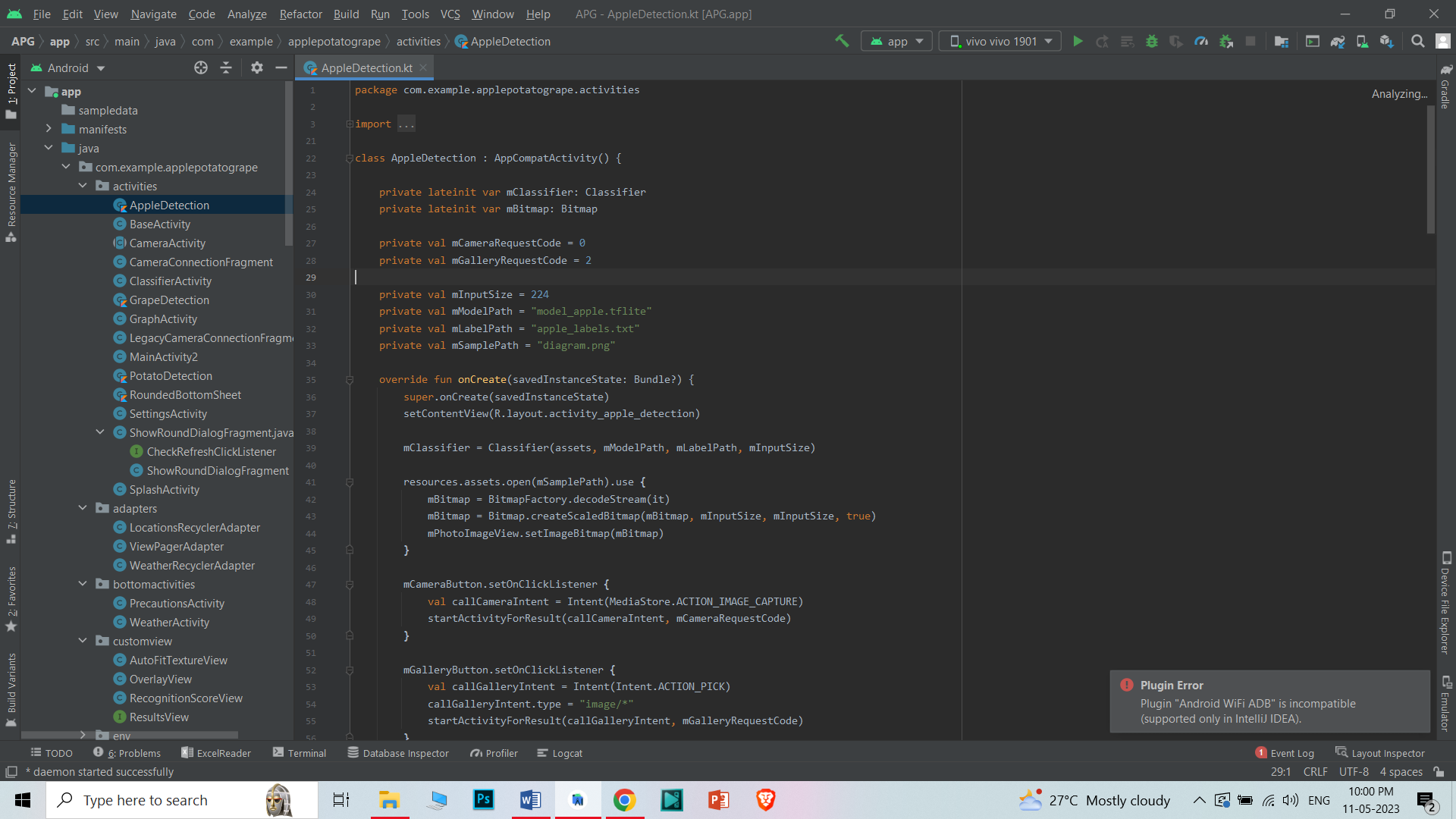


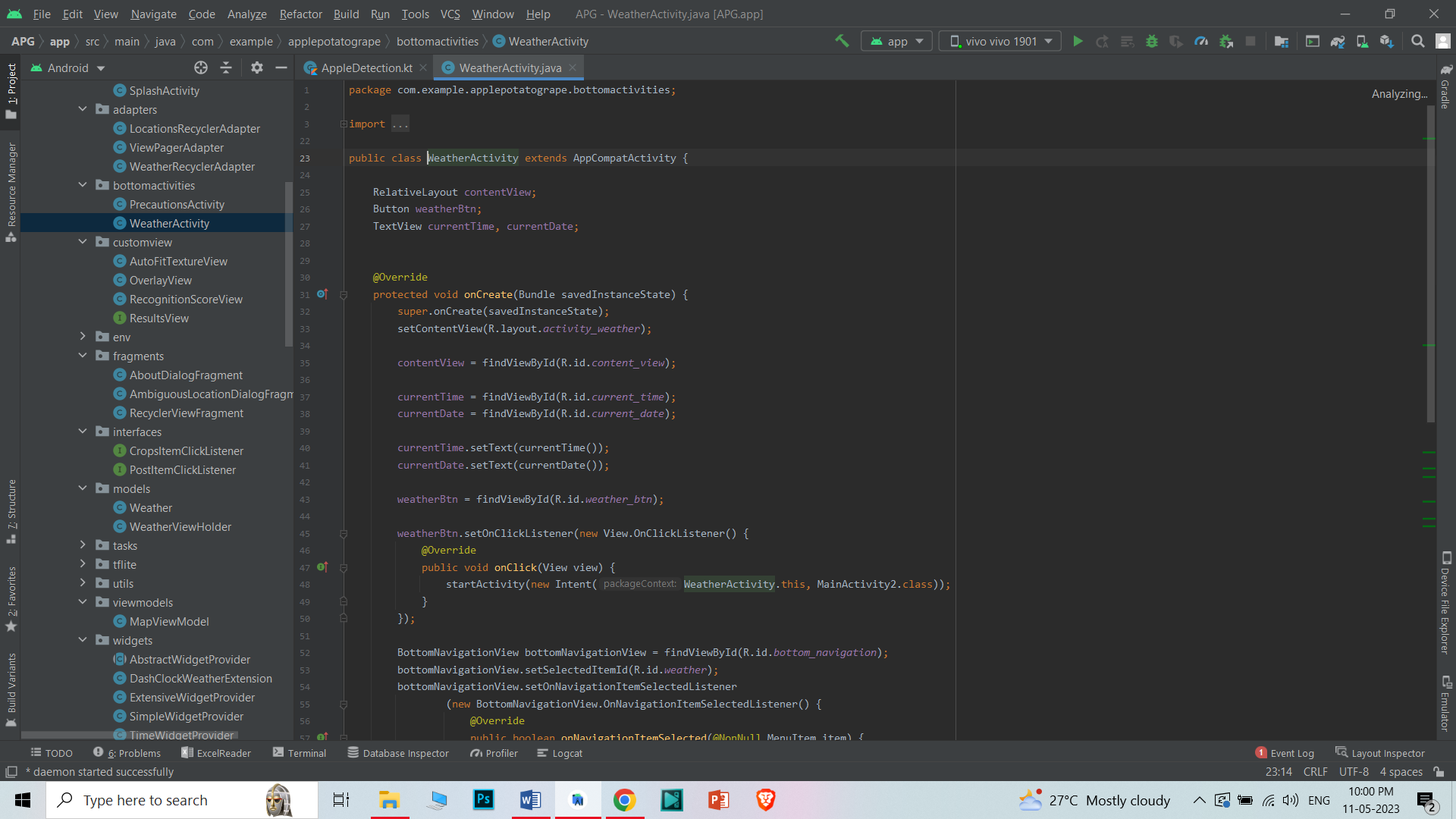
**Fig. 4.10:** Confusion Matrix without Normalization

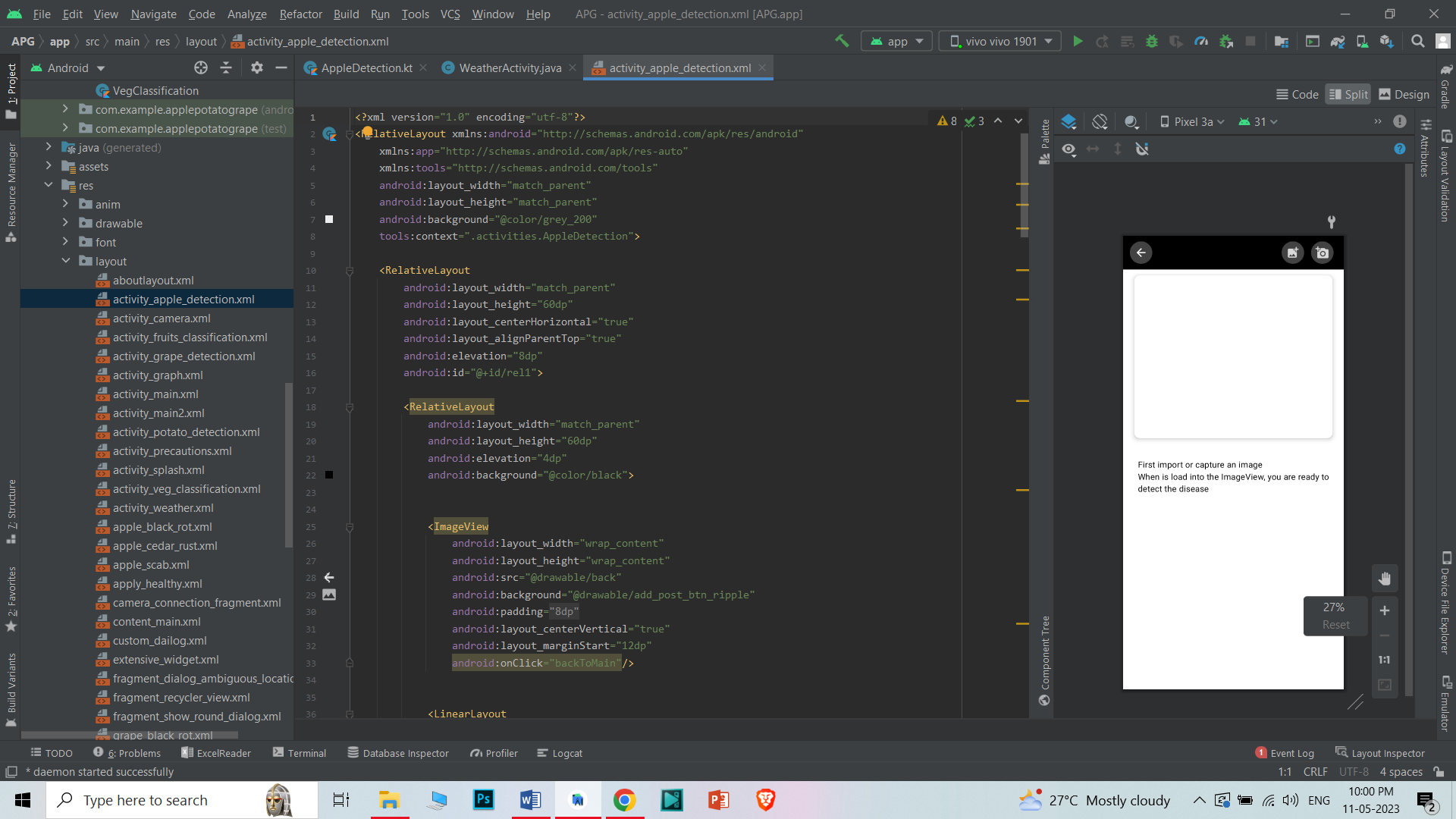
**Disease Detection for Apple and Grape is same as the Potato, Follow the same steps.**

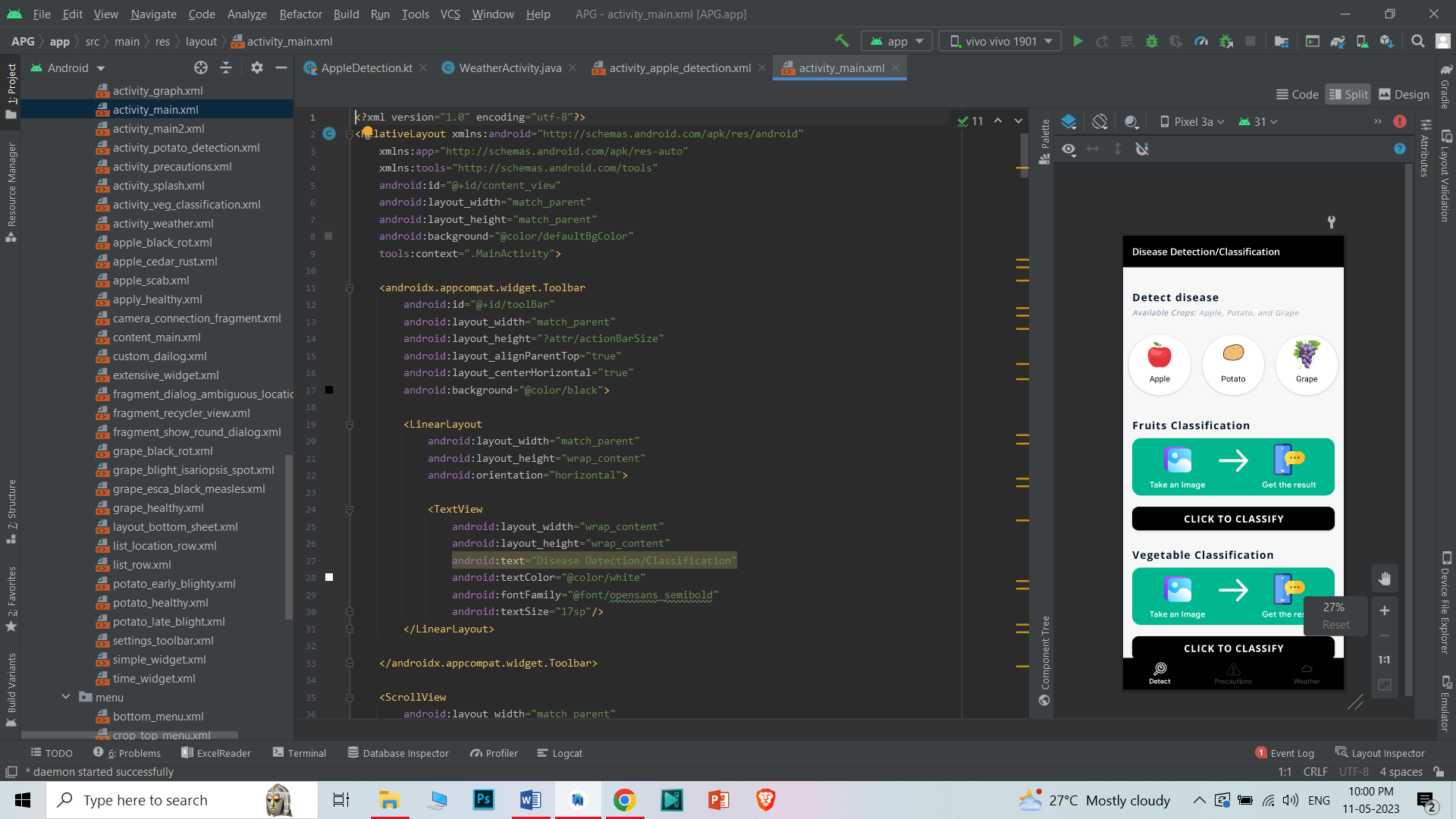
### MOBILE APPLICATION

An Android Application is build for Disease Detection and Fruits and Vegetables classification. With Weather API. Android Studio is used in the Front-end. It is Java code based application.







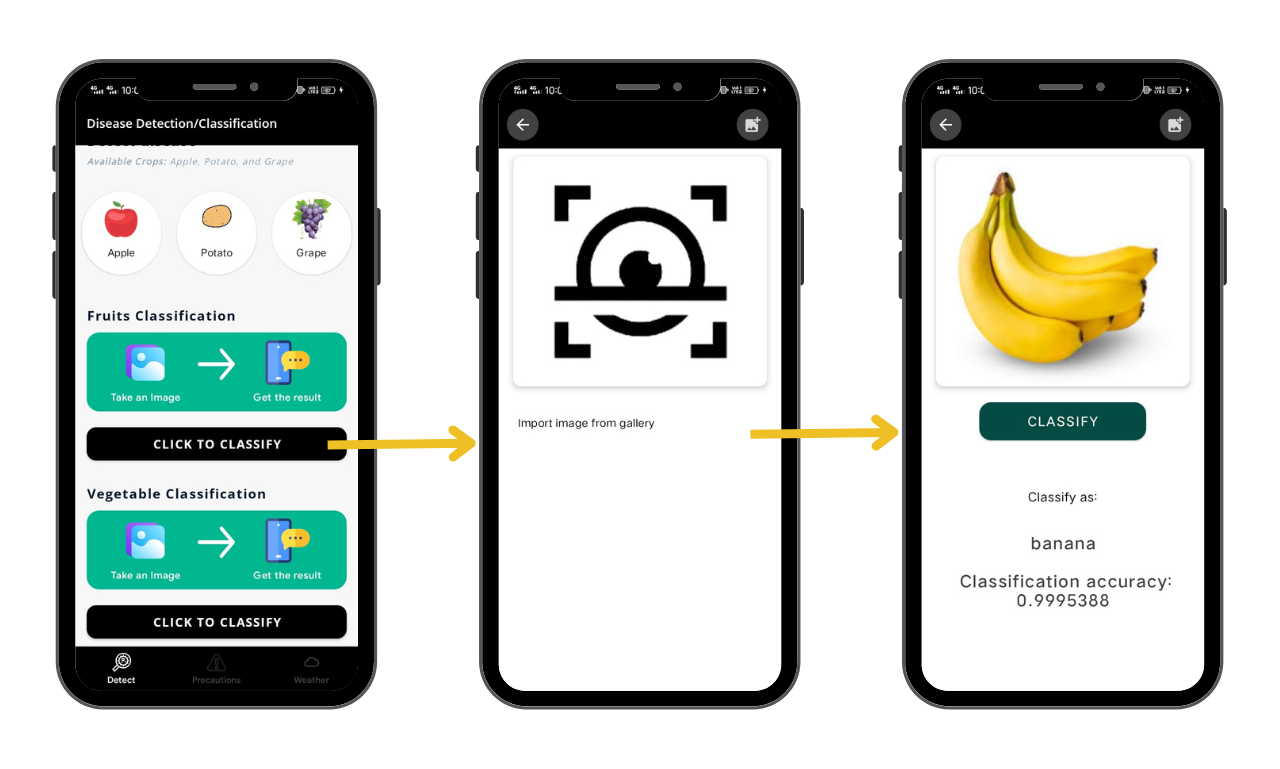


# CHAPTER 5

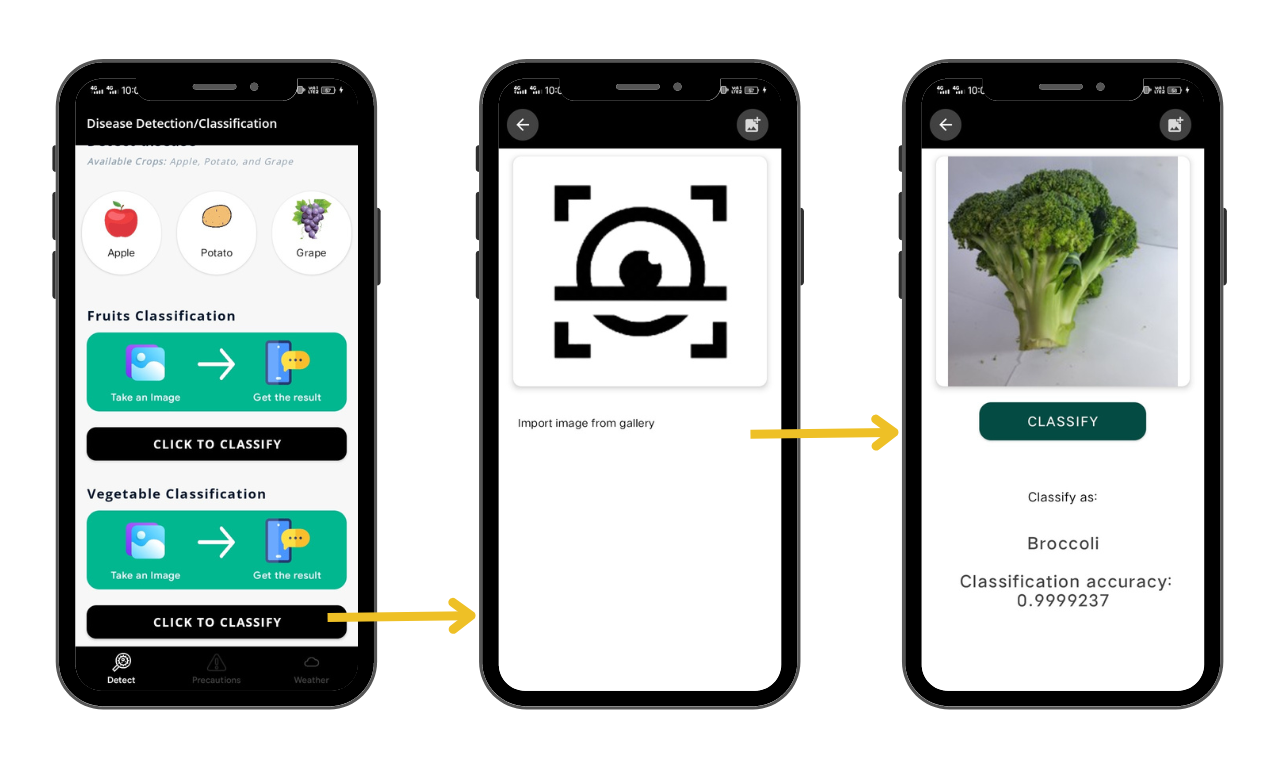
# RESULT

This chapter consists of a presentation and analysis of the data collected in the study.

## Result –Classification of Fruits and vegetables

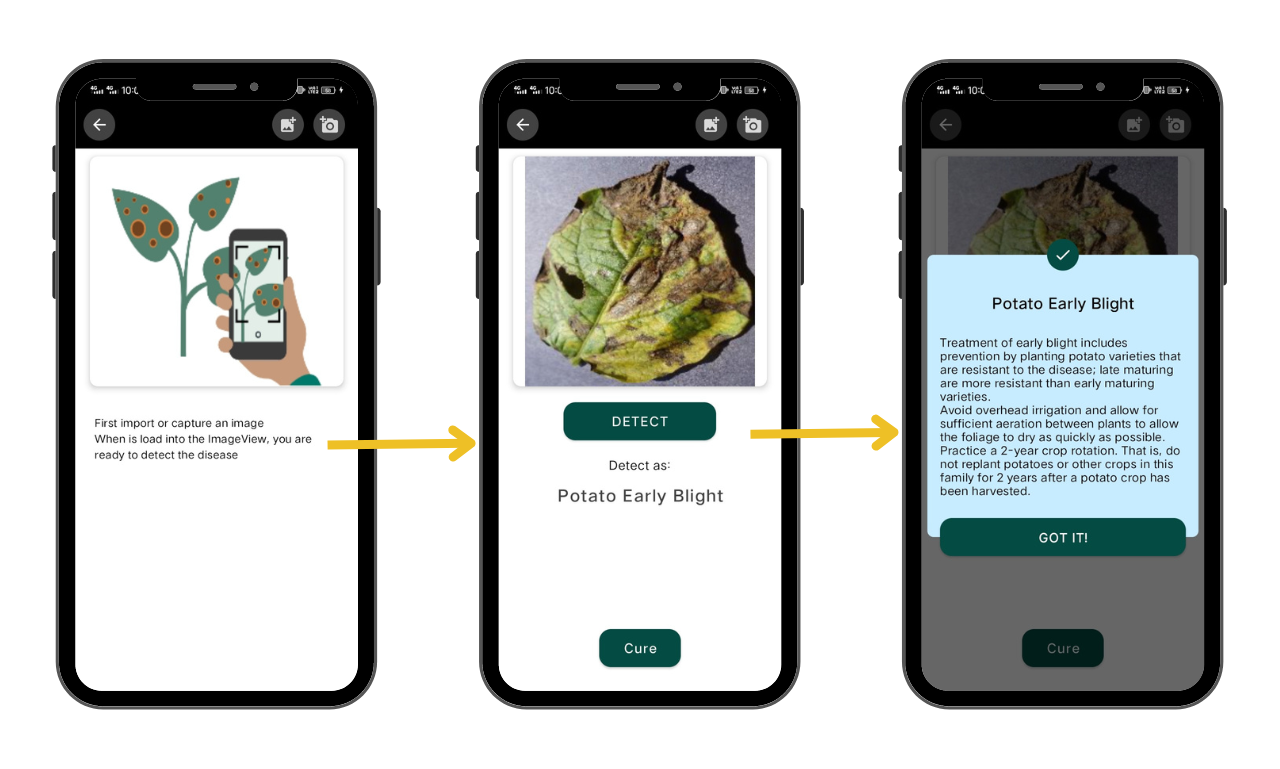


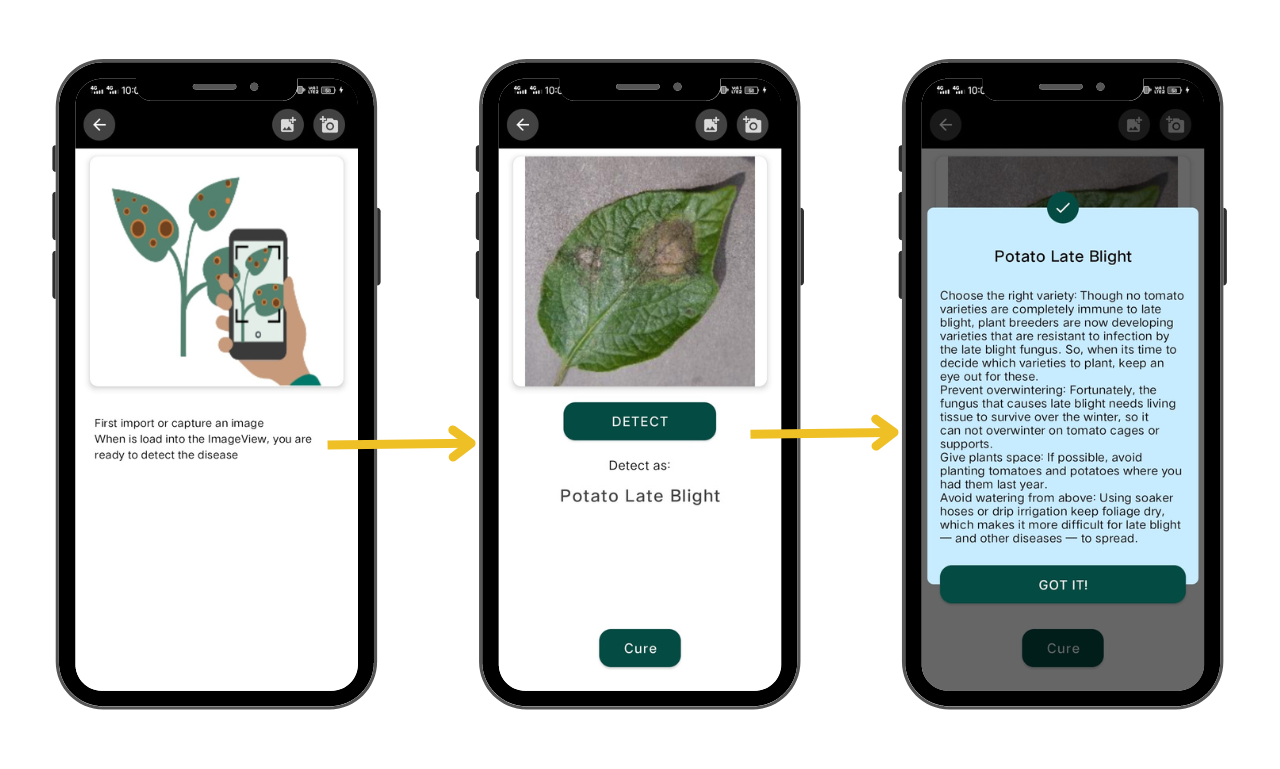
**Figure 1: Fruits Classification**



**Figure 2: Vegetable Classfication**

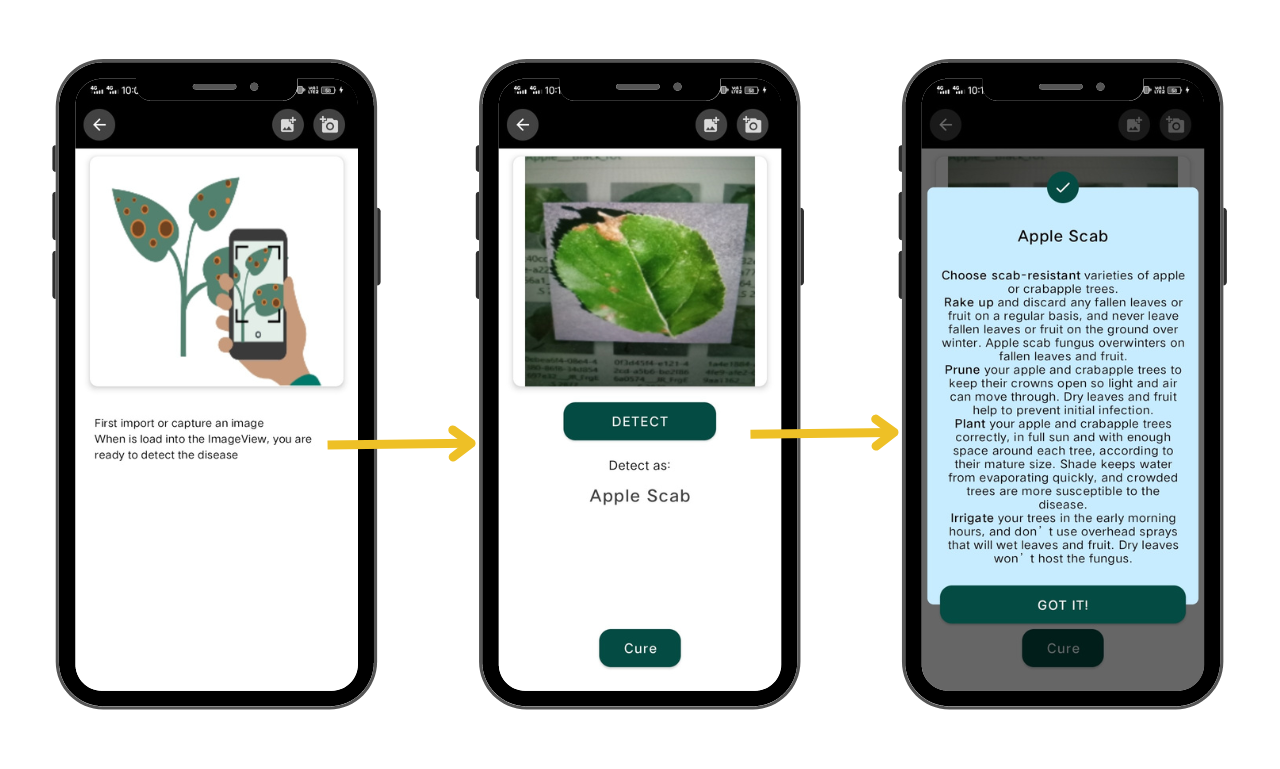
## Result – Potato disease Prediction





**Figure 3: Potato Disease Detection**

## Result – Apple disease Prediction



**Figure 4: Apple Disease Detection**

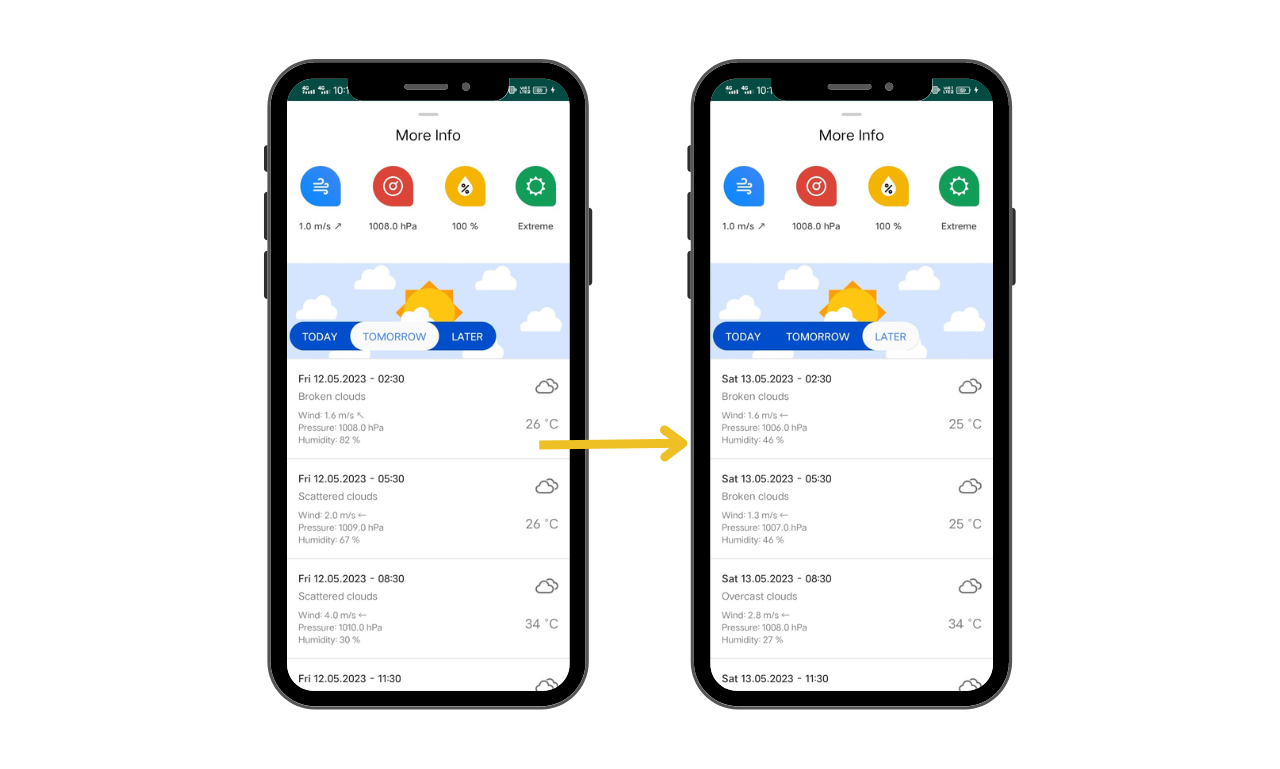
## Result – Grape disease Prediction



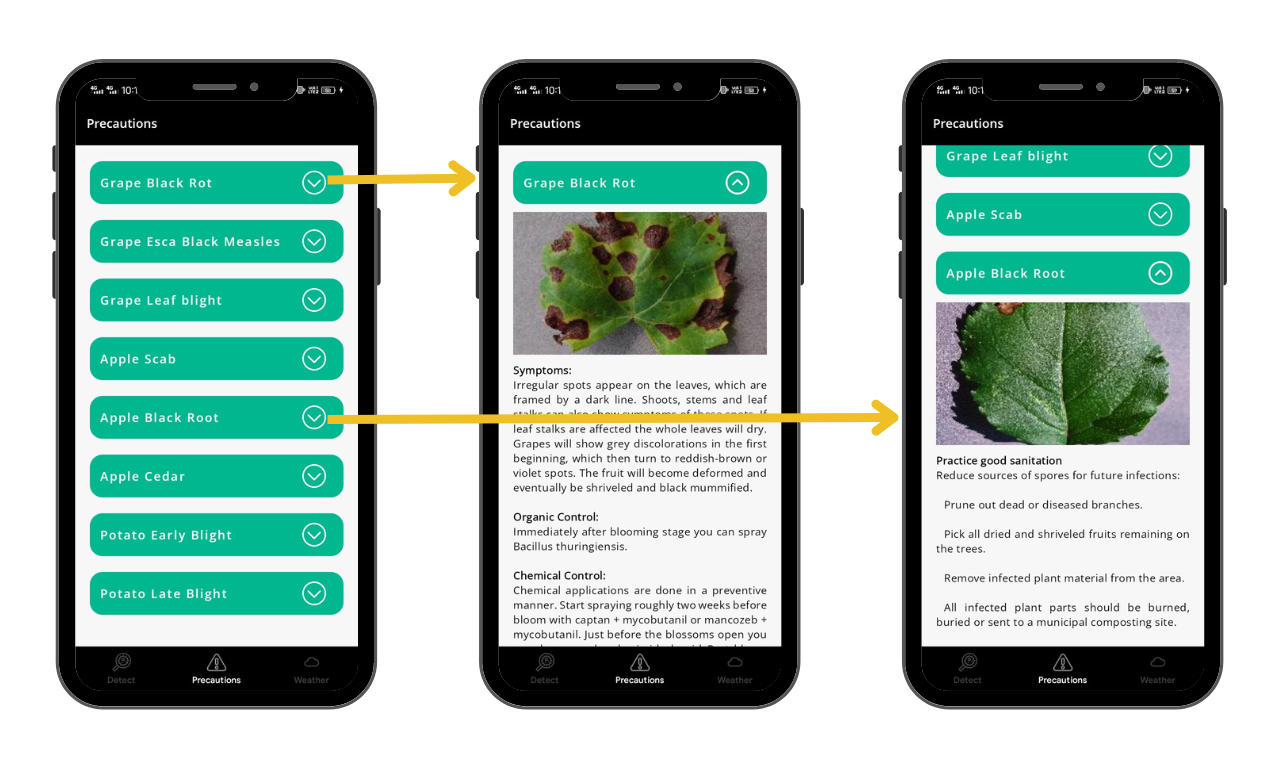
**Figure 5: Grape Disease Detection**

## Result – Weather Report





**5.6 Precautions Page**



# CHAPTER 6

**CONCLUSION & FUTURE SCOPE**

This chapter summarizes the key findings and implications of the project. Future scope outlines potential directions for future research, identifies research gaps, and suggests solutions or research methods to address these issues.

## CONCLUSION

In conclusion, the development of a smart farming system for plant disease and quality detection using machine learning techniques has been presented in this project. The system has been designed to provide accurate and efficient identification of plant diseases, which can help farmers to take timely measures to prevent the spread of diseases. Moreover, the system has also been designed to provide fruit and vegetable quality grading based on size and shape.

The implementation of this system has shown promising results in the detection and classification of various plant diseases and the grading of fruits and vegetables. The system has achieved high accuracy in disease detection and grading, which is expected to lead to increased productivity and cost savings for farmers.

The development of a user-friendly mobile application for the system is also a significant achievement in this project. The application allows farmers to use the system conveniently and access the results of disease detection and quality grading on their smartphones.

In the future, the proposed system can be further improved by incorporating additional features such as the detection of nutrient deficiencies, the prediction of yield, and the integration of sensor data from the farm environment. These enhancements can provide farmers with a comprehensive understanding of their crops' health, which can lead to better decision-making and improved crop management.

Overall, the smart farming system for plant disease and quality detection using machine learning techniques has the potential to revolutionize the agriculture industry by providing farmers with a cost-effective and efficient solution for crop management.

## FUTURE SCOPE

We believe that there is a vast potential for the future scope of our project. With the increasing demand for sustainable agriculture, the implementation of technology- based solutions has become a necessity. Our project can provide a cost-effective and efficient solution for plant disease detection and crop yield prediction. The following are some potential future scopes for our project:

Firstly, we can integrate our app with drones for monitoring large agricultural fields. The app can be used to control and guide the drones over the fields, capturing high- resolution images of crops. The images can then be processed using our algorithms to identify plant diseases and predict crop yield. This can significantly reduce the time and effort required for manual monitoring and increases the accuracy of results.

Secondly, we can expand our project to include more crop varieties and diseases. Currently, our project focuses on a few specific crops and their diseases. By including more crop varieties and diseases, we can cater to a wider audience and provide more comprehensive solutions. We can also collaborate with agricultural research institutions to gather more data and improve the accuracy of our algorithms.

Thirdly, we can incorporate weather forecasting into our app. Weather conditions have a significant impact on crop growth and health. By integrating weather forecasting, our app can provide more accurate predictions of crop yield and help farmers take necessary precautions to protect their crops from weather-related damage.

Fourthly, we can develop a feedback system for our app. The feedback system can allow farmers to report their experiences and provide feedback on the accuracy of our predictions. This can help us improve our algorithms and provide more accurate results.

Lastly, we can explore the potential of using blockchain technology to ensure the transparency and security of data. Blockchain technology can provide a decentralized system for storing and managing data, which can improve data privacy and security. It can also provide a transparent system for sharing data and transactions, which can help build trust between farmers and other stakeholders.

### References

* + 1. R. Sharma, "Classification of Carrots based on Shape Analysis using Machine Learning Techniques," *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2021, pp. 1010- 1014, doi: 10.1109/ICICCS49897.2021.9483109.
    2. A. S. Bandsode, "Automated Sorting And Grading of Fruits Using Image Processing," *in Proceedings of National conference on Changing Technology and Rural Development*, 2019, pp. 130-133.
    3. S. Gowid, "An Intelligent Hybrid Experimental-Based Deep Learning Algorithm for Tomato-Sorting Controllers," in *IEEE Access*, vol. 9, pp. 131384-131400, 2021, doi: 10.1109/ACCESS.2021.3111117.
    4. M. H. Saleem, M. Shahbaz, M. N. Jaffar and M. Farooq, "Plant Disease Detection and Classification by Deep Learning," *2020 International Conference on Frontiers of Information Technology (FIT)*, Islamabad, Pakistan, 2020, pp. 1-6, doi: 10.1109/FIT49408.2020.00008.
    5. G. Tonguç, "Fruit Grading Using Digital Image Processing Techniques,"

*Journal of Agricultural Machinery Science*, vol. 15, no. 1, pp. 1-8, 2019.

* + 1. M. Pushpavalli, "Image Processing Technique for Fruit Grading," *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 9, no. 1, August 2019, doi: 10.35940/ijeat.A1373.109119.
    2. K. Patil, "Fruits and Vegetables Grading Using Digital Image Processing Techniques," *International Journal of Scientific Research in Engineering and Management*, vol. 6, no. 1, pp. 12-17, Jan. 2022. doi: 10.46373/ijsrem.v6i1.23.
    3. S. K. Behera, "Identification, Classification & Grading of Fruits using Machine Learning & Computer Intelligence: A Review," in *Advances in Intelligent Systems and Computing*, vol. 1129, pp. 216-223, 2020, doi: 10.1007/978-981-15-5059-3\_26.
    4. Z. Buribayev, "Development of a Weed Detection System Using Machine Learning and Neural Network Algorithms," *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 8 (111), pp. 31-39, Dec. 2021.
    5. A. S. Priya, "Weed Detection using Machine Learning," *International Journal of Advance Research, Ideas and Innovations in Technology*, vol. 7, no. 2, pp. 1942-1947, 2021. doi: 10.35940/ijarait.b1141.1272
    6. M. A. Mercado Espinoza, "Using Machine Learning for Weed Identification and Yield Prediction of Strawberries," 2020.
    7. Islam, N “Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm” *International Journal of Advance Research*, 2021.
    8. C Jackulin , S. Murugavalli *“*A comprehensive review on detection of plant disease using machine learning and deep learning approaches” *Science Direct*,30 August 2022. [doi:10.1016/j.measen.2022.100441](https://doi.org/10.1016/j.measen.2022.100441)
    9. A. Almadhor, "AI-Driven Framework for Recognition of Guava Plant Diseases through Machine Learning from DSLR Camera Sensor Based High Resolution Imagery," *MDPI*, 1 June 2021.doi:[10.3390/s21113830](https://doi.org/10.3390/s21113830)
    10. Ms. Deepa, “A Machine Learning Technique for Identification of Plant Diseases in Leaves,”*6th International Conference on Inventive Computation Technologies (ICICT)*, 2021. doi:[10.1109/ICICT50816.2021.9358797](http://dx.doi.org/10.1109/ICICT50816.2021.9358797)
    11. Punitha Karthikeyan “Review on Emerging Trends in Detection of Plant Diseases using Image Processing with Machine Learning”, *International Journal of Computer Applications*,11th January 2021.doi:[10.5120/ijca2015905789](http://dx.doi.org/10.5120/ijca2015905789)