VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Jnana Sangama, Belagavi-590018, Karnataka

PROJECT Work (17ECP85)

REPORT on

"DISEASE DETECTION IN FRUITS USING DEEP LEARNING"

Submitted in partial fulfillment of the requirements for the award of degree of

BACHELOR OF ENGINEERING
in
ELECTRONICS AND COMMUNICATION ENGINEERING
by

RAJU S HOSAKOTI	**1BG17EC061**
SOMA PAVAN KUMAR	**1BG17EC079**

Under the guidance of
Mrs. PADMAJA JAIN
Assistant Professor

Vidyayāmruthamashnuthe

B.N.M. Institute of Technology

Approved by AICTE, Affiliated to VTU, Accredited as Grade A Institution by NAAC.
All UG branches – CSE, ECE, EEE, ISE &Mech.E Accredited by NBA for academic years 2018-19 to 2020-21 & valid upto 30.06.2021
Post box no. 7087, 27th cross, 12th Main, Banashankari 2nd Stage, Bengaluru- 560070, INDIA
Ph: 91-80- 26711780/81/82 Email: principal@bnmit.in, www.bnmit.org

Department of Electronics & Communication Engineering

2020 – 2021

B.N.M. Institute of Technology

Approved by AICTE, Affiliated to VTU, Accredited as Grade A Institution by NAAC.
All UG branches – CSE, ECE, EEE, ISE &Mech.E Accredited by NBA for academic years 2018-19 to 2020-21 & valid upto 30.06.2021
Post box no. 7087, 27th cross, 12th Main, Banashankari 2nd Stage, Bengaluru- 560070, INDIA
Ph: 91-80- 26711780/81/82 Email: principal@bnmit.in, www.bnmit.org

Department of Electronics and Communication Engineering

Vidyayāmruthamashnuthe

CERTIFICATE

Certified that the project work entitled **"Disease Detection In Fruits Using Deep Learning"** carried out by **Raju S Hosakoti (1BG17EC061), Soma Pavan Kumar (1BG17EC079)** bona fide students of **VIII semester** in partial fulfillment for the award of **Bachelor of Engineering** degree in **Electronics and Communication Engineering** of the Visvesvaraya Technological University, Belagavi during the year **2020-2021**. It is certified that all corrections/suggestions indicated for Internal assessment have been incorporated in the report deposited in the department library. The project report has been approved as it satisfies the academic requirements in respect of Project work prescribed for the said degree.

Mrs. Padmaja Jain	**Dr. P.A. Vijaya**	**Dr. Krishnamurthy. G.N**
Designation	Professor and Head	Principal
Dept. of ECE	Dept. of ECE	BNMIT
BNMIT, Bangalore	BNMIT, Bangalore	Bangalore

ACKNOWLEDGMENT

We would like to place on record our sincere thanks and gratitude to the concerned people, whose suggestions and words of encouragement has been valuable.

We express our heartfelt gratitude to **BNM Institute of Technology**, for giving us the opportunity to pursue Degree of Electronics and Communication Engineering and helping us to shape our career. We take this opportunity to thank **Prof. T. J. Rama Murthy**, Director, **Dr. S.Y. Kulkarni,** Additional Director**, Prof. Eishwar N Maanay,** Dean and **Dr. Krishnamurthy G.N.,** Principal for their support and encouragement to pursue this project. We would like to thank **Dr. P.A. Vijaya,** Professor and Head, Dept. of Electronics and Communication Engineering, for her support and encouragement.

We would like to thank our guide **Mrs. Padmaja Jain,** Assistant Professor**,** Dept. of Electronics and Communication Engineering**,** who has been the source of inspiration throughout our project work and has provided us with useful information at every stage of our project.

We would also like to thank our project coordinator **Mrs. Anuradha JP,** Assistant Professor**,** Dept. of Electronics and Communication Engineering, who is kind enough to extend her help for our project whenever the need arose.

Finally, we are thankful to all the teaching and non-teaching staff of Department of Electronics and Communication Engineering for their help in the successful completion of our project. Last but not the least we would like to extend our sincere gratitude to our parents and all our friends who were a constant source of inspiration.

Raju S Hosakoti

Soma Pavan Kumar

ABSTRACT

Crop cultivation plays an essential role in the agricultural field. Presently, the loss of food is mainly due to infected crops, which reflexively reduces the production rate. To identify the plant diseases at an untimely phase is not yet explored. The main challenge is to reduce the usage of pesticides in the agricultural field and to increase the quality and quantity of the production rate. It is very difficult to monitor the plant diseases manually. It requires tremendous amount of work, expertise in the plant diseases, and also require the excessive processing time. Hence, image processing is used for the detection of plant diseases. Disease detection involves the steps like image acquisition, image pre-processing, image segmentation, feature extraction and classification.

This project aims at providing cost effective and real time solution to detect the Fruit diseases. Here image processing is used for the detection of plant diseases using their Fruit images. Convolutional Neural Network (CNN) algorithm is used for feature extraction and classification.

This user-friendly system uses the concept of deep learning, that give the results with more accuracy. Hence this results in predicting the diseases so that the necessary actions to cure them can be taken immediately.

Table of Contents

List of Figures

CHAPTER - 1

INTRODUCTION

CHAPTER – 1

1.1 INTRODUCTION

An Image is a two-dimensional signal. Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. Advent of new technologies such as Digital image processing and Image analysis technology has many applications in the biological field. About 78% of the farmers are small and marginal in the country and they are poor in resources. Therefore, they are not in the position to use optimum quantity of inputs in their fruits which are essential for increasing the productivity. Most of farmers may not know the amount of fertilizer required for their fruits and thus it may lead to unbalanced use of fertilizer and they may also not know which pesticide/insecticide to be used for the diseased fruit. Hence the yield gets affected. A user-friendly software may to some extent help farmer solve the problem of a farmer to detect a disease. The farmer clicks the image of the fruit and sends it. The image is processed using the Image Processing techniques and the disease is detected. The disease is detected by our image processing software that helps the farmers to take some precautions. This may prove benefits in monitoring large orchards of fruits, and thus automatically detects the diseases as soon as they appear on fruits.

India is eminent for Agriculture that means most of the people are engaged towards agriculture industry. The agriculture industry act as a significant role in the economic sectors. Most of the plants are infected by variant fungal and bacterial diseases. Due to the exponential inclination of population, the climatic conditions also cause the plant disease. The major challenges of sustainable development are to reduce the usage of pesticides, cost to save the environment and to increase the quality. Precise, accurate and early diagnosis may reduce the usage of pesticides.

The classical approach for detection and identification of fruit diseases is based on the naked eye observation by the experts. In some developing countries, consulting experts are expensive and time consuming due to the distant locations of their availability. Automatic detection of fruit diseases is essential to automatically detect the symptoms of diseases as early as they appear on the growing fruits. Apple fruit diseases can cause major losses in yield and quality appeared in harvesting. To know what control factors to take next year to avoid losses, it is crucial to recognize what is being observed. Some disease also infects other areas of the tree causing diseases of twigs, leaves, and branches. Some common diseases of apple fruits are apple scab, apple rot, and apple blotch. Apple scabs are gray or brown corky spots. Apple rot infections produce slightly sunken, circular brown or black spots that may be covered by a red halo. Apple blotch is a fungal disease and appears on the surface of the fruit as dark, irregular, or lobed edges.

The classical approach for detection and identification of fruit diseases is based on the naked eye observation by the experts. In some developing countries, consulting experts are expensive and time consuming due to the distant locations of their availability. Automatic detection of fruit diseases is essential to automatically detect the symptoms of diseases as early as they appear on the growing fruits. For example, in the case of Apple fruit, diseases can cause major losses in yield and quality appeared in harvesting. To know what control factors to take next year to avoid losses, it is crucial to recognize what is being observed. Some disease also infects other areas of the tree causing diseases of twigs, leaves, and branches. Some common diseases of apple fruits are apple scab, apple rot, and apple blotch.

K-means clustering algorithm used to cluster and extract the surface defect features of apples to classify the quality of the apples. The disadvantage of this method is its low accuracy; Fruit quality detection based on traditional machine vision has disadvantages such as low accuracy, slow speed, and weak generalization ability. With the rapid development of deep learning, deep convolutional networks have been widely used in the field of target detection and have shown great advantages in detection speed and accuracy.

1.2 MOTIVATION

There is a lack of usage of modern technology in the agricultural sector in India. This project aims to introduce in its own dimension the usage of modern technology in agriculture in India. A lot of yield is unnecessarily destroyed due to the lack of early detection of diseases in crops. Existing technologies are less inaccurate in detecting the diseases; hence this project uses Deep Learning in which the accuracy would be more. Fruits such as apples, grapes, mangoes, etc., are considered as cash crops and bring large profits to farmers in our country. Hence disease detection in fruits will help the farmer in a very significant way in early detection of the disease in the fruit and hence help the farmer in procuring a profit.

In India there has been a steady increase in the production of fruits by farmers. Hence there is a need to use image processing techniques to increase the actual yield in fruits, that is to decrease the wastage due to diseases caused by pathogens and pests. The graph in the Fig 1.1 below shows the steady increase in the production of fruits in India.

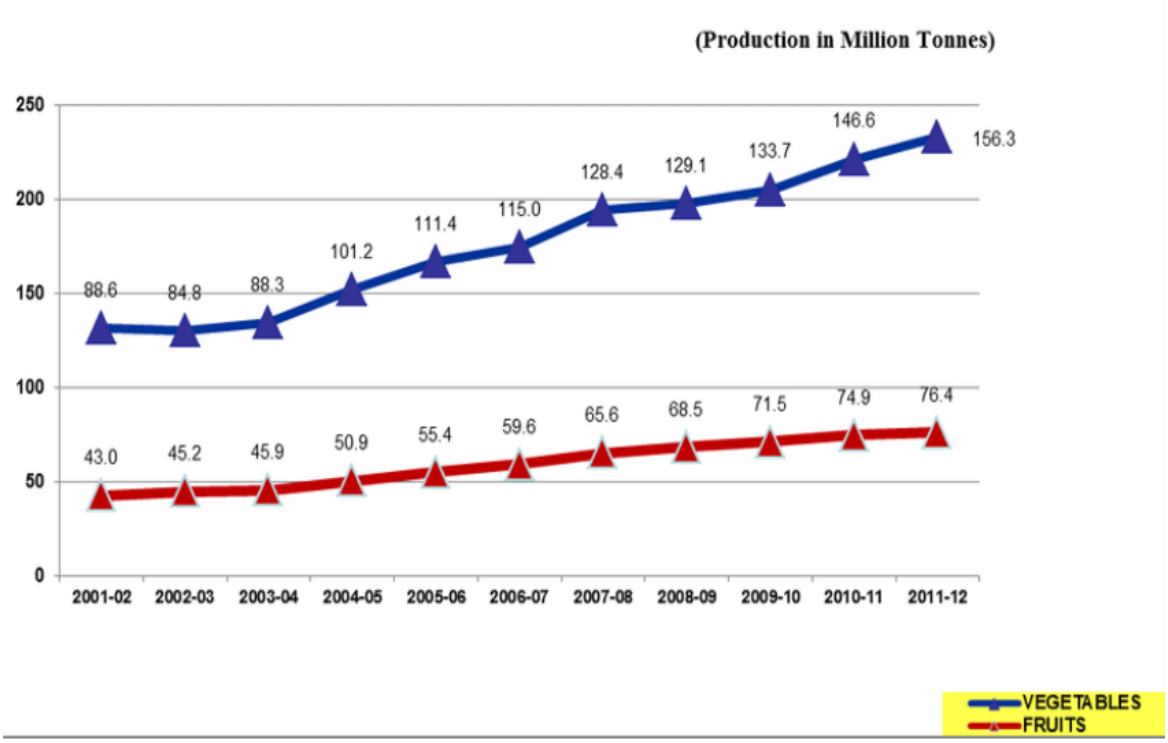

Fig 1.1 Production of Fruits and Vegetables in India

1.3 OBJECTIVE AND SCOPE

Disease detection in fruits is very important as it helps the farmer in getting a good yield. This project is focused on the early detection of diseases in fruits and hence increase the options available to the farmer in taking precautions to protect his produce. Usage of modern technology in agriculture is the main objectives of this project and the implementation of Image Processing using Deep Learning helps in obtaining accurate results or detection.

"Agriculture is the backbone of the Indian Economy"- said Mahatma Gandhi seven decades ago. Even today, the situation is still the same, with almost the entire economy being sustained by agriculture. It contributes 16% of the overall GDP and accounts for the employment of approximately 52% of the Indian population. Rapid growth in agriculture is essential not only for self-reliance but also to earn valuable foreign exchange. It is felt that with provision of timely and adequate inputs such as fertilizers, seeds, pesticides and by making available affordable agricultural credit /crop insurance, Indian farmers are going to ensure food and nutritional security to the Nation. Hence the detection of diseases in fruits is an important aspect in increasing the yield of the fruit. By detecting the fruit disease, the farmer can take the safety precautions that increases the fruit yield which leads in growth of country's economy.

CHAPTER – 2

LITERATURE SURVEY

CHAPTER-2

LITERATURE SURVEY

PAPER-1

Hongjun wang, Qisong Mou, Youjun Yue, Hui Zhao: "Research on Detection Technology of various Fruit Disease Spots Based on Mask R-CNN", Tianjin University of Technology, Tianjin China, 2016

This research talks about the implementation of a special type of CNN called Mask R-CNN in order to increase the accuracy of spot and rot detection in apples, pears, oranges and pears. Mask R-CNN is an improved target detection algorithm based on Faster R-CNN. According to the Hongjun, Qisong, Youjun and Hui- "Although the detection model proposed in the article has high detection accuracy, the detection speed is slightly slower."

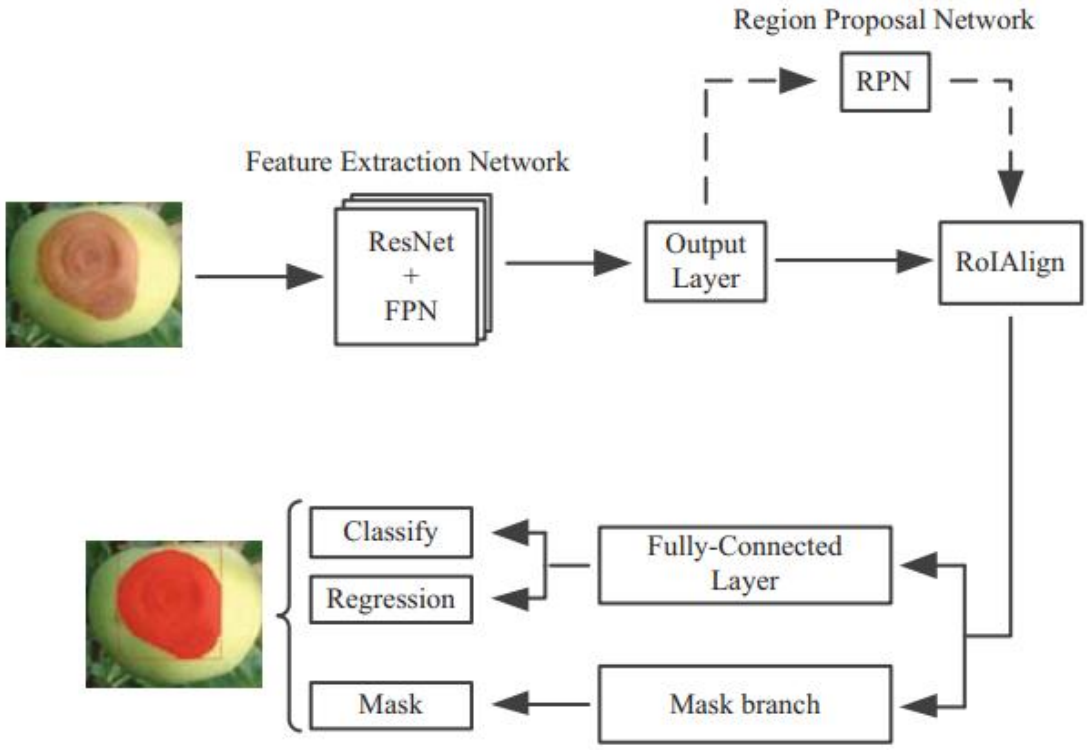

Fig 2.1 Mask R-CNN Model Architecture

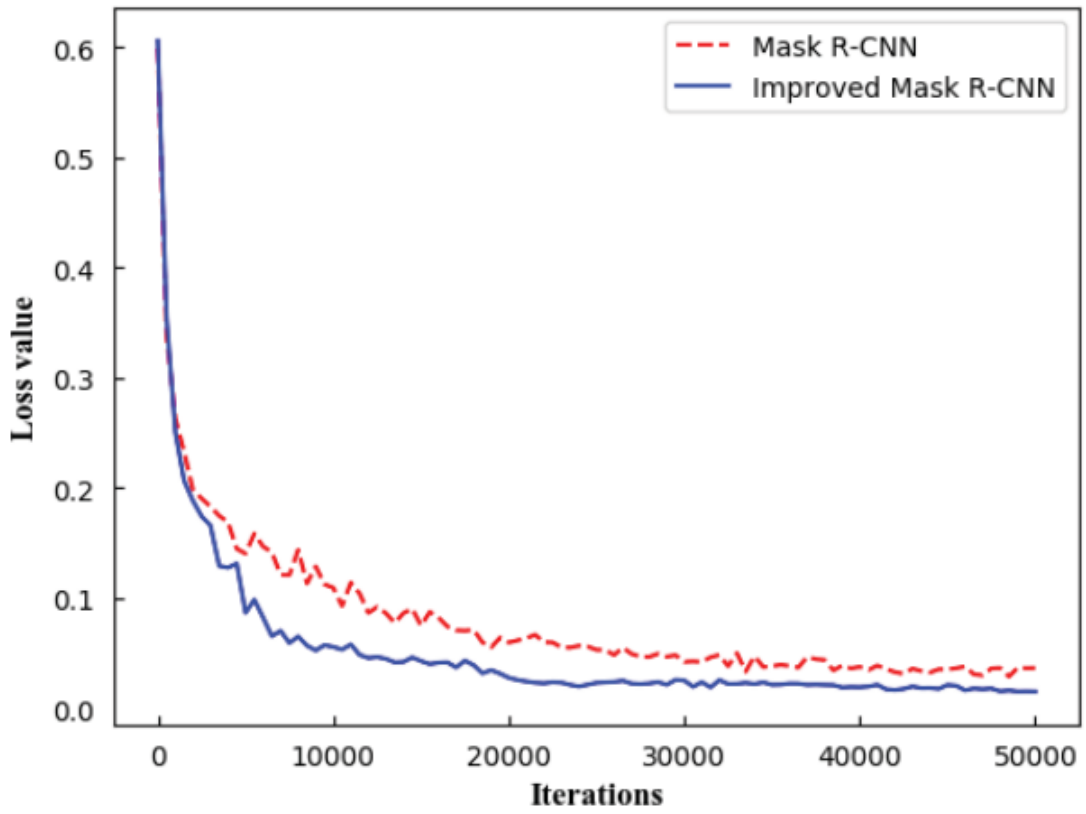

Fig 2.2 Convergence curve of loss function

PAPER-2

Shiv Ram Dubey, A.S Jalal: "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Pattern", December 2012.

In this research, image processing is used to detect diseases in apples. The processing is divided in to two steps- 1.K-means clustering technique (used for segmentation) 2. Some state-of-the-art features are extracted from the image and are classified into various classes using Multi class support vector machines. The classification accuracy of the proposed solution is achieved up to 90%. K-Means clustering technique is used for the image segmentation. Images are partitioned into four clusters in which one cluster contains majority of the diseased part of the image. K-Means clustering algorithm was developed by J. MacQueen (1967). The k-means clustering algorithms classify the objects (pixels in our problem) into K number of classes based on a set of features. The

classification is carried out by minimizing the sum of squares of distances between the data objects and the corresponding cluster. In this experiment, squared Euclidean distance is used for the K means clustering.

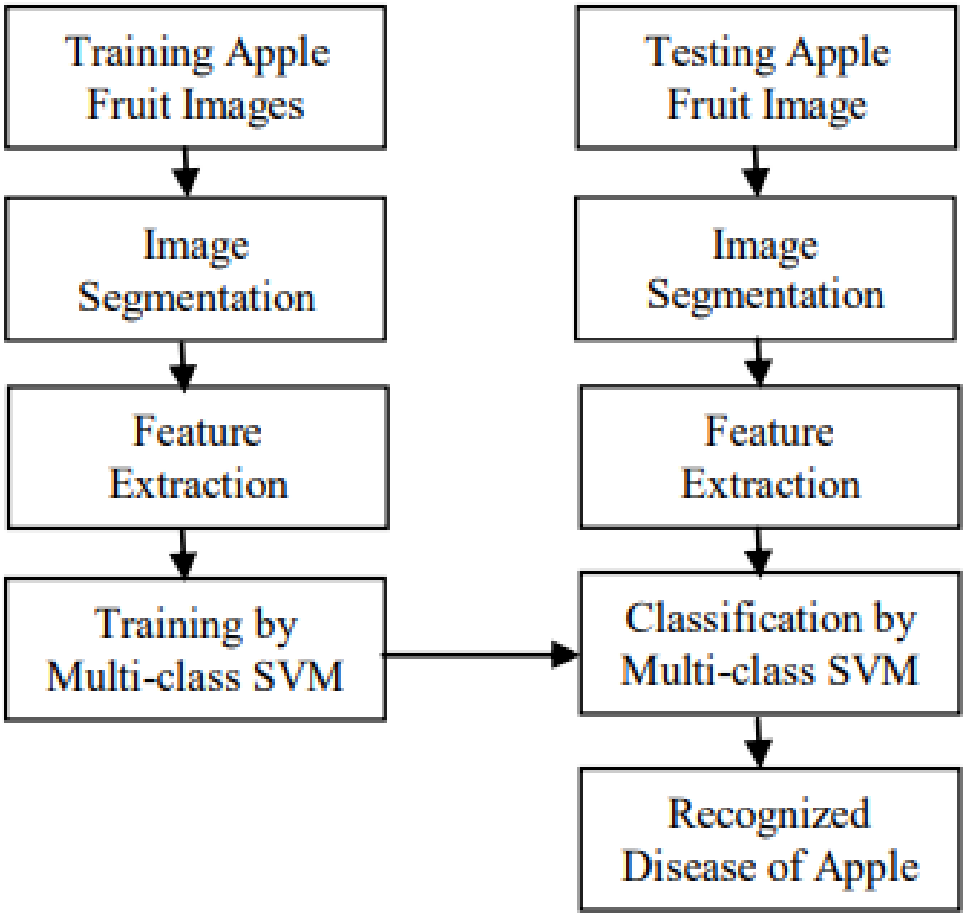

Fig 2.3 Framework of the proposed approach

PAPER-3

Nivedita.R., Kakade, Dnyaneswar.D.Ahire: "Real Time Grape Leaf Disease Detection", International Journal of Advance Research and Innovative Ideas in Education, Vol-1 Issue-4, 2015.

Here they developed a system for automatic disease detection in grape leaves. Here the system is tested on 5 diseases which effect on plant, they are: Black rot, Downy mildew,

Powdery mildew, normal and leaf roll. In this approach the leaf images used are captured using high resolution camera. The image color conversion, histogram, histogram equalization techniques are applied on the images to improve quality and clarity before feature extraction. Later the images are segmented using canny edge detection methods. The Grey co-occurrence Matrix (GLCM) method is used for feature extraction, finally back propagation neural network is used to classify the images for different diseases. This approach consumes more time because for complex computations in canny edge detection, hence it fails in real time response.

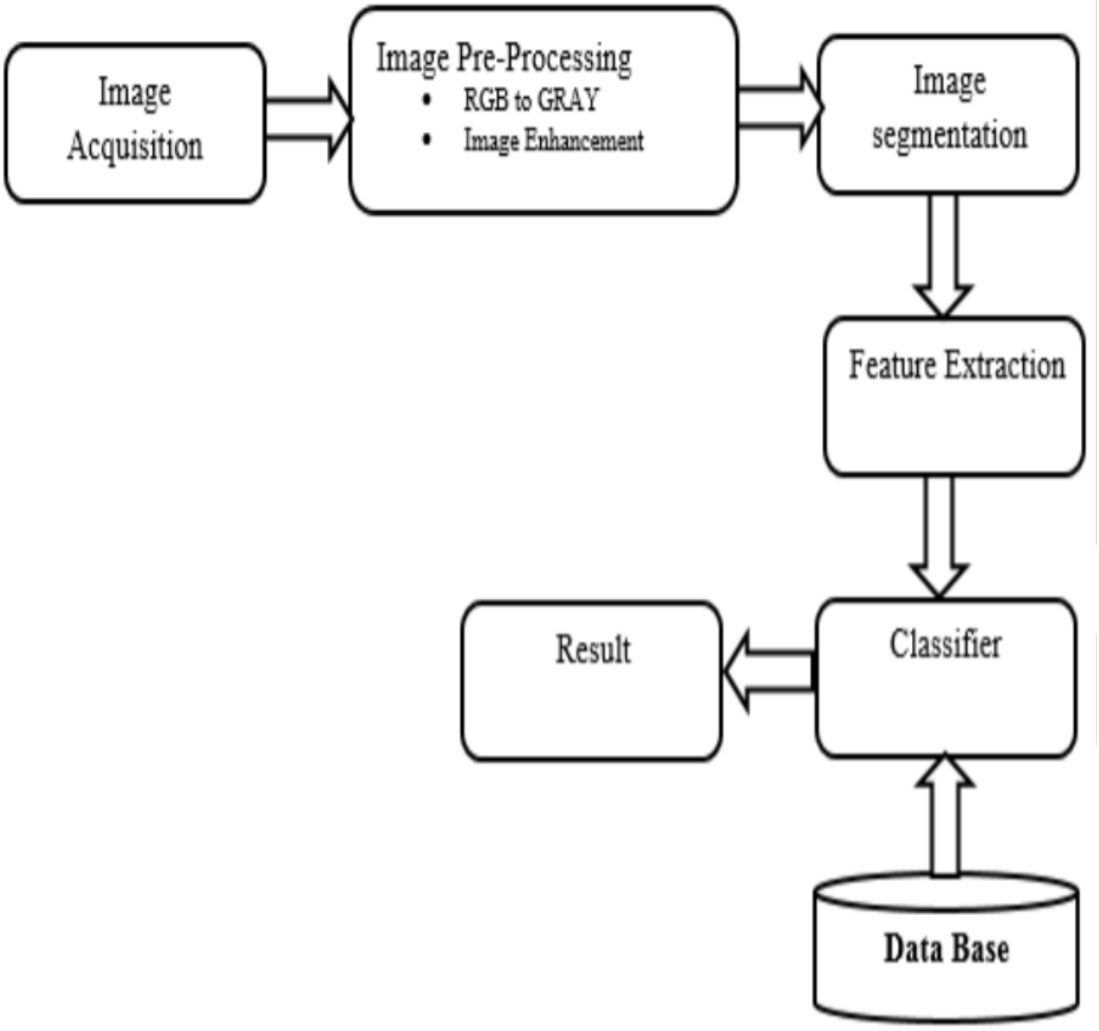

Fig 2.4 Basic procedure for the software approach

PAPER-4

Kulkarni Anand H, Ashwin Patil RK: "Applying image processing technique to detect plant diseases". Int J Mod Eng Res, 2012

This paper presents a methodology for early and accurately plant diseases detection, using artificial neural network (ANN) and diverse image processing techniques. As the proposed approach is based on ANN classifier for classification and Gabor filter for feature extraction, it gives better results with a recognition rate of up to 91%. An ANN based classifier classifies different plant diseases and uses the combination of textures, color and features to recognize those diseases. This paper presents an algorithm for image segmentation technique which is used for automatic detection and classification of plant leaf diseases. It also covers survey on different diseases classification techniques that can be used for plant leaf disease detection. Image segmentation, which is an important aspect for disease detection in plant leaf disease, is done by using genetic algorithm.

PAPER-5

Sue Han Lee, Chee Seng Chan, Paul Wilkin, Paolo Remagnino: "Deep-plant: plant identification with Convolutional neural networks", rearXiv**,** 2015

Here they use a convolutional neural networks (CNN) in order to learn unsupervised feature version for 44 various types of plant classes, collected at the Royal Botanic Gardens, Kew, England. To grow perception on the selected features from the CNN model, a visualization method based on the de-convolutional networks (DN) have been used. Experimental results by means of these CNN features with number of classifiers demonstrate reliability and supremacy compared to the fantasy solutions which solely depend upon hand-crafted features. For development of accurate image classifiers system for the plant disease recognition, we are in need of a large, confirmed dataset of images of diseased and healthy plants.

PAPER-6

Halil Durmu, Ece Olcay Güne, Murvet Kirci: "Disease Detection on the Leaves of

Tomato Plants by Using Deep Learning", International Journal of Computer Applications, vol. 52, no. 2, 2012.

use deep learning to detect the various diseases on the leaves of tomato plants. The aim of this work is to run deep learning algorithm in real time on the robot. So, the robot will be able to detect the diseases of the plants while wandering manually or autonomously on the field or in the greenhouse. The examined diseases in this study cause physical changes in the leaves of the tomato plant. These changes on the leaves can be seen with RGB cameras. Deep learning architecture selection was the key issue for the implementation. Two different deep learning network architectures were tested first Alex Net and then Squeeze Net. For both of these deep learning networks training and validation were done on the Nvidia Jetson TX1. Tomato leaf images from the Plant Village dataset has been used for the training. Ten different classes including healthy images are used. Trained networks are also tested on the images from the internet.

PAPER-7

Sachin D. Khirade and A. B. Patil: "Plant disease detection using image processing," in Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on, IEEE, 2015

They developed a system for plant disease detection using image processing technique. Various operations applied on the plant images include pre- processing to remove noise in image or other object removal, Image clipping. Image smoothing is done using the smoothing filter. Image enhancement is carried out for increasing the contrast. Later pre-processed images are segmented using K-means and Otsu threshold algorithms. The infected leaf shows the symptoms of the disease by changing the color of the leaf. Hence the greenness of the leaves can be used for the detection of the infected portion of the leaf. The R, G and B component are extracted from the image. The threshold is calculated using the Otsu's method. Then the green pixels are masked and removed if the green pixel intensities are less than the computed threshold. The image features are extracted

using grey level co-occurrence matrix (GLCM). Finally, images are classified using ANN.

PAPER-8

S. B. Ullagaddi and S.Viswanadha Raju: "Disease Recognition in Mango Crop Using Modified Rotational Kernel Transform Features". 2017 International

Here they use Modified Rotation Kernel Transformation (MRKT) based directional feature extraction scheme is presents to resolve the issues occurring due to shape, color, or other deceptive features during plant disease recognition. The MRKT based scheme is used to calculate the directional features and histograms for plant parts like leaf and fruits. These histograms and the directional feature set with use of artificial neural network lead to a better recognition technique of Anthracnose disease which occurs in form of black spots on fruits and leaves of mango plant.

PAPER-9

Jamil Ahmad, Bilal Jan, Haleem Farman, Wakeel Ahmad, and Atta Ullah: "Disease Detection in Plum using Convolutional Neural Network under true field conditions".28 September 2020.

In this paper, they propose an efficient convolutional neural network-based disease detection framework in plum under true field conditions for resource-constrained devices. As opposed to the publicly available datasets, images used in this study were collected in the field by considering important parameters of image-capturing devices such as angle,

Scale, orientation, and environmental conditions, Furthermore, extensive data augmentation was used to expand the dataset and make it more challenging to enable robust training. Variations in scale, orientation, and position of the objects of interest in the images make predictions manner. One of the most popular approach that they used was data augmentation. It is a process of generating more images from existing ones by applying similarity -preserving transformations.

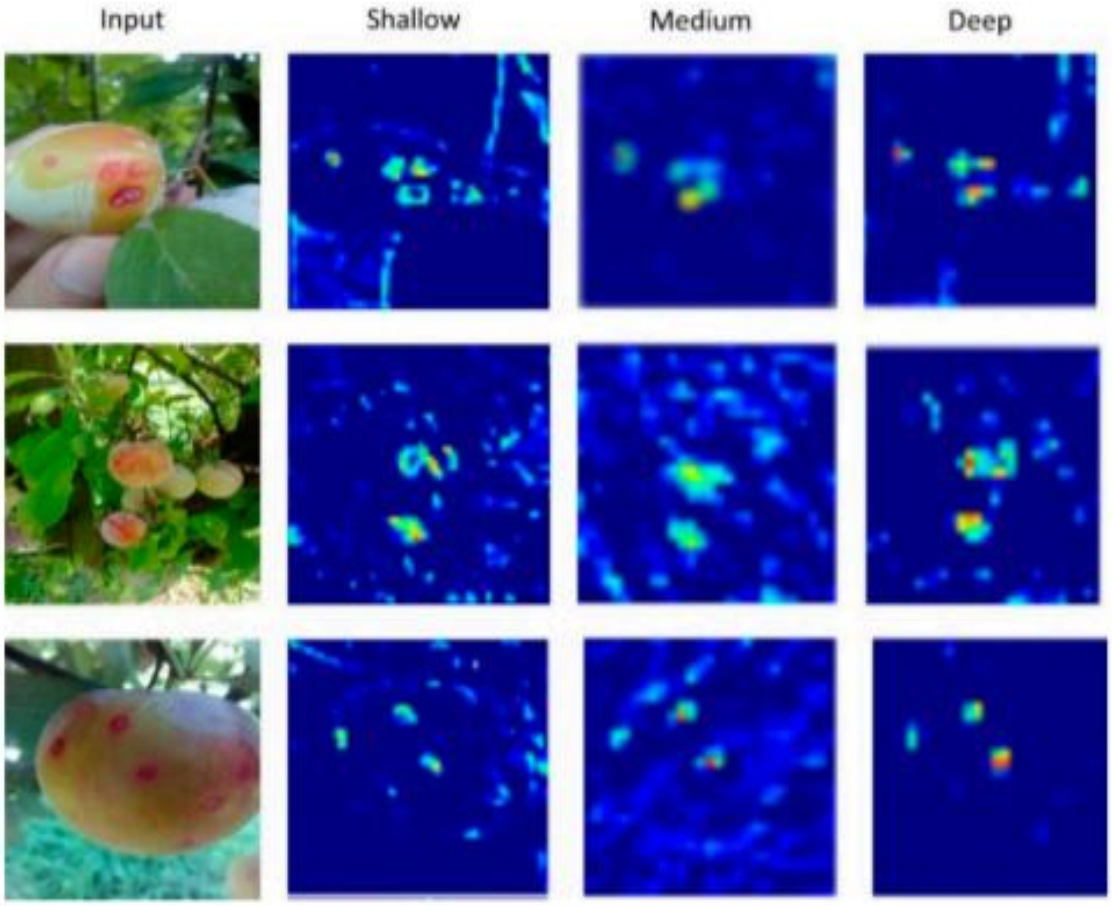

Fig 2.5 Activation response of various convolutional layers to lesions

Convolutional neural networks (CNNs) have been widely used for visual recognition tasks like image classification, object detection and recognition, image matching, image in-painting, and a variety of other similar challenging tasks. Its superior capability to automatically identify useful features from raw image data makes it a compelling choice for computer vision researchers. As a result, significant research has been conducted on the development and use of CNNs for various challenging visual recognition tasks.

CHAPTER – 3

SYSTEM DESIGN

CHAPTER-3

SYSTEM DESIGN

3.1 FLOW DIAGRAM

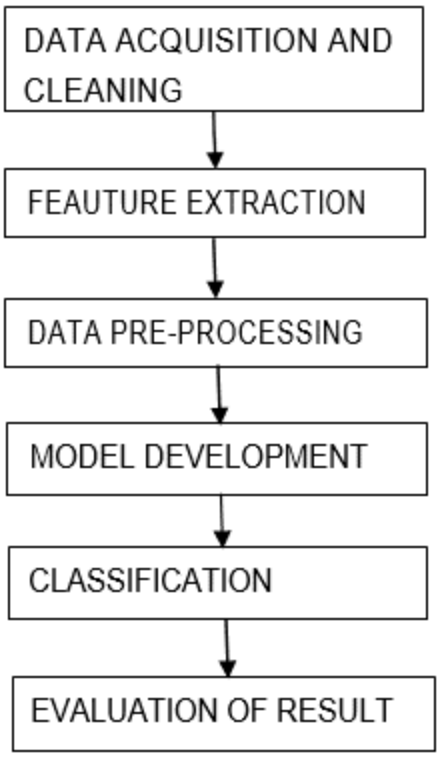

Fig 3.1 Proposed Flow Chart

3.1.1 Data Set Preparation

A data set (or dataset) is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds to a given record of the data set in question. The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set. Each value is known as a datum. Data sets can also consist of a collection of documents or files.

In the open data discipline, data set is the unit to measure the information released in a public open data repository. The European Open Data portal aggregates more than half a million data sets. In this field other definitions have been proposed, but currently there is not an official one. Some other issues (real-time data sources, non-relational data sets, etc.) increases the difficulty to reach a consensus about it. These datasets are used for machine-learning research and have been cited in peer-reviewed academic journals. Datasets are an integral part of the field of machine learning. Major advances in this field can result from advances in learning algorithms (such as deep learning), computer hardware, and less-intuitively, the availability of high-quality training datasets. High-quality labelled training datasets for supervised and semi-supervised machine learning algorithms are usually difficult and expensive to produce because of the large amount of time needed to label the data. Although they do not need to be labelled, high-quality datasets for unsupervised learning can also be difficult and costly to produce.

To demonstrate the performance of the proposed approach, we have used a data set of normal and diseased apple fruits, which comprises four different categories: Apple Blotch, Apple rot, Apple scab, and Normal Apple. Fig. 6 depicts the classes of the data set. Presence of a lot of variations in the type and colour makes the data set more realistic.

Apple scabs are gray or brown corky spots it is caused by the ascomycete fungus Venturia inaequalis. This disease affects several fruit genera, including sorbus, cotoneaster, and pyrus. Apple rot infections produce slightly sunken, circular brown or black spots that may be covered by a red halo, the rot spreads progressively and releases enzymes. The principle enzymes involved in this are pectic enzymes – those that break down the gel-like pectic compounds that cement the apple cells together. Apple blotch is a fungal disease and appears on the surface of the fruit as dark, irregular, or lobed edges. The blotches are cosmetic damage and are unacceptable to costumers, it downgrades the fruit from premium fresh-market grade to processing use (reduces its market value), however the fruit development is not affected.

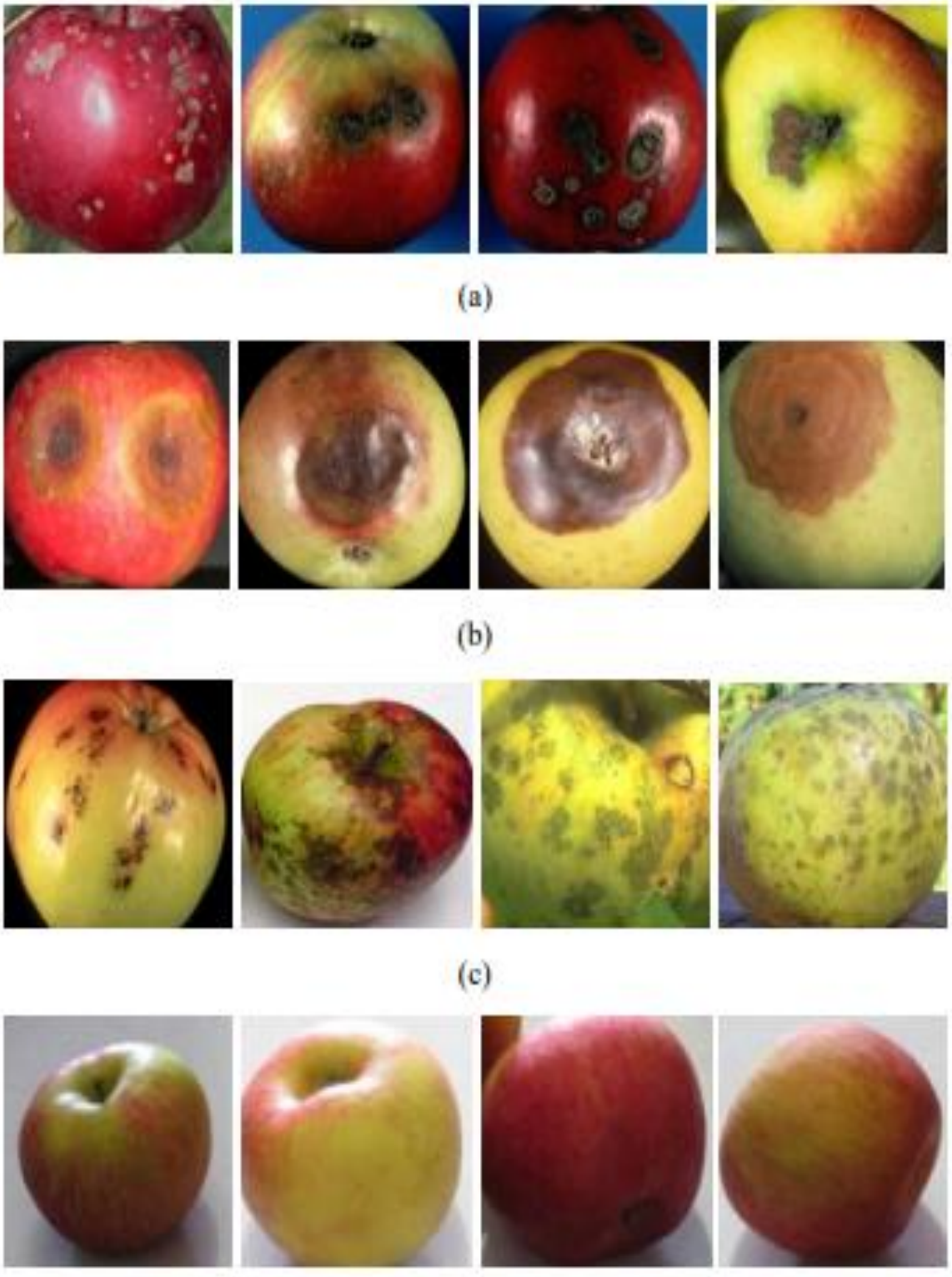

Fig 3.2 Sample Images from Dataset (a) Apple scab, (b) Apple rot, (c) Apple blotch, and (d) Normal apples

3.1.2 Image Acquisition

Image acquisition is the first method of digital image processing and it is described as capturing the image through digital camera and stores it in digital media for further operations. In our work, we collected images of fruits(apples) for image processing. The general aim of image acquisition is to transform an optical image (real world data) into an array of numerical data which could be later manipulated on a computer, before any video or image processing can commence an image must be captured by camera and converted into a manageable entity.

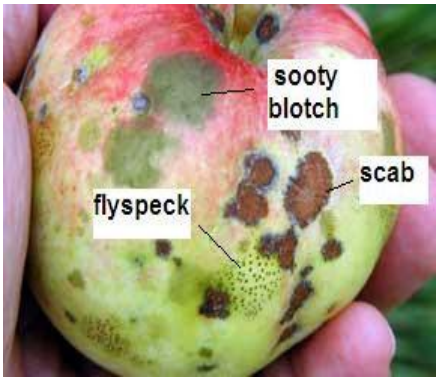

Fig 3.3 Diseased Fruit (Apple)

Some common diseases of apple fruits are apple scab, apple rot, and apple sooty blotch as shown in Fig 3.3. Apple scabs are gray or brown corky spots. Apple rot infections produce slightly sunken, circular brown or black spots that may be covered by a red halo. Apple blotch is a fungal disease and appears on the surface of the fruit as dark, irregular, or lobed edges.

3.1.3 Image Pre-processing

The main purpose of image pre-processing is to improve the image data contained unwanted distortions or to enhance some image features for further processing. Pre-processing method uses various techniques such as changing image size and shape, filtering of noise, image conversion, enhancing image and morphological operations. In this work, we used Python code to resize image and RGB to grayscale conversion.

Pre-processing is a common name for operations with images at the lowest level of abstraction - both input and output are intensity images. These iconic images are of the same kind as the original data captured by the sensor, with an intensity image usually represented by a matrix of image function values (brightnesses). The aim of pre-processing is an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing, although geometric transformations of images (e.g. rotation, scaling, translation) are classified among pre-processing methods here since similar techniques are used. Image pre-processing methods are classified into four categories according to the size of the pixel neighbourhood that is used for the calculation of a new pixel brightness. Some authors classify image pre-processing methods differently into image enhancement, covering pixel brightness transformations (local pre-processing in our sense), and image restoration. Image pre-processing methods use the considerable redundancy in images. Neighbouring pixels corresponding to one object in real images have essentially the same or similar brightness value, so if a distorted pixel can be picked out from the image, it can usually be restored as an average value of neighbouring pixels.

Image pre-processing methods use the considerable redundancy in images. Neighbouring pixels corresponding to one object in real images have essentially the same or similar brightness value, so if a distorted pixel can be picked out from the image, it can usually be restored as an average value of neighbouring pixels.

If pre-processing aims to correct some degradation in the image, the nature of a priori information is important:

• A first group of methods uses no knowledge about the nature of the degradation; only very general properties of the degradation are assumed.

• A second group assumes knowledge about the properties of the image acquisition device, and the conditions under which the image was obtained. The nature of noise (usually its spectral characteristics) is sometimes known.

• A third approach uses knowledge about objects that are searched for in the image, which may simplify the pre-processing very considerably. If knowledge about objects is not available in advance it can be estimated during the processing.

3.1.4 Image Segmentation

Image segmentation is the method for conversion of digital image into several segments and rendering of an image into something for easier analysis. Using image segmentation is used for locating the objects and bounding line of that image. In segmentation, we used Convolutional Neural Network (CNN) algorithm for partitioning of images into segments in which at least one part of segment contains image with major area of diseased part. The CNN algorithm is applied to classify the objects into 2 classes i.e. diseased or a healthy fruit.

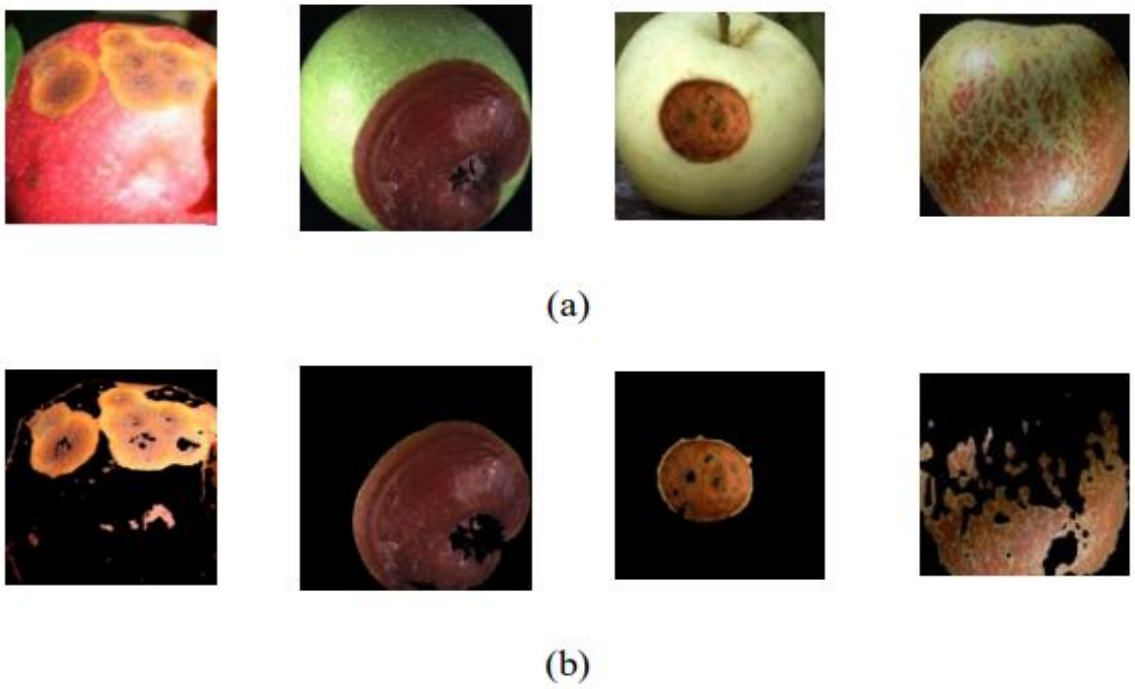

Fig 3.4 Image Segmentation Results (a)Images before segmentation (b)Images after segmentation

3.1.5 Feature Extraction

In feature extraction desired features such as color, texture, morphology, and structure are extracted. Feature extraction is method for involving number of resources required to describe a large set of data accurately. The technique of extracting the features is useful when you have a large data set and need to reduce the number of resources without losing any important or relevant information. Feature extraction helps to reduce the amount of redundant data from the data set.

3.1.6 Classification

Classifying images is a complex problem in the field of computer vision. The deep learning algorithm is a computerized model simulates the human brain functions and operations. Training the deep learning model is a costly process in machine resources and time.

Classification is the process of recognizing, understanding, and grouping ideas and objects into preset categories or "sub-populations." Using pre-categorized training datasets, machine learning programs use a variety of algorithms to classify future datasets into categories.

3.1.7 Verification

Verification is the process of checking that a software system meets specifications and that it fulfills its intended purpose. Using the statistical measures of the performance like sensitivity, specificity, accuracy, precision, and loss values finding out which classifier detects disease more accurately.

CHAPTER – 4

SOFTWARE DETAILS

CHAPTER-4

4.1 SOFTWARE DETAILS

4.1.1 Language: python

Python is an interpreted, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python becomes an apt choice for such Image processing techniques. This is due to its growing popularity as a scientific programming language and the availability of many State-of-The-Art Image Processing tools in its ecosystem.

Why is python used for this deep learning project?

AI projects differ from traditional software projects. The differences lie in the technology stack, the skills required for an AI-based project, and the necessity of deep research. To implement your AI aspirations, you should use a programming language that is stable, flexible, and has tools available. Python offers all of this, which is why we see lots of Python AI projects today.

From development to deployment and maintenance, Python helps developers be productive and confident about the software they're building. Benefits that make Python the best fit for machine learning and AI-based projects include simplicity and consistency, access to great libraries and frameworks for AI and machine learning (ML), flexibility, platform independence, and a wide community. These add to the overall popularity of the language.

Python offers concise and readable code. While complex algorithms and versatile workflows stand behind machine learning and AI, Python's simplicity allows developers to write reliable systems. Developers get to put all their effort into solving an ML problem instead of focusing on the technical nuances of the language. Additionally, Python is appealing to many developers as it's easy to learn. Python code is understandable by humans, which makes it easier to build models for machine learning.

Many programmers say that Python is more intuitive than other programming languages. Others point out the many frameworks, libraries, and extensions that simplify the implementation of different functionalities. It's generally accepted that Python is suitable for collaborative implementation when multiple developers are involved. Since Python is a general-purpose language, it can do a set of complex machine learning tasks and enable you to build prototypes quickly that allow you to test your product for machine learning purposes.

4.1.2 OS: windows 10 (64bits)

Windows 10 is a series of operating system developed by Microsoft and released as part of its windows NT family of operating systems. It is the successor to windows 8.1, released nearly two years earlier, and was released to manufacturing on July 15, 2015, and broadly released for the public on July 29, 2015.

4.1.3 IDE: Anaconda Navigator

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda distribution that allows you to launch applications and easily manage conda packages, environments, and channels without using command-line commands. Navigator can search for packages on Anaconda.org or in a local Anaconda Repository. It is available for Windows, MacOS, and Linux.

Anaconda distribution comes with over 250 packages automatically installed, and over 7,500 additional open-source packages can be installed from PiPy as well as the conda package and virtual environment manager. It also includes a GUI, Anaconda Navigator, as a graphical alternative to the command line interface (CLI).

The big difference between conda and the pip package manager is in how package dependencies are managed, which is a significant challenge for Python data science and the reason conda exists.

When pip installs a package, it automatically installs any dependent Python packages without checking if these conflict with previously installed packages. It will install a package and any of its dependencies regardless of the state of the existing installation.

Because of this, a user with a working installation of, for example, Google TensorFlow, can find that it stops working having used pip to install a different package that requires a different version of the dependent numpy library than the one used by TensorFlow. In some cases, the package may appear to work but produce different results in detail.

4.1.4 CNN: Convolutional Neural Network

A Convolutional Neural Network (CNN, or ConvNet) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. It is a multilayered neural network with a special architecture to detect complex features in data. CNNs have been used in image recognition, powering vision in robots, and for self-driving vehicles.

Once a CNN is built, it can be used to classify the contents of different images. All we have to do is feed those images into the model. Just like ANNs, CNNs are inspired by the workings of the human brain. CNNs are able to classify images by detecting features, similar to how the human brain detects features to identify objects.

Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. Although fully connected feedforward neural networks can be used to learn features and classify data, this architecture is impractical for images. It would require a very high number of neurons, even in a shallow architecture, due to the very large input sizes associated with images, where each pixel is a relevant variable. For instance, a fully connected layer for a (small) image of size 100 x 100 has 10,000 weights for each neuron in the second layer. Instead, convolution reduces the number of free parameters, allowing the network to be deeper. For example, regardless of image size, tiling 5 x 5 region, each with the same shared weights, requires only 25 learnable parameters. Using regularized weights over

fewer parameters avoids the vanishing gradient and exploding gradient problems seen during backpropagation in traditional neural networks.

4.1.5 Max Pooling Layer

Max pooling is a pooling operation that selects the maximum element from the region of the feature map covered by the filter. Thus, the output after max-pooling layer would be a feature map containing the most prominent features of the previous feature map. The pooling operation involves sliding a two-dimensional filter over each channel of feature map and summarizing the features lying within the region covered by the filter.

4.1.6 TensorFlow

TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. It is a symbolic math library based on dataflow and differentiable programming. It is used for both research and production at Google. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications. Tensor flow can help to build neural network models to classify images. Commonly, these will be convolutional neural networks (CNN). TensorFlow is a powerful framework that lets you define, customize, and tune many types of CNN architectures. Missing Link's deep learning platform provides an additional layer for tracking and managing TensorFlow projects.

4.2 Keras

Keras runs on top of open-source machine libraries like TensorFlow, Theano or Cognitive Toolkit (CNTK). Keras is based on minimal structure that provides a clean and easy way to create deep learning models based on TensorFlow or Theano. Keras is designed to quickly define deep learning models. Well, Keras is an optimal choice for deep learning applications.

4.2.1 Features

Keras leverages various optimization techniques to make high level neural network API easier and more performant. It supports the following features -

1. Consistent, simple and extensible API.

2. Minimal structure - easy to achieve the result without any frills.

3. It supports multiple platforms and backends.

4. It is user friendly framework which runs on both CPU and GPU.

5. Highly scalability of computation.

6. Numerous implemented layers and parameters

7. Multiple methods for Data Preprocessing

8. Model Evaluation

9. Modularity

10. Evaluation and Prediction

11. Pre-trained Models in Keras

12. Large Dataset

4.2.2 Benefits

Keras is highly powerful and dynamic framework and comes up with the following advantages -

- Larger community support.
- Easy to test.
- Keras neural networks are written in Python which makes things simpler.
- Keras supports both convolution and recurrent networks.
- Deep learning models are discrete components, so that, you can combine into many ways.

4.3 Softmax Function

The softmax function is a function that turns a vector of K real values into a vector of K real values that sum to 1. The input values can be positive, negative, zero, or greater than one, but the softmax transforms them into values between 0 and 1, so that they can be interpreted as probabilities. If one of the inputs is small or negative, the softmax turns it into a small probability, and if an input is large, then it turns it into a large probability, but it will always remain between 0 and 1.

The softmax function is sometimes called the softargmax function, or multi-class logistic regression. This is because the softmax is a generalization of logistic regression that can be used for multi-class classification, and its formula is very similar to the sigmoid function which is used for logistic regression. The softmax function can be used in a classifier only when the classes are mutually exclusive.

Many multi-layer neural networks end in a penultimate layer which outputs real-valued scores that are not conveniently scaled and which may be difficult to work with. Here the softmax is very useful because it converts the scores to a normalized probability distribution, which can be displayed to a user or used as input to other systems. For this reason, it is usual to append a softmax function as the final layer of the neural network.

4.3.1 Softmax Formula

The softmax formula is as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

$\vec{z}$ ⇨ The input vector to the softmax function, made up of (z0, ... zK)

z_i ⇨ All the zi values are the elements of the input vector to the softmax function, and they can take any real value, positive, zero or negative. For example a neural network could have output a vector such as (-0.62, 8.12, 2.53), which is not a valid probability distribution, hence why the softmax would be necessary.

e^{z_i} ⇨ The standard exponential function is applied to each element of the input vector. This gives a positive value above 0, which will be very small if the input was negative, and very large if the input was large. However, it is still not fixed in the range (0, 1) which is what is required of a probability.

$\sum_{j=1}^{K}$ ⇨ The term on the bottom of the formula is the normalization term. It ensures that all the output values of the function will sum to 1 and each be in the range (0, 1), thus constituting a valid probability distribution.

K ⇨ The number of classes in the multi-class classifier.

4.4 Flatten Layer

Flatten is used to flatten the input. For example, if flatten is applied to layer having input shape as (batch_size, 2,2), then the output shape of the layer will be (batch_size, 4)

Flatten has one argument as follows:

keras.layers.Flatten (data_format = None)

4.5 Dense Layer

Dense implements the operation: output=activation (dot (input, kernel) + bias) where activation is the element-wise activation function passed as the activation argument, kernel is a weight matrix created by the layer, and bias is a bias vector created by the layer (only applicable if use_bias is True). These are all attributes of Dense.

If the input to the layer has a rank greater than 2, then Dense computes the dot product between the inputs and the kernel along the last axis of the input and axis1 of the kernel(using tf.tensordot), if input has dimensions (batch_size, d0, d1), then we create a kernel with shape (d1, units), and the kernel operates along axis2 of the input, on every sub-tensor of shape(1, 1, d1) (there are batch_size * d0 such sub-tensors).The output in this case will have shape(batch_size, d0, units).

CHAPTER – 5

RESULTS

CHAPTER-5

5.1 RESULTS

This project is mainly concentrating on Disease detection in fruits using deep learning Concepts. By using this concept, the disease identification is done for all kinds of fruits the user can rectify the problem very easy and with less cost. Convolutional neural network and Max pooling layer is used to train the images.

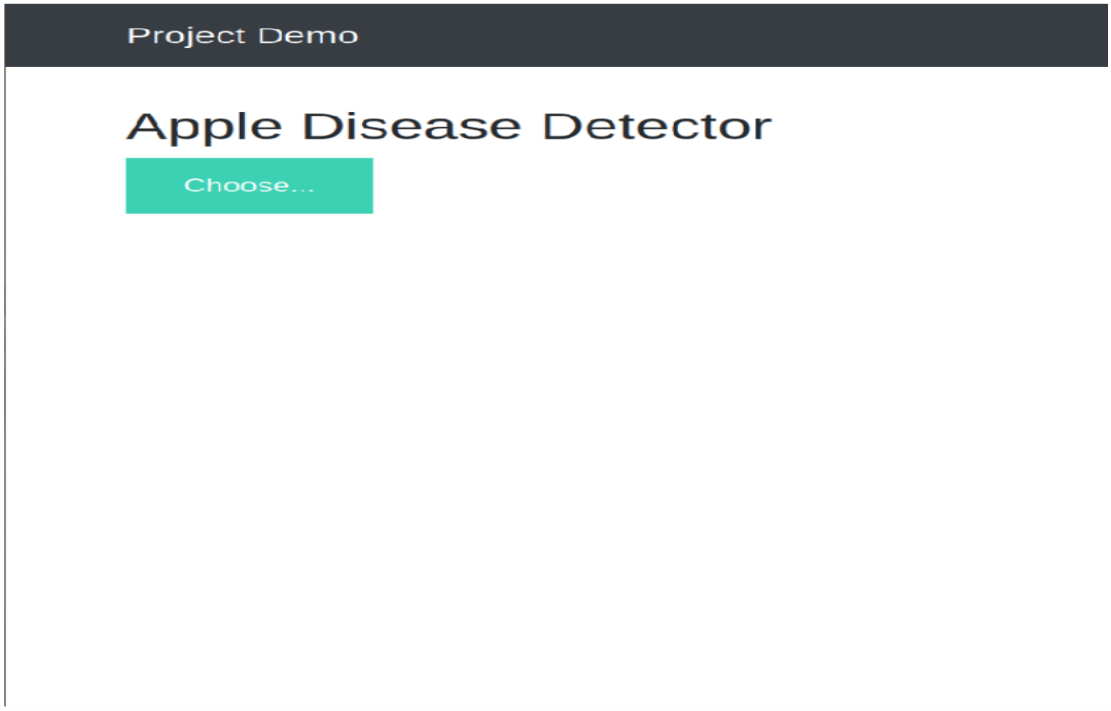

Fig 5.1 Demo Page

The above page will indicate that the Demo page. Here we just need to choose the Image and the application will starts working on that.

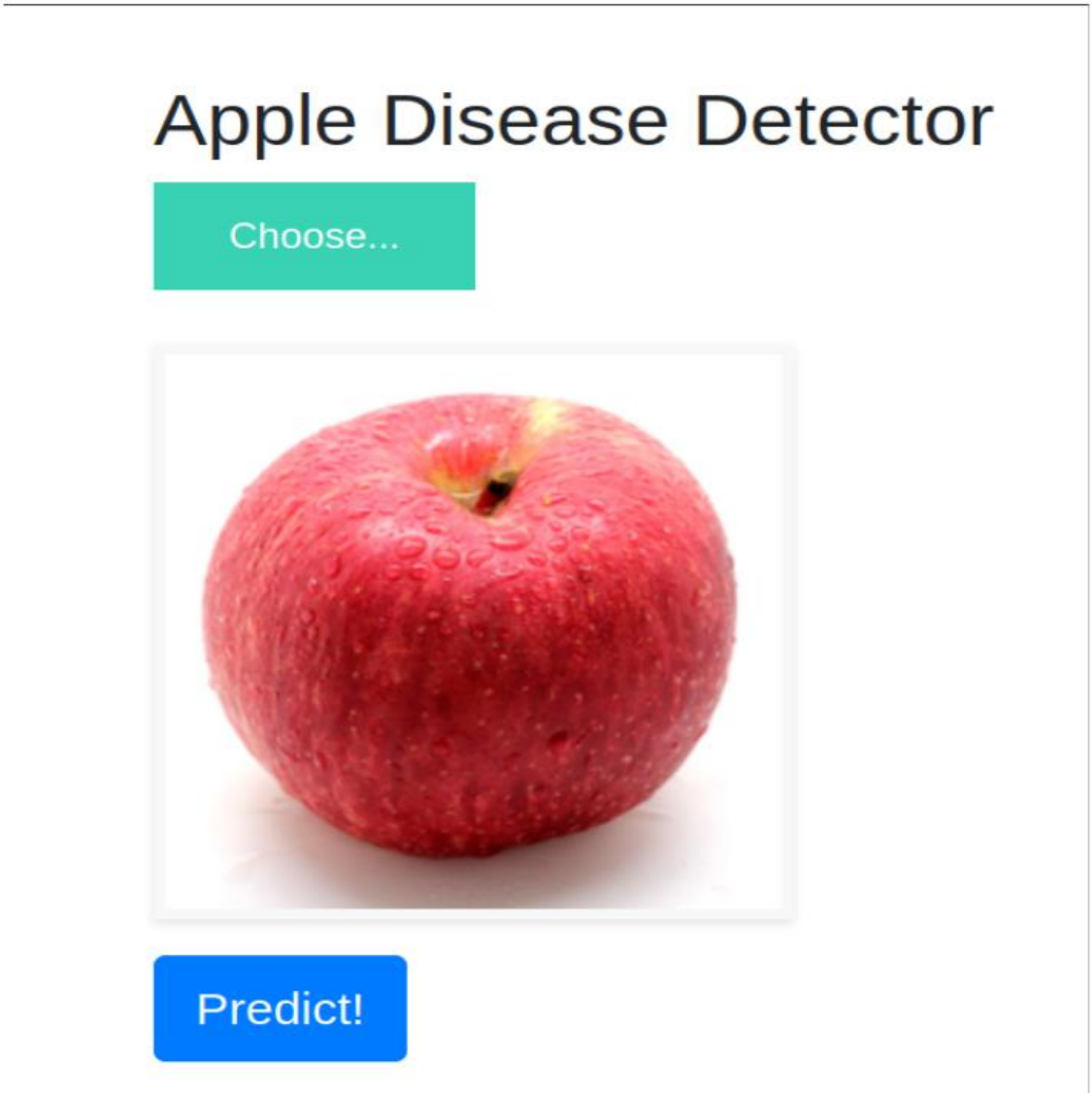

Fig 5.2 Image of a fresh apple fruit

.

The above figure is the sample image of a fresh fruit. Here we need to select the sample image of a fruit. The application will accept the image as shown above. Then finally we need to click on Predict! then the application starts working on it.

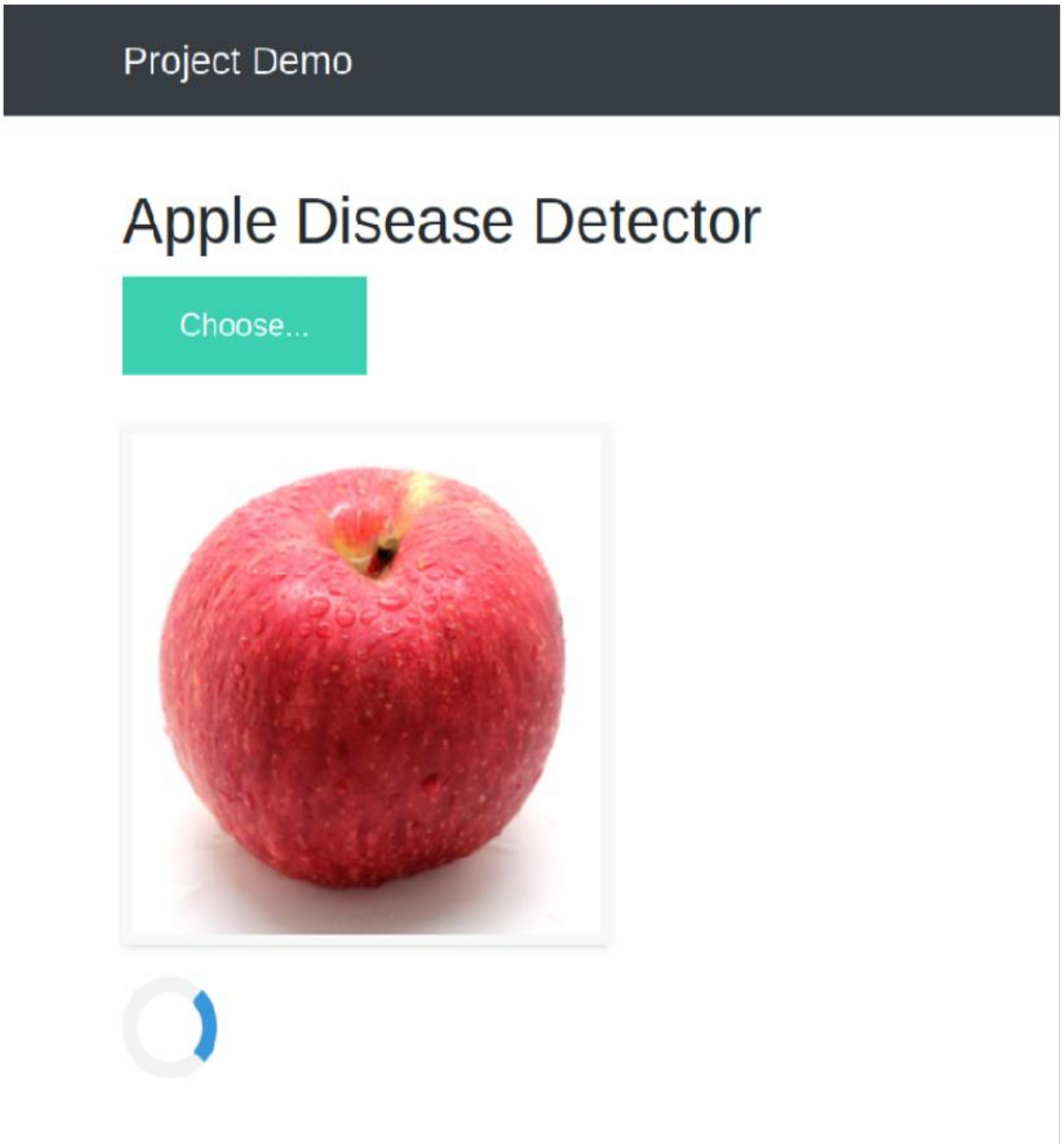

Fig 5.3 Processing of the image uploaded

The above image indicates the Processing of the image uploaded. Here the fresh image uploaded will perform the processing. Here the sample image is compared with the trained image and finally the result is displayed.

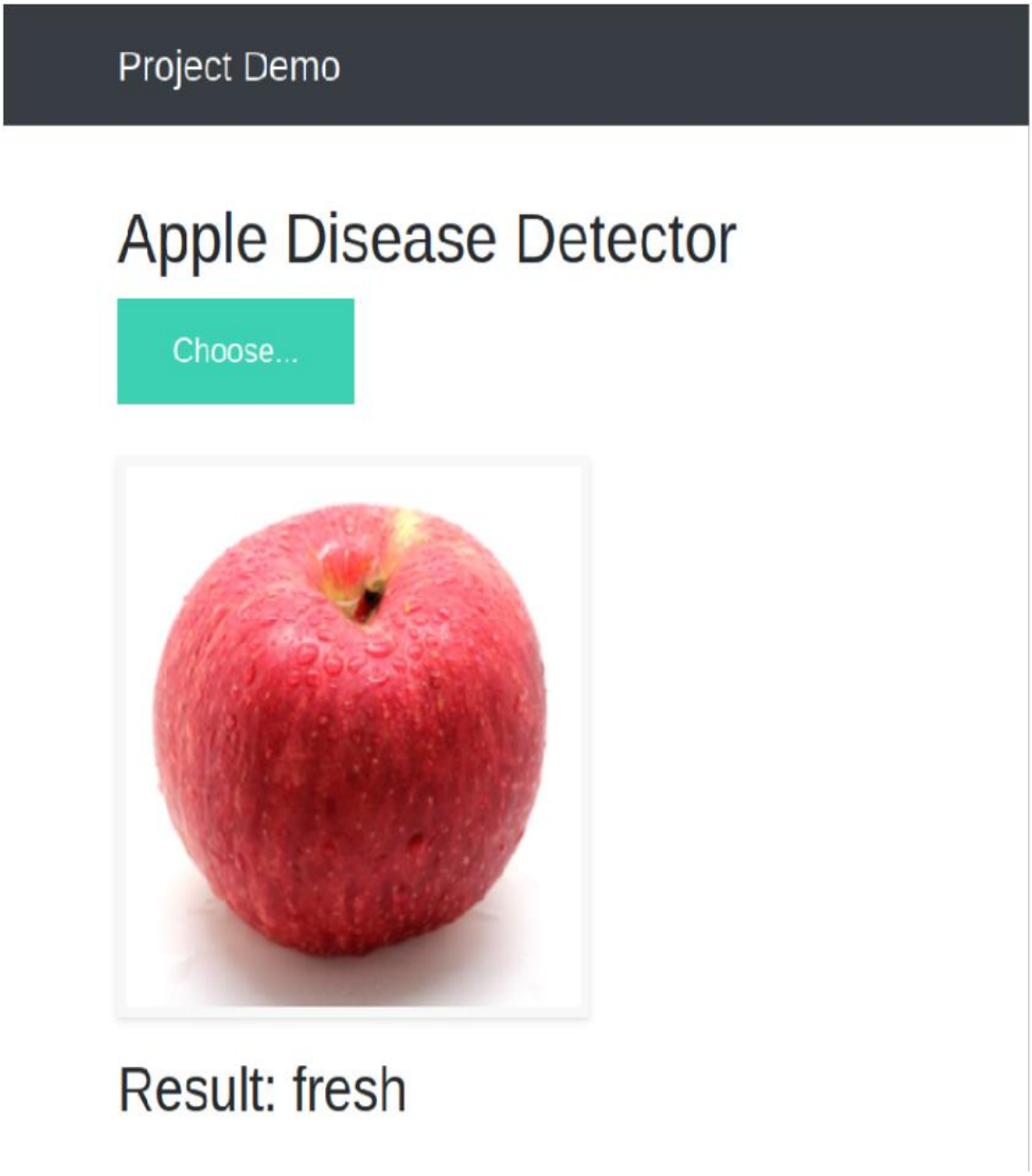

Fig 5.4 Result for a fresh apple fruit

The above image indicate the Result for a fresh apple fruit. Once the processing is done then the application will display the result as above. The sample image given is fresh hence it will display the result as fresh. And finally the application will also display the accuraccy graph.

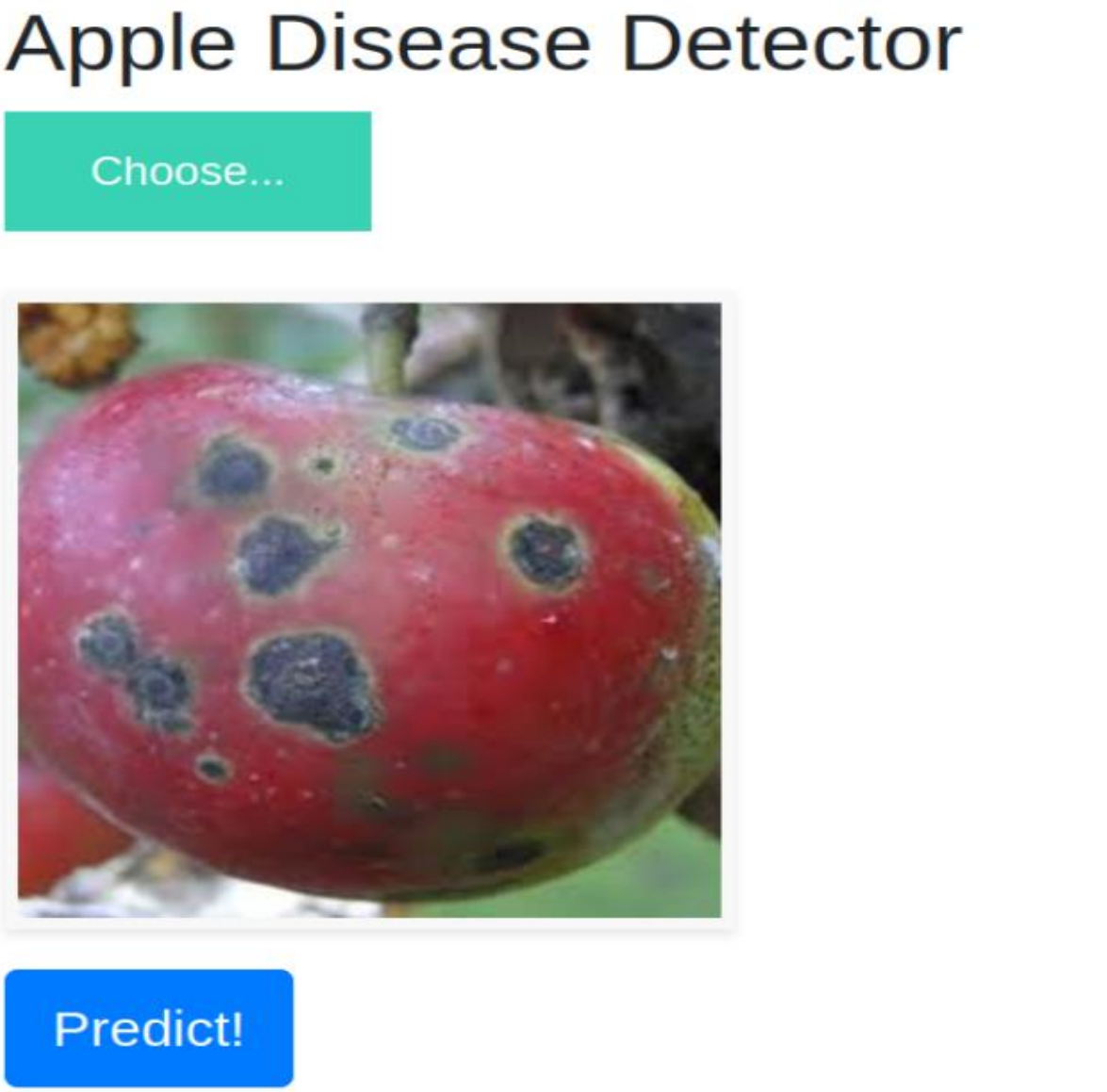

Fig 5.5 Image of a diseased apple

The above figure is the sample image of a diseased fruit. Here we need to select the sample image of a fruit. The application will accept the image as shown above. Then finally we need to click on Predict! then the application starts working on it.

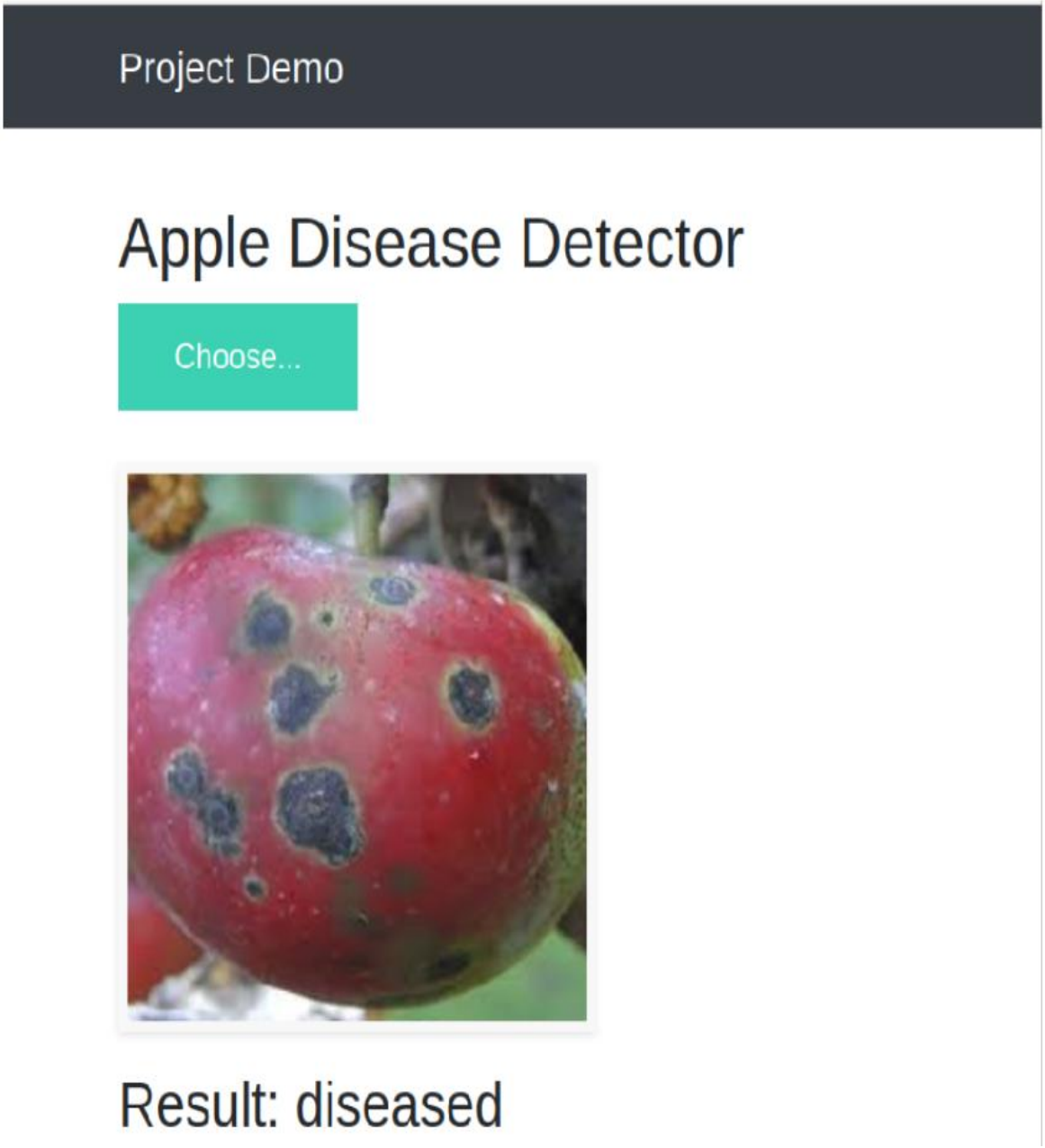

Fig 5.6 Result for a diseased apple

The above image indicate the Result for a diseased apple fruit. Once the processing is done then the application will display the result as above. The sample image given is diseased hence it will display the result as diseased. And finally the application will also display the accuraccy graph.

CHAPTER – 6

CONCLUSION

CHAPTER-6

6.1 CONCLUSION

This project uses innovative methods to detect diseases in fruits and is also a prototype in implementation of the usage of modern technology in agriculture, especially in cash crops. This project also aims at the thorough implementation of Digital Image Processing techniques. The disease is detected in a very cheap and fast way and measures to stop the spread of disease can be taken.

By using this concept, the disease identification is done for all kinds of Fruits the user can rectify the problem very easy and with less cost. Here we use the concept of Deep learning, this gives the results with more accuracy, hence that helps to take the early precautions.

REFERENCES

1] Hongjun wang, Qisong Mou, Youjun Yue, Hui Zhao: "Research on Detection Technology of various Fruit Disease Spots Based on Mask R-CNN", Tianjin University of Technology, Tianjin China, 2016

2] Shiv Ram Dubey, A.S Jalal: "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Pattern", December 2012.

3] Nivedita.R. Kakade, Dnyaneswar.D.Ahire. Ahire: "Real Time Grape Leaf Disease Detection", International Journal of Advance Research and Innovative Ideas in Education, Vol-1 Issue-4, 2015.

4] Kulkarni Anand H, Ashwin Patil RK: "Applying image processing technique to detect plant diseases". Int J Mod Eng Res, 2012

5] Sue Han Lee, Chee Seng Chan, Paul Wilkin, Paolo Remagnino: "Deep-plant: plant identification with Convolutional neural networks", rearXiv**,** 2015

6] Halil Durmu, Ece Olcay Güne, Murvet Kirci: "Disease Detection on the Leaves of Tomato Plants by Using Deep Learning", International Journal of Computer Applications, vol. 52, no. 2, 2012.

7] Sachin D. Khirade and A. B. Patil: "Plant disease detection using image processing," in Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on, IEEE, 2015

8] S. B. Ullagaddi and S.Viswanadha Raju: "Disease Recognition in Mango Crop Using Modified Rotational Kernel Transform Features". 2017 Internationals

9] Jamil Ahmad, Bilal Jan, Haleem Farman, Wakeel Ahmad, and Atta Ullah: "Disease Detection in Plum using Convolutional Neural Network under true field conditions".28 September 2020.