MID WEST UNIVERSITY

GRADUATE SCHOOL OF ENGINEERING

BIRENDRANAGAR, SURKHET

[Engineering project IV (CO521)]

A PROGRESS REPORT OF FINAL YEAR PROJECT

ON

“PLANT LEAF DISEASE DETECTION”

Submitted by:

Roshani Poudel [026]

Amrita Pant [002]

Paras Ghimire [018]

Subash Ghimire [029]

Submitted to:

Department of Computer Engineering

Graduate School of Engineering

Date: Ashar ,21

"PLANT LEAF DISEASE DETECTION "

Submitted by:

Roshani Poudel [026]

Amrita Pant [002]

Paras Ghimire [018]

Subash Ghimire [029]

Supervised by:

Er. Sonalal Yadav

(Assistant professor)

Submitted to:

Department of Computer Engineering

Graduate School of Engineering

Birendranagar, Surkhet

Ashar,21

2080

ACKNOWLEDGEMENT

It gives us immense pleasure to express our deepest sense of gratitude and sincere thanks to our highly respected and esteemed guide Er. Sonalal Yadav our project Supervisor, for his valuable guidance, encouragement and help for completing this work. His useful suggestions for this whole work and co-operative behavior are sincerely acknowledged.

We would like to express our sincere thanks to our supervisor Er. Sonalal Yadav with full designation and department for giving us this opportunity to undertake this project. We would also like to thank head of department of computer Engineering (HOD) Er. Prabin Bhusal for whole hearted support .We are also grateful to our faculty teachers for their constant support and guidance.

At the end we would like to express our sincere thanks to all our friends and others who helped us directly or indirectly during this project work.

Roshani Poudel [026]

Amrita Pant [002]

Paras Ghimire [018]

Subash Ghimire [029]

ABSTRACT

The detection of plant diseases is crucial to avert the losses in the productivity and in the yield. It is not obvious to monitor the plant diseases manually as the act of disease detection is very critical. It needs a huge effort, along with knowledge of plant diseases and extensive processing times. Therefore, image processing technology is used to detect the plant disease, this is done by capturing the input image that undergoes the process and is compared with the dataset. This dataset is composed of diverse plant leaves in the image format. This project aims at designing a standalone application that will provide the framer with the necessary information about the type of disease. The purpose of this project is to assist and provide efficient support to the monoculture farmers. In this, we propose a system that will use the different techniques of the image process to both analyze and detect the plant diseases. The results of the implementation show that the designed system could give a successful result by detecting and classifying the plant diseases., we are focusing on mango leaf to detect the disease.

The key to reducing losses in agricultural product output and quantity is early detection of plant diseases. The investigation of visually observable patterns on the plant is required for the research of plant diseases. Plant health monitoring and disease detection are crucial for long-term agriculture. Manually monitoring plant diseases is quite tough. It necessitates a great deal of effort, as well as knowledge of plant diseases and long processing times. As a result, image processing is utilized to detect plant illnesses. Image acquisition, image pre-processing, image segmentation, feature extraction, and disease identification are all processes in the disease detection process. Image acquisition, picture pre-processing, image segmentation, feature extraction, and classification are all processes in the disease detection process. The approaches for detecting plant illnesses using photographs of their leaves were addressed in this research. In addition, certain segmentation and feature extraction algorithms for plant disease diagnosis were discussed in this study. Image capture, segmentation, and feature extraction are some of the terms used in this paper.

Keywords: Image processing, disease detection, smart detection.

Table of Contents

TABLE OF FIGURE

1. INTRODUCTION

Nepal is a cultivated country, with agriculture employing over 66% of the people. Farmers have a wide range of options when it comes to choosing appropriate crops and insecticides for their plants. Plant disease causes a considerable decrease in the quality and quantity of agricultural goods. Plant disease research is concerned with the examination of visually discernible patterns on plants. The monitoring of plant health and disease is critical to the effective cultivation of crops on the farm. In the beginning, plant disease monitoring and analysis were done manually by an expert in the field. This necessitates a great deal of effort as well as a long processing time. Plant disease detection can be done using image processing techniques. Symptoms of disease can be noticed on the leaves, stems, and fruits in the majority of cases. The illness symptoms are shown on the plant leaf, which is used to detect the sickness. This study provides an overview of image processing techniques for detecting plant diseases.

Plants play a crucial role in our ecosystem, agriculture, and food production. However, they are vulnerable to various diseases that can significantly impact crop yield and quality. Early detection and accurate diagnosis of these diseases are essential for timely intervention and effective management strategies. In recent years, advancements in image processing and machine learning techniques have opened up new avenues for plant disease detection and monitoring.

The objective of this project is to develop a robust and automated system for plant disease detection using image processing and machine learning algorithms. By harnessing the power of computer vision and artificial intelligence, this project aims to provide farmers and agricultural experts with a reliable tool to identify and classify diseases affecting plants.

The proposed system will leverage image processing techniques to extract meaningful features from

"The plant disease detection" project follows the following process:

- Image acquisition
- Preprocessing
- Feature extraction
- Model evaluation
- Disease classification
- Integration
- Validation
- Iterative refinement

However, challenges exist in developing an effective plant disease detection system. These challenges include variations in image quality, environmental factors affecting disease manifestation, and the need for robust algorithms that can handle diverse plant species and diseases. Addressing these challenges will be a crucial aspect of this project, ensuring the reliability and accuracy of the developed system. In conclusion, this project aims to leverage the capabilities of image processing and machine learning to revolutionize plant disease detection. By providing an automated and accurate tool for farmers and agricultural experts, it has the potential to enhance crop management practices, optimize resource allocation.

Identification of the plant diseases is the key to preventing the losses in the yield and quantity of the agricultural product. The studies of the plant diseases mean the studies of visually observable patterns seen on the plant. Health monitoring and disease detection on plant is very critical for sustainable agriculture. 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 and Machine learning techniques are used for the detection of plant diseases. Disease detection involves the steps like image acquisition, image pre-processing, image segmentation, feature extraction and classification

MACHINE LEARNING:

Machine learning (ML) is the study of algorithms and mathematical models that computer systems use to progressively improve their performance on a specific task. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.

Plant disease identification by visual way is more laborious task and at the same time, less accurate and can be done only in limited areas. Whereas if automatic detection technique is used it will take less efforts, less time and become more accurate. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral and bacterial diseases. Image processing is used for measuring affected area of disease and to determine the difference in the color of the affected area. Image segmentation is the process of separating or grouping an image into different parts. There are currently many different ways of performing image segmentation, ranging from the simple thresholding method to advanced color image segmentation methods. These parts normally correspond to something that humans can easily separate and view as individual objects. Computers have no means of intelligently recognizing objects, and so many different methods have been developed in order to segment images. The segmentation process is based on various features found in the image. This might be color information, boundaries or segment of an image. Image is segmented using the K-means clustering technique. Then unnecessary part (green area) within leaf area is removed. After that we calculate the texture features for the segmented infected object. Finally, the extracted features are passed through a pretrained neural network.

IMAGE PROCESSING:

Image processing is any form of processing for which the input is an image or a series of images or videos, such as photographs or frames of video. The output of image processing can be either an image or a set of characteristics or parameters related to the image. It also means "Analyzing and manipulating images with a computer". Image processing is performed these three steps:

• First, import images with an optical device like a scanner or a camera or directly through digital processing.

• Second, manipulate or analyse the images in some way. This step can include image improvement and data summary, or the images are analyzed to find rules that aren't seen by the human eyes. For example, meteorologists use this processing to analyse satellite photographs

. • Last, output the result of image processing. The result might be the image changed by some way or it might be a report based on analysis or result of the images

CONVOLUTION NEURAL NETWORK (CNN):

When it comes to machine learning Artificial Neural Networks perform really well. Artificial Neural Networks are used in various classification tasks like image, audio, words. Different types of Neural Networks are used for different purposes, for example for predicting the sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolutional Neural Network. In this Blog, we are going to build basic building block for CNN. ConvNets derive their name from the “convolution” operator. The primary purpose of Convolution in case of a Convent is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small filters. There are many important parts of the Convolution Network. These includes the following properties

1. Depth: Depth corresponds to the number of filters we use for the convolution operation.

2. Stride: Stride is the number of pixels by which we slide our filter matrix over the input matrix.

3. Zero-padding: Sometimes, it is convenient to pad the input matrix with zeros around the border, so that we can apply the filter to bordering elements of our input image matrix.

4. Non-Linearity: ReLU is an element wise operation (applied per pixel) and replaces all negative pixel values in the feature map by zero. The purpose of ReLU is to introduce non-linearity in our ConvNet, since most of the real-world data we would want our ConvNet to learn would be non-linear (Convolution is a linear operation element wise matrix multiplication and addition, so we account for non-linearity by introducing a non-linear function like ReLU).

5. Spatial Pooling: Spatial Pooling (also called subsampling or down sampling) reduces the dimensionality of each feature map but retains the most important information. Spatial Pooling can be of different types: Max, Average, Sum etc. These networks have grown in the number of layers leading to architectures such as RestNet and Alex Net that have been trained on images such as Cifar-10 and then fined tune to other problems, such as plant classification.

1.1 Objective:

Here are the main objectives for plant leaf disease detection project:

- Develop an automated system that can accurately detect and classify plant leaf diseases using computer vision and machine learning techniques.
- Classify the healthy and affected leaf part by feature extraction and classification.
- Train the model by using testing data for accurate result.

1.2 Scopes:

The scope of plant leaf disease detection project can vary depending on the resources and time available. Here are some potential scopes to consider:

- Agriculture Sector related web-based application.
- Crop detection.
- It helps the farmer.

2. LITERATURE REVIEW

Image segmentation is dividing an image into more manageable, smaller pieces. This method is typically used to identify things in digital images. There are numerous methods for segmenting images, including thresholding, color-based, transform-based, and texture-based approaches. By removing only the most significant and eye-catching features from an image, a method known as "feature extraction" decreases the amount of pixels in the image. With this method, picture matching or retrieval can be sped up by using small representations and a high image size. "Image classification" refers to the labeling of photos into one of a variety of defined categories. The classification has two subcategories: supervised and unsupervised [4].

1.Detection of Plant Leaf Disease Using Image Processing Approach .[5]

Author: Sushil R. Kamlapurkar

The objective of the project is to Develop a system to accurately identify and classify Grape fruit diseases by automating the process using pre-processing, feature extraction, and ANN classification.

Proposed system improves disease identification accuracy through image resizing, color space conversion, segmentation, feature extraction (major axis, minor axis, eccentricity), and BPNN classification.

This project is limited to specific Grape fruit diseases, lacks evaluation with human experts, dataset details undisclosed, and limited assessment of robustness to environmental variations.

Overall: Further research needed to address limitations and validate system's applicability in real-world agricultural settings.

2. Plant Disease Detection Using Machine Learning [6]

Author: Shima Ramesh Maniyath,Vinod PV, Hebbar Ram

The objective of the project is to Develop a system to recognize abnormalities in plants, whether in greenhouses or natural environments, using preprocessing, HOG feature extraction, classifier training, and classification.

The proposed methodology utilizes HOG feature extraction for effective plant abnormality detection. Preprocessing ensures uniform image sizes. The system differentiates between diseased and healthy leaves through classifier training.

Limitation: Lack of information on the dataset used and classifier employed, requiring further research for comprehensive evaluation and performance assessment.

3.Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN.[7]

Author: Usha Kumari

The objective of the project is to develop a system to identify leaf spot diseases using image processing techniques, including image acquisition, segmentation, feature extraction, and classification with a neural network (NN) classifier.

The proposed system accurately segments disease-affected regions using K-means clustering and extracts relevant features for disease classification. The NN classifier is utilized for accurate disease identification.

Limitation: Lack of specific dataset details and performance metrics for the NN classifier, requiring further research for comprehensive evaluation and comparison with other approaches.

4.Multi-resolution mobile vision system for plant leaf disease diagnosis.[8]

Author: Shitala Prasad

The objective of Shitala Prasad's research is to propose a mobile-based client-server design for leaf disease detection using Gabor wavelet transform (GWT). The system involves color conversion, mobile pre-processing, human perception enhancement, K-means unsupervised algorithm for leaf image analysis, and Gabor wavelet conversion for feature extraction.

The proposed system shows potential for leaf disease detection. Color conversion, pre-processing, and human perception enhancement improve accuracy. K-means and Gabor wavelet enable effective analysis and feature extraction. However, limited details provided on dataset, performance metrics, and system effectiveness.

The review lacks specific details about the experimental setup, dataset characteristics, and performance metrics used. It does not provide information on the accuracy or effectiveness of the proposed system. Additionally, the focus on a homemade dataset raises concerns about the generalizability of the results to real-world scenarios and different plant species.

Further research is needed to address these limitations by conducting experiments with diverse datasets, evaluating system performance, and testing the proposed approach in complex backgrounds and varying lighting conditions encountered in practical settings.

5. Plant disease leaf image segmentation based on super pixel clustering and EM algorithm. [9]

Author:Shanwen Zhang, Zhuhong You, Xiaowei Wu

The objective of the project is to discuss a hybrid clustering method for leaf segmentation in plant disease detection, utilizing superpixel clustering based on common feature characteristics and suggesting the use of Expectation Maximization (EM) algorithm for color image segmentation.

The proposed hybrid clustering method shows promise for leaf segmentation in plant disease detection. Superpixel clustering reduces image complexity while preserving relevant information. The suggestion of EM algorithm for color image segmentation is valuable. However, the review lacks specific implementation details, performance evaluation, and consideration of limitations.

Limitation: Lack of specific implementation and evaluation details, and limited discussion on potential limitations and real-world applicability, requiring further research for validation and optimization of the proposed method.

6.Feasibility study on plant chili disease detectiion using image processing techniques[10]

Author Name: Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook

The objective of the project is to determine the health status of chilli plants by capturing leaf images and applying chemical treatments to diseased plants. Use MATLAB for feature extraction and image recognition. Pre-process images using Fourier filtering, edge detection, and morphological operations. Implement a GUI using LABVIEW software.

The proposed methodology aims to determine chilli plant health using image processing and selective chemical treatment. MATLAB is used for feature extraction, while pre-processing involves Fourier filtering, edge detection, and morphological operations. The integration of computer vision and LABVIEW enables user-friendly image analysis. However, limited details are provided on performance evaluation and scalability.

Limitation: The review does not provide specific details regarding the performance, accuracy, or validation of the proposed methodology. It lacks information on the dataset used and the effectiveness of the chemical treatment on diseased plants. Additionally, potential limitations related to varying lighting conditions, image variations, and scalability are not discussed.

7. An artificial neural network approach to identify fungal diseases of cucumber Plants using digital image processing[12]

Author: Keyvan Asefpour Vakilian and Jafar Massah.

The objective of the project is to develop an ANN approach for identifying P.cubensis and S.fuliginea infections in cucumber plant leaves using digital image processing.

The study successfully demonstrates the effectiveness of the ANN model in accurately identifying fungal diseases in cucumber plants. Digital image processing enables the detection and classification of P.cubensis and S.fuliginea infections.

This project is limited to two specific diseases of cucumber plant.

8. Detection of plant leaf diseases using image segmentation and soft computing Techniques[13]

Author:Vijai Singh ,A.K. Misra

The objective of the project is to detect plant leaf diseases using image segmentation and soft computing techniques. The authors aim to develop an algorithm for segmenting plant leaf images and propose an image recognition and segmentation process.

This project present an algorithm for segmenting plant leaf images, utilizing different segmentation methods. They capture images of various types and apply the segmentation process. Feature extraction is performed using the color co-occurrence method, and the experiments are conducted in MATLAB. The results are demonstrated for beans, leaf, lemon, and banana leaf.

The limited focus on a few types of leaves and the absence of comprehensive results for all leaf types restrict the generalizability of the proposed algorithm. Further research should involve a wider range of leaves and comprehensive evaluation to enhance the algorithm's applicability in practical scenarios.

9. Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification[14]

Author: Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic,"

The objective of the project is to utilize deep convolutional neural networks (CNN) for the recognition of plant diseases through leaf image classification.

This review demonstrates the effectiveness of deep CNN for leaf disease recognition. They highlight the impact of climate change on disease development and train a deep neural network to differentiate the surroundings of leaves. The images are manually cropped to focus on the region of interest, and an augmented process is applied to increase the dataset size. The use of the Caffe framework for deep CNN is presented.

The focus on a limited number of plant diseases and leaf types restricts the generalizability of the results. Additionally, the project does not address the challenges of real-time disease detection or the robustness of the proposed approach under different environmental conditions. Further research is needed to address these limitations and enhance the applicability of the system in practical agricultural settings.

10.Tomato leaves diseases detection approach based on support vector machines[15]

Author: Usama Mokhtar, Mona A. S. Ali, Aboul Ella Hassenian, Hesham Hefny

The objective of the project is to detect tomato leaf diseases using a support vector machine (SVM) approach. They aim to extract tomato leaf features using Gabor wavelet transform techniques and train an SVM model for disease detection.

The project successfully applies Gabor wavelet transform for feature extraction and SVM for disease detection in tomato leaves. Early blight and powdery mildew are observed and classified. Preprocessing involves background subtraction and image resizing. The SVM model is trained and tested using a kernel function.

The project is limited by the specific focus on tomato leaf diseases and may not cover all possible diseases. The fixed image resolution during preprocessing and lack of specific SVM details are limitations. Further research is needed with a larger and diverse dataset to improve accuracy and generalize the approach.

11. Image Processing Based Leaf Rot Disease, Detection of Betel Vine[16]

Author:Amar Kumar Dey, Manisha Sharma, M.R.Meshram

The objective of the project by Amar Kumar Dey, Manisha Sharma, and M.R. Meshram is to detect leaf rot disease in betel vine using image processing algorithms. They propose a vision-based method to observe and identify peripheral disease features based on color features of the rotted leaf area.

The project successfully applies image processing algorithms to detect leaf rot disease in betel vine. Color features and Otsu thresholding method are utilized for disease segmentation. The study focuses on Bangla desi varieties of betel vine and utilizes a Canon scanner for detection.

The project is limited to specific varieties of betel vine and may not be applicable to other varieties. The use of a specific scanner and resolution restricts generalizability. Further research is needed to validate the approach on a wider range of betel vine varieties and real-world scenarios.

12. Detection of Plant (soybeans)Diseases[18]

Author: S. Arivazhagan , A. R. Bhagat Patil1, Lokesh Sharma

The objective of the project is to detect plant diseases, specifically soybean rust, using image processing techniques and artificial neural networks (ANN).

The project propose a method for detecting plant diseases, particularly soybean rust, using image processing and ANN. The method employs a combination of ANN classification and Gabor filter-based feature extraction. The proposed approach shows promising results, achieving an identification rate of up to 91% after feature extraction.

The scope of the study focuses primarily on soybean rust and may not cover a wide range of plant diseases. Further research is needed to evaluate the method's performance on a broader range of plant diseases and optimize the detection accuracy.

13.Mango leaf disease recognition using neural network and support vector machine.[19]

Author: Md. Rasel Mia, Sujit Roy, Subrata Kumar Das & Md. Atikur Rahman

The objective of the project is to develop a system for mango leaf disease recognition using neural network and support vector machine (SVM). The aim is to provide an automated and accurate method for identifying mango leaf diseases, which can be challenging to detect with the naked eye.

This project propose a neural network ensemble (NNE) approach for mango leaf disease recognition. The system utilizes machine learning techniques to monitor and identify different symptoms of mango leaf diseases. Trained data are collected through classification techniques using images of diseased leaves. The proposed system achieves an average accuracy of 80% in detecting and classifying the examined diseases, which can significantly benefit mango plantations by saving time and facilitating proper disease treatment.

Limitation: The system's performance may vary depending on the diversity and complexity of mango leaf diseases. Further research is needed to enhance the accuracy and robustness of disease detection and classification, considering a broader range of mango leaf diseases.

14.Crop disease detection using image processing technique[20]

Author:Yashpal Sen, Chandra Shekar Mithlesh, Dr. Vivek Baghel

The objective of the project is to develop an automated system for crop disease detection using image processing techniques. The aim is to identify and classify different diseases of contaminated plants to prevent heavy losses in agriculture.

The paper presents an effective approach for detecting and classifying crop diseases, focusing on the region of interest, which is the leaf where most diseases occur. The system uses image processing techniques such as histogram equalization and K-means clustering for accurate disease detection, providing valuable information to farmers for disease analysis and prevention.

The system's performance may vary depending on the diversity and complexity of leaf diseases.

15. Detection and classification of leaf diseases using K- means-based segmentation and neural-networks-based classification[21]

Author: K. Elangoran, S. Nalini

The objective of the project is to detect and classify leaf diseases using image segmentation and neural networks-based classification.

The paper presents a concept for plant disease classification using image segmentation and support vector machine (SVM) techniques. The proposed approach involves analyzing colored images of plant leaves to identify visual symptoms of diseases. The software program developed in LABVIEW captures the RGB color model of the leaf images, and MATLAB is used for the recognition process to determine the plant disease. The color model is utilized to reduce the impact of illumination and efficiently distinguish between leaf colors, and clustering is performed to group color pixels in the images.

The system's performance may vary depending on the diversity and complexity of leaf diseases.

3. METHODLOGY

To develop plant leaf disease detection system among various machine learning algorithm we will be using CNN with Convolution and Max Pooling Layers. Through image processing and ML algorithms we aim to classify plant diseases and generate a model that would provide easy and accurate way of determining plant diseases though image of affected plant leaf.

To achieve our objectives, we are planning on using following workflow:

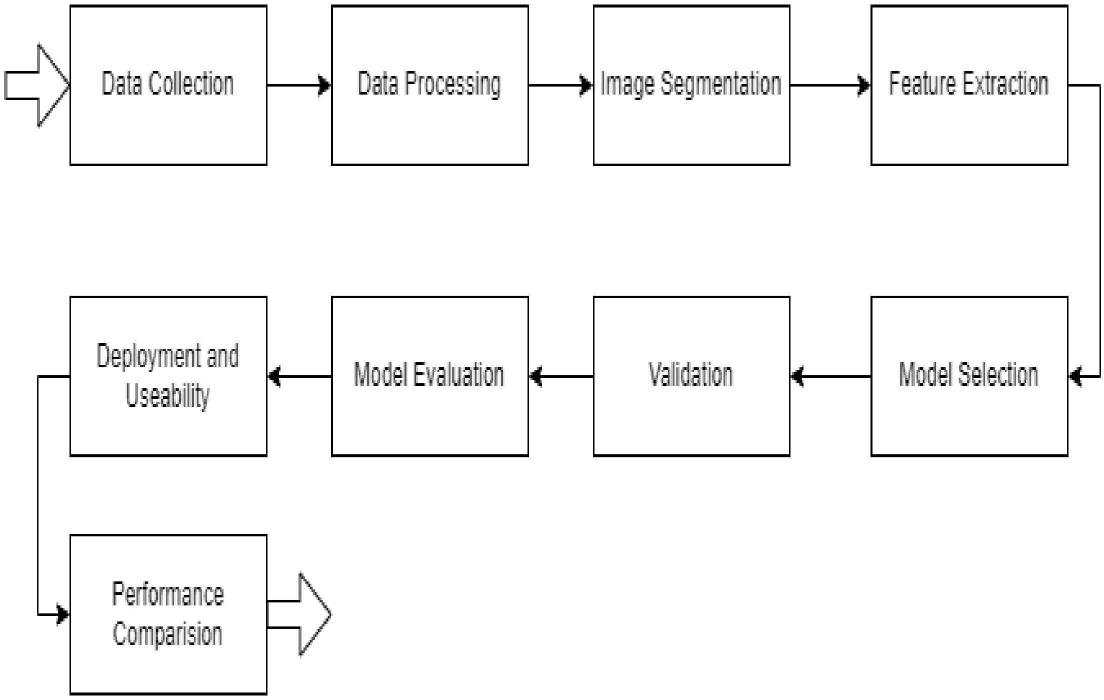

Figure 1. Block diagram of workflow of project

3.1 SYSTEM ARCHITECTURE:

The proposed System architecture comprises of data acquisition from a huge dataset, processing at different convolutional layers and then the classification of plant diseases which declares if the plant image is of a healthy class or diseased class.

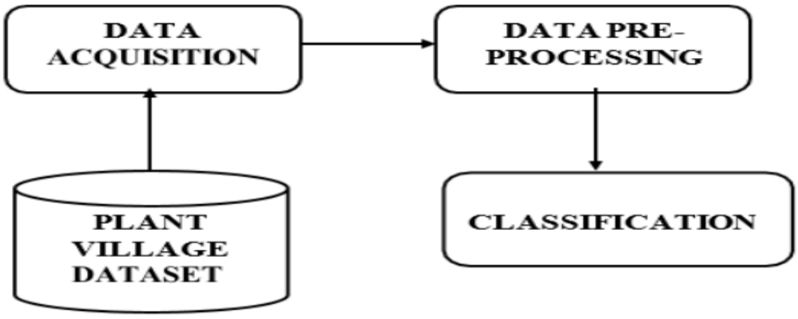

Figure 2: System Architecture

3.2 SOFTWARE REUQIREMENTS SPECIFIACTION

The software requirement specification of our project will have the entire necessary requirement which will be a baseline of our project. The software requirement specification will be software development life cycle, system architecture, data flow diagrams, UML diagrams, experimental setup requirements and performance metrics.

3.2.1 Software development Life cycle:

We use waterfall model software development life cycle in our project.

Waterfall Model: The waterfall model is a sequential software development process that follows a linear and rigid approach. It is called the "waterfall" model because the development flows downward through several phases, much like a waterfall flowing from one step to the next. Here is a short description of the waterfall model

The waterfall model consists of distinct and sequential phases, which typically include requirements gathering, system design, implementation, testing, deployment, and maintenance. Each phase has specific deliverables and objectives, and they are completed one after the other without overlapping or iteration.

General overview of “waterfall model”

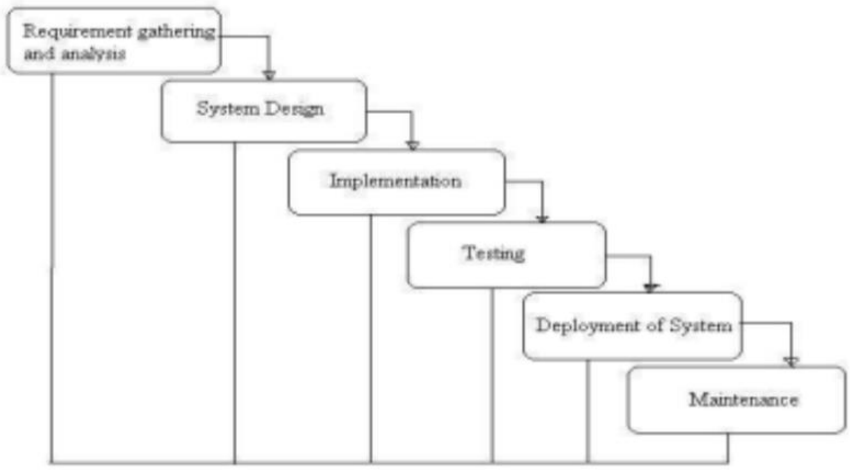

Figure 3: Analysis models: SDLC model to be applied

System implementation plan:

1. **Requirement gathering and analysis**:

In this step of waterfall we identify what are various requirements are need for our project such are software and hardware required, database, and interfaces.

2. System Design:

In this system design phase we design the system which is easily understood for end user i.e. user friendly. Web design some UML diagrams and data flow diagram to understand the system flow and system module and sequence of execution.

3. Implementation:

In implementation phase of our project we have implemented various module required of successfully getting expected outcome at the different module levels. With inputs from system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality which is referred to as Unit Testing.

4. Testing:

The different test cases are performed to test whether the project module are giving expected outcome in assumed time. All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.

5. Deployment of System:

Once the functional and nonfunctional testing is done, the product is deployed in the customer environment or released into the market.

6. Maintenance:

There are some issues which come up in the client environment. To fix those issues patches are released. Also to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment. All these phases are cascaded to each other in which progress is seen as flowing steadily downwards like a waterfall through the phases. The next phase is started only after the defined set of goals are achieved for previous phase and it is signed off, so the name "Waterfall Model". In this model phases do not overlap.

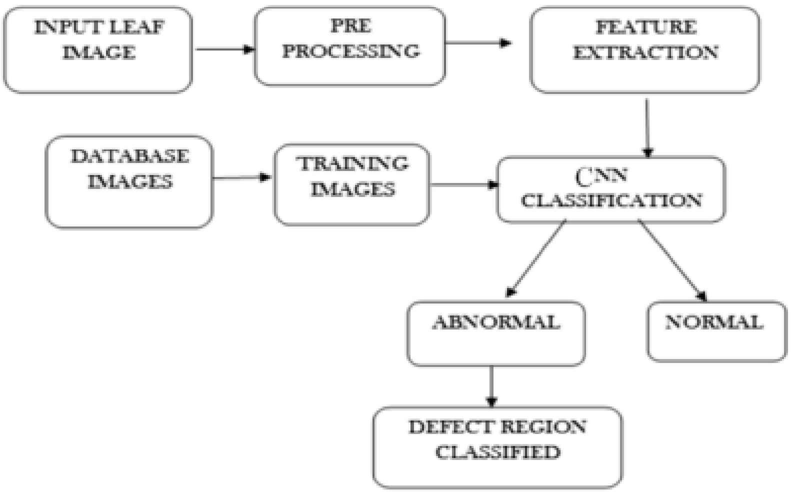

Figure 4: System flow

3.2.2 Data flow diagram (DFD):

Data Flow Diagrams(DFDs) describe the processes of how the transfer of data takes place from the input till prediction of the corresponding output.

1. **Data Flow Diagram – Level 0**

The DFD Level 0 depicts the users to input the image of the plant leaves. The system in turn detects and recognizes the plant leaf disease.

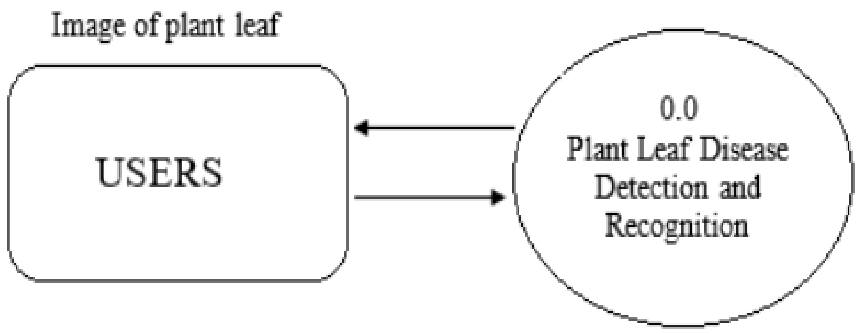

Figure 5: DFD Level 0

2. **Data Flow Diagram – Level 1:**

The Figure Displays the DFD Level 1, where the CNN model takes the image from the training dataset and then CNN model predicts the type of disease of the leaf.

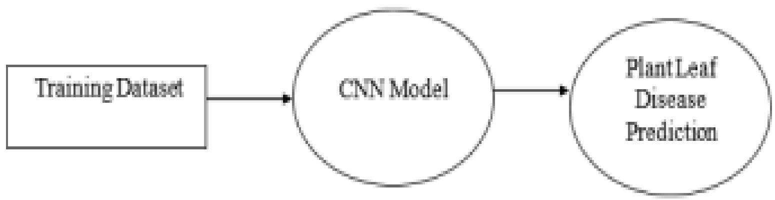

Figure 6: DFD Level 1

3.Data Flow Diagram – Level 2:

DFD Level 2 goes one step deeper into parts of 1-level DFD. It can be used to plan or record the specific/necessary detail about the system's functioning.

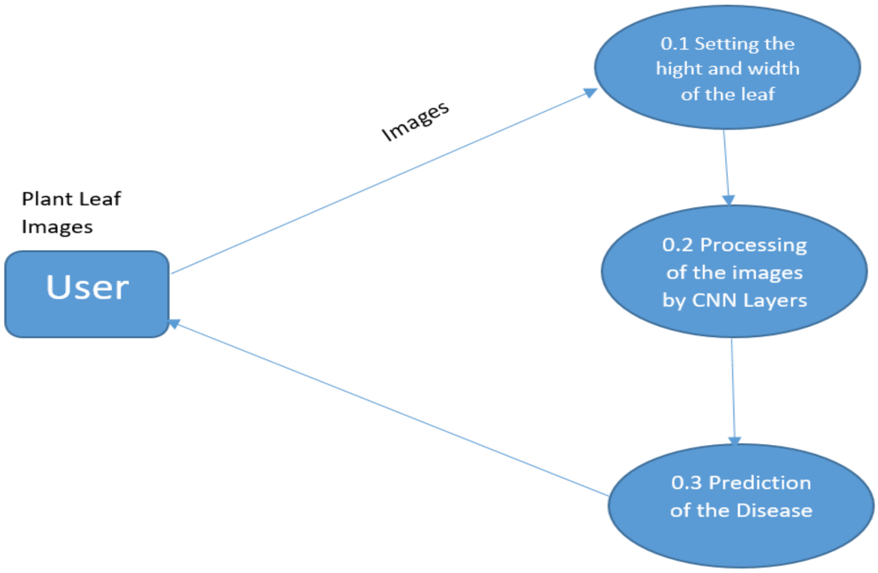

Figure 7: DFD Level 2

3.2.3 Use -case diagram:

Use-case diagrams describe the high-level functions and scope of a system. These diagrams also identify the interactions between the system and its actors. The use cases and actors in use-case diagrams describe what the system does and how the actors use it, but not how the system operates internally

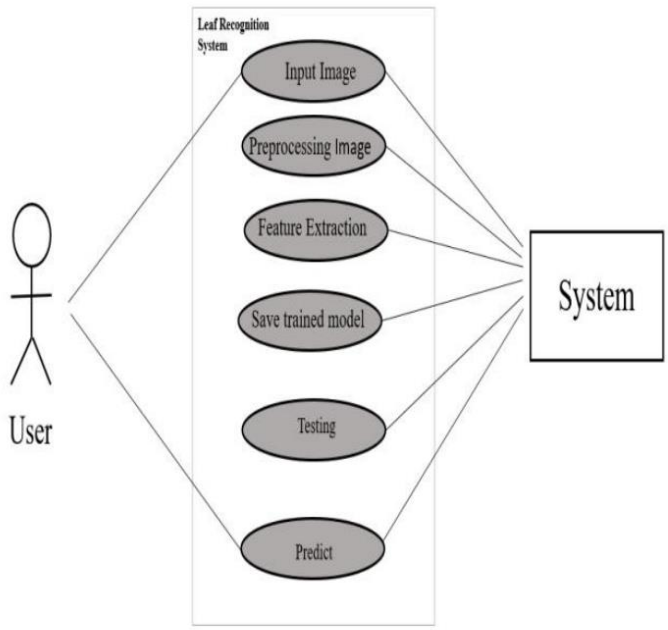

Figure 8: Use case Diagram

3.2.4 Sequence Diagram:

Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows as parallel vertical lines (lifelines), different processes or objects that live simultaneously and as horizontal arrows, the messages exchanged between them in the order in which they occur.

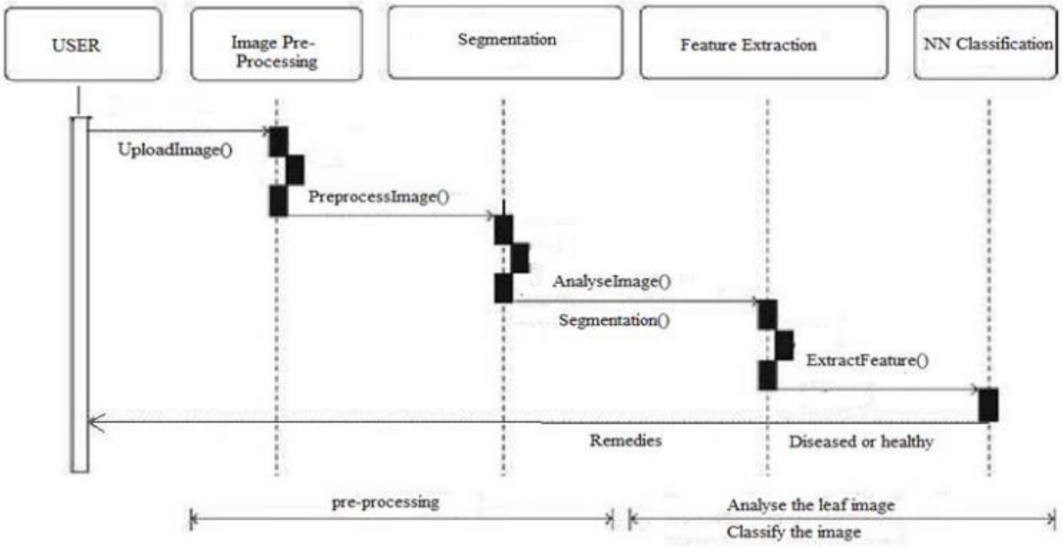

Figure 9: Sequence Diagram

4.DATA ANALYSIS

The dataset that is used in this proposed system project is the Plant leaf dataset and it was downloaded from different website like: Kaggle the dataset consisted of images of diseased and healthy plant leaf images. Upon exploration, we found that the dataset did not have any missing values. The dataset was further explored to understand the various spices and diseases of plant leaf. In this process we approach employed to classify the leaves into diseased or healthy and if the leaf is diseased, name of the disease is mentioned along with the remedies. Our methodology primarily revolves around the following steps.

4.1 Data collection:

Gather a diverse dataset of plant images containing both healthy plants and plants affected by various diseases. You can collect images from public databases, collaborate with local farmers, or capture your own images. The dataset used in our system is collection of images of mango plant leaf that either are healthy or fall.

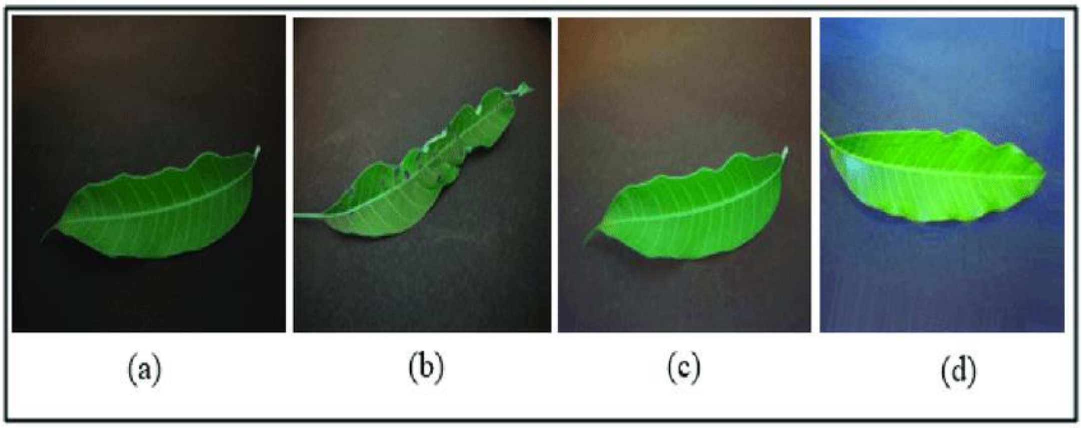

Figure 10: Mango leaf dataset

4.2 Image Processing:

Clean and preprocess the collected images to enhance their quality. This may involve resizing, cropping, color normalization, and noise reduction techniques. It is a whole big mechanism comprising of various stages and algorithms. If we look from the entire process as a whole, the function of this stage is to [22]:

1.Perform training of the images that are pre-composed and collected by us and stored in the form of our dataset

2. Perform testing on the image of the plant leaf captured by us to know if the leaf is diseased or not. The dataset collected forms the training data which is to

test the images taken in by the camera. The Web interface comes into requirement when we need a user interface wherein a user needs to upload the captured picture into the front-end and the model is pre-trained by the dataset of images in the back-end. The result is of the system. This interface renders easy and smooth flow of control and the user does not need to know the entire mechanism behind the Plant Disease Identification System

4.3 Procedure:

The Digital Image Processing is a combination of a number of steps and algorithms that go together in a controlled flow manner. The flowchart as shown below depicts the various stages that the images goes through before the final result is produced.

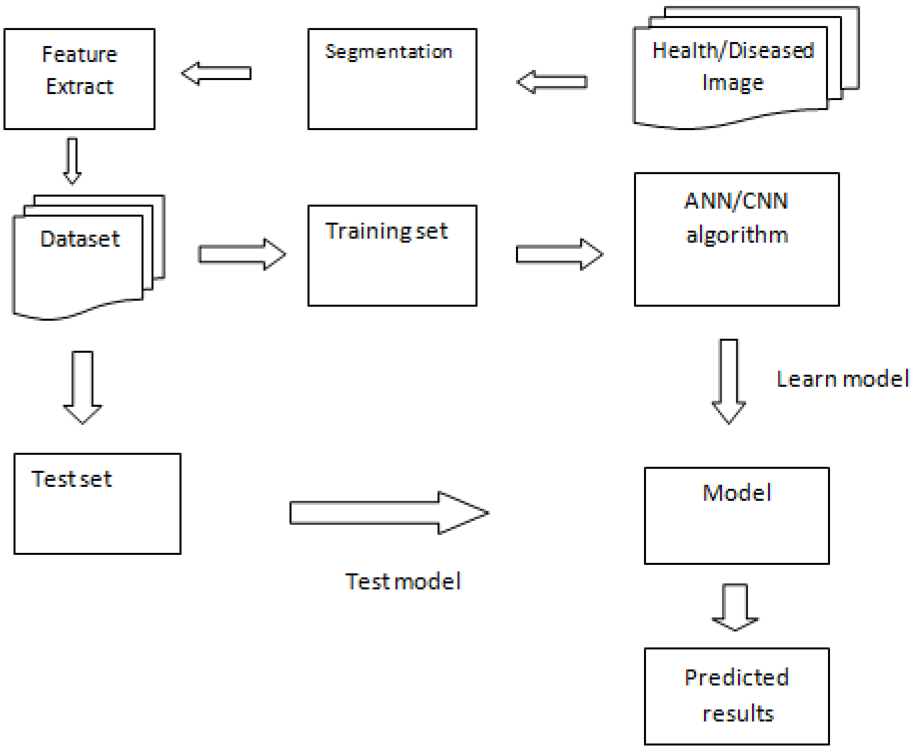

Figure 11: Execution Flow

4.4 Image segmentation:

Apply image segmentation techniques to isolate the plant regions from the background. This step helps in focusing the analysis on the relevant areas.

4.5 Feature extraction:

Extract meaningful features from the segmented plant images. Features can include color histograms, texture descriptors, or shape characteristics. These features should capture the distinctive patterns associated with healthy and diseased plants Image segment. The segmented images are clustered into different sectors using Otsu classifier and k-mean clustering algorithm. Before clustering the images, the RGB color model is transformed into Lab color model. The advent of Lab color model is to easily cluster the segmented images.

4.6 Model selection:

Choose a suitable machine learning or deep learning model for disease classification. Convolutional Neural Networks (CNNs) are commonly used for image-based classification tasks. You can explore pre-trained models such as VGG, Res Net, or Inception, or design your own architecture.

4.7 Training and validation:

Split the dataset into training and validation sets. Train the selected model on the training set and fine-tune its parameters to optimize performance. Monitor the model's performance using validation metrics such as accuracy, precision, recall, and F1-score.

4.8 Model evaluation:

Evaluate the trained model on a separate test set that was not used during training. Calculate the model's performance metrics to assess its accuracy and generalization capabilities. Consider using confusion matrices to analyze the model's classification performance for each disease class.

4.8 Deployment and usability:

Develop a user-friendly interface or application to demonstrate the practical use of your model. This could be a mobile app or a web-based tool where users can upload plant images for disease detection.

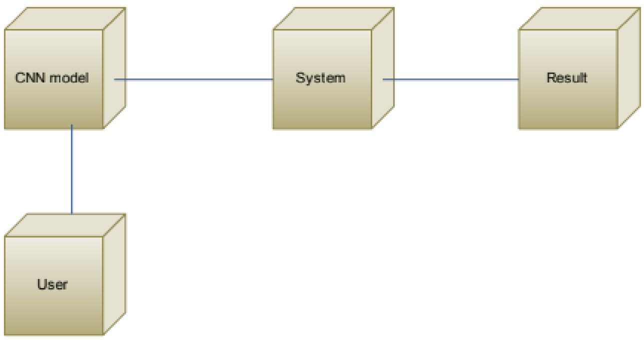

Figure 12: Deployment Diagram

4.9 Performance comparison:

Optionally, compare the performance of your model with other existing methods or models in terms of accuracy, speed, or resource requirements. This step provides a basis for evaluating the effectiveness of your approach.

5.TECHINICAL PROGRESS

We have developed a prototype model for identifying diseases in leaves, which currently distinguishes between healthy and unhealthy leaves. However, the categorization of specific diseases has not been incorporated into the prototype yet. The prototype undergoes several processes and steps as outlined below:

a. Dataset Collection:

For the collection of the mango leaf image dataset, we obtained a significant portion of the data from Kaggle, a popular online platform for sharing datasets and data-related projects. Kaggle provides a wide range of datasets contributed by individuals and organizations worldwide. In our case, we specifically searched for mango leaf image datasets that were publicly available and aligned with our project objectives.

After thorough evaluation and verification of the datasets, we selected a suitable mango leaf dataset from Kaggle. The dataset included a substantial number of labeled images of mango leaves, covering various health conditions, diseases, and disorders. These images were contributed by the Kaggle community, including researchers, enthusiasts, and experts in the field.

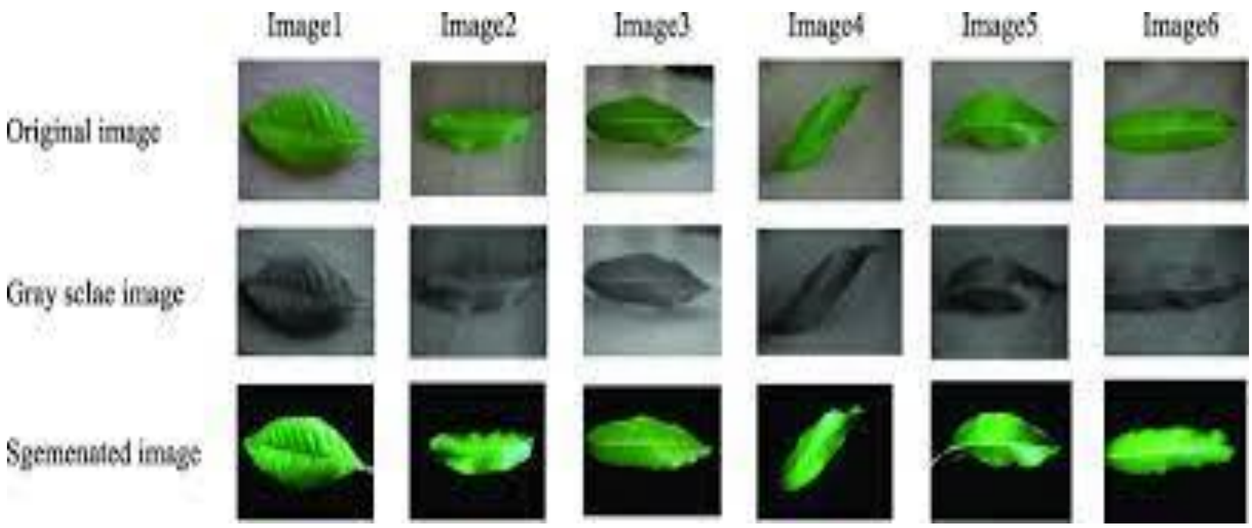

Figure 13: Mango leaf

b. Model Development:

A deep learning model based on convolutional neural networks (CNN) has been designed and implemented. The model architecture consists of several convolutional and pooling layers, followed by fully connected layers for classification. The model is trained on the collected dataset using state-of-the-art optimization algorithms.

Create Model

```
num_classes = 2
model = Sequential([
  Conv2D(16, 3, padding='same', activation='relu',input_shape=(img_height, img_width ,3)),
  MaxPooling2D(),
  Conv2D(32, 3, padding='same', activation='relu'),
  MaxPooling2D(),
  Conv2D(64, 3, padding='same', activation='relu'),
  MaxPooling2D(),
  Dropout(0.2),
  Flatten(),
  Dense(128, activation='relu'),
  Dense(num_classes)
])
model.compile(optimizer='adam',loss=tf.keras.losses.BinaryCrossentropy(from_logits=True),metrics=['accuracy'])
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 180, 180, 16)      448

 max_pooling2d (MaxPooling2D  (None, 90, 90, 16)       0
 )

 conv2d_1 (Conv2D)           (None, 90, 90, 32)        4640

 max_pooling2d_1 (MaxPooling  (None, 45, 45, 32)       0
 2D)

 conv2d_2 (Conv2D)           (None, 45, 45, 64)        18496

 max_pooling2d_2 (MaxPooling  (None, 22, 22, 64)       0
 2D)

 dropout (Dropout)           (None, 22, 22, 64)        0

 flatten (Flatten)           (None, 30976)             0

 dense (Dense)               (None, 128)               3965056

 dense_1 (Dense)             (None, 2)                 258

=================================================================
Total params: 3,988,898
Trainable params: 3,988,898
Non-trainable params: 0
```

Figure 14:Create Model

c. Model Training:

The designed model has been trained using the collected dataset. The training process involved feeding the labeled images to the model and iteratively adjusting its parameters to minimize the prediction errors. This step is crucial for the model to learn the distinguishing features of healthy and diseased mango leaves.

Train Model

```
rlr = ReduceLROnPlateau(patience=10, verbose=1)
es = EarlyStopping(patience=24, restore_best_weights=True, verbose=1)
mc = ModelCheckpoint('/content/mangodetection/model.h5', save_best_only=True, verbose=1)

epochs = 15
history = model.fit(
  train_data_gen,
  callbacks=[rlr, es, mc],
  validation_data=validation_data_gen,
  epochs=epochs
)
```

```
Epoch 1/15
11/11 [==============================] - ETA: 0s - loss: 0.6243 - accuracy: 0.6182
Epoch 1: val_loss improved from inf to 0.51945, saving model to /content/mangodetection/model.h5
11/11 [==============================] - 195s 18s/step - loss: 0.6243 - accuracy: 0.6182 - val_loss: 0.5
Epoch 2/15
 1/11 [=>............................] - ETA: 2:14 - loss: 0.5370 - accuracy: 0.6562
```

Figure 15:Train model

d. Normal Inference of Trained Model:

Once the model is trained, it undergoes normal inference, where it takes in new, unseen images and predicts whether the leaves are healthy or diseased. The inference process involves passing the images through the trained model and generating predictions with corresponding confidence scores. This allows the system to provide accurate and real-time results to the users.

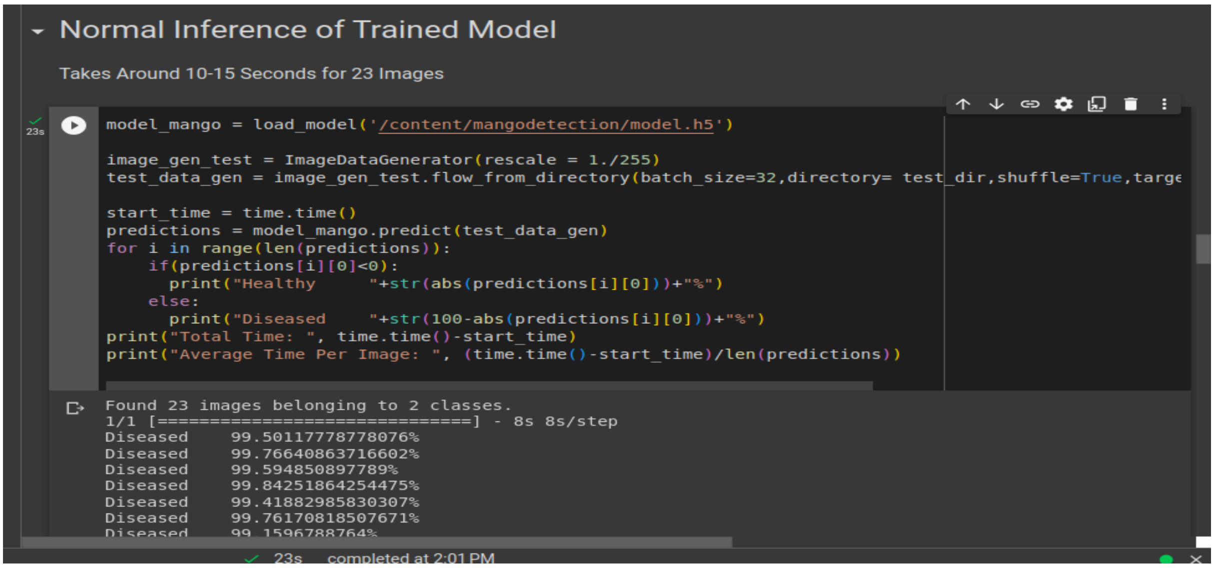

Figure 16:Normal interface and train model

6. CONCLUSION

The proposed methodology in the following plant leaf disease detection system focus on generating an advance and efficient system which makes the process of creating high yield of plants much easier for the farmers. The project aims to detect the most common diseases occurring on a leaf, using image processing technique under up bringing technology i.e., machine learning. In easier terms, the farmer will be able to accurately detect the leaf is healthy or not a particular plant is having using the image of the plant. The proposed system is based on important modules namely:

- Pre-processing.
- Segmentation
- Feature extraction.

In this study, we describe the comparison of our system with preexisting systems with proper methodology and implementation. The proposed systems functionality is better than existing disease detection system as it is able to generate a more accurate and precise result with easier and faster implementation. It aims to make the life of farmers easier. The system can be a boon to the agricultural sector as it advances the crop production and management process, as agriculture is of the major reason to facilitate growth of per capita income of our country [22].

REFERENCE

[1] Sankaran, S.; Mishra, A.; Ehsani, R. “A review of advanced techniques for detecting plant diseases”. Comput.Electron. Agric. 2010, 72, 1–13. [CrossRef]

[2] Chaerani, R.; Voorrips, R.E. “Tomato early blight (Alternariasolani): The pathogens, genetics, and breeding for resistance”, J. Gen. Plant Pathol. 2006, 72, 335–347. [CrossRef]

[3] Alvarez, A.M.” Integrated approaches for detection of plant pathogenic bacteria and diagnosis of bacterial diseases”, Annu. Rev. Phytopathol. 2004, 42, 339–366. [CrossRef] [PubMed]

[4]] Abu Sarwar Zamani, L. Anand et.al, “performance of Machine Learning and Image Processing in plant leaf Disease Detection”, Journal of Food Quality, 26 Apr 2022.

[5]Sushil R. Kamlapurkar,” Detection of Plant Leaf Disease Using Image Processing Approach,” International Journal of Scientific and Research Publications, Volume 6, Issue 2, pp.1-6. February 2016

[6] Shima Ramesh Maniyath, Vinod P V, Hebbar Ram, "Plant Disease Detection Using Machine Learning,"in Research Gate, April 2018. Availabe https://www.researchgate.net/publication/327065422 . Accessed on: May 2023

[7] Usha Kumari, " Leaf Disease Detection: Feature Extraction with K-means clustering and Classification," in Research Gate, March 2019. Available: https://www.researchgate.net/publication/335495508. Accessed on: May 2023

[8] Shitala Prasad, Sateesh K. Peddoju, Debashis Ghosh ,“Multi-resolution mobile vision system for plant leaf disease diagnosis,” pp. 379–388, Springer-Verlag London 2015

[9] Shanwen Zhang, Zhuhong You, Xiaowei Wu,” Plant disease leaf image segmentation based on superpixel clustering and EM algorithm,” Springer, June 2017.

[10] Zulkifli Bin Husin, Abdul Hallis Bin Abdul Aziz, Ali Yeon Bin Md Shakaff Rohani Binti S Mohamed Farook, “Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques”, 2012 Third International Conference on Intelligent Systems Modelling and Simulation.

[11] Wenjiang Huang, Qingsong Guan, Juhua Luo, Jingcheng Zhang, Jinling Zhao, Dong Liang, Linsheng Huang, and Dongyan Zhang,“New Optimized Spectral Indices for Identifying and Monitoring Winter Wheat Diseases”, IEEE journal of selected topics inapplied earth observation and remote sensing, Vol. 7, No. 6, June 2017

[12] Keyvan Asefpour Vakilian & Jafar Massah,” An artificial neural network approach to identify fungal diseases of cucumber (Cucumis sativus L.) Plants using digital image processing,” Vol. 46, No. 13,1580–1588, Taylor &Francis, 2013

[13] : Vijai Singh, A.K. Misra,” Detection of plant leaf diseases using image segmentation and soft computing Techniques,”Information Processing In Agriculture 4 (2017) 41–49 , science direct, 2017

[14] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic,”Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification,”

Hindawi Publishing Corporation Computational Intelligence and Neuroscience,Vol 2016, Article ID 3289801, 11 pages
[15] Usama Mokhtar, Mona A. S. Ali, Aboul Ella Hassenian, Hesham Hefny," Tomato leaves diseases detection approach based on support vector machines," IEEE,2015
[16] Amar Kumar Dey, Manisha Sharma, M.R.Meshram,"Image Processing Based Leaf Rot Disease, Detection of Betel Vine (Piper BetleL.)," Procedia Computer Science 748 – 754, science direct, 2016
[17] no.4, 299–315, Taylor & Francis, 2017[15] H.Al-Hiary, S. Bani-Ahmad, M.Reyalat, M.Braik & Z.AlRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", International Journal of Computer Applications, Vol.17,No.1, pp.31-38.March 2011.
[18]A. R. Bhagat Patil1, Lokesh Sharma, "A Literature Review on Detection of Plant Diseases", European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 7, Issue 07, 2020.
[19] Md. Rasel Mia1 · Sujit Roy1,2 · Subrata Kumar Das2 · Md. Atikur Rahman1 Received: 4 February 2019 / Accepted: 21 March 2020
[20] A survey on crop disease detection using image processing technique for economic growth of rural area. Yashpal Sen1, Chandra Shekhar Mithlesh2, Dr. Vivek Baghel3
[21]K. Elangoran, S. Nalini, 2011 "Detection and classification of leaf diseases using K- means-based segmentation and neural-networks-based classification." Inform. Technol. J., 10: 267-275. DOI: 10.3923/itj.2011.267.275.
[22] Plant Disease Detection and Classification using CNN Rinu R, Manjula S H
[23]Detection of Plant Leaf Disease Using CNN Algorithm,Nageshwar Jaiswal1, and Vivek Sarnaik2