# **PROFORMA FOR THE APPROVAL**

**PROJECT PROPOSAL**

PRN No:- **2021016400296187** Seat No**:-1172759**

1. Name of the Student:- **Farooqui Mohd. Arshad Mohd Sharmad**
2. Title of the Project:- **Plant Disease Detection Using Artificial . Intelligence**

1. Name of the Guide:- **Assistant Prof. Simran Shaikh**

1. Is this your first submission? Yes

.

Signature of the Student Signature of the Guide

Date: Date:

.

Signature of the Coordinator

Date:

**Plant Disease Detection**

**Using Artificial Intelligence**

## **A Project Report**

Submitted in partial fulfillment of the

Requirements for the award of the Degree of

**BACHELOR OF SCIENCE (INFORMATION TECHNOLOGY)**

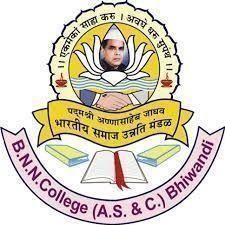
**By**

**Farooqui Mohd Arshad Mohd Sharmad**

**Seat Number : 1172759**

**Under the esteemed guidance of**

***Assist. Prof. Simran Shaikh***



**DEPARTMENT OF INFORMATION TECHNOLOGY**

**B.N.N COLLEGE (ARTS, SCIENCE AND COMMERCE)**

***(Affiliated to University of Mumbai)***

**BHIWANDI, 421302**

**MAHARASHTRA**

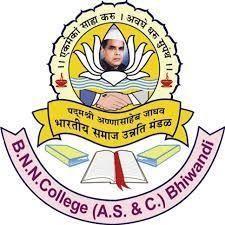
**2023-2024**

**B.N.N COLLEGE (ARTS, SCIENCE AND COMMERCE)**

***(Affiliated to University of Mumbai)***

**BHIWANDI-MAHARASHTRA-421302**

**DEPARTMENT OF INFORMATION TECHNOLOGY**



**CERTIFICATE**

This is to certify that the project entitled, **"Plant Disease Detection Using Artificial Intelligence"**, is bonafied work of “**Farooqui Mohd Arshad”** bearing Seat No. **“1172759”**submitted in partial fulfillment of the requirements for the award of degree of BACHELOR OF SCIENCE INFORMATIONTECHNOLOGY from University of Mumbai.

**.**

**Internal Guide Coordinator**

**.**

**External Examiner**

**Date: College Seal**

**ACKNOWLEDGEMENT**

This project arose in part out of months of efforts that have been done since we came to final year. By that time, we have worked with a great number of people whose contributions in different ways to the research and the making of this thesis deserved special mention. It is our great

pleasure to convey gratitude to them all in our humble acknowledgment. In the first place we would like to record my gratitude to our guide, Assist. Prof.Simran Shaikh, Department of Information Technology, B.N.N college, Bhiwandi, Maharashtra for her supervision, advice, and valuable guidance. She provided us unflinching encouragement and support in various ways.

It is because of her that we could synchronize our effort her support and dynamic supervision, which exceptionally inspire and enrich our growth as a student.

We also thank to respected Prof. Pramod Shivale, HOD, B.N.N college,Bhiwandi, Maharashtra, for their constant support.

We are also thankful to all the teaching and non-teaching staff and our batch mates who have directly or indirectly contributed in the word

**DECLARATION**

We hereby declare that the project entitled, “**Plant Disease Detection Using Artificial Intelligence**” done at **“BNN COLLEGE BHIWANDI”** Thane – 421302, has not been in any case duplicated to submit to any other university for the award of any degree. To the best of our knowledge other than me, no one has submitted to any other university.

The project is done in partial fulfillment of the requirements for the award of degree of **MASTER OF SCIENCE (INFORMATION TECHNOLOGY)** to be submitted as a final semester project as part of our curriculum.

**Farooqui Mohd Arshad**

**Contents**

Introduction.....................................................................................................................................7

1.1 Abstract...............................................................................................................................8

1.2 Objectives.......................................................................................................................9

1.3 Purpose……………………………...................................................................................9

1.4 Problem Definition.............................................................................................................10

2.1Survey of Technology.........................................................................................................11

RequirementsAnalysis...............................................................................................................12

3.1Software and Hardware Requirements………………………………………………..12

3.2 Requirements Specification.........................................................................................13

3.3 Conceptual Models.........................................................................................................13

Literature Survey…………………………………………………………………………………….14

4.1Research……………………………………………………………………………………….14

Materials and Methods…………………………………………………………………………….16

5.1 Data Acquisition………………………………………………………………………………….16

5.2 B. Data Pre-Processing…………………………………………………………………….17

6.1Classification by CNN…………………………………………………………………………18

Sample Images From The Datasets……………………………………………………………..21

7.1Processing method:.....................................................................................................23

8.1Sample screen shot :.......................................................................................................24

9.1Results…………………………………………………………………………………………..26

10.Conclusion………………………………………………………………………………………27  
11. Refrences………………………………………………………………………………………..28

**INTRODUCTION**

## **1.1** **Background**

The most widely used method for plant disease detection is simply naked eye observation by experts through which identification and detection of plants diseases are done. For doing so, a large team of experts is required, which costs very high when farms are large.

At the same time, in some countries, farmers don’t have proper facilities or even idea that they can contact to experts. Due to which consulting experts even cost high as well as time consuming too. In such a condition, the suggested technique proves to be beneficial in monitoring large fields of crops. And automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it by easier as well as cheaper.

Plant disease identification by the visual way is a 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 fewer efforts, less time and more accurately. In plants, some general diseases are bacterial, black spotted, and others are Rust, viral and Red cotton Leaf.

Image processing is the technique which is used for measuring the affected area of disease, and to determine the difference in the color of the affected area [1]. 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. The segmentation process is based on various features found in the image. This might be color information, boundaries or segment of an image.

**1.2 ABSTRACT**:

Plant disease is an ongoing challenge for smallholder farmers, which threatens income and food security. The recent revolution in smartphone penetration and computer vision methods has created an opportunity for image classification in agriculture.

Convolutional Neural Networks(CNN) are considered state of the art in image recognition and offer the ability to provide a prompt and definite diagnosis.

In this paper, the performance of a pre-trained ResNet34 model in detecting crop disease is investigated. The developed model is deployed as a web application and is capable of recognizing 7 plant diseases out of healthy leaf tissue.

The project focuses on the approach based on image processing for detection of diseases of plants. In this paper, we propose an Android application that helps farmers for identifying plant disease by uploading a leaf image to the system.

The system has a set of algorithms which can identify the type of disease. Input image given by the user undergoes several processing steps to detect the disease and results are returned back to the user via android application.

**1.3 Objective:**

The primary objective of plant disease detection is to identify and diagnose plant diseases accurately and quickly. This is important to prevent the spread of the disease and to minimize crop loss.

Early detection of plant diseases is crucial to prevent the spread of the disease and to minimize crop damage. Disease detection using Machine Learning Algorithms should be reliable and produce consistent results.

**1.4 Purpose of Plant Leaf Disease:**

A disturbance in the stock situation of a herb that destroys or alters crucial. All types of herbs, both unbroken and educated, are susceptible to illness. Each family is prone to specific diseases, but each of these is relatively rare.

The incidence and prevalence of plant illness vary from prime to prime, pivoting on the company of pathogens, territory conditions, and the supply and cultivars grown. Some plant varieties are particularly susceptible to disease epidemics, while others are more resilient. See also list of herb illness

**1.5 Definitions of plant disease** :

Generally, when plants are consistently disturbed, they might catch illness. pathogens that result in abnormal physiological processes that disrupt the normal structure, growth, function, or other activities of the plant. The essential physiological or biochemical processes of the plant are disrupted, which results in the typical diseased states or symptoms. Depending on whether the primary cause of the disease is infectious or noninfectious, plant diseases may be broadly categorised. Infectious plant diseases are caused by pathogens such as fungi, bacteria, mycoplasma, viruses, viroids, nematodes, or parasitic flowering plants. Within or on a host, infectious organisms can grow and spread from one vulnerable host to another.

Unfavorable growth circumstances, including as excessive temperatures, unfavorable moisture-oxygen ratios, soil and air pollutants, and an abundance or shortage of vital minerals, are the root causes of non**-infectious plant illnesses**

**2.1 Survey of Technology**

The technologies used to make this project are:

1. **Python :**

• Python is an interpreted, high-level, general-purpose programming language.

• Python is easy to learn and use, and its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

• Python is widely use for web development, software development, data science, machine learning and more.

1. **Tensorflow-python :**

* TensorFlow is an open-source library developed by Google primarily for deep learning library
* It also supports traditional machine learning.
* TensorFlow can be used to develop models for various tasks, including natural language processing, image recognition, handwriting recognition, and different computational-based simulations such as partial differential equations.
* Tensorflow is a powerful tool that can be used to explore and understand data.

1. **Flask-python :**

* Flask is a small and lightweight Python web framework that provides useful tools and features that make creating web applications in Python easier
* It gives developers flexibility and is a more accessible framework for new developers
* There is a built-in development server and a fast debugger provided. Lightweight.

**Requirements Analysis**

**3.1 Software and Hardware Requirements**

**Hardware Requirements:**

**Process:** 11th Gen Intel Core i5-11320 @ 3.20GHz or above

**RAM:** 8 or 16 GB or above

**Storage:** 500 GB or above

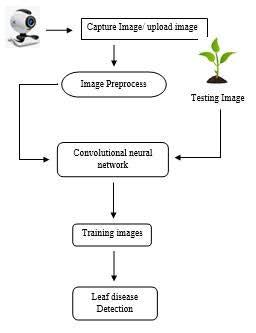
**Software Requirements:**

**Operating System:** Windows 11

**Programming Language:** Python

**Code Editor:** Visual Studio Code or Jupyter Notebook or any python idle

**3.2 Conceptual Models :**



**Literature Survey**

**4.1 Research**

Extensive research has been conducted to explore various methods for automated identification of plant diseases. The disease can manifest in various parts of the plant such as roots, stem, fruit or leaves. As stated before, this work concentrates, particularly on leaves.

Discussed a methodology for recognition of plant diseases present on leaves and stem. The proposed work is composed of K-Means segmentation technique and the segmented images are classified using a neural network. They developed a method for detecting the visual signs of plant diseases by using the image processing algorithm. The accuracy of the algorithm was tested by comparing the images, which were segmented manually with those automatically segmented.

Discussed various techniques to segment the diseased part of the plant. This paper also discussed some Feature extraction and classification techniques to extract the features of infected leaf and the classification of plant diseases.

The use of ANN methods for classification of disease in plants such as self-organizing feature map, back propagation algorithm, SVMs, etc. can be efficiently used. From these methods, we can accurately identify and classify various plant diseases using image processing techniques.

An approach based on image processing is used for automated plant diseases classification based on leaf image processing the research work is concerned with the discrimination between diseased and healthy soybean leaves using SVM classifier. They have tested our algorithm over the database of 120 images taken directly from different farms using different mobile cameras.

The SIFT algorithm enables to correctly recognize the plant species based on the leaf shape. The SVM classifier can help in recognizing normal and Diseased soybean leaves with an average accuracy as high as 93.79%.

The main aim of the proposed work is to provide inputs to an autonomous DSS which will provide necessary help to the farmers as and when required over the mobile.

This system will provide help to the farmer with minimal efforts. The farmer only needs to capture the image of the plant leaf using a mobile camera and send it to the DSS, without any additional inputs.

The work represents groundnut leaf disease extraction and classification using color imagery. The color imaginary transform, color co-occurrence matrix, feature extraction will be done and get an efficiency output with a neural network, Back propagation gives efficient groundnut leaf detection with a complex background, in this work we classified only four different diseases with 97 Al % of efficiency. But in the future, the work carried out more diseases by using this method.

Contain the study of detection of plant diseases and the detection of the infected part of plants. Initially, input images are taken and then image processing is started. Background and Black pixels are both segmented in the first step. Then Hue and Saturation part of the image is also separated. And finally infected part and infected area % and a name of the disease is acquired which is main work using our proposed methodology.

The main aim of this work is to provide the advancement and enhancement in computing classifiers of a neural network approach and provide better results. This study contains a unique work that is it will calculate the % of an infected area of plants.

**MATERIALS AND METHODS**:

This section describes the steps involved in creating and deploying the classifier. Classification by CNN is divided into three phases which tackle separate tasks.

All work involved in this research was completed on one machine, with specifications listed.

**5.1 Data Acquisition**

All Potato and Tomato imagery derive from ‘The PlantVilllage Dataset’ [35], an open-access repository which contains in total 54,323 images. All Rice imagery originates from the “Rice Diseases Image Dataset” Kaggle dataset [36].

For each species, a select number of classes are chosen, with details viewable in Table II.

All images are captured in a controlled environment. Due to this, model bias is expected. To access this, a test dataset containing 50 images, sourced from Google is also established. These images contain additional plant anatomy, infield background data and varying stages of disease.

**5.2 Data Pre-Processing**.

The dataset is divided into 80% for training and 20% for validation.

First, augmentation settings are applied to the training data. These are generated ‘on the fly’, with each operation carrying a weighted probability of appearing in each epoch.

Each directory of the plant disease dataset folder varies in the number of images.

Instead of taking them all, we select the first N\_IMAGES from each directory to train our model. Finally, we set the path of the dataset in the root\_dir to access plant images.

The settings applied include flipping (random), padding mode (reflection) and zoom with crop (scale = (1.0,1.5)). ‘Zoom with crop’ was later omitted after discovering that it had inappropriately cropped areas of infected leaf.

Finally, all images are re-sized and normalized. Resizing is carried out using a compress function, to 150 x 150. As a pretrained model is used, the RBG ImageNet statistics are used to normalize. A sample of the final pre-processed images is viewable.

**Classification by CNN**

**1) Phase One – Trialling of Image size**

To begin, the Resnet34 pre-trained weights are downloaded. As a default of transfer learning, all layers with the except of the final two layers are frozen. These contain new weights and are specific to the plant disease classification task. Freezing allows these layers to be disease separately trained, without backpropagating the gradients. In exactly this way, the 1cycle policy is used to train the final layers.

With this complete, the remaining layers are released. To aid the fine-tuning process, a plot displaying learning rate vs loss is generated and analysed. From this, a suitable learning is selected, and the model is run.

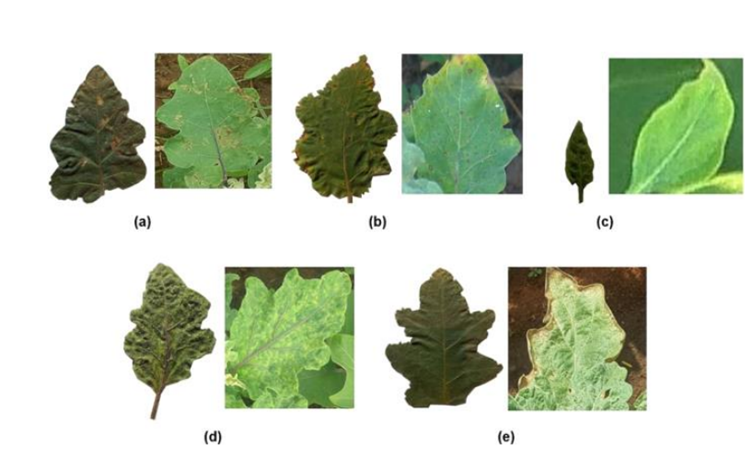
**2) Phase Two** –

Model Optimisation Using the most suitable image size, the ResNet34 model is optimised. To further improve the model’s performance, additional augmentation settings are added . Operations include brightness changes

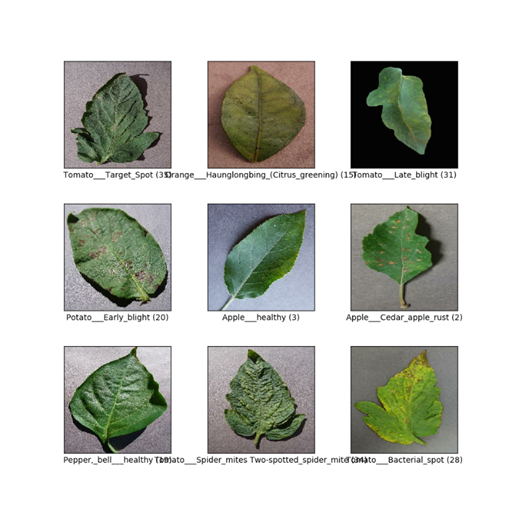
(0.4,0.7) and warp (0.5).

Next, the final two layers are isolated and trained at the default learning rate. With this complete, fine tuning is performed, running multiple trials to test a series of learning rates and number of epochs. Pre-processed image - augmentation settings = flipping (random), padding mode (reflection).

(I) Disease Classification in Eggplant Using Pre-trained VGG16 and MSVM

****

(a) Epilachna bettle, (b) Cercospora leaf spot, (c) Little leaf disease, (d) Tobacco Mosaic Virus (TMV), (e) Two spotted spider mite.



**3) Phase Three**:

Visualisations For the purpose of interpretation, a series of visualisations are generated based on the validation and test datasets. Additionally, the model is deployed to create a web application.

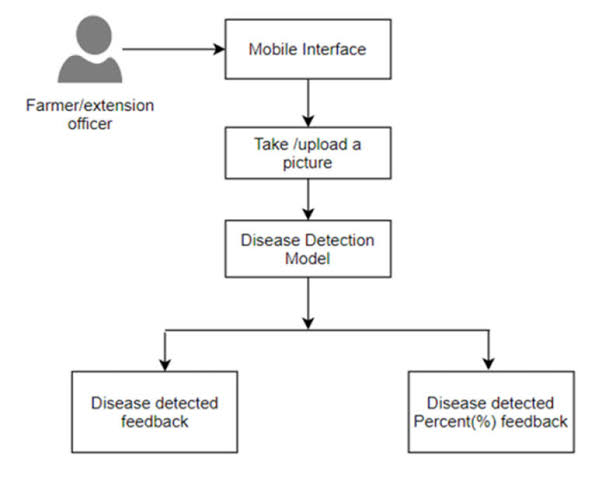
To achieve this, the completed essential files are stored in a GitHub repository and the model is exported as a pickle file.

To deploy the model, the repository is connected to the unified platform; Render. In carrying out this task, the ‘Render Examples’ GitHub repository was used as a guide

**4.3.3 Activity Diagram :**

Another crucial UML diagram for describing the system's dynamic elements is the activity diagram. An activity diagram is essentially a flowchart used to show how an activity flows from one to the next. One could refer to the activity as a system operationThe control flow moves from one function to the next.

This flow may occur concurrently, forked, or sequentially. Activity diagrams include a variety of elements, including join, fork, and others, to handle various forms of flow control.



**1.8 SAMPLE IMAGES FROM THE DATASETS**

****

**Processing method:**

For this project we have used plant village dataset. The Plant Village dataset consists of more than 54,000 healthy and unhealthy leaf images divided into 38 categories by species and disease. However, the original dataset is no longer available from the original source, their website (plantvillage.org).

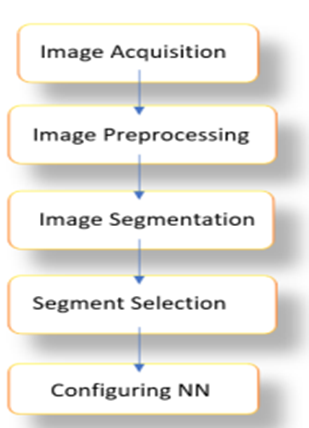
Therefore, in this study, the not augmented dataset from Pandian et al. was used. In this dataset, 39 different classes plant leaf and background images are available.

The dataset contains 61,486 images. Some of its classes are the various diseases of plants like apple, blueberry, cherry, grapes, corn, orange, peach, pepper, potato, raspberry and soybean. Sample images are shown below.  
  
For this project we have used CNN method to detect the leaf diseases. In training the model, only the most available plant leaves in the Philippines were utilized. The training images include strawberry, peach, bell pepper, Tomato, and cherry. 20 images were used to represent each plant disease and healthy images for specificplant.

In total, there are approximately 500 images used for training the model. Included in the dataset are the non-leaf images to enhance the recognition feature of the model.

Configuring Convolutional Neural Network: Convolutional neural networks are great for photo tagging and recognizing patterns in image data. Here we will build Convolutional Neural Network in TensorFlow Framework.

There is total five steps involved in Disease detection they are as follows: - 1) image acquisition 2) image preprocessing 3) Image Segmentation 4) Segment Selection 5) Configure NN

****

**Image Acquisition**: We need various images of leaves to train the model. So, we acquire high quality images from different sources and categorize them into different categories. For now, we will be using Plant Village dataset from Kaggle. It has up to 15 different directories and having up to 20K images. Hence, we can directly train our model with this dataset and use for predictions. Later we also can add different images and train model again to improve model

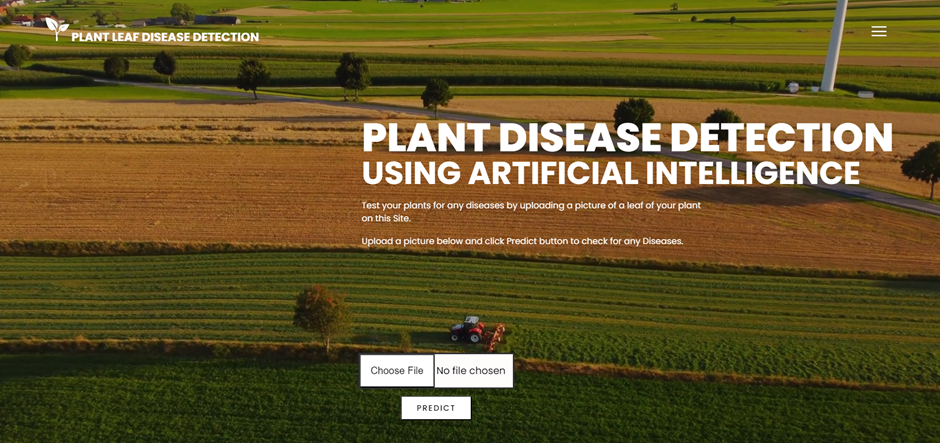
**Image Preprocessing:** The aim of Image pre-processing is an improvement of the image data that suppresses unwilling distortions or enhances some image features important for further processing, although geometric transformations of images (e.g., rotation, scaling, translation) are classified among pre-processing methods.

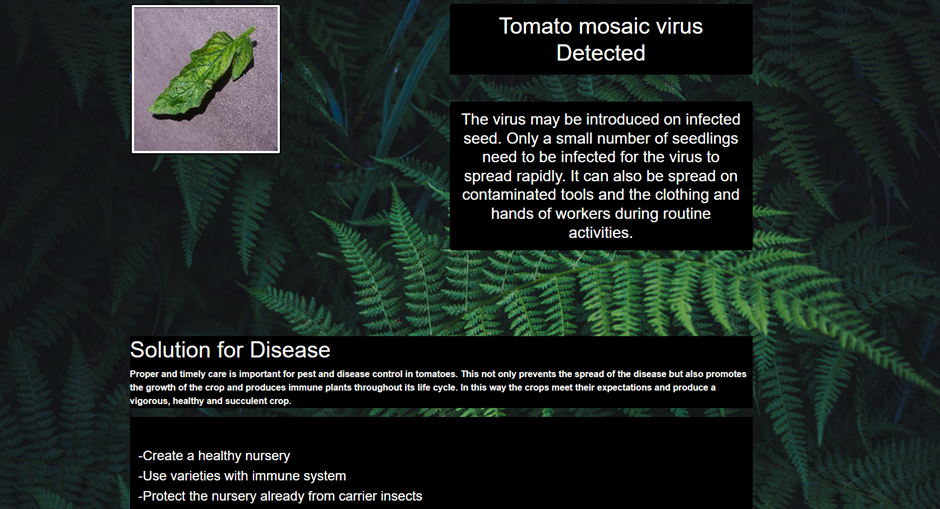
The aim of image processing is to enhance the image data (features) by suppressing unwanted distortions and/or enhancement of some important image features in order that our models can benefit from this improved data to make computation on data.

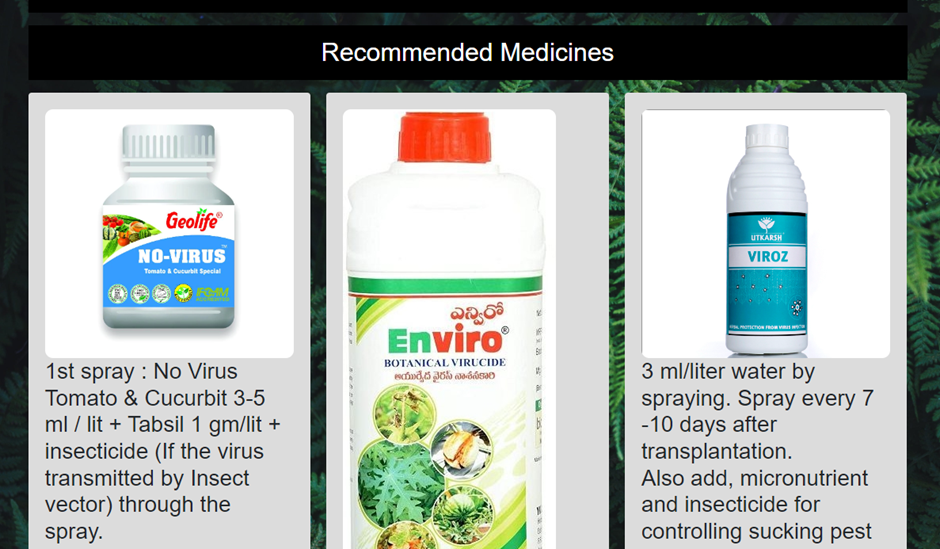
In this phase we usually improve image quality and we make the color correction to highlight feature so that they could be simply catch by the classifier. We also reduce the size of image to certain limit so that we can convert the image to array and feed it to neural network Image is pre-processed to enhance the image data that suppress undesired distortions, enhances some image features important for further processing and analysis task. It includes color space conversion, image enhancement.

**Image Segmentation:** Image segmentation is process of finding boundaries of objects in images and enhance those boundaries in such way that they can become more highlighted and could be capture immediately by classifier.  
Image segmentation is best technique to improve image and characteristics of image. It also works well on low resolution and low-quality images.

**1.9 Sample screen shot :**

****

****



**RESULTS:**

All the experiments are performed in MATLAB. For input data disease, samples of plant leaves like rose with bacterial disease, beans leaf with bacterial disease, lemon leaf with Sun burn disease, banana leaf with early scorch disease and fungal disease in beans leaf are considered. Original images which are followed by output segmented images. Segmented image can be classified into different plant diseases.

The input and output image where input image is a banana leaf with early scorch disease and output image shows the classification of disease using feature extraction method.

The co-occurrence features are calculated after mapping the R, G, B components of the input image to the thresholded images. The co-occurrence features for the leaves are extracted and compared with the corresponding feature values that are stored in the feature library.

The classification is first done using the Minimum Distance Criterion with K-Mean Clustering and shows its efficiency with accuracy of 86.54%. The detection accuracy is improved by proposed algorithm. Therefore, compared with the new learning, this paper uses the transfer learning to converge faster and achieve better model identification effect.

It can meet the requirements of smart agriculture for low hardware resources, fast training time, and high training efficiency. Then, the image is input into transfer learning model based on the segmentation of Chan–Vese algorithm. As a contrast, the image that has not been processed in this paper is input.

**CONCLUSION :**

With very less computational efforts the optimum results were obtained, which also shows the efficiency of proposed algorithm in recognition and classification of the leaf diseases. Another advantage of using this method is that the plant diseases can be identified at early stage or the initial stage.

Banana, beans, jackfruit, lemon, mango, potato, tomato, and sapota are some of those ten species on which proposed algorithm is tested. Therefore, related diseases for these plants were taken for identification. T

he model not only adapts to complex environments, but also increases the accuracy of identification. Compared with the traditional model, the model proposed in this paper not only guarantees the robustness of the convolutional neural network, but also reduces the number and quality requirements of the convolutional neural network on the data set and obtains better results.

The model which overcomes the problem of environment complexity can get an accurate identification result in practical application. Furthermore, this study enriches the existing theory and helps to improve the accuracy.

At the same time, it is of great significance for the study of plant disease identification in the field of environmental complexity and helps researchers pay attention to the important role of environmental complexity in plant disease identification.

Therefore, the model applies information technology to agricultural production and is favorable to sustainable development of smart agriculture.

**REFERENCES**:

1. Savita N. Ghaiwat, Parul Arora Detection and classification of plant leaf diseases using image processing techniques: a review Int J Recent Adv Eng Technol, 2 (3) (2014), pp. 2347-2812 ISSN (Online)

2. Sanjay B. Dhaygude, Nitin P. Kumbhar Agricultural plant leaf disease detection using image processing Int J Adv Res Electr Electron Instrum Eng, 2 (1) (2013)

3. R. Badnakhe Mrunalini, Prashant R. Deshmukh An application of K-means clustering and artificial intelligence in pattern recognition for crop diseases Int Conf Adv Inf Technol, 20 (2011) 2011 IPCSIT

4. S. Arivazhagan, R. Newlin Shebiah, S. Ananthi, S. Vishnu Varthini Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features Agric Eng Int CIGR, 15 (1) (2013), pp. 211-217

5. Anand H. Kulkarni, R.K. Ashwin Patil Applying image processing technique to detect plant diseases Int J Mod Eng Res, 2 (5) (2012), pp. 3661-3664

6. J. G. A. Barbedo, “Factors influencing the use of deep learning for plant disease recognition,” Biosystems Engineering, vol. 172, pp. 84–91, 2018.

7. G. Geetharamani and J. Arun Pandian, “Identification of plant leaf diseases using a nine-layer deep convolutional neural network,” Computers & Electrical Engineering, vol. 76, pp. 323–338, 2019.

8. P. F. Konstantinos, “Deep learning models for plant disease detection and diagnosis,” Computers & Electrical Engineering, vol. 145, pp. 311–318, 2018.