ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY

PROJECT PROPOSAL

INTEGRATED ENGINEERING TEAM PROJECT (IETP4115)

# Plant Disease Detection Using Image Processing and Machine Learning

By: Fransi Ayele 2023

# Table of Contents

EXECUTIVE SUMMARY ........ 3
INTRODUCTION ........ 3
PROBLEM STATEMENT ........ 4
PROJECT BACKGROUD ........ 4
PROJECT OBJECTIVE ........ 4
SCOPE AND LIMITATION OF THE PROJECT ........ 5
SIGNIFICANCE OF THE PROJECT ........ 5
BENEFICIARIES OF THE PROJECT ........ 6
METHODOLOGY ........ 6
Fig:2 Architecture of the system ........ 8
Timeline ........ 8
Budget ........ 9
RESPONSIBILITIES AND WORK ALLOCATIONS ........ 10
DOCUMENTATION AND REFERENCE ........ 11

## EXECUTIVE SUMMARY

One of the important and tedious tasks in agricultural practices is detection of disease on plants. It requires huge time as well as skilled labor. This paper proposes a smart and efficient technique for detection of plant disease which uses computer vision and machine learning techniques. The proposed system helps in identification of plant disease and provides remedies that can be used as a defense mechanism against the disease. The dataset obtained from the Internet is properly segregated and the different plant species are identified and are labeled. The proposed system is able to detect 15 different diseases of 3 common plants (Pepper, Tomato and Potato) with average 90.4% of accuracy.

## INTRODUCTION

Identification of the plant diseases is important in order to prevent the losses within the yield. It's terribly troublesome to observe the plant diseases manually. It needs tremendous quantity of labor, expertise within the plant diseases, and conjointly need the excessive time interval. Hence, image processing and machine learning models can be employed for the detection of plant diseases. In this project, we have described the technique for the detection of plant diseases with the help of their leaf's pictures. Image processing is a branch of signal processing which can extract the image properties or useful information from the image. Machine learning is a sub part of artificial intelligence which works automatically or give instructions to do a particular task. The main aim of machine learning is to understand the training data and fit that training data into models that should be useful to the people. So, it can assist in good decisions making and predicting the correct output using the large amount of training data. The color of leaves, amount of damage to leaves, area of the leaf, texture parameters are used for classification. In this project we have analyzed different image parameters or features to identifying different plant leaves diseases to achieve the best accuracy. Previously plant disease detection is done by visual inspection of the leaves or some chemical processes by experts. For doing so, a large team of experts as well as continuous observation of plant is needed, which costs high when we do with large farms. In such conditions, the recommended system proves to be helpful in monitoring large fields of crops. Automatic detection of the diseases by simply seeing the symptoms on the plant leaves makes it easier as well as cheaper. The proposed solution for plant disease detection is computationally less expensive and requires less time for prediction than other deep learning-based approaches since it uses statistical machine learning and image processing algorithm. The main reason that we chose this project is the huge loss of time and resource due to the lack of technological advancements in plant farms and related industries.

## PROBLEM STATEMENT

Manual examination of the disease signs found on various plant components, particularly leaves, is used to identify plant diseases. Farmers overuse pesticides, insecticides, and other chemicals to cure plant illnesses due to the method's complexity and lack of awareness of the diseases that harm plants. Therefore, it is imperative that diseases be diagnosed using the ever-evolving technology, so that the proper quantity and kind of chemicals may be administered. As the majority of plant diseases only affect leaves, this project discusses various machine learning and image processing-based techniques that have been proposed in various literature for treating plant diseases. The methods and techniques covered in this project include those that can identify plant diseases through images of the leaves of affected plants.

## PROJECT BACKGROUD

Detecting plant diseases is a key area of research in the science of machine vision. It is a technique that takes plant photos and predict whether they contain diseases using the algorithm called convolutional neural network (CNN). Plant diseases detection tools based on machine vision are currently being used in agriculture and have partially replaced the old-fashioned naked eye identification methods.

Common datasets are presented, and the effectiveness of previous studies is contrasted. On the basis of this, this project examines potential difficulties in deep learning-based plant disease identification in real-world settings. In addition, various recommendations are made as well as potential research directions and remedies for the problems.

## PROJECT OBJECTIVE

### General Objective of the project

The main objective is to identify the plant diseases using image processing. It also, after identification of the disease, suggest the remedy and treatment to be used. It completes each of the process sequentially and hence achieving each of the output. The systems' main purpose is to detect the plant disease and inform the user with supplementary precautions and treatment mechanisms in order to save time, expenses and environment related issues

### Specific Objective of the project

1. To design such system that can detect plant disease accurately

By the use of deep learning technique and a well-defined collection of data we set out to achieve a system that is able to capture an image of a specimen plant under analysis, and from its organized data, can accurately identify the specific disease (if any).

2. To provide remedy for the disease that is detected

3. After the accurate identification of the plant disease, the system will recommend a treatment viable to the case of disease detected in the first session.

## SCOPE AND LIMITATION OF THE PROJECT

The scope of this project extends to mostly farmers and large-scale plantations, but with some future modifications and advancements we can reach a much larger user base. The main beneficiaries from this technology will be, local farmers, small in-house gardens, distributers, health and inspection organs, the day-to-day consumer, and the general society as a whole.

Some of the major limitations faced in the rising of this project will be mostly related to sufficiency of acquiring dataset, difficulty in comprised more and more species, expenses to develop the whole system, time taken to training the system as well as the user. As the project commences, we will also presume to face some technical limitations and difficulties related with system operations and some functionalities.

## SIGNIFICANCE OF THE PROJECT

This project is a huge step up towards modern farming and health stability to any environment it is introduced to as the mechanisms and principles are designed to be eco-friendly and beneficiary to all subjects involved.

At the end of the project, we are aiming to solve the listed goals:

1. Good health and wellbeing

Plant diseases are well known to reduce the food available to humans by ultimately interfering with crop yields. This can result in inadequate food to humans or lead to starvation and death in the worst cases. For example, late blight disease of potato, which is caused by Phytophthora infestans , destroyed potatoes which were the main crop in Ireland during 1845–1850. This resulted in the Great Famine (or Great Hunger), where about one million people died and another million emigrated to Canada, the USA and other countries.

With the help of our plant disease detection system, we can overcome the complications of this tragic events by detecting the disease in its early stage and providing the necessary treatment mechanisms to prevent outcomes of disaster that are listed above.

2. Climate action

It is a fact that plant is mostly directly related to the climate, global warming and the shearing of the ozone layer is one of the huge merits of how devastating the plant life will sequentially bring up on a negative impact on the climate.

Plant disease plays its own role in changing the climate in a negative manner, it will prevent the plant from fulfilling the necessary environmental attributes and hence affecting the ecological balance as a whole.

Our aim is to implement our knowledge as engineers to prevent and perhaps control the relation between climate and plant disease in a positive way of manipulation.

## BENEFICIARIES OF THE PROJECT

This project mostly benefits farmers and large-scale plantations, but with some future modifications and advancements we can reach a much larger user base. The main beneficiaries from this technology will be, local farmers, small in-house gardens, distributers, health and inspection organs, the day-to-day consumer, and the general society as a whole.

As a main aim in solving health, wellbeing and climate issues, we are also concerned in benefitting those greatly affected by the previous lack of the know-how and right tools to prevent the disasters that occurred. We focus in helping those directly and indirectly affected by the merits of plant disease as explained in the previous section.

## METHODOLOGY

1. Dataset

For this project we have used public dataset for plant leaf disease detection called PlantVillage contributed by **Tairu Oluwafemi Emmanuel ,Bulent Siyah** and **saisravan medicherla**
The dataset consists of 20,000 RGB images of healthy and unhealthy plant leaves having 16 classes for experimentation of our algorithm These classes are

| Plant | Disease Name | No of image |
|---|---|---|
| | Pepper__bell___Bacterial_spot | 997 |
| | Pepper__bell___healthy | 1478 |
| Pepper | Potato___Early_blight | 1000 |
| Potato | Potato___healthy | 152 |
| Tomato | Potato___Late_blight | 1000 |
| | Tomato__Target_Spot | 1404 |
| | Tomato__Tomato_mosaic_virus | 373 |
| | Tomato__Tomato_YellowLeaf__Curl_Virus | 3209 |
| | Tomato_Bacterial_spot | 2127 |
| | Tomato_Early_blight | 1000 |
| | Tomato_healthy | 1591 |
| | Tomato_Late_blight | 1909 |
| | Tomato_Leaf_Mold | 952 |
| | Tomato_Septoria_leaf_spot | 1771 |
| | Tomato_Spider_mites_Two_spotted_spider_mite | 1676 |

2. Data preprocessing and feature extraction

Data preprocessing is important task in any computer vision-based system. Fig. 2 illustrates the preprocessing steps for each image. To get precise results, some background noise should be removed before extraction of features. So first the RGB image is converted to greyscale and then

Gaussian filter is used for smoothening of the image. Then to binaries the image, Otsu's thresholding algorithm is implemented. Then morphological transform is applied on binarised image to close the small holes in the foreground part. Now after foreground detection, the bitwise AND operation on binarised image and original color image is performed to get RGB image of segmented leaf. Now after image segmentation shape, texture and color features are extracted from the image. By using contours, area of the leaf and perimeter of the leaf is calculated. Contours are the line that joins all the points along the edges of objects having same color or intensity. Mean and standard deviation of each channel in RGB image is also estimated. To obtain amount of green color in the image, image is first converted to HSV color space and we have calculated the ratio of number of pixels having pixel intensity of hue (H) channel in between 30 and 70 and total number of pixels in one channel. Non green part of image is calculated by subtracting green color part from 1. After extracting color features from the image, we have extracted texture features from grey level co-occurrence matrix (GLCM) of the image [7].

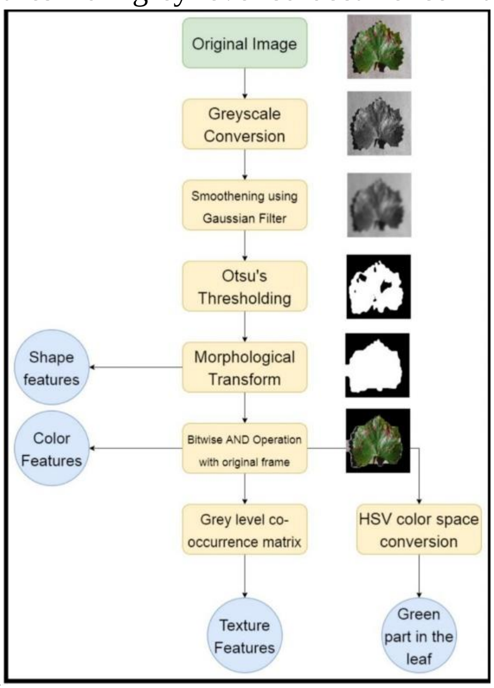


Fig:1 Data preprocessing and feature extraction

3. Classification Algorithm

Convolutional neural network classifier has been used for classification or detection task a CNN is a kind of network architecture for learning algorithms and is specifically used for image recognition and tasks that involve the processing pixel data. Generally, to achieve higher accuracies, different architectures could be used. This can reduce the overfitting and improves the accuracy of the classifier. We are going to split the dataset into train set (80%) for fitting the model and test set (20%) for validation. K-fold cross validation technique is implemented to find the accuracy score. This method can find the accuracy on whole dataset without any bias. After fitting the data, f1 score, precision, recall, accuracy could be calculated.

4. Prototype development

After the collection and training is successfully completed, we move on to assembly of the final product, which consists of both mechanical and virtual parts. The mechanical part includes a detection system mainly consisting a camera and a raspberry pi. Virtual side includes a processing software and a user interface (UI). By the combination of the two we get a working prototype with an input variable and an output data info.

5. Advancements and future works

Our work will be an open frame system which consistently learns and provides data accordingly, so the system can function in a wider range of species. So as to achieve this goal we shall design a website that acts as both the interface and future advancement platform, this will be a medium for communication with similar and related projects and a major source of data as well.

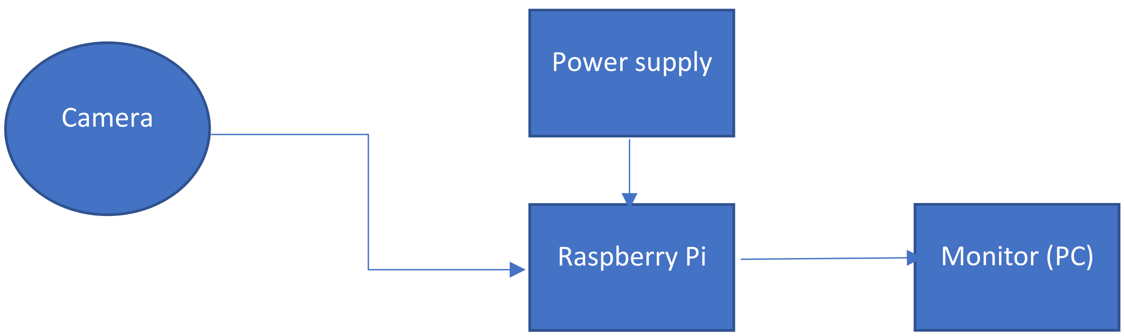


Fig:2 Architecture of the system

## Timeline

| Project Steps | Major Activities and Milestones | Deliverables | Start Date | Expected End Date |
|---|---|---|---|---|
| 1. Data collection and Curation | - Gathering disease affected plant images | Data collection | Right after submission of project proposal | 1 week before training |
| 2.Training and Testing | - Splitting the dataset into train set (80%) for fitting the model and test set (20%) for validation K-fold cross validation | Trained model | 1 week after data collection | 1 week before Prototype development |

| | | | | |
|---|---|---|---|---|
| | technique is implemented to find the accuracy. | | | |
| 3. Prototype development | - Prototype design<br>- Prototype construction<br>- Prototype testing | Physical working prototype | Right after successful training | 2 weeks before final report |
| 4. advancements and future work | - Website development<br>- UI advancements<br>- Communication medium | Final product | After prototype development | 1 week before final report |

## Budget

There are two alternatives budget plans depending upon the price of raspberry pi and materials in which prototype made up of.

Option 1

If the plastic material is designed and developed by us, the cost needed is as follows

| S/N | Material list | Individual price |
|---|---|---|
| 1 | Raspberry pi | 8500br |
| 2 | Raspberry camera | 1000br |
| 3 | Plastic | 1000br |
| | Total cost | 10,500br |

Option 2

If we used composite, the cost needed is detailed as follows:

| S/N | Material list | Individual price |
|---|---|---|
| 1 | Raspberry pi | 8000br |
| 2 | Raspberry camera | 1000br |
| 3 | Composite | 1000br |
| | Total cost | 10,000br |

## RESPONSIBILITIES AND WORK ALLOCATIONS

The project is subdivided in to many efforts, which in turn come together as one final team product.

**Data gathering and research:**

In this step all engineers are expected to contribute in data taking and data manipulation, this is done by referring to different articles on the internet as well as searching previous data points.

**Fetching data and training/testing:**

Using a sequence of methodologies and Python programs, this area falls in the hand of software engineers

**Prototype design and construction:**

Design of the prototype falls to the mechanical engineer, and in cooperation with the electrical and software engineers the construction will commence

**Treatment and prevention:**

This field requires extensive research and study on the specific disease and methods of treatment, this will be done with the collaborative work of the environmental, chemical and civil engineer.

## Project flow and contribution

1. **Mechanical engineer**

- Design of the prototype using 3D design software (Catia v5)
- Data collection
- Prototype construction

2. **Electrical engineer**
3. Algorithm development

- Circuit design
- Collecting data
- Prototype construction

**4. Software engineer**
- Algorithm development
- Website and user interface
- Data collection and organization

**5. Chemical engineer**
- Data collection
- Treatment analysis
- Future advancements

**6. Environmental engineer**
- Treatment analysis
- Data collection
- Future work analysis

**7. Civil engineer**
- Data collection
- Treatment analysis
- Prototype construction

## DOCUMENTATION AND REFERENCE


1. Savary, s., ficke, A, Aubertort,j-n.,et al crop losses sue to diseases and their implications for global food production losses.
2. S. C. Madiwalar and M. V. Wyawahare, "Plant disease identification: A comparative study," 2017 International Conference on Data Management, Analytics and Innovation
3. P. Moghadam, D. Ward, E. Goan, S. Jayawardena, P. Sikka and E. Hernandez, "Plant Disease Detection Using Hyperspectral Imaging," 2017 International Conference on Digital Image Computing: Techniques and Applications .
4. S. D.M., Akhilesh, S. A. Kumar, R. M.G. and P. C., "Image based Plant Disease Detection in Pomegranate Plant for Bacterial Blight.
5. International Conference on Communication and Signal Processing
6. G. Shrestha, Deepsikha, M. Das and N. Dey, "Plant Disease Detection Using CNN," 2020 IEEE Applied Signal Processing Conference .