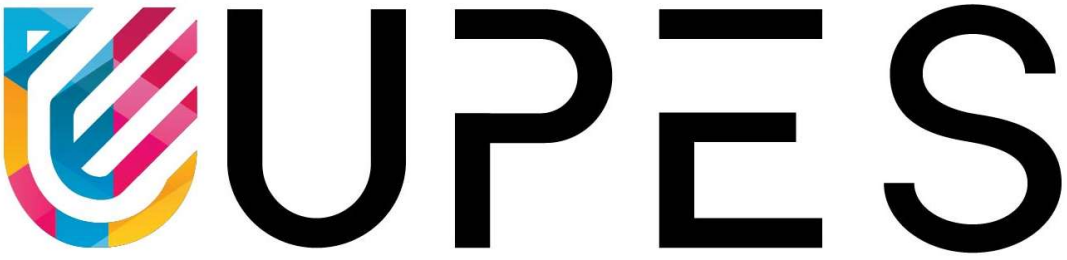
Software Requirements Specification

For
Plant Disease Recognition using Machine Learning

31st March 2023

Prepared by

Specialization	SAP ID	Name
BAO	500083441	Hiya Chopra
BAO	500083281	Shivam Gupta
AIML	500084169	Shivam Rathi
BAO	500083687	Shrey Gilotra



Department of Informatics
School Of Computer Science
UNIVERSITY OF PETROLEUM & ENERGY STUDIES,
DEHRADUN- 248007. Uttarakhand

Table of Contents

1. INTRODUCTION:

Plant diseases can seriously harm the crops around us, and farmers who use inaccurate methods to detect diseases reduce crop yield and incur greater financial losses. Therefore, early disease detection is essential for efficient disease management. The need for effective and precise methods for detecting plant diseases has become more pressing due to the rising demand for agricultural production and the threat of climate change.
In this study, we train machine learning models to recognize and detect patterns in plant images as well as to pinpoint their symptoms. Image processing is used to classify new images and correctly identify the disease present. It involves extracting features from the images, training machine learning models to recognize patterns, and using the learned patterns to classify new images.

1.1 PURPOSE:

The purpose of this project is to use Machine Learning for plant disease recognition to provide farmers and agriculture experts with an efficient and accurate tool to detect the disease associated at an early stage. Also, Machine Learning algorithms can analyze a large amount of data and identify patterns that are difficult for humans to detect, making it possible to automate the detection and diagnosis. The use of machine learning algorithm is when we need to generalize and detect plant diseases. The technology can also help to mitigate the shortage of agricultural experts in certain parts of the world, making it possible to benefit more farmers in the world.

1.2 TARGET BENEFICIARIES:

The target of detection of plant diseases is mainly farmers, agricultural experts, home farmers, and other people interested in plant cultivation and care. Farmers will benefit significantly from this technology as it provides them with an efficient and accurate way to detect plant diseases at an early stage. Agricultural professionals, including researchers and consultants, will also benefit from this technology, providing them with a tool to support their work in crop management and disease control. In addition to farmers and agricultural experts, other stakeholders involved in crop production and maintenance, such as decision-makers, input suppliers and processors, will also benefit from this technology, reducing the economic losses and environmental impacts caused by plant diseases.

1.3 PROJECT SCOPE:

The scope of the machine learning detection of plant diseases project can be divided into the following phases:

Data collection: The first step of the project is to collect an extensive database of images of plants with known diseases. This dataset must cover a wide range of plant species and diseases to ensure model accuracy and reliability.

Data pre-processing: Collected data must be pre-processed to remove noise or irrelevant information that may affect the accuracy of the model. This includes image enhancement, feature extraction, and data normalization. Also, various image processing techniques are used to develop image classification models but we are using neural network technique.

Model training and testing: The pre-processed data is then used to train a machine learning model using algorithms such as Convolutional Neural Networks (CNN). The trained model should be tested on a separate dataset of plant images with known diseases to ensure its accuracy and generalizability.

Implementation of the model: Once the model is tested and validated, it should be implemented in a web-based platform or mobile application for use by farmers. The platform must be user-friendly and easily accessible to farmers in remote areas.

1.4 REFERENCES:

[1] P.Chaitanya Reddy et al stated that improving the technical and mechanized support will increase agriculture productivity. They employed various Machine learning algorithms like SVM Classification and Random Forest for the detection of leaf diseases. Various Performance metrics like Root mean square Error (RMSE), Peak Signal Noise Ratio is compared to benefit the farmers with less time and resources.

[2] Ahmed et al presented the design and implementation of a machine learning-powered plant disease detector that enables farmers to diagnose the most common 38 diseases in 14 species. They trained a CNN Model using an imaginary dataset consisting of 96,206 photos. They designed a mobile application to increase the system's usability.

[3] Mrs. Shruthi U et al reviewed a comparative study on five types of machine learning classification techniques for the recognition of plant diseases. SVM Classifier is used by many researchers but the result showed that CNN Classifier detects more diseases with high Accuracy.

2. PROJECT DESCRIPTION:

2.1. Data/Data Structure-

We use a publicly available and quite famous, the Plant Village Dataset. The dataset was published by crowdAI during the "Plant Village Disease Classification Challenge".

The dataset consists of about 54,305 images of plant leaves collected under controlled environmental conditions. The plant images span the following 14 species:

Apple, Blueberry, Cherry, Corn, Grape, Orange, Peach, Bell Pepper, Potato, Raspberry, Soybean, Squash, Strawberry, and Tomato.

The dataset contains a total of 38 classes of plant diseases listed below:

Apple Scab

Apple Black Rot
Apple Cedar Rust
Apple healthy
Blueberry healthy
Cherry healthy
Cherry Powdery Mildew
Corn Gray Leaf Spot
Corn Common Rust
Corn healthy
Corn Northern Leaf Blight
Grape Black Rot
Grape Black Measles
Grape Leaf Blight
Grape healthy
Orange Huanglongbing
Peach Bacterial Spot
Peach healthy
Bell Pepper Bacterial Spot
Bell Pepper healthy
Potato Early Blight
Potato healthy
Potato Late Blight
Raspberry healthy
Soybean healthy
Squash Powdery Mildew
Strawberry Healthy
Strawberry Leaf Scorch
Tomato Bacterial Spot
Tomato Early Blight
Tomato Late Blight
Tomato Leaf Mold
Tomato Septoria Leaf Spot
Tomato Two Spotted Spider Mite
Tomato Target Spot
Tomato Mosaic Virus
Tomato Yellow Leaf Curl Virus
Tomato healthy

Note: The dataset also consists of an additional class background to differentiate between leaves and their background features.

The dataset file is stored in Google Drive (GDrive) and can be shared via a URL that provides a unique id. Using this unique id we download the dataset zip file from the GDrive and unzip it.

2.2. SWOT ANALYSIS-

Strengths -

- Accessibility: The technology can be implemented through smartphones, making it easily accessible to farmers in remote areas.
- Scalability: Once a machine learning model is trained, it can be easily scaled to analyze large amounts of data from multiple farms and regions.
- Efficiency: Detection of plant diseases based on machine learning is faster and more efficient than traditional manual diagnostic methods, reducing time and labor costs.
- Accuracy: Machine learning algorithms can accurately classify and diagnose plant diseases, resulting in better crop management and increased yields.

Weaknesses -

- Complexity: Developing a machine learning model for plant disease detection requires specialized knowledge and expertise in computer science and agricultural science.
- Data dependency: Machine learning algorithms require large amounts of data to train, and model accuracy depends on the quality and diversity of the data set.

Opportunities -

- Increased productivity: Accurate and timely disease diagnosis can increase yields and improve food safety.
- Sustainability: Technology can reduce pesticide use by enabling targeted and timely use, improving environmental sustainability.
- disease control, reduce disease spread and minimize crop losses.

Threats-

- Lack of awareness: Many farmers may not be aware of the technology and its benefits, which may limit its adoption and impact.
- Cost: The cost of developing and deploying machine learning-based plant disease detection systems can be a barrier, especially for small farmers.

2.3. Project Features -

The project mainly focuses on features like Speed, Accessibility, Versatility, User-friendly, and Cost-effectiveness. Other than this, we have applied the approach as follows-

We have created a sequential model for the classification of the task. For, the model we first create a 2D Convolutional layer with 32 filters of 3*3 kernel and a RELU (Rectified Linear Unit) activation. Then performed in batch normalization, Max Pooling, and 25% dropouts (25 epochs) operations in the following layers.

2.4. Design and Implementation Constraints-

To run the model, we need to specify all of these steps, the processing techniques and working model are represented in the data flow diagram as shown below in Fig. 1.

Step 1: Collecting the dataset

The first step in our project is to collect image data of various plant leaves. We are using Plant Village Data which is available on Kaggle

Step 2: Preprocessing and Augmentation:

Pre-Processing and Augmentation are done on the collected dataset by using Keras. We will improve the quality and quantity of images using Data Augmentation by rotating and changing the brightness of the image to help the model train faster.

Step 3: Building CNN

There are three components of CNN:

A. Conv Layer
B. Pooling Layer
C. Fully Connected Layer

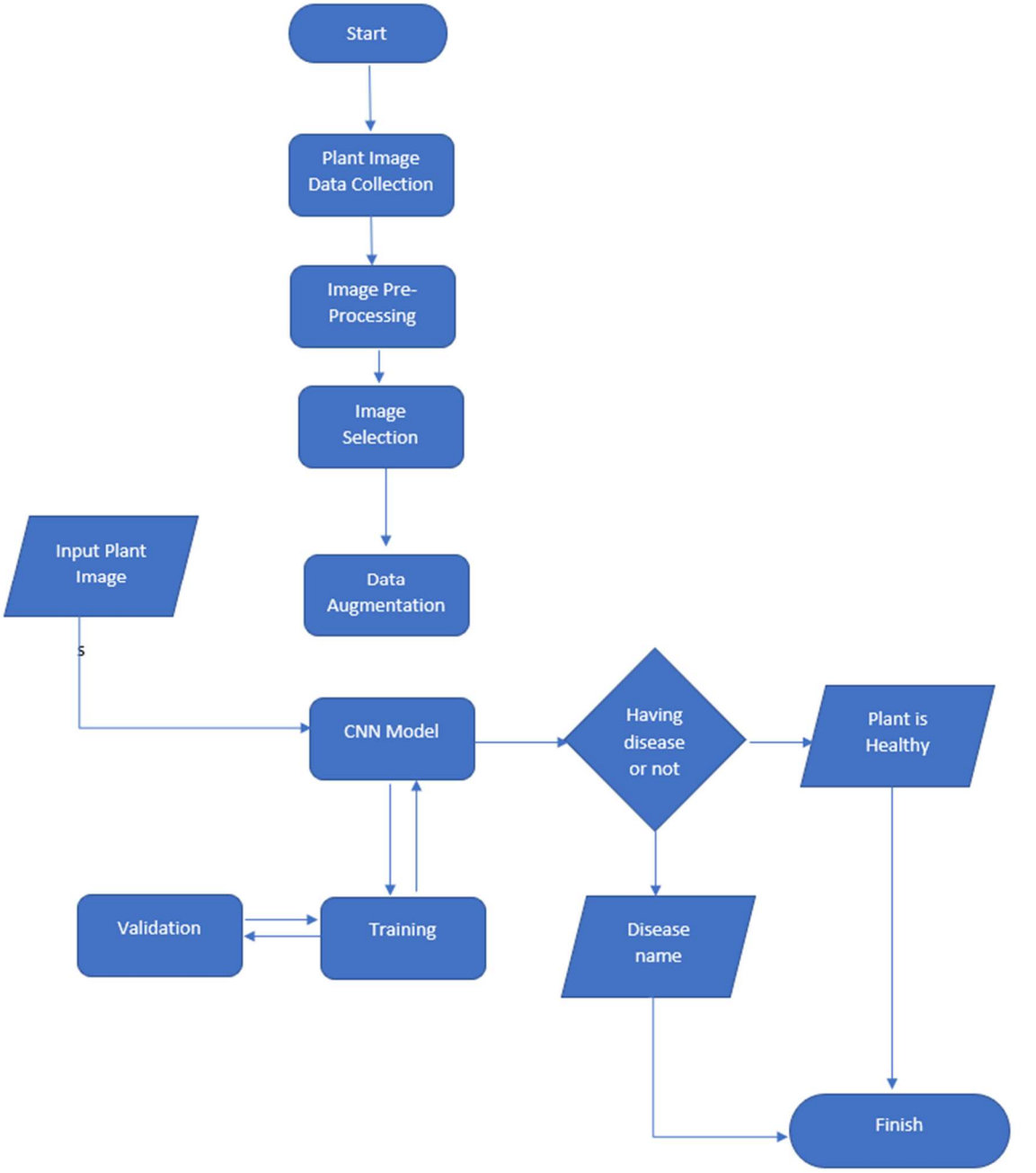


Fig. 1. Plant disease detection system.

3. SYSTEM REQUIREMENTS:

3.1. Hardware Requirement-

1. Minimum 8 GB RAM
2. Minimum 500 GB internal storage
3. Minimum Intel core i5

3.2. Software Requirement-

* System with Python installed
1. IDE to run the program.(developer)
2. Andriod mobile phone(User)

4. User Interface

The user interface of our project is to create an android platform where a user gives input as an image of the plant leaf. And the model classifies it whether it is infected or not.

4.1. Software Interface

Implementation is in Python programming language that will execute in any system which will require SDE's compilers like Python.exe and an IDE for executing the code. And deployed on google cloud for the android platform.

5. NON-FUNCTIONAL REQUIREMENTS

5.1. Performance Requirements

We are trying to develop an efficient model with nearly 100% accuracy. The goal of performance model is to be more robust and not be overfitted due to which the problem of biasness could be solved.

5.2. Security Requirements

The dataset is stored in the Google Cloud Platform database management system. An android Mobile is required for accessing the model through an Android application.

5.3. Software Quality Attributes

5.3.1. Usability: The usability of our model is very high allowing the end users to use our model easily.
5.3.2. Reliability: The reliability of our model is quite high as it is able to predict grades on basis of different attributes provided.
5.3.3. Availability: It is easily available for every end user for better analysis of their health.
5.3.4. Robustness: Our model is highly robust and effective as the code used for its development is really optimized.

5.3.5. Adaptability: The model developed is highly adaptive in nature. This means that the model can adapt to any changes if required in the future. For instance, if any new attribute is required for classification, then we can easily train our model and it can take that into consideration the next time we use the model.

5.3.6. Portability: Our whole project is developed completely using Python. The requirements of our model are present in a standard system running python.exe which allows us to group together the libraries and our personal settings into one unit. Python makes it easier to write portable and universal programs that can run on a variety of operating systems like Windows, Mac, Linux etc.