Wollo University

Kombolcha Institute of Technology (KIoT)

College of Informatics Department of Software Engineering

Title: Tomato plant disease detection and prediction system using Machine learning for Ethiopian Agricultural Sector

A project Submitted to Department of Software Engineering in Partial Fulfillment of the Requirement for the Degree of Bachelor in Software Engineering

Prepared By

No.	Name	Id
1.	Hiluf Meressa	WOUR/1463/10
2.	Getnet Begashaw	WOUR/1315/10
3.	Bewket Dereje	WOUR/0965/10

Adviser Name : Tilhahun Ayalew (MSc)

Date of submission : 13/11/2014 EC

Kombolcha, Ethiopia

Declaration

We hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. We also declare that, as required by these rules and conduct, we have fully cited and referenced all material and results that are not original to this work.

Group Members :

Student Name : Hiluf Meressa Id : 1463/10 Signature: ______Date: ______

Student Name : Getnet Begashaw Id : 1315/10 Signature: ______Date: ______

Student Name : Bewket Dereje Id : 0965/10 Signature: ______Date: ______

This BSc. project has been submitted for examination with our approval as a Project adviser.

Adviser Name : Tilahun Ayalew Signature: ________Date:______

Approval of the board of examiner

Adviser : Tilahun Ayalew (MSc) Signature : ___________ Date : ___________

Chairperson : _______________________Signature : ___________ Date : ___________

Examiner 1 : ________________________Signature : ___________ Date : ___________

Examiner 2 : ________________________Signature : ___________ Date : ___________

Examiner 3 : ________________________Signature : ___________ Date : ___________

Examiner 4 : ________________________Signature : ___________ Date : ___________

Acknowledgment

First of all, Above all we would like to thank God for giving us courage, unconditional support not only in the good situations but also in the bad situations of our life, and strength to complete this work.

Second we would like to express our deepest gratitude to our adviser Tilahun Ayalew (MSc) for his supervision, persistent advice, inspiration, patience and continued guidance right from the moment of becoming our adviser, topic selection, statement of problem formulation to the completion of the work. We also want to thank him for his comprehensive lectures on Machine Learning which helped us solve the problem of the project.

Finally we would like to sincerely thank all of our friends, colleagues, classmates for their assistance, encouragement, and inspiration in this project work.

Keywords: Convolution Neural Network(CNN);Machine learning;Tomato plant.

Executive summery

The quality and quantity of the crop are significantly affected by numerous diseases in plants. In this regard, an early detection of such diseases is highly effective. Tomato is one of the important crops that is produced in large quantities with high commercial value. Several types of tomato

diseases affect the tomato crop at an alarming rate. Several methodology are carried out in this project to be develop. such as observation,Open dataset and custom datasets and we use different developments tools to develop system, such as TensorFlow,Keras,Opencv,NumPy and Pandas. we will deploy Convolution Neural Network (CNN) based models for tomato plant disease detection and prediction. The proposed system aims to find the best solution to the problem of tomato leaf disease detection using a machine learning approach.

Table of Contents

List Tables

Table 1 Parameters used in the CNN model......9

Table 2 Project Budget break-down and cost analysis......10

Table 3 Technology requirements......10

Table 4 Schedule of the project......11

Table 5 Team composition of the project......11

Table 6 Document convention......12

List of Figures

Figure 1 structure of conventional neural network..8
Figure 2 Proposed System...14
Figure 3 Essential use case Diagram..18
Figure 4 use case diagram...20
Figure 5 Sequence diagram...23
Figure 6 Activity diagram..24
Figure 7 Images of different type of tomato plant diseases...25
Figure 8 Proposed system architecture...29
Figure 9 subsystem decomposition..30
Figure 10 Hardware and software mapping...31
Figure 11 Component diagram..32
Figure 12 User interface for Default page of the system...33
Figure 13 user interface for uploading image...34
Figure 14 User interface for prediction feedback and percentage..35

List of Abbreviation

ML : Machine learning

LAN : Local area network

CNN : Convolutional neural network

CUDA : Compute Unified Device Architecture

GPU : Graphical Processing Unit

JSON : Java Script Object Notation

CSV : Comma Separated Value

RAM : Random Access Memory

UML : Unified Model Language

ONNX : Open Neural Network Exchange

GDP : Gross Domestic Product

BSc : Bachelor of Sciences

Opencv : Open Source Computer Vision Library

KG/Hac : Kilogram per Hectare

FTC : Farmer Training Center

CSS : Cascading style

HTML : Hypertext Markup Language

Chapter one

Introduction

Tomato plant diseases have always been a thorny problem in agricultural production and one of the main factors restricting the sustainable development of agriculture. As a common vegetable and important cash crop in Ethiopia, tomato is widely cultivated in various regions. According to the current statistical data, there are 9 types of tomato plant diseases, which have seriously affected the yield and quality of tomatoes and caused huge economic losses. In the past, disease diagnosis mainly uses artificial recognition methods. Includes, subjectively judge disease types based on years of planting experience of farmers or consult books on agricultural knowledge, Obtain disease specimen pictures and search on the Internet for judgment, Consult experts to undertake an analysis of the disease symptoms. In general, farmers are not highly educated and do not have the necessary professional knowledge. They usually have a high misjudgment rate of tomato plant diseases. The rest of the paper is organized as follows: Chapter one presents the introduction part, chapter two presents description of existing system, chapter three presents proposed system, chapter four presents system methodology and design and the last chapter presents conclusion and recommendation. We strongly recommend the user of the document to give a detail look for every entities we use. If there is any misrepresentations of facts, figures and system related concepts, we would like to give our greatest apology on behalf of the preparation team.

1.1 Background of the organization

Ethiopian agriculture sector is the oldest profession. Humans started cultivation even science there was no civilization. With the development of science, plants were identified as living-thing. It could also respire, reproduce, and even get prone to various diseases. These are different types of diseases by various microorganisms may it be bacteria, viruses, or fungi. Plant diseases can damage crops to a great extent. It can even be fatal to human beings. One such situation emerged in 1840 when a large amount of tomato crop was destroyed due to a disease called the Late blight of tomato. We all know that plants are very important in our lives so we need to protect them not only from deforestation but also from various plant diseases. Still, two-third of the population relies upon agriculture directly or indirectly. Majority of Ethiopia people depend on agriculture for their livelihoods and has contributed about 32.6% of nation's GDP alone by the agricultural sector in the year 2012/13[2]. In the year 2012/13, the average economic growth was confined to 0.77%. Tomato is one of the major cash crops cultivated in Ethiopia. Throughout the year tomatos are cultivated in 25.49 hectors of area and harvesting rate is about 13419 KG/Hac. In disease recognition in such plants , the current system relies on visual observation which is a time consuming process. In Ethiopia large population is engaged in agriculture. It ranks 3rd in Africa in terms of agriculture production. Tomato crops are damaged in Ethiopia because of various factors one of the major causes is natural calamities and other is the microbial diseases. As humans do not have any control over natural disasters but we can control microbial infections in tomato plants. Once a crop is infected by some disease it is difficult for farmers to find out the real cause of the disease. pathogens and pests are affecting the crops badly. In the Ethiopia context where agriculture contributes to 16% of GDP and engages almost 60% of the population requires great measures to be taken to avoid tomato plant diseases. According to the Ministry of Food Processing Industries in the year 2009 agricultural loss was 13 billion dollars. One of the helpful measures in tomato plant disease detection and prediction methods can be done with the help of image processing and neural network.

1.2 Motivation

Tomato plant diseases possess a very devastating threat to the agriculture industry and have the potential of pushing the whole of human society into starvation if not detected early. With the implementation of machine learning models in the domain of plant pathology, the detection of tomato plant diseases will become easier and cheaper helping many farmers in the timely detection of tomato plant diseases, preventing wastage of tomato plants and protecting the transmission of diseases from diseased to healthy tomato plants. The motivation to develop this system is :-

- ❖ Tomato production loss due to diseases is estimated to be approximately 30-50%.
- ❖ Tomato plant disease makes the farmers less productive.
- ❖ The importance of machine learning in farmland of tomato plant.
- ❖ Highly transmission of tomato plant disease from one tomato plant to the other.

1.3 Statement of the problem

For increasing the tomato crop production and productivity farmers contact experts to ask for their advice concerning the action towards the occurrence of the diseases to their crops and recommendation for control. Some times they have to go long distances to approach the experts, even though they go such long distances the expert whom the farmer wants to contact may not be in a position for advising the farmer. In these cases, looking for expert advice is very expensive and time consuming. In addition, tomato plant disease or deficiency detection and classification are usually carried out by farmers, agricultural extension workers, and pathologies by continues monitoring of tomato plants. In small scale farm, early detection and prediction of tomato plant disease are very much simple and the disease could be easily controlled by organic pesticides or by the use of a minimum amount of chemical pesticides. In large scale farm, continuous monitoring and early detection and classification of disease are very difficult and it results in a severe outbreak of the disease and pest growth could not be controlled by organic means. In this situation, farmers are urged to use chemicals pesticides so as to control the disease and to retain the crop yield. This problem is happened due to lack early detection and prediction of tomato plant disease and consumes large amount of man power. In existing system the early detection and prediction of tomato plant disease is carried out by agricultural extension workers was based on different traditional methods. Such as: judge tomato plant disease types based on years of planting experience of agricultural extension workers and Obtain disease specimen pictures and search on the Internet for judgment. This is leads to misjudgment of tomato plant disease detection and prediction.

1.4 Objective

1.4.1 General Objectives

The general objective of the project is to develop a model that can automatically detect and predict tomato plant diseases based on machine learning and to evaluate its performance

1.4.2 Specific Objectives

- ❖ Explore the different types of tomato plant diseases.
- ❖ To verify the reliability and accuracy of the developed system.
- ❖ Analyze and compare results with different machine learning methods
- ❖ Draw conclusions and propose a set of future improvements.

1.5 Scope and limitation of the project

1.5.1 Scope

- ❖ Gathering tomato plant leaf datasets.
- ❖ Working pre processing (image labeling and enhancement)
- ❖ Creating of trained model.
- ❖ Testing for 9 types of tomato plant disease (Detect and predict with a certain level of accuracy.
- ❖ Develop user friendly web interface.

1.5.2 Limitation

The intended system will have the following limitations.

- ❖ Limits to only one language i,e English.
- ❖ Not support mobile applications.

1.6 Assumptions made in the project

Our system needs high processing power. To address these problems using one of free cloud service is the solution i,e co-laboratory. The alternative solution for co-laboratory is using high performance computing device with higher processing power i,e supercomputer.

1.7 Feasibility Study

To tell if it is feasible or not to automate tomato plant diseases prediction and detection system it should check (seen) from different factors.

1.7.1 Technical Feasibility

The proposed system can be easily maintained and repaired; technically. The system will be powerful to be applied by low skilled users as much as possible. There is no need for the expertise developer involvement in almost all implementation of the entire system.

1.7.2 Operational Feasibility

The proposed system will provide best services for user. The system will be operationally feasible and it will be operationally acceptable to users. The system give better user interface.

1.7.3 Economic Feasibility

The purpose of economic feasibility assessment is to determine the positive economic benefit to the organization. Some of economic benefits are reduction of costs for pen, paper, printer, transportation, Reduction of human power, Reduce wastage of time, Reduction in loss of file and increase efficiency etc.

1.8 Significance of the project

- ❖ Improve early detection and prediction of tomato plant disease at farmlands.
- ❖ This study will help pathologist understand the importance of artificial intelligence in general and machine learning in particular area.
- ❖ It will help to achieve high quality and quantity production because it lets detail disease management techniques incurring the disease appeared on the tomato plant for pathologist, users and agricultural extension workers.
- ❖ The study will help farmers reduce the cost of production that brings huge loss due to excessive use of fungicides on their plants.
- ❖ This study would serve as a reference resource for other researchers who would seek to conduct further studies into the problems related to plant disease detection and classification

1.9 Beneficiary of the project

For Clients (Farmer)

- ❖ generate highly cultivated plants.
- ❖ Gets high products.

For developer (for you)

- ❖ Develop his/her carrier.
- ❖ To develop academic skills Generally on Artificial Intelligence.

For System users(Service center)

- ❖ Get Satisfaction.
- ❖ Early detection and prediction of tomato plant disease.
- ❖ Minimize the human power.
- ❖ minimize detection and prediction time.

1.10 Methodology

1.10.1 Open datasets

We have 15,012 images of diseased and healthy tomato plants of the selected tomato plants were downloaded from the plant village website[1] to apply the proposed methodology and conventional machine learning models. The dataset contains thousands of images of healthy and diseased crop plants that are open and available on the web. The data-set bacterial spot contains 2127 images, late blight contains 1909 images, leaf mold contains 952 images, septoria leaf spot contains 1771 images contains ,two spotted spider mite contains 1676 images, target spot contains 1404 images, tomato yellow leaf curl virus contains 3209 images, tomato mosaic virus contains 373 images and healthy contains 1591. Furthermore, the datasets have been partitioned in training and testing. The 80% data, 12,009 datasets have been used for training, and 20%, which consist of 3003 have been used for testing.

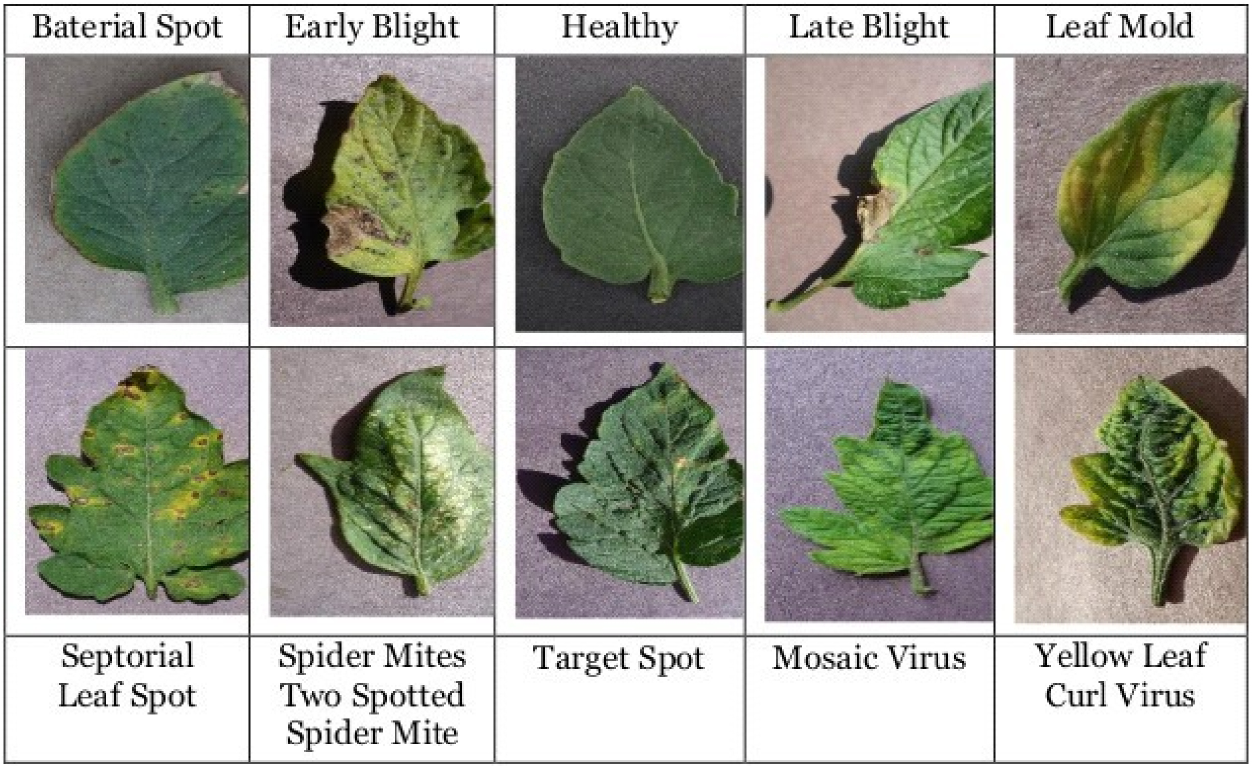

Figure 7 Images of different type of tomato plant diseases

1.10.2 Observation

We take observation in farmland to take measure how the service center can detect and predict the tomato plant disease and what could happen when they take this action. During the observation, we try take look what gap could be made by service center. In addition, take the consideration how the tomato plant diseases detection and prediction project solve the problem.

1.10.3 Image pre processing techniques

The pre-processing step for any machine learning model is of great importance and ideally shapes the performance and results of the models chosen. In this project, the following were the steps that were carried out in order to make sure that the models produced optimal results.

- Read images
- Resize images
- Convert the images into an array
- Split images into two different sets, namely, train and test set.

1.10.4 Development tools

- TensorFlow

It helps in training and building of our model.

- Keras

It provides essential abstractions and building blocks for developing and shipping machine learning solutions with high iteration velocity.

- NumPy

It is used for performing mathematical and logical operations on Arrays.

- Pandas

It is used for data cleaning and analysis.

- Opencv

It is used to solve computer vision problems such as reading or re-sizing an image

- Matplotlib

It is used for creating static and visualization.

1.10.5 Model Training Techniques

The process of training an ML model involves providing an ML algorithm (I,e the learning algorithm) with training data to learn from. The term ML model refers to the model artifact that is created by the training process. The need for huge amount of data is one of the drawbacks of machine learning models, to tackle this problem we use a technique called fine-tuning. In this tomato plant leaf disease detection and prediction system using Machine Learning for Ethiopian Agricultural sector method we use pertained model weight and bias of a model must be adjusted very precisely in order to fit with certain observations. Also, the need for high processing power is another drawback back of machine learning. To address these problems using one of free cloud service is the solution. Co-laboratory, or "Co-lab" for short, is a product from Google Research. Co-lab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education.

1.10.6 Model Deployment Techniques

Serving is exposing the internal function to the end user. Serving the model is very different from model training, the skill, the technology tools are different. In addition, with the model need to deployed in real-time, that makes it little complex and computationally intensive. Task queuing and parallelization is a solution for this kind of real-time system. Celery helps to run code asynchronously or on a periodic schedule which are very common things you'd want to do in most web projects.

1.10.7 Algorithm

Training using Convolutional neural network (CNN)[3]. CNN is deep learning technique that plays a vital role in image processing applications such as object identification, recognition, detection, and classification. machine Learning techniques are capable of learning feature representations from the data. A CNN model takes images as the input, applies convolution operations and extracts the image features, resulting in dimensional reduction of the input. The accuracy of image processing is directly influenced by those extracted features. CNN model consists of layers such as Convolution, ReLU Pooling, Fully Connected, Flatten, and Normalization. Using CNN, the images would be compared piece by piece. Each piece is known as a feature or filter. From the input image, CNN uses the weight matrix and extract the specific features without misplacing the information about its spatial arrangement.

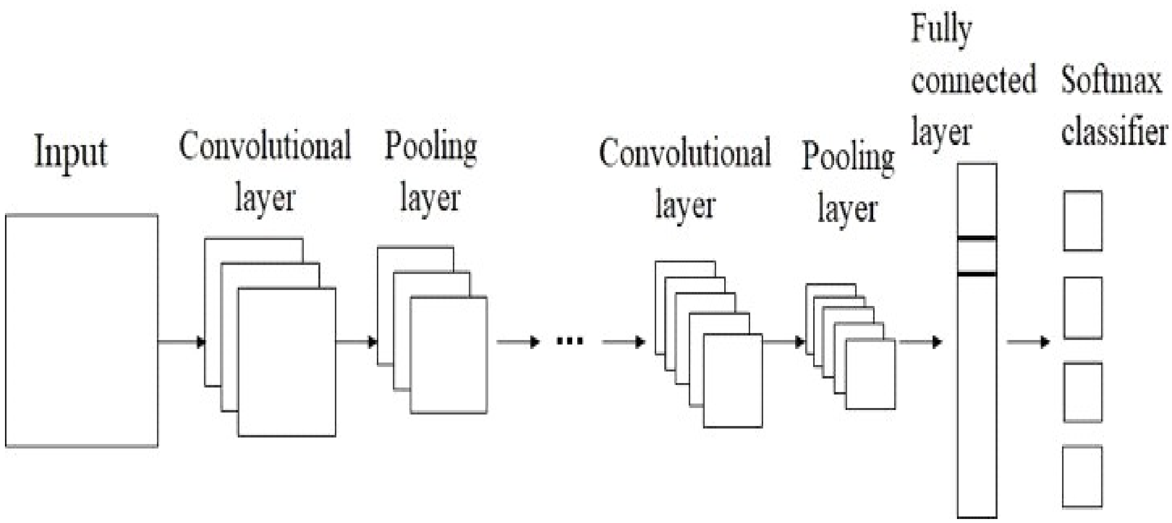

Figure 1 structure of conventional neural network

The reason to choose CNN model is uses partially connected layers, size of connection depends on the size of filter,used for any images (for any amount of data) , considers only the part of the image which is in the receptive filed of the filter,number of parameters are very less compared to the others,learn hierarchy of patterns I,e higher level of CNN are formed by combining lower layers this helps in identifying the patterns more efficiently than others.

1.11 Project Budget break-down and cost analysis

Table 1 Project Budget break-down and cost analysis

No.	Type	Quantity	Unit(ETB)	Total price(ETB)
1	Paper and pen	100 single paper and 3 pens	1 birr for paper 10 birr for pen	130
2	Hard disk	500GB	2500	2500
3	Flash disk	1	300	300
6	Laptop	2	50,000	100,000
7	Transport	4*(3 members)	20	480
8	Total cost			103,410 ETB

1.12 Technology requirements

Table 2 Technology requirements

Software Tools	Activities	Why?
co-lab notebook,Anaconda	For writing and editing the source code.	Very modern and easy to user.
LibreOffice writer, LibreOffice presentation.	For writing the document and presentations	Easy to user and it is free
Edrawmax	For editing and drawing the designs.	It has free trial version
Python, HTML, CSS, JavaScript	Programming languages	Easy to learn and use.
Flask, bootstrap	Web frameworks	Easy to learn and use
TensorFlow,Open-CV, Keras,NumPy,pandas	Machine learning tools	Used for machine model training and deployment
Chrome, Firefox	Web browsers	Free easy to use

1.13 Schedule of the project.

Table 3 Schedule of the project

No_	Task	Description	Start date	Finishing date	Duration	Adviser Feed back
1	Requirement gathering	Gathering requirement in Kombolcha farmlands	may 15	may 20	One weeks	Good
2	Requirement analysis	Identifying functional requirement	may 21	June 26	One weeks	Good
3	Design	System design	June 27	July 20	Three	Good

					weeks	
4	Coding	Full implementation				

1.14 Team composition of the project.

Table 4 Team composition of the project.

Student name	Id	Responsibility
Hiluf Meressa	1463/10	In all activity
Bewket Dereje	0965/10	In all activity
Getnet Begashaw	1315/10	In all activity

1.15 Intended Audience and Suggested Readings

The document is intended for requirements engineer,domain expert, developer and project manager. Before reading this document it is highly recommended to read the abstract of the document to get an overview of the project.

1.16 General Constraints

A constraint that slows a system down or prevents it from achieving its goal. Some of software constraints in our project are :

- Shortage of RJ45 cable for internet connection.
- Failure of hard disk
- The processor speed of the machine is low.

Chapter Two
Existing System

2.1 Description of existing system

Currently, the system being used at the counter time is an internal system which is manually used in prediction and detection of tomato plant disease.

2.2 Major function of existing system.

The main function of existing system is detect and predict of tomato plant disease based on some criteria. Such as: judge tomato plant disease types based on years of planting experience of service center and Obtain disease specimen pictures and search on the internet for judgment.

2.3 Users of current System

The users of current system are farmers and FTC that concentrated their works on tomato crop production agricultural area.

2.4 Drawback of current System

The drawback of current System is facing that the service centers losses a lot of time to detect and identify the tomato plant disease, difficult to identify the exact disease of the tomato plant disease and also the system works manually due to these factors productivity of tomato plant is leads to decrease.

2.5 Business Model Rule

Name of the rule : Service center must upload correct tomato plant image format.

Goal of the rule :

The goal of the above business rule is to get the correct tomato plant disease feedback and result.

Example :-

If service center select Leaf spots turn from light green to dark brown, rounded shape, the system displays the result of the tomato plant disease. I,e tomato bacterial spot.

Source

The origin of the business rule is team.

Chapter Three
Proposed System

Overview

The proposed system is to detect and predict the diseases of tomato plant by using feature extraction methods where features such as color, shape and texture are taken into consideration. Convolutional neural network (CNN), a machine learning technique is used in classifying the tomato plant leaves into healthy or diseased and if it is diseased , CNN will give the name of that particular disease. First the images of various leaves are acquired using high resolution camera so as to get the better results & efficiency. Then image processing techniques are applied to these images to extract useful features which will be required for further analysis. The intended system is capable of checking whether the tomato plant is affected or not and also it is capable analyzing and reporting the result. In addition, the system uses very cutting edge technologies which research and knowledge transfer value are significance.

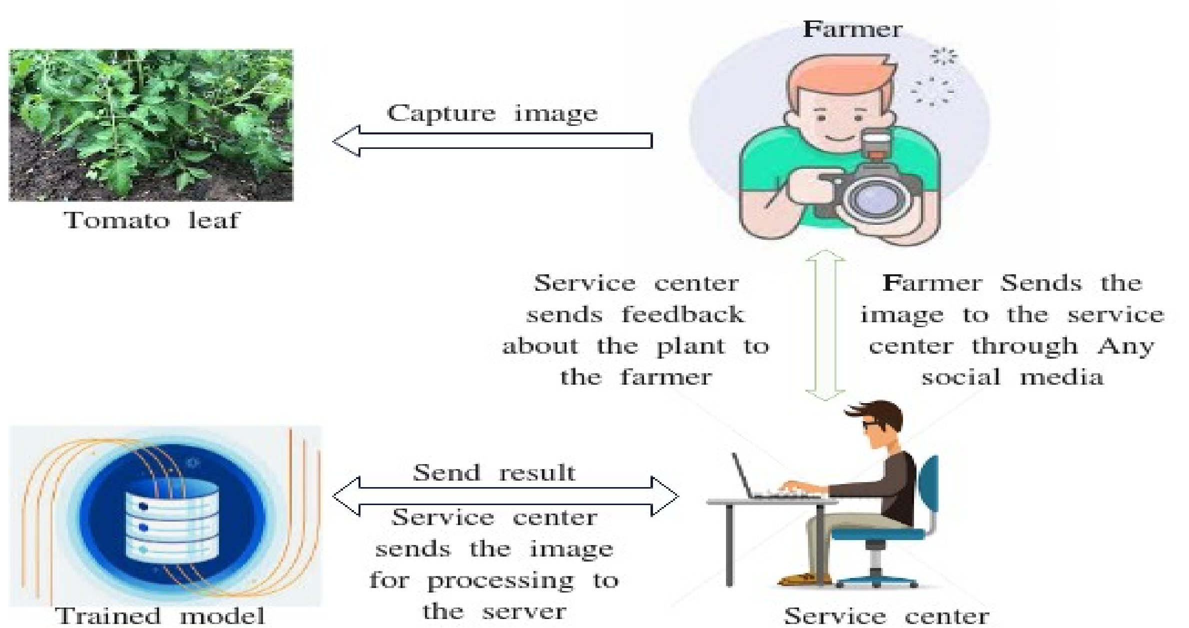

Figure 2 Proposed System

3.1 Functional Requirement

Table 7 Functional Requirement

FRID	Requirement Description	Use case	Priority
FRID-001	The system should be enhance the quality of tomato plant image.	Image Enhancement	High
FRID-002	The system should be allow to the service center to upload tomato plant image to be detect and predict.	Upload image	High
FRID-003	The system should be process the uploaded tomato plant image.	Process image	High
FRID-004	The system should be extract feature of the input tomato plant image.	Feature extract	High
FRID-005	The system should be detected and predicted tomato plant image disease.	detect and predict	High
FRID-006	The system should be allow to the service center to view the result about the tomato plant image.	View result	High

3.2 Non-functional Requirement

Table 8 Non Functional Requirement

NFRID	Requirement Description	Requirement Group
NFRID-001	The load time for the user interface screens and detection and prediction of the disease shall take no longer than two seconds.	Performance
NFRID-002	The system is easily maintainable,since components are modulated.	Maintainability
NFRID-003	The system should be user friendly due to users are use the system with graphical interface.	Usability
NFRID-004	The system works with out failure.	Reliability

3.3 External Interface requirements

3.3.1 User Interfaces

Since tomato plant disease detection and prediction system is web-based application it directly interacts with service center. In this system case we can separate user interface in to two, the server user interface and the client side user interface. In the server user interface the interaction (user to system) is based on command line user interface. The server admin start the server using predefined configuration file or optional parameters which pass with server run command. The other user interface is client side user interface (graphical user interface or web interface). The web user interface needs to be very modern, responsive and attract to look to work on.

3.3.2 Hardware Interfaces

The neural network will run on locally deployed GPU server. The external hardware beside the client web browser is camera, the camera connect to the server using the same local network with port and IP address.

3.3.3 Software Interfaces

Ubuntu 21.04 is the operating system for the system server. The following are specific to deep learning libraries dependencies.

- ❖ NVIDIA® GPU drivers CUDA 11.0 requires 450.x or higher.
- ❖ CUDA® Toolkit TensorFlow support CUDA 11 (TensorFlow >= 2.40)
- ❖ CUPTI ships with the CUDA Toolkit.
- ❖ cuDNN SDK 8.0.4 cuDNN versions

3.4 System Model

3.4.1 Essential use case model

Essential use case diagram helps to show the general overview of the system. It is used to identify the functions of the system and anything or anyone that interacts with the system

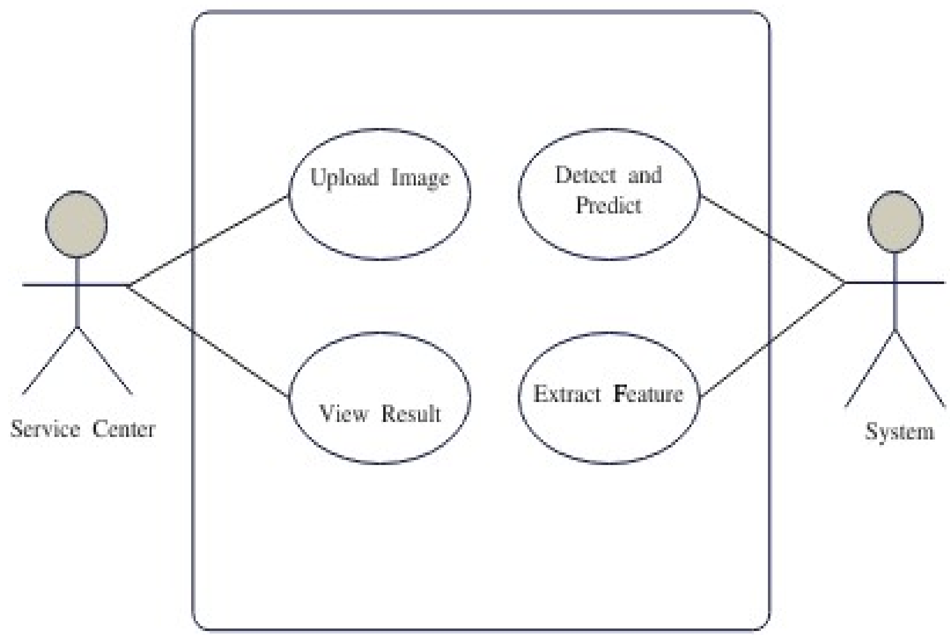

Figure 3 Essential use case Diagram

3.4.2 Use case Diagram

Actors in use case diagram represents users that interactive a system. They can be a person, an organization or an outside system that interact with our application. They must be external objects that produce or consume data. The actors our system are System and Service center.

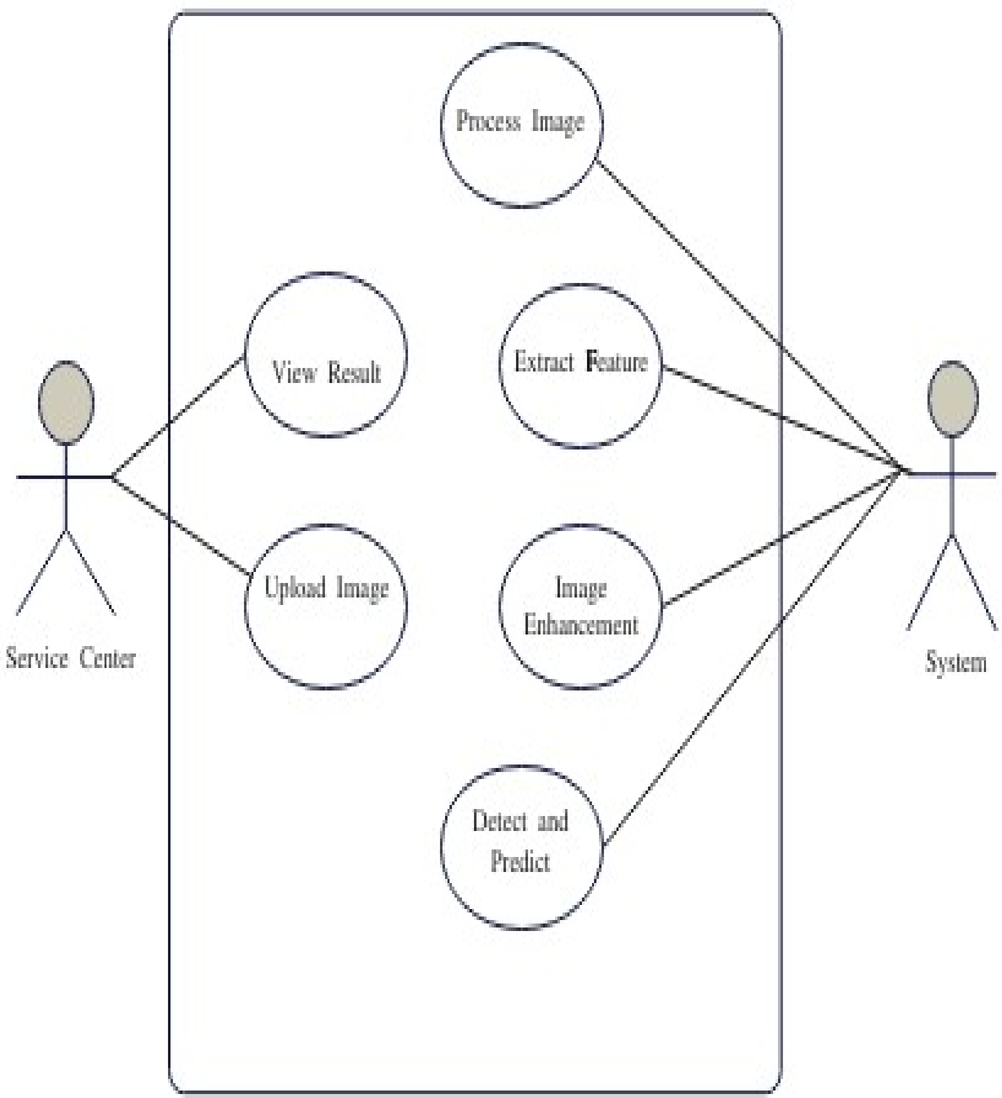

Figure 4 use case diagram

3.4.3 Use case Description

3.4.3.1 Use case description for Upload image

Table 9 use case description for upload image

Use case ID	0001
Use case name	Upload image

Date created	22/10/2014
Actor	Service center
Description	The service center wants to upload an image from the file server to be detected and predicted by the system.
Triggers	The use case is being initiated when the service center wants to upload image.
Preconditions	Images must exist in the file server.
Normal flow	Service center click upload image button. If the image format is correct the image is uploaded to the image panel.
Post conditions	Image is successfully uploaded the page.
Alternative flows	If service center upload wrong file (image format), the system reminds the service center to upload correct image format.
Exceptions	If the service center upload wrong file (image format) , the system displays an error message to the service center.

3.4.3.2 Use case description for Detect and predict

Table 10 use case description for detect and predict

Use case ID	0002
Use case name	Detect and predict
Date created	22/10/2014
Actor	System
Description	The system wants to to detect and predict tomato plant disease.
Triggers	The use case is being initiated when the service center clicks the detect and predict button.
Preconditions	Image must be uploaded by the service center before.

Normal flow	❖ Get tomato plant image from the service center. ❖ Pre process the tomato plant image. ❖ Extract features for the input tomato plant image. ❖ Compare the the extracted feature of tomato plant image with the classifier model. ❖ Finally show the result to the service center on the user interface
Post conditions	System shows the disease type and percentage for the service center.
Alternative flows	If the file format uploaded is incorrect that is before uploaded by the service center, the system reminds to the service center to upload correct image format.
Exceptions	If the service center upload wrong file (image format) : The system displays an error message to the service center.

3.5 Sequence Diagram

3.5.1 Sequence Diagram for image processing

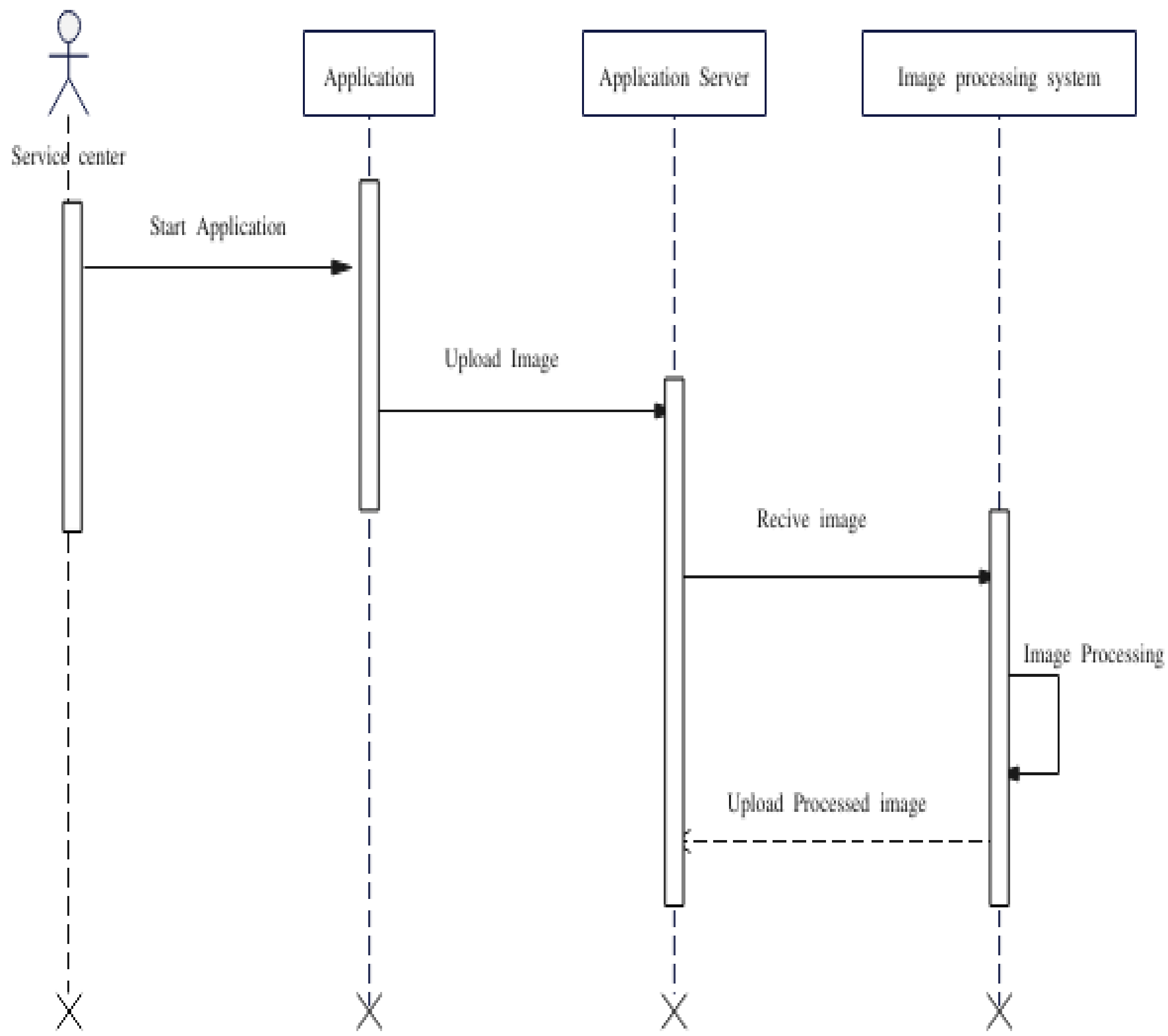

Figure 5 Sequence diagram for image processing

3.5.2 Sequence Diagram for detection and prediction

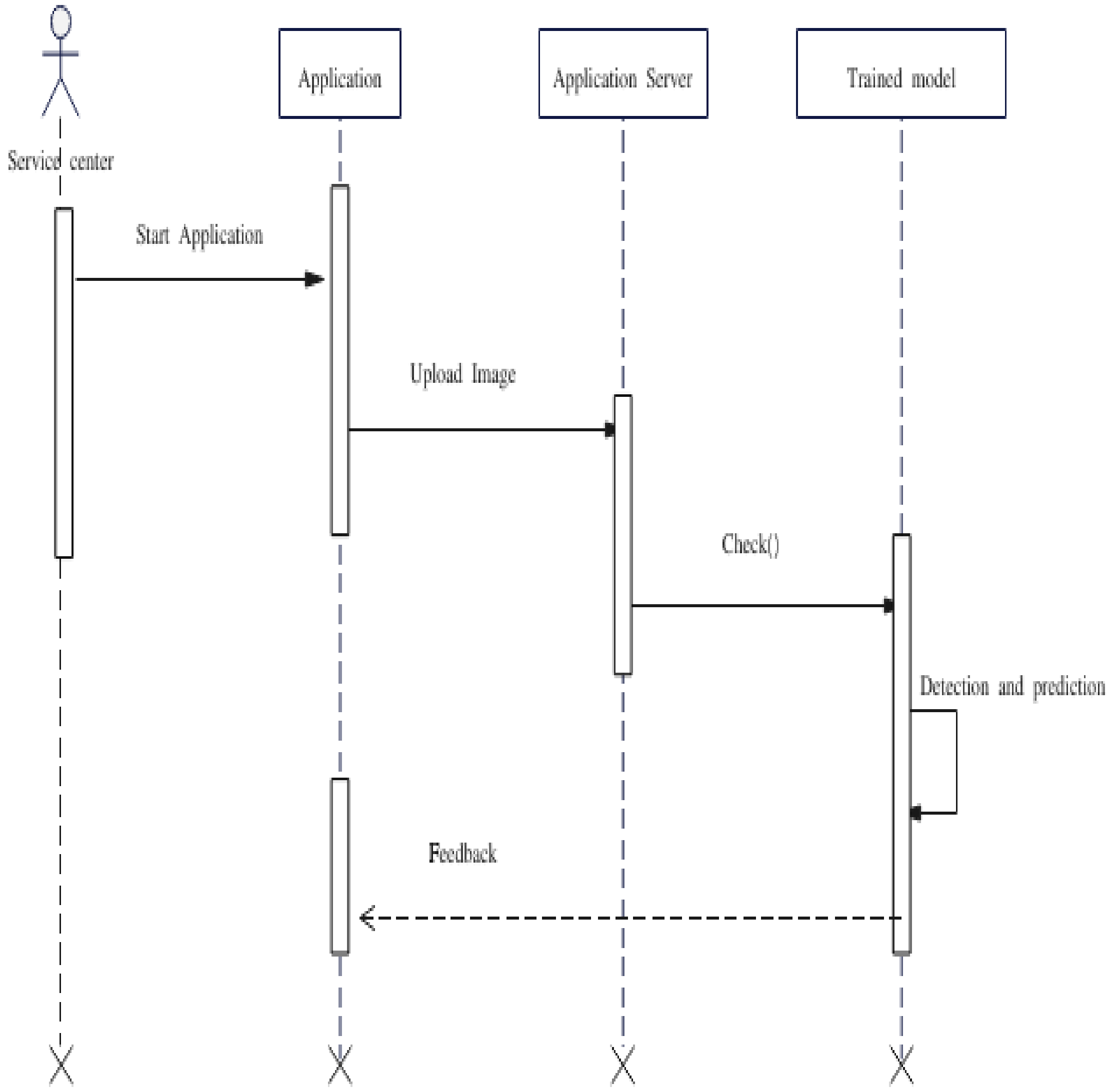

Figure for detection and prediction

3.6 Activity Diagram

Activity diagram is another important diagram in UML to describe dynamic aspects of the system. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as in operation of the system. The control flow is drawn from one operation to another

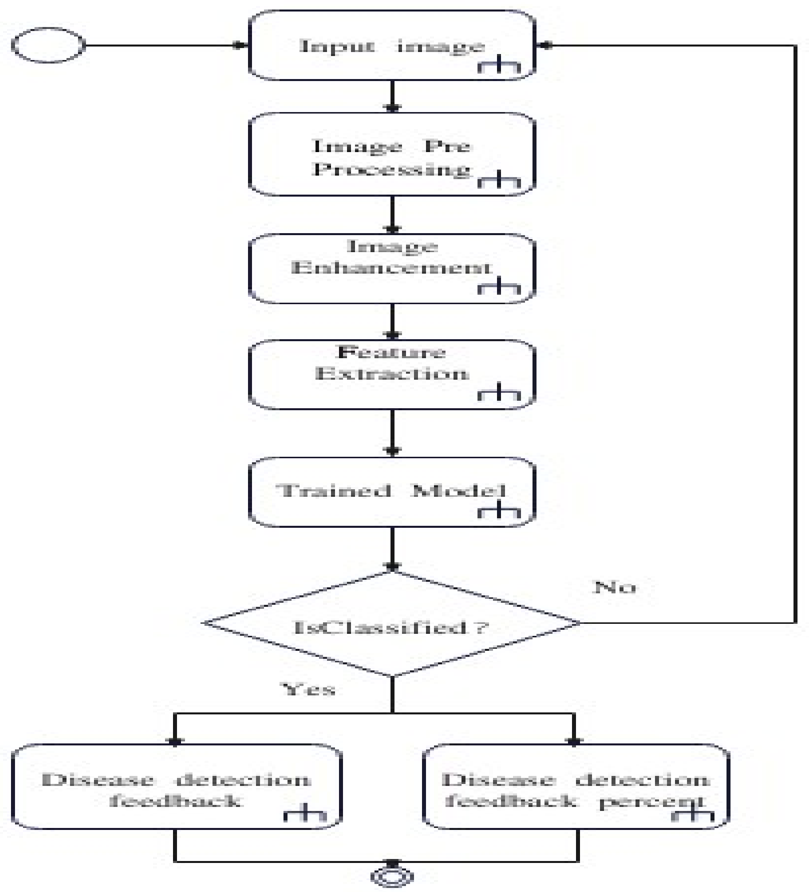

Figure 6 Activity diagram

The above flow represents the flow from one activity to another activity, the activity starts from input leaf image, and then input leaf is per-processed and extract the features. Now, the processed image is classified as Normal or Abnormal, if Abnormal is found in the leaf, then the result is display.

Chapter Four

System Method and Design

4.1 Design goal

4.2.1 Develop New Architecture

The proposed solution must leverage the current manual work flow. In the process of tomato plant disease detection and prediction system, automation the manual system shall be improved. Typically separate work activities into well-defined tasks, roles, rules, and procedures which regulate most of the work in tomato plant disease detection and prediction, this simplicity must be kept or improved in the new system.

4.2.2 New Model Development

At the end of this system new tomato plant disease detection and prediction model will be develop. The model can be used for other systems or it can be used to develop new models (can be used to fine tuning or transfer learning).

4.2.3 Collect Training Data

Data collection and pre processing is a major task in machine learning. The challenging part of machine learning is not about developing the models, having the right data is a very crucial point. In order to develop the model at least 15,032 labeled plant images will be need, after the model training the data can be open sourced and can be used for further research.

4.3 Purpose of the system design

4.3.1 Performance

The main attention of this system is performance. In real-time system performance is fundamental issue more important than quantitative measures. The major qualitative performance criteria are compiled. Systems fulfilling the qualitative aspects may then be compared on the basis of costs. To improve performance we can do many technique in our development but this are the most necessary get more data, invent more data, Rescale the data, transform the data and feature selection.

4.3.2 Dependability

For many computer-based systems, the most important system property is the dependability of the system. The dependability of a system reflects the user's degree of trust in that system. It reflects the extent of the user's confidence that it will operate as users expect and that it will not 'fail' in normal use. Dependability covers the related systems attributes of reliability, usability and availability. These are all inter-dependent.

4.3.3 Maintenance

There are always the considerations of time complexity, cost and functionality in a design. Considering these factors during the design process provides a meaningful basis to balance the needs of maintenance as we attempt to restore a system to service. There are some factors to consider when designing a system that will require maintenance of our system. Such as Interchangeability,Fault isolation and Modularization.

4.4 End user

In general, this System can be used by anyone who is involved in Ethiopia agricultural sector growing tomato plants.

4.5 Proposed system architecture

Figure 7 Proposed system architecture

4.6 System process

4.6.1 Planning and Projects Setup

To start the software development process first the software requirement specification document must be accepted. In the document, project requirements must be clearly stated, from the requirement we can decide the general trade offs with respect to cost, time, accuracy and speed.

4.6.2 Recognition Model

By the nature of the system, the models should work in real-time. The inference of the model should be very fast and accurate at the same time; to accomplish the models need to be highly optimized. Also we need to know that in real-time system accuracy and speed are inversely proportional, so we need to trade off each other.

4.7 Subsystem decomposition

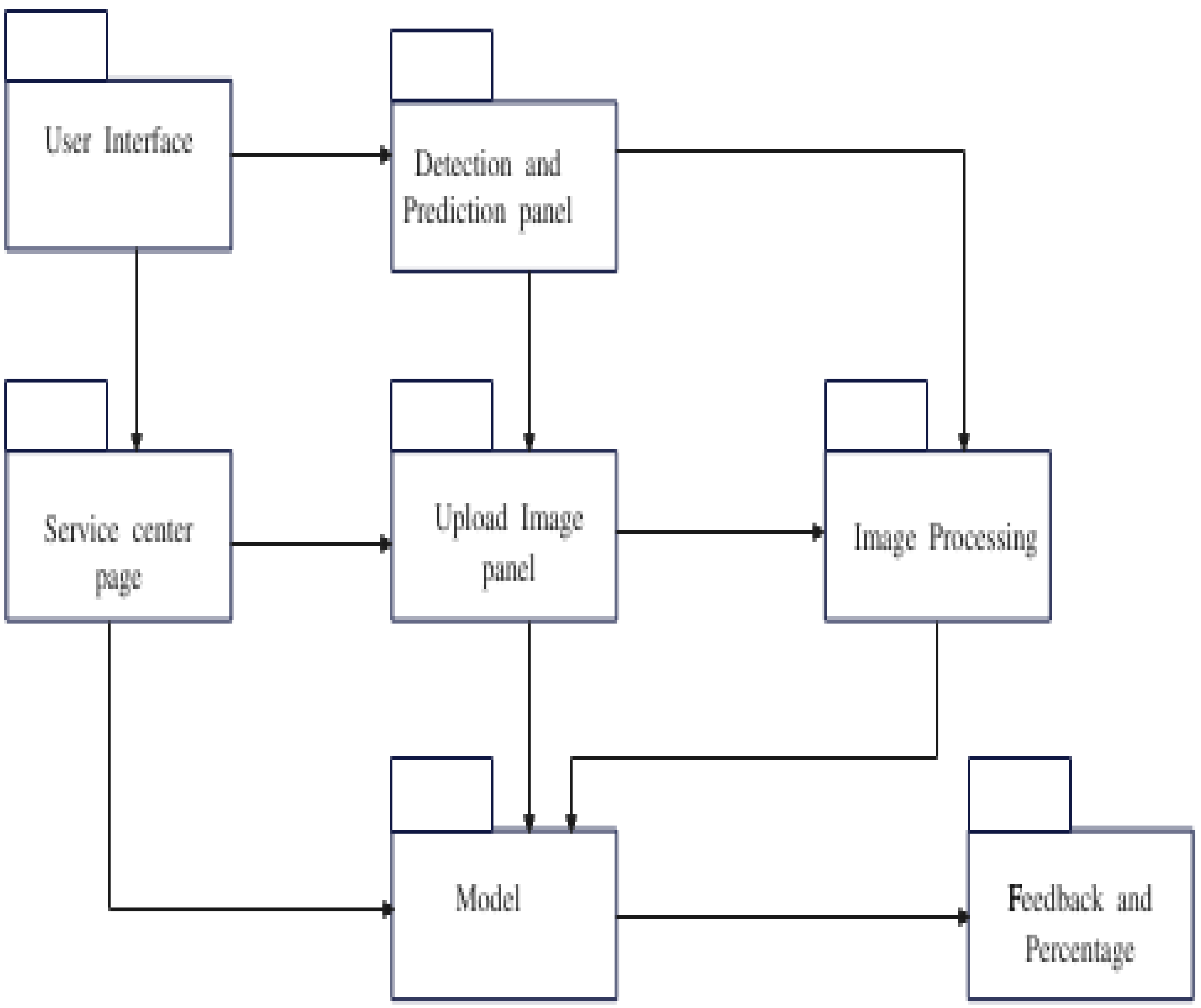

Figure 8 subsystem decomposition

4.8 Hardware and software mapping

The Hardware Software mapping is described to indicate the various hardware devices and equipments used in our system and its interaction with the software components.

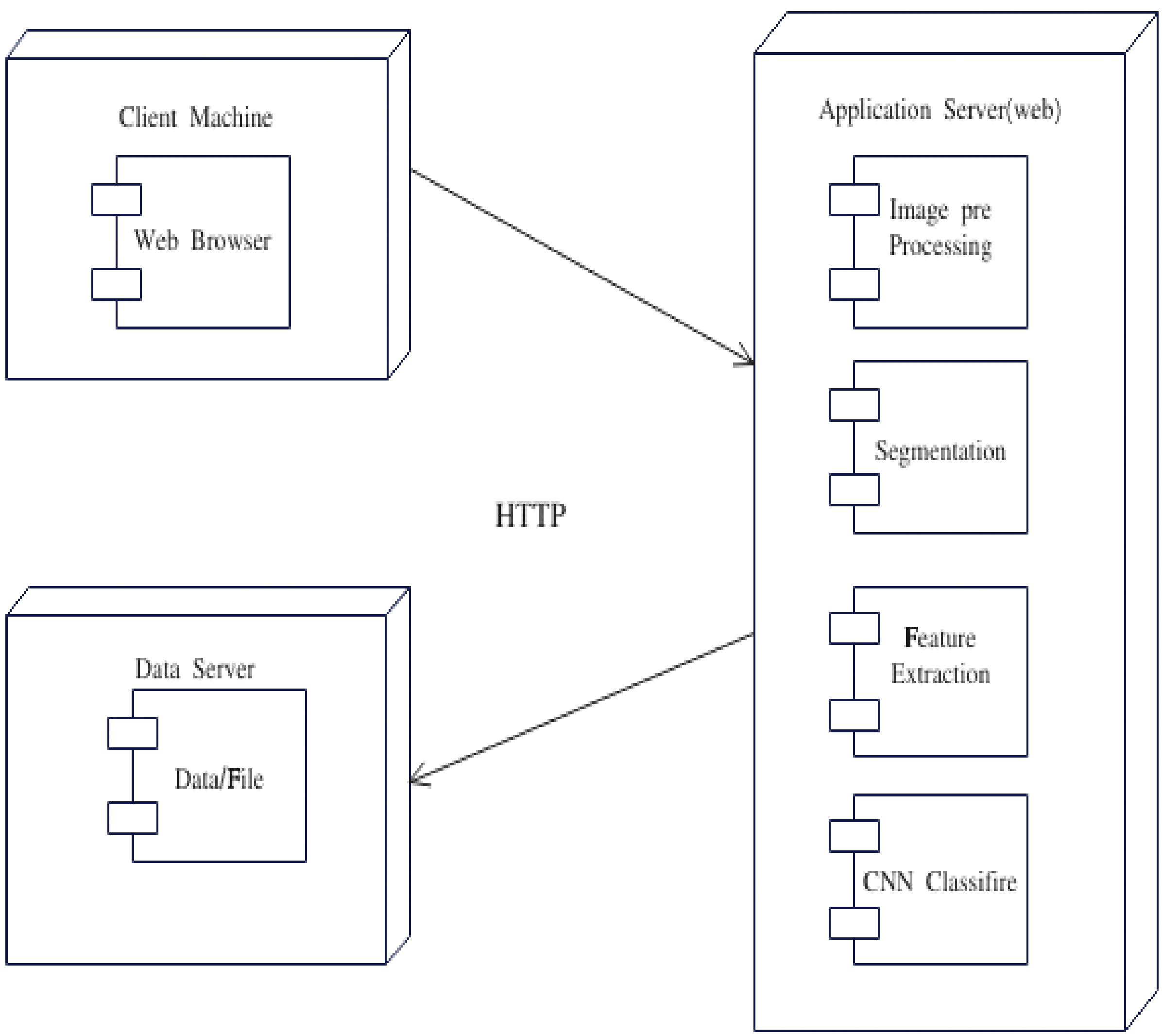

Figure 9 Hardware and software mapping

4.9 Component diagram

A component diagram, also known as a UML component diagram, describes the organization and wiring of the physical components in a system. Component diagrams are often drawn to help model implementation details and double-check that every aspect of the system's required functions is covered by planned development.

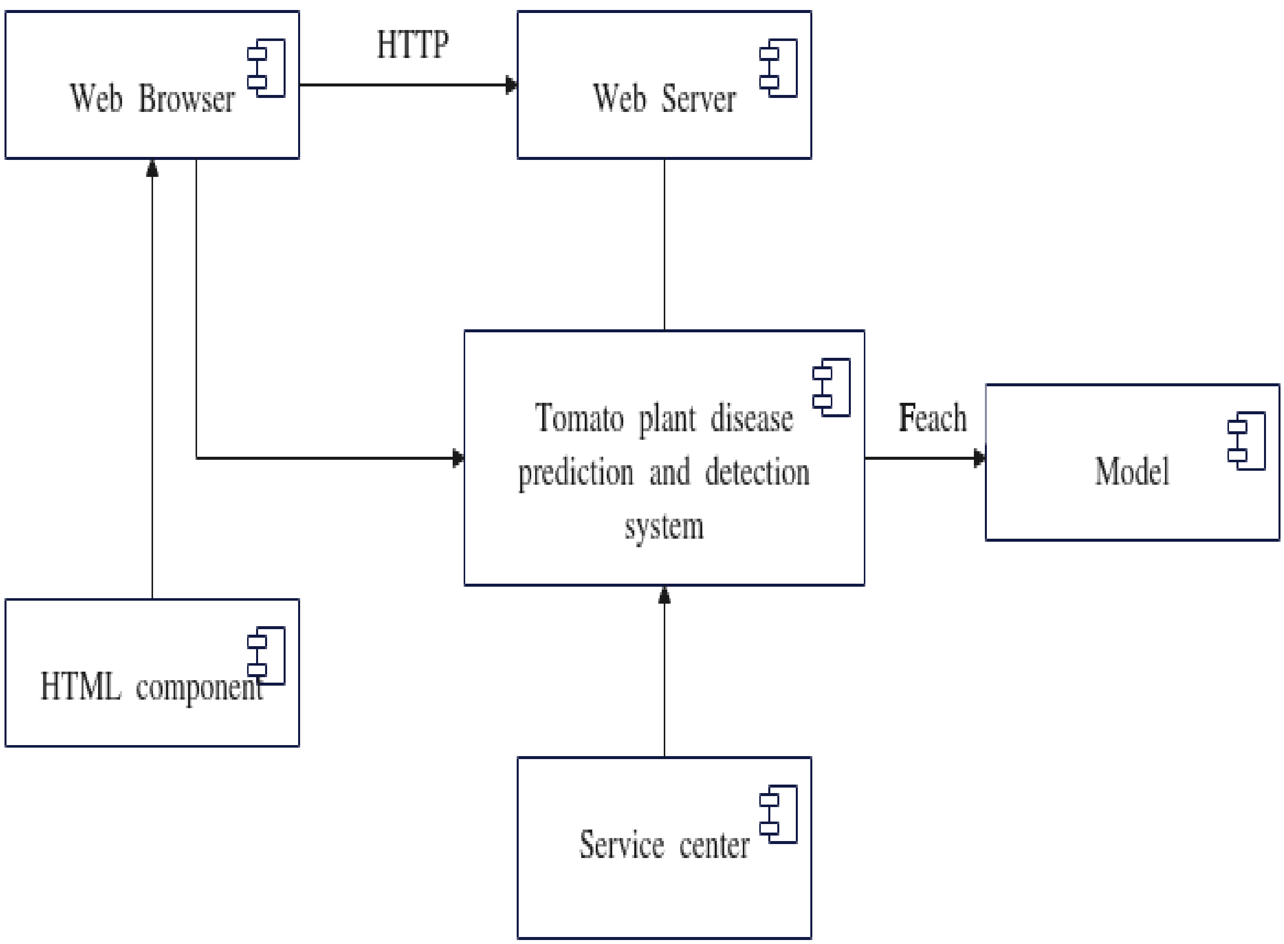

Figure 10 Component diagram

4.10 User Interface Design

4.10.1 User interface for Default page of the system

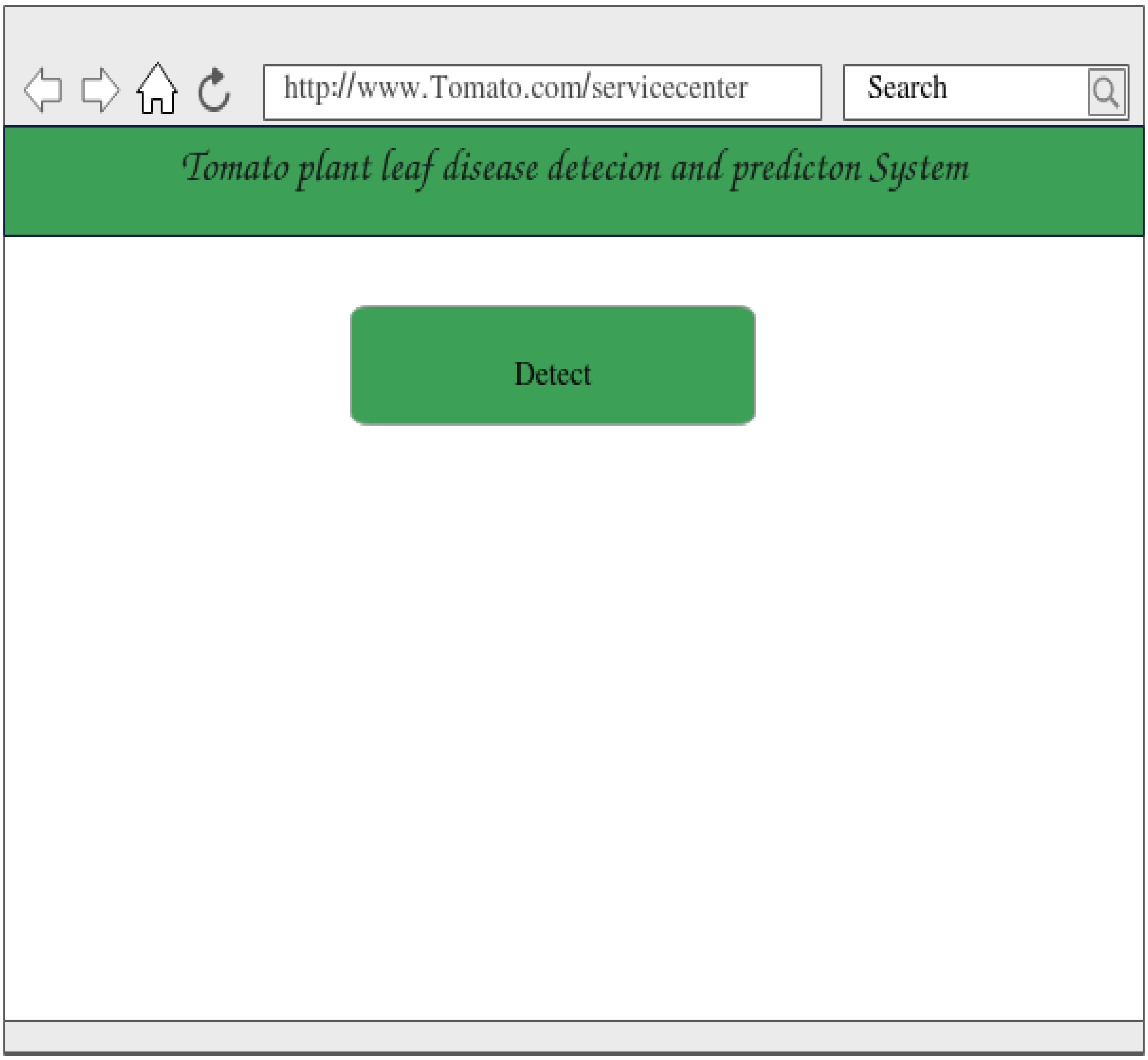

Figure 11 User interface for Default page of the system

4.10.2 user interface for uploading image

Figure 12 user interface for uploading image

4.10.3 User interface for prediction feedback and percentage

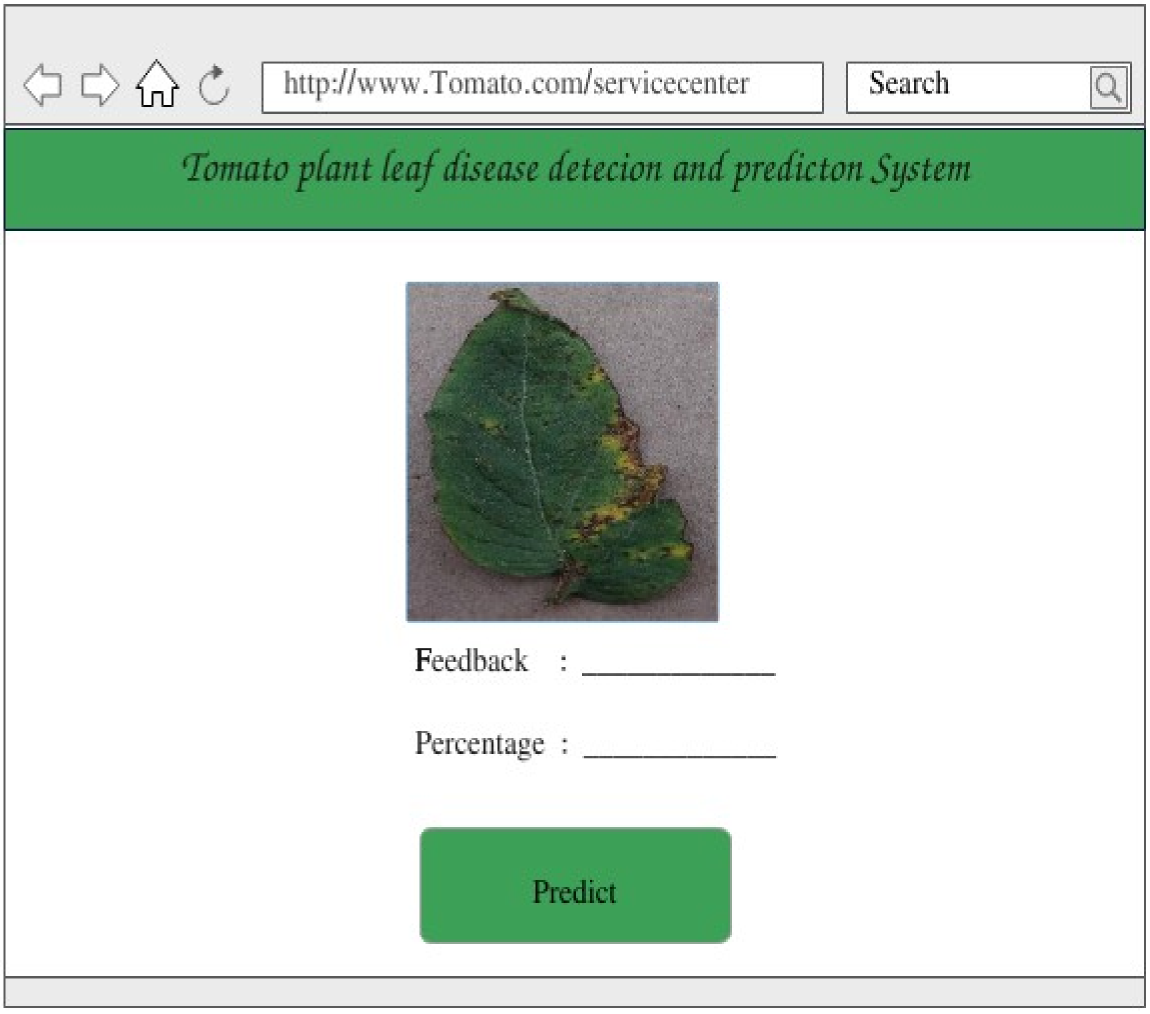

Figure 13 User interface for prediction feedback and percentage

Chapter Five
Conclusion and Recommendation

5.1 Conclusions

In this project, we expressed our idea to automatically detect and predict tomato plant disease using machine learning frameworks in order to reduce the human effort and cost. We tested the ability of the model to detect and predict tomato plant disease with high accuracies. The results highlight the importance of machine learning technology in detection and prediction of tomato plant disease. Hence, such frameworks can be used to reduce human effort and cost.

5.2 Recommendation

In the proposed system we we tire our best to get the desired output and the accuracy value is quite good for this method, yet there is still room for improvement as long as the accuracy is not exactly 100 percent. As a future work , researchers should increase the number of the images and the classes used in the dataset which will have an impact on the detection and prediction. In addition through this, a conclusion were drawn that machine learning techniques heavily rely on the input parameters. By tuning the CNN input parameters and layers, the accuracy might be improved. The future work should be to bring the plant disease detection and prediction accuracy over 99.7%.

References

[1] Agarwal, M. et al. (2020) 'ToLeD: Tomato Leaf Disease Detection using Convolution Neural Network', Procedia Computer Science, 167, pp. 293–301. doi:10.1016/j.procs.2020.03.225

[2] Ministry of State Farms Development (MSFD), 1984. State farm development, its role, organizations, present and future activities. MSFD, Addis Ababa.

[3] Dan C. Ciresan, Ueli Meier, Jonathan Masci, et al. Flexible, High Performance Convolutional Neural Networks for Image Classification[J]. PROCEEDINGS OF THE TWENTY-SECOND IN TERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 2011