System Design Specification
Of
Portable Soil Detection System based on Image Processing for Agriculture

Team Colon

Index number	Registration number	Name
3946	ICT/16/17/083	K. G. Rajapakse
3917	ICT/16/17/050	G.T. S. Perera
3884	ICT/16/17/008	W. M. D. H. Bandara
3929	ICT/16/17/064	J. P. R. Thakshila
3948	ICT/16/17/086	T. L. U. Sudarshana

2021/04/08

Abstract

There are different types of soil in the world, and the properties of each type of soil are different. Also, the diverse properties of the soil directly affect plant growth. They are properties that exist both physically and chemically. If the most important of these factors can be easily and accurately identified, the most suitable crops for these different soil types can be proposed. We propose a system that is easy to use to determine the properties of the soil and thereby identify the soil. Using machine language concepts, we analyze the data set. Data pre processing is done before training the model in order to enhance the accuracy. After that we will be using that data to train the AI model. With the use of this model, we develop a mobile application to deliver the output to the user.

Acknowledgement

This is a project carrying out as our third-year final project of Faculty of Applied sciences, Rajarata University of Sri Lanka. We, as group members are much thankful to the supervisors of our group, Mr. N. M. A. P. B. Nilwakke, who is a lecturer of the faculty of Applied Sciences and Miss. Malithi Jayasooriya, temporary lecture of the faculty for the guidance and the assistance given to us in the project work.
And special thanks must be given to the Natural Resources Management Center, Peradeniya for providing us with the soil maps of Sri Lanka and for the guidance given on the procedures and the knowledge given to us.

Other than them, we need to thank all our batch-mates who helped us to research about image processing techniques and soil sample collection process, other technologies and methodologies that we have to use when implementing the system

Table of content

CHAPTER 1 - INTRODUCTION

Agriculture could be considered as the key livelihood of Sri Lanka. For the inhabitants in Sri Lanka especially in the dry zone their daily income depends on the harvest they obtain from their cultivations. Hence their cultivations should have a profitable harvest. One of the major reasons for the lack of yield in their plantations is the incapability of the farmers to select the best crop which suits their land. Due to this reason there is a rapid drop in the number of people who are engaged in the agricultural sector. By the year 2020 the percentage of farmers whose main occupation is agriculture was reported as 23.73%. [1] Therefore, it is important to introduce a technique for the cultivators to select the best and the most profitable crop to be cultivated in their land.

Soil is an important influence to plants as it not only provides an anchor for the plants, but the different properties of soil which also provides several other physical and chemical factors which affects the plant growth. Among such factors are the porosity, moisture content, cation exchange rate, electrical conductivity, pH level etc. [2] These factors are manually identified using several chemical processes. If these factors could be identified precisely, easily and accurately, by combining these factors the best suitable crop/plant for a specific soil type can be determined.

Determining the properties of soil to estimate the strength of soil is one prime objective of geotechnical engineering. For this a range of traditional lab testing methods have been employed. But the data obtained from these tests are highly subjective, which even further complicates the existing variability of soil. [2]

Image processing has been incorporated with agriculture since recently and is widely expanding. Therefore, through the proposed Portable soil detection system we are planning to use image processing techniques to analyze a soil sample and suggest the best suitable crop for a given soil type.

1.1 Problem to be addressed

Even to date agriculture contributes to the national economy of our country at large scale. Cultivation of tea, coconut, paddy, vegetables and fruits, oilseed crops, spices etc. are among the most cultivated. When selecting a crop to be grown in a specific land most farmers tend to select a crop which commonly grows in the surrounding. This might lead to poor growth of that crop, pronging the crop to various deficiencies, reduction of harvest and they’re by spending money to overcome these problems result in extra expenditure, wastage of resources and time and overall a loss.

To minimize these problems farmers are educated to test their soil sample and determine the best suitable crop for that soil type as the variety in soil is immeasurable. The traditional testing methods used up to date in Sri Lanka are laboratory mechanical testing which occurs when the farmer takes a soil sample to the lab. For an example in order to determine the ph value of a soil sample the amount of hydrogen ions were calculated. And in order to find the electrical conductivity of the soil sample the salinity value was found by using an electrical conductivity meter. [2] But these methods are expensive, time consuming, involves much labor and also can be error prone as it's a mechanical process and is not very accurate. Hence the need of a method to make this process simple, more convenient and accurate for the farmers is seen as a gap.

1.2 Objectives of the project

Most farmers now have access to a mobile phone with a camera. If they can take a photograph of the soil sample and if the crop is suggested by the app at that instance itself, much time and effort in taking them to a lab will be saved, accuracy of the result will be higher, the extra expenses which have to be borne in selecting an unsuitable crop will not have to be spent and overall increase the harvest and profit for the farmer.

The main aim of this project is to develop a mobile application to address the above problem so that once the farmer takes a photo of the soil, he's planning to cultivate it automatically analyses the soil and suggests the most suitable crop to be grown in that soil so that the problems mentioned above are minimized.
We plan to fulfill this aim under two objectives using digital image processing.

1. Determine the type of soil by analyzing the soil sample
    I. Determine the soil type using color
    II. Determine the soil type using porosity
    III. Determine the soil type using grain size
    IV. Determine the soil type using moisture

2. Suggest the best suitable crop for that soil type instantly.

1.3 Project Deliverables

A project deliverable is any particular kind of output element that is a result of a well planned work done during a project. It is essential that a deliverable is within the scope of the project and that is has a definite role in accomplishing the projects objectives.

1.3.1 Deliverables of the documentation

Deliverable No	Deliverable Name
D 1.3.1.1	Team Registration
D 1.3.1.2	Project Proposal Submission and Presentation (Hardcopy + Softcopy)
D 1.3.1.3	Revised Project Proposal Submission (Hardcopy + Softcopy)
D 1.3.1.4	System Requirement Specification (SRS) Document Submission and Presentation
D 1.3.1.5	Interim Report Submission
D 1.3.1.6	System Design Specification (SDS) Document Submission and Presentation.
D 1.3.1.7	Scrum Report Submission and Viva + Presentation
D 1.3.1.8	Progress Report Booklet 01 Submission
D 1.3.1.9	Progress Report Booklet 02 Submission

D 1.3.1.10	Final Dissertation Submission
D 1.3.1.11	Research Paper Write-Up Completion

1.3.2 Deliverables expected in the implementation of the project.

Deliverable No	Deliverable Name
D 1.3.2.1	Organizing and carrying out a survey on existing image processing techniques to detect soil types to gain further knowledge
D 1.3.2.2	Collect the dataset
D 1.3.2.3	Preprocess the data to identify the key soil attributes that are useful to train the model
D 1.3.2.4	Analyzing the key soil attributes and categorizing the soil types
D 1.3.2.5	Developing the mobile application
D 1.3.2.6	Final product delivery

1.4 System Design Approach

A software process model is a conceptual illustration of a process that presents a description of a process from a particular perspective. Software Process is a logical set of activities for specifying, designing, implementing and testing software systems. According to the concept of process modeling this proposed system is component based. In order to obtain a premier outcome, the proposed system stands with the Agile model which is one of the process models of the software development life cycle. The process of this application is planned and oriented to be complete under this agile scrum model.

Scrum is a framework which can be used by people to address complex and adaptive problems, while delivering high quality products productively and creatively. The Scrum process model is one of the most popular ways to implement Agile process model. It is an iterative software model which follows a fixed set of roles and responsibilities.

As the portable soil detection system for agriculture using image processing is a team project which has to be completed within a limited time frame and as there are many changes that will have to be made the team has chosen the scrum model to monitor the progress towards the sprint goal and to inspect how the project is trending towards the completion of work in the Sprint Backlog. Generally there are about 5 to 9 members in a scrum team. However in the proposed project the team consists of 5 members.

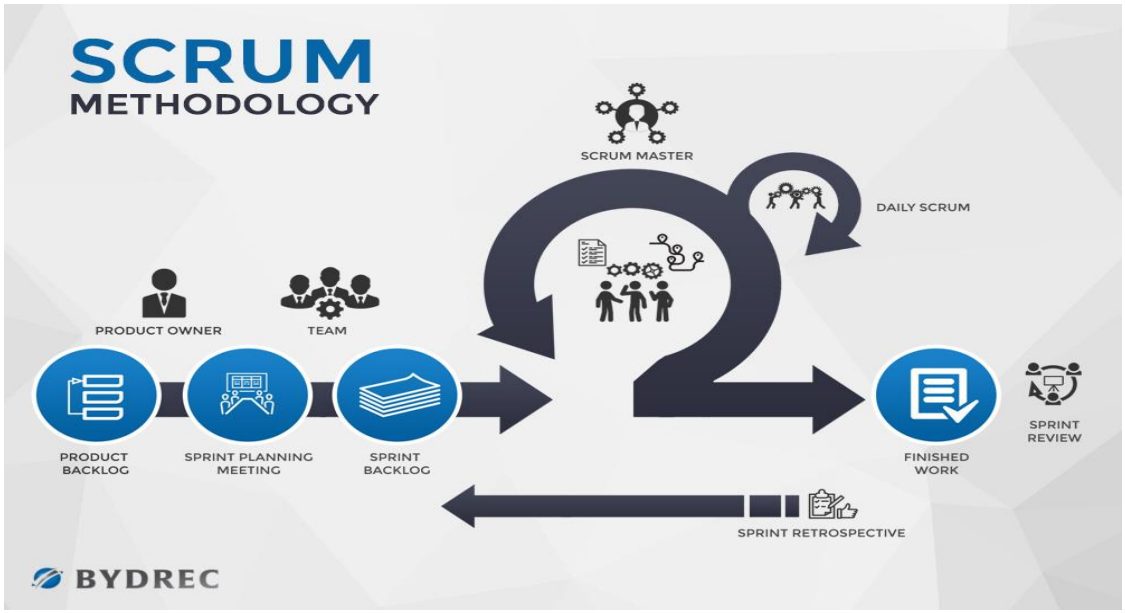


Figure 1-1 Scrum Framework

Image source - https://blog.bydrec.com/a-comprehensive-comparison-between-the-agile-scrum-and-waterfall-methodology

1.4.1 The Scrum Events

Once a Sprint begins, its duration is fixed and cannot be shortened or lengthened. The remaining events may end whenever the purpose of the event is achieved and ensuring a sufficient amount of time is spent without allowing a waste of time in the process. The Scrum Events are:

- Sprint – a sprint is a short time boxed period in which a scrum team decides to complete a set amount of work. In the proposed system each sprint is going to be 2 weeks long.
- Sprint Planning – this occurs at the beginning of the sprint. Here the team determines the sprint backlog items which they plan to work out during that spring.
- Daily Scrum – a face to face meeting among the team members held on every working day of the sprint to discuss the daily work and progress and it normally lasts for about 15 minutes/
- Sprint Review – it is an informal meeting which the development team, product owner, scrum master and the stakeholders attend for the team to show the customers and stakeholders what they have achieved during that sprint.
- Sprint Retrospective – it is an opportunity for the scrum team to inspect themselves and o create plan for the improvements which need to be enhanced for the next sprint.

Prescribed events are used in Scrum to create regularity and to minimize the need for meetings that are not defined in the Scrum model. All the events are time-boxed.

1.4.2 Scrum Artifacts

Scrum's artifacts present a work or value to provide transparency and opportunities for inspection and adaptation. Artifacts defined by Scrum are specifically designed to maximize transparency of key information so that everybody has the same understanding of the artifact.

The Scrum Artifacts are:

• Product Backlog – it is a prioritized list of work for the development team throughout the entire process

• Sprint Backlog – it is a prioritized list of work for the development team during a specific sprint

• Increment – it is a collection of the completed sprint

The scrum model is chosen over the other traditional models such as waterfall model, V model and spiral model because they are not suitable for this project as there is a limited time frame and frequent changes have to be made throughout the project.

1.5 Standards to be followed

1.5.1 Standards for system design specification

1016-2009-IEEE Standard for Information Technology—Systems Design—Software Design Descriptions is a standard which describes software designs and sets up the content and organization of a SDD. SDD stands for the Software Design Description which is used for recording design information and communicating that information to the key design stakeholders.

1.5.2 Standards for Image Processing

ISO/IEC12087-This standard is used to provide an application program interface and an image interchange representation in order to increase the portability of application software. This model defines an API which is applicable to all application areas that involve processing, manipulation or transfer of image data.

1.5.3 Standards for Coding

General coding standards will be used in this application such as naming conventions, commenting and documenting. Meaningful names and comments were used as much as possible to enhance the maintainability of the project. The code is written in a standard format using proper indentation. The code is written using python language as it is one of the standard languages when developing an AI model.

1.5.4 Standards for Database

In this application SQL is used as the database language as it is one of the standard database languages according to ISO/IEC 9075 standard:"Information technology – Database languages-SQL".

1.6 Organization of the SDS

1. **Introduction**

In this Chapter we discuss the problem to be addressed, objectives of the project and the project deliverables. The system design approach, The Scrum model is briefly introduced in the middle section of this Chapter. Furthermore in this section, we have discussed about the standards to be followed in developing this application. Finally the organization of the SDS is conversed at the latter part of this Chapter.

2. **Architectural Design**

In the second Chapter, System architecture, objects and communication will be discussed. The processes and the special algorithms that are used in this application will be briefly introduced in the middle section of the second chapter. Furthermore in the second chapter we will discuss about the state machines, tools, techniques, libraries, third part tools and the implementation environment.

3. **UI Design**

At the beginning of the third chapter, the PACT analysis will be described followed by the UI design consideration and approaches. The design tools, techniques and templates will be discussed in the middle part of the third chapter. Furthermore in this chapter we will discuss the input, output and dialogue design aspects. Finally the hosting and installation environment will be conversed at the latter part of the third chapter.

4. **Data Management**

In this chapter initially we discuss the data requirements, design tools and techniques. Towards the middle part of this chapter conceptual database design, logical database design, schema refinement and physical database design will be conversed. In the latter part of the chapter security design will be briefly introduced.

5. **Research Design**

In this chapter we study the objectives based on the literature review that we have conducted and formalize high level implementation components. Extraction of data, samples design, testing data sets and training data sets are addressed towards the middle of this chapter. Non functional aspects, the proposed validation methods and measurements are briefly elaborated at the end of this chapter.

6. Approval

In the final chapter the recommendation of the supervisor and the details of the team members are presented.

CHAPTER 2- ARCHITECTURAL DESIGN

In this chapter we intend to discuss about the architectural design of the portable soil detection system using image processing for agriculture. Under this Chapter we elaborate the system architecture, objects and communication, pre processes and special algorithms, state machines and tools, techniques, libraries, third party tools and implementation environments in detail.

2.1 System architecture

2.1.1 Component diagram

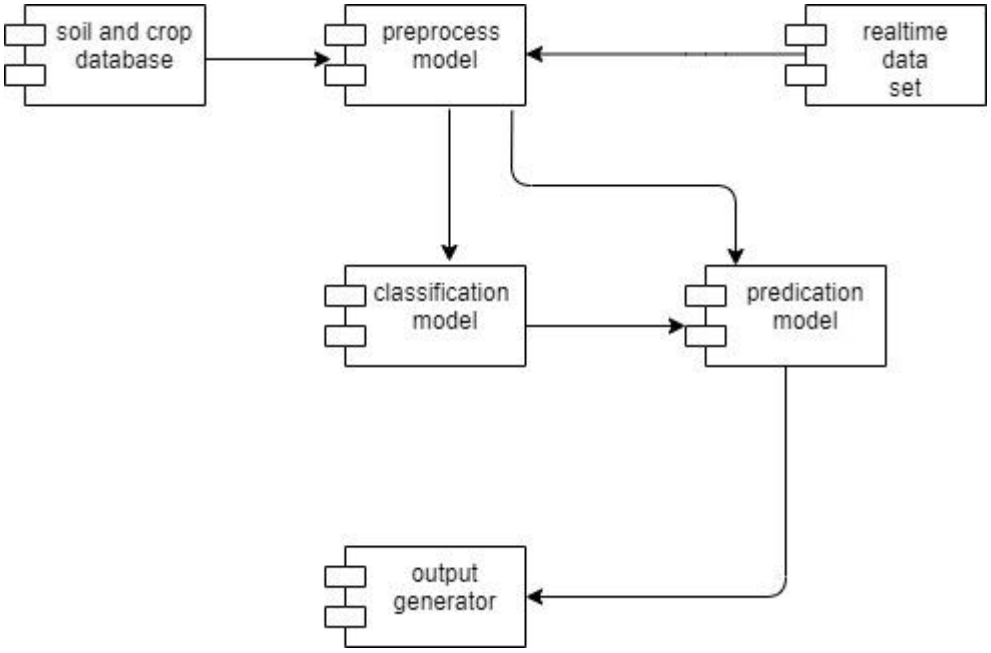


Figure 2-1 Component Diagram

2.1.2Components

2.1.2.1 Preprocess model

The format of the data has to be changed in to a proper format in order to achieve best result from the AI model. Therefore prior inserting the data in to the model the data has to be pre processed. The soil types and their properties in the soil database and the real time data are preprocessed in this module. Preprocessing plays a vital role in AI as the data may contain missing values, null values and some incomplete data. By preprocessing the data will be cleaned and the missing values will be filled. Through pre processing the key soil attributes could be identified which is useful for model training.

2.1.2.2 Classification model

The classification model will conduct statistical analysis on the dataset and create graphs on the dataset to give a more clear idea. After analyzing the dataset the patterns and trends of the soil attributes can be identified.

2.1.2.3 Predication model

In this component the prediction model will predict the soil type using the preprocessed and the classification dataset. For this purpose the AI model will be trained using the preprocessed and the classification data.

2.1.2.4 Output Generator

With the use of the predicted data of from the predication model, the output generator will generate it's output. This output will either be the soil type or suggested crop depending on the result requested by the user and it will be displayed on the screen.

2.1.2.5 Realtime Dataset

The real-time dataset of this application involves with the different soil types and crop types updated to the database by the admin. This will be done occasionally by the admin with the advent of an unexisting soil type or crop type in the database.

2.2 Objects and communication

2.2.1. Sequence diagrams

2.2.1.1 Select Location

In this sequence diagram, the location selection process is represented. The application initially displays two options to select the location. They are selecting the location manually and through GPS technology. If the user chooses to select the location manually, the application will display the available provinces, districts and divisional secretariat divisions respectively to the user. Once the user selects the location it will be stored in the database. The alternate path is selecting the location through GPS technology. Here the user simply has to switch on the location settings of the mobile device and then the GPS facility of the device automatically detects the location of the user and returns it to the application.

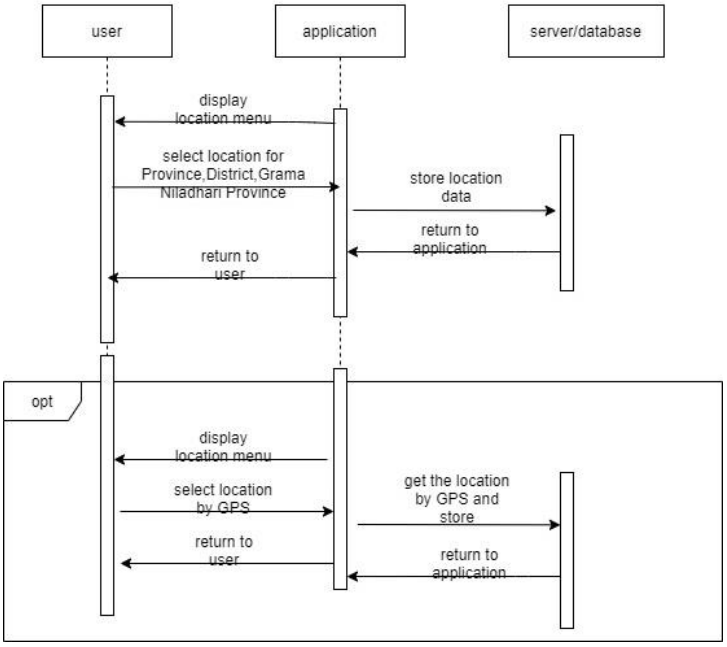


Figure 2-2 Select Location

2.2.1.2 Upload Photo

In this sequence diagram the process of uploading the photograph is represented. The application initially displays two options to upload the photo. They are capturing a photograph through the application and uploading an image already existing in the gallery. If the user chooses to capture a photograph through the application, the user will be directed to the device camera with a square in the middle so that the user will have to capture the soil sample within that square frame. When the photo is successfully taken within the given resolution boundaries it will be stored in the database. The alternative path is to upload an existing image from the gallery. If the user chooses that option, he will be redirected to the gallery interface of the device and the user can select the relevant image from the gallery. Thereafter, the image will be cropped according to the relevant dimensions and if the image is with the required dimension and resolution then the image will be successfully stored in the database.

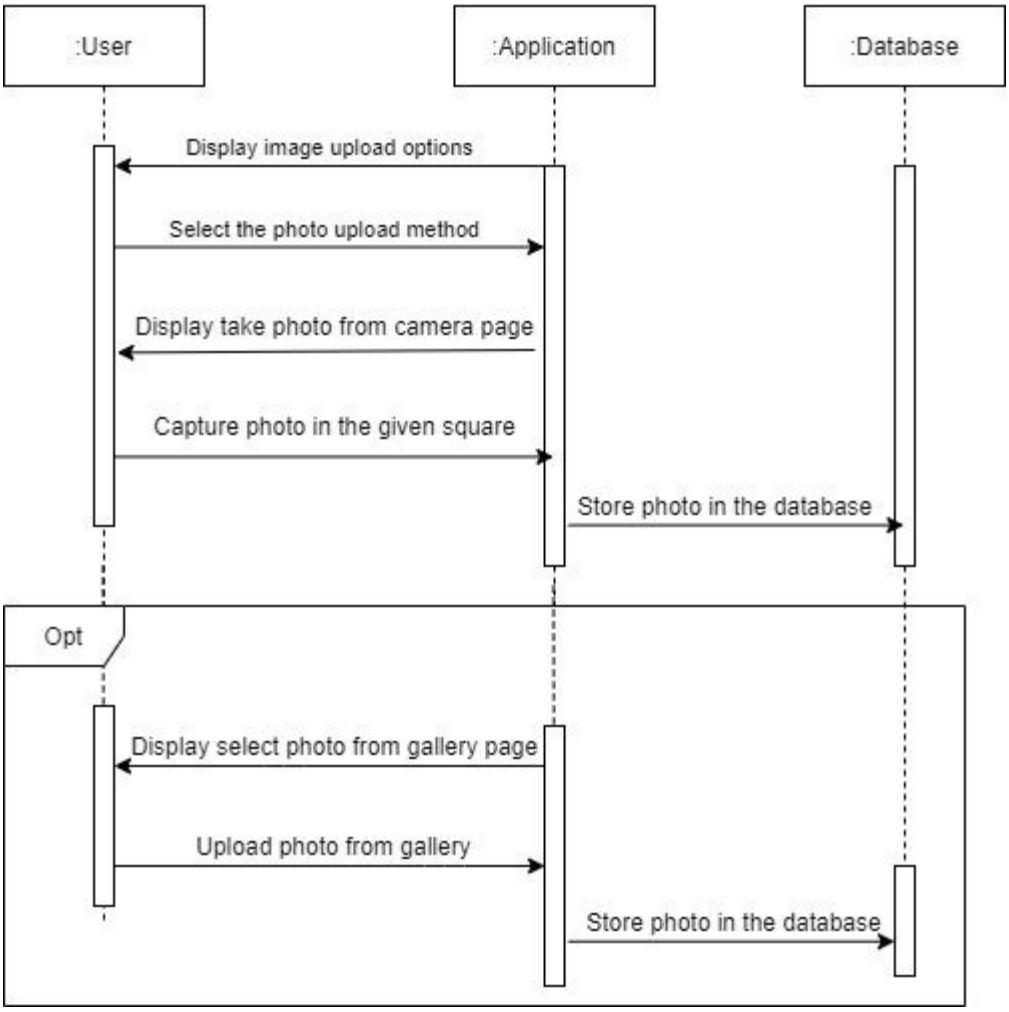


Figure 2-3 Upload Photo

2.2.1.3 Analyze the soil type

This sequence diagram represents the process of analyzing the soil type using image processing techniques. The application will analyze the previously uploaded photo by means of soil color, porosity, moisture and grain size. After the photo is being analyzed, the result will be displayed to the user.

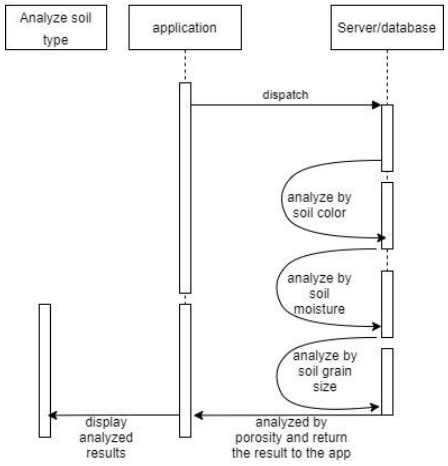


Figure 2-4 Analyze the soil type

2.2.1.4 Display the suitable crop

This sequence diagram represents the process of finding and displaying the most suitable crop to be cultivated in the previously analyzed soil. Here once the application displays the type of the analyzed soil sample, if the user wishes to know the best suitable crop to be cultivated in that soil then the user must click on the option button to view the suitable crop types. Then the user will be directed to an interface consisting of the suitable crop types after being compared with the database.

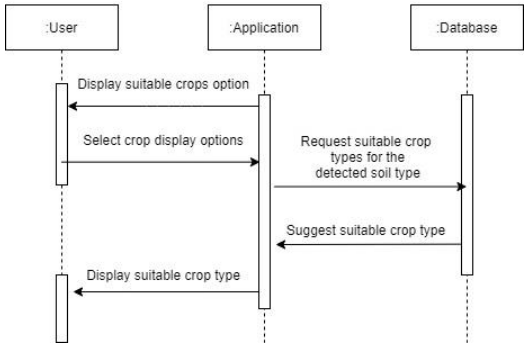


Figure 2-5 Display most suitable crop

2.2.1.5 Login as administrator

This sequence diagram represents the process of logging in as the administrator. The application will initially provide an interface for the administrator to login. Once the administrator clicks on the login option, he will be directed to the login form where he has to enter the user name and the password which will be validated through the application. If the data is successfully validated the application will display a login success message.

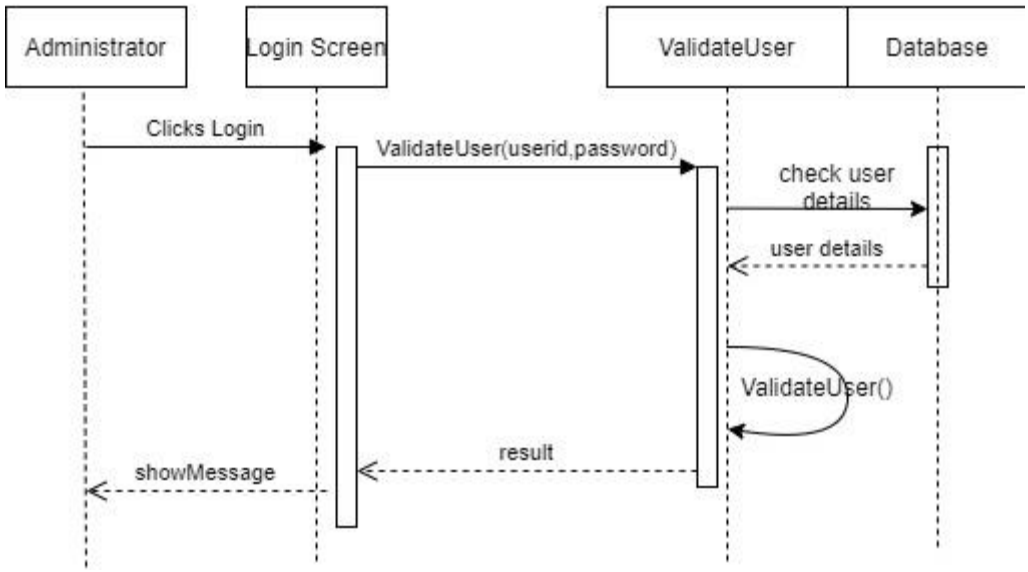


Figure 2-6 Login as Administrator

2.2.1.6 Update database

This sequence diagram represents the process of updating the database. In order to update the database, the administrator must initially login to the system. There after the administrator can either update an existing record in the database or can add a new record. The update database process involves either updating or adding a crop type or soil type.

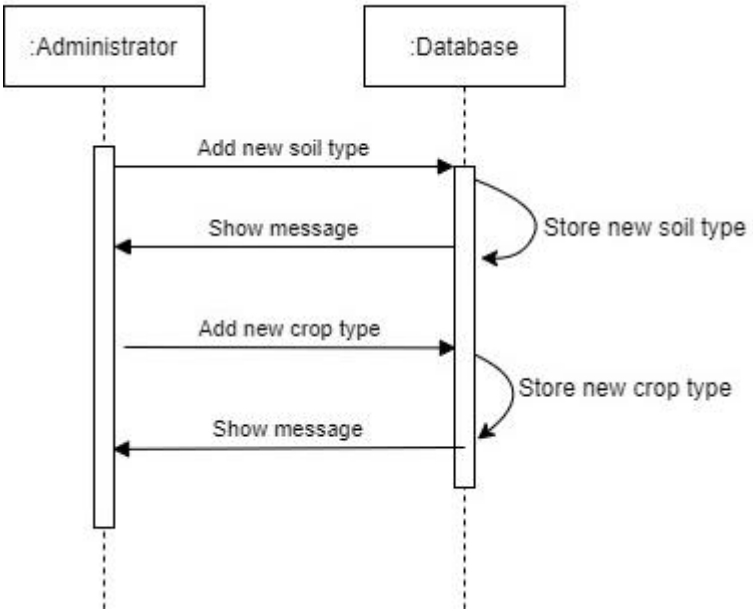


Figure 2-7 Update Database

2.3 Processes and special algorithms

In processes and special algorithms include the algorithms that are going to use for the system device. And also the processes which we are using in the system. There are several algorithms and processes related to our system. In this section we describe one of the main processes and one of the main algorithms that are important to our device.

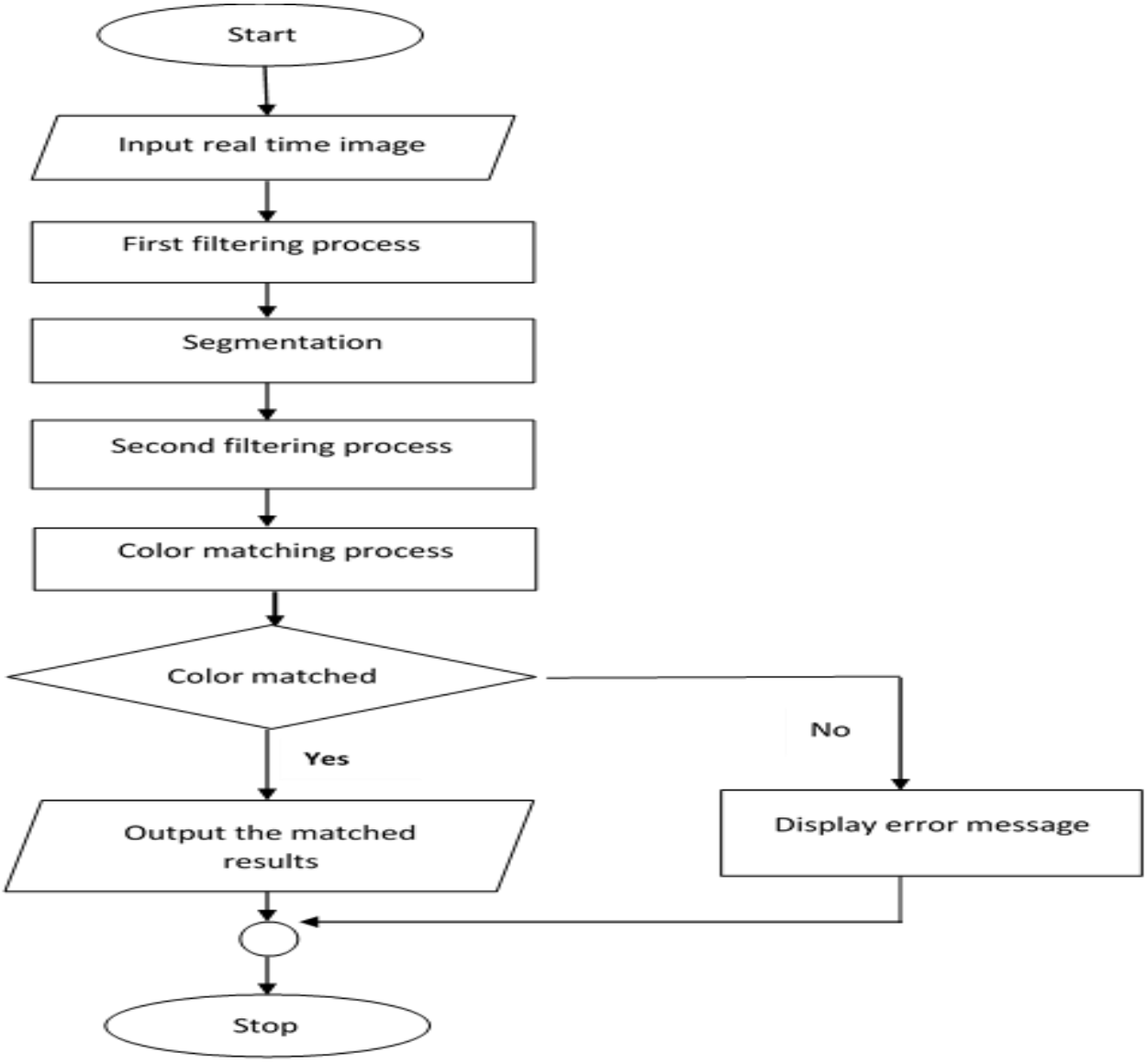


Figure 2-8 Algorithm for color comparison

In the above flow chart shows the process of matching a color of a soil from the previously stored data. It takes an input of real time image of a soil sample and enhance that image by using filtering processes and compare it with the database images. This algorithm mainly consists of five steps. Using this algorithm we detect the soil type by matching the soil color.

We use K-Nearest Neighbor (KNN) as their soil color classification algorithm for soil color classification. . It is a supervised machine learning technique. KNN considers the k nearest cases from an instance and decides which the most frequent class in the set is. The most repeated class is assumed to be the class of that instance (x), KNN system adopts a distance metric in order to

determine nearest instance. Various distance metrics can be adopted including Euclidean which is used at this research as it gives the best results.

2.4 Tools, Techniques, Libraries, 3rd party tools and implementation environment

2.4.1 Microsoft Visio

Microsoft Visio is software for drawing a variety of diagrams. These include flowcharts, org charts, building plans, floor plans, data flow diagrams, process flow diagrams, business process modeling, swim lane diagrams, 3D maps, and many more. Visio can be used in a variety of setting to create professional-looking diagrams. Visio includes a big library of shapes/symbols used in dozens of diagram types. Using Visio, we use for drawing sequential diagrams and flow chart. 12

2.4.2 Python

Python is an interpreted, interactive, object-oriented programming language. It incorporates modules, exceptions, dynamic typing, very high level dynamic data types, and classes. We hope to use python as our implementing environment.

2.4.3 Machine learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. We are going to use some machine learning approach to build our AI model.

2.4.4. Open CV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products

CHAPTER 3 – UI DESIGN

User interface is a bridge between a human and computer interaction. Enabling a user to effectively control a computer or machine they are interacting with and receiving feedback to communicate effective completion of tasks is the main purpose of a user interface. A successful interface should be intuitive, efficient and user friendly which means that the operator should be able to achieve the desired output by providing a minimum input or in other words the machine should minimize undesired outputs to the user.

3.1 PACT analysis

The PACT analysis is a framework which is used to analyze with whom, with what and where a user interacts with a user interface. The acronym PACT stands for people, activities, context and technologies. In this section of the documentation the UI design of the portable soil detection system based on image processing for agriculture is analyzed as a summary in terms of people, activities, context and technologies. PACT analysis could be considered as a fundamental requirement analysis technique which is used by most of the traditional software engineering techniques such as scrum. This could be done using various techniques such as brain storming, observations, interviews and workshops. In our system we have used observation technique.

3.1.1 People

There are many different ways in which people differ from each other. That mainly depends on the five senses hearing, vision, smell, taste and sight. Hence it is important to ensure that the system is user-friendly, usable and simple for everyone.

In our application, portable soil detection system based on image processing for agriculture our target audient will be the farmers in the North central province of Sri Lanka. As the main livelihood of the people living in that region depends mainly on agriculture it is critical that they select the best suitable crop to be cultivated when preparing their land to reap an optimal harvest with a minimum cost. Here the stakeholders and the users are the end users (farmers), administrators, design, development, testing and maintenance team.

There are a variety of differences among the users and their interactions which can be discussed under the following criteria.

Physical differences

One of the major physical distinctions among people is the difference in their five senses. Through our observations we have recognized that there are several people who have vision impairments such as color blindness and long sightedness. Hence we have done improvements to our interfaces such that the colors red and green are not used to represent important content and an average font size was used for an efficient delivery of the content to the user.

Psychological differences

Since the native language of Sri Lanka is Sinhala in addition to English we have used Sinhala language also to ensure that the native farmers understand the important facts in our user interfaces. Also as many traditional farmers are not tech savvy we have made the interfaces simple and attractive as much as possible.

Social differences

As this project is designed for the farmers it is targeted at a homogeneous group. Hence mostly the entire audience expects the same things from the system. Also as this system is designed for the android operating system the users should be able to afford an android mobile phone.

3.1.2 Activities

Frequency

This application is used every time when the farmer decides to cultivate a new crop. Therefore it is not used very frequently by a single user.

Tasks

When the application is downloaded the users should select his current location and upload a photograph of the soil sample to be detected. Then once the system analyze and provide the type of the soil the user can further check the most suitable crops to be cultivated in that soil type.

The tasks must possess special characteristics in order to have a better interactive application. Some of the main characteristics of the tasks that need to be considered are,

1. **Load time of the application**

   The application should be simple enough to be able launch within a short period of time and we have reduced it to 2 seconds in our mobile application.

2. **Processing time of the application**

   The efficiency and user friendliness of our application is enhanced by reducing the processing time to less than 3 seconds which leads to a higher customer satisfaction.

3. **Photo upload time of the application**

   The photo upload time of our application is reduced to 10 seconds so that the user won't get exhausted and change his mind to quit the process.

3.1.3 Context

The physical environment of this application is varied together with the climatic changes and it has a severe impact on the application as well. In order to deal with this we are using the weather information such as the time period this region gets rainfall.

All the activities could be performed both indoors and outdoors. When the application is used outdoors the brightness of the screen should be higher than that of the environment so that the users can see the content clearly. But it should not be too high so that the user's eye does not feel any discomfort. This portable soil detection system is made for the personal use hence it is not affected by any organization or an external user.

When this application is used outdoors there is a higher probability that the user's hands could be sweaty. Therefore we have minimized the number of swiping options as when the user swipes the screen the since the hands are slippery the finger movement will be unable to detect.

The select location activity has its own context. Here as the application gets the users current location through GPS most users are reluctant to provide the permission for it. Therefore it is important to explain the purpose of obtaining the current location to them and also that the location is not shared with any third party.

3.1.4 Technologies

Input

Drop down menus- when the user needs to select the location, the locations are suggested by a drop down menu so that the user doesn't have to remember anything.

Map view – as the user can select his current location through the GPS a map view will be provided.

Camera – the camera is provided through the application to upload the photo of the soil sample. Here a rectangular frame will be used to capture the photograph according to the required measurements.

Output

Phone display – almost all the outputs are given from the phone display hence the user will have to rely on his vision.

Communication – an internet connection is required by the user to upload the photo and receive the outputs. However all smart phones comes with a Wi-Fi connection or 3G support.

The technology used for the UI design is Android studio. Android studio is an integrated development environment which provides new tools for application development, making it possible to view any changes done to the application in real time and view how the output look like in different devices. The version of this software we are using is Android Studio 3.5.

3.2 UI Design Considerations and Approaches

- User familiarity

  The interfaces should be designed with words and phrases which are familiar to the targeted audience of the application.

- Consistency

  The interfaces should be designed such that comparable operations should be activated in the same way and the similar elements in all interfaces should have a similar look and should stick in to single theme.

- Recoverability

  The interfaces should include mechanisms to allow users to recover from errors.

- User guidance

  When errors occur, the interfaces should provide with meaningful feedbacks to recover from them and provide context- sensitive user help facilities.

- User diversity

  The interfaces should appropriately respond to different types of application users.

- Responsiveness

  The responsiveness of the mobile application should be quick

- User friendly

  The application should be user friendly and easy to manage.

3.3 Design tools and techniques

3.3.1 Design tools

- Flowchart maker and online diagram software (https://app.diagrams.net/) – This software was used to create design diagrams such as activity diagrams, ER diagrams, Sequence diagrams etc.
- Ms Word – This software is used to create the documentation
- Ms PowerPoint – This software is used to create slides used or the presentations
- Android Studio – This software is used to develop the mobile application
- Open CV – This software is used to apply the image processing techniques required in our system
- XAMPP server – This is used to run the server side processes
- Ms Visual Studio 2019 – This is used as the IDE for Open CV
- pycharm

3.3.2 Techniques

- Set user interface design standards and stick to them
- Rather than explaining in detail exactly how to use each feature in the application step by step a clear set of instructions are explained at the beginning once.
- Using full words and sentences for messages and labels will enable the users to interact with the application successfully.
- Design the interfaces as much as simple to understand to the client easily.
- Using colors in the interfaces which are pleasant to the user's eye

3.4 Input Design Aspects

3.4.1 Select location

In the first input screen there are two buttons. One button is to set the location by GPS where as the other button is to set the location manually. If the user selects the set location manually button the user is directed to an interface with three drop down menus each to select the province, district and the Grama Niladhari division respectively. The user then has to press the confirm button to set the location. If the user selects the set location by GPS button the user is directed to an interface containing a map which shows the current location. Then the user has to press the confirm button to set the current location by GPS.

Figure 3-1 Home page

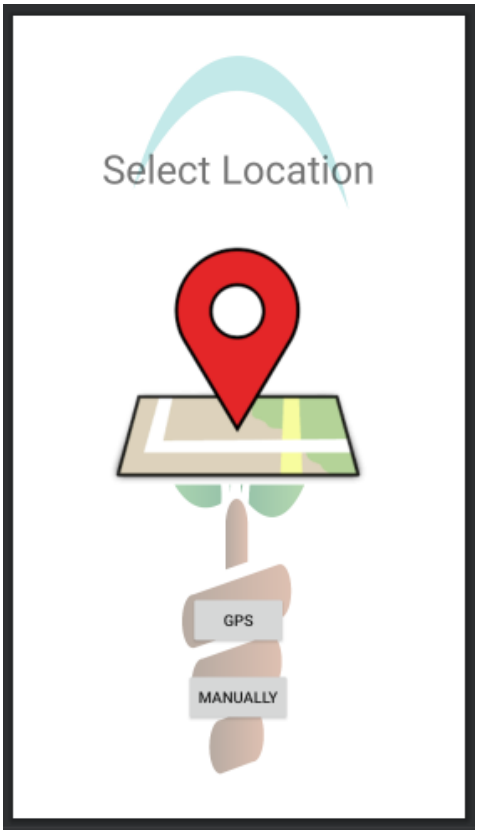


Figure 3-2 select location page

Figure 3-3 Get Location by GPS interface

Figure 3-4 Get Location by Manually interface

3.4.2 Upload photo of the soil sample

There are two buttons in the first input screen of the upload photo of the soil sample interface. One button is to upload the photo from the gallery whereas the other button is to take a photo from the camera through the application. If the user presses the upload photo from gallery button the user is directed to an interface containing all the albums available in the phone gallery from which he has to select the relevant photo then it is uploaded to the application. The user then has to press the upload photo button to confirm the upload. If the user presses the take photo from the camera button the phone camera is opened with a rectangular frame of the required measurement, through the application where the user captures the photo of the soil sample and presses the upload photo button to confirm the upload.

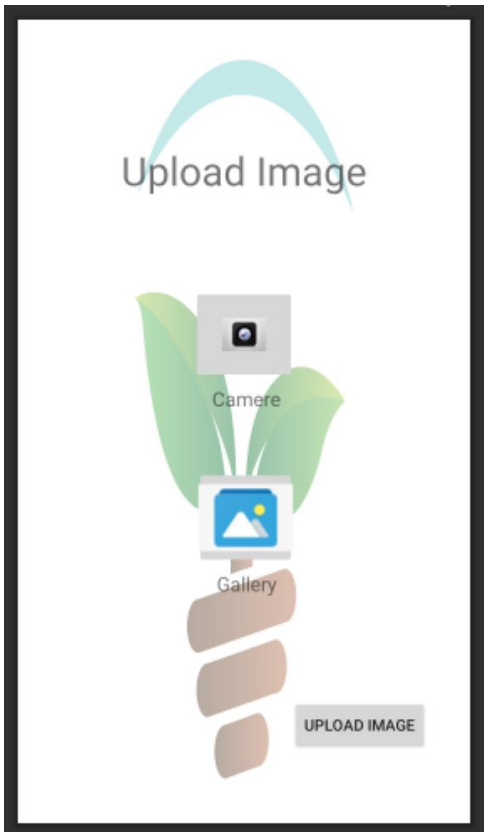


Figure 3-5 Upload image interface

3.5 Output Design Aspects

This application contains two output interfaces, one which outputs the detected soil type and the other which outputs the type of suitable crops for that soil type. In the displaying detected soil type interface there is a button to be pressed only if the user wishes to display the suitable crops for that detected soil type. There are two more buttons in the displaying detected soil type interface, one which directs the user to the upload image interface where as the other button terminates the application.

Once the user uploads the image, the image is analyzed and together with the details saved in the server the final result of the soil type or the suitable crop types are displayed in a new interface. There are two buttons in the displaying suitable crops interface, one which directs the user to the set location interface where as the other terminates the application.

3.6 Hosting/ Installation Environment

The portable soil detection system for agriculture based on image processing is hosted in a web server which consists of the database and the model to detect the soil type.

This application is supported by any device that has the android version 4.4 or above.

3.7 Summary

At the beginning of the chapter, the PACT analysis was described followed by the UI design consideration and approaches. The design tools, techniques and templates were discussed in the middle part of the chapter. Furthermore in this chapter we discussed the input, output and dialogue design aspects. Finally the hosting and installation environment was conversed at the latter part of this chapter.

Chapter 4 - Data Management

4.1 Introduction

Data management is a very important feature in our project. We will store main details about the soil types, its attributes and suitable crops for each soil type. Since we are handling our project according to these details, we need to consider about huge amount of data. We need to store our datasets in a powerful database system. To access, store and retrieve the data in more efficiently and accurately from database, it needed some powerful data management techniques. This chapter will show about the data management of our system using EER diagram and how will be database designed.

4.2 Data requirement

Main objective of the data management section is to collect details about soil type, their attributes and store them in the database. Image processing techniques will be applied to the images we get as inputs and compare them with the data that we stored in the database. To get these results precisely, data requirements should be done by very carefully. Here are the data requirements of the project.

- Store the details of the soil types and their attributes.
- Store the details of the suitable crops to be cultivated in each soil type.
- Retrieve, update data of soil and crops.
- Store the details of the locations

4.3 Design tools and techniques

MYSQL Workbench has been selected as our database designing tool. MySQL Workbench is a unified visual tool for database architects, developers, and DBAs. MySQL Workbench provides data modeling, SQL development, and comprehensive administration tools for server configuration, user administration, backup, and much more. MySQL Workbench delivers visual tools for creating, executing, and optimizing SQL queries

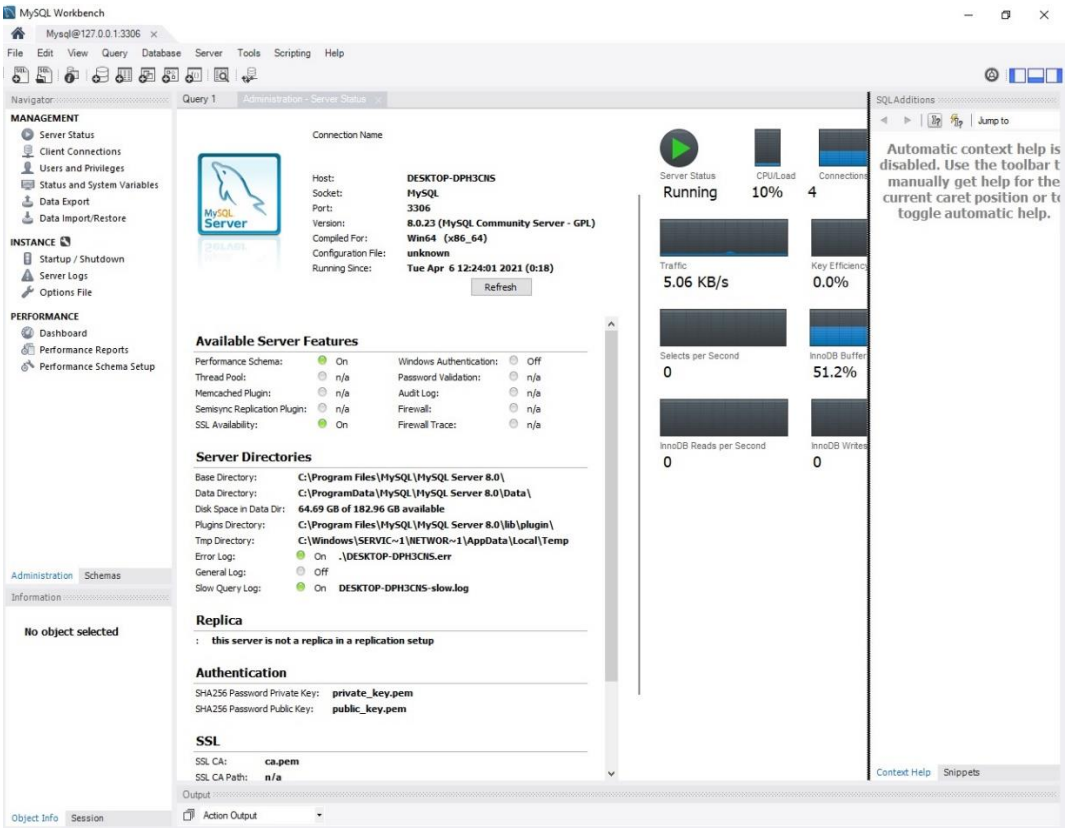
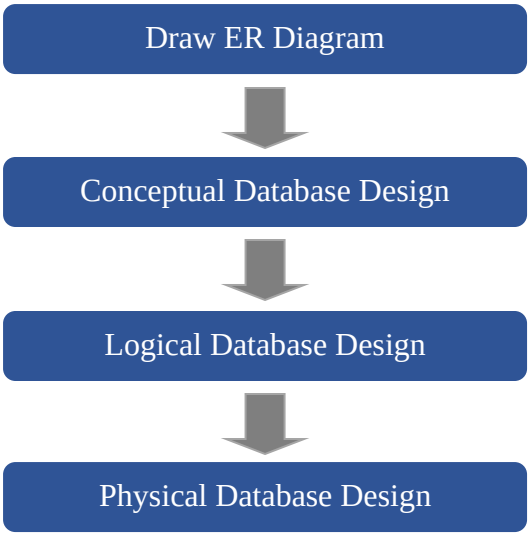


Figure 4-1

MySQL workbench 8.0

Following steps are the design approaches of the proposed database



4.3.1 ER Diagram

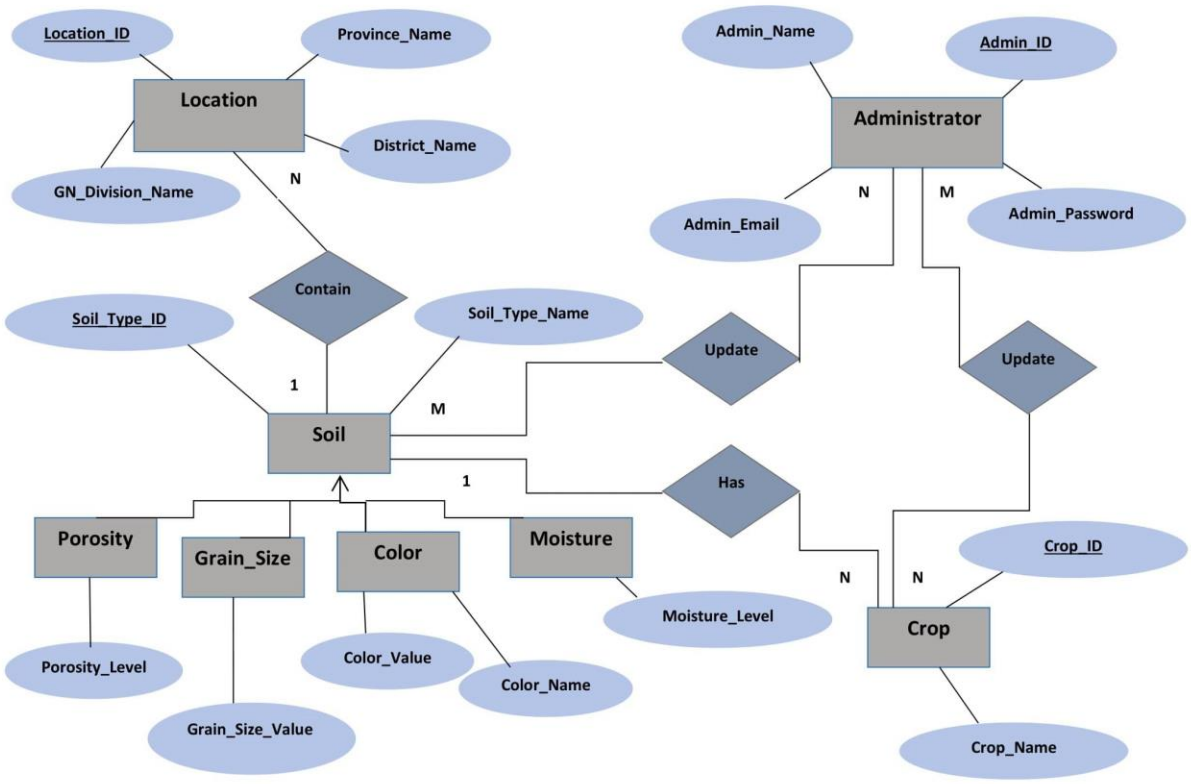


Figure 4-2 ER Diagram

4.3.2 Conceptual database design

The goal at this stage is to design a database that is independent of database software and physical details. The output of this process is a conceptual data model that describes the main data entities, attributes, relationships, and constraints of a given problem domain

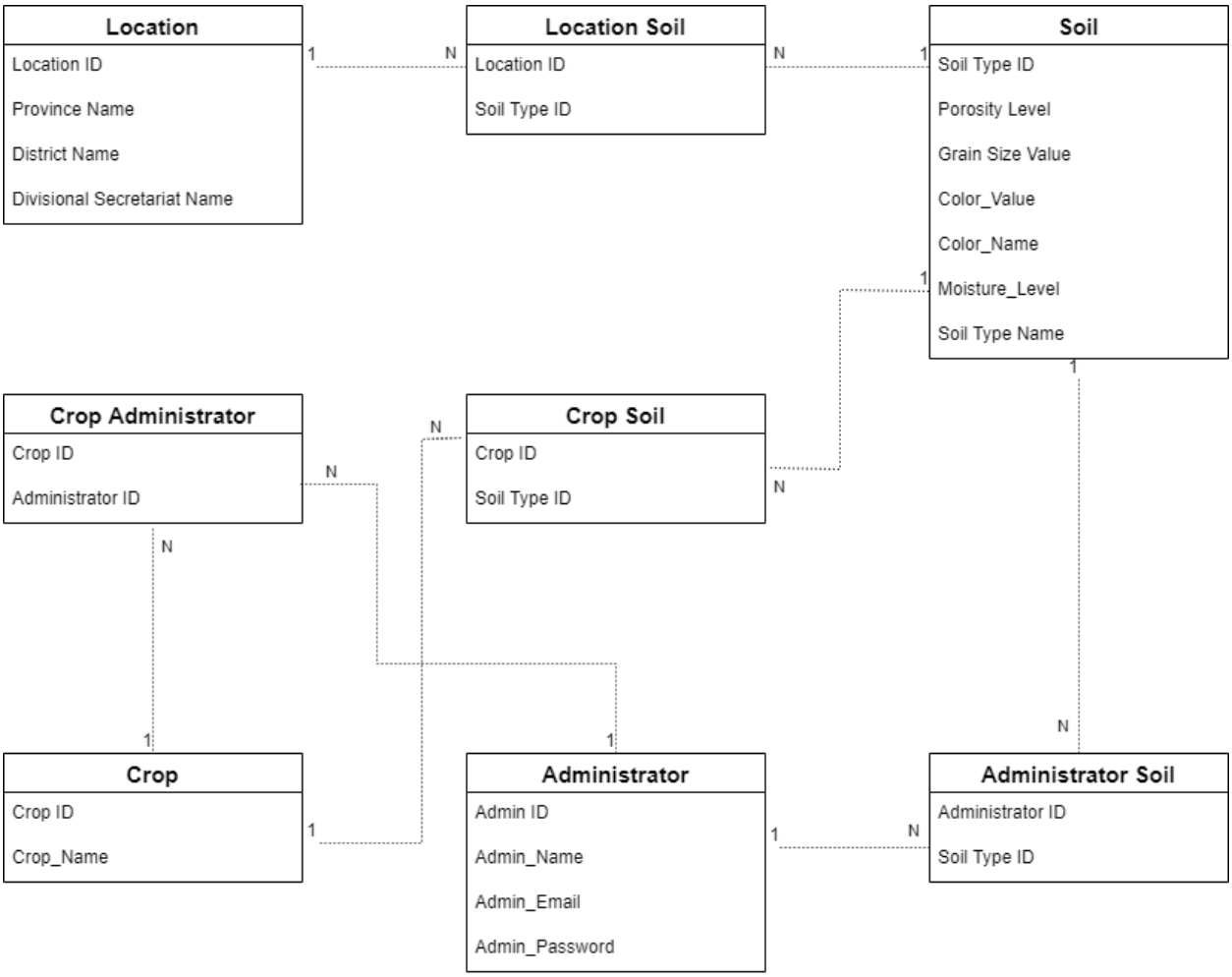


Figure 4-3 Conceptual database design

4.3.3 Logical database design

Logical data models add further information to the conceptual model elements. The process of logical design involves arranging data into a series of logical relationships called entities and attributes.

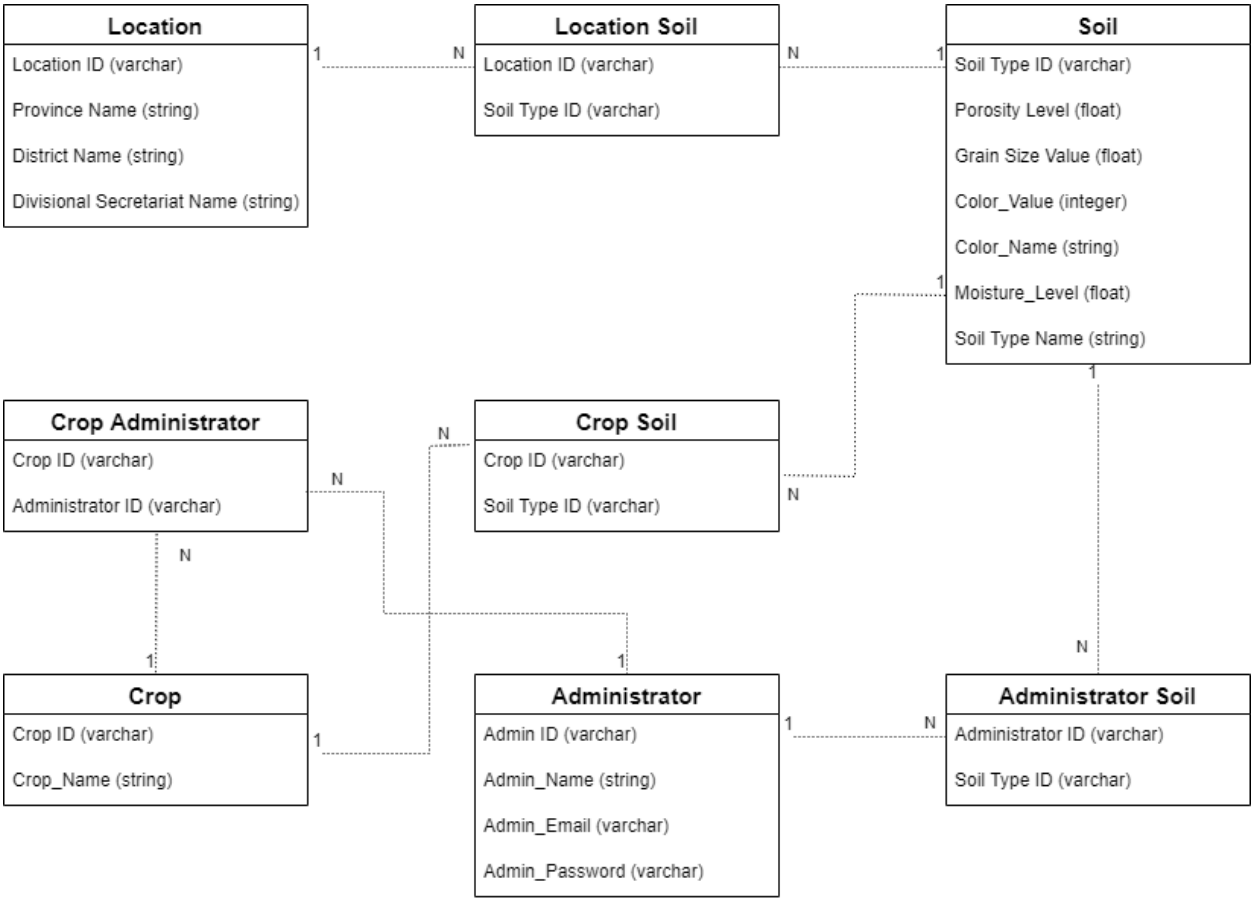


Figure 4-4 Logical database design

4.3.4 Physical database design

Physical database design represents the materialization of a database into an actual system. While logical design can be performed independently of the eventual database platform, many physical database attributes depend on the specifics and semantics of the target DBMS.

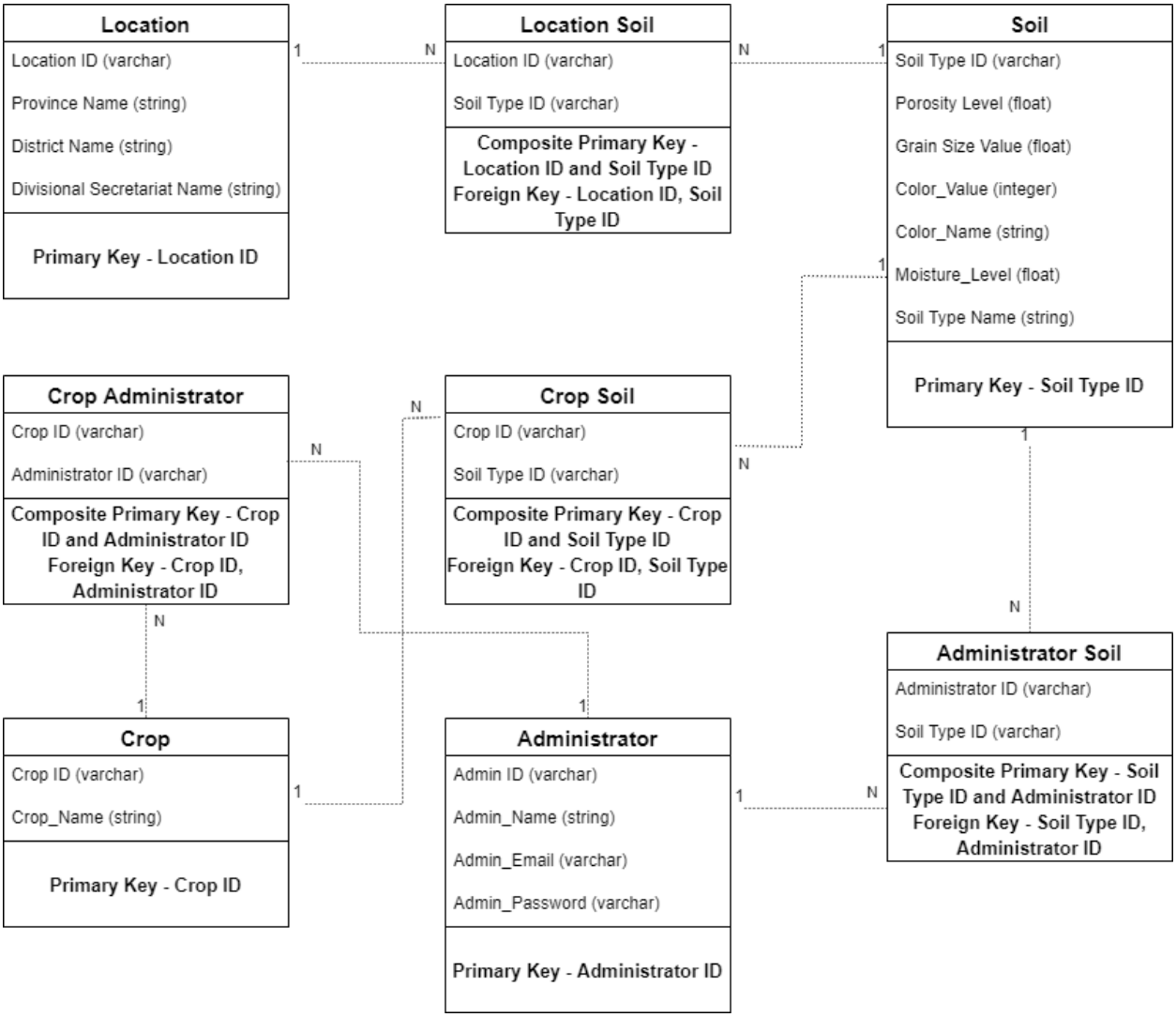


Figure 4-5 Physical database design

4.4 Schema refinement

Schema refinement means polishing tables. It is the last step before considering physical design/tuning with typical workloads. In this process mainly checking tables for redundancies and anomalies. To minimize redundancies and anomalies, normalization was used. It will organize the data in the database.

A normalized database is closer to the relational model because it does not have any dependencies.

Normalization techniques involves,

- Identify functional dependencies
- Avoiding anomalies
- Normalizing multi value attributes

4.5 Security design

Database security encompasses a range of security controls designed to protect the Database

Management System from illegitimate use and malicious threats and attacks. System security design will

protect all the data that we stored in the database.

CHAPTER 5 -RESEARCH DESIGN

The Research design chapter consists primarily of a review of the literature survey carried out on the proposed system. The literature review section provides a summary of the technologies, methodologies involved in previous researches under the project idea. Also, this chapter includes a brief description of the high-level implementation components that the project consists of and the non-functional aspects of the project

5.1 Objective based Literature review

The literature review has done depend on the project objectives. We have categorized the literature review according to the objectives of our project.

5.1.1 Determine the soil type using color

A research done by scientists Pengcheng Han, Daming Dong, Xiande Zhao, Leizi Jiao, Yun Lang had developed a Smartphone-based soil color sensor for soil type classification. In this research they mainly compared and analyzed the roles of the visible spectrum and machine vision adopted in soil classification. They directly used the CMOS device of the mobile phone as the sensor and the flashgun of the phone as the light source. In this method they used peripheral components included external lens, shading devices, and color calibration card, which were assembled on the phone directly. The colors of soil and proofread cards were acquired by the smart phone and converted into RGB signals. With RGB signals, after simple processing, they achieved rapid soil classification. But for image acquisition of machine vision they used D7000 SLR (Single Lens Reflex) kit (Nikon Corporation, Japan) in their experiment. After they captured pictures, the pictures of each soil type were treated with the Matlab7.0 as their data set. By doing more spectral curves they obtained the visible reflectance spectral curve of each soil type. Different types of soils were different in the visible spectrum characteristics. Using these visible spectrum reflection characteristics they managed to detect the soil colors of each soil type. [6]

Within this study the researchers have used the munsell soil chart to categorize the initial images which were collected to create the database. In the proposed research study the munsell chart will be used to categorize the soil types in terms of color. However, as peripheral devices are used to capture the image, when experimenting (Shading bucket, external lenses, calibration card are attached to the smart phone) it is not convenient for a farmer to use these devices when obtaining the images in their cultivating land. Hence the objective of this research is not efficient in the practical aspect. Also in this research visible Spectroscopy method was used to classify the soil types. And also as mentioned in their research paper this method is not suitable for the rapid classification of soil samples.

Shima Ramesh Maniyath, Mr. Ramachandra Hebbar,Dr. S Rama Subramoniam, Akshatha K N, Architha L S had developed a method to detect the soil color using digital image processing. The authors bring some information about the background of the problem and they use the munsell chart to compare the soil colors using image processing. According to this research they had generated a database about the soil colors using soil munsell chart. And this literature shows that

even the comparing soil sample image was not clear which means that it contains noises, that they had used median filter as their noise removing filter and used thresholding segmentation to do the segmentation. And they used K-Nearest Neighbor (KNN) as their soil color classification algorithm. It is a supervised machine learning technique. KNN considers the k nearest cases from an instance and decides which the most frequent class in the set is. The most repeated class is assumed to be the class of that instance (x), KNN system adopts a distance metric in order to determine nearest instance. Various distance metrics can be adopted including Euclidean which is used at this research as it gives the best results.KNN classifier could be taken as a successful technique to detect the soil color because of its simplicity and it is relatively speed. [7]

According to this research study the researchers have used the munsell soil chart to categorize the initial images which were collected to create the database. In the proposed research study the munsell chart will be used to categorize the soil types in terms of color. Also in this study a mobile phone camera was used to take the images which will enhance the practical usability of the objective of this research. A mobile phone camera is used in the proposed system as well to elevate the convenience of the user to use the application. Another plus point which was observed in this research methodology was that the median filter was used here as the best filter to remove the salt and pepper noises. However in order to identify the most suitable crops to be cultivated in a farmers land which is the indirect practical objective of this research cannot be achieved by identifying the color alone. Several other characteristic of soil such as porosity, moisture and grain size also effect the growth of a crop.

FitriUtaminingrum and Ihwanudien Hasan Robbani had proposed an algorithm to detect the soil color using image processing, The Soctect Algorithm which means soil color detection algorithm. This research paper shows that this algorithm is consisting of five main steps, namely, creating database, first filtering process, segmentation, and second filtering process and color matching process. In this algorithm they create a database and they used mode of RGB values to get representation data. A median filter method is used to get the clear image. And they segmented the image by using K-means segmentation method. Furthermore the segmented image will be filtered again by using median filter method. In this method they have used Euclidean distance to match the layers of image soil with the color in the database. They use this algorithm to detect the soil colors of different layers in soil sample. [8]

In this research the researchers have used the munsell soil chart to categorize the initial images which were collected to create the database. In the proposed research study the munsell chart will be used to categorize the soil types in terms of color. Also in this study a mobile phone camera was used to take the images which will enhance the practical usability of the objective of the research. A mobile phone camera is used in the proposed system as well to elevate the convenience of the user to use the application. Another benefit which was observed in this research methodology was that the median filter was used here as the best filter to remove the salt and pepper noises. Nevertheless the result of this technique is highly dependent on the original image, as when obtaining images from different cameras the results will differ and if the images are collected using the same camera a higher accuracy could be reached. Also as there

are two filtering processes this approach is highly time consuming and larger neighborhoods might make the image blurry.

5.1.2 Determine the soil type using porosity

In a research study which was conducted by Shufang Jiang, Yaohu Kang1, Zeqiang Sun the digital image method used for analysis of soil pores was examined. They have used a personal computer and flatbed scanner, along with commercial remote sensing software (ENVI, PCI, ERDAS, etc.) and Geographic Information System software (ArcGIS, ArcView, SuperMap, etc.) to complete this method. They have invented some equations to get the results to complete their process. They had Input thin section images of the samples to be investigated into Remote Sensing software. Using the statistical function in the Remote Sensing software, they obtained a table consisting of statistical parameters. Subsequently, they determined the total porosity of the thin section of each investigation sample by using the spreadsheet column showing the cumulative percentage Digital Number, along with the image porosity threshold from the corresponding output table. They Converted original digital images of all investigation samples to binary images, using the image porosity threshold and the classify function in the Remote Sensing software. Then Input the binary images of all the investigation samples into GIS software, and apply the region-group function to output the attribute table to an excel file, and divided the pores into capillary pores and air pores according to the capillary pore threshold. Finally they calculated the porosity of capillary pores, porosity of the air pores, diameter of any pore and the porosity of any size pore using their predefined equations. [9]

Within this research study the researchers have used the Intrusion method to predetermine the total porosity to create the database. In the intrusion method the total porosity is determined by the volume of the liquid that just fills the open spaces and completely saturates a soil sample, divided by a sample volume. In the proposed research study the Intrusion method will be used to predetermine the total porosity to create the database. However, as external devices are used in the experimental process (flatbed scanner, commercial remote sensing software, and GIS software) it is not convenient for a farmer to use these devices when obtaining the images in their cultivating land. Hence the objective of this research is not efficient in the practical aspect.

5.1.3 Determine the soil type using grain size

The scientists G.H.A. Janaka J. Kumara, KimitoshiHayano and Keita Ogiwara has conducted a research study under the title of Image Analysis Techniques on Evaluation of Particle Size Distribution of Gravel. The researchers have adopted a methodology where the particles were arranged on a transparent sheet without touching or overlapping each other. Then the images were captured using a Nikon D7000 camera. The experimental procedure of this research study

was done in two approaches. In the first approach 100 particles were placed on a white color sheet and in the second approach the particles were placed on a black color sheet. Shadow effect on the gradation curves were examined comparing the results of the first and second approaches. An alternative approach is used in the first approach as well. That is by using a transparent sheet to eliminate shadow effects. Then the images were analyzed by using the ImageJ software. Initially the pixel values are converted to millimeters using a scale factor. Then the images were converted into binary images and image processing techniques such as erode, dilate and fill holes were applied on the images respectively. Area, minor and major axes and shape characteristics such as roundness, perimeter and circularity were measured in the image analysis. The volume of the particles was determined by the results of the image analysis. [5]

In this research the morphological operations such as erosion and dilation was used to extract the image components. These techniques will be used in the proposed research study when determining the grain size of the soil particles. As this research is done by using an experimental apparatus that is not a practical approach for a farmer to determine the grain size of their soil samples. The farmers should consult an image processing expert to find out the grain size of that soil sample. Hence it is not a convenient method. Also in this method the individual gravel particles are arranged in a white and black sheet in order to obtain the photograph to be analyzed. But in most of the soil samples it cannot be done practically. Furthermore in order to identify the most suitable crops to be cultivated in a farmers land which is the indirect practical objective of this research cannot be achieved by identifying the grain size alone. Several other characteristic of soil such as porosity, moisture also effect the growth of a crop.

5.1.4 Determine the soil type using moisture

In a research study which was conducted by Gheorghe C., Dean T.A. and Filip N. under the title of Image processing techniques used in soil moisture analysis the researchers have adopted a methodology by which they have obtained several pictures of a large field located nearby Cluj - Napoca city through a quadcopter and processed them in MATLAB software. The intensity of the original RGB image was determined by using MATLAB codes and that was compared with the black and white image pixel's intensities. In this methodology the researchers have converted the original RGB image to a gray scale image and that image was converted to a high contrast gray scale image. Then the coordinate matrices were obtained from the above images and the intensity level was set on the z axis using image processing algorithms. By applying this algorithm on the MATLAB software the researchers have obtained a surface plot with three axes, x (number of pixels on the length), y (number of pixels on the width), z (number of pixels in the height). From the plotted graph the researchers were able to observe that the number of pixels in z was higher as the color of the image is brighter. Hence it was concluded that when the number of pixels in z is lower that places contain dry soil and when the number of pixels in z is higher the soil is wet as the grass is grown in the moist soil and that is the green color which is giving a higher number of pixels in z. It was also found out that the contrast enhanced black and white images can produce easier to see results at first glance and also the contours of the grass parcels are better defined in this image than in the other two images. [3]

According to this research study it was found out that the black and white images has a pixel distribution and a contour higher than the RGB images. Hence that finding will be used in our research study when applying the image processing techniques. As in this research a drone is used to take the aerial view of the soil it is not practically supporting the agriculture in general as the farmers does not possess a drone to take aerial images and they have to go to an expert to analyze the soil moisture content of their land. This is not a convenient method to the farmer and also is time consuming. Furthermore in order to identify the most suitable crops to be cultivated in a farmers land which is the indirect practical objective of this research that cannot be achieved by identifying the moisture alone. Several other characteristics of soil such as porosity, grain size also effect the growth of a crop.

Yuanjun Zhu, Yungiang Wang, Mingan Shao and Robert Hortan has published a research article under the title of Estimating soil water content from surface digital gray level measurements under visible spectrum. The researchers have adopted a methodology where the soil samples were initially passed through a 2 mm sieve and dried at 108 ^{0}C. The soils were then packed into small cylinders and water saturated. The images were taken using a digital camera. The packed soil samples were allowed to dry. At selected times during drying, the samples were weighed to determine the soil water content and surface images were obtained. Two color checkers of black and white were used in two corners of the image when taking the digital photograph to minimize the effect from the background light. All the images taken were converted to gray scale images and the gray level was determined. It was observed by the researchers that in dry soil the surface gray level is increased and in the wet soil the gray level is very less. For saturated soil the gray level is at its minimum. In this research statistical characteristics of the surface gray levels were described by using different mathematical models and formulas. [4]

In this methodology they have dried a soil sample and has saturated it with water and allowed them to dry and has taken the photographs of it. This approach could be used in the proposed research study as well. However in the proposed research study it was not planned to saturate the sample at once but gradually saturate it by increasingly adding 20 ml of water until it is saturated. These researches have also adopted a technique to minimize the effect of the background light in the images that are taken. Here they have used two black and white checkers. Hence that technique will also be helpful in the proposed research study to minimize the effect of the background light. However in this study the soil sample has to be packed in a cylinder with 3 cm diameter and 1.5 cm deep. This cannot be practically found by the farmers. Also as this study has used a mathematical basis to find the moisture level of soil the farmer has to consult an image processing expert to determine the moisture level of the soil. This is not a convenient and a practical method for a farmer.

5.2 Formalizing High-level implementation components

In our research we plan to train an AI model to suggest the soil type of the uploaded image of the soil sample. For this we use different machine learning techniques which can develop programs to access data and to learn for themselves. One such machine learning algorithms that can be used for in our program is the K-Nearest Neighbours algorithm which is used for classification and regression of the entire data set.

5.3 Data extractions, sample design, test data sets, training data sets

5.3.1 How to extract data

Initial data set was collected from the soil maps of Natural Resources Management Center and by photographing the soil samples taken from different areas of the North Western Province.

5.3.2 Test dataset and training data set

According to our project, the database is updated when a new soil type or crop type is discovered. We planned to use the same data set for training the module and as test data.

5.4 Non-functional aspect

Non-functional requirements deal with the behavior of a system or the technical aspects. The Nonfunctional requirements are the limitations on the functions available by the system which are limitations on timing, limitations on the development process and standards. Followings are the non-functional requirements in this project.

5.4.1 Product requirements

- Usability

  This mobile application is specially designed to the farmers those who are working in outdoors. This mobile application is having a simple, appropriate user interface. So, to make the full use of this device, the farmer no need to have a previous practice of using such a device. The colors were chosen sensibly as the farmer is using the app outdoor.

- Efficiency requirements

  The system is proposed to implement in order to give the responses immediately to the end user. Since the end user is a farmer, the main purpose is to provide the soil type results, without making him waited or confused. The system is providing the results quickly in real time.

- Portable requirements

  Proposed mobile application will be portable and can handle at any place in any device that has android version 4.4(KitKat) or above.

5.4.2 Organizational Requirements

- Delivery requirements

  This proposed system should be delivered to user according to the time period in the project proposal.

- Implementation Requirements

  As a software application the programming language is a very important requirement. Selecting best suitable programming language makes the implementation much easier. As we are working with artificial intelligence python is the most recommended programming language for making these applications.

5.4.3 External Requirements

- Ethical requirements

  Our application requests for the location of the user which is a sensitive data. And thereby that data should not be misused and not be used for any unethical activity.

- Privacy requirements

  The location of the user is confidential. Therefore it is important that this data is not leaked to the outsiders and ensuring security within the application is important. For the above reasons, admission to the database is restricted to authorized persons only.

5.5 Proposed validation methods and measurements

Validation is used in projects to see whether the project has met it's requirements and specifications or the desired outcomes. There are several steps used to validate our project activities.

- Soil sample photos are enhanced and pre processed before model training.
- Outputs are tested with the existing data set /test data.

5.6 Summary

This chapter describes the objective-based literature review of the project and then describes the high-level implementation aspects of the system. Further the non- functional aspect regarding the implementing project. Further in this section it has included the validation methods of the system which is to be implemented.

References

[1]Gheorghe C. ,Deac T.A. , Filip N. ofTechnical University of Cluj - Napoca, Faculty of Automotive, Mechatronics and Mechanics / Romania,"IMAGE PROCESSING TECHNIQUES USED IN SOIL MOISTURE ANALYSIS",2019.

[2]Yuanjun Zhu, Yunqiang Wang, Mingan Shao, and Robert Horton,*Estimating soil water content from surface digital image gray level measurements under visible spectrum,*<https://cdnsciencepub.com/doi/full/10.4141/cjss10054>

[3]G. H. A. Janaka J. Kumara, KimitoshiHayano and Keita Ogiwara of Department of Civil Engineering, Yokohama National University, Japan,"Image Analysis Techniques on Evaluation of Particle Size Distribution of
Gravel",2012.

[4]Pengcheng Han, DamingDong ,Xiande Zhao, Leizi Jiao, Yun Lang,"A smartphone-based soil color sensor: For soil type classification",2016.

[5]Shima Ramesh Maniyath, Mr. Ramachandra Hebbar,Akshatha K N, Architha L S, Dr. S Rama Subramoniam," Soil Color Detection Using Knn Classifier",2018.

[6]FitriUtaminingrum and Ihwanudein Hassan Robbani, "Scotect algorithm: A novel approach for soil color detection using five step algorithm",2016.

[7] Jiang, S., Kang, Y. and Sun, Z., 2009, in IFIP International Federation for Information Processing, Volume 294, Computer and Computing Technologies in Agriculture II, Volume 2, eds. D. Li, Z. Chunjiang, (Boston: Springer), pp. 1029–1038.

[8] Schema refinement - http://blog.dancrisan.com/intro-to-database-systems-schema-refinement-functional-dependencies

[9] Security design - https://looker.com/definitions/database-security

Approval

Signature of team members:

Name	Registration Number	Index Number	Signature
K.G.Rajapaksha	ICT/2016/2017/083	3946	
G.T.S.Perera	ICT/2016/2017/050	3917	
W.M.D.H. Bandara	ICT/2016/2017/008	3884	
J.P.R. Thakshila	ICT/2016/2017/064	3929	
T.L.U. Sudarshana	ICT/2016/2017/086	3948	

Date: 08/04/2021

Recommendation of the supervisors

Name: Mr. N.M.A.P.B. Nilwakke

Department: Department of Computing

Signature: ……..

Name: Ms. Malithi Jayasooriya

Department: Department of Computing

Signature: ……..