LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CNN	Convolutional Neural Network
GPS	Global Positioning System
K-NN	K-Nearest Neighbors
PH	Power of Hydrogen
RGB	Red Green Blue
SRDI	Soil Resources Development Institute
SVM	Support Vector Machine

VGG 16

Visual Geometry Group 16 **xii**

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

Agriculture is the primary source of livelihood for about 58% of the population of India. Continuous efforts have been taken to develop this sector as the whole nation depends on it for food. For thousands of years, we have been practicing agriculture but still, it remained underdeveloped for a long time. After the green revolution, we became self-sufficient and started exporting our surplus to other countries.

Earlier we used to depend completely on monsoon for the cultivation of food grains but now we have constructed dams, canals, tube-wells, and pump-sets. Also, we now have a better variety of fertilizers, pesticides, and seeds, which help us to grow more food in comparison to what we produce during old times. With the advancement of technology, advanced equipment, better irrigation facilities agriculture started improving. Furthermore, our agriculture sector has grown stronger than many countries and we are the largest exporter of many food grains.

In recent years, farmers are suffering financially and are facing many hardships. This is due to various reasons such as urbanisation, globalisation, pollution, water scarcity, less rainfall, low fertility of soil, drastic climatic changes, political and economic reasons, poverty, lack of technological assistance etc.

Addressing their needs through technology is the need of the hour.

Though we have very less to contribute to improvise the natural factors to help agriculture, we have a lot to contribute to this sector through computer science and technology. Internet of Things(IoT), Artificial Intelligence, smart agriculture, Agricultural Engineering, Irrigation Engineering are some of the fields that contributed to the development of agriculture in recent years.

With large scale increase in the availability of data, machine learning, deep learning, big data analytics can help in solving various problems. Machine learning has emerged with big data technologies and high-performance computing to create new opportunities for data intensive science in the multi-disciplinary agritechnologies domain. The works can be categorized as (a) crop management, including applications on yield prediction, disease detection, weed detection crop quality, and species recognition; (b) livestock management, including applications on animal welfare and livestock production; (c) water management; and (d) soil management. By applying machine learning to sensor data, farm management systems are evolving into real time artificial intelligence enabled programs that provide rich recommendations and insights for farmer decision support and action.

There are many ways to suggest crops suitable for a farm land. It can be based on the climate or soil or the crop that produces high profit in that region. We want to suggest crops considering all these factors. We also want soil classification to be done easily with android camera images so that the laboratory tests can be avoided to identify the type.

1.2 PROBLEM STATEMENT

While analysing the various problems faced by farmers,choosing crops for their land appears to be a concerning problem.Crops must be chosen not only based on the soil and climate but also on various other factors like usage of the crop in the particular area, cost, revenue, how much the crop is exported or imported. In this project, our aim is to suggest crops to farmers such that it leads to maximum production and profit.The problem statement is to provide a user friendly application that classifies the soil into four types (Alluvial,Black,Clay,Red) with a simple camera image and suggests the best crops which will give higher yield and profit.

1.3 OBJECTIVES

- To classify soil image into one of the four categories precisely (red, alluvial, black, clay).
- To implement different models and find the best suitable model for soil image classification.
- To suggest crops for a region considering weather and past production and profit rate.
- To give success rate for each crop cultivable in that soil and region

1.4 OVERVIEW OF THESIS

Chapter 2 elaborates on the related work in soil classification and crop suggestion domain. The issues in the previous work are analysed and it has been rectified. The requirements of the system (both functional and non-functional) are identified and specified in Chapter 3. The architecture diagram, use case diagram and activity diagram for the entire project is given in Chapter 4. In chapter 5, all the three modules in the project are discussed elaborately. The implementation of the system, the intermediate outcomes, evaluation are specified in Chapter 6. Results, screenshots and test cases are given in Chapter 7. In chapter 8, conclusion and future work of our project are specified.

CHAPTER 2

LITERATURE SURVEY

2.1 SOIL CLASSIFICATION

Srunitha K, S.Padmavathi created a soil classification model that uses Support Vector Machine based classification.Almost all countries export their products, Countries which export agricultural products depend on soil characteristics.Hence classifying soil, based on their characteristics is very important to reduce the product quantity loss.The nature of soil is influenced by many factors. Some of them are power of hydrogen (PH), Exchangeable sodium percentage, moisture content etc. depending on their amount in soil they show different characteristics and that varies for different region. The manual segmentation and classification methods are time consuming, require efficient people and expensive also .With the emerging of image processing and machine learning we can efficiently classify the soil sample in to groups and hence we can automate the classification process.Soil classification includes steps like image acquisition, image pre processing, feature extraction and classification.

2.1.1 SVM(Support Vector Machine)

SVM models are mainly used for analyzing the data for regression and classification. For a set of training examples it belongs to either one of the two categories, a support vector machine algorithm for training generates a model which tells the new thing falls into which category by a non-probabilistic binary classifier.

The SVM model is the depiction of points in space which are mapped.

Thus, the data of different types are separated by as wide as possible.

2.1.2 BASIC SEGMENTATION METHOD

The segmentation process splits the region of interest from that of non-interest regions.A two class classifier is required for classifying pixels in feature space considering segmentation as a two class problem. Method of segmentation includes,

1) Training data with one or a few images having objects. Traditional segmentation or by manually foreground and background regions are splitted. Pixels in objects are marked using I and I-which produces RGB color histogram. Color values are also marked.

2) Prepare for SVM the training data,

(xi,yi) , + + +l,if Xi€[. xi is -1,if X,EI. +- Xi€[VI , yi= a color vector.

algorithms such as weighted k-Nearest Neighbor (k-NN), Bagged Trees and Gaussian kernel based Support Vector Machines (SVM) are used for soil classification.The method involves two phases: training phase and testing phase. Two datasets are used: Soil dataset and crop dataset. Soil dataset contains class labeled chemical features of soil which include salinity,pH values and iron,magnesium content etc. This system mainly uses three methods namely, Weighted K-NN, Gaussian Kernel based SVM, and Bagged Tree.

2.2.1 WEIGHTED K-NN

It is a refinement of the k-NN classification algorithm.It weighs the contribution of each of the k neighbors according to their distance to the query point, giving greater weight wi to closer neighbors. It makes use of all training examples not just k if weighting is used.The algorithm then becomes a global one.

The only disadvantage is that the algorithm will run more slowly.

2.2.2 SVM

SVM is a supervised machine learning algorithm which works based on the concept of decision planes that defines decision boundaries. A decision boundary separates the objects of one class from the object of another class. Kernel function is used to separate non-linear data by transforming input to a higher dimensional space.

The Gaussian radial basis function kernel is used in this method.

2.2.3 BAGGED TREE

Here they have used a bagged decision tree ensemble classifier which consists of 30 trees. Bagging generates a set of models each trained on a random sampling of the data. The predictions from those models are aggregated to produce the final prediction using averaging.