CHAPTER 1

Introduction

India is one of the world's leading agriculture producers. It has created attention in the fastest growing economy worldwide. As per report of forecasting institutions, India will play a enormous role in the world agriculture market by year 2020. India gets the rank in the top ten Gross Domestic Product (GDP) countries and as per the United Nation's Food and Agriculture Organization (UN FAO) report; its rank is third after China and Japan [Anon (2013)]. The agricultural sector cannot be overlooked because more than 50% of the manpower in India is employed only in this sector from last 30-40 years as per the UN FAO. Figure 1.1 is showing the distribution of manpower involved in agriculture sector worldwide. It is evident that the India is among one of the countries, where 40-84.8% of manpower is involved in agriculture business. Despite the major workforce involvements and huge area of production, the trend in GDP contribution by this sector in India is the decreasing from 23% in 2001 to 17% in 2011 (UN FAO). The agricultural yield is not very much distinctive for the developing countries like India and other and thus giving a scope for research in this area.

1.1 Motivation:

Soybean is among one of the most profitable cash-crops which is grown in many countries, such as the United States of America, China, India, Brazil, and Argentina. The
above mentioned countries contribute more than 90% of global soya production [Database (2013) and Miles et al.(2003)]. As per the report published by united nations food and agriculture organization in 2013, the contribution by Brazil, USA and Argentina makes up more than 82% of global production, while China, India and Paraguay contribute about 13% only. [Database (2013)]

![Figure 1.1 Employment in agriculture, share of total employment (%), 2005-2010](source: FAO http://faostat.fao.org/site/567/default.aspx#anchoraccess at 31/03/14.)

One more reason for giving the so much attention to Soybean crops, as it forms the main part of food and a source of protein for a huge vegetarian population of the country. Soybean is mostly grown on marginal lands under rain-fed circumstances and is cautious as a poor farmer’s crop. The satisfactory quantity of water on soybean harvest is necessary for getting high yields. From establishment to crop, successful irrigation system management is important at each stage of soybean growth. As per the figure-1.2 it can be seen that India ranks on top among the 20 countries which are having the best potential for the irrigation worldwide so definitely no problem with the point of view of irrigation.
India is leading across worldwide in irrigation, but soybean yield is still constant and low compared to other countries [Karnz and Specht (2012)].

![Irrigation potential](image)

*Figure 1.2 Irrigation potential, top 20 countries (2012)*


Soya beans have proven their importance and commercial nature due to its use in industries to produce the products like sauces, salads, mayonnaises, chocolates, baby food, and baked goods. Soya beans are not only used for human consumption, but are very useful for animals also. Today, soya beans are also used as an alternative fuel called bio-diesel [Maurya (2011)]. Soybean meal gets the second rank in the Indian agriculture export market [IFPRI (2013) and Anon (2007)]. Compare to other agriculture produces, soybean meal is stable on world agriculture market and it contributes 6% of share [IFPRI (2013) & Anon (2007)].

The reported data show that there is a huge regional disparity in soya production. A comparative analysis has been performed among five major soya producing countries and made known that the total year on year production is increasing in most of the countries but the yield remains almost constant [Singh (2010)]. Indirectly, the cost of production is increasing in the form of required manpower, land and other necessary
resources. The soya yield is significantly higher in countries like the USA, Brazil, and Argentina, comparatively low in China, and the lowest in India as per the report. On the basis of reported data there are a few obvious questions to be answered: “Why there is a huge gap in soya yield among various countries?” and “What is the reason for low yield? Both the questions can be answered by considering two factors; namely, technological usage and environmental conditions. The following figure-1.3 (a)-(c) are showing variations in the area of production (Ha), production (Tones) and yield (Hg/Ha) against the respective years between 2001 and 2012. An appropriate technological application can certainly help in controlling bugs and infections, which are indirectly caused by the environmental conditions. The report, published by soybean processors association (SOPA) in 2010 is clearly showing that the production area for soybean in India has dramatically increased from 608000 Ha in 1980 to 12.6 M-Ha in 2013. A massive increase in the area of cultivation is the main reason for increase in gross production, whereas the yield remains almost constant. The soya bean production, area and yield against the respective years as per the report by SOPA have been shown in figure 1.4
Figure 1.3: Year on year data between 2001 and 2010 (a): Area of soybean cultivation (b): Soybean production, and (c): Yield of soybean cultivation (Data Source: Food and Agriculture organization of United Nations, http://faostat.fao.org/site/567/default.aspx#anchor)
Figure 1.4 Year on year Data Between Year 1980 to 2013 in Indian Soybean Cultivation, (a) Area of Soybean cultivation in India, (b) Production of Soybean cultivation in India, and (c) Yield of Soybean cultivation in India. (Data Source: Food and Agriculture Organization of United Nations, Retrieve at 01/07/2014 http://faostat.fao.org/site/567/default.aspx#ancor)
The annual increment in the cost of production is due to the increase in the area of [Anon (2011)] soybean cultivation business.

The main reasons for low yield as reported by Shrivastava et al. (2014) are plant infections, bug or insect attacks, environmental conditions and a lack of technological applications in agriculture sectors and of course the necessary knowledge among our farmers. The knowledge among the farmers can be reinforced with the help of information technology and associated tools & techniques in various sectors of agriculture like watershed management, natural disaster management, and bug & disease management to enhance the productivity.

The commercial farmers in developed countries are able to handle difficult situations up in the best way with the high input cost due to the increasing area of production, manpower and others by the application of proper knowledge base developed with the latest technology, whereas it is not so in the developing countries like India.

This research is targeted to fill the gap between commercial farmers in developed and developing countries and to provide with the latest technological assistance to farmers especially for soybean cultivation.

Computer Technology has a key role to play in all facets of Indian agriculture. Nowadays there are lots of technologies like remote sensing, geo-informatics, wireless sensor networks, neural networks and digital image processing etc. which are providing the support to Indian agriculture and farming sector. The latest technological application in this most deprived sector can help the Indian farmers for improving the efficiency and productivity of agriculture. The technology may provide the E-powering tool for Indian farmers in decision making, who lives in the rural areas. Image processing plays a major role in the Indian agriculture sector.
Mobile telephony is among one of the technological choices to be adopted to provide the technical knowledge and support to our farmers. The appropriateness of this technology can be established with the help of data published by UN FAO. The figure-1.5 is giving the idea of mobile phone users' distribution worldwide. Again the India is among the top ranked countries using the mobile telephony. So the mobile phones may establish its utility as the best prospective sensors and processors to be used by the Indian farmers.

![Image of world map with mobile phone usage data]

**Figure 1.5 Access to mobile cellular phones subscriptions, annual growth (%) 2000-2010**


A mobile telephone with a reasonably good camera becoming affordable with low price, size and weight. With the help of captured video and images it may be provided the computer based demonstration and knowledge to the rural farmers. It provides the necessary knowledge about the disease, insects, plant stress, fertilizers and growth of the agriculture plants. Computer technology provides a various type of device to provide the fastest and accurate necessary knowledge to rural farmers like output devices, software, networking, transmission media and mobile technology, etc. Mobile technology provides the information sharing services by the SMS, like a market price of the crop, information
about weather, information about the yield, information about the loan and insurance policies, information about the storage of grains and many more. [Anon (2011)]

In international scenario World Trade Organization (WTO) also provide the data awareness, decision support system, information about the new techniques, farming technologies, quality of seeds and fertilizers etc. for the rural Indian farmers.

The requirement of mobile phone based application for targeting the Indian farmer specific requirements and needs in rural India which depends on agriculture sector. This requirement provides a prospect chance to research community to study for agro-economic development of Indian rural areas. The mobile application solutions based on Indian farmers needs and demands are a major challenge for technical and social research community with rural agro-economic constraints. [Rao and Sonar (2013) & Mittal (2011)]

As per report of Telecom Regularity Authority of India (TRAI) show the wireless subscribers in the urban areas declined from 556.99 million at the end of February, 2014 to 555.26 million at the end of March, 2014 whereas, subscribers in rural areas increased from 374.96 million to 377.73 million during the same period.

India has a large number of universities and institutions who plays a critical role to design and analysis of data and necessary applications to the rural farmers. These facilities to achieving the task with control process and generate the quality of the result. Some institutes are already developed the application based system to easily demonstrate the necessary information to huge number of audience in rural India [Jiggins et al.(2007)].

Various researchers have been applying digital image processing methods and pattern recognition techniques for agriculture applications, namely: weed-control, foliar infection identification, and vegetable and fruit grading [Lee et al. (2010) & Sankaran (2010)]. Image processing techniques can be used to detect the plant disease at an early
stage with the help of captured images by any imaging device. An expert system can be
designed to provide the best care at the early stages with the help of type of the disease
present and its severity level. In developing countries most of the farmers identify the
disease on the basis of their experience, but some time, even expert farmers and plant
pathologist are not able to recognize the exact disease present in the agricultural product.
This problem can be solved using a mobile based expert system in which the farmer only
needs to capture the image of the diseased portion of the leaf and on the basis of visual
features the developed system will identify its type and severity.

Various image processing techniques were used in agriculture like fruits and
vegetable recognition, weed and soil segmentation, disease quantification, disease
classification, disease identification. The need of the rural farmers is the exact
information about the disease, seeds, fertilizer, insects and cures [Pydipati (2004)].

Factor affecting the cultivation can be diagnosed by different techniques like
Hyper Spectral Remote Sensing, Thermograph techniques, Machine vision and learning,
Spectral radiometer, RFDI, and Digital image processing.
The following subsections presenting the introduction about the various steps involved in
a general pattern recognition problem. The work reported in this thesis is considered as a
special case of pattern recognition along with the expert system.

1.2 Pattern Recognition System:

The general pattern recognition process has been represented via figure 1.6. To
understand the difficulties while designing such systems, it is required to understand the
functions / objectives / problem which can be performed / archives / solved by each
subsection [Duda et al. (2000)].
(a) **Input:** Now days the pattern recognition is applied to solve, the variety of problems from diversified areas. The value of the input parameter/type is responsible to decide.

(b) **Feature Extraction:** The purpose of any feature extraction is to characterize an object to be recognized while having less intra class variation and high inter-class variation. The best feature extraction method would generate a representation which makes the job of classification very trivial. On the other hand the best classifier may not require the sophisticated feature extractor. Sometimes the generated features should be quite robust against various conditions. So the extracted features must be invariant to translation, rotation, scale, flip, occlusion, projective distortion and deformation etc.

(c) **Classification:** The task assigned to classify is to assign the object category by using the feature vectors generated by the extraction. The abstraction provided by the feature vector representation of the input data enables the development of the largely domain independent theory of classification if the variation among the feature vector of the same class are more relative to the different class objected then it is very difficult to classify correctly.

(d) **Post Processing:** The purpose of the post processor is to use the output of the classifier to decide as the recommended action. The performance of classification measured on the basis of error rate. The best classifier is expected to show minimum error rate the post processor improves the performance of the system either by exploiting some contextual information on multiple classifier fusion techniques.
Figure 1.6 Pattern Recognition Systems
1.3 Design Cycle:

The figure 1.7 represents the design cycle of a typical pattern recognition system.

(a) **Data Collection**: The largest part of the cost designing a pattern recognition system may be accounted for data collection. Lots of data are required to assure the performance of the system. The evaluation result may require for repetition of various steps in order to get the optimum performance.

(b) **Feature and model choice**: To choose the most distinguishing features is the most critical design step. The feature and model choice characterize the problem domain. While the feature and model selection the prior knowledge plays an important role in designing of the classifier by suggesting the most promising feature.

(c) **Evaluation**: Evaluation is important to measure the performance and identifying the need for improvements if performance is not up to the mark.

(d) **Learning Methods**: A pattern recognition system that gathers information from training sample uses learning. Learning process can be divided into three categories, namely a) supervised b) unsupervised c) Reinforcement learning. Supervised learning involves the class or cost for each pattern in the training set itself. Whereas in unsupervised learning or clustering no explicit teacher or expert and the system terms the cluster or natural groups of input patterns. Reinforcement learning is a special case of supervised learning where only binary feedback passes to classifier to minimize the error.
Figure 1.7 Design Cycle
As a part of thesis research a detailed study is developed on the methods to identify and classify the soybean disease using the digital image processing techniques. Six different types of disease class namely soybean rust, downy mildew, brown spot, bacterial blight, frog eye, sudden death syndrome were investigated using digital image processing method to identify the plant foliar diseases, their feature extraction [Stark and Bowyer (1996), Tveten (1998) & Watanabe (1969)].

1.4 Thesis Organization

Chapter.1 gives the introduction covering brief account of Soybean, contribution of soybean in Indian GDP, Problems and difficulties in Indian farmer, the contribution of information technology in agriculture, detailed data analysis of various factors like Soybean yield, area, production, mobile consumers, detailed study of Food and Agriculture organization report which is published by the United Nations. The various sensing techniques for disease identification and classification are discussed and their metrics and measurement are explained. In chapter 2 practical and theoretical concepts of disease detection and classification method are discussed. The limitation and difficulties of classification and detection method are explained. Brief analysis of six soybean disease and it behaviour with cures is given. At last the four major problems are identified.

Chapter.3 deals with the experimental layout of automatic disease detection. The main motive of this chapter is to find out the infected lesion in soybean leaves. Infected lesion is found with the help of image pre-processing, background separation and color filtering operation. The color filtering result is compared to K-mean clustering and Otsu’s method for validation of results. The remaining problems of this chapter, namely classification and identification of diseases are explained in chapter no-4. In chapter. 4 the classification of disease is discussed and developed. A ST-NDCT feature vector for classification of
disease is also developed. Validity of result is compared with histogram analysis and clustering method. Linear discriminate analysis based classification method is used for identification of disease and the results obtained are quite responsive.

Chapter 5 deal with the classification and retrieval based method. The result is compared by various classifiers and descriptors. Chapter 6 consists of discussion about the quantification of disease and development of various new parameters for calculation. In last chapter 7 conclusions and future scope of this research are enumerated in detail.