Chapter 1

INTRODUCTION

Recent developments in Information Technology modernize many disciplines of health care especially Biomedicine. Sophisticated methods have been proposed to automatically extract, useful information from radiology images leading to the discovery of new knowledge. The accuracy of any diagnosis method using medical imaging technologies depends on the quality of medical images and expertise of radiologist [1]. Computer Aided Diagnosis (CAD) aims the identification and localization of abnormalities at an early stage, which prevents the further spread of abnormality with the help of proper clinical management. The work in this dissertation consists of two phases, phase 1- the classification of mammogram and placental images, phase 2- their retrieval. These two phases are preceded by preprocessing and feature extraction methods. The architecture is given in Fig.1.1. This study is a novel approach in placental grading and the contributions in digital mammogram analysis, having scope for further research. This chapter gives an introduction of the problem domain, challenges, motivation, significance and major contributions of the work.

Breast cancer is the second major cause of cancer death in women [2]. It affects the health and lives of millions and millions of women world over. In recent years we notice a rapid growth in the number of breast cancer
patients in all countries irrespective of development. Recent study by ICMR (Indian Council for Medical Research) says that one in 22 women in India is at the risk of breast cancer. The number of breast cancer cases reported increases by one in every 2 minutes [3]. Most of the breast cancer cases are detected only at advanced stages. In the rural areas of the developing and under developed countries women are unaware of the fact that breast cancer is fully curable if detected at an initial stage. Use of Computer Aided Diagnosis (CAD) helps in early detection of breast cancer. This study aims at the automated detection of breast cancer using techniques in Machine Learning and image retrieval.

This study also focuses on the classification and retrieval of similar grade placental images. Placenta connects the growing foetus to the uterine wall and allows nutrient intake, waste elimination, and gas exchange via mother's blood supply [4]. The normal degenerative processes in placenta result in many subtle changes. One such change is the presence of calcification. Placental development begins by around 4 to 5 weeks of gestation. According to Grannum et al. [5] placenta can be grouped into 4 grades, grade 0 to grade 3. The different grades are observed from late first trimester to 39 weeks of gestational period. This work also analyses the characteristics of placenta during this gestational period. As the accuracy of manual grading depends on several subjective factors, automated grading becomes important.
1.1. Challenges

- Medical images cannot be precisely segmented due to their low contrast and high noise content.
- Image cross sections of objects lack clear shape and boundary.
- Biovariability exist for most of the anatomical parts.
- Not many techniques are available to deal with the semantic gap and sensory gap [6].
  - Semantic gap: is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation [7].
  - Sensory gap- is the gap between the object in the world and the information in a (computational) description derived from a recording of that scene.
1.2. Motivation of the Work

Recently, in Medical Imaging, many development oriented studies have been made by scientists to assist radiologists. Early detection of cancerous lesions play a vital role in the diagnostic process and is important for the complete cure of breast cancer [2]. Prevention of breast cancer is difficult, but if detected at an initial stage it can save the life of thousands of patients. The different methods that exist for diagnosing breast cancer are mammogram, ultrasound, Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Cost effectiveness, minimum and optimum radiation, early identification of abnormalities are the reasons for the wide usage of mammogram in breast cancer screening. The recent decrease in mortality rate shows the importance of early detection techniques. Factors like difficulty in identifying suspicious region, position of cancer tissue, the large volume of mammograms given to each radiologist and the repeated nature of work adversely affect the correct interpretation of mammograms. In certain cases, superimposed tissues cause obscure cancerous lesions. To overcome these problems Computer Aided Diagnosis (CAD) techniques can be used [8].

The advances in healthcare over decades have resulted in the development of various computerized methods and tools to support foetal monitoring. Ultrasound is the ideal and most widely used tool by clinicians to capture foetal images. Obstetrics ultrasound examination analyses the anatomy, growth, lung maturity and placental maturity of the foetus. If growth parameters are less compared to gestational age, and if placental maturity is more it indicates an increased probability of IUGR (Intrauterine Growth Restriction). Placental grading can diagnose intrauterine growth restriction. Also grading is an alternate way to predict gestational age and
lung maturity [9]. Placenta appears to be different when the overlapping tissues are more. This makes the manual grading difficult. Moreover the diagnosis result varies among examiners [10]. As the accuracy of manual grading depends on several subjective factors, automated grading is important. This study also focuses the automated classification of placenta into different grades and their retrieval.

The possibility for assisted diagnosis, interpretation and decision making is motivated by factors such as time constraints on readers, disparities between readers based on perceptual errors, lack of training and fatigue. Considerable inter-observer variation has been documented in number of studies [11]. This distinction results partly from the complexity of processing the immense collection of knowledge needed to interpret image findings. Access to appropriate information is a basic necessity in medical field especially in diagnosis. The rapid growth in the quantity, easy availability and accessibility of medical records motivate research into automated image retrieval. However, apart from conventional algorithms used in image retrieval process, the proposed work retrieves required images with the help of multiple image queries using logical operators, AND, OR and NOT.

1.3. Significance of the Work

- Automated classification of abnormalities helps radiologists for quick decision making.
- Screening and early detection, together with improved therapy have resulted in a striking improvement in survival.
• Identification of cancerous lesions which may be missed due to limitations in human eye/brain visual system and the occurrence of vast number of normal cases in screening programs.

• In placental sonograms, diagnostic results vary with different examiners and machine conditions. It is very important to overcome the variability in manual grading to arrive at correct conclusions.

• Repeated scanning during pregnancy increases the volume of data, which highlights the need for automation.

• Content Based Image Retrieval (CBIR) system overcomes difficulties such as manual annotation, subjectivity, language dependency and incompleteness that exist in traditional text based search engines.

• CBMIR (Content Based Medical Image Retrieval) systems assist radiologists in diagnosis, by learning from prior known cases.

• Easy analysis of disease specific information among patients using different modalities is possible.

1.4. Objectives of the Work

The primary focus of this research work is to design and develop methods and algorithms for improving the performance of classification and content based retrieval of digital mammograms and placental sonograms. The following objectives are set to accomplish this.

• Explore the use of classical as well as new Machine Learning techniques in the study of breast cancer and placental maturity.

• Identify the major challenges in the automated analysis of digital mammograms and placental sonograms and solve these issues.
• Develop new preprocessing techniques for improving the quality of digital mammograms and placental sonograms.

• Identification and extraction of relevant features for improving the accuracy in classification and retrieval.

• Overcome difficulties in manual grading of placental sonogram through automated grading.

• Develop new techniques in Content Based Medical Image Retrieval.

1.5 Contributions

1.5.1 Technical contributions

• Development of a new linear filter based on local binary pattern for removing speckle noises in ultrasound images.

• Collected ultrasound images of placenta and developed database using it.

• Use of hybrid feature extraction methods for the study of mammogram & placental images gives good result in classification and retrieval.

• Automated classification of placental images helps to overcome the variability in manual grading.

• Introduction of a new classifier, Extreme Learning Machines for the classification of placental sonograms.

• Design of multiple image query system for expressing the user’s requirement in a better way in retrieving similar images.

• Use of Content Based Image Retrieval techniques for retrieving similar grade placenta images.
1.5.2 Social Contributions

- The proposed work is a contribution in the health care especially in the growth of a healthy foetus and for the general health and well being of women.
- The contribution of the study helps in foetal monitoring and provides medical assistance if required.
- Computer Aided Diagnosis using new algorithms avoids inter/intra variability in diagnosis, and helps the radiologists to improve the accuracy and reliability of their diagnosis.
- A Primary mammogram screening can be conducted in rural areas. An automated system can easily screen the huge volume of data which can be send for further diagnosis if required.

1.6. Chapter Summary

This work is an integration of Computer Science and Medical Science, and highly contributes to the specified problem domains. This chapter gives a gist about the thesis, system framework, challenges faced, motivation for this work and major contributions. In the coming chapters more detailed technical explanation and experimental analysis are given.