CHAPTER 1 - INTRODUCTION

The modern computation is carried in the different applications from space analysis to agriculture process. The medical application is one among the major field of computing and data process is used for clinical decision making. The Medical image analysis is used to detect the diseases as well as the patient’s curability level using the medical treatments.

Many research processes are carried out in the integration of computing and the medical application is for the betterment of the human health. Most of the medical application which is used with the computations is aided to the practitioners to take the decision on their drug recommendations and the identification of the diseases [R.Agrawal et al 1993]. Computing and knowledge processing applications improve the living condition in many areas and many ways. Medical science is one among the potential domain where computing process is used for the enhancement of human lifestyle. The computing technologies are well developed, as well as the medical science also using the technology but they determine suitable ways to adopt the technology and the processing depends on the requirement of the medical applications that are challenging task for the researchers.

The computation process is used in the Medical science field by experts as a tool, but the tools are developed by the computer experts with various technical procedures. The technical experts derive the concept and integrate the same in various medical science applications [D.Brazokovic et al 1993]. The Medical field computational process is used for data and images analysis. The research process involves analysis, design of solution and implement the identified existing algorithms in medical science .This analysis process involves to adopt the latest technology such as data mining [H.Li et al 1995, Maria-Luiza et
The data mining process involves various techniques such as data cubic representation, classification, clustering and decision support system [Prather J.C et al 1997, J.Han et al 2000, T. Wang et al 1998, and Jiawei Han et al 2001]. This research deals with the determination of cancer stage using medical image analysis using univariate analysis.

1.1 OVERVIEW OF THE PROJECT

The data mining approach is to determine the level of breast cancer and its level using equal interval value approach in the image analysis process. Breast cancer is a malignant growth that begins in the tissue of the breast. This can be identified using digital mammograms. The digital Mammogram determination and analysis process are strengthened using data mining tools such as clustering, classification which leads to prevention from breast cancer and other health care applications. This research is initiated and determined in the early stage detection of breast cancer using Univariate classification from the digital mammography using multi view analysis.

The Univariate equal interval classification is used to identify the stage of a mammogram using multiple view of the image. The classified mammogram digital numbers [index value] are used to adopt common value in different image layer to determine the stage and suggest the level of prevention of disease using a common index value of digital mammogram. The identification process of mammogram techniques such as density, type and stage specification to be strengthened using the algorithms to prevent the impacts.

The trends of computation mostly used for data and image analysis process. The data representation, analysis and predication are implemented with data mining tools. Identify and implement a data mining tool for classification method. It is used to determine features
from an image analysis process relevant to the scope of the domain. Identify a suitable image processing techniques to aid classification process. The index value classification is used to identify the stage of a mammogram value. The factors involved in the cancer (Breast cancer) symptoms and identification process with latest determination technology form a major study.

1.2 SCOPE

The medical image analysis process is used for various prediction and pharmaceutical analysis of medical applications. The medical applications are focused for the women issues. The breast cancer for women is identified one among the potential medical applications. The cancer detection and its status is a challenging process with easy way of calculation. Medical tools and the user approach concepts and techniques are developed by the researchers. The medical application is a process in which the captured image provides the unique result with limited preprocess approach. The preprocessed images are manipulated till the identification of suitable image and the result. But all the probability geometric approach on the preprocessing image are not considered for preprocess and the deterministic process. Therefore the researcher fixed the scope to identify the possible geometric preprocess model for the obtained digital mammographic image and simple computation to determine the level of cancer with the following objectives

- **Study the computation process to determine the clinical decision support system and to determine the level of cancer.**

- **Identification algorithmic approach using data mining classification and clustering algorithm for the preprocessed digital mammographic image**
To identify the level of cancer using geometrical preprocess and simple computation algorithms

The computation process of digital image processing for medical applications

Classify the image and its impact level of the tissues via computing density and the occurrence

Role of density computation from the digital values of pre processed geometrical digital image

Classify the mammographic image using different range process in preprocessed image and calculate the common affected level of cancer

Ensure the identified cancer level using different captured image and confirm the level or stage of cancer

The occurrence value and the level of density is used to identify the stage of breast cancer from the preprocessed geometrical image of the digital values

The above specified objectives are derived from the following explained motivation

1.3 MOTIVATION

As per the preliminary student in the field of medical science, women issues and possible computation, it is understood from Indian Council of Medical Research (ICMR) report, majority of women suffered from breast cancer in India. The death rate is also high. The reason is unavailable technology, interpretation, un awareness and not able to detect at the proper time. Indian Council of Medical Research (ICMR) is reported that, In India’. Breast cancer is becoming the number one cancer in women and it pushes the cervical cancer
to the second spot. In rural areas many cases go unnoticed when compared with metros. It is also reported that one in 22 women in India is suffering from breast cancer. But in America one in eight is suffering. Studies have indicated that death rate dramatically can be decreased if the breast cancer is detected at an earlier stage. After the identification of medical issues such as cancer, we continued to observe the various latest technical issues that are used to identify the disease.

Mammography is one among the method of choice for early detection of breast cancer in women. Abnormalities in mammogram can be varied by breast tissue densities. The interpretation of mammogram is also very difficult to the radiologist. During screening routine, the radiologist missing 10% – 30% of breast Cancer [Kononeko, I.Kukar, 1995]. After viewing the causes and the latest technical application, We are motivated to do some contribution in the field of Medical Application(cancer) using the latest technology in the specialized area such as data mining to increase the efficiency to the betterment of women community. As a result of motivational factors, this research is executed.

We are interested to integrate and develop some social contributed application in the field of women related medical issues with the help of computational algorithms. In this research, several image properties are applied to the image which gives many views of color images. Based on these images the density level on digital mammogram is to be identified using Univariate layer based approach. Studies show that the images taken for breast cancer detection uses only grayscale image. In our research, using the high density pixels, accumulative mathematical results can be calculated to find the stage of the breast cancer.
1.4 METHODOLOGY

The research methods are starting from the mammographic image and proceeds with the determination and verification of stage of cancer. The collected digital mammogram image is processed and converts the pixel value into corresponding Digital Numbers or Index Values. The digital Mammogram values are classified according to the Univariate using multi view analysis. According to clustering and classification process, the density level of digital Mammogram area is identified.

The clustering analyses are implemented with direct image and geometrical conversion method. In the geometrical conversion method the image is processed with ascending red16,Auxctq16, BWlnVLog16, BW Parabolic16, correction16,Cyclic16, Descending red16, Design16,Grayscale16,Hot body16, Hot metal16, Isocount16, Heart16, Rainbow16, Red16, Spectrum, Parathyroid16 and Warm metal 16 attributes. The attribute names 16 represents for the bit process and the attribute represented for the preprocessed image. The captured mammographic image is converted and stored in the file as 16 bit representation therefore all the analysis is carried out in 16 bit process. In the rotational property we study Mirror +90, Mirror _90, +90.-90, magnify + and Magnify – properties.

After processing all the above methods, the average and frequency of the image is computed to determine the stage level. The stage of cancer is processed and the prevention possibilities could be recommended from the practitioners.
1.5 ORIGIN OF THE RESEARCH

There are several imaging techniques for examination of the breast, including magnetic resonance imaging, ultrasound imaging, and X-ray imaging. Mammography is a specific type of imaging that uses a low-dose X-ray system to examine the breast, and is currently the most effective method for detection of breast cancer before it becomes clinically palpable [K. H. Ng, M. Muttarak, 2003]. Mammography offers high-quality images at a low radiation dose, and is currently the only widely accepted imaging method used for routine breast cancer screening. Current guidelines of the American Cancer Society (ACS) recommend that women aged 40–49 years have a routine mammogram every one to two years, with the first beginning at age 40 [C. Lewis, 1999].

Currently, there are two types of mammography [D. Gur, 2007, E. D. Pisano et al 2007]: one is film mammography and the other is digital mammography. In film mammography, the image is created directly on film, whereas digital mammography takes an electronic image of the breast and stores it directly on a computer [Online, 56, 57]. Although both types of mammography have their own advantages and disadvantages, digital mammography has some potential advantages over film mammography. Compared to digital mammography, screen film mammography has some limitations, which include [W. Yang, 2006]:

1) limited range of X-ray exposure;

2) Image contrast cannot be altered after the image is obtained;

3) The film acts as the detector, display, and archival medium; and

4) Film processing is slow and introduces artifacts.
All of these limitations have pushed researchers further to develop advanced techniques for digital mammography. Digital mammography is overcoming and will continue to overcome the limitations of film mammography described before, and will have the following potential advantages [W. Yang, 2006]:

1) Wider dynamic range and lower noise;

2) Improved image contrast;

3) Enhanced image quality; and

4) Lower X-ray dose.

Although digital mammography has many potential advantages over traditional film mammography, clinical trials show that [E.D.Pisano et al 2005] the overall diagnostic accuracy levels of current digital and film mammography are similar when used in breast cancer screening. However, digital mammography may be more effective than screen-film mammography for certain women [D.Spurgeon, 2005, M.Del et al 2007]. For example, Spurgeon [D.Spurgeon, 2005] showed that digital mammography depicts more tumors than screen-film mammography, especially lesions seen as micro calcifications (MCs). Pisano et al [E. D. Pisano et al 2005] showed that digital mammography is more accurate in women under the age of 50, women with radiographically dense breasts, and premenopausal women.

There are two types of examinations performed using mammography: screening mammography and diagnostic mammography. Screening mammography is performed to detect breast cancer in an asymptomatic population [NCI Cancer Fact Sheets (2007)]. Screening mammography generally consists of four views, with two views of each breast: the cranio caudal (CC) view and the mediolateral oblique (MLO) view.
The diagnostic mammography is to examine a patient who has already demonstrated abnormal clinical findings, such as a breast lump [NCI Cancer Fact Sheets (2007)]. Similar to screening mammography, each breast examined using diagnostic mammography may also have two views. Additional diagnostic mammography may offer an in-depth look at suspicious areas. Diagnostic mammography is often performed as a follow up examination of an abnormal screening mammography in order to determine whether the area of concern on the screening examination needs additional breast imaging or a biopsy to determine whether the woman has breast cancer [NCI Cancer Fact Sheets (2007)].

The adoption of mammographic examinations, especially screening mammography, has been proven to increase the rate of detection of cancer and reduce the rates of morbidity and mortality [K. H. Ng, M. Muttarak, 2003]. One of the difficulties with mammography [T. Wang, N. Karayiannis, 1998] is that mammograms generally have low contrast. This makes it difficult for radiologists to interpret the results. Studies [R. Bird, 1992, K. Kerlikowske et al 2000] have shown that mammography is susceptible to a high rate of false positives as well as false negatives, causing a high proportion of women without cancer to undergo further clinical evaluation or breast biopsy, or miss the best time interval for the treatment of cancer. Several solutions have been proposed to increase the accuracy, specificity, and sensitivity of mammography and reduce unnecessary biopsies.

Double reading of mammograms [J. Brown, S. Bryan, 1996, R. Warren, S. Duffy, 1995] has been advocated to reduce the proportion of missed cancers. The basic idea of double reading is to have two radiologists read the same mammograms. According to Warren and Duffy [R. Warren, S. Duffy 1995], double reading can contribute significantly to high sensitivity and effective screening. However, the workload and cost associated with double
reading are high. Instead of double reading, CAD, which is referred to as the “second pair of eyes of the radiologists,” is aimed to be used to aid radiologists in their interpretation of mammograms. With a CAD system, only one radiologist is needed to read each mammogram rather than two. The adoption of a CAD system could reduce the experts’ workload. It has been proven that CAD systems can improve the detection rate of cancer in its early stages. For example, research by Morton et al. [M. Morton, et al 2006] indicates that the use of CAD improved the detection of breast cancer with a 7.62% increase in the number of breast cancers detected, with a small but acceptable increase of 0.93% in the recall rate, and a minimal increase in the number of biopsies with benign or negative results. Brem et al. [R. Brem et al 2003] reported that use of a CAD system significantly improved the detection of breast cancer by increasing the radiologist’s sensitivity by 21.2%.

The double reading process is increased but it will required a machine process. Instead of executing the double process, the captured image could be preprocessed with geometrical analysis is implemented and the result are obtained and presented in this research work.

1.6 ORGANIZATION OF THESIS

The research work is summered and presented into seven chapters.

The first chapter covers the scope, objectives, overview of the research, motivation and origin of the research work.

The chapter two describes about the different medical image analysis methods using computational techniques and possible impacts on the applications which are collected during the review of literature.
The chapter three describes data mining techniques for medical image classification and clustering. The geometrical image preprocessing techniques for the selection of the image and its attributes also discussed.

The proposed Multi view univariate classification is explained in chapter four. The classification algorithm and the image conversion samples are presented.

The classification of the range values and the obtained clusters range values are presented in chapter five.

Based on the observations of clustered range values, all the findings and the interpretations are discussed under chapter six.

Chapter seven deals with the impact of the research, learning and the future work. All the referred documents, research works are listed under bibliography. The sample reports of the observed cluster images, its values are listed on appendix.

1.7 SUMMARY

The first chapter provides the introduction to the research along with the scope and objectives. The research is initiated to determine the level of breast cancer with single scanning but increasing the strength of the computation process using geometrical view of the images as a preprocess. The geometrical view of the process differs and the digital values also differ but the determination results are obtained on density of the Digital values and the data mining algorithms and techniques.