Chapter 1

Introduction

Prevention of cancer by the early detection of its presence is the main goal of researchers. Although our knowledge of cancer is ever increasing but we still not fully understand the cause of cancer. It is at this amiable situation of increasing knowledge and decreasing mortality, we must make use of a computer for the analysis of a patient data to assist a doctor. The amount of medical knowledge is such that today no physician can access or memorize all the necessary information in his daily practice. The Computer Assisted Medical Diagnosis system (CAMD) needs to be developed. The CAMD has two main objectives: help physicians for medical diagnosis and offer a rapid access to medical information by using telemetric networks. The knowledge base may contain information on more than 10000 diseases from all pathological fields, using more than 100000 signs or symptoms. In this proposal, breast cancer and brain tumor are considered for study.

1.1 Breast Cancer

World widely the breast cancer is the leading and the second most fatal disease in women [1]. It is one of the most deadly cancers among the middle-edged women. There have been a host of consistent efforts to mitigate the malady. Successful treatment is a key to reduce the high death rate. Early detection of breast cancer avoids the risk. Cancers in the early stages are permissible to treat while those in the advanced stages are usually impossible to treat. X-Ray mammography is the most important modality for detecting the early-stage breast cancer. Currently, X-ray mammography is the clinical Gold Standard for the detection of breast cancer. It is a well understood and standardized procedure. It works fairly well in postmenopausal women and is inexpensive. Consequently, breast cancer is an intensely researched field [2-7].

Micro-calcifications and masses are the two most important indicators of malignancy and their automated detection is very valuable for early breast cancer diagnosis. Micro-calcifications
are calcium deposits as small as 0.1–0.3 mm in diameter and they can be identified in mammograms as tiny areas that are slightly brighter than their surrounding tissues. Only clusters of three or more particles within a region of approximately 1 cm are considered as clinically suspicious. Micro-calcifications clustered are considered a strong indicator of malignancy and they appear in 30-50% of the mammographically diagnosed cases. Since masses are often indistinguishable from the surrounding parenchymal, automated mass detection and classification is even more challenging [8]-[12].

In the proposed research, segmentation is carried out separately for masses and micro-calcifications. Mass segmentation of mammograms by both entropy thresholding and Ostu’s method is attempted. Other approaches such as marker control watershed and level set are also compared for mass segmentation. In addition, micro-calcifications segmentation is performed with Foveal adaptation segmentation algorithm and local background subtraction technique. After segmentation, feature extraction is essential for effective classification. Texture and shape based features are extracted from images for masses and micro-calcifications respectively. The rough set theory is used for feature selection and finally classification results are presented using support vector machine (SVM) and neural network classifiers.

1.1.1 Literature Survey on Breast Cancer

Image segmentation aims at partitioning the image into physically meaningful regions. The breast region is identified by the presence of higher gray values than that of the non-breast region. Therefore, thresholding is often employed for the segmentation of the breast region. Masses are more difficult to detect than micro calcifications because the features of a mass bear semblance to those of the normal breast parenchyma. A mass is demarcated as a space-occupying lesion seen in more than one projection and is usually characterized by its shape and margin. A mass with regular shape has a higher probability of being benign whereas a mass with an irregular shape has a high probability of being malignant. There are innumerable numbers of segmentation techniques, mention may be made of: thresholding, edge detection or gradient-based. However, region growing is preferred for the segmentation of suspicious portions from the mammograms [13-16].

Bick et al. [17] have explored a segmentation method for the breast region based on the morphological gradient calculation and the modified global histogram analysis. Ball and Bruce
[9] present an automated mammographic computer aided diagnosis system to detect and segments spicules. Mendez et al. [20] describe an automatic algorithm that computes the gradient of the gray levels. Wirth and Stapinski [21] make use of the snakes and fuzzy approach [22] for the segmentation. Both Cheng et al. [23] and Rangayyan et al. [24] quantitatively have dealt with enhancement, detection, characterization and classification of masses. Elter and Horsch [25] have focused their attention on approaches for the diagnosis of mass and micro-calcification, covering the segmentation of region of interests for extracting shape and contour features and their posterior classification [26]. In particular neural networks have demonstrated their efficacy in the clinical domain on diseases such as cancer where there is a weak relationship between the classes forming a benign or malignant diagnosis [27-29]. Hassanien [30] has proposed a hybrid scheme that combines the advantages of fuzzy sets and rough sets in conjunction with statistical feature extraction techniques. An application of breast cancer imaging has been chosen and hybridization scheme applied to see its ability and accuracy to classify the breast cancer images into two outcomes: cancer or non-cancer [30]. Du et al. [31] have presented a framework for the improvement of mammogram classification, which includes a new preprocessing methodology for segmenting unique associative rule discovery based algorithm for the classification and an evaluation of efficacy of the derived features using fuzzy K-nearest neighbor method and agglomerative clustering of associative features. A co-occurrence analysis [32] is applied to identify the statistically significant differences in pathology co-occurrence patterns between the premenopausal and postmenopausal women. A proper combination of the kernel function and the training-test partition can maximize the performance of SVM classifier up to a classification accuracy figure of 99.385% [33].

1.2 Brain tumors

The brain serves as the control center for the functions of the body and allows us to cope with our environment. It is made up of many types of cells. When cells lose the ability to control their growth, they are divided too often without any order. The extra cells form a mass of tissue called a tumor. A tumor is classified as either benign or malignant. A benign tumor is not cancerous and does not spread to other parts of the body. In contrast, a malignant tumor is cancerous; it can penetrate and destroy healthy body tissues, as well as travel to other parts of the body.
imaging many methods are available such as CT scan, Ultrasound and MRI etc. The MR imaging method is the best due to its higher resolution than the other methods. MR imaging is currently the method of choice for the early detection of brain tumor in human brain. However, the interpretation of MRI is largely based on radiologist's opinion. Generalization of brain screening programs requires efficient double reading of MR image, which allows reduction of false negative interpretations, but it may be difficult to achieve. Computer aided detection systems are dramatically improving and can now assist in the detection of suspicious brain lesions, suspicious masses. The task of manually segmenting brain tumors from MR imaging is generally time consuming and difficult. An automated segmentation method is desirable because it reduces the load on the operator and generates satisfactory results [34].

Brain tumors are the second leading cause of cancer death in children under 15 years and young adults up to the age of 34. These tumors are also the second fastest growing cause of cancer death among humans older than 65 years. Early detection and correct treatment based on accurate diagnosis are important steps to cure the disease. Currently, magnetic resonance imaging (MRI) is an important tool to identify the location, size and type of brain tumor. A tissue usually becomes dense when diseased. Often, surrounding tissues are pulled toward the cancerous region, resulting in distortion. Masses are examined for location, shape, density, size and definition of margins. Higher density is usually an indicator of malignancy, while lucent-centered lesions are usually benign. Cancerous lesions generally have a more irregular shape than benign lesions. Most benign masses are circumscribed, compact and roughly elliptical. Malignant lesions usually have a blurred boundary and an irregular appearance [35-38].

In the proposed research brain tumor segmentation is carried out by the level set, marker controlled watershed and modified gradient region growing techniques and results are compared with the manual segmentation of an expert radiologist. Also volume calculation is performed with the seeded region growing and level set evaluation methods for 3D MRI tumor and validated with the manual expert segmentation results.

1.2.1 Literature Survey on Brain Tumors

Brain tissue and tumor segmentation in MR images has been an active area of research today [39-41]. In general the problem of image segmentation involves clustering of similar feature vectors [42-43]. Extraction of good features is thus fundamental to successful image
The segmentation task becomes more challenging when one wants to derive common decision boundaries on different object types in a set of images. Owing to the complex structure of different tissues such as white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF) in the brain images, extraction of useful feature is a fundamental task. Intensity is an important feature in discriminating different tissue types in brain MR images. However, using intensity feature alone to segment complex brain tissue and tumor in a single modality MR image has been found to be insufficient [40-46].

Nowadays, detection of anatomical brain structures with their exact location and orientation has become an extremely important task in the diagnosis of brain tumor [47]. Detection of anatomical brain structures plays an important role in the planning and analysis of various treatments including radiation therapy and surgery [48]. Because of this, development of efficient and accurate MRI segmentation technique has become one of the most important aspects of research today, worldwide. These days, in most of the hospitals, radiologists performs the diagnosis of brain tumor manually on MR images, and it is time consuming and error prone process, in particular because of large number of image slices of single patient and due to the large variation in the intensity of various images representing different brain structures.

Owing to the involvement of various kinds of abnormalities, pathology, radiologist’s perception and image analysis [49] at the diagnosis stage, manual segmentation of brain tumor from MR image seems to be a difficult and time consuming task. All of these lacunas concerned with manual segmentation make a computer aided segmentation tool most desirable. In the last 20 years, several techniques have been developed by researchers to identify anatomical brain structures/brain tumors. But most of them have their own limitations. So, none of them has gained wide popularity in the field of image segmentation. Some of them are based on edge detection, clustering and basic watershed segmentation. The edge detection technique works effectively on high contrast images. This method fails in detecting the edges in low contrast noisy images due to the weak gradient magnitude [50]. Similarly, the clustering based method such as K-means algorithm has a fast speed which allows it to run on large datasets. But its main disadvantage is that it does not produce the same result with each run, because the resulting clusters depend on the initial random assignments [51-52]. Another method is morphological watershed segmentation, but major problem with this method is that it produces over segmentation. A new marker based watershed algorithm which requires less processing time and
minimizes the over segmentation problem up to a large extent has been proposed [53-54]. Also, level set approach is used as another powerful tool for MRI brain tumor segmentation to achieve accurate estimation of area.

Manual method is gold standard approach for MRI quantitative measurements. The main disadvantage of this method is that it is labor intensive and time consuming. In the case of MRI segmentation, uncertainty is introduced due to factors such as partial volume effects, integration of multi-protocol image data and observer variability. Segmentation of region of interest in medical images is still a challenging problem. Current survey proves that region growing is an effective approach for image segmentation especially for the homogenous regions. The disadvantage of region growing is the partial volume effect [55]. The partial volume effect limits the accuracy of MRI brain image segmentation. It blurs the intensity distinction between tissues classes at the border of the two tissues types because voxel may represent more than one kind of tissue types. M. Sato et al. [55] developed a suitable modification in region growing technique. This modification is called modified gradient magnitude region growing technique (MGMRGT) used to remove the partial volume effects and to incorporate gradient information for more accurate boundary detection and filling holes occurred after segmentation [55-56].

1.3 Motivation

Mammography is currently the method of choice for the early detection of breast cancer in women. However, the interpretation of mammograms is largely based on radiologist's opinion. Generalization of breast screening programs requires efficient double reading of mammograms, which allows reduction of false negative interpretations, but it may be difficult to achieve. CAD (Computed Aided Detection) systems are dramatically improving and can now assist in the detection of suspicious mammographic lesions, suspicious micro-calcifications, masses or architectural distortion. Characterization of the lesions is improving as well. CAD mammography might compete with or substitute the human double reading.

Tumor volume is considered useful in evaluating disease progression and response to therapy and in assessing the need for changes in treatment plans. Delineation has been found to agree with the operators' visual inspection most of the time except in some cases when the tumor is close to the boundary of the brain. The developed semi-automatic segmentation methodologies
are rapid, robust, consistent, yielding highly reproducible measurements and are likely to become part of the routine evaluation of brain tumor patients in our health system.

Magnetic resonance imaging (MRI) is currently the method of choice for the detection of brain tumor. Segmentation of 3-D tumor structures from MRI is a very challenging problem due to the variability of tumor geometry and intensity patterns. Manual method is gold standard approach for MRI quantitative measurements. The main disadvantage of this method is that it is labor intensive and time consuming. In the case of MRI segmentation, uncertainty is introduced due to factors such as partial volume effects, integration of multi-protocol image data and observer variability. Manual segmentation of these abnormal tissues can not be compared with the modern day's high speed computing machines which enable us to visually observe the volume and location of unwanted tissues. Automated tumor segmentation in mammograms poses many challenges with regard to characteristics of an image. Preliminary comparisons demonstrate that the semi-automatic segmentation comes close to the manual expert segmentation. Therefore, a semi-automatic/automatic method is more desirable for segmentation of brain tumor from MR images to reduce the workload of radiologists to improve the accuracy of segmentation.

1.4 Tasks of the Thesis

The following are the tasks that are taken up in this thesis for consideration.

1) To develop algorithms to extract shape based and texture based features from mammograms.
1) To develop a system for the feature selection in mammograms.
3) To develop SVM and neural network classifiers.
4) To develop breast cancer mammogram segmentation techniques for masses and micro-calcifications.
5) To study the design of CAD-PACS system and its integration.
6) To develop techniques for the segmentation of MRI brain tumors.
7) To study the risk of brain tumors from wireless phone use.
1.5 Organization of the Thesis

The thesis is organized as follows:

Chapter 2 discusses two important thresholding techniques and a novel approach for the segmentation of masses. Also, level set and marker-controlled watershed algorithms are used for the segmentation and characterization of masses.

In Chapter 3, we describe an approach for the segmentation of micro-calcifications and compare the results with other techniques.

Chapter 4 presents the feature extraction for both masses and micro-calcifications for classifying them into benign or malignant categories.

Chapter 5 describes the current CAD and PACS technologies in medical imaging.

Chapter 6 details out various brain tumor segmentation techniques used in this research work.

Chapter 7 examines the risk of brain tumors from wireless phone use.

In Chapter 8 conclusions of the thesis work and suggestions for the future work are provided.