Abstract

Methodologies for early detection of breast cancer still remain an open problem in the Research community. Breast cancer continues to be a significant problem in the contemporary world. Nearly 25% of breast cancer deaths occur in women diagnosed between 40 to 49 years. Early detection and treatment are currently the only proven means to reduce breast cancer related mortality rates. Computer aided detection is suggested as an adjunct to screening mammogram to decrease errors in perception. Medical Image Processing tools have been demonstrated as effective tools for helping radiologists to identify suspicious tissues in different medical imaging modalities such as Mammograms, Magnetic resonance imaging (MRI), and Ultrasound.

It is found that there are several types of abnormalities in breast. Among those, signs of breast cancer are normally associated with asymmetry between images of left and right breasts. Other type of abnormalities related to breast tumors is presence of micro-calcifications in the breast, presence of masses in the breast and Architectural Distortion (AD). Architectural Distortion refers to, disruption of the normal arrangement of the tissue strands of the breast resulting in a radiating or haphazard pattern without an associated visible centre. Micro-calcifications (MC) are tiny deposits that range from 50 to several hundred microns in diameter, which usually appear in clusters. Masses are signs of breast cancer. Masses with spiculated margins have a high likelihood of malignancy. Architectural distortion (AD) is the third most common mammographic finding of breast cancer. Literature informs that about 81% of spiculated mass and 48-60% of AD is malignant and it is estimated that 12-45% of cancers not found in mammographic screening are AD. The detection sensitivity of the current computer systems for
spiculated mass and AD is not as effective as micro-calcification detection algorithms and thus there is a pressing need for improvements in their detection. This distortion of architecture may be the only visible evidence of the malignant process. Though various methods have been proposed for detection of AD and spiculated masses in mammograms, the accurate and robust detection still remains a technical challenge since the spiculated patterns are often subtle and varied in appearance. Accurate detection of suspicious tissue patterns such as architectural distortion, spiculations and micro calcification in breast images reveal early signs of breast cancer. In the thesis we have devised many techniques to detect these suspicious patterns accurately.

State of art says that the demand for breast screening services is increasing because there are greater numbers of women older than 40 years in the population and there will be a need for greater number of interpreting physicians. At the same time, particularly in environments where the requirement for interpreting large volumes of screening, which may lead to fatigue and distraction, there is a threat for misinterpretation and implicit litigation. Hence, to lessen the burden on the physicians who face such problems and extra tasks it is imperative to facilitate more expertise, and to train sufficient numbers of residents to interpret mammograms or other modalities in the future is required. Difference of opinion between readers was greater than the differences between imaging techniques. There are many image acquisition, display and processing parameters, and their effects on optimizing images for human interpretation are largely unknown. An expert system with some complementary information would prove valuable in minimizing the troubles that he or she might face due to inexperience. It is also a valuable teaching tool for the junior radiologists. An expert system for this application would make diagnostic expertise wider and makes it readily available in the clinical community. In the dissertation we have developed new techniques that can support radiologists/physicians for better interpretation, which in turn helps in accurate and early diagnosis of breast cancer.
Mammography is a most effective imaging modality in early breast cancer detection. The radiographs are searched for signs of abnormality by expert radiologists. But mammograms are complex in appearance and the signs of early disease are often small or subtle. That is the main reason for many missed diagnoses that can be mainly attributed to human factors. In order to improve the accuracy of interpretation or to prompt with the locations of possible abnormalities, a variety of computer systems have been proposed. Mammograms generated by screening of population must be interpreted and diagnosed by relatively few radiologists. In order to improve the accuracy of interpretation, a variety of computer-aided diagnosis (CAD) systems have been proposed. Boundary detection and image segmentation constitute a crucial initial step before performing high-level tasks such as abnormality detection. Segmentation in digital mammogram includes two stages: Breast contour identification and Pectoral muscle segmentation. The pectoral muscle represents a predominant density region in most medio-lateral oblique (MLO) views of mammograms, and can affect the results of image processing. In our work, we propose two preprocessing algorithms, one for the breast boundary extraction and the other for pectoral muscle detection and elimination. To detect the abnormalities like AD and spiculated masses in mammograms we have followed a model based approach which separates the abnormal patterns of AD and spiculated masses from normal breast tissue. The model parameters are retrieved from Gabor filters which characterize the texture features and synthetic patterns generated to retrieve specific patterns of abnormalities in mammographic images. In addition, eight discriminative features are extracted from region of interest (ROI) which describes the patterns representing AD and spiculated masses. Support vector machine (SVM) and Multi-layer Perceptrons (MLP) classifiers are used to classify the discriminative features of AD and spiculated masses from normal breast tissue. Our proposal is based on the texture pattern that represents salient features of AD and spiculated mass.
Ultrasound or Sonography is the primary adjunctive breast cancer screening modality to mammography. It determines whether the area under consideration consists of a cancerous tissue or not. In addition, Ultrasound imaging can be made from any angle or orientation. Diagnosis using ultrasound is noninvasive, it is also safe and does not have any cumulative biological side effects. Mainly, strength of this imaging technique is low acquisition time, portability and relatively inexpensive. Ultrasound images are used to evaluate lumps that are hard to see on a mammogram. Therefore, nowadays the ultrasound images are markers for the early detection of some breast cancers. Ultrasound images have identical textures at different depths, different brightness, and the images are further corrupted by speckle noise due to the attenuation characteristic of ultrasound. Therefore, the first step in breast ultrasound image CAD system is preprocessing that suppresses the speckle noise. Preprocessing and segmentation separates the suspicious regions that may contain masses from the background parenchyma. Thus, denoising and segmentation is essential to maintain high sensitivity and accuracy for the abnormality detection and classification. In our work we have devised: image denoising (Gaussian smoothing and anisotropic diffusion) to remove additive and multiplicative noise of ultrasound images, and image segmentation (active contours). Characterization of tissues in ultrasound images is done by extracting speculation, shape and shadow features. Spiculations of mass are found by measuring the angle of curvature at each pixel of contour. For shape feature we consider shape of the mass because it is clear from the literature that most of the benign masses tend to be wider and roughly ellipse. A mass is said to produce acoustic shadow if the ultrasound is attenuated when crossing through it. When a mass generates acoustic shadow it is considered as malignant. Shadowing features are computed by detecting acoustic shadow appearance to discriminate malignant mass from benign one. Based on these features the spiculated malignant mass can be significantly discriminated from the benign masses by the classifier.

Mammography and ultrasound are currently the most sensitive noninvasive modalities for detecting breast cancer, but they have their own limitations. In
mammograms, glandular tissues look dense and white, much like cancerous tumour. As mammographic images have noisy nature and underlying structures obscure the features, mammographic image analysis leads to high miss rate and low specificity. Whereas the cancers found on ultrasound are almost all small invasive cancers that have not yet spread to the lymph nodes and therefore have good prognoses. However ultrasound itself has some limitations: low resolution, low contrast, blurry edges and speckle noise. Hence it becomes cumbersome for a radiologist to read and interpret an ultrasound image. The above argument justifies that features retrieved from one modality are not sufficient to detect the abnormalities of breast cancer in early stages.

Integration of multimodalities have been widely used for generating more diagnostic and clinical values in medical imaging. As early detection of cancer is probably the major contributor to a reduction in mortality for certain cancers, images guided and targeted minimally invasive therapy has the promise to improve the outcome and reduce collateral effects. The existing methods concentrate on only features retrieved from one modality. Extracting features from dual modality such as mammography and ultrasound helps us to retrieve structural and functional features of breast to detect the abnormal tissue patterns of breast cancer more accurately.

One of the fundamental issues in the design of a multimodal system is to determine the type of information that should be fused. Depending on the type of information that is fused, the fusion scheme can be classified as feature level, score level and decision level fusion. We have used feature level fusion which is expected to be the best type of fusion because the feature sets constitute the richest source of information. Feature level fusion refers to combining different feature sets that are extracted from multiple sources. In feature level fusion, the feature sets of each modality are generated separately, and then fused together to produce a single multimodal template. This template is then used as the input to
the matcher or classifier, which is then fed to the decision making module. Our work includes methods to detect spiculated masses and architectural distortions (AD) in mammogram and ultrasound images. Distinctive and discriminative features related to these abnormalities are retrieved from these two modalities are normalized and fused for better classification. For classifying the mass as benign or malignant Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) classifiers are used. The proposed techniques for mammograms are applied on standard Mammographic Image Analysis Society, London, U.K. (Mini MIAS) data base and DDSM data base. The thesis is structured according to the mentioned objectives and is organized as follows:

- An extensive survey on mass segmentation and detection methods (mammogram and ultrasound).
- Preprocessing and segmentation of mammographic images. Segmentation in digital mammogram includes two stages: Breast contour identification and Pectoral muscle segmentation.
- Analysis and classification of mammographic images. Where we have used methods for retrieving texture features and nodal features Algorithms are presented to reduce false positives and to extract distinctive features for characterizing AD and spiculated mass from normal breast tissue. Finally these extracted features are sent as an input for MLP and SVM classifiers and the results are discussed.
- Preprocessing methods for ultrasound images and novel methods to detect spiculated mass in ultrasound image. We have proposed new methods to retrieve discriminative features form suspicious lesions from ultrasound images and finally have applied classifiers.
- Information retrieved from above two modalities are normalized and fused using different methods to analyze the suspicious tissue patterns for early detection of breast cancer.
To summarize, in the proposed research work we have conducted detailed study to understand the suspicious patterns such as AD, spiculations and micro calcification in mammography and ultrasound images. We have improved the suspicious patterns detection in mammography by devising novel techniques to overcome some of the bottlenecks in the existing methods. New features are designed to improve the spiculated tissue detection in ultrasound images. Further to leverage the complimentary information available in mammography and ultrasound images we have proposed novel fusion methodology to improve the overall suspicious tissue detection which in turn helps in early detection of breast cancer.