Chapter 1

Introduction

Breast cancer is the second most common malignancy that affects women worldwide and is the leading cause among non-preventable cancer death [1]. The American Cancer Society (ACS) estimates that on an average, in every 15 minutes five women are diagnosed with breast cancer. It is also estimated that one in eight women will be diagnosed with this disease in her lifetime, and 1 in 30 will die from it [2]. Breast cancer is the second most prevalent cancer among Indian women, the first being cervical cancer [3]. In the age group of 30-70 years, one in fifty eight women are affected by this disease and the occurrence is mainly seen in the urban areas.

Mammography is the best technique for reliable detection of early, non-palpable, potentially curable breast cancer [4]. As a result of the increasing utilization of mammographic screening, the mortality rate due to this disease was observed to decrease for the first time in 1995 [5]. Since the interpretation of mammograms is a repetitive task that requires much attention to minute details, the opinion of radiologists may vary. To overcome this difficulty, during the past decade, the use of image processing techniques [6], [7], [8], [9], [10] for Computer Aided Diagnosis (CAD) in digital mammograms has been initiated. This has increased diagnostic accuracy as well as the reproducibility of mammographic interpretation.
Chapter 1. Introduction

1.1 Digital Image Processing

Digital image processing is a rapidly evolving field with growing applications in the fields of science and Engineering. Interest in Digital Image Processing stems from two principal application areas: improvement of pictorial information for human interpretation and processing of scenic data for machine perception. It finds application in a wide range of areas like image transmission and storage for remote sensing via satellites, automated inspection of industrial parts, industrial machine vision for product assembly, automatic character recognition, automatic processing of finger prints, RADAR, SONAR and acoustic image processing, Medical image processing etc.

Images have their information encoded in the spatial domain. In other words, features in images are represented by edges, not by sinusoids. Hence, the spacing and number of pixels are determined by how small a feature need to be seen, rather than by the formal constraints of the sampling theorem. A digital image can be considered as a matrix whose row and column indices identify a point in the image and the corresponding matrix element value identifies the gray level at that point.

Processing of digital images involve procedures that are usually expressed in algorithmic form. Thus with the exception of image acquisition and display, most image processing functions can be implemented in software. Transforms are the fundamental tools that are used in most of the image processing applications. The wavelet based multiresolution analysis is found to be one of the best tools for this.

The various realms of image processing are briefly described below [11], [12]:

1.1.1 Image enhancement

The principal objective of image enhancement techniques is to process a given image to make it more suitable than the original for some specific application. These techniques do not increase the inherent information content in the data but emphasize certain image characteristics. Enhancement is useful for feature extraction, image analysis and display of visual information. The enhancement techniques fall into two broad categories:
frequency domain methods and spatial domain methods. The former is based on the modification of the Fourier Transform of an image and the latter refers to the direct manipulation of pixels in an image. Image enhancement operations include contrast and edge enhancement, pseudo coloring, sharpening, magnifying and noise filtering.

1.1.2 Image Restoration

Image restoration is the process that reconstructs or recovers a degraded image, using some apriori knowledge of the degrading phenomenon. The ultimate goal of restoration is to improve a given image in some sense, as in image enhancement. The difference between enhancement and restoration is that the former is concerned with accentuation and extraction of image features while the latter restores degradations.

1.1.3 Image compression

Digital representations of images usually require a very large number of bits. In many applications it is important to consider techniques for representing an image or the information contained in it using fewer number of bits. Image compression addresses this problem. Image data compression methods fall into two categories: Predictive coding and Transform coding. In predictive coding compression is achieved by exploiting the redundancy of the data. Techniques such as delta modulation, differential pulse code modulation etc. fall into this category. In transform coding the given image is transformed into another domain such that a large amount of information is packed into a small number of samples. The compression process inevitably results in some distortion due to the removal of relatively insignificant information.

1.1.4 Image segmentation

Image segmentation is an essential preliminary step in most automatic pictorial pattern recognition and scene analysis problems. It is the process that subdivides an image into its constituent parts or objects. The concept of segmenting an image is generally based
on the similarity or discontinuity of the gray level values of its pixels and can be applied to both static and dynamic images.

1.1.5 Image description and representation

Representation and description of objects or regions of interest, that have been segmented out of an image are the initial steps in the operation of most automated image analysis systems. After segmentation, the resulting aggregates of pixels are represented and described in a form suitable for further processing. Generally, an external representation is chosen when the primary focus is on morphological features. When one is interested in reflectivity properties such as color and texture, an internal representation is selected. The choice is dictated by the problem under consideration, so as to capture the essential differences between objects or class of objects, maintaining as much independence as possible to changes in factors such as location, size and orientation.

1.2 Medical Image Processing

The advent of medical imaging is one of the milestones in the progress of medical science. It serves as a beneficial tool for the medical practitioners during diagnosis of ailments. The application of image processing techniques to medical imaging has made the results accurate and reliable. In many cases it is possible to eliminate the necessity for invasive surgery, thus avoiding trauma to the patient as well as an inevitable element of risk.

One of the early applications of image processing in the medical field is the enhancement of conventional radiograms. When converted to digital form, it is possible to remove noise elements from X-ray images, thereby enhancing their contrast. This aids interpretation and removes blurring caused by unwanted movement of the patient. This form of representation also enables the physicians to measure the extent of tumors and other significant features accurately.
1.3. Tools for Image Processing

The basic image processing operations on medical images are conveniently placed in four categories: filtering, shape modeling, segmentation and classification [13]. Filtering includes linear and non-linear enhancement, deblurring and edge detection techniques using local operators or classification techniques. Shape modeling includes three-dimensional representation and graphics manipulation such as three-dimensional contours of the spinal column, coronary artery or shaded images. Clustering, object detection, and boundary detection are the main operations that come under segmentation. Simple histogram or thresholding techniques are used to segment objects of interest. When adequate prior information is available matched filters can be used effectively. Heuristic techniques are useful for tracing contours in the presence of highly structured background such as chest radiographs. Feature selection, texture characterization and pattern recognition are the major operations in classification. [14], [15].

Another application of digital image processing in medical imaging is ‘tomography’, the generation of images of a slice through the body [16] involving the reconstruction of two-dimensional images.

**1.3 Tools for image processing**

The first step after obtaining the image in any digital image processing system is preprocessing that image. The key function of this is to improve the image in ways that increase the chances of success of other processes. Wavelet Transform (WT) techniques are found to be a very effective processing tool for this purpose.

Neural Networks are found to be efficient tools for classification applications. They are rough models of human mental processes with powerful learning, memorization, and associative recall capabilities of pattern formatted information.

A brief introduction to these two image processing tools are provided in the sections below:
1.3.1 The Wavelet Transform

Perhaps the most prominent signal analysis technique is Fourier analysis, which breaks down a signal into its constituent sinusoids of different frequencies or transforms our view of the signal from a time-based one to a frequency-based one. But, this has the serious drawback of loss of time information while transforming into the frequency domain. This is not very prominent for stationary signals. However, Fourier analysis become inadequate when the local frequency contents of the signal are of interest or when it contains non-stationary or transitory characteristics like drift, trends, abrupt changes, etc.

In an effort to correct this, Dennis Gabor [17] adapted the Fourier transform to analyze only a small section of the signal at a time — a technique called windowing the signal. Gabor’s adaptation, called the Short-Time Fourier Transform (STFT), maps a signal into a two-Dimensional (2-D) function of time and frequency. While the STFT’s compromise between time and frequency information can be useful, the drawback is that once a particular size is chosen for the time window, it remains the same for all frequencies.

Wavelet analysis, a windowing technique with variable-sized regions, represents the next logical step. It allows the use of long time intervals where more precise low frequency information is needed and shorter intervals where high frequency information is needed. One major advantage offered by wavelets is the ability to analyze a localized area of a larger signal. Further, because it offers a different view of data than those presented by traditional techniques, wavelet analysis can often compress or de-noise a signal without appreciable degradation. Indeed, in their brief history within the signal processing field, wavelets have already proven themselves to be an indispensable addition to the analyst’s collection of tools and continue to enjoy a burgeoning popularity today.

Wavelets are oscillatory functions that exist for a few cycles only and satisfy certain properties. Most of the wavelets are associated with a scaling function. There are
various kinds of wavelets like compactly supported wavelets, symmetric and non-symmetric wavelets, orthogonal and biorthogonal wavelets and smooth wavelets.

1.3.1.1 History of Wavelets

From a historical point of view, wavelet analysis is a new method, though its mathematical underpinnings date back to the work of Joseph Fourier in the nineteenth century [18]. Fourier laid the foundations of frequency analysis with his theories, which proved to be enormously important and influential. When it became clear that an approach measuring average fluctuations at different scales might prove less sensitive to noise, the attention of researchers gradually turned from frequency-based analysis to scale-based analysis. The first recorded mention of the term “wavelet” was in 1909, in a thesis by Alfred Haar [19]. Morlet and the team working under Alex Grossmann at the Marseille Theoretical Physics Center in France first proposed the concept of wavelets in its present theoretical form [20]. The main algorithm for WT computation dates back to the work of S. Mallat in 1988 [21]. Since then, research on wavelets has become international and is particularly active in the United States, spearheaded by veteran scientists Ingrid Daubechies, Ronald Coifman, and Victor Wickerhauser [22].

1.3.1.2 The Continuous Wavelet Transform (CWT)

The WT of a signal represents the signal as a linear combination of scaled and shifted versions of the wavelets and scaling functions. When the scale and shift parameters are continuous, the transform under consideration is called a CWT. In the CWT a function \( \psi \), which in practice looks like a little wave, is used to create a family of wavelets \( \psi(at + b) \) where \( a \) and \( b \) are real numbers, \( a \) dilating (compressing or stretching) the function \( \psi \) and \( b \) translating or displacing it. The word continuous refers to the transform, not to the wavelet. The CWT turns a signal \( f(t) \) into a function \( W_\psi f \) of two variables, scale and time as:
$W_{\psi} f(a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \psi^*(at + b) dt$  \hspace{1cm} (1.1)

where $\psi^*$ is the complex conjugate of $\psi$. This transformation in theory is infinitely redundant, but it can be useful in recognizing certain characteristics of a signal.

1.3.1.3 The Discrete Wavelet Transform (DWT)

The CWT maps a signal of one independent variable $t$ into a function of two independent variables $a$ and $b$. The highly redundant nature of this transform makes it inefficient from a computational point of view. One way to eliminate the problem of redundancy is to sample the CWT on a 2-D dyadic grid. That is, use wavelets only of the form $\psi (2^k t + l )$ with $k$ and $l$ being whole numbers. The resulting WT is called DWT. DWT is still the transform of a continuous time signal, with discretization performed in the $a$ and $b$ variables only. Hence it is analogous to the Fourier series, and also referred to as a continuous time wavelet series [23], [24].

1.3.1.4 The Multiplexed Wavelet Transform (MWT)

MWT is an alternate method for the time-scale representation of pseudo periodic signals with constant period, first proposed by Evangelista [25]. This transform simplifies the analysis of a pseudo periodic signal by decomposing it into a regular asymptotically periodic signal and a number of fluctuations over this.

Images can be treated as oscillatory signals, although they are not periodic in a strict mathematical sense. Contrary to the case of one-Dimensional (1-D) signals, no period detection is required in the case of images. When treated as quasi-periodic signals, the periods along the horizontal and vertical directions respectively are the width and length of the image segment. Hence, the DWT of the rows of the image gives the MWT of the image taken as a 1-D signal along the vertical direction and that of the columns corresponds to the MWT of the image taken as a 1-D signal along the horizontal direction.
1.3. Tools for Image Processing

1.3.1.5 WT in Two Dimensions

When the input signal is 2-D, it is necessary to represent the signal components by 2-D wavelets and 2-D approximation function. Often this is done by using separable products of 1-D wavelets and scaling functions which make it possible to use the Fast Wavelet Transform (FWT) algorithms.

For any scaling function and its corresponding wavelet function, we can construct three different 2-D wavelets and one 2-D approximation function using the tensor product approach. Each new wavelet measures the variations along a different direction; vertical, horizontal and diagonal. As a result the 2-D extension of the wavelet transform is achieved by applying the 1-D algorithm along the rows and columns of the image. That is, the image is decomposed row wise first, for every row and then this is repeated column wise for every column.

1.3.1.6 Computation of DWT

The DWT of a signal is determined by finding the detail in the signal at each level of resolution; that is, for each successive value of the dilation variable. In essence, this is done by convolving the input signal with the appropriately dilated wavelet function at each translation. As the dilation increases, the number of translation points for which values must be determined drops; at the highest resolution the wavelet is being used to measure the difference between successive samples while at the lowest resolution the wavelet is comparing the first half of the signal with the second half. When the wavelet family is orthogonal, adding the detail at all levels of resolution yields the original signal.

Stephane Mallat [21] has shown how a scaling function and a wavelet function can be used in a recursive algorithm to compute the orthogonal forward and inverse WT of a signal in $O(n \log n)$ time. This is considered as the standard algorithm for WT computation. The scaling and wavelet functions are in effect low and high pass filters; at each level of recursion wavelet function is used to extract the details at that level of
resolution and scaling function is used to construct a coarser version of the signal for analysis at the next level. The process is repeated on successively coarser representations of the signal, until only the steady-state (average) value of the signal remains.

1.3.1.6.1 Sectioned computation

Generally, the sequences involved in real time implementations are quasi-infinite and processing of such data is done after segmenting it to smaller blocks or frames. The DWT and Inverse Discrete Wavelet Transform (IDWT) are recursive-filtering processes. Hence, WT is not a block transform and due to the lack of data beyond block boundaries, edge artifacts will be produced on block boundaries in the reconstructed signals. For correct computation near the data boundaries each processor would need to access data allocated to other processors. This demands frequent data exchange between processors or requires large buffer storage for intermediate transform coefficients.

1.3.1.7 WT in Biomedical Image Processing

In the past few years, researchers in applied mathematics and signal processing have developed powerful wavelet methods for the multiscale representation and analysis of signals [23], [26]. These new tools differ from the traditional Fourier techniques by the way in which they localize the information in the time-frequency plane. They are capable of trading one type of resolution for the other, which makes them suitable for non-stationary signal analysis. One important area where these properties are found relevant is biomedical engineering.

The main difficulty in dealing with biomedical signals is their extreme variability and the necessity to operate on a case-by-case basis. Often there is no apriori knowledge about the pertinent information and/or at which scale it is located. Frequently, the deviation of some signal feature from the normal is the most relevant information for diagnosis. Another important aspect of biomedical signals is that the
information of interest is often a combination of features that are well localized spatially or temporally (e.g. microcalcifications in mammograms) and others that are more diffuse (e.g. texture). This requires the use of sufficiently versatile analysis methods, to handle events that can be at opposite extremes in terms of their time-frequency localization.

The applications of wavelets in biomedical field include performing image processing tasks like noise reduction, enhancement, detection and reconstruction, acquisition techniques for X-ray tomography and MRI and statistical methods for localizing patterns of activity in the brain using functional imaging.

1.3.1.7.1 Computer Assisted Mammography

Image enhancement is especially relevant in mammography where the contrast between the soft tissues of the breast is inherently small and a relatively small change in the mammary structure can signify the presence of a malignant breast tumor. Because of the current interest in mammographic screening, wavelet based enhancement methods have been recently designed with that application in mind [27], [28], [29]. All these approaches invariably use reversible redundant or non-redundant wavelet decomposition and perform the enhancement by selective modification of WT coefficients. These enhancement techniques are not fundamentally different from the noise reduction techniques, since in the former case certain features of interest are amplified while in the latter some unwanted features are suppressed.

One of the key issues in computer-assisted mammography is the detection of clusters of fine granular microcalcifications, which are one of the primary signs of breast cancer. Individual calcifications typically range from 0.05-1mm in diameter. The detection of microcalcifications is closely related to the enhancement task described earlier, except that detection is typically performed by thresholding in the wavelet domain. The detection results so far reported suggest that wavelet techniques perform better than the best available single scale methods [30], [31], [32], [33].
1.3.1.7.2 Computer Assisted Tomography (CAT)

In 2-D computerized X-ray tomography image of an object is reconstructed from the measured values of its angular projections. These measurements are described by the Radon transform. The primary motivation for using wavelets for tomography is that the wavelet reconstruction formulas tend to be localized spatially and can be applied to obtain partial reconstructions when only a portion of the Radon transform is available (limited angle tomography). The WT also appears to have some merits for noise reduction in tomography.

1.3.1.7.3 Magnetic Resonance Imaging (MRI)

One of the major applications of the WT in medical imaging is the noise reduction in MR images. One approach proposed is to compute an orthogonal wavelet decomposition of the image after applying a soft thresholding rule on the coefficients [34]. A more sophisticated approach is an over complete wavelet decomposition followed by a reconstruction from the retained significant WT maxima by exploiting the correlation between adjacent scales [35], [36]. When applied to MR images this method compared favorably with the optimal Wiener filter and produced images with much sharper edges and did not induce any ringing artifacts [36], [37].

1.3.1.7.4 Functional Image Analysis

Functional neuro-imaging is a fast developing area aimed at investigating the neuronal activity of the brain in vivo. Positron Emission Tomography (PET) and fMRI are the two modalities that are used to obtain functional images. PET measures the spatial distribution of certain function specific radiotracers injected into the blood stream prior to imaging. A typical example is the measurement of cerebral glucose utilization with the tracer 2-fluoro-2-dioxy-D- Glucose (FDG). fMRI allows for a visualization of local changes in blood oxygenation induced by neuronal activation. It is substantially faster than PET and also offers better spatial resolution.
The functional images obtained with these two modalities are extremely noisy and variable and their interpretation requires the use of statistical analysis methods. The first step in this analysis is the registration of various images, which compensates for intersubject anatomical variability or intrasubject movement in the scanner. Efficient multiresolution solutions to this problem have been proposed resulting in much faster and robust algorithms compared to single scale counterparts [38].

The second step is the computation of difference between the aligned group averages and performing the statistical analysis. Direct testing in the image domain is difficult because of the amount of residual noise and the necessity to use a very conservative significance level to compensate for multiple testing. Testing in the wavelet domain has the advantage that the discriminative information, which is smooth and well localized spatially, becomes concentrated into a relatively small number of coefficients while the noise remains evenly distributed among all coefficients.

1.3.2 Neural Network

Traditional DSP is based on algorithms, changing data from one form to another through step-by-step procedures. Most of these techniques also need parameters to operate. For example, recursive filters using recursion coefficients, feature detection implemented by correlation and thresholds, image display depending on the brightness and contrast settings, etc. Algorithms describe what is to be done while parameters provide a benchmark to judge the data. The proper selection of parameters is often more important than the algorithm itself. Neural networks take this idea to the extreme by using very simple algorithms, but many highly optimized parameters. They replace the traditional problem-solving strategies with trial and error pragmatic solutions, and a "this works better than that" methodology.

A neural network structure can be defined as a collection of parallel processors connected together in the form of a directed graph, organized such that the network structure tends itself to the problem being considered [39]. It is radically different from
the notions of ordinary serial computing strategy and forms a powerful tool for applications where the processing is to be done in parallel. They offer the following advantages:

i) Adaptive learning: This is learning to perform specific tasks by undergoing training with illustrated examples. This feature eliminates the need of elaborating *apriori* models or specifying probability distribution functions.

ii) Self-organization: Neural networks use self-organizing capabilities to create representations of distinct features in the presented data, which leads to the generalization of features.

iii) Fault tolerance: Networks can learn to recognize noisy and incomplete data and also exhibit graceful degradation when part of the network itself is destroyed.

iv) Real-time operation: Due to its parallel distributive structure most networks operate in the real time environment and the only time consuming operation is training the network.

Neural networks have been applied in many fields, some of which are mentioned below. In Aerospace applications it is used for high performance aircraft autopilot, flight path simulation, aircraft control systems, autopilot enhancements, and aircraft component simulation and fault detection. In automotive industry it is used for automobile automatic guidance system and warranty activity analysis. It is used in banking sector for cheque and other document reading and credit application evaluation. In the field of communication, neural network finds extensive applications in image and data compression, automated information services, real-time translation of spoken language and customer payment processing systems. In medical field neural networks are employed for breast cancer cell analysis, EEG and ECG analysis, prosthesis design, optimization of transplant times, hospital expense reduction and hospital quality improvement. It is also used in the fields of defense, entertainment, finance, manufacturing, oil and gas exploration, robotics, transportation etc [40], [41].
1.3.2.1 Target detection

Scientists and engineers often need to know if a particular object or condition is present. For instance, geophysicists explore the earth for oil, physicians examine patients for disease, astronomers search the universe for extraterrestrial intelligence, etc. These problems usually involve the comparison of the acquired data against a threshold and if the threshold is exceeded, the target is deemed present. The conventional approach to target detection (sometimes called pattern recognition) is a two-step process. The first step is called feature extraction, which uses algorithms to reduce the raw data to a few parameters, such as diameter, brightness, edge sharpness, etc. These parameters are often called features or classifiers. Feature extraction is needed to reduce the amount of data and to distill the information into a more concentrated and manageable form.

In the second step, an evaluation is made of the classifiers to determine if the target is present or not. This is quite straightforward for one and two-parameter spaces; the known data points are plotted on a graph and the regions separated by eye. As the number of parameters increases this cannot be done by the human brain and dedicated networks are required to carry out this task. The neural network is the best solution for this type of problems. Some of the important neural classifiers include Perceptrons, Backpropagation network, Self-organizing map, Competitive networks, Learning Vector Quantization (LVQ) and Probabilistic Neural Network (PNN).

1.4 Objective of the work

Cancer is not preventable, but early detection leads to a much higher chance of recovery and lowers the mortality rate. Considering the incidence of breast cancer and the favorable prognosis associated with early detection, it is surprising to note that only 15 to 30% of eligible women have ever had a mammogram [42] and even fewer are involved in a regular screening program. Reasons for this are high cost, skepticism about reliability and the physical discomfort of the process.
The high cost of a mammography-screening program can be partly attributed to the fact that the mammographic images are difficult to interpret even for skilled radiologists with years of experience. One reason for this is that a mammographic image is a highly textured 3-D structure, which has been projected onto a 2-D plane. Additionally the images are often of low contrast, in order to maintain low radiation dose to the patients. It can be assumed that less than 10 percent of the mammograms from a screening population contain some type of abnormality. The visual fatigue of reading numerous mammograms, most of which are negative, and the existence of a wide variation of breast tissue structures lead to inconsistent readings between radiologists, and even by a single radiologist at different times.

CAD and automated pre-screening by computer makes it easy to interpret the multitudes of mammographic readings. Even if there is no large screening program computerized mammogram image analysis could be used to improve the quality of conventional mammography. In a CAD scenario, computerized image analysis is used to suggest possible suspicious regions in the image so that a radiologist can then examine these regions more carefully. Evidence is mounting that prompting the radiologist with computer detection results of mammographic images leads to an increased sensitivity without affecting specificity [5], [43], [44].

Cancer treatment is most effective when it is detected early and the progress in treatment will be closely related to the ability to reduce the proportion of misses in the cancer detection task. The effectiveness of algorithms for detecting cancers can be greatly increased if these algorithms work synergistically with those for characterizing normal mammograms. This research work combines computerized image analysis techniques and neural networks to separate out some fraction of the normal mammograms with extremely high reliability, based on normal tissue identification and removal.

The presence of clustered microcalcifications is one of the most important and sometimes the only sign of cancer on a mammogram. 60% to 70% of non-palpable breast carcinoma demonstrates microcalcifications on mammograms [44], [45], [46].
WT based techniques are applied on the remaining mammograms, those are obviously abnormal, to detect possible microcalcifications. The goal of this work is to improve the detection performance and throughput of screening-mammography, thus providing a 'second opinion' to the radiologists.

The state-of-the-art DWT computation algorithms are not suitable for practical applications with memory and delay constraints, as it is not a block transform. Hence in this work, the development of a Block DWT (BDWT) computational structure having low processing memory requirement has also been taken up.

1.5 Layout of the Thesis

The thesis is organized in the following way:

A brief review of the previous research works in the field of computer-aided breast cancer detection is presented in chapter 2. Special stress is given to microcalcification detection and neural network based classification of normal/abnormal tissue in mammograms. Different methods of computation of both 1-D and 2-D WT are also reviewed in this section.

Chapter 3 summarizes the features of different types of breast lesions in digital mammograms, namely, microcalcifications, circumscribed lesions, and spiculated lesions.

Chapter 4 describes the basic theory for classification using neural networks and detection of microcalcifications using WT. An overview of neural networks for classification purposes and multiresolution representations of signals using wavelets are provided. One-dimensional wavelet analysis is discussed; including the orthogonal and biorthogonal wavelet representations and is extended to 2-D. Different types of WTs are also considered in this chapter.

Chapter 5 describes the algorithms developed for block-wise computation of both 1-D and 2-D DWT. The conventional method and its computational complexity are
described in detail. The computational complexity of the BDWT algorithm is evaluated and compared against the standard methods.

Chapter 6 presents the classification of mammograms into normal and abnormal classes using neural networks. First the features of normal mammograms are explained followed by the derivation of different features for classification purpose. Finally results and conclusion are presented.

The new MWT based algorithms for automatic detection of microcalcifications is presented in chapter 7. The microcalcification detection problem is represented as an edge detection operation and different WT based edge detection methods are discussed in detail. Experimental results on mammographic data and discussions are also provided.

Chapter 8 is the concluding chapter, wherein the observations and inferences already brought out in the previous chapters are summarized. The suggestions for further work are also given.

This thesis includes one appendix, which describes a line detector that is capable of extracting linear mammographic features. This line detector is used to find out and remove normal linear markings from mammograms.