CHAPTER 1

INTRODUCTION

1.1 CANCER

Cancer is a group of diseases involving abnormal growth of cells. Cancer may appear and affect any part of the human body. It is made up of numerous of cells. Cancerous tumors fall into two main categories, benign or malignant. While benign (non-cancerous or non-malignant) tumors do not spread to other parts of the body. The malignant ones can spread into, or invade, nearby tissues. There are more than 100 types of cancer, for example breast cancer, skin cancer, lung cancer, colon cancer, prostate cancer, and lymphoma, bladder cancer, endometrial cancer, kidney cancer (Renal cell cancer), thyroid cancer, bone cancer, cervical cancer, eye cancer, liver cancer, ovarian cancer etc. Mostly cancer to develop in the skin, lungs, breast, prostate, colon and rectum.

The main types are Lung (1.3 million deaths/year), Stomach (8,03,000 deaths), Colorectal (6,39,000 deaths), Liver (6,10,000 deaths), Breast (5,19,000 deaths). More than 70% of all cancer deaths have occurred in developing countries. It is expected that cancer deaths in the world might be around 11.5 million in 2030.

Approximately, 1.5 million cancer deaths were averted during the past two decades, while 16, 58,370 new cancer cases were diagnosed and 5, 89,430 cancer deaths occurred. In cancers, some are more serious than others, some are more easily treated than others (particularly if diagnosed at an early stage), and some have a better outlook (prognosis) than others. Thus cancer is of many types: In each case it is important to know exactly what type of cancer has developed, how large it has become, and whether it has spread [127], [30] & [91].

There is a prediction that half of all men and one-third of all women will get cancer at some time in their life. This prediction includes skin cancers, most of which are easily treated. Millions of people have cancer today or had in the past. All age groups are prone to it. It is not contagious. Neither air nor contact nor blood transfusions can spread cancer. The main risk of cancer can be inherited. That is to say that it can be inherited by children from parents. Thus
inheritance is a main risk factor for cancer. There are several risk factors of getting cancer such as Tobacco use, High fat diet and being overweight, Excessive exposure to sunlight, Drinking too much alcohol, X-rays and other source of radioactivity, Geographic area, Chemicals and other substances in the environment, Unsafe sexual practice such as HIV and Family members who have cancer, Chronic infections from helicobacter pylori, Hepatitis B Virus (HBV), Hepatitis C Virus (HCV) and some types of Human Papilloma Virus (HPV) [1].

The common symptoms of cancer are thickening or lump in the body, Cough or Hoarseness that persists, Obvious change in a wart or mole, Changes in bowel or bladder habits, unexplained bleeding or discharge, any sore that does not heal, unusual upset stomach or difficulty in swallowing. Nevertheless, these symptoms just do not mean that the patient has cancer. Only a doctor can make a diagnosis.

Doctors can use various cancer diagnosing methods such as Physical Examination, Laboratory tests such as blood and urine tests, imaging – X-Ray, CT Scan, MRI and Biopsy etc. In a biopsy, tissue is examined directly to see if it has acquired the characteristics of cancer. Needling or a surgical procedure is adopted to extract a tissue. Biopsy helps establish the precise diagnosis of cancer.

More than half of all cancers diagnosed are now curable. Even with cancers that cannot be cured, symptoms are often greatly reduced by treatment. Options of treatment depend on the stage and type of cancer. The different treatment methods of cancer are Surgery, Radiation therapy, Chemo therapy, Biological therapy and Hormone therapy.

1.2 BREAST CANCER

Breast cancer is one of the leading causes of cancer mortality among women. According to American Cancer Society (ACS) breast cancer is the second leading cause of cancer. The most common disease found in women which causes second highest rate of death after lung cancer is breast cancer. However, with increasing rates of breast cancer diagnosis, mortality has been on the decline. The decrease in breast cancer associated mortality may be attributable to several factors falling under the broad categories of treatment and detection. Both have been significantly advanced over the past decades, due to the vast amount of research in the field.
According to a recent study, in 2015, women in the United States, around 60,290 breast carcinoma cancer cases were diagnosed; 83% was ductal carcinoma in situ breast cancer (DCIS) and 12% was lobular carcinoma in situ breast cancer (LCIS).

Breast Cancer (malignant breast neoplasm) is cancer which starts in the cells of the breast tissue, most commonly from the inner lining of milk ducts or the lobules that supply the ducts with milk. Those that originate from ducts are known as ductal carcinomas and the ones that begin from lobules are known as lobular carcinomas. Breast cancer is a disease of humans and other mammals [61]; the disease occurs almost entirely in women, but can also as a rare case develop in men. In the United States, breast cancer dominates and it has been found, to be the most common cause of cancer death in women while it is the main cause of death in women aged between 20 & 50 years. Fig.1 show different parts of breast tissue [16].

Fig.1   Basic Structure of the Female Breast

The most common type of breast cancer is Ductal Carcinoma in Situ (DCIS). It is found in the cells of the ducts. Each breast has 15 to 20 sections called lobes, which in turn get divided into many smaller sections called lobules. It is the thin tubes called ducts that link the lobes and the lobules. The cancers started in these breast cells are called Lobular Carcinoma In Situ (LCIS). Cancers also are classified as Non- Invasive (in situ) and Invasive (infiltrating) Ductal Carcinoma (IDC). The term In-situ refers to cancer that has not spread past the area where
initially it developed. Invasive breast cancer is likely to spread (metastasize) to other tissues of the breast and/or other regions of the body. A less common type of breast cancer is inflammatory breast cancer characterized by general inflammation of the breast. Other rare types of breast cancer are medullar carcinoma (an invasive breast cancer that forms a distinct boundary between tumor tissue and normal tissue), mucinous carcinoma (formed by the mucus-producing cancer cells), tubular carcinoma, etc [3].

Survival rates are significantly increasing when the breast cancer is detected in earlier stage. According to five-year survival rate for people with breast cancer the survival rate is 99% if the cancer is located only in the breast. If, however, it has spread to the regional lymph nodes, the five-year survival rate is 85%. If it spreads to a remote part of the body, the five-year survival rate is 25%. For all stages of cancer combined, the five-year survival rate for breast cancer is 89%, the 10-year rate is 83%, and the 15-year rate is 78% [21] & [56].

The most common and effective early-detection screening tests include mammography, clinical breast examination and breast MRI. Most of the cancers detected in clinical examination and mammography are proven and significant tools in early detection [27] & [103]. Mammography and clinical breast examination are the most effective ways to detect breast cancer for earlier stage. Clinical breast examination is also relatively very simple and more cost inexpensive.

If abnormal finding is detected, a woman is usually required to undergo further testing which may include:

1. Flat Panel X-Ray Detectors (FPXDs) were first used to examine breast tissue a century ago.

2. Ultrasound also known as sonography that uses sound waves to look inside the parts of the body.

3. Magnetic Resonance Imaging (MRI) uses magnets and radio waves. Instead of x-rays powerful magnets linked to a computer are used to create detailed pictures of areas inside the breast. Breast MRI is mainly used for 2 purposes (a).To measure the size of the
cancer and look for any other tumours in the breast (b). To as certain high risk in certain women.

4. **Needle biopsies** remove tissue or cells.

5. Core-Needle biopsies remove a small amount of suspicious tissue.

6. **Surgical biopsies**: a surgeon makes a cut to remove breast tissue.

The most important prognostic factor in breast cancer is the stage at which it is diagnosed. Cancer diagnosed in the early stages has a much higher probability of being cured than those that have metastasized. Screening programs like digital mammography are widely used to detect potential breast cancers.

### 1.3 FILM Vs DIGITAL MAMMOGRAPHY

To study the human breast, digital mammography is used as a diagnostic and a screening tool that uses X-rays. The digital mammogram is a modern computerized device that makes the analysis, interpretation and diagnosis of breast cancer easier and more precise. The goal of mammography is the rapid information of breast cancer, usually through detection of characteristic Microcalcifications (MCs) and masses. It helps to identify breast cancer in the early stage through radiographs of high precision. The essential visual clues of breast cancer include preliminary signs of masses and Microcalcifications (MCs) clusters. The diagnosis falls under three categories, viz., normal, benign and malignant. Normal mammogram shows no sign of cancer in cells. "Benign" means not harmful. Other tumors are "malignant" and are cancerous. Breast Cancer in the beginning, is too small to be felt. It slowly spreads over the breast or to other parts of the body. It causes serious health problems and eventually death [134]. According to Breast Imaging Reporting and Data System (BIRADS) benign and malignant can be differentiated using its shape, size and density. This is how a radiologist visualizes the mammograms. According to BIRADS mass shape characteristics, benign masses tend to have round, oval, lobular in shape and malignant masses are lobular or irregular in shape. Measuring regular and irregular shapes mathematically is difficult as there is no single measure to differentiate various shapes.
Film and digital mammography have the same ability to identify cancer. It is digital mammography that dominates now. Digital images are viewed on a computer. Digital mammography is more accurate than film mammography images can be well controlled on the computer. In general, digital mammography is better at finding breast cancer in women who fall into one of following groups: a) premenopausal or peri-menopausal b) under age 50 c) have dense breast tissue. For others, film and digital mammography are equally good in finding breast cancer early.

1.3.1 Screening for Disease

The main goal of any cancer screening test is to correctly identify those people who have cancer (called the sensitivity of the test). When sensitivity is high, very few cases are missed. However, this means some healthy people will be misidentified as having cancer (a false positive result). For example, a sensitivity of 90 percent means that 90 percent of people tested who truly have cancer are correctly identified as having cancer.

An ideal cancer screening test would also be able to correctly identify all the people who do not have cancer as not having it (called the specificity of the test). When specificity is high, there are fewer false positive results, but more cases of true cancer are missed. For example, a specificity of 90 percent means that 90 percent of the people who are healthy are correctly identified as not having cancer.

No screening test has perfect sensitivity and perfect specificity. There is a trade-off between the two for all types of screening tests. That is, when a test gains sensitivity, it loses some specificity.

The quality of a screening test is described in terms of:

- How well the screening test tells who truly has a disease (sensitivity).
- How well the screening test tells who truly does not have a disease (specificity).

The goal of any screening test is to correctly identify everyone who has a certain disease (100 percent sensitivity).
An ideal screening test would also be able to correctly identify everyone who does not have the disease (100 percent specificity). A perfect test would correctly identify everyone with no mistakes. There would be no false negatives (when people who have the disease are missed by the test) and no false positives (when healthy people are incorrectly shown to have the disease) [46].

By incorporating the expert knowledge of radiologists, the computer-based systems provide a second opinion in detecting abnormalities and making diagnostic decisions. Such a diagnostic procedure is called Computer-Aided Diagnosis (CADx) [82]. Computer-Aided Diagnosis (CADx) is usually used as a second opinion by the radiologists [63]. The proposed (CADx) system improves the accuracy of interpreting mammograms by providing an important opinion to the radiologist. The diagnosis result of tissue is classified into three categories: normal, benign and malignant. Normal represents mammogram without any cancerous cell, benign represents mammogram showing a tumour, but not produced by cancerous cells and malignant represents mammogram showing a tumour with cancerous cells. It is difficult to distinguish a benign MCs/mass from one that is malignant. Consequently, many unnecessary biopsies are often undertaken due to the high positive false rate.

1.4 OBJECTIVE OF THE THESIS

The main objective is classification of abnormal lesions in breast radiographic images to differentiate between normal or abnormal in stage 1, classify mass or microcalcifications (MCs) in stage 2 and finally the abnormal severity in mass or microcalcifications (MCs) either classified into benign or malignant in stage 3.

1.5 SCOPE OF COMPUTER AIDED DIAGNOSIS

The aim of the research is to develop an efficient and automated (CADx) system for the classification of digital mammogram images as normal or abnormal, mass or microcalcifications (MCs) given the abnormality of images and further subdivided into benign or malignant. Also performance of CAD system is measured by accuracy, sensitivity, specificity.
The scope of the thesis is restricted to the digital mammography for the classification of mammogram images even though there are many other techniques such as film-screen mammogram, digital breast tomosynthesis mammography (DBT), biopsy etc., for detecting the breast cancer. Also the scope is restricted to Dual Tree M-Band Wavelet Transform (DTMBWT) technique used for feature extraction even though there are many other feature extraction techniques available. Also the scope is restricted to Support Vector Machine (SVM) classifier used for the classification for mammogram classification.