CHAPTER 1

INTRODUCTION

1.1 OUTLINE OF RESEARCH

Breast cancer is the major killer disease for women and men all around the world. It can be treated only if it has been identified in its earlier stage. Computer aided detection (CAD) and Computer Aided Diagnosis (CADi) is the main aids used by the radiologists in detecting and diagnosing the breast cancer. CAD consists of three main stages. They are preprocessing, segmentation and detection of suspicious regions. Each of the stage has been designed by many techniques. CADi consists of techniques for the diagnosis of cancer presence (benign or malignancy), type and stages of the breast cancer. An x-ray like image has been observed from the human breast are called mammogram image, which is the input to CAD system.

Many researchers proposed and tested many algorithms for the techniques used in the CAD and CADi. The detection efficiency of the existing methods is still to be improved to conclude about, whether the sample mammogram image is having cancer information or not. In this research work, new techniques for preprocessing and segmentation have been proposed for the improvement of detection efficiency of CAD and compared with the existing techniques.
1.2 BREAST CANCER

Breast cancer is a type of cancer and it originating from breast tissue in the inner area of milk ducts. Cancers originating from ducts are known as ductal carcinomas and those originating from lobules are known as lobular carcinomas. Breast cancer occurs in humans and also in other mammals. The majority of human causes occur in women and male breast cancer can also occur. Men can get breast cancer and the symptoms include breast lump, pain in breast area and tenderness. It is very rare disease in men and less than 2% of all cancer occurs in men has been discussed by Sheshadri & Kandaswamy (2005). Every year millions of women and thousands of men were diagnosed with breast cancer in all around the world. The term breast cancer refers to a malignant tumor that has developed from cells in the breast.

Breast cancer begins in the cells of the lobules, which are the milk producing glands and passage that drain milk from the lobules to the nipple, it can also begin in the stromal tissues, which include the fatty and fibrous connective tissues of the breast. Cancer cells can invade nearby healthy breast tissues and make their way in to the lymph nodes. If cancer cells get in to the lymph nodes, then they have pathway in to other parts of the body.

The description of the breast cancer was discovered in Egypt approximately in 1600 BC. For centuries, physicians concluded that, there is no treatment for this disease. Greeks believed that, imbalances in the fundamental fluids that controlled the body is the main cause for breast cancer. Also, patients treated it as divine punishment. In eighteenth century peoples believed that this disease is because of sexual activity. In the nineteenth century, the Scottish surgeon John Rodman said that fear of cancer caused cancer and that this anxiety learned by example for the mother, accounted for breast cancer tendency to run in families, most women had died too young to have developed breast cancer. Additionally, early and frequent
child bearing and breast feeding probably reduced the rate of breast cancer
development in those women who did survive to middle age. Breast cancer
prognosis and survival rates vary depending on the type of cancer, stage of
cancer, treatment and health condition of patient has been discussed by Zhang
et al 2002. Survival rates in the western countries are much higher because of
their treatment methods. Report of cancer survival in England shows more
than eight out of ten women in England diagnosed with breast cancer survive
for at least five years. In developing countries the survival rates were much
poorer in last decade and it is good now a days.

Type of treatment depends on the size, stage and rate of growth of
breast cancer. Treatment includes biopsy, surgery, medicines like hormonal
therapy, chemotherapy and radiation therapy. Many chemotherapy regimens
are given in addition to the surgery for increasing the possibility of cure.
Radiation therapy is used after surgery and substantially improves local
relapse rates and to increase the survival rate in many circumstances. Breast
cancer accounts for more percentage of all cancer for women. It is more than
hundred times more common in women than in men and a report from
national cancer institute of U.S. mentioned, men tend to have poorer survival
rate due to delays in detection and diagnosis.

Tumor can be benign (not dangerous) or malignant (dangerous). Benign
tumors are not considered cancerous because their cells are closed to
normal in appearance, they grow slowly and they do not invade nearby tissues
Cancer symptoms when it is grown are breast lump or hard lump in the armpit which has uneven edges and different from rest of breast tissue. A report from cancer fact and figures of American cancer society said that, more than 80% of breast cancer cases are discovered when the women feels a lump as shown in Figure 1.1 (a). Indications of breast cancer other than a lump may include difference in both of the breasts in size, shape and one breast becoming larger or lower.

![Figure 1.1 Early signs of breast cancer](image)

Next noticeable symptoms are skin dimpling as shown in Figure 1.1 (b), change in skin color as shown in Figure 1.1 (c), pulling in of nipple or inverted nipple as shown in Figure 1.1 (d) and fluid leaks out of nipple as shown in Figure 1.1 (e). The fluid may be bloody, yellow, green and so on like Whirley beaded string dimpling of atrial hypertrophy swelling in the
some women may asked to continue medications and will be instructed to test their blood frequently, mammograms and other tests required after the treatment is over. Main factors involved in the causes of breast cancer are female sex and older age. Smoking tobacco for long term will increase the risk. Aceves et al (2005) studied that, high alcohol intake and iodine deficiency may also the risk factors. Other risk factors are lack of breast feeding, higher hormone levels, diet and obesity (Yager 2006). Radiation, chemicals and pesticides have also been the risk factors. Some genetic susceptibility probably is the risk factors for most cases. Breast cancer occurs because of an interaction between the environment and a defective gene has been studied by Gage et al (2012). Normal cells divided as many times as needed and stop. They attach to other cells and stay in correct place in tissues. Cells become cancerous when mutations destroy their ability to stop dividing, attach to other cells and to stay where they belong and deposited are called as microcalcification. This makes the risk of cancer development. Male and female, both the genders have breast tissue and the tissue may increase in size based on the level of hormones in the human. But sometimes men can develop real breast gland tissue because they take certain medicines or have abnormal hormone levels.

1.3 MAMMOGRAPHY

Mammography is the input and main aid required by the CAD system in detecting breast cancer. It has been obtained by positioning the breast between two plates that flatten the breast to spread out the tissue and help to identify any abnormal areas. Most standard mammographic process includes two views of each breast taken from different angles. This way the breast can be compared and checked for abnormalities. Breast cancer is a major killer disease for women and it happens to over 8% women during their lifetime. Every year, lots of women have lost their life because of breast cancer. From the study of Sheshadri & Kandaswamy (2005), it is clearly
noted that, breast cancer affects the women over 15 years, especially between 35-55 years and the men over the age of 40.

Early detection of breast cancer increases both survival rate and the treatment options. Screening mammography and x-ray imaging of the breast are currently the most effective tools for early detection of breast cancer. Screening mammographic examinations are performed on asymptomatic women to detect early, clinically unsuspected breast cancer and it reduced the mortality rate by 30-70% has been studied by Rangaraj M.Rangayyan et al (2007). The detection and diagnosis of breast cancer with mammography are composed of two steps. The first is asymptomatic screening, where suspicious areas in a mammogram are identified. And the second is diagnostic mammography where symptomatic women with an abnormal mammogram having abnormality like palpable lump have been examined. The examination has been done by using two special viewing methods of mammogram. They are ultrasound and Magnetic Resonance Imaging (MRI). Suspected areas obtained from mammogram are called breast lesions and they are described and reported according to the Breast Imaging Reporting And Data (BIRAD™) system discussed by Armando Bazzani et al (2000).

BIRADS™ is a mammography lexicon developed by the American College of Radiology (ACR), for the description of mammographic lesions. The BIRADSTM lexicon includes descriptors such as the margin of a mass and the distribution of calcification and it defines final assessment categories to describe the radiologist’s level of suspicion about the mammographic abnormality. BIRADS™ final assessment rating is an indicator of the possibility of malignancy has been tested by Harris Georgiou et al (2007). If a suspicious abnormality is detected, a diagnostic mammographic examination is carried out to decide the future course of action required. Based on the level of suspicion of the abnormality following the diagnostic examination, a
A recommendation is made for routine follow up, short term follow up or biopsy.

The ultimate aim of obtaining a diagnostic mammogram is to determine whether a woman should have a biopsy. If the biopsy result confirms malignancy, immediate chemotherapy and surgery to remove the cancerous area in the breast followed by required level of radiotherapy has to be done. Figure 1.2 shows the mammogram image mdb219 referred from Mammogram Image Analysis Society (MIAS) database. Breast region and non breast region are the two main components of a mammogram image. The breast region contains pectoral muscle and breast tissues with mass. The non breast region contains dark background and background object as indicated in the Figure 1.2.

![Figure 1.2 Components of a mammogram image](image)

Images in Figure 1.3 are the left side breast image numbered 57 and have been referred from Digital Database for Screening Mammography (DDSM) database. It shows the two important views of left side breast. Figure 1.3 (a) CranioCaudal (CC) view, which is a top to bottom view and Figure 1.3 (b) MedioLateral Oblique (MLO) view, which is a side view taken at an angle.
Ultrasound and MRI are the other tests that are performed additionally in specific situations where the mammography is not able to find the suspicious areas in the breast. Ultrasound is used to further evaluate abnormalities found in mammography. An ultrasound uses sound waves to create image and recognize irregularities in the breast. MRI utilizes a magnetic field to visualize the features of soft tissues with highly detailed
resolution, even in cases of dense breast. Unfortunately, not all cancers are visible on MRI and some cancers, such as ductal carcinoma in-situ and lobular carcinoma are more difficult to detect using MRI.

The molecular imaging techniques are effective in the early detection of cancer and can be used to help distinguish malignant and benign tumors through the cellular absorption of a tracing agent. Molecular Breast Imaging (MBI), also known as Breast Specific Gamma Imaging (BSGI), is a procedure that shows the metabolic activity of breast lesions. For molecular imaging, tests specific to breast imaging such as BSGI/MBI/ breast Scintigraphy and Positron Emission Mammography (PEM) reveal tumors in dense or surgically altered breasts. BSGI/MBI uses a small injection of radioactive material, from that the breast cancers are revealed as hot spots in the image. This is because cancerous cells have a higher rate of metabolic activity than normal cells and absorb more of the tracing agent at a faster rate than healthy tissue.

The approximate likelihood of cancer presence can be indicated by screening mammography and clinical breast examination. If a patient had a mammogram in the past, the radiologist will compare that old mammogram to the new one to look for changes. Many national health organizations recommend that women over the age of 40 receive annual mammography screenings. Women under the age of 40 with either a family history of breast cancer or other concerns about their personal breast health should consult their doctor about when to start screenings. It is widely recommended that women get a baseline mammogram at age 35 for comparison with all later images. During the procedure of mammography, the breast is compressed using a dedicated unit. Parallel plate compression evens out the thickness of breast tissue to increase image quality by reducing the thickness of tissue that x-rays must penetrate, decreasing the amount of scattered radiation because of
the scatter degrades image quality, reducing the required radiation dose and holding the breast still for preventing motion blur. In screening mammography, both CC view and MLO view of the breast are to be considered for detection has been tested by Aize Cao et al (2008). Diagnostic mammography may include these and other views, including geometrically magnified and spot compressed views of the particular area of concern has been discussed by Abdel-Mottaleb et al (1996). Deodorant, talcum powder or lotion may show up on the x-ray as calcium spots and women are discouraged from applying these on the day of their exam.

Mammography was typically performed with screen film cassettes in olden days. Now days, mammography is undergoing transition to digital detectors, known as digital mammography or Full Field Digital Mammography (FFDM). The first FFDM system was approved by the Food and Drug Administration (FDA) in the U.S. in 2000. This progress some years later than in general radiology due to several factors. They are, the higher spatial resolution demands of mammography, significantly increased expense of the equipment, concern by the FDA that digital mammography equipment demonstrate that it is at least as good as screen film mammography at detecting breast cancers without increasing breast dose or the number of women recalled for further evaluation.

Some foreign countries, in order to encourage the use of mammograms as a screening measure for breast cancer, a number of hospitals, cancer centers and other healthcare groups have started mobile mammography vans to bring affordable, accessible and convenient mammograms to their communities. In the past several years, the process has become quite formalized. It generally consists of screening mammography, diagnostic mammography and biopsy when necessary, often performed via stereotactic core biopsy or ultrasound guided core biopsy. After a screening
mammogram, some women may have areas of concern which can’t be resolved with only the information available from the screening mammogram. They would then be called back for a diagnostic mammogram. This phrase essentially means a problem solving mammogram. During this session, the radiologist will be monitoring each of the additional films as they are taken by a technologist. Depending on the nature of the finding, ultrasound may often be used at this point, as well. Changes in breast tissue makes very difficult to differentiate the normal tissue and cancerous tissue. The breast of young women is composed of firm tissue. In mammography, this firm tissue appears as bright white spot. Cancerous tissues are also appearing as bright white dense spots on a mammogram. So, finding a cancer in a dense breast is tough task. Due to their age factor, breast density decreases and tissue becomes fattier. Fatty tissue appears dark on a mammogram and making it easier to view the bright white areas of cancer cells.

Common methods for evaluation of breast start with some additional imaging techniques other than mammography and then proceed to biopsy for extraction of tissues if necessary. Biopsy is the first process of diagnosis and it is used to extract the small part of breast tissue from the suspected breast area. The removal of a section of breast lump called core biopsy and removal of entire lump is called excisional biopsy has been studied by Yu et al (2010).

All type of breast cancer is easy to diagnose by simple microscopic analysis of the biopsy. If these tests failed to conclude, a test which confirms the presence of cancer by extract a small portion of fluid from the lump is called Fine Needle Aspiration and Cytology (FNAC). If the extracted fluid is clear then there is no possibility of cancer, if it is bloody fluid then that may be sent for microscope examination for cancerous cells. All these tests together or any one of these tests can be used to diagnose breast cancer with a
good degree of detection accuracy. There are however, some of the rare types of breast cancer required specific laboratory tests.

The use of mammography as a screening tool for the detection of early breast cancer in otherwise healthy women without symptoms continues to be debated. Since from 1990, the death rate from breast cancer has decreased by almost 30% and points to studies in Sweden and the Netherlands that show two thirds of the decrease in cancer deaths is due to mammography screening (Kopans 2009). This result has to be seen against the negatives of errors in diagnosis, overtreatment and radiation exposure. The goal of any screening procedure is to examine a large population of patients and find the small number most likely to have a serious condition. These patients are then referred for further testing. Thus a screening exam is not intended to be definitive. It is intended to have sufficient sensitivity to detect a usual proportion of cancers. The cost of higher sensitivity is a large number of results that would be regarded as suspicious in patients without disease.

The patients without disease who are called back for further testing from a screening session are sometimes referred to as False Positive (FP). There is a trade off between the number of patients with disease found and the much longer number of patients without disease that must be rescreened. Research of Brewer et al (2007) shown that FP mammograms may affect women’s well being and behaviour. Some women who receive FP results may be more likely to return for routine screening or perform breast self examination more frequently. However, some women who receive FP results become anxious, worried and distressed about the possibility of having breast cancer, feelings that can last for many years. FP also means greater expense, both for the individual women and for the screening program. Since follow up screening is typically much more expensive than initial screening, more FP that must receive follow up means fewer women may be screened for a given
amount of money. Thus as sensitivity increases, a screening program will cost more, or be able to screen a smaller number of women.

Mammograms also have a rate of missed tumors or False Negative (FN). Accurate data regarding the number of FN are very difficult to obtain, simple because mastectomies cannot be performed on every woman who has had a mammogram to determine the FN rate accurately. Estimates of the FN rate depend on close follow up of a large number of patients for many years. This is difficult in practice, because many women do not return for regular mammography making it impossible to know if they ever developed a cancer. Researchers have found that breast tissues are denser among younger women, making it difficult to detect tumors. For this reason, FN is twice as likely to occur in premenopausal mammograms. This is why the screening program in the UK does not start calling women for screening mammograms until the age of 50. The importance of these missed cancers is not clear, particularly if the woman is getting yearly mammograms. Research on a closely related situation has shown that small cancers that are not acted upon immediately, but are observe over periods of even several years, will have good outcomes.

The radiation exposure associated with mammography is a potential risk of screening. The risk of exposure appears to be greater in younger women. Feig & Hendrick (1997) investigated that, the largest study of radiation risk from mammography concluded that for women 40 years of age or older, the risk of radiation induced breast cancer was minuscule, particularly compared with the potential benefit of mammographic screening, with a benefit to risk ratio of 48.5 lives saved for each life lost due to radiation exposure. Organizations such as the national cancer institute and preventive task force of U.S. take such risks into account when formulating screening guidelines. The majority of health experts agree that the risk of breast cancer for asymptomatic women under 35 years is not high enough to
warrant the risk of radiation exposure. For this reason and because the radiation sensitivity of the breast in women under the age of 35 is possibly greater than in older women, most radiologists will not perform screening mammography in women under the age of 40.

However, if there is a significant risk of cancer in particular patient, mammography may still be important. Often, the radiologist will try to avoid mammography by using ultrasound or MRI imaging. The statistics about mammography and women between the ages of 40 and 55 are the most contentious. An estimate of this over diagnosis is 10 breast cancers diagnosed and unnecessarily treated per life saved when 2000 women are screened for 10 years. While screening between 40 and 50 is still controversial, the preponderance of the evidence indicates that there is some small benefit in terms of early detection. Currently, the American cancer society, the national cancer institute and the ACR encourage mammograms every two years for women ages 40 to 49. In 2012, the U.S. preventive services task force recommended that screening of those ages 40 to 49 be based on individual’s risk factors and that screening should not be routine in this age group. Their report says that the benefits of screenings before the age of 50 don’t outweigh the risks. While the cost of mammography is relatively low, its sensitivity is not ideal, with reports listing the range from 45% to about 90% depending on factors such as the density of the breast. However, the use of alternate technologies for breast evaluation is still in development.

From the elaborate study, it is concluded that the mammography still remains the first and best step for breast cancer screening and diagnosis, even though it is less accurate in patients with dense breast tissue. Hence, to overcome this inability of mammography, many testing options like ultrasound for the effective detection and diagnosis of breast cancer has been practiced. But still mammography proved that, it is one of the best aids for the
radiologists in detection and diagnosis of breast cancer only when it is supported by preprocessing techniques (Noise removal and enhancement techniques). Mammogram is examined for selecting the Region of Interest (RoI), which is the successful detection of suspicious area. Detection algorithm plays major role in concluding about the suspicious area whether cancerous (malignancy) or non cancerous (benign). This thesis is about discussion of techniques used in CAD of breast cancer with the help of mammogram image.

1.4 TREATMENT AND SURGERY

After successful detection of cancer using the mammogram, treatment plays major role in saving one’s life having breast cancer. Radiation therapy and surgery are the local treatments for breast area itself. The treatment to be given for the entire body is called systemic treatment which is also called as Chemotherapy. Most women need a combination of these treatments. Physical removal of the cancer area is called the surgery. Surgery for removing breast area has been done from eighteenth century. Standard surgeries are the removal of whole breast called Mastectomy, one quarter of the breast removal called Quadrantectomy and small part of the breast removal called Lumpectomy.

Drugs used after and in addition to surgery are called adjuvant therapy. Chemotherapy or other types of therapy prior to surgery are called neoadjuvant therapy. The French surgeon Jean Louis Petit (1674- 1750) and later the Scottish surgeon Benjamin Bell (1749-1806) were the first to remove the lymph nodes, breast tissue and underlying chest muscle. There are currently three main groups of medications used for adjuvant breast cancer treatment studied by Christoyianni et al (2002). They are hormone blocking therapy, chemotherapy and monoclonal antibodies. Chemotherapy treatment uses medicine to weaken and destroy cancer cells in the body, including cells
at the original cancer site and any cancer cells that may have spread to another part of the body.

Chemotherapy or chemo is a systemic therapy, which means it affects the whole body by going through the bloodstream. In many cases, a combination of two or more medicines will be used as chemotherapy treatment for breast cancer. And it is used to treat early stage invasive breast cancer to get rid of any cancer cells that may be left behind after surgery and to reduce the risk of the cancer coming back. Then for advanced stage breast cancer to destroy or damage the cancer cells as much as possible. In some cases, chemotherapy is given before surgery to shrink the cancer area. It also has drawbacks of producing side effects. While destroying cancer cells, it also destroys some normal cells in the blood, mouth, intestinal tract, nose, nails, vagina and hair.

In radiation therapy or radiotherapy, high energy rays are used to damage cancer cells and stop them from growing and dividing. A specialist in radiation therapy is called a radiation oncologist. Like surgery, radiation therapy is also a local treatment and it affects cancer cells only in the treated area. Radiation can come from a machine and called as external radiation. It can also come from a small container of radioactive material placed directly into or near the tumor and called as internal radiation. Some patients receive both kinds of radiation therapy. With radiation therapy, the side effects depend on the treatment dose and the part of the body that is treated. The most common side effects are tiredness, skin reactions such as a rash or redness, permanent pigmentation and scarring in the treated area. This therapy can also cause a decrease in the number of white blood cells that help to protect the body against any infection.

Hormonal therapy medicines treat hormone receptor positive breast cancers in two ways. The first way is by lowering the amount of hormone
estrogen in the body and blocking the action of estrogen on breast cancer cells. Most of the estrogen in women’s bodies is made by the ovaries. Estrogen makes hormone receptor positive breast cancers grow. So reducing the amount of estrogen or blocking its action can reduce the risk of early stage hormone receptor positive breast cancers coming back after surgery. The second way by providing hormonal therapy medicines can also be used to help shrink or slow the growth of advanced stage or metastatic hormone receptor positive breast cancers.

1.5 PREVENTION

Women may reduce their risk of breast cancer by maintaining a healthy weight, drinking less alcohol, being physically active and breastfeeding their children. These modifications in their life style might prevent 38% of breast cancers in the US, 42% in the UK, 28% in Brazil, 20% in China and 17% in India. Moderate exercise such as brisk walking and good diet at all age groups including postmenopausal women can also prevent them from breast cancer. Screening is also main prevention method involves in helping to know the presence of breast cancer.

Finding abnormal tissue or cancer as early as possible is crucial, as it may be easier to treat. By the age of 20, women should self examine their breasts every month for any changes in breast tissue. If something that seems unusual, then immediate discussion with a doctor is required as soon as possible. During a clinical breast exam, the physician will examine the breasts, feeling for any lumps or unusual tissue. The pink ribbon is a symbol to show support for breast cancer awareness and it is the most prominent symbol of breast cancer awareness too.

Screening is looking for cancer before any sign of symptoms. If abnormal tissue is found during a screening, additional diagnostic tests are
conducted to determine the presence of breast cancer. Regular screenings make it possible to catch potential problems early. Breast cancer screening refers to testing otherwise healthy women for breast cancer in an attempt to achieve an earlier diagnosis under the assumption that early detection will improve outcomes.

There are three important types of screenings to help women catch breast cancer early. They are clinical and self breast exams, mammography and genetic screening, ultrasound and magnetic resonance imaging. A clinical or self breast exam involves feeling the breast for lumps or other abnormalities. Clinical breast exams are performed by health care providers, while self breast exams are performed by the person themselves. Evidence does not support the effectiveness of either type of breast exam, as by the time a lump is large enough to be found it is likely to have been growing for several years and thus soon be large enough to be found without an exam. Mammographic screening for breast cancer uses x-rays to examine the breast for any uncharacteristic masses or lumps. During a screening, the breast is compressed and a technician takes photos from multiple angles. A general mammogram takes photos of the entire breast, while a diagnostic mammogram focuses on a specific lump or area of concern.

1.6 BREAST CANCER STATISTICS

Breast cancer effects have been felt by U.S. peoples higher than the peoples in rest of the world. Hence, the breast cancer statistics of U.S. and India has been discussed in this thesis. The statistics discussed in this section have been referred from official publication of the Population Based Cancer Registry (PBCR) and the Hospital Based Cancer Registry (HBCR) of the National Cancer Registry Program (NCRP) in India.

In 2012, 230,480 new cases of invasive breast cancer were diagnosed in women in the U.S. along with 57,650 new cases of non invasive
breast cancer and about 2,140 new cases of invasive breast cancer were diagnosed in men in 2012. A man’s lifetime risk of breast cancer is about 1 in 1,000. Breast cancer incidence rates in the U.S. decreased by about 2% per year since from 1990.

Many women in the U.S. were expected to die in 2013 from breast cancer, though death rates have been decreasing since 1990 especially in women under 50. These decreases are thought to be the result of treatment advances, earlier detection through screening and increased awareness. For women in the U.S., breast cancer death rates are higher than those for any other cancer, besides lung cancer and skin cancer. In U.S., 30% of cancers diagnosed were breast cancers and most commonly diagnosed cancer among women. White women are slightly more likely to develop breast cancer than African American women.

However, in women under the age of 45, breast cancer is more common in African American women than white women. Overall, African American women are more likely to die of breast cancer. Asian, Hispanic and Native American women have a lower risk of developing and dying from breast cancer. In 2011, there were more than 2.6 million breast cancer survivors in the U.S. and this rate doubles almost because of the current development in early detection and treatment. A woman’s risk of breast cancer approximately doubles if she has a first degree relative like mother, sister and daughter who has been diagnosed with breast cancer. About 15% of women who get breast cancer have a family member diagnosed with it. Sotiriou & Pusztai (2009) investigated that, about 5-10% of breast cancers can be linked to gene mutations inherited from one’s mother or father.

Women with gene mutations have up to an 80% risk of developing breast cancer during their lifetime and they are more likely to be diagnosed at a younger age. In men, about 1 in 10 breast cancers are believed to be due to
gene mutations. About 85% of breast cancers occur in women who have no family history of breast cancer. These occur due to genetic mutations that happen as a result of the aging process and life in general, rather than inherited mutations. The most significant risk factors for breast cancer are gender that is female and the age that is growing older.

Breast cancer is the commonest cancer in women in India. Its effect has been felt in cities, towns and villages. In Mumbai, it accounts to about 30.1% of all cancers in women. Approximately 2000 to 2500 new cases of breast cancer are detected every year in Mumbai and the number is ever rising. Also, the numbers in this registry tend to be lower than the real incidence, because a majority of breast cancer patients are managed at small hospitals and nursing homes, which may not be included in the count. Hence the actual number is higher, may be even 2 to 3 times. A few decades back, breast cancer was much more common after the age of 50.

The number of young women suffering from this disease was lesser and almost 65% to 70% patients were above 50 years and only 30 to 35% women were below 50 years of age. However, presently, breast cancer is more common in the younger age group and 52% of all cases are below 50 years of age. A significant number of patients are below 30 years. There has been a significant rise in the number of breast cancer cases as compared to earlier in 1980’s, breast cancer accounted for about 20% of all cancers in women. And it is reaching almost 30% in 2013.

In New Delhi, a few decades back, breast cancer was much more common after 50 years of age and the number of young women suffering from this disease was lesser. Almost 65% to 70% patients were above 50 years and only 30 to 35% women were below 50 years of age. Presently, breast cancer is more common in the younger age group and 50% of all cases are in the 25 to 50 years age group.
In Chennai, breast cancer accounts to about 26.5% of all cancer in women. About 48% of women suffering from breast cancer are in the age of 25 to 50. So similar to other cities, breast cancer is more common in younger age groups. There has been a significant rise in the number of breast cancer cases as compared to earlier in 1980’s, breast cancer accounted for about 16% of all cancers in women. But, it is reaching almost 30% in 2013.

In Bengaluru, breast cancer accounts to about 26.9% of all cancers in women. About 54% of women suffering from breast cancer are in the 25 to 50 year age group. So similar to other cities, breast cancer is more common in younger age groups. There has been a significant rise in the number of breast cancer cases as compared to earlier in 1983, breast cancer accounted for about 14% of all cancers in women, whereas now, it is reaching almost 26%.

In the U.S., breast cancer is the most common cancer amongst women and 1 in 8 women in the U.S. have a chance of developing breast cancer in their life time. In India, the overall incidence of breast cancer is less as compared to the U.S. and it is around 1 in 30. In the year 2008, there were about 1,82,000 breast cancer cases reported in the U.S., whereas in India, 1,15,000 new cases were diagnosed. This implies that, though, because of India’s population, the percentage of total women affected seems less, the breast cancer burden in India has almost reached about two times that of the U.S. and is steadily rising.

Ratio between the incidence and mortality gives a good idea of how long do the patients survive on an average. For U.S., 1,82,000 new cases and about 40,000 deaths with a ration of about 1 death for 4.5 new cases detected. For India, 1,15,000 new cases and 53,000 deaths. This means a ratio of about 1 death for every 2 cases detected. As the treatment of a cancer evolves, as more and more early cancers are detected, the mortality becomes less and less, though the incidence may rise, leading to a change in ratio. India’s ratio
tells, it has a long way to go. Since more patients turn up in later stages, they do not survive long irrespective of breast treatment they may get and hence the mortality is fairly high. Hence, it is recommended to use the facility of CAD system in identifying the presence of cancer for the early treatment.

1.7 COMPUTER AIDED DETECTION AND DIAGNOSIS

CAD systems have been developed to aid radiologists in detecting mammographic lesions that may indicate the presence of breast cancer. In the recent years, the CAD has become a part of the routine clinical work for the detection of breast cancer in mammograms at many screening sites and hospitals (Kunio Doi 2007). These systems act only as a second reader and the final decision is made by the radiologist. Recent studies have also shown that CAD systems, when used as an aid, have improved radiologist’s accuracy of detection of breast cancer. The impact of a CAD system on the detection efficiency of an experienced and relatively inexperienced radiologist has been investigated and in both cases the CAD system has proven an efficient support tool for the detection, though, its autonomy remains in doubt. Therefore, due to the complexity of the CAD autonomy, automatic or semi automatic systems still play only the role of a signalling tool for the radiologist.

Mammogram images are the input to CAD system for the detection of breast cancer. It is important to realize that mammographic image analysis is also an extremely challenging task for a number of reasons. First, since the efficacy of CAD system can have very serious implications, there is a need for near perfection. Second, the large variability in the appearance of abnormalities makes this a very difficult image analysis task. Finally, abnormalities are often occluded or hidden in dense tissue, which makes detection difficult.

Some of the important signs of breast cancer that radiologists wanted to know are the clusters of microcalcifications, masses and
architectural distortions. Calcifications are tiny deposits of calcium, which appear as small bright spots on the mammogram. They are characterized by their type and distribution properties. A mass is defined as a space occupying lesion seen in at least two different projections. Masses are described by their shape and margin characteristics. It is more difficult to detect masses than microcalcifications because their features can be obscured or similar to normal breast parenchyma. Brzakovic et al (1990) investigated that the masses are quite subtle and often occurred in the dense areas of the breast tissue. It has smoother boundaries than microcalcifications and has many shapes. The circumscribed ones usually have distinct boundaries, 2-30mm in diameters and are high density radiopaque. The speculated ones have rough, star shaped boundaries and the lobulated ones have irregular shapes.

Lai et al (1989) investigated about the classification of masses as benign and malignant. In general, masses with radiopaque and more irregular shapes are usually malignant and those combined with radiolucent shapes are benign. A mammogram is basically distinct with four levels of the intensities. They are background, fat tissue, breast parenchyma and calcifications with increasing intensity. Women who have breast cancers can easily get contra lateral cancers in the other side breast. Asymmetry of breast parenchyma between the two sides has been one of the most useful signs for detecting primary breast cancer. Architectural distortions are the effects produced by mammographic image itself. It’s because of noise and unwanted particles added in the mammogram image while the scanning process is done.

Detection of suspicious abnormalities is a repetitive and fatiguing task. For every thousand cases analyzed by a radiologist, only 3 to 4 are cancerous and thus an abnormality may be overlooked. As a result, radiologists fail to detect 10-30% of cancer. Approximately two thirds of these false negative results are due to missed lesions that are evident retrospectively (Robert M.Nishikawa 2007). Due to the considerable amount of overlap in the appearance of malignant and benign abnormalities,
mammography has a Positive Predictive Value (PPV) of less than 35% found by Rangaraj et al (2007), where the PPV is defined as the percentage of lesions subjected to biopsy that were found to be cancer. Thus, a high proportion of biopsies are performed on benign lesions. Avoiding benign biopsies would spare women from anxiety, discomfort and expense.

Masses appear as dense regions of varying sizes and properties. Figure 1.4 has been referred from Christoyianni et al (2002) and has been characterized as circumscribed in Figure 1.4 (a), speculated in Figure 1.4 (b) and ill defined in Figure 1.4 (c). On the other hand, microcalcifications in Figure 1.4 (d) appear as small bright arbitrarily shaped regions on the large variety of breast texture background. Finally, asymmetry in Figure 1.4 (e) and architectural distortion in Figure 1.4 (f) are also very important and difficult to detect. The great variability of the mass appearance along with the other abnormalities in digital mammograms is the main obstacle of building a unified mass detection method has been investigated by Christoyianni et al (2002).

![Figure 1.4 Types of breast cancer](image)

Computer aided methods in the field of digital mammography are divided into two main categories. They are CAD methods that are capable of pinpointing Regions of Suspicion (RoS) or RoI in mammograms for further analysis by expert radiologist and CADi methods which are capable of making a decision whether the examined RoS consist of abnormal or healthy
tissue and distinguishing between malignant and benign RoS. It is important to restate that one of the aims of computerized classification is to reduce the number of women undergoing biopsy for benign disease that is to increase the PPV without reducing the sensitivity of breast cancer detection. Specifically, CAD system for detecting masses or microcalcifications in mammograms have already been used and proven to be a potentially powerful tool.

However, the development of methods for recognizing the identity of a RoS, especially in the case of all kinds of abnormalities is still very limited. Processings involved in general CAD and CADi systems are given in Figure 1.5.

![Figure 1.5 Processing of general CAD and CADi system](image)

The goal of CAD is to improve radiologist’s performance by indicating the sites of potential abnormalities, to reduce the number of missed lesions and by providing quantitative analysis of specific regions in an image to improve diagnosis. CAD systems typically operate as semi automated
systems that indicate lesion location. Individual human observes overlook different findings and it has been shown that double reading increases the detection rate of breast cancer by 5-15%. Many CAD systems for mammographic screening have been studied by Qian et al (1999).

Most CAD algorithms consist of two stages. In stage one the aim is to detect suspicious regions at a high sensitivity. In stage two, the aim is to reduce the number of False Positive (FP) or wrong selections without decreasing the sensitivity drastically. The steps that are involved in designing algorithms for both stages are shown in Figure 1.5. They are preprocessing (Noise removal and enhancement), segmentation and detection of abnormality.

In radiology, CAD and CADi are the procedures involved to assist doctors in the interpretation of medical images. Imaging techniques in X-ray, MRI, and Ultrasound diagnostics yield a great deal of information, which the radiologist has to analyze and evaluate comprehensively in a short time. CAD is a relatively interdisciplinary technology combining elements of artificial intelligence and digital image processing with radiological image processing. A typical application of CAD in this thesis is the detection of breast cancer. For instance, some hospitals use CAD to support preventive medical checkups in mammography (diagnosis of breast cancer), the detection of many cancers in human body. CAD systems are usually confined to marking conspicuous structures and sections.

CADi systems evaluate and analysis the conspicuous structures. For example, in mammography CAD highlights mircocalcificiation clusters and hyperdense structures in the soft tissue. This allows the radiologist to draw conclusions about the condition of the pathology. Another application is CADq, which quantifies (q) the size of a tumor or the tumor's behavior. Computer Aided Simple Triage (CAST) is another type of CAD, which
performs a fully automatic initial interpretation and triage of studies into some meaningful categories. CAST is particularly applicable in emergency diagnostic imaging, where a prompt diagnosis of critical, life threatening condition is required. At the present stage of the technology, CAD cannot and may not substitute the doctor, but rather plays a supporting role. The radiologist is always responsible for the final interpretation of a medical image.

Steps include preprocessing for reduction of artifacts, image noise reduction and leveling of image quality for clearing the image’s different basic conditions. Segmentation for differentiation of different structures in the image. From the RoI, every detected region is analyzed individually for special characteristics like compactness, form, size and location, reference to close by structures, average grey level value analyze within a RoI and proportion of grey levels to border of the structure inside the RoI. Evaluation and classification after the structure is analyzed, every RoI is evaluated individually. This procedure includes nearest neighbor rule, minimum distance classifier, cascade classifier, Bayesian classifier, multilayer perception, Radial Basis Function (RBF) network and Support Vector Machine (SVM). If the detected structures have reached a certain threshold level, they are highlighted in the image for the radiologist. A good CAD should have reasonable sensitivity and specificity. That is, CAD systems seek to highlight suspicious structures. Today’s CAD systems cannot detect 100% of pathological changes. The hit rate (sensitivity) can be up to 90% depending on system and application. A correct hit is termed a True Positive (TP), while the incorrect marking of healthy sections constitutes a FP. The less FP indicated, the higher the specificity. A low specificity reduces the acceptance of the CAD system because the user has to identify all of these wrong hits. In CAST system the FP rate must be extremely low (less than 1 per examination) to allow a meaningful study triage. A good CAD should also
have absolute detection rate of the radiologist is an alternative metric to sensitivity and specificity.

CAD is used in screening mammography (X-ray examination of the female breast). Screening mammography is used for early detection of breast cancer. CAD is especially established in US and the Netherlands and is used addition to human evaluation, usually by a radiologist. The first CAD system for mammography was developed in a research project at the University of Chicago. However, while achieving high sensitivities, CAD system tend to have low specificity and the benefits of using CAD remain uncertain. Some studies suggest a positive impact on mammography screening programs, but others show no improvement. Main requirement of CAD tool in breast cancer analysis is mammography. It is the process of using low energy X-rays to examine the human breast and is used as a diagnostic and a screening tool. The goal of mammography is the early detection of breast cancer, typically through detection of characteristic masses and microcalcifications. Generally the cause of the unusual appearance is found to be benign. If the cause cannot be determined to be benign with sufficient certainty, a biopsy will be recommended. The biopsy procedure will be used to obtain actual tissue from the site for the pathologist to examine microscopically to determine the precise cause of the abnormality. In the past, biopsies were most frequently done in surgery, under local or general anesthesia. CAD systems are being tested to decrease the number of cases of cancer that are missed in mammograms. Destounis et al (2004) have tested the CAD twice and found that, from the first test, the CAD system identified 71% of the cases of cancer that had been missed by physicians. However, the CAD system also flagged twice as many non cancerous masses than the physicians did. In a second test of a larger set of mammograms, CAD recommended six biopsies that physicians did not. All six turned out to be cancers that would have been missed.
Generally, CAD systems in screening mammography have poor specificity and compare poorly to double reading has been studied by Taylor et al (2005). While data are accumulating suggesting that CAD can find a few additional cancers, this should be put in perspective. Welch & Frankel (2011) studied and concluded that, women whose breast cancer was detected by screening mammography before the appearance of a lump or other symptoms commonly assume that the mammogram saved their lives.

Welch & Frankel (2011) investigated that, the screening mammography produces no benefit to any of the remaining 87% to 97% of women. Commercially available CAD and CADi systems had been approved by FDA of United States (U.S.) to aid radiologists for the detection and diagnosis of mammographic abnormalities is presented in Table 1.1.

Table 1.1 Some of the existing CAD and CADi systems with their detection accuracy

<table>
<thead>
<tr>
<th>S.No.</th>
<th>CAD&amp;CADi System</th>
<th>Detection Accuracy of Microcalcification</th>
<th>FP</th>
<th>Detection Accuracy of Masses</th>
<th>FP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R2 Technology image checker</td>
<td>98.5%</td>
<td>0.74</td>
<td>85.7%</td>
<td>1.32</td>
</tr>
<tr>
<td>2</td>
<td>Intelligent system software</td>
<td>89.3%</td>
<td>2.32</td>
<td>87.4%</td>
<td>3.32</td>
</tr>
<tr>
<td>3</td>
<td>CADi Medical systems</td>
<td>85%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>GPCALMA</td>
<td>96%</td>
<td>4.2</td>
<td>80%</td>
<td>0.3</td>
</tr>
</tbody>
</table>

FP- False Positive (Sensitivity)
1.8 OBJECTIVE AND METHODOLOGY

From the detailed study, it is noted that, the CAD system plays major role in early identification and detection of breast cancer. But the detection accuracy improvement is always a major requirement of a good CAD.

- Hence, the main objective of this thesis is to propose new methods for the enhancement of main methods (preprocessing, segmentation and detection) involved in the existing CAD system.
- Curvelet based noise removal (UnequiSpaced Fast Fourier Transform - USFFT) and Modified Local Range Modification (MLRM) contrast enhancement methods have been proposed for preprocessing of mammogram images.
- The MLRM has been combined with Laplacian of Gaussian (LoG) and utilized for segmentation of cancer area in mammogram images.
- Fuzzy C Means clustering (FCM) has been used for finding microcalcification clusters.
- Statistical moments calculation has been done for classification.

The detailed descriptions of these methods are available in the following chapters.

1.9 MAMMOGRAM DATABASES

Mammogram images used by researchers for the purpose of analysis, detection and diagnosis of breast cancer and its treatment have been taken from the databases given in Table 1.2. Among this, mammogram images from the Mammogram Image Analysis Society (MIAS) and the Digital Database for Screening Mammography (DDSM) are widely used by the researchers.
Table 1.2 Mammogram databases

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Database Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DDSM</td>
<td>Digital Database for Screening Mammography (DDSM) was created by Massachusetts General Hospital, the University of South Florida and Sandia National Laboratories.</td>
</tr>
<tr>
<td>2</td>
<td>LLNL/UCSF</td>
<td>Database was created by Lawrence Livermore National Laboratories (LLNL) and the Radiology Department at the University of California at San Francisco (UCSF).</td>
</tr>
<tr>
<td>3</td>
<td>MIAS</td>
<td>Database was created by the Mammographic Image Analysis Society (MIAS), United Kingdom.</td>
</tr>
<tr>
<td>4</td>
<td>Nijmegen</td>
<td>Database was created by the National Expert and Training Centre for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, the Netherlands.</td>
</tr>
<tr>
<td>5</td>
<td>CALMA</td>
<td>Computer Aided Library in Mammography (Italian National screening program of Mammography).</td>
</tr>
<tr>
<td>6</td>
<td>HLMCCIP</td>
<td>H Lee Moffitt Cancer Center Imaging Program.</td>
</tr>
<tr>
<td>7</td>
<td>MG</td>
<td>MammoGrid (Pan European MammoGrid)</td>
</tr>
</tbody>
</table>

1.10 ORGANIZATION OF THESIS

Information about the types, symptoms, treatments and prevention of breast cancer has been discussed. Mammographic imaging and other techniques involved in imaging the cancer area has been discussed. Information about the CAD and CADi has been discussed. Mammogram image databases have been collected and tabulated. Many detection techniques have been studied and included in this thesis. Performance of
proposed techniques has been analyzed by plotting Receiver Operating Characteristics (ROC) curve. Chapter one consists of information about breast cancer, types, detection, treatment, prevention, information about CAD and CADi. Chapter two is preprocessing and it consists of existing preprocessing techniques and proposed preprocessing techniques. Chapter three is segmentation and having information about existing edge detection methods, segmentation methods and proposed segmentation method. Chapter four consists of information about the detection methods involved in CAD. Results, comparisons and discussions of all the methods proposed in this thesis are provided in Chapter five.