Chapter

3. Formulation of Problem

3.1 Diagnosis without CAD

There are essentially two types of errors a radiologist can make in a cancer diagnosis without CAD systems. The first and foremost error is of (wrong) perception and it is the most common one. The radiologist simply does not see the evidence of suspicious lesion indicating disease on the mammogram. The second error is of (miss) interpretation. The radiologist sees the cancerous lesion on the mammogram, but interprets that lesion as something other than what actually it is. The technical issues such as imaging quality and human error are more taxing while radiologists’ interpretation during the diagnosis of breast cancer. In such a situation, the second opinion by another radiologist plays a vital role in final diagnosis. But it is expensive and time consuming.

The possibility of missing a cancer reduces by about 10-15 percent when mammograms are interpreted by two radiologists. This type of double reading is a routine practice at many mammography centres. Of course, a wide inter-observer variability (interpretation of the same image by two observers is different; especially while marking lesion contour) is observed while interpreting a high volume of mammograms. This variability is observed in both, accuracy of abnormality detection as well as uniformity [47]. Here, uniformity means, an observer interpret the image ‘I’ first time and after interpreting few images again if he interprets the image ‘I’ with the same way as that of first. This is not guaranteed. A few studies on breast cancer have shown that the diagnosis with mammography is susceptible to high false positives leading to unnecessary biopsies and false negatives resulting in missed cancer [48].

3.2 Diagnosis using CAD

The CAD systems can be envisaged as a second pair of eyes to the expert radiologists if such CAD can identify and characterize the suspicious lesions on mammograms with accuracy as that of the radiologists [49]. The advancements in the machine learning algorithms including classifiers such as k-NN, SVM are useful in building
more efficient CAD systems based on merging of different computer-extracted image features.

Published data indicates that the breast cancer mortality has been declined in several countries [50]. This fall in mortality is mainly possible because of earlier detection, perhaps for large population base through CAD-assisted screening and timely treatment services. Especially, cancer detection through CAD is more advantageous because of reduction in overall interpretation duration. A vast appropriate training of CAD systems might improve diagnostic performance of cancer detection [51].

However, during screening mammography programs, the overall rate of false negatives is about 25% in young women in their 40s and 10% among older women. Mammographic findings also include false positives interpretations leading to biopsies which turn out to be benign or negative [52]. The single view CAD systems for mass detection have good diagnostic accuracy but less positive predictive value; and hence are not as accurate and reliable as the experienced radiologist [53]. Another research revealed that the decision support of CAD is of no use [54]. Moreover, the commercial CAD systems based on two views are not used in clinical practice yet.

### 3.3 Rationale and Motivation

Although the use of CAD is costly, its apparent effectiveness in detecting early stage breast cancer makes it a desirable tool. For this reason, the research on improving the efficiency and accuracy of CAD is still continued. There is scope for improvement in the sensitivity, specificity, accuracy etc. of the CAD systems to make them more effective and useful in clinical practice [55]. This improvement can be achieved through the algorithms for (i) drawing radiologists' attention to suspicious regions on a mammogram that could be missed; (ii) reducing the false positives and (iii) classifying the detected suspicious lesions as either benign mass or malignant tumors; avoiding unnecessary biopsies.

The CAD systems have potential to decrease the false positives and maintain the proper balance of sensitivity and specificity. The CAD systems are less time consuming, more economic than second opinion of radiologist. Single view diagnosis and two views diagnosis can be combined together to yield even better and reliable results in breast cancer diagnosis.
At present CAD applications have been developed for breast cancer detection using screening mammograms. The next-generation CAD systems are expected to go for further characterization and then classification of abnormal lesions. The CAD systems can help radiologists to improve their efficiency in terms of sensitivity, specificity, accuracy and consistency. The faster decisions from CAD can help them to increase their productivity.

In short, improvement in the detection performance of CAD with proper balance of sensitivity and specificity is crucial for accurate diagnosis and treatment. The CAD systems combining the advantages from single view and two views systems can be useful as a mammographic workstation tool for the radiologists. The radiologists are now enthusiastic about CAD systems and they are raising their interest in better decision support from CAD [56]. These facts have motivated the researchers to undertake the proposed research.

### 3.4 Gap Identified

The following problems are identified from the literature.

- Existing automatic techniques for accurate delineation of breast border, extraction of pectoral muscle on MLO view and localizing the nipple are not efficient and are open research problems.
- Existing methods of FFDM segmentation are not having proper balance of sensitivity and specificity to aid the radiologist in breast cancer diagnosis. Designing efficient algorithms for FFDM segmentation is yet an open research problem.
- Extracting the appropriate features of the suspicious lesions and classifying them with expected level of accuracy is another challenging task.
- Detection and diagnosis using ipsilateral views of mammograms with improved case based lesion detection and diagnosis at a comparable FP rate is still an open challenge.

### 3.5 Problem Statement

Considering the aforementioned aspects and gap identified, the present research work establishes the following basic hypothesis: “is it possible to create an automatic
method for detecting breast abnormalities to assist diagnosis of breast cancer using digital image database?"

The primary purpose of this research would be to design and develop algorithms for automatic detection of breast abnormalities indicating cancer and classify them as either benign mass or malignant tumor using full field digital mammograms.

3.6 Specific Objectives of Research

The work presented in this thesis focuses on making a useful contribution to improve the performance of breast cancer CAD systems. The major objectives of the work include design and development of algorithms for

i. Automatic and accurate delineation of breast border on CC / MLO views and extraction of pectoral muscle on MLO views in an efficient manner to reduce the computational burden on further image analysis stages.

ii. Segmentation of FFDM images to identify the abnormal lesions accurately with minimum possible false positives.

iii. Extracting the appropriate features of the suspicious lesions and classifying them as benign mass or malignant tumor with proper balance of sensitivity and specificity.

iv. Detection and diagnosis using ipsilateral views of mammograms with improved case based lesion detection and diagnosis at a comparable FP rate.

The work presented in this thesis focuses on the design and development of the algorithms for automatic detection and diagnosis of breast cancer using single and two views of FFDM images from Tata Memorial Centre, Mumbai.

3.7 Chapter Summary

This chapter presented the various issues of breast cancer detection with and without CAD revealed from literature study. The motivation for the research is then explained which is followed by clear problem definition. The detailed objectives are then enlisted. The next chapter covers the preprocessing task of the work.