Computer-Aided Detection and Diagnosis of Breast Abnormalities in Digital Mammography

Jelena Bozek, Kresimir Delac, Mislav Grgic
University of Zagreb, Faculty of Electrical Engineering and Computing
Department of Wireless Communication
Unska 3/XII, HR-10000 Zagreb, Croatia
E-mail: jelena.bozek@fer.hr

Abstract - Successful treatment of breast cancer depends on early detection and diagnosis of breast abnormalities and lesions. Mammography is the best available examination for the detection of early signs of breast cancer such as masses, calcifications, bilateral asymmetry and architectural distortion. Because of the limitations of human observers, computers have major role in detecting early signs of cancer. Wide range of features that define abnormalities and the fact that they are often indistinguishable from the surrounding tissue makes the computer-aided detection and diagnosis of breast abnormalities a challenge. This paper discusses breast lesions and their features. The paper briefly presents some of the developed computer-aided detection and diagnosis methods for each lesion.

Keywords - Digital Mammography, Breast Abnormalities, CAD

1. INTRODUCTION

Mammography is, at present, the best available examination for the detection of early signs of breast cancer and it can reveal pronounced evidence of abnormality, such as masses and calcifications, as well as subtle signs such as bilateral asymmetry and architectural distortion [1]. Mammography is a special type of x-ray imaging used to create detailed images of the breast. Mammography uses low dose x-ray, high contrast, high-resolution film and an x-ray system designed specifically for imaging the breasts [2].

During mammography, each breast is carefully positioned on a special film cassette and then gently compressed with a paddle. This compression flattens the breast so that the maximum amount of tissue can be imaged and examined. Special energy and wavelength of the x-rays allow them to pass through the compressed breast and create the image of the internal structures of the breast onto a film cassette positioned under the breast. Differences in absorption and the corresponding varying exposure level of the film create the images which can clearly show normal structures such as fat, fibroglandular tissue, breast ducts and nipples. Further, abnormalities such as masses, calcifications, architectural distortion and bilateral asymmetry are also visible. Fat appears as black regions on a mammogram and everything else (glands, connective tissue and abnormalities) appear as levels of white on a mammogram.

The advance in x-ray mammography is digital mammography. With digital mammography, the breast image is captured using a special electronic x-ray detector, which converts the image into a digital picture for review on a computer monitor. The digital mammogram is then stored on a computer. With digital mammography, the magnification, orientation, brightness and contrast of the image may be altered after the exam is completed to help the radiologist more clearly see certain areas [2]. Digital mammography has the potential to offer several advantages over traditional film mammography, including: faster image acquisition, shorter exams, easier image storage, easy transmission of images to other physicians and computer processing of breast images for more accurate detection of breast cancer [2].

To date, studies have shown that digital mammography is "comparable" to film mammography in terms of detecting breast cancer. Digital mammography did, however, perform significantly better than the film method in women less than 50 years of age, in those having heterogeneously dense or very dense breasts and premenopausal or perimenopausal women. The digital and film methods performed equally well in women age 50 years and older, those with fatty breasts or scattered fibroglandular densities and those who were postmenopausal [3].

In this paper, we describe lesions that are possible to detect and diagnose using developed computer-aided detection and diagnosis methods. Masses, calcifications, architectural distortion and bilateral asymmetry are defined with wide range of features and can indicate malignant changes but can also be a part of benign changes. Most of the features such as shape, margin, distribution, size etc. can be detected by using developed algorithms. However, there are some problems in detection and diagnose of breast abnormalities specific for
particular lesion. Some of the problems are visibility of the lesion, possibility to differ it from surrounding tissue and appropriate classification of the change as malignant or benign.

This paper is organized as follows. In Section 2, we briefly portray typical steps in computer-aided detection and computer-aided diagnosis algorithms. Standardized method BI-RADS for breast imaging reporting is described in Section 3. The BI-RADS system includes terms to describe lesions such as mass, calcifications, architectural distortion and bilateral asymmetry. Brief description of mass, calcifications, architectural distortion and bilateral asymmetry, methods to detect and diagnose each lesion and the problems in detection and diagnosis are presented in subsections 3.1, 3.2, 3.3 and 3.4, respectively. Finally, Section 4 concludes the paper.

2. COMPUTER-AIDED DETECTION AND DIAGNOSIS

The use of computers in processing and analyzing biomedical images allows more accurate diagnosis by a radiologist. Humans are susceptible to committing errors and their analysis is usually subjective and qualitative. Objective and quantitative analysis facilitated by the application of computers to biomedical image analysis leads to a more accurate diagnostic decision by the physician [4]. Computer-aided detection (CADe) is designed to provide the radiologist with visual prompts on series of mammograms. It works by marking a mammogram with marks that indicate regions where the detection algorithm recognizes a suspicious entity that warrants further investigation, thereby complementing the radiologists’ interpretation. Findings in a number of studies have demonstrated that CADe has the ability to detect and prompt mammographic signs of cancer with the potential to increase cancer detection rates by approximately 20% [5]. If a patient's medical history and radiologist's findings are taken into account, together with computer-aided detection data that provides diagnostic output, a computer-aided diagnosis (CADx) system exists. Sometimes, both computer-aided detection and computer-aided diagnosis are referred to as CAD.

In most developed CADe and CADx programs, there are some common steps that have to be fulfilled in order to find the suspect lesions. Typical steps needed for CADe and CADx program are depicted in Fig. 1.

Most detection algorithms consist of two stages. In stage 1, the aim is to detect suspicious lesions at a high sensitivity. In stage 2, the aim is to reduce the number of false positives without decreasing the sensitivity drastically. The steps that are involved in designing algorithms for both stages are shown in shadowed box in Fig. 1. In some approaches some of the steps may involve very simple methods or be skipped entirely. Most diagnosis algorithms (CADx) begin with a region of interest (ROI) containing the abnormality. Again, the steps typically involved in design such a system are shown in the same shadowed box in Fig. 1. The output of a CADx system may be the likelihood of malignancy or a management recommendation. Different research groups have worked on different components of the problem and human interaction may occur at various stages. For example, many CADx algorithms start with manually segmented ROIs.

**Fig. 1.** Typical scheme showing the main steps involved in the computer-aided detection (CADe) and computer-aided diagnosis (CADx) of mammographic abnormalities [6]

In the preprocessing step the breast is segmented in order to limit the search for abnormalities without undue influence from the background of the mammogram and some filtering or normalization is accomplished in order to improve the quality of the image and reduce the noise. The next step, feature extraction is one of the most important factors that affects the CAD performance. Basically, researchers have investigated two types of features: those traditionally used by radiologists (gradient-based, intensity-based and geometric features) and high-order features that may not be as intuitive to radiologists (e.g. texture features). Critical issue in CAD design is the choice of the best set of features for detecting or classifying the suspect lesions. The whole set of features may include redundant or irrelevant information. One feature taken alone might not be significant for classification but might be very significant if combined with other features. In order to decide which features are best suited for classification, feature selection is used. Feature selection is defined as selecting a smaller feature set...
subset of size $m$ from a set of $d$ features, that leads to the largest value of some classifier performance function [5]. Finally, a classification (false-positive reduction) step is performed, where on the basis of the mentioned features false signals are separated from the suspect lesions by means of a classifier. In the other words, the candidate lesions are first located and then further analyzed in a feature analysis and classification phase to determine the final classification of each candidate [5].

### 3. BI-RADS

The ACR (American College of Radiology) Breast Imaging Reporting and Data System (BI-RADS®) suggests a standardized method for breast imaging reporting [7]. Terms have been developed to describe breast density, lesion features and lesion classification. Depending on the amount of fibroglandular tissue, breast tissue seen on mammogram can be divided into four categories shown in Fig. 2.

![Fig. 2. Examples of mammograms, each of different category of breast tissue: (a) fat breast tissue, (b) scattered fibroglandular dense breast tissue, (c) heterogeneously dense breast tissue and (d) extremely dense breast tissue [8]](image)

The breast is almost entirely fat when there is less than 25% fibroglandular tissue. Scattered fibroglandular dense breast tissue has between 25% and 50% fibroglandular tissue and heterogeneously dense breast tissue has between 51% and 75% fibroglandular tissue. When the breast is consisting of more than 75% fibroglandular tissue the breast is extremely dense. In the latter case sensitivity of mammography exam is decreased and the diagnosis of malignant lesions is more difficult.

Many lesions (masses, calcifications, architectural distortion, bilateral asymmetry) are defined with wide range of features. The features determine lesions shape, size, distribution, margins etc. Some of the lesions can be easily overlooked because of the poor feature visibility.

One of the problems that appear in diagnosis of malignant lesions is incorrect classification of lesions. Final assessment and classification of mammograms is made using ACR BI-RADS categories [7]. A negative diagnostic examination is one that is negative, with a benign or probably benign finding (BI-RADS 1, 2 or 3) and a positive diagnostic examination is one that requires a tissue diagnosis (BI-RADS 4 or 5) or the one with biopsy proof of malignancy (BI-RADS 6). If the finding can not be assessed, an additional imaging evaluation and/or prior mammograms are needed for comparison (BI-RADS 0).

#### 3.1. Mass

A mass is defined as a space occupying lesion seen in at least two different projections [7]. If a potential mass is seen in only a single projection it should be called 'Asymmetry' or 'Asymmetric Density' until its three-dimensionality is confirmed. Masses have different density (fat containing, low density, isodense, high density), different margins (circumscribed, microlobular, obscured, indistinct, spiculated) and different shape (round, oval, lobular, irregular).

Fat-containing radiolucent and mixed-density circumscribed lesions are benign, whereas isodense to high-density masses may be of benign or malignant origin [9]. Benign lesions tend to be isodense or of low density, with very well defined margins and surrounded by a fatty halo, but this is certainly not diagnostic of benignancy. The halo sign is a fine radiolucent line that surrounds circumscribed masses and is highly predictive that the mass is benign.

Circumscribed (well-defined or sharply-defined) margins are sharply demarcated with an abrupt transition between the lesion and the surrounding tissue [10]. Without additional modifiers there is nothing to suggest infiltration. A mass with circumscribed margin is shown in Fig. 3(a). Lesions with microlobular margins have wavy contours. Obscured (erased) margins of the mass are erased because of the superimposition with surrounding tissue. This term is used when the physician is convinced that the mass is sharply-defined but has hidden margins. The poor definition of indistinct (ill defined) margins raises concern that there may be infiltration by the lesion and this is not likely due to
superimposed normal breast tissue. The lesions with spiculated margins are characterized by lines radiating from the margins of a mass shown in Fig. 3(b). A lesion that is ill-defined or spiculated and in which there is no clear history of trauma to suggest hematoma or fat necrosis suggests a malignant process [9].

Fig. 3. Examples of (a) circumscribed mass and (b) spiculated mass [11]

Mass detection and diagnosis

As it is already said, a typical benign mass has a round, smooth and well-circumscribed boundary. On the other hand, a malignant tumor usually has a spiculated, rough and blurry boundary. However, there exist atypical cases of macrolobulated or spiculated benign masses, as well as microlobulated or well-circumscribed malignant tumors [1]. The detection of masses requires the segmentation of all possible suspicious regions, which may then be subjected to a series of tests to eliminate false positives.

Masses can have a range of sizes. Cancerous lesions are stochastic biologic phenomena that manifest in images as having various structures occurring at different sizes and over ranges of spatial scales [6]. The boundaries of masses require a localized approach, although the sharpness and hence the scales of interpretation of the lesion boundaries, can vary considerably. Moreover, the spiculations that are associated with many cancerous lesions occur with different widths, lengths and densities, which suggests that their characterization will require analysis over scales.

Some of the researchers have used texture features to discriminate between mass and normal tissue [13-15]. Others have defined a number of features that were designed to capture image characteristics like intensity, iso-density, location and contrast [16, 17].

Most diagnosis algorithms (CADx) begin with a region of interest (ROI) containing a suspicious mass. In the preprocessing step of CADx algorithm depicted in Fig. 1, the mass is segmented from the background normal tissue. Then the features that capture the difference between malignant and benign masses are extracted. Most features are designed to capture the shape and margin characteristics of masses. These features can be organized into morphologic features and texture features. Finally, masses are classified as malignant or benign. Some researchers have also proposed classification of masses into other categories, such as round, nodular or stellate [12], or such as fibroadenoma, cyst, or cancer [18].

3.2. Calcifications

Calcifications are tiny granule like deposits of calcium and are relatively bright (dense) in comparison with the surrounding normal tissue [5]. Calcifications detected on mammogram are an important indicator for malignant breast disease. Unfortunately, calcifications are also present in many benign changes. Malignant calcifications tend to be numerous, clustered, small, varying in size and shape, angular, irregularly shaped and branching in orientation [5]. Benign calcifications are usually larger than calcifications associated with malignancy. They are usually coarser, often round with smooth margins, smaller in number, more diffusely distributed, more homogeneous in size and shape and are much more easily seen on a mammogram. One of the key differences between benign and malignant calcifications is the roughness of their shape. Typically benign calcifications are skin calcifications, vascular calcifications, coarse popcorn-like calcifications, large rod-like calcifications, round calcifications, lucent-centered calcifications, eggshell or rim calcifications, milk of calcium calcifications, suture calcification and dystrophic calcifications. Malignancy suspicious calcifications are amorphous and coarse heterogeneous calcifications. Malignancy highly suspicious calcifications are fine pleomorphic, fine-linear and fine linear-branching calcifications.

While observing calcifications it is important to consider their distribution (diffuse, regional, cluster, linear, segmental). In diffuse distribution calcifications are diffusely dispersed in the breast. Calcifications in regional distribution are distributed in larger breast tissue volume (> 2 cm³) and are very
often part of the benign changes. Cluster of calcifications is indicated if five or more calcifications are present in small breast tissue volume (< 1 cm³) and it is shown in Fig. 5. Linear distribution of calcification indicates malignant disease. Segmental distribution of calcifications also indicates malignant disease, but if the calcifications in segmental distribution are larger, smooth and rod-like they indicate benign changes.

![Figure 5. Example of cluster of calcifications [11]](image)

An analysis of the calcifications as to their distribution, size, shape or morphology, variability, number and the presence of associated findings, such as ductal dilatation or a mass, will assist one in deciding which are benign, which should be followed carefully and which should be biopsied [9]. The size of individual calcifications is less important than their morphology for deciding their classification and potential etiology. Variability in size, shape and density of microcalcifications is a worrisome feature, but variability must be assessed in conjunction with morphology. Those calcifications with sharp, jagged margins that are variable in appearance are much more likely to be malignant than are variably sized and shaped but smoothly marginated calcifications.

**Calcification detection and diagnosis**

One of the main characteristics to consider in the detection of calcifications is that they are generally very small. Their size varies from 0.1 mm to 1 mm and the average diameter is 0.3 mm [6]. Small calcifications may be missed due to the overlapping breast parenchyma. Another issue is that in regions where the background tissue is dense, it is very difficult to localize the calcifications. Finally, calcifications sometimes have a low contrast to the background and can be mistaken as noise in the inhomogeneous background.

A number of different approaches have been applied for the detection of calcifications. Calcifications represent high spatial frequencies of the image using wavelet transform [19-23]. Other non-wavelet-based methods try to make maximum use of the fact that calcifications have much higher intensity values than the surrounding tissue in a mammogram [24-27]. These methods are more likely to fail when the calcifications are present in dense background tissue [6].

Most diagnosis algorithms (CADx) begin with an ROI containing a cluster of calcifications. First step, segmentation, is the most difficult step in the computer-aided diagnosis of calcifications. Small size of calcifications makes the problem worse. Features that are extracted can be organized in two terms whether they describe properties of the cluster as a whole, calcification cluster features, or of the individual calcifications that make up the cluster, individual calcification features. Some of the common cluster features include the number of microcalcifications, the mean microcalcification area, standard deviation of the microcalcification contrast and the number of microcalcifications per unit area [6]. The features used for the diagnosis of calcification can be viewed as capturing morphologic or texture information. Finally, calcifications are classified as benign, malignancy suspicious or malignancy highly suspicious.

### 3.3. Architectural distortion

Although architectural distortions are less prevalent than masses or calcifications, they are the third most common mammographic sign of cancer and are strongly suggestive of malignancy [6]. A mammogram with architectural distortion is shown in Fig. 6.

![Figure 6. Mammogram with architectural distortion (dashed circle) [1]](image)
Architectural distortion is defined as distortion of the normal architecture with no definite mass visible, including spiculations radiating from a point and focal retraction or distortion at the edge of the parenchyma [7]. Architectural distortion of breast tissue can indicate malignant changes especially when integrated with visible lesions such as mass, asymmetry or calcifications. Architectural distortion can be classified as benign when including scar and soft-tissue damage due to trauma.

Detection of architectural distortion

Methods for detection of architectural distortion are often included in mass detection algorithms. However, methods designed exclusively for the detection of spiculated masses, which may rely on the presence of a central mass [1]. In order to detect architectural distortion some methods are based on the detection of spiculated lesions [28], on the detection of architectural distortion around the skin line and within the mammary gland [29] and some are texture-based [30]. Accurate detection of architectural distortion could be the key to efficient detection of early breast cancer, at pre-massformation stages.

3.4. Bilateral asymmetry

Asymmetry of breast parenchyma between the two sides has been one of the most useful signs for detecting primary breast cancer [31]. Asymmetric density is shown in Fig. 7.

A focal asymmetry is circumscribed area of asymmetry seen on two views, but it lacks the borders and conspicuity of a mass. It is usually an island of healthy fibroglandular tissue that is superimposed with surrounding fatty tissue. Focal asymmetries are often benign but occasionally may be a sign of breast cancer.

Asymmetries of concern are those that are changing or enlargeing or new, those that are palpable and those that are associated with other findings, such as microcalcifications or architectural distortion [9]. If a palpable thickening or mass corresponds to an asymmetric density, the density is regarded with a greater degree of suspicion for malignancy.

Detection of bilateral asymmetry

The evaluation of bilateral breast asymmetry based on density, shape and size is usually the first stage in the mammographic evaluation process [33].

Before performing asymmetry analysis it is necessary to apply some kind of alignment of the breasts. However, alignment procedures applied to mammograms have to confront many difficult problems such as the natural asymmetry of the breasts, absence of good corresponding points between left and right breast images to perform matching and distortions inherent to breast imaging [5].

Some of the developed methods for detection of bilateral asymmetry are texture-based [34]. Others are based on measures of shape, topology and distribution of brightness in the fibroglandular disk [35] or are based on measures of brightness, roughness and directionality [36]. More methods are desirable in this area to analyze asymmetry from multiple perspectives, including pattern asymmetry in the fibroglandular tissue as well as morphological and density measures related to the breast and the fibroglandular disk [1].

4. CONCLUSION

In this paper we have described typical steps in computer-aided detection and diagnosis of breast abnormalities. The paper presents brief overview of breast lesions and their features. Most significant abnormalities that may indicate breast cancer are described (mass, calcifications, architectural distortion and bilateral asymmetry). Wide range of features and their low visibility within the surrounding tissue makes the computer-aided detection and diagnosis of breast abnormalities a challenge. We outlined some of the developed CAD algorithms and showed that further developments are required to improve the detection and diagnosis of breast abnormalities using computers.
REFERENCES


[10] H. Zonderland: "BI-RADS - Introduction to the Breast Imaging Reporting and Data System", Available at: www.radiologyassistant.nl


[32] The mini-MIAS Database of Mammograms, available at: peipa.essex.ac.uk


