Handbook of X-ray Imaging
Physics and Technology
Paolo Russo

Computer Analysis of Mammograms

Publication details
Chisako Muramatsu, Hiroshi Fujita
Published online on: 14 Dec 2017


PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
60

Computer Analysis of Mammograms

Chisako Muramatsu and Hiroshi Fujita

CONTENTS

60.1 Introduction .................................................................................................................................................. 1231
60.2 Lesion Segmentation and Classification .................................................................................................... 1231
60.3 Interactive CAD ........................................................................................................................................ 1233
60.4 Image Retrieval ....................................................................................................................................... 1233
60.5 Density Measurement and Risk Assessment .............................................................................................. 1236
60.6 Digital/Simulated/Mathematical Phantoms ............................................................................................... 1237
60.7 3D Phantom ........................................................................................................................................... 1240
60.8 Deep Convolutional Neural Network ........................................................................................................ 1240
60.9 Radiomics and Radiogenomics .................................................................................................................. 1243
60.10 Summary ............................................................................................................................................. 1244

References ......................................................................................................................................................... 1244

60.1 Introduction

Studies on computer analysis of medical images began in the 1960s (Winsberg et al. 1967). However, extensive research under the concept of computer-aided diagnosis (CAD) on mammograms and other medical images started in the 1980s (Chan et al. 1987, 1988) (see Section IV, Chapter 59). Since the first FDA approval of the commercial system and the Medicare reimbursement, computer-aided detection (CADe) systems for breast masses and microcalcifications on mammograms have been widely used in clinical practice. Many studies have been published on the utility of CADe in prospective studies (Freer and Ulissey 2001; Birdwell et al. 2005). On the other hand, research topics have been shifted from lesion detection to lesion classification, density measurement for cancer risk analysis, prognostic estimation for treatment response, and development of digital and 3D phantoms. In this chapter, we introduce recent broad research topics on computer analysis of mammograms.

60.2 Lesion Segmentation and Classification

One of the problems in breast cancer diagnosis is the difficulty in distinguishing between benign and malignant lesions on mammograms. Assisting radiologists in reduction of false positive recalls and biopsies is a challenge medical image analysts have faced. A large number of studies have been published on computer-aided diagnosis/classification (CADx) of breast lesions. There are several review papers on CADx schemes (Cheng et al. 2006; Rangayyan et al. 2007; Elter and Horsch 2009; Jalalian et al. 2013). Readers should refer to these articles and the original papers for specific approaches. A general schematic flow includes the lesion segmentation, feature extraction and lesion classification. For classification of breast masses, the shape and margin characteristics are two important features, whereas, for microcalcifications, the shape of individual microcalcifications and their variation as well as the cluster distribution are important characteristics. Naturally, lesion segmentation is one of the major components in classification schemes.

There have been many studies investigating the computerized schemes for mass segmentation on mammograms. Oliver et al. (2010) provided a comprehensive review on such algorithms. Advantages and disadvantages of different segmentation methods are described in Cheng et al. (2006). In their review, the segmentation techniques are grouped into four types, that is, region-based methods, contour-based methods, clustering methods and model-based methods. Typical region-based techniques include region growing methods (Zucker 1976) and watershed methods (Beucher and Lenteuejoul 1979). The region growing method is a classic algorithm which expands a seed region by including the neighbor pixels with a specific criterion (generally pixel values). In region growing methods for mass segmentation, gradient or edge information was often used as an additional stopping criterion (Barman and Granlund 1994; Kupinski and Giger 1998; Petric et al. 1999). In some studies, a probabilistic model approach was also incorporated (Kupinski and Giger 1998; Marti et al. 2003; Kinnard et al. 2004). Other groups investigated improving the seed point placement (Zheng et al. 1995; Qi and Snyder 1998; Zhang et al. 2004).

Active contour models (Snakes) (Kass et al. 1988) and level sets (Osher and Sethian 1988) are the major contour-based
segmentation algorithms. An initial rough outline is iteratively adjusted to the probable location. Several groups investigated such optimization algorithms for precise segmentation of masses (Kobatake et al. 1999; Sahiner et al. 2001; Allen et al. 2003; Nakagawa et al. 2004; Timp and Karssemeijer 2004; Yuan et al. 2007; Shi et al. 2008). The common clustering methods include the k-means (MacQueen 1967) and fuzzy c-means (FCM) (Bezdek 1981) algorithms. These rather simple methods, sometimes combined with other techniques, were employed for the mass segmentation in relatively earlier studies (Sahiner et al. 1996; Li et al. 1997).

The model-based techniques by the definition by Oliver et al. (2010) (Cheng et al. 2006) include ones involving a training stage. One of the most common techniques is template matching. The templates can be the actual sample images or the probabilistic or statistical models. The studies using such an approach, however, generally aimed at detecting the masses rather than segmenting the lesions. Similarly, other model-based techniques involving the training of classifiers, such as neural networks, were intended to classify regions of interest as containing an abnormality or not.

More recently, Tao et al. (2010) proposed a learning-based segmentation scheme involving pixel classification and a graph-cut algorithm. The image features from sub-patches of mass and non-mass images were used to train a linear discriminant analysis classifier along with Gaussian mixture model to obtain a pixel-level probability map. By applying the Otsu thresholding method to the probability map, the mass candidate region is obtained. On the other hand, spicules were detected using Gaussian templates. Finally, the graph cuts algorithm was employed to integrate all information to obtain the segmentation mask. Cordeiro et al. (2016) proposed a modified GrowCut algorithm to segment masses. Unlike the regular GrowCut, one seed point at the center of mass was selected by an operator, and Gaussian functions were employed as fuzzy membership degrees in updating the cells.

Rahmati et al. (2012) used gamma functions to model the intensity of foreground (mass) and background of a region of interest. Using a level set algorithm, mass contour is iteratively adjusted to maximize the likelihood function, which is the product of the inner and outer probability distributions modeled by the gamma function. Abbas et al. (2013) proposed a mass segmentation algorithm based on multiscale features and maximum a posteriori (MAP) estimation. From a multiscale decomposition image by a steerable pyramid transform, local texture features are determined by a Gaussian Markov Random Field and employed as prior probabilities. For posterior probability, the fuzzy c-mean clustering method is applied to the energy features extracted from the decomposition image. Finally, the mass outline is obtained by MAP optimization.

Some investigators continue to explore improving classical methods. Berber et al. (2013) proposed a region growing method with an adaptive threshold based on size restriction. Using the result of Otsu’s thresholding, the threshold value of the region growing is adjusted to prevent under- or oversegmentation. Xu et al. (2011) proposed a watershed method controlled by internal and external markers, in which a lesion boundary is limited in a band region determined by a template matching.

Microcalcification segmentation methods are reviewed in Elter and Horsch (2009). Some of these methods are designed to enhance calcifications for detection rather than to obtain precise outlines, which can be difficult because of their size. For classification of microcalcification clusters, the shape of calcification, such as round or linear, and the variation in shape for a cluster (pleomorphic shape) are important features. In earlier studies, individual calcifications were delineated by region growing using manual seed points (Shen et al. 1994; Jiang et al. 1996). With the proposed various enhancement techniques, the enhanced microcalcifications can be segmented by rather simple approaches. Nakayama et al. (2004) applied a filter bank to enhance microcalcifications, which were then segmented by a thresholding technique.

Similar techniques used for mass segmentation have also been applied for microcalcification, such as the adaptive thresholding (Moradmand et al. 2011), region growing (Shanmugavadivu and Narayanan 2013) and watershed (Moradmand et al. 2011; Marrocco et al. 2012). Fu et al. (2005) proposed a segmentation method by lesion enhancement using a top-hat filter followed by edge detection using Sobel and Canny filters. Use of active contour methods was investigated by Liu et al. (2015) and Arikidis et al. (2015).

After a mass lesion is segmented, various features are determined. Tan et al. (2014) investigated different types of features and their potential utilities, based on the frequency of the features being selected in a 10-fold cross-validation experiment. In their study, the features that were frequently found useful were related to mass shape, isodensity and the presence of fat. In general CADx schemes for masses, shape-related features can include, but are not limited to the area, average diameter, largest diameter, ratio of two largest perpendicular diameters, circularity, irregularity, compactness and eccentricity. Although definitions may vary among studies, they describe mass shapes as radiologists describe them, by using the Breast Imaging Reporting and Data System (BI-RADS).

The features characterizing the lesion margin are often considered for detection and classification of spiculated lesions, such as radial edge gradient (Huo et al. 1995), line orientation (Karssemeijer and te Brake 1996), and linear structures (Zwiggelaar et al. 2004). Density-related features describe another important characteristic, which may include contrast, variance in pixel values, and histogram-based features. Various texture features have been investigated by a number of researchers, which include ones based on the gray level occurrence matrix (Sahiner et al. 1998; Mudigonda et al. 2000; Lim and Er 2004; Gupta and Markey 2005; Mavroforakis et al. 2006; Mohanty et al. 2013), gray level run length (Sahiner et al. 1998; Mavroforakis et al. 2006; Mohanty et al. 2013), and local binary patterns and their variants (Choi and Ro 2012; de Sampaio et al. 2015; Gardezi and Faye 2015; Liu and Zeng 2015; Masnouidi et al. 2015; Muramatsu et al. 2016).

For calcification classification, the calcification distribution, as mentioned earlier, is one important feature. It can be described by cluster shape features, such as the area, perimeter, circularity, eccentricity and orientation. The number of microcalcifications and their concentration, or the number per unit of area, are other possible features. For describing individual microcalcifications, the statistics are often calculated, which include the means and standard deviations of sizes, pixel values, contrasts, and irregularities. Note that these features are only some examples.
Finally, lesions can be classified into benign or malignant, based on the extracted features, using various classifiers. A number of classifiers and combinations of them have been applied to breast lesion classification. Some of these algorithms include linear and quadratic discriminant analyses, artificial neural network (ANN), support vector machine (SVM), k-nearest neighbor (kNN), decision tree (DT), random forest, and adaboost. A feature selection step may be added before entering them into the classifiers. Some of the common feature selection methods are forward or backward selection, step-wise selection, genetic algorithm, and principal component analysis.

60.3 Interactive CAD

There are different types of errors in image reading. CADe systems are intended to reduce perception or recognition errors; radiologists should not dismiss their initially suspected lesions just because the system did not mark them. On the other hand, radiologists may make interpretation errors, in which radiologists recognize the lesions but fail to recall them. Even when a system prompts the lesions, the correct CADe marks may be dismissed by radiologists. One reason for this could be due to too many false positive marks provided by the CADe systems, especially for mass lesions, so that radiologists lose confidence in them. In general, CADe algorithms output probabilities or continuous values. However, these outputs are dichotomized to “present” or “absent” by a prespecified threshold value. A very suspicious lesion and a marginal lesion are presented in the same way to users.

In assisting to reduce interpretation errors, some commercial CAD systems now provide (with limited distributions) additional information, such as the suspiciousness of lesions, size of lesions, morphology of masses, and the number of microcalcifications. For the same objective, Hupse et al. (2013) proposed a system that provides detection marks only when requested by a user’s click on queried regions with the suspiciousness scores, which they called an interactive CAD, as shown in Figure 60.1. The utility of the system was evaluated by the observer study, in which radiologists interpreted mammograms unaided, with regular CADe prompts and with interactive CAD. The results using localization receiver operating characteristic (ROC) analysis showed the superiority of the performance by the radiologists with interactive CAD compared to the unaided and with regular CAD performances.

60.4 Image Retrieval

Another type of supplemental information potentially useful in the CAD framework is reference images. Although CADx systems which provide the likelihood of malignancy are considered helpful, as supported by previous studies (Chan et al. 1999; Jiang et al. 1999; Huo et al. 2002b), their potential is not fully appreciated. The CADx systems are not yet matured enough to win the confidence of radiologists, and radiologists may be reluctant to use the probability without evidence. One solution is to provide the morphology information, as mentioned earlier, another is to present similar cases with known outcomes from a reference library.

Content-based image retrieval methods have been actively studied for a couple of decades in the field of computer vision and medical informatics. For assisting breast lesion classification, Qi and Snyder (1999) investigated an image retrieval of masses on
mammograms based on feature vector distance between a query and images in the archive. Sklansky et al. (2000) proposed a mapped-database system, which provides a biopsy recommendation and a relational map of a query and images in the database using a visual neural network for microcalcification clusters, as shown in Figure 60.2. Giger et al. (2002) developed a similar image retrieval system, called an intelligent workstation, as illustrated in Figure 60.3, which provides the likelihood of malignancy of a query mass and selects similar images on the basis of a single feature, multiple features or the likelihood of malignancy.

For retrieving visually similar and clinically relevant images, Muramatsu et al. (2005, 2008) proposed a machine learning method, in which an ANN was trained to learn the relationship between the image features and the similarities of pairs of masses and pairs of microcalcification clusters. The Gold Standard of image similarities was the subjective similarity ratings determined by expert radiologists and the estimated similarity outputs from the trained network were called the psychophysical similarity measures. Figure 60.4 illustrates the schematic diagram of the training procedure. In their later study, a multidimensional scaling (MDS) (Muramatsu et al. 2013) analysis was incorporated to map the similarity relationship of sample lesions and the ANN was employed to model this similarity space, as illustrated in Figure 60.5, for improving the similarity estimation. Observer study results indicated the potential usefulness of presenting reference images and similarity map as a visual aid (Figure 60.6) for distinction between benign and malignant masses.

Similarly, El-Naqa et al. (2004) investigated the machine learning approach to retrieve similar clustered microcalcification lesions. They proposed a two-stage procedure, in which coarse selection was made by a regular classifier in the first stage and more refined selection was performed on the basis of similarity learning in the second regression stage. A Fisher discriminant and SVM classifiers were considered for the first stage. For the second stage, ANN and SVM were tested to learn the relationship between the image features and subjective scores, based on the calcification distribution. Their experimental result showed that the use of SVM in both stages provided the best performance among other combinations and single-stage SVM.

Zheng et al. (2006) proposed an image retrieval scheme, which included an “interactive” step. For selecting visually similar images, reference images were preselected on the basis of the closeness of subjective margin ratings (±1 of nine scales) before the final selection was made by using a kNN algorithm. More recently, they proposed the new interactive CAD system, which combines interactive cueing and image retrieval approaches (Wang et al. 2012). The system first outputs the conventional detection prompts and users can query any regions, regardless of the presence of the original prompts, where the system then analyzes and searches for similar images. Based on the retrieved images, the system determines the malignant likelihood score. The final outputs are the detection score, the likelihood score,
**FIGURE 60.3** User interface of intelligent workstation for breast CAD. (Figure provided courtesy of M. Giger.)

**FIGURE 60.4** Schematic diagram of the training of an artificial neural network for determination of psychophysical similarity measure.
and the retrieved similar images, which could be malignant masses, benign masses or CAD-detected false positive regions, as a visual aid (Figure 60.7).

Kinoshita et al. (2007) proposed visual features to retrieve similar mammograms with comparable breast density levels. Tourassi et al. (2007) used the retrieved similar images based on mutual information for classification of masses and normal breast tissues. In their later study (Habas et al. 2007), they proposed the use of retrieved similar images for assessing the reliability of CAD results. Jing et al. (2012) proposed to use retrieved images in optimizing the classifier, by putting more emphasis on those that are similar to a query case for classification of clustered microcalcifications.

60.5 Density Measurement and Risk Assessment

It has been reported that the mammographic density is strongly associated with the breast cancer risk (Byng et al. 1998; Boyd et al. 2007). According to BI-RADS, mammographic density patterns can be classified into four categories, namely, almost entirely fatty, scattered fibroglandular densities, heterogeneously dense and extremely dense, based on the breast tissue composition. Hence, researchers have been investigating the computerized methods to segment fibroglandular tissue, estimate the percentage of dense tissue and/or classify mammograms into four density categories for estimating the breast cancer risk.

Byng et al. (1996) proposed a computerized method to analyze mammograms using the brightness histogram and fractal dimension. They found that these objective measures strongly correlate with the subjective classification by radiologists. Later they proposed an automatic measurement of volumetric breast density using a step wedge calibration phantom based on a 2-component model, in which breast is composed of fibroglandular tissues, including skin and blood vessels, and fatty tissues (Pawluczyk et al. 2003).

Zhou et al. (2001) proposed an automated method to estimate breast density by histogram analysis. First, they segmented the breast region from mammograms and applied an adaptive dynamic range compression technique to enhance the gray level histogram of the breast regions. Based on the histogram features, mammograms were classified into four classes. Finally, a threshold value for each image was automatically determined to segment the dense tissue for estimating the percentage of the dense tissue area, as shown in Figure 60.8.

Saha et al. (2001) investigated an automatic method to segment dense tissue using a scale-based fuzzy connectivity method for quantification of breast tissue density. Oliver et al. (2008) developed a breast tissue classification method that includes segmentation of dense tissue using fuzzy c-means clustering, extraction of morphological and texture features and a classification based on Bayesian combination of kNN and DT classifiers. Apart from density measurement, Huo et al. (2002a) and Li et al. (2005) analyzed mammographic density patterns by extracting the texture.
features for identifying women with breast cancer risk who were BRCA1 and BRCA2 gene mutation carriers.

In the US, breast density notification laws have been passed in 27 states and are in the process of being put into effect in eight states (as of May 2016). These laws basically require physicians to notify women with dense breasts who have undergone mammography examination that they may be associated with an increased risk of breast cancer. These laws may have played a part in increased need for automated density analysis software. Several vendors now provide commercial software, such as Quantra by Hologic, Inc., Marlborough, MA, VolparaDensity by Volpara Solution, Wellington, New Zealand and iReveal by iCAD, Inc., Nashua, NH.

60.6 Digital/Simulated/Mathematical Phantoms

Optimization of image quality and patient dose is an important issue in evaluating new imaging devices. The modern devices, such as breast tomosynthesis and dedicated breast CT, have a number of parameters that should be optimized. For a new device approval, a large clinical imaging trial is required, which is quite laborious. In the preclinical laboratory studies, various phantoms, such as ACR phantom (a mammography phantom accredited by American College of Radiology) and contrast detail phantom, are used for evaluating different systems and parameters (see Section IV, Chapter 56). Although obtaining real clinical images is preferable, such studies are limited by the cost, time, and radiation to subjects. Similarly, for development and evaluation of new CAD software, it is desirable to have large datasets. For assessing the clinical utility, observer studies are often conducted, in which case selection is essential. However, collecting cases having specific types of lesions with various subtleties at arbitral locations in breasts with different densities is extremely difficult. An alternative approach would be designing realistic simulation phantoms (see Section IV, Chapter 57).

Bakic et al. (2002) proposed a mammography simulation consisting of three components: a 3D breast phantom consisted of two ellipsoidal regions of large-scale adipose tissue and fibroglandular tissue and the internal structures of adipose compartments and ductal network; a compression model based on tissue...
elasticity and a breast deformation model; and an image acquisition model assuming monochromatic X-rays with a parallel beam geometry. They later proposed an efficient algorithm to generate high resolution phantoms (Pokrajac et al. 2012). Further, they incorporated a partial volume simulation to allow voxels with multiple tissue compositions for reducing artifacts (Chen et al. 2015). Examples of digital phantoms with different breast densities are shown in Figure 60.9 (Barufaldi et al. 2016).

Erickson et al. (2016) created a large dataset of realistic computational breast phantom based on dedicated breast computed tomography (CT) data. Breast volume on CT was first segmented using a 3D bias-corrected FCM clustering algorithm. After denoising and nonuniformity correction, the FCM method was applied again to segment the breast tissue into background, adipose, fibroglandular, and skin. The 3D breast phantom was defined in the same format as the 4D extended cardiac-torso phantom (Segars et al. 2010), in which a nonuniform rational B-splines surface was used to model each anatomical object whose shape can be modified by manipulating its control points. The fibroglandular boundaries were represented by the surface meshes. After digital compression and simple ray-tracing, simulated mammograms are obtained as shown in Figure 60.10.

Bliznakova et al. (2003) proposed a breast model including the breast surface, duct system, Cooper’s ligaments, pectoral muscle,
and mammographic background. The images were synthesized assuming monoenergetic fan beams. The algorithm was later improved in 3D mammographic background texture generation (Bliznakova et al. 2010). Youn et al. (2011) proposed a procedure to generate mammograms considering the image characteristics of a detector. The phantom consisted of soft tissue generated by low-pass-filtered white-spectrum noise with a parameter assigning the voxel to glandular tissue or adipose tissue and ductal networks with random angles and lengths. The image was then modified regarding the resolution and noise characteristics of a detector.

Shaheen et al. (2014) proposed a method to build 3D mass models using MRI data. Having manually segmented masses as nonspiculated models, spicules were grown using an iterative branching algorithm. The created mass models were inserted into 2D mammograms and projection images of tomosynthesis data followed by reconstruction. de Sisternes et al. (2015) also proposed a method to produce 3D simulated masses by embedding them within mammograms. The masses were generated using a modified stochastic Gaussian random sphere model and spicules were added by an iterative fractal branching algorithm.

FIGURE 60.9 Examples of digital phantoms with breast densities of (a) 15%, (b) 21%, and (c) 26%. (With kind permission from Springer Science + Business Media: International Workshop on Digital Mammography, Lecture Notes in Computer Science, The effect of breast composition on a no-reference anisotropic quality index for digital mammography. Barufaldi et al. 2015.)

FIGURE 60.10 Simulated projection images with volumetric breast densities of 11.8%, 24.7% and 34.6% from left to right. (From Erickson, D. W. et al. 2016. Medical Physics 43:23–32. American Association of Physicists in Medicine. With permission.)
**60.7 3D Phantom**

With recent technology development, it is now possible to create a realistic 3D physical phantom using 3D printing. Sikaria et al. (2016) developed an anthropomorphic breast phantom based on dedicated breast CT data using a commercial 3D printer. In their second generation phantom, they were able to incorporate a high resolution voxelized printing with desired propositions of two mixed materials (Figure 60.11). Simulated lesions can be inserted by inkjet printing on a sheet of paper. Clark et al. (2016) introduced an open source code, mammoreplicator, which creates a printable 3D object based on an input 2D mammogram. The phantom only replicates the attenuation profile of the mammogram and cannot be used for evaluating other imaging modalities at this point.

These phantoms can be used to evaluate imaging systems. On the other hand, there has been growing interest in application of 3D printing to surgical assistance. Customized breast phantoms have potential utility in assisting cancer removal and breast reconstruction surgery (Kelil et al. 2015). Such an application can be expected more in the future.

**60.8 Deep Convolutional Neural Network**

Currently one of the hottest topics in medical image analysis is deep learning. Since its remarkable success in pattern recognition competitions, applications of deep learning methods to medical image analysis have been extensively examined.

Petersen et al. (2014) and Kallengerg et al. (2016) proposed the use of a convolutional sparse autoencoder (CSAE) which consisted of three unsupervised convolutional layers and one unsupervised pooling layer, as shown in Figure 60.12. The last two layers were trained to output the (1) segmentation result of background, pectoral muscle and breast tissue, (2) density score or (3) texture score for evaluation of cancer risk. Multiresolution patches from mammograms are input to the CSAE. Although the areas under the ROC curves (AUCs) for predicting cancer cases were not high, they were slightly better than those of other state-of-the-art methods, and the segmentation result was similar to the manual result.

Sun et al. (2016) employed a deep convolutional neural network (DCNN) for cancer risk analysis. Each mammogram was divided into 100 region of interests (ROIs) and these patches were input to a network consisting of four pairs of convolution and pooling layers and a fully connected layer. They obtained the case-based AUC of 0.7173 for separating low risk and high risk cases.

Kooi et al. (2016, 2017) applied a DCNN for classifying ROIs with malignant masses and normal ROIs detected by a conventional CAD system. They trained a network similar in architecture to one by the Oxford group, so called VGG net, with five convolutional layers, each of which were followed by a pooling layer, except for the fourth layer and two fully connected layers, as shown
in Figure 60.13. The positive training samples were augmented by translation and scaling for reducing the overfitting effect. It was reported that the DCNN outperformed the conventional system at a low sensitivity level and performed comparably at a high sensitivity level. The DCNN also obtained the comparable AUC with the mean AUC by radiologists in the observer study.

Arevalo et al. (Arevalo et al. 2016) reported that even a neural network with two convolutional layers followed by pooling layers and one fully connected layer can perform better than a conventional system with hand-crafted features for classification of benign and malignant masses. In their experiment, when the convolutional neural network (CNN) was combined with the hand-crafted features, no improvement was obtained, suggesting the capability of the CNN in learning such information from the images.

Carneiro et al. (2015) investigated the use of DCNN on multiview mammograms for estimating the BI-RADS score. First, they built a single view model that takes a full mammographic view as input and output probabilities for negative (BI-RADS 1), benign (BI-RADS 2 and 3), and malignant (BI-RADS 4, 5 and 6). Such a model was built for each of six views, including craniocaudal (CC) and mediolateral oblique (MLO) views, mass segmentation and microcalcification segmentation in each view. The network consisted of four sets of convolution and pooling layers, two fully connected layers and a multinomial logistic regression layer. They also built a multiview model, which takes the features from the final fully connected layers from six models and trains a softmax layer, as shown in Figure 60.14. They tested the models with and without transfer learning, which was pretrained with Imagenet, and with and without data augmentation. Using two publicly available databases (InBreast and DDSM), they concluded that the multiview model performs better than single view models in terms of volume under the ROC surface for three-class classification and AUC for benign–malignant classification with cases having at least one lesion. They also found that the pretrained model without data augmentation performs better than the randomly initialized model with data augmentation.

Mordang et al. (2016) applied a DCNN approach to microcalcification detection. Small patches of 13 × 13 pixels were obtained at the centers of individual microcalcifications for positive samples and at random locations on mammograms for negative samples. Because of the large imbalance in positive and negative samples, two CNNs were trained, in which the first CNN was trained with all positive training samples multiplied by 10-times and the equal number of samples randomly selected from the training set. The second CNN was trained with all the positive samples with data augmentation by flipping and rotating

![FIGURE 60.13 Illustration of DCNN architecture employed by Kooi et al. (2016).](image)

![FIGURE 60.14 Illustration of single view CNN (left) and multiview CNN (right) models proposed by Carneiro et al. (2015).](image)
and the balanced “difficult” negative samples according to the classification score obtained by the first CNN. The network consisted of two sets of two repeated convolutional layers followed by a pooling layer and three fully connected layers. They found the superior performance compared with the state-of-the-art cascade classifier in terms of the mean sensitivity in the specificity range from 0.9–0.999999 on a logarithmic scale.

### 60.9 Radiomics and Radiogenomics

Other emerging fields that have been receiving growing interest are radiomics and radiogenomics. These studies on breast cancer so far are limited to magnetic resonance imaging (MRI), probably because of the limited application of 3D X-ray imaging in breasts.

Yamamoto et al. (2012) reported a preliminary study on radiogenomic analysis of breast cancer with MRI. They first analyzed gene expression data of 353 breast cancer patients. Using a small subsample of 10 patients who underwent MRI exams, they studied the relationship between MRI features and the selective genes for potential imaging biomarkers.

More recently, Wang et al. (2015) investigated MRI features for predicting triple-negative (TN) breast cancers, which lack expression of the estrogen receptor, progesterone receptor and human epidermal growth factor 2 receptor and, therefore, generally have a poor prognosis. Using 84 patients with 88 invasive cancers, in which 11 are TN, a SVM trained with MRI features achieved an AUC of 0.878 in differentiating TN and non-TN cancers. The unsupervised clustering analysis showed that the majority of TN cancers were clustered into the group with similar radiomic

![Diagram](image)

**FIGURE 60.16** Overview of identified significant association between genomic features and radiomic phenotypes. (a) A graph showing association between radiomic features and genomic features. Circular nodes specify radiomic or genomic features while lines indicate identified statistically significant associations. (b) A table showing the numbers of statistically significant associations between genomic features of different platforms and radiomic phenotypes of different categories. (Reprinted by permission from Springer Nature Scientific Reports, Zhu, Y. et al. 5:17787, Copyright 2015. Nature Publishing Group, permission by Creative Commons.)
expression patterns, although non-negligible numbers of non-TN cases also had similar expression patterns (Figure 60.15).

Zhu et al. (2015) studied the association between various genomic features and MRI-based radiomic phenotypes using 91 breast cancer data from the Cancer Imaging Archive and the Cancer Genome Atlas by the U.S. National Institutes of Health. They found a statistically significant association between many of the genomic features and radiomic phenotypes (Figure 60.16).

One of the difficulties in radiomic and radiogenomic studies is data acquisition. With the effort of collecting standardized large “-omics” data, especially by national programs, further progress in these research fields and their application in clinical decision-making are expected in the future.

60.10 Summary

In this chapter, we introduced the computerized image analysis methods primarily on mammograms. For more details of each methods, readers should refer to the original articles. The clinical aspects of CAD systems and the technology application on breast tomosynthesis are discussed in Chapter 59. With the availability of big data and faster machines with high processing capability, the imaging research trend toward machine intelligence and data mining will likely continue. The effort in improving image analysis methods is expected to contribute to advance clinical practice.

REFERENCES


Computer Analysis of Mammograms


