Maryellen L. Giger, PhD

Computer-aided Diagnosis in Diagnostic Mammography and Multimodality Breast Imaging[1]

Useful interpretation in mammography depends on the quality of the mammographic images and the ability of the radiologists who interpret them. Improvements in radiographic technique, as well as mandatory accreditation programs, have made the early signs of breast cancer more apparent on mammograms. However, radiologists still sometimes miss cancer on a mammogram. One approach to improving performance is replicated interpretations, in which more than one observer reviews the images. For example, investigators have shown that detection of early breast cancer can be increased with double reading by two radiologists (1). In addition, results of studies have shown that interpretation performance varies greatly among radiologists (2–4).

An alternate approach is to use a computer as the second reader. Use of output from a computerized analysis of an image by radiologists may help them in the tasks of detection or diagnosis and potentially improve the overall interpretation of breast images and the subsequent patient care. Many factors motivate the attempts to aid or automate radiologic diagnosis. Inadequacies in interpretation performance may be due to the presence of image noise or normal anatomic structures, as well as to known limitations in the human search and perception process. Ultimately, computer-aided diagnosis (CAD) may become an integrated tool in the diagnostic work-up of suspect breast lesions by using multimodality images.

This chapter reviews various CAD methods in breast imaging (mammography, ultrasound [US], and magnetic resonance [MR] imaging), which are focused on the characterization of lesions and the estimation of the probability of malignancy for use in the diagnostic work-up of suspect lesions. CAD systems in diagnostic work-up usually involve having the computer extract the margin of the lesion from the surrounding parenchyma, extract characteristics (features) of the lesions, merge these computer-extracted features into an estimate of the probability of malignancy, and, as an option, automatically retrieve similar lesions from an online reference library. The aim of CAD in diagnostic work-up is to increase classification sensitivity and specificity, as well as to reduce intra- and interobserver variability. Various reviews have been written about CAD in breast imaging (5–11).
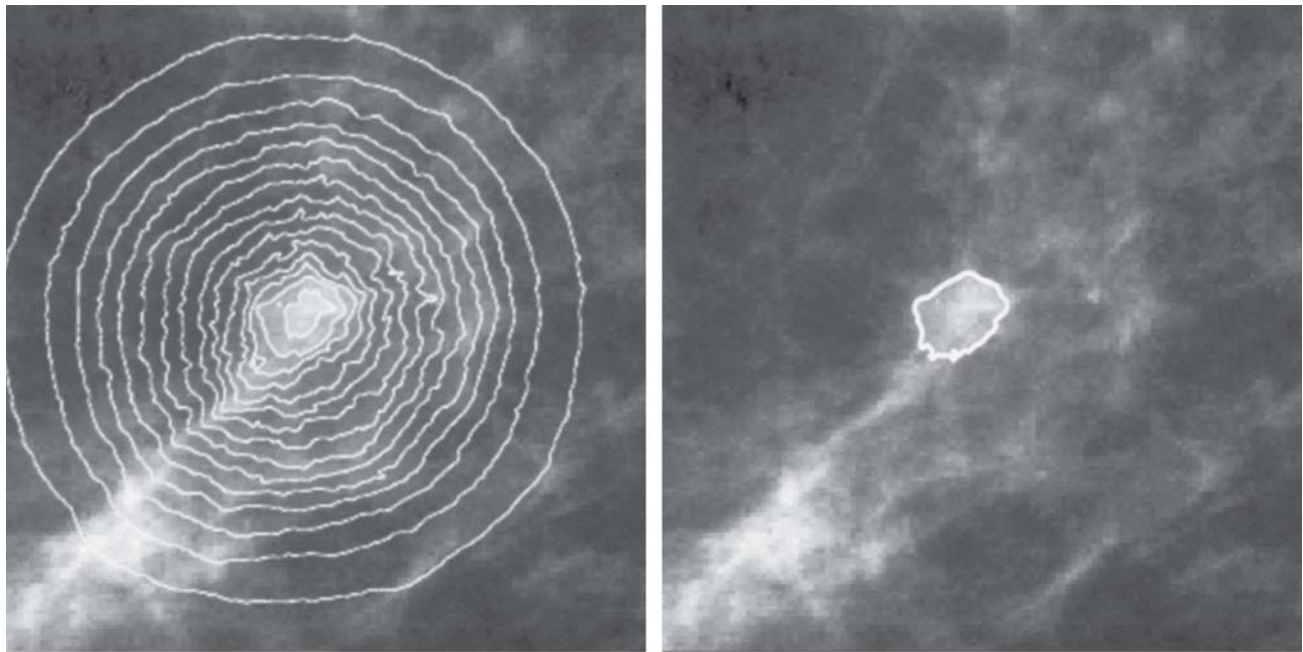
RSNA'04

a.

b.

Figure 1. **(a)** Segmentation contours resulting from region growing by using radial gradient index segmentation and **(b)** segmentation result. (Reprinted, with permission, from reference 16.)

Computer-aided detection has already been incorporated into clinical screening mammography, and its status is reviewed elsewhere in this RSNA syllabus (see chapter by Chan et al). There are currently computer-aided *detection* systems approved by the Food and Drug Administration, with many insurance carriers, including Medicare, providing coverage for such technology. Development of promising computer-aided *diagnosis* prototypes is also underway, and aspects of these future systems are reviewed in this chapter.

The general techniques employed in the computer analysis of images can be broadly categorized as computer vision and artificial intelligence (12,13). Computer vision involves having a computer extract from a digital image features that may or may not be otherwise perceived by a human reader. The development of computer vision schemes requires a priori information about the medical image (eg, the mammogram) and knowledge of various computer processing techniques and decision analysis methods. The required a priori knowledge includes the physical imaging properties of the digital image acquisition system and morphologic information concerning the abnormality (eg, mass lesion or cluster of microcalcifications), along with its associated anatomic background. That is, a sufficient database is needed to cover the entire range of abnormal and normal findings. Computer vision techniques include, in general, image processing, image segmentation, and feature extraction (12,13).

Computer vision algorithms can be initially used *(a)* to isolate, or segment, the breast from the remainder of the image and so limit the computer search re-gion or *(b)* to enhance the peripheral breast border region to compensate for reduced breast thickness at the edge. For example, investigators have described methods that use computer-defined unexposed and direct-exposure image regions to generate a border around the breast region (14,15). Next, segmentation techniques can be employed to separate the image into regions with similar attributes (eg, regions exhibiting high contrast that might reflect the presence of calcifications) or to isolate a lesion from its surrounding parenchymal background, as demonstrated in Figure 1 (16). Once segmented, each region is then further analyzed with feature extraction and feature analysis techniques, which yield mathematical descriptors of the radiographic features.

Radiologists seem to extract and interpret simultaneously many radiographic image features corresponding to signs of malignancy. For example, a high degree of spiculation exhibited by a mass is a strong sign of malignancy, and many computerized methods have been developed to quantitate spiculation (17–21). Thus, computer vision methods involve determining the mathematical descriptors of image features, along with the selection of which individual computer-extracted features are clinically important (7,22).

Much as the radiologist weighs different aspects of a mammographic finding, artificial intelligence techniques can be used to merge and/or select image features obtained with computer vision into a diagnostic decision (22–32). Various classifiers have been explored as means to merge the computer-extracted features, including rule-based methods, discriminant

analysis, artificial neural networks, Bayesian classifiers, fuzzy logic, and combination methods, and these are summarized elsewhere (7). Analogous to radiologist training, classifiers are "trained" on cases in which the actual disease status is known. This training determines the setting internal to the computer algorithm in a way that provides output that reflects the possible disease status most accurately.

It should be noted that some of the limitations of the human visual system that are based on image quality would also limit computerized image analyses (33). One should not expect computer image analysis systems to "interpret" poor-quality mammograms having low signal levels or high noise levels. However, the analysis of mammograms by computers is expected to improve if the mammograms are of high quality.

Validation ultimately requires assessment of the performance of a computerized method for a specific task on an independent set of cases. The assessed performance of a computer system can be based on the system's performance alone, or it can be based on radiologists' performances when they are using the computer output as an aid in their interpretations (34).

The output of the various detection and classification methods varies depending on the details of the specific methods that are available, the database employed to evaluate the computerized approach, and the level of expertise of the human user. It is also important to note that it is not possible to compare different computerized methods because of the use of different databases. That is, it cannot be assumed that a computerized scheme that achieves a high sensitivity with one database of mammograms will achieve a similar performance level with another database or with an actual patient population. In addition to the database, the method of evaluation will influence performance and expectations of a specific computerized scheme. Some of the computerized schemes have been tested, or merely demonstrated, with only a few images, whereas others have been tested with moderate-sized databases and statistical testing methods. Performance levels in the latter situations are usually given for the detection schemes in terms of the sensitivity (true-positive rate) for detection and the number of false-positive detections per image and are given for the classification schemes in terms of the sensitivity for classification and the specificity (ie, 1 – the false-positive fraction).

CAD IN SCREENING MAMMOGRAPHY: OVERVIEW

For completeness, CAD in screening mammography will be briefly mentioned here. In screening mammography, radiologists are limited in their detection ability by quantum noise and the presence of overlapping structures, that is, "structured noise." Although mammography is currently the best method for the detection of breast cancer, radiologists may still fail to detect cancer that is evident retrospectively. The missed detections may be due to the subtle nature of the radiographic findings (ie, low conspicuity of the lesion), poor image quality, eye fatigue, or oversight by the radiologist. Humans also routinely incur oversight errors because of nonsystematic search patterns and lapses in perception. Such oversight error is exacerbated in mammographic screening, in which more than 99.5% of the cases are normal.

In addition, it has been suggested that double reading (by two radiologists) may increase sensitivity (1). Thus, one aim of CAD is to increase the efficiency and effectiveness of screening procedures by using a computer system, as a "second reader" (like a spell-checker), to aid radiologists by directing their attention to regions on mammograms that the computer deems to have features associated with cancer. Note that the final decision regarding the likelihood of the presence of a cancer and patient management is left to the radiologist. Because mammography is a high-volume x-ray procedure routinely interpreted by radiologists and because radiologists do not detect all cancers that are visible on images in retrospect, it is expected that the efficiency and effectiveness of screening will be increased with CAD.

Computer detection systems identify locations of suspect abnormalities and convey this information to radiologists by means of marks on display monitors, hard-copy film, or paper. The computer output for the detection of potential lesions indicates both actual lesions and false-positive findings. During the years, investigators have worked to increase detection sensitivity while decreasing the number of false marks per image.

The ultimate test of a CAD system is its additive value, that is, the improved performance of radiologists when the system is used in the clinical interpretation process. The clinical usefulness of computer-aided detection for screening mammography has been studied with various laboratory observer studies (35–37), retrospective analysis of cases of "missed" lesions (38–41), and, more recently, actual prospective clinical usage of commercially available systems (42–44). Studies are also being conducted to determine the necessary sensitivity and false mark rate for computerized detection methods to allow for improvement in radiologists' performance levels (45). If the sensitivity is too low or if there are too many false marks per image, the computer aid may be either not useful or detrimental to observer performance. Most recently, studies are being performed to assess the performance of computer detection methods for full-field digital mammography systems, for which computer analysis becomes a "push-button" implementation without the need for film digitization.

CAD IN DIAGNOSTIC BREAST IMAGING: OVERVIEW

Once a possible abnormality is detected, its characteristics are simultaneously extracted and assessed by a radiologist to decide if the abnormality is likely to be malignant or benign and to recommend the subsequent patient management (eg, return to screening, short-term follow-up, additional images, biopsy). Although some lesions may show obvious signs of malignancy or the lack thereof, many lesions present confusing findings. Many patients are referred for surgical biopsy on the basis of a radiographically detected mass lesion or cluster of microcalcifications. Although there are general rules for differentiating between benign and malignant breast lesions (46), radiologists' interpretations of lesions with current radiographic techniques vary considerably (47). Thus, an aim of CAD is to extract and analyze objectively the characteristics of benign and malignant lesions seen on images, aiding radiologists by increasing their diagnostic accuracy and reducing false-positive diagnoses of malignancies and thereby decreasing patient morbidity as well as the number of surgical biopsies performed and their associated complications.

Characteristics of a lesion can also be assessed further with multiple imaging techniques, including special-view mammography, US, and MR imaging. Incorrect management decisions resulting from inaccurate image interpretation will typically decrease the positive predictive value for biopsy recommendations. The advent of more multimodality analyses of the breast furthers the role of CAD, both for computerized image analysis of the suspect lesion with different modalities and for the merging of multimodality information.

CAD systems in diagnostic work-up are expected to extract the margin of the lesion from the surrounding parenchyma, extract characteristics (features) of the lesions, merge these computer-extracted features into an estimate of the probability of malignancy, and, as an option, retrieve similar lesions automatically from an online reference library (7). The aim of CAD in diagnostic work-up is to increase classification sensitivity and specificity and to reduce intra- and interobserver variability. It is important to restate that one of the aims of computerized classification is to reduce the number of benign cases sent for biopsy. Such a reduction, however, will be clinically acceptable only if it does not result in malignant cases not being sent for biopsy, because the "cost" of a missed cancer is much greater than misclassification of a benign case. Thus, computer classification schemes should be developed to improve specificity, but not at the loss of sensitivity.

Computerized image analysis schemes to aid in diagnosis may use features that are extracted either by computers or by radiologists. The benefit of computer-extracted features is the objectivity and reproducibility of the measurement of the specific feature. However, radiologists employ many radiographic image features, which they seem to extract and interpret simultaneously and instantaneously. Thus, as discussed earlier with computer detection methods, the development of computer diagnostic methods that use computer-extracted features requires both the determination of which individual features are clinically important and the computerized means for the extraction of each such feature (22–32). Computer vision and artificial intelligence techniques are presently being developed for the full range of modalities encountered in diagnostic breast imaging: diagnostic mammography, US, and MR imaging.

CAD in Diagnostic Mammography

The initial investigations into the use of computers in diagnostic mammography involved artificial intelligence techniques to merge observations of image features made by radiologists into useful diagnostic predictions (48–50). Such methods are dependent on the subjective identification and interpretation of the mammographic data by human observers. Recently, BI-RADS ratings provided by human readers have been analyzed by computers for lesion characterization (51). However, to eliminate the subjectivity of human ratings and to automate lesion classification more fully, features extracted with computer vision have been investigated as computerized diagnostic aids. Such mathematical descriptors may be based on those visually used by radiologists, such as mass spiculation or the spatial distribution of microcalcifications, or they may be features that are not visually apparent to a human observer.

One of the earliest investigations into using computers to extract mammographic features and merge them into a decision on malignancy was in 1972, when Ackerman and Gose (52) evaluated the use of digital xeroradiographs and the computer extraction of four measures of malignancy (calcification, spiculation, roughness, and shape) from radiologist-identified suspicious regions. Calcification was characterized by using a derivative histogram of the suspect region. Spiculation was measured by examining the lines tangent to the edges of lesions. Roughness consisted of the sum of the absolute differences between points along a vertical line, whereas shape was determined by using a perimeter-to-area ratio. The features were then merged into a decision on malignancy. The results, expressed in terms of receiver operating characteristic (ROC) curves, were obtained with a database of 60 images (30 having malignant and 30 having benign lesions) and indicated that the computer performed as well as an expert radiologist.

Various reviews of computer-extracted features of mammographic lesions have been written elsewhere (5–11), and only some will be presented here for
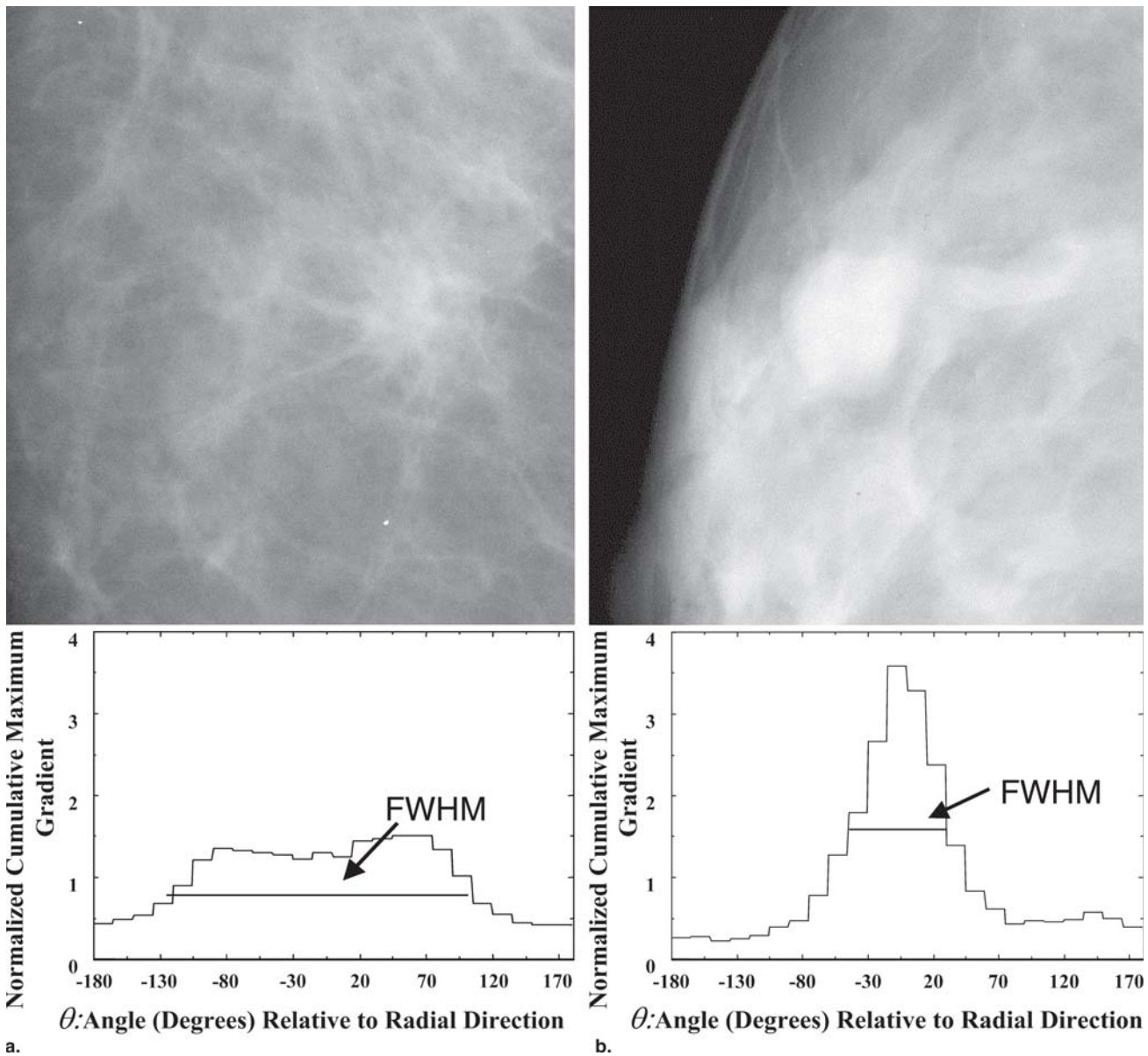
Figure 2. Examples of **(a)** a mammographic spiculated mass and its corresponding cumulated edge-gradient-orientation distribution and **(b)** a mammographic circular smooth mass and its corresponding cumulated edge-gradient-orientation distribution. *FWHM* = full width at half maximum. (Reprinted, with permission, from reference 18.)

illustration. Various investigators have characterized lesions by using mass features, such as size, degree of spiculation, average gray level, contrast, smoothness of the interior, and texture, and microcalcification features, such as the shape of individual calcifications and their distribution within a cluster (7,17,18,53–66). These features were subsequently merged by using discriminant functions to establish a probability that a region contained malignant, benign, or normal tissue.

For example, in the late 1980s, a characterization method was developed that was based on the degree of spiculation exhibited by a mass, because malignant masses often can be distinguished from benign masses by their more spiculated appearance at mam-

mography (17,18). In the classification scheme for mass lesions (18,53), the mass is first extracted from the anatomic background of the mammogram with automatic region-growing techniques. Features extracted are then obtained by using cumulative edge-gradient histogram analysis, in which the gradient is analyzed relative to the radial angle. In the cumulative edge-gradient analysis, the maximum gradient and angle of this gradient relative to the radial direction are calculated.

Figure 2 (18) illustrates the calculation of the full width at half maximum from the cumulative gradient orientation histogram for a smooth mass and a spiculated mass. Other features include gray-level and geometric measures. In the associated classifier,
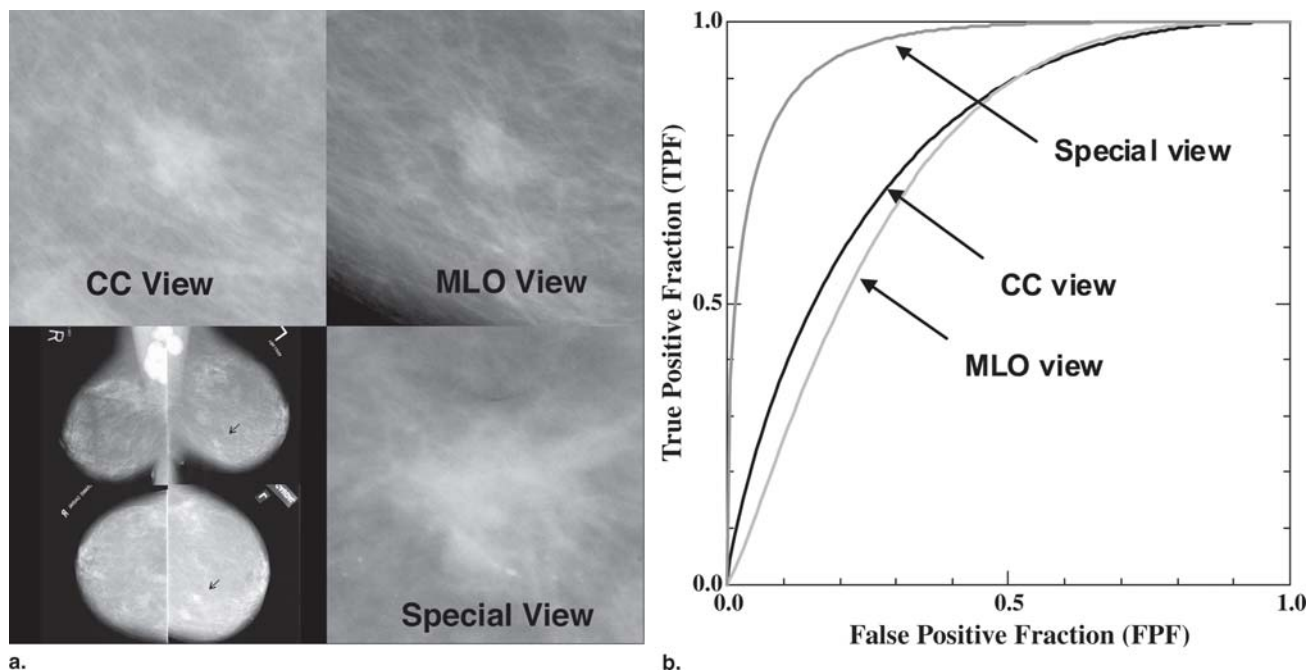
Figure 3. **(a)** Example of a malignant mass lesion, including the conventional mammograms (four-on-one view), the region of interest (ROI) for the craniocaudal *(CC)* view, the ROI for the mediolateral oblique *(MLO)* view, and the special-view ROI. **(b)** ROC curves obtained for the computer analysis of single-view mammographic images, that is, the CC, MLO, and special views. (Reprinted, with permission, from reference 56.)

spiculation was shown to dominate the learning, thus leading to the use of a combined rule-based/artificial-neural-network classifier (54). In addition, the effects of variations in both case mix and film digitization technique on the performance of the method were assessed (55). Computer categorization of lesions as malignant or benign achieved an A_z value (area under the ROC curve) of 0.90 for a prior training database (one commercial scanner's digitization) in a round-robin evaluation and A_z values of 0.82 and 0.81 for an independent database analyzed with two other commercial digitization formats. In the statistical comparison of these performances, however, the investigators failed to show a significant difference between the performances with the training database and the independent validation database ($P > .10$). Thus, the computer-based method for the classification of lesions on mammograms was shown to be robust to variations in case mix and film digitization technique (55).

The analysis of robustness was extended when the investigators applied their algorithm to special-view mammography, which resulted in a substantial improvement in computer performance, as demonstrated in Figure 3 (56). In a result that parallels human observer performance, the investigators demonstrated improved computer performance in diagnosing lesions on special-view mammograms compared with that on standard views (craniocaudal and mediolateral oblique views).

In 1991, in one of the earliest investigations of computer characterization of microcalcifications, Patrick et al (60) reported on a network of trained expert learning systems. The network included three systems—one for features of individual microcalcifications, one for features of the clusters, and one for clinical presentations—that classify clusters of calcifications as benign or malignant. Calculated geometric parameters included center of gravity, length, breadth, orientation, area, perimeter, moments, and convex perimeter. Features were then constructed from these measurements. With 25 clusters of calcifications that had been referred for biopsy, the system had a classification sensitivity of 90% (nine of the 10 cancers were correctly recognized) and a specificity of 60% (nine of the 15 benign cases were correctly recognized), for an overall classification accuracy of 72%.

Various other methods for characterizing clustered microcalcifications have been reviewed elsewhere (7). For example, in the mid-1990s, a computerized method for classifying clustered microcalcifications was developed; it involved extracting features that correlate qualitatively with radiologists' experience (61,62). The computerized classification system estimated the likelihood of malignancy on the basis of the appearance of clustered microcalcifications on mammograms. The computer automatically extracted features of the microcalcifications that were related to their size, shape, quantity, and spatial distribution and then used an artificial neural network to merge the features to produce an estimate of the likelihood of malignancy. This likelihood can then be used by a radiologist to help decide whether the microcalcifications are malig-
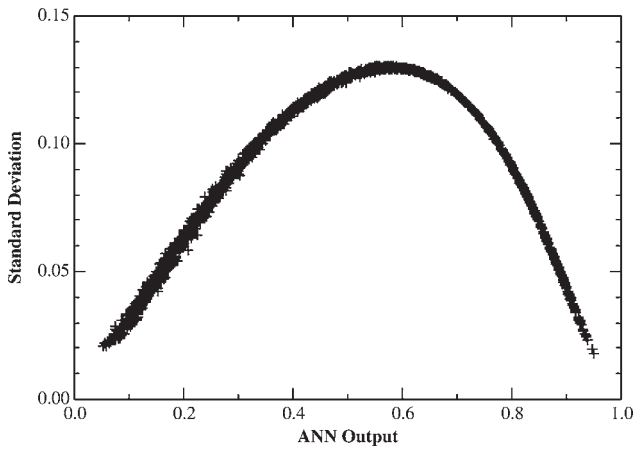
Figure 4. Standard deviation in the output of eight artificial neural networks (ANNs) as a function of the average ANN output. The ANNs were trained with a single training data set and identical network parameters except for an arbitrary but different seed value to a random number generator. (Reprinted, with permission, from reference 63.)

nant or benign. Further studies on the classifier used in the computerized diagnosis method demonstrated the variability in the outputs of multiple neural networks that are trained and "optimized" from the same training set (63). This is demonstrated in Figure 4 (63), which shows the standard deviation in the output of eight artificial neural networks as a function of the average artificial neural network output. Interestingly, the largest variations were when the case was difficult to interpret, as indicated by the midrange output.

In laboratory observer studies, computerized diagnostic methods have been shown to aid radiologists in distinguishing between malignant and benign lesions (62,64,65). Use of a computer diagnostic aid therefore has the potential to increase sensitivity, specificity, or both in the work-up of breast lesions. Investigators have demonstrated that radiologists showed an increase in both sensitivity and specificity in the characterization of clustered microcalcifications and in the associated recommendation for biopsy (62). In addition, it has been shown that improved performance can be obtained by both expert mammographers and community-based radiologists who use CAD information, with the increase greater for the nonexperts (64). However, methods for combining radiologist interpretations and computer outputs are still being investigated, and these methods will depend on the performances of the computer and the radiologist alone, as well as on the method by which the radiologist can incorporate computer output into his or her final decision-making process. However, early studies show the potential for using computer output to reduce the variability between radiologists' interpretations (66).

Future directions for CAD in diagnostic mammography include connecting the computerized image

analysis for diagnosis to the output of computerized mammographic detection methods. Toward this end, Edwards et al (67) have investigated methods for three-way classification to incorporate computer classification methods with computer detection output and thus distinguish between malignant lesions, benign lesions, and false-positive detections.

CAD in Breast Ultrasound

The diagnostic work-up of a suspected breast lesion may also include imaging with multiple modalities, such as US and MR imaging. Breast US is used as an important adjunct to diagnostic mammography and is typically performed to evaluate palpable and mammographically identified masses and determine whether they are cystic or solid. Although US is used by many radiologists to distinguish between solid lesions and cysts, it is difficult to distinguish between benign and cancerous solid lesions. Stavros et al (68) used various features to characterize masses as benign, indeterminate, and malignant. Their classification scheme had a sensitivity of 98.4% and a negative predictive value of 99.5%.

US is a digital modality that is amenable to the application of CAD techniques that ultimately could be used in a real time, at the time of examination, to improve diagnostic accuracy. Given that US interpretation is a subjective process and that criteria have been developed that may allow differentiation between benign and malignant solid breast masses, it is reasonable to assume that CAD techniques applied to US images would improve radiologists' performance, particularly when combined with corresponding mammographic data.

CAD methods in breast US are being explored by various researchers (69–74). As with mammographic lesions, the lesion is first extracted from the breast background (72), and lesion features are then extracted. Figure 5 demonstrates lesions extracted from US images by means of gradient information (16,72,75). Assuming that lesions are roughly spherical, lesion extraction is possible even in the presence of acoustic shadowing. Mass lesions visible at US can be classified with a computer by using a variety of mathematical descriptors of texture, margin, shape, and acoustic shadowing criteria (69–76).

As with CAD systems for mammography, the effect on radiologists' performance must also be assessed to determine the additive value of CAD systems for characterizing breast lesions at US. In one study, investigators demonstrated that the use of computer classification schemes to distinguish between benign and malignant masses helped inexperienced operators avoid misdiagnosis (71). In another study, involving six expert mammographers and six community-based radiologists, investigators demonstrated that the use of computer-estimated probabilities of malignancy
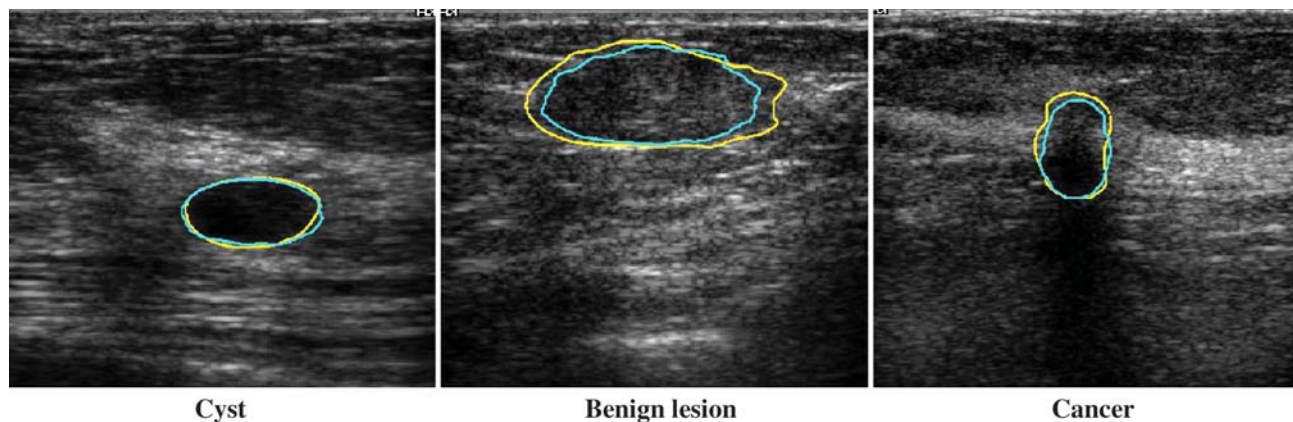
Figure 5. Manual (yellow) and computer (blue) segmentation of three lesions at US. (Courtesy of K. Horsch and K. Drukker, University of Chicago, Chicago, Ill.)

significantly improved radiologists' performance in interpreting breast US images (76). Interestingly, the performance of the computer alone was similar to that of the radiologists (76).

Current research trends in US CAD include classification of multiple lesion types (malignant, solid benign, and cysts) and incorporation of three-dimensional US imaging. For US CAD, a computerized characterization method needs to be evaluated for its ability to distinguish between malignant lesions, solid benign lesions, and cysts; there is a three-way classification problem (67,77). The advent of three-dimensional US allows breast masses to be characterized on the basis of three-dimensional US volumes. Investigators have used two- and three-dimensional active contour models for the automated segmentation of mass volumes (74). On a data set of three-dimensional volumes of 102 mass lesions subjected to biopsy, investigators evaluated their two- and three-dimensional segmentation methods, reporting A_z values of 0.87 and 0.92, respectively.

Future directions in breast US CAD also include *(a)* further development and evaluation in three-dimensional US, *(b)* assessment of the role of CAD as an aid to radiologists, and *(c)* studies of the robustness of US CAD. US is quite an operator-dependent modality, and computer performance must therefore be assessed with databases from different institutions and across acquisition platforms. Drukker et al (75,78) developed a method for detecting and characterizing masses with two-dimensional breast US. Results from independent validation tests indicated that the methods for distinguishing actual lesions from false-positive detections and distinguishing between malignant and benign lesions were robust across different institutions and across two different manufacturers' acquisition platforms.

CAD in Breast MR Imaging

MR imaging may aid in the diagnosis of breast cancer as a complement to mammography. MR imaging has the benefit of yielding three-dimensional spatial information on the breast as well as temporal information. MR imaging of the breast with contrast material enhancement is used to indicate differences between lesions and normal tissue caused by the increased vascularity and capillary permeability of tumors. Computerized analysis of MR images has potential benefits because radiologists vary in their assessment of lesions at MR imaging and lack standard imaging protocols for breast MR imaging. Mussurakis et al (79) reported that there is considerable variability in human assessment of lesions on MR images. With the variation in criteria and the absence of guidelines for interpretation being major obstacles to applying MR imaging in routine clinical practice, investigators aim to develop CAD for breast MR imaging in order to increase the objectivity and reproducibility of interpretation. With MR imaging, both temporal and spatial features of lesions, as well as combined features, may be useful in the discrimination task.

One computerized method employed a decision-tree approach to interpreting mammographic architectural features (80), with lesion margin being the primary determinant of malignancy. Others have investigated use of the fractal dimension of the mass boundary to characterize margin roughness (81). Combining fractal analysis with prior architectural features (81) substantially improved lesion characterization.

At the initial stage, quantitative analyses of MR lesions involve manual or automatic lesion segmentation, which can be performed in either two or three dimensions. In one study, examples of temporal features included maximum uptake of contrast material (82). Spatial features included lesion contrast with respect to surrounding background, inhomogeneity of voxel values in the lesion, mean sharpness and variation of sharpness along the lesion margins, circularity and irregularity of the lesion, texture, and directional analysis of the gradients in the lesion. To take full advantage of the data, the spatial features can be computed in three dimensions.

In an initial study, the performance of a radial gradient analysis and margin sharpness (combined by means of a linear discriminant function) were used to distinguish between malignant and benign lesions, yielding an A_z of 0.96 in round-robin ROC analysis (83). Additional studies involved the investigation of three categories of features: morphology, enhancement kinetics, and variance dynamics over the lesion (84). In a database of 121 cases (77 malignant and 44 benign), stepwise feature selection was employed to select an effective subset of features. These were then combined by means of linear discriminant analysis into a discriminant score, which was related to the likelihood of malignancy. With radiologist-delineated lesion contours, stepwise feature selection yielded four features and an A_z of 0.80 for the linear discriminant analysis in leave-one-out cross-validation testing, compared with six features and an A_z of 0.86 for automatically computer-segmented lesion volumes.

CAD in Multimodality Breast Imaging

Although some features encountered in breast imaging, such as margin sharpness, can be seen across modalities, others are particular to the imaging modality, such as the computer characterization of the posterior acoustic behavior at US and the inhomogeneity of contrast material uptake in breast MR imaging. Radiologists, for example, rarely interpret a US image without a corresponding mammogram, and thus it is important to assess the usefulness of CAD on multimodality images in combination. Combining computer analyses from multiple views or modalities requires *(a)* additional computer intelligence to determine the correspondence of lesions across views and modalities (85) and *(b)* an efficient means of conveying multimodality information to the radiologist, as discussed in the next section.

Human-Computer Interface for CAD and Indexes of Similarity

For diagnostic output, a computer-determined estimate of the probability of malignancy for a specific case can be given to the radiologist. This estimated probability could then be used by the radiologist in deciding on the next step in patient management. Swett and Miller (86) reported on an expert system that used a critiquing approach combined with an expert system to control the image display. Sklansky et al (87) designed a mapped database diagnostic system in which two-dimensional relation maps of known cases are accessible for comparison to the feature values of some unknown case. Giger et al (88,89) presented a CAD system with output in the form in three modes: a numerical mode giving values corresponding to the probability of malignancy, a pictorial mode displaying similar images of known diagnoses, and a graphic mode that indicates on a distribution plot the unknown lesion relative to all lesions in a known library (an online atlas).

Output from multimodality computer image analyses parallels the interpretations of radiologists who interpret whole cases rather than individual images. Such multimodality output, however, requires effective and efficient displays to communicate the output to the radiologist. For example, Figure 6 shows an interface for the output of a CAD system that is presented for different imaging modalities in terms of *(a)* computer-estimated probabilities of malignancy, *(b)* the display of similar images of known diagnoses, and *(c)* a graphic representation of the unknown lesion relative to all lesions in a known database (an online atlas) (88–90).

A known image atlas can be searched according to individual features or the likelihood of malignancy. The similarity index used in such a search can be selected by the radiologist from a variety of possibilities, such as the computer-estimated probability of malignancy or computer-extracted features (eg, margin definition, spiculation, contrast). Note that in Figure 6, color coding is used to expedite the extraction of information from the interface, with the malignant lesions from the online reference library outlined in red and the benign lesions outlined in green. The similarity index used in the search can be chosen by the radiologist to be based on a single feature, multiple features, or the computer estimate of the likelihood of malignancy. The probability distributions of the malignant and benign cases in the known database are shown along with the "location" of the unknown case relative to the two distributions. The intelligent search workstation combines the benefit of CAD with prior knowledge obtained from confirmed cases. The potential of this interface has been demonstrated as an aid in the diagnostic interpretation of lesions by radiologists for both mammography alone (90) and combined display of mammography and US (91).

In summary, the potential for improved diagnostic performance by radiologists in the task of interpreting mammograms provides strong motivation for CAD methods. The performance of computer algorithms continues to improve, and attention is now also directed to integrating computers in the clinical arena. Careful evaluation of system performance will require *(a)* comparisons of sensitivity and specificity across similar databases, *(b)* ultimate improvement in clinical practice, and *(c)* ease of integration into the breast imaging practices (92). Although the roles of computers in mammographic detection and diagnosis are still being clarified and validated, the ultimate effect on practice may one day be both substantial and commonplace. While the breadth and depth of CAD in breast cancer imaging are increasing, continued and expanded efforts are needed for establishing and confirming databases and providing means for clinical validation.

Figure 6. Display interface for a multimodality workstation that displays computer outputs in numerical, pictorial, and graphic modes. **(a)** Mammography CAD output *(Fig 6 continues)*.

References

1. Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. Radiology 1994; 191:241–244.
2. Elmore JG, Wells CK, Lee CH, et al. Variability in radiologists' interpretations of mammograms. N Engl J Med 1994; 331:1493–1499.
3. Beam C, Layde PM, Sullivan D. Variability in the interpretation of screening mammograms by U.S. radiologists. Arch Intern Med 1996; 156:209–213.
4. Beam C, Sullivan D, Layde P. Effect of human variability on independent double reading in screening mammography. Acad Radiol 1996; 3:891–897.
5. Giger ML. Computer-aided diagnosis. In: Haus AG, Yaffe MJ, eds. Syllabus: a categorical course in physics—technical aspects of breast imaging. 2nd ed, updated. Oak Brook, Ill: Radiological Society of North America, 1993; 283–298.
6. Giger ML. Computer-aided diagnosis. In: Haus AG, Yaffe MJ, eds. Syllabus: a categorical course in diagnostic radiology physics—physical aspects of breast imaging: current and future considerations. Oak Brook, Ill: Radiological Society of North America, 1999; 249–272.
7. Giger ML, Huo Z, Kupinski MA, Vyborny CJ. Computer-aided diagnosis in mammography. In: Sonka M, Fitzpatrick MJ, eds. Handbook of medical imaging: medical imaging processing and analysis. Vol 2. Bellingham, Wash: SPIE, 2000; 915–1004.
8. Vyborny CJ, Giger ML. Computer vision and artificial intelligence in mammography. AJR Am J Roentgenol 1994; 162:699–708.
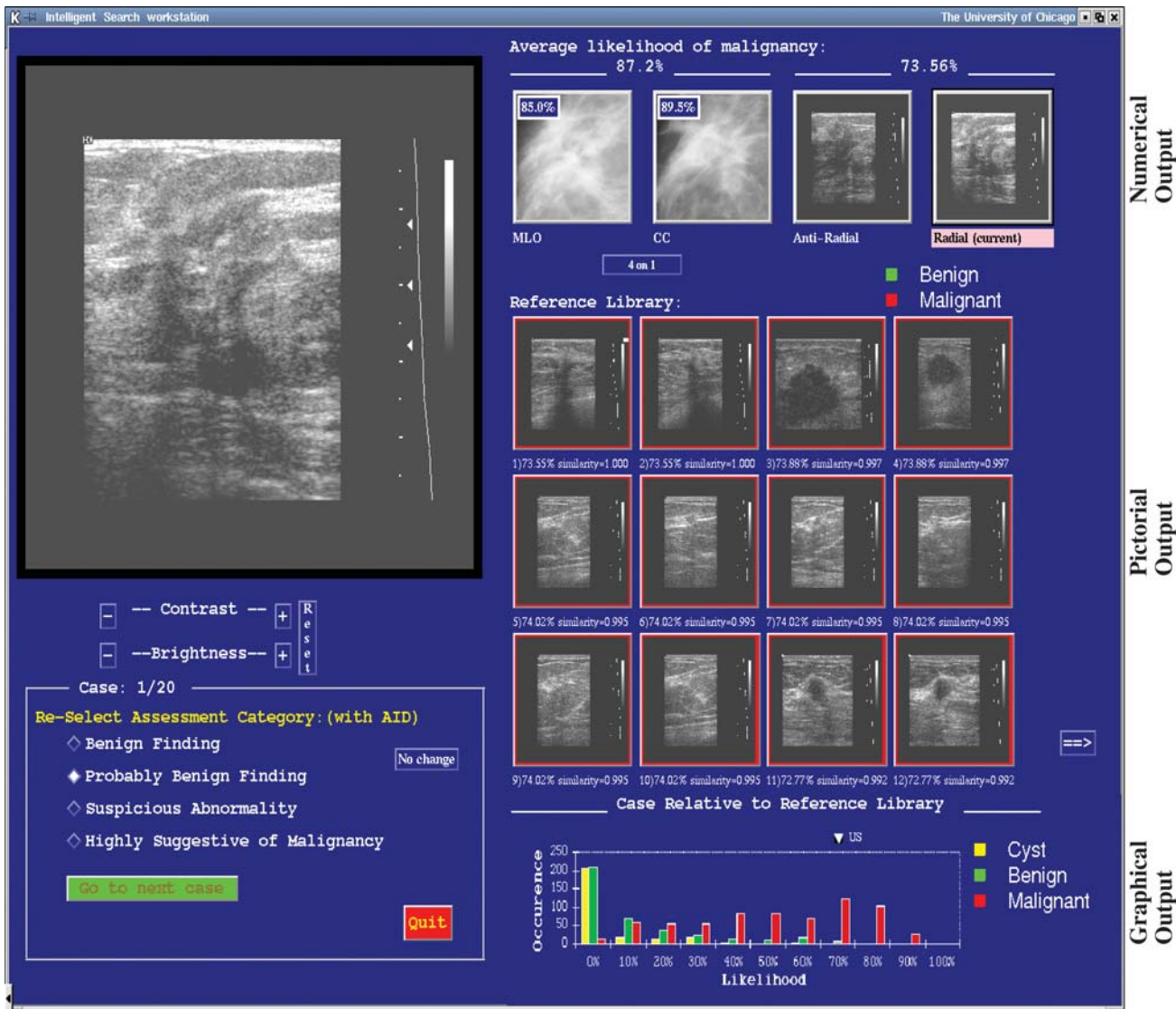9. Vyborny CJ, Giger ML, Nishikawa RM. Computer-aided detection and diagnosis. Radiol Clin North Am 2000; 38:725–740.

Figure 6 *(continued).* **(b)** US CAD output.

10. Jiang Y. Classification of breast lesions in mammograms. In: Bankman I, ed. Handbook of medical imaging: processing and analysis management. New York, NY: Academic Press, 2000; 341–358.

11. Giger ML. Computerized image analysis in breast cancer detection and diagnosis. Semin Breast Dis 2002; 5:199–210.

12. Ballard DH, Brown CM. Computer vision. Englewood Cliffs, NJ: Prentice-Hall, 1982.

13. Sonka M, Hlavac V, Boyle R. Image processing, analysis, and machine vision. Pacific Grove, Calif: Brooks/Cole, 1999.

14. Bick U, Giger ML, Schmidt RA, Nishikawa RM, Wolverton DE, Doi K. Automated segmentation of digitized mammograms. Acad Radiol 1995; 2:1–9.

15. Mendez AJ, Tahoces PG, Lado MJ, Souto M, Correa JL, Vidal JJ. Automatic detection of breast border and nipple in digital mammograms. Comput Methods Programs Biomed 1996; 49:253–262.

16. Kupinski MA, Giger ML. Automated seeded lesion segmentation on digital mammograms. IEEE Trans Med Imaging 1998; 17:510–517.

17. Giger ML, Vyborny CJ, Schmidt RA. Computerized characterization of mammographic masses: analysis of spiculation. Cancer Lett 1994; 77:201–211.

18. Huo Z, Giger ML, Vyborny CJ, et al. Analysis of spiculation in the computerized classification of mammographic masses. Med Phys 1995; 22:1569–1579.

19. Karssemeijer N, te Brake GM. Detection of stellate distortions in mammograms. IEEE Trans Med Imaging 1996; 15:611.

20. Kegelmeyer WP Jr, Pruneda JM, Bourland PD, Hillis A, Riggs MW, Nipper ML. Computer-aided mammographic screening for spiculated lesions. Radiology 1994; 191: 331–337.

21. Sahiner B, Chan HP, Petrick N, Helvie MA, Goodsitt MM. Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis. Med Phys 1998; 25:516–526.

22. Kupinski MA, Giger ML. Feature selection with limited datasets. Med Phys 1999; 26:2176–2182.

23. Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA. Computer-aided detection of clustered microcalcifications on digital mammograms. Med Biol Eng Comput 1995; 33: 174–178.

24. Polakowski WE, Cournoyer DA, Rogers SK, et al. Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency. IEEE Trans Med Imaging 1997; 16:811–819.

25. Bankman N, Christens-Barry WA, Kim DW, Weinberg IN, Gatewood OB, Brody WR. Automated recognition of microcalcification clusters in mammograms. Proc SPIE 1993; 1905.

26. Kalman BL, Reinus WR, Kwasny SC, Laine A, Kotner L. Prescreening entire mammograms for masses with artificial neural networks: preliminary results. Acad Radiol 1997; 4:405–414.

27. Kupinski MA, Giger ML. Investigation of regularized neural networks for the computerized detection of mass lesions in digital mammograms. In: Proceedings of the IEEE Engineering Medicine and Biology Conference. 1997; 1336–1339.

28. Nagel RH, Nishikawa RM, Doi K. Analysis of methods for reducing false positives in the automated detection of clustered microcalcifications in mammograms. Med Phys 1998; 25:1502–1506.

29. Cheng HD, Lui YM, Freimanis RI. A novel approach to microcalcification detection using fuzzy logic technique. IEEE Trans Med Imaging 1998; 17:442–450.

30. Sahiner B, Chan HP, Petrick N, Helvie MA, Goodsitt MM. Design of a high-sensitivity classifier based on a genetic algorithm: application to computer-aided diagnosis. Phys Med Biol 1998; 43:2853–2871.

31. Anastasio MA, Yoshida H, Nagel R, Nishikawa RM, Doi K. A genetic algorithm-based method for optimizing the performance of a computer-aided diagnosis scheme for detection of clustered microcalcifications in mammograms. Med Phys 1998; 25:1613–1620.

32. Kupinski MA, Giger ML, Doi K. Optimization of neural network inputs with genetic algorithms. In: Doi K, Giger ML, Nishikawa RM, Schmidt RA, eds. Digital mammography 96. Amsterdam, the Netherlands: Elsevier, 1996; 401–404.

33. Vyborny CJ. Image and observer issues in mammography. Semin Breast Dis 2002; 5:190–198.

34. Nishikawa RM. Assessment of the performance of computer-aided diagnosis and computer-aided diagnosis systems. Semin Breast Dis 2002; 5:217–222.

35. Kegelmeyer WP Jr, Pruneda JM, Bourland PD, et al. Computer-aided mammographic screening for spiculated lesions. Radiology 1994; 191:331–337.

36. Chan HP, Doi K, Vyborny CJ, et al. Improvements in radiologist's detection of clustered microcalcifications on mammograms: the potential of computer-aided diagnosis. Invest Radiol 1990; 25:1102.

37. Roehrig J, Doi T, Hasegawa A, et al. Clinical results with R2 imagechecker system. In: Karssemeijer N, et al, eds. Digital mammography 1998. Dordrecht, the Netherlands: Kluwer Academic, 1998; 395.

38. Schmidt RA, Nishikawa RM, Osnis RB, et al. Computerized detection of lesions missed by mammography. In: Doi K, Giger ML, Nishikawa RM, Schmidt RA, eds. Digital mammography 96. Amsterdam, the Netherlands: Elsevier Science, 1997; 105.

39. Warren Burhenne LJ, Wood SA, D'Orsi CJ, et al. Potential contribution of computer-aided detection to the sensitivity of screening mammography. Radiology 2000; 215:554–562.

40. Birdwell RL, Ikeda DM, O'Shaughnessy KF, Sickles EA. Mammographic characteristics of 115 missed carcinomas later detected with screening mammography and the potential utility of computer-aided detection. Radiology 2001; 219: 192–202.

41. te Brake GM, Karssemeijer N, Hendriks JH. Automated detection of breast carcinomas not detected in a screening program. Radiology 1998; 207:465–471.

42. Freer TW, Ulissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology 2001; 220:781–786.

43. Gur D, Sumkin JH, Rockette HE, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. J Natl Cancer Inst 2004; 96:185–190.

44. Feig SA. Clinical evaluation of computer-aided detection in breast cancer screening. Semin Breast Dis 2002; 5: 223–230.

45. Zheng B, Ganott MA, Britton CA, et al. Soft-copy mammographic readings with different computer-assisted detection cuing environments: preliminary findings. Radiology 2001; 221:633–640.

46. Tabar L, Dean PB. Teaching atlas of mammography. New York, NY: George Thieme, 1983.

47. Ciccone G, Vineis P, Frigerio A, Segnan N. Inter-observer and intra-observer variability of mammogram interpretation: a field study. Eur J Cancer 1992; 28A:1054–1058.

48. Ackerman LV, Mucciardi AN, Gose EE, et al. Classification of benign and malignant breast tumors on the basis of 36 radiographic properties. Cancer 1973; 31:342–352.

49. Getty DJ, Pickett RM, D'Orsi CJ, Swets JA. Enhanced interpretation of diagnostic images. Invest Radiol 1988; 23:240–252.

50. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. Radiology 1993; 187:81–87.

51. Baker JA, Kornguth PJ, Lo JY, Williford ME, Floyd CE Jr. Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon. Radiology 1995; 196:817–822.

52. Ackerman LV, Gose EE. Breast lesion classification by computer and xeroradiography. Cancer 1972; 30:1025–1035.

53. Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K. Automated computerized classification of malignant and benign mass lesions on digitized mammograms. Acad Radiol 1998; 5:155–168.

54. Huo Z, Giger ML, Metz CE. Effect of dominant features on neural network performance in the classification of mammographic lesions. Phys Med Biol 1999; 44:2579–2595.

55. Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Metz CE. Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness. Acad Radiol 2000; 7:1077–1084.

56. Huo Z, Giger ML, Vyborny CJ. Computerized analysis of multiple-mammographic views: potential usefulness of special view mammograms in computer-aided diagnosis. IEEE Trans Med Imaging 2001; 20:1285–1292.

57. Chan HP, Wei D, Helvie MA, et al. Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. Phys Med Biol 1995; 40:857–876.

58. Chan HP, Sahiner B, Lam KL, et al. Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces. Med Phys 1998; 25:2007–2019.

59. Sahiner B, Chan HP, Petrick N, Helvie MA, Goodsitt MM. Design of a high-sensitivity classifier based on a genetic algorithm: application to computer-aided diagnosis. Phys Med Biol 1998; 43:2853–2871.

60. Patrick EA, Moskowitz M, Mansukhani VT, Gruenstein EI. Expert learning system network for diagnosis of breast calcifications. Invest Radiol 1991; 26:534–539.

61. Jiang Y, Nishikawa RM, Wolverton DE, et al. Malignant and benign clustered microcalcifications: automated feature analysis and classification. Radiology 1996; 198:671–678.

62. Jiang Y, Nishikawa RM, Schmidt RA, Metz CE, Giger ML, Doi K. Improving breast cancer diagnosis with computer-aided diagnosis. Acad Radiol 1999; 6:22–33.

63. Jiang Y. Uncertainty in the output of artificial neural networks. IEEE Trans Med Imaging 2003; 22:913–921.

64. Huo Z, Giger ML, Vyborny CJ, Metz CE. Breast cancer: effectiveness of computer-aided diagnosis—observer study with independent database of mammograms. Radiology 2002; 224:560–568.

65. Chan HP, Sahiner B, Helvie MA, et al. Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. Radiology 1999; 212:817–827.

66. Jiang Y, Nishikawa RM, Schmidt RA, Toledano AY, Doi K. Potential of computer-aided diagnosis to reduce variability in radiologists' interpretations of mammograms depicting microcalcifications. Radiology 2001; 220:787–794.

67. Edwards DC, Lan L, Metz CE, Giger ML, Nishikawa RM. Estimating three-class ideal observer decision variables for computerized detection and classification of mammographic mass lesions. Med Phys 2004; 31:81–90.

68. Stavros AT, Thickman D, Rapp CL, Dennis MA, Parker SH, Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology 1995; 196:123–134.

69. Garra BS, Krasner BH, Horii SC, Ascher S, Mun SK, Zeman RK. Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. Ultrason Imaging 1993; 15:267–285.

70. Giger ML, Al-Hallaq H, Huo Z, et al. Computerized analysis of lesions in US images of the breast. Acad Radiol 1999; 6:665–674.

71. Chen DR, Chang RF, Huang YL. Computer-aided diagnosis applied to US of solid breast nodules by using neural networks. Radiology 1999; 213:407–412.

72. Horsch K, Giger ML, Venta LA, Vyborny CJ. Automatic segmentation of breast lesions on ultrasound. Med Phys 2001; 28:1652–1659.

73. Horsch K, Giger ML, Venta LA, Vyborny CJ. Computerized diagnosis of breast lesions on ultrasound. Med Phys 2002; 29:157–164.

74. Sahiner B, Chan HP, Roubidoux MA, et al. Computerized characterization of breast masses using three-dimensional ultrasound volumes. Med Phys 2004; 31:744–754.

75. Drukker K, Giger ML, Horsch K, Kupinski MA, Vyborny CJ, Mendelson EB. Computerized lesion detection on breast ultrasound. Med Phys 2002; 29:1438–1446.

76. Horsch K, Giger ML, Vyborny CJ, Huo Z, Venta LA. Performance of computer-aided diagnosis in the interpretation of lesions on breast sonography. Acad Radiol 2004; 11:272–280.

77. Drukker K, Edwards DC, Giger ML, Metz CE. Computerized 3-way classification of lesions using breast ultrasound and mammography. In: Pisano E, ed. Proceedings of International Workshop on Digital Mammography (in press).

78. Drukker K, Giger ML, Vyborny CJ, Mendelson E. Computerized detection and classification of cancer on breast ultrasound. Acad Radiol 2004; 11:526–535.

79. Mussurakis S, Buckley DL, Coady AM, Turnbull LW, Horseman A. Observer variability in the interpretation of contrast enhanced MRI of the breast. Br J Radiol 1996; 69:1009–1016.

80. Nunes LW, Schnall MD, Orel SG, et al. Breast MR imaging: interpretation model. Radiology 1997; 202:833–841.

81. Penn AI, Bolinger L, Schnall MD, Loew MH. Discrimination of MR images of breast masses with fractal-interpolation function models. Acad Radiol 1999; 6:156–163.

82. Gilhuijs KG, Giger ML, Bick U. Automated analysis of breast lesions in three dimensions using dynamic magnetic resonance imaging. Med Phys 1998; 25:1647–1654.

83. Gilhuijs KG, Giger ML, Bick U. Automated feature extraction and classification of breast lesions in magnetic resonance images. Proc SPIE 1998; 3338:294–300.

84. Chen W, Giger ML, Lan L, Bick U. Computerized interpretation of breast MRI: investigation of enhancement-variance dynamics. Med Phys 2004; 31:1076–1082.

85. Wen HH, Giger ML, Horsch K, Hendrick RE, Vyborny CJ, Lan L. Correlation of lesions from multiple images for CAD. Proc SPIE Med Imaging 2004; 5370:93–96.

86. Swett HA, Miller PA. ICON: a computer-based approach to differential diagnosis in radiology. Radiology 1987; 163: 555–558.

87. Sklansky J, Tao EY, Ornes C, Disher AC. A visualized mammographic database in computer-aided diagnosis. In: Doi K, MacMahon H, Giger ML, Hoffmann KR, eds. Computer-aided diagnosis in medical imaging. Amsterdam, the Netherlands: Elsevier, 1999; 215.

88. Giger ML, Huo Z, Lan L, Vyborny CJ. Intelligent search workstation for computer-aided diagnosis. In: Lemke HU, Vannier MW, Inamura K, et al, eds. CARS 2000: computer assisted radiology and surgery. Amsterdam, the Netherlands: Excerpta Medica, 2000; 822–827.

89. Giger ML, Huo Z, Vyborny CJ, et al. Intelligent CAD workstation for breast imaging using similarity to known lesions and multiple visual prompt aids. Proc SPIE Med Imaging 2002; 4684:768–773.

90. Giger ML, Hui Z, Vyborny CJ, Lan L, Nishikawa RM, Rosenbourgh I. Results of an observer study with an intelligent mammographic workstation for CAD. In: Peitgen HO, ed. Digital mammography IWDM 2002. Berlin, Germany: Springer, 2003; 297–303.

91. Giger ML, Horsch KJ, Vyborny CJ, Lan L. Performance of multimodality CAD for breast cancer diagnosis (abstr). Radiology 2003; 229(P):216.

92. Giger ML, Karssemeijer N, Armato SG III. Computer-aided diagnosis in medical imaging. IEEE Trans Med Imaging 2001; 20:1205–1208.