J Digit Imaging (2012) 25:570–579
DOI 10.1007/s10278-012-9451-0

An Interactive System for Computer-Aided Diagnosis of Breast Masses

Xingwei Wang · Lihua Li · Wei Liu · Weidong Xu · Dror Lederman · Bin Zheng

Published online: 11 January 2012

Abstract Although mammography is the only clinically accepted imaging modality for screening the general population to detect breast cancer, interpreting mammograms is difficult with lower sensitivity and specificity. To provide radiologists "a visual aid" in interpreting mammograms, we developed and tested an interactive system for computer-aided detection and diagnosis (CAD) of mass-like cancers. Using this system, an observer can view CAD-cued mass regions depicted on one image and then query any suspicious regions (either cued or not cued by CAD). CAD scheme automatically segments the suspicious region or accepts manually defined region and computes a set of image features. Using content-based image retrieval (CBIR) algorithm, CAD searches for a set of reference images depicting "abnormalities" similar to the queried region. Based on image retrieval results and a decision algorithm, a classification score is assigned to the queried region. In this study, a reference database with 1,800 malignant mass regions and 1,800 benign and CAD-generated false-positive regions was used. A modified CBIR algorithm with a new function of stretching the attributes in the multi-dimensional space and decision scheme was optimized using a genetic algorithm. Using a leave-one-out testing method to classify suspicious mass regions, we compared the classification performance using two CBIR algorithms with either equally weighted or optimally stretched attributes. Using the modified CBIR algorithm, the area under receiver operating characteristic curve was significantly increased from 0.865 ±0.006 to 0.897±0.005 ($p<0.001$). This study demonstrated the feasibility of developing an interactive CAD system with a large reference database and achieving improved performance.

Keywords Computer-aided detection and diagnosis (CAD) · Content-based image retrieval (CBIR) · Breast cancer · Mammograms

X. Wang · D. Lederman · B. Zheng (✉)
Department of Radiology, University of Pittsburgh,
3362 Fifth Avenue,
Pittsburgh, PA 15213, USA
e-mail: zhengb@upmc.edu

L. Li · W. Liu · W. Xu
College of Life Information Science and Instrument Engineering,
Hangzhou Dianzi University,
Hangzhou 310018, China

Introduction

Breast cancer is one of the leading cancers in women over 40 years old worldwide [1]. Scientific research has found that earlier detection of breast cancer could reduce not only treatment cost but also patients' mortality and morbidity rates [2]. Although a number of screening tools such as breast magnetic resonance imaging [3, 4] have been approved and used as adjunct screening tools for the high-risk women identified by epidemiology-based risk models [5], periodic screening with mammography is the only clinically accepted imaging modality for screening the general population to date. However, interpreting mammograms by radiologists is difficult and time-consuming due to the low cancer prevalence in the screening environment (i.e., less than three to five cancers detected per every 1,000 screening examinations), as well as a large variability of depicted breast abnormalities and overlapped dense fibro-glandular tissue on the two-dimensional projected images [6, 7]. As a result, both detection sensitivity and specificity of screening mammography are not satisfactory [8] particularly in younger women with dense breasts and in other high-risk groups.

For example, the reported sensitivity levels ranged from 16% to 40% for women carrying one of two breast cancer susceptibility genes namely BRCA1 and BRCA2 [9]. Meanwhile, a large fraction of women are likely to be recalled for additional imaging work-up and/or biopsy due to the indecisive interpretation in the screening environment [10]. Among the recalled cases, only a small fraction of biopsies (15–30%) prove to be malignant [11]. Hence, to help radiologists improve detection and diagnosis performance, a large number of computer-aided detection and/or diagnosis (CAD) schemes have been developed and tested [12]. The commercialized computer-aided detection systems are currently available and routinely used as "a second reader" in a large number of medical institutions.

While studies have showed that using CAD helped radiologists detect more cancers associated with micro-calcification clusters [13, 14], current CAD has no or little impact in helping radiologists detecting more subtle cancers associated with mass-like abnormalities [15, 16]. Studies have shown that due to the inability to explain the reasoning of the CAD decision making (the "black-box" type approach) and the higher false-positive cueing rates, radiologists in general had very low confidence in accepting or considering CAD-cued mass-like abnormalities [17]. As a result, although CAD is able to detect a fraction of cancers missed by radiologists albeit with a relatively high false-positive rates [18], most of these subtle masses are cued by CAD only on one view and thus are discarded by radiologists as false-positives [19, 20]. As important, improperly using CAD (e.g., as "the first reader") could not only increase radiologists' recall rates but also reduce their detection sensitivity in the non-cued regions [21].

To improve the clinical utility of CAD, a number of research groups have been actively working on developing a new type of CAD systems that enable to provide radiologists the reasoning information of CAD decision making and "visual aid" [22]. For this purpose, the content-based image retrieval (CBIR) methods previously applied in picture archiving and communication systems and medical informatics [23, 24] have also been integrated into CAD schemes [25–29]. Unlike conventional CAD schemes that detect and cue the suspicious abnormalities based on a "global" optimization function trained using the entirely available image database, the new CAD schemes using a CBIR approach applies an adaptive approach to generate each detection and/or diagnostic result based on the selection of local approximations as the target function for each unknown query of a suspicious testing region. Specifically, for each initially detected suspicious "lesion" depicted on one testing image, the CBIR-based CAD scheme segments the suspicious region and computes the likelihood of this "lesion" being positive (i.e., cancer). Each likelihood score is computed based on comparing the queried region with a set of "similar lesions" retrieved by CBIR algorithm from a pre-established reference database in which the status of all selected references ("similar lesions") has been previously verified. In addition, using the retrieved "similar" reference images, researchers have also developed the interactive CAD systems that can provide radiologists "visual aid" by displaying both CAD-generated detection and/or diagnostic scores as well as the retrieved "similar" reference images on the workstation screens. The feasibility of using CBIR algorithms to select visually similar breast lesions has also been previously demonstrated [30]. Such "a visual aid" approach aims to increase radiologists' confidence in accepting or considering CAD-cued results in their decision making and thus to improve radiologists' performance in diagnosis of breast cancer using mammograms. In this article, we presented a new interactive CAD system with several unique functions and tested a new approach to optimize CBIR algorithm to further improve CAD classification performance. The detailed description in system design, experimental procedure, along with the study results and conclusions, is presented in the following sections.

Materials and Methods

An Overview of an Interactive CAD System

Figure 1 is a simplified dataflow diagram showing the concept and design of our new interactive CAD system. A conventional CAD scheme [31] is pre-installed in the system, which can preprocess the images of interest stored in the system. For each initially detected suspicious region, this CAD scheme uses a multi-feature-based artificial neural network (ANN) to generate a detection score indicating the likelihood of the region being associated with a true-positive mass. If the detection score is greater than a predetermined CAD operating threshold, the region is cued (marked) on the image. Similar to current commercialized CAD systems, the detection score of this cued region is not provided. If the detection score is smaller than CAD operating threshold, the initially detected suspicious region is discarded (not cued). One primary innovation (unique function) of our interactive CAD system is that it allows observers (e.g., radiologists) to query any undetermined suspicious regions either cued or not cued by the conventional CAD scheme. The new CAD scheme implemented in our interactive CAD system will conduct more detailed analysis and classification of the queried region. For this purpose, our interactive CAD system includes the following

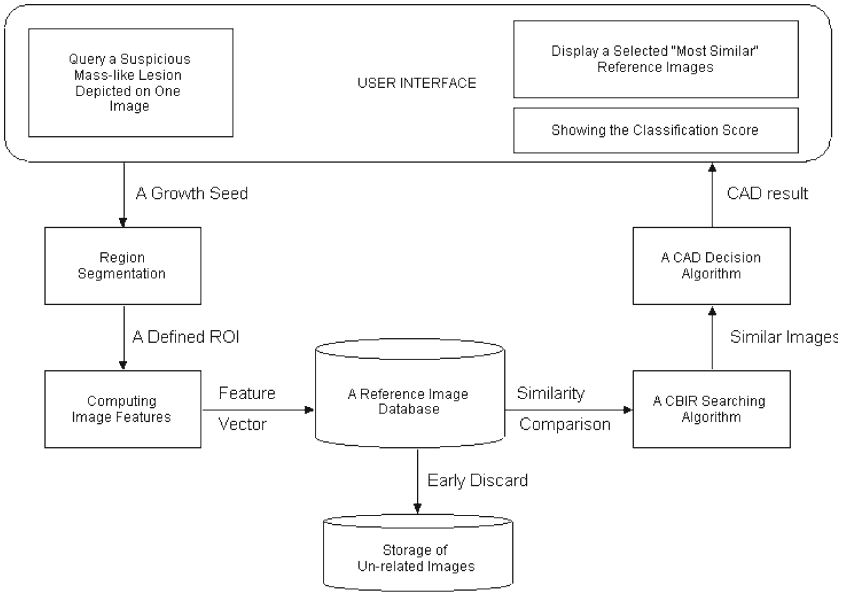

Fig. 1 Illustration of an interactive system for computer-aided detection and diagnosis of mass-like breast abnormalities

primary components (functions) and processing steps to response the queries generated by the observers.

1. After reading and interpreting images of a test case and then viewing regions cued by the conventional CAD scheme, an observer can query any identified undetermined suspicious mass-like regions depicted on the images (no matter whether these are CAD-cued regions or not) by moving the computer mouse cursor inside the region and clicking the mouse button to send the command of processing and analyzing the targeted region to CAD scheme.
2. Once a suspicious region is queried, CAD scheme automatically searches for and detects a pixel with the minimum intensity value around the queried point as an initial seed of a suspicious region. From the detected growth seed, CAD scheme applies a region growth algorithm that combines a multi-layer topographic region growth and an active contour approach [28] to segment the suspicious mass region. The computed boundary contour of the segmented region will be displayed on the image. Since the accuracy of suspicious region segmentation determines the accuracy of extracted and computed image features used in CBIR algorithm, if the observer does not satisfy with the automated segmentation results, he/she can reject the automated results and manually draw the boundary contour of the suspicious region.
3. CAD scheme extracts and computes a set of 14 image features from the segmented region and its surrounding background. A pre-optimized artificial neural network based classifier [31] implemented in the conventional CAD scheme is called again to process this region and compute a detection score. This score is not provided and recorded in current CAD-cueing. In our interactive CAD system, this detection score will also provide to the observers showing the likelihood of this region associating with a true-positive mass based on a globalized optimization function. If the queried suspicious region is not originally cued by CAD, providing this detection score can also explain why this "subtle" mass region is initially detected but not cued (discarded). For example, the ANN-generated detection score of this suspicious region is lower than CAD-cueing threshold.
4. CAD scheme uses an early discard function reported in our previous study [32] to remove a substantial fraction of reference region of interests (ROIs) stored in the reference image database from the detailed analysis of using CBIR algorithm. Although the performance and robustness of all CAD schemes heavily depends on the size and diversity of reference databases, the CBIR-based CAD scheme uses an adaptive and "local" approximation method in which only a small number of references have direct impact on the detection or classification of the specifically queried ROI. Thus, early discarding the majority of reference ROIs with relatively lower correlation to the queried ROI is an important step to improve the computational efficiency of the CAD scheme as well as to reduce the "semantic gap" between computer vision and human vision [33].
5. CAD scheme applies a CBIR algorithm to compare and retrieve a set of K reference images (ROIs) that are considered the "most similar" to the queried ROI. The searching and retrieving result of the CBIR algorithm depends on the effectiveness of the summary index or criterion (i.e., distance metric) to measure the

"similarity" level among the selected images. A new approach to measure similarity level was tested in this study, which will be discussed later in "A CBIR Algorithm".

6. Based on the similarity levels of the retrieved reference ROIs to the queried ROI and their verified outcomes (i.e., either positive or negative), CAD scheme computes a new classification score of the queried ROI. It indicates the likelihood of this region associating with a malignant mass. This new classification score often does not correlate with the detection score computed by the ANN of the conventional CAD scheme due to the use of different global- and local-based optimization approaches [34].
7. The interactive CAD system then displays a set of retrieved "most similar" reference images and two scores of the queried region including the detection score generated by the conventional CAD scheme and the classification score generated by the new CBIR-based CAD scheme on workstation monitor screen. These displayed CAD results provide the radiologist a "visual aid" to interpret mammograms and diagnose the targeted suspicious or undetermined mass regions.

A Reference Image Database

One of the most important components in an interactive CAD system is to build a large and diverse reference image database (or library). In our research laboratory, we have gradually assembled and expanded a reference database by selecting and extracting regions of interest from the digitized screen-film-based mammograms. The detailed description of the method to build the reference database and the image characteristics of the selected ROIs have been reported in our previous studies [32, 33]. In brief, the original screen-film mammograms were collected from several medical institutions and they were digitized using several film digitizers with a pixel size of approximately 50×50 μm and 12-bit gray-level resolution. From the center of an identified suspicious mass region, a region of interest with a fixed size of 512×512 pixels was extracted. Using the ROI center pixel as a mass region growth seed, a CAD scheme was applied to segment the mass region and define region boundary contour. For each true-positive mass (either malignant or benign) region, the automated segmentation result was visually examined. If the noticeable segmentation error was identified, the mass boundary contour was manually corrected (re-drawn). Approximately half of verified masses were subjectively rated by radiologists as "subtle" to "very subtle" (namely, 4 and 5 in a five rating scale). Each non-mass depicted negative ROI selected in our reference database contains a false-positive mass that is automatically segmented and cued by the CAD scheme. In this study, the selected reference database includes 3,600 ROIs. Among them, 1,800 are true-positive regions that depict a verified malignant mass each and the remaining 1,800 ROIs are false-positive (cancer-free) regions in which 300 depict benign masses and 1,500 contain CAD-cued false-positive mass-like regions.

A computer scheme was then applied to compute 14 morphological and intensity (pixel value) distribution features from each segmented mass-like region (including both true-positive and false-positive regions). Specifically, these features include three global features computed from the whole breast area segmented from the image and 11 local features computed from the segmented mass region and its surrounding background inside the extracted ROI (as shown in Table 1). The detailed feature definition and computing methods have been previously reported [32]. These 14 computed image features were saved in a reference feature data file that directly links to all extracted and selected ROIs in our reference image database.

A CBIR Algorithm

We initially developed and tested a content-based image retrieval algorithm using a multi-feature-based k-nearest neighbor (KNN) classifier to search for similar breast masses depicted on the reference database. The similarity is measured by the Euclidean distance (d) between a queried mass region (y_T) and each of the reference

Table 1 List of image features computed from each reference ROI

Global features	1. Average pixel value in the breast area
	2. Average local pixel value fluctuation in the breast area
	3. Standard deviation of the local pixel value fluctuation in the breast area
Local features	1. Region conspicuity
	2. Normalized mean radial length of a region
	3. Standard deviation of radial length
	4. Skew of radial length
	5. Shape factor ratio
	6. Standard deviation of pixel value inside the mass region
	7. Standard deviation of the gradient of boundary pixels
	8. Skew of the gradient of boundary pixels
	9. Standard deviation of pixel values in the surrounding background
	10. Average local pixel value fluctuation in the surround
	11. Normalized central position shift

regions (x_i) in a multi-dimensional space with F selected image features (f_r) [32].

$$d(y_T, x_i) = \sqrt{\sum_{r=1}^{F} (f_r(y_T) - f_r(x_i))^2} \quad (1)$$

A smaller distance indicates a higher degree of "similarity" between two compared regions. Using a distance weight:

$$w_i = \frac{1}{d(y_T, x_i)^2} \quad (2)$$

KNN classifier computes a classification score indicating the likelihood of the queried region being a true malignant mass:

$$P_{TP} = \frac{\sum_{i=1}^{N} w_i^{TP}}{\sum_{i=1}^{N} w_i^{TP} + \sum_{j=1}^{M} w_j^{FP}} \quad (3)$$

where $N+M=K$, w_i^{TP} and w_j^{FP} are the distance weights for the true-positive (i) and false-positive (j) mass regions, respectively. N is the number of cancer-verified true-positive (TP) mass regions, and M is the number of cancer-free including the benign and CAD-cued false-positive (FP) regions. In this study, K=15 as selected in a genetic algorithm-based optimization process [32].

However, there are two major disadvantages of using this conventional KNN-based instance learning approach in an interactive CAD system. First, the cost of classifying new queried instance can be high in particular as the increase of reference database size to include or cover more different types of subtle mass-like lesions [35]. As a result, the queried region needs to be compared with a large number of reference regions but majority of them are irrelevant to the queried region. Second, the conventional approach of weighting all attributes (features) equally in the similarity search is also not an optimal approach [36]. To reduce the impact of these two disadvantages, we adopted two approaches in our CBIR algorithm. The first approach has been tested in our previous study [32] to eliminate a large number of unnecessary searches (the detailed comparison between the queried and reference regions). In brief, based on our observation and analysis on improving "visual similarity" of the retrieved ROIs, we added two boundary conditions on size and circularity difference between a queried region (S_T and C_T) and a reference region (S_r and C_r) in the KNN algorithm. These two empirically determined conditions were: (1) $\frac{|S_r - S_T|}{S_T} \leq \frac{1}{3}$, and (2) $|C_r - C_T| \leq 0.15$. Circularity was defined as $C = \frac{N_{GR} \cap N_C}{N_{GR}}$, the number of pixels located both inside the growth region (N_{GR}) and the equivalent circle (N_C) divided by the number of pixels located inside the growth region alone. As a result, the KNN algorithm was restricted to selecting "similar" regions each with a reasonably comparable size and an overall shape. This "early" discard process does not only improve computational efficiency and also help to reduce "semantic gap" between computer vision and human vision on evaluating the similarity between the queried and the retrieved reference regions [28].

In this study, we added and tested second approach to our KNN-based CBIR algorithm. To optimally use each attribute in a multi-dimensional feature space, we weighted each attribute differently by stretching (multiplying) a factor z_r to each feature in computing Euclidean distance (d):

$$d(y_T, x_i) = \sqrt{\sum_{r=1}^{F} z_r (f_r(y_T) - f_r(x_i))^2} \quad (4)$$

Similar to our previous study that using a genetic algorithm (GA) to search for the optimal feature set and K number [32], we also applied the same GA-based optimization approach to search for the optimal stretching factors z_r, r=1,…14, used in the KNN algorithm. In our experiment, a binary coding method was applied to create GA chromosomes. The stretching factor of each attribute (feature) corresponds to three genes of a chromosome in which "000" indicates the smallest weight (z_r=1) and "111" represents the greatest weight (z_r=8). Hence, each GA chromosome includes 42 genes in this experiment. Using a leave-one-out testing method involving our reference database with 3,600 ROIs, GA iteratively performs cross-over and mutation operation to find the composition of genes improving the performance of the CBIR algorithm. At each of GA iterations, the area under receiver operating characteristic (ROC) curve was computed based on the corresponding composition of genes selected by GA. When there is no performance improvement in the new generation or the searching generation reaches the predetermined maximum number (i.e., 100 in our studies), the GA optimization was terminated and the selected stretching factor set was applied to the KNN-based CBIR algorithm implemented in our interactive CAD system.

Performance Assessment

We first tested the functionality of our interactive CAD system and then assessed the potential CAD scheme performance in classifying between the malignant mass regions and benign or CAD-cued false-positive regions. Specifically, when applying the optimized CBIR scheme, a retrieved reference image (ROI) is considered to be clinically relevant if it belongs to the same class (cancer-verified or cancer-

free) of the query image (ROI). Since CBIR schemes use the instance-based machine learning methods that depend only on the nearest neighbors and/or locally weighted regression to approximate real-valued or discrete-valued target functions, no pre-training process is needed to construct a general and explicit target function. Thus, to test and evaluate the performance of our CBIR-based CAD scheme using a leave-one-out testing method, each of 3,600 ROIs in our reference database was separately used once as a queried (testing) ROI. In this iterative process of performance evaluation, once a testing ROI is selected, the CBIR scheme searched for $K=15$ ROIs that are considered the most similar to this queried ROI through the remaining reference database with 3,599 ROIs (excluding the queried region itself). As a result, a set of K similar reference ROIs and a corresponding classification score is generated for the queried ROI. Based on the detection scores for both true-positive and false-positive ROIs, we applied a maximum likelihood (binormal) ROC data fitting and analysis program (ROCKIT, http://www-radiology.uchicago.edu/krl) to generate ROC curve, compute the area under ROC curve (AUC) and its 95% confidence interval. The AUC value was used as an index to assess the classification performance of CBIR-based CAD scheme in selecting clinically relevant reference ROIs. The performance levels were then compared between CAD schemes using two KNN algorithms. One used equally weighted and another used optimally stretched attributes in the multi-dimensional feature space. The significant difference in the classification performance (p value) was computed.

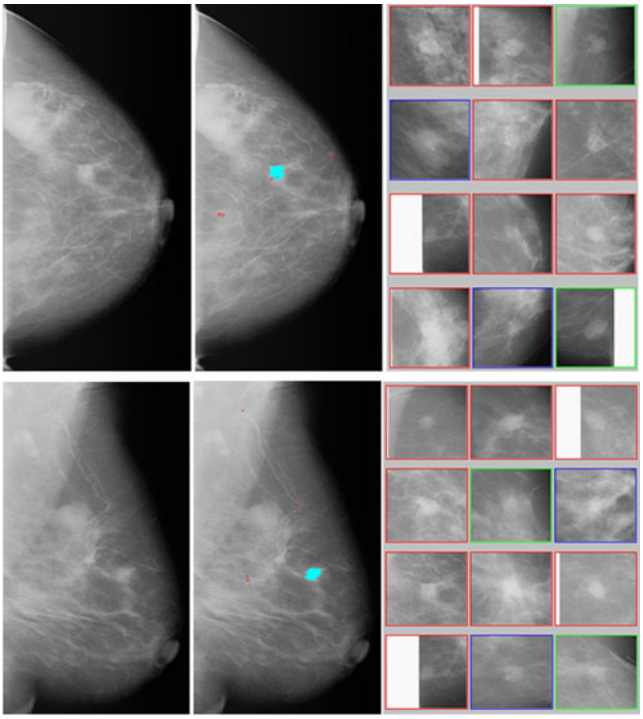

Fig. 2 An example of applying the interactive CAD system to process two images (CC and MLO view of one breast) where the *left column* shows the originally digitized mammograms, the *middle column* shows the images with CAD-cueing results (with one cued mass and several cued micro-calcification clusters), the *right column* shows two sets of 12 CBIR-selected reference ROIs depicting the similar mass regions to the CAD-cued mass regions. By counting horizontally row-by-row from the *top left* ROI (1) to the *bottom right* ROI (12), the reference image set of CC view includes eight ROIs (1, 2, and 5–10) depicting malignant masses, two ROIs (3 and 12) depicting benign masses, and two ROIs (4 and 11) depicting CAD-cued false-positive regions, while the reference set of MLO view includes eight malignant ROIs (1–4 and 7–10), two benign ROIs (5 and 12), and two CAD-cued false-positive ROIs (6 and 11)

Results

An example of how, using our interactive CAD system, to process an image is shown in Figs. 2 and 3. When applying the conventional CAD scheme to independently process the two images (CC and MLO view of one breast) as displayed in left column of Fig. 2, the scheme detects and cues one suspicious mass region on each image and several micro-calcification clusters (as shown in middle column of Fig. 2). If an observer queries a CAD-cued mass region, our interactive CAD scheme automatically segments the region, computes image features, applies the conventional CAD scheme to process this region again to acquire the original detection score (that is not provided in the original cueing), and uses a CBIR algorithm to search for a set of $K=15$ most similar suspicious mass regions stored in our reference database and compute a new classification score of this region being malignant. Among the 15 retrieved reference ROIs, 12 most similar ROIs in the sorting list are displayed (as shown in right column of Fig. 2). The interactive CAD system will also show and report the observer two likelihood scores that represent the original detection score generated by the conventional CAD scheme with an ANN classifier and the classification score computed by the new CAD scheme using the CBIR approach, respectively. In CC view image, the detection score is 0.96 and the classification score is 0.74, while in MLO view two scores are 0.91 and 0.80, respectively. Due to the difference of tissue overlapping in the CC and MLO view, one mass projected on two images often have different image features. As a result, two sets of CBIR-selected reference ROIs as well as two CAD-generated scores could also be different for two queried regions of the same mass.

Our interactive CAD system also allows the observer queries any visually identified suspicious regions (no matter whether they are cued or not cued by the conventional CAD scheme). Assuming the observer queries two regions associated with one suspicious mass depicted on CC and MLO

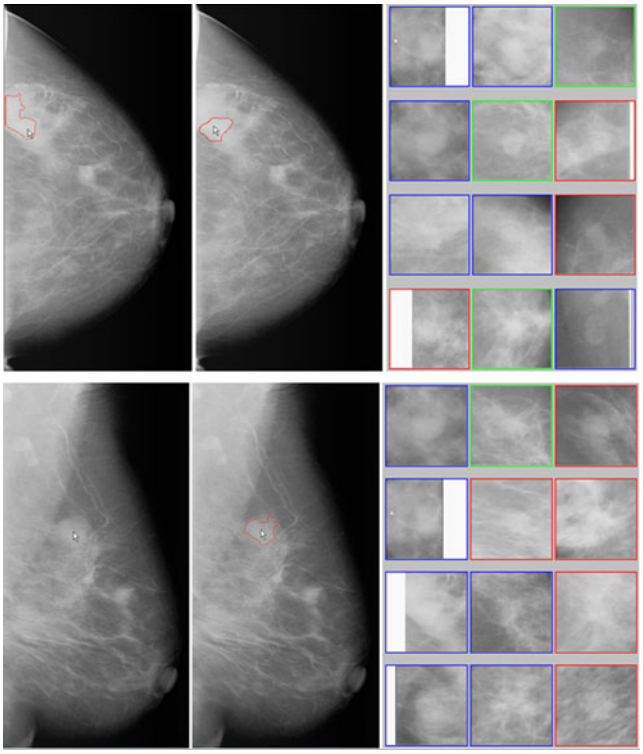

Fig. 3 An example of manually querying any suspicious mass regions not cued by the conventional CAD scheme. Two images in the *left column* show queried regions (pointed by the *arrows*) and the failed automated segmentation result in CC view image. Two images in *middle column* show the manually defined (CC view image) and automated segmented (MLO view image) suspicious region boundary contour. The *right column* show two sets of 12 CBIR algorithm selected most similar reference ROIs. By counting horizontally row-by-row from the top left ROI (1) to the bottom right ROI (12), the reference image set of CC view includes three ROIs (6, 9, and 10) depicting malignant masses, three ROIs (3, 5, and 11) depicting benign masses, and six ROIs (1, 2, 4, 7, 8, and 12) depicting CAD-cued false-positive regions, while the reference image set of MLO view includes five malignant ROIs (3, 5, 6, 9, and 12), one benign ROI (2), and six CAD-cued false-positive ROIs (1, 4, 7, 8, 10, and 11)

view images (as pointed by the arrow in left column of Fig. 3), our CAD scheme fails to correctly segment this suspicious region in CC view image as shown by the segmented boundary contours on the top left image due to the partial dense tissue overlapping. The observer can reject the segmentation result and manually draws the boundary contour as shown on the top middle image of Fig. 3. CAD scheme is forced to process and analyze the manually defined region in this CC view image. The results displayed on the workstation screen include two sets of 12 CBIR-selected ROIs depicting the suspicious mass regions that are considered similar to the queried one (as shown in right column of Fig. 3) and CAD-generated scores. For the suspicious region queried in CC view, the likelihood of this region being a mass is 0.54 (detection score) and the likelihood of the region being malignant is 0.22 (classification score). For the queried region in MLO view, the CAD-generated detection and classification scores are 0.52 and 0.40, respectively. Based on CAD analysis, this is an uncued mass-like area (because the detection scores of two regions depicted on CC and MLO view images are smaller than CAD operating cueing threshold of 0.56 [31]) and the probability of this area associated with malignancy is also low based on the comparison of the CBIR-selected similar reference regions.

Figure 4 shows and compares two ROC-type performance curves of our interactive CAD system evaluated using two KNN algorithms with the same reference database of 3,600 ROIs and a leave-one-out testing method. The original KNN algorithm uses the equally weighted 14 attributes (features) in computing similarity distance index, while in new KNN algorithm 14 attributes are stretched based on the optimization coefficients generated using genetic algorithm. The selected stretching coefficients on 14 attributes are 3, 3, 4, 7, 3, 6, 1, 7, 4, 5, 2, 5, 7, and 4, respectively. Using the original KNN, the area under ROC curve is AUC=0.865 and its 95% confidence interval (CI) is [0.853, 0.876]. When using new KNN, the classification performance level of the interactive CAD system is significantly increased ($p<0.001$) to AUC=0.897 with the corresponding 95% CI of [0.887, 0.907].

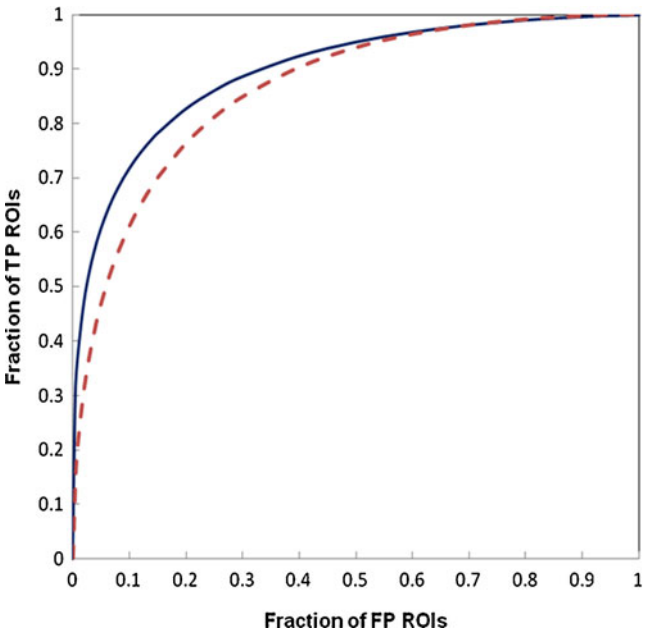

Fig. 4 Comparison of two ROC-type performance curves of our interactive CAD scheme evaluated using a leave-one-out testing method involving 3,600 ROIs in the reference database. One curve (*solid line*) was generated using a new KNN algorithm with an attribute stretching function (AUC=0.897±0.005) and another curve (*dashed line*) was generated using the original KNN algorithm with equally weighted attributes (AUC=0.865±0.006)

Discussions

Since the current conventional CAD schemes have lower performance (including both sensitivity and specificity) than radiologists in detecting mass-like breast abnormalities or cancer and the CAD is used as "a second reader," using a "black-box" type cueing concept adopted by current conventional CAD schemes is very difficult to convince radiologists accepting CAD-cueing results (if they are different from the original interpretation of the radiologists) [17]. Hence, to improve clinical utility of mammographic CAD schemes in mass detection and help radiologists more accurately detect/diagnose the subtle mass-like abnormalities, the interactive CAD concept and the real-time use of a CBIR algorithm with a large and diverse reference database may ultimately play an important role in future CAD development and application. Specifically, the interactive CAD system enables to retrieve similar images or ROIs with verified diagnostic truth and compute a new classification-based score. These results may provide radiologists useful "visual aid" in their decision-making process of detecting or diagnosing subtle cancers. In this article, we demonstrated a unique interactive CAD system. Although many approaches to improve performance of clinical relevance and visual similarity of our interactive CAD schemes have been reported in a number of our previous studies [28, 32, 33, 35, 37–39], this article demonstrated following three innovative characteristics or unique functions that have not been reported in previous studies.

1. Since conventional CAD could miss or be unable to cue a high fraction of subtle masses [31], we added a function in our interactive CAD system, which allows an observer to query any suspicious regions (no matter whether it is cued or not cued by CAD). Thus, the new CAD integrated with a region segmentation algorithm, a global feature-based classifier (ANN), and a local region-based learning method (CBIR) will process and analyze the queried region and provide the observer feedback or evidence of why CAD does or does not detect (or cue) this region. The observer can view two CAD-generated scores and compare a set of retrieved similar reference regions. Such "visual aid" information may help the observer more accurately decide whether this queried suspicious region is associated with a subtle mass or not.
2. Since the complex overlapping of dense fibro-glandular breast tissue on a two-dimensional projection image (mammogram) can make the boundary or partial boundary of some subtle masses quite vague, automatically segmenting these mass regions can be difficult and often inaccurate. This is one of the important reasons that the conventional CAD schemes are unable to detect these suspicious mass regions. To increase the flexibility of our interactive CAD system, we also added a function that allows the observer to manually correct and re-segment the suspicious mass regions. There is no difference for our CAD scheme to process the automated or manually segmented regions.
3. Since the features computed from the reference ROIs in a multi-dimension (e.g., 14) space domain are unlikely to be uniformly distributed, the conventional KNN algorithm using an equally weighted distance index may not be an optimal approach [36]. Thus, we modified our KNN algorithm by adding a new optimal function of stretching the attributes (features) in the multi-dimensional space. The preliminary experimental result using our expanded reference database indicated that this new modification was able to further improve classification performance of the KNN-based CBIR algorithm. The reliability of the evaluated CAD classification performance could also be increased with smaller standard deviation of computed AUC value (as shown in Fig. 4).

In summary, we reported in this article our continuous effort to improve the potential clinical utility of applying the interactive CAD systems by implementing more user-friendly functions to control the interaction between the observer (radiologist) and CAD schemes, as well as to improve CAD classification performance by using a modified KNN algorithm. Although we demonstrated several advantages of our interactive CAD system, this is a preliminary technology development study with a number of limitations. First, we only evaluated the CAD classification performance (Fig. 4) on the entire database. Since our current reference image database was composed by integrating images acquired from a number of independent small databases, many radiographic features (e.g., mass boundary types and mammographic density BIRADS ratings) are not available for all images or masses in our reference database. Hence, the classification performance differences of our CBIR-based CAD scheme on different sub-groups of cases (e.g., four density BIRADS groups) have not been evaluated to date. Second, the ROIs used in our reference database were extracted from the digitized mammograms. Currently, the full-field digital mammography (FFDM) has been widely used in the clinical practice of many medical centers to replace screen-film-based mammography. Due to the different image characteristics of digital and digitized mammograms, the CAD schemes developed using digitized images need to be converted and re-optimized using the FFDM images in the future studies. Third, whether and how using this interactive CAD system could actually help increase radiologists' confidence in accepting or considering the CAD-cued suspicious mass regions and thus improve their performance in detecting more subtle mass-like cancers at

an early stage has not been tested to date. Despite the limitations in this early development stage, we believe that due to the large image repositories in the clinical environment, the complexity of retrieving the clinically relevant images, and the need of the reference images to compare the similar cases with previously verified results in the image diagnostic practice, developing the similar interactive CAD systems integrated with the advanced CBIR algorithms for detecting breast cancer or other types of cancers and diseases will continue attracting wide research interest in the medical imaging and informatics fields. The success of this approach will substantially improve the clinical utility of CAD technologies and better help radiologists interpret or diagnose medical images in the future clinical practice and applications.

Acknowledgment This work is supported in part by grants CA77850 to the University of Pittsburgh from the National Cancer Institute, National Institutes of Health, USA, and the National Distinguished Young Research Scientist Award (60788101) from National Natural Science Foundation of China.

References

1. Jemal A, Siegel R, Xu J, Ward E: Cancer statistics, 2010. CA Cancer 60:277–300, 2010
2. Cady B, Michaelson JS: The life-sparing potential of mammographic screening. Cancer 91:1699–1703, 2001
3. Leach MO, Boggis CR, Dixon AK, et al: Screening with magnetic resonance imaging and mammography of a UK population at high familiar risk of breast cancer: a prospective multicentre cohort study (MARIBS). Lancer 365:1769–1778, 2005
4. Saslow D, Boetes C, Burke W, et al: American cancer society guidelines for breast screening with MRI as an adjunct to mammography. CA Cancer J Clin 57:168–185, 2007
5. Amir E, Freedman OC, Seruga B, Evans DG: Assessing women at high risk of breast cancer: a review of risk assessment models. J Natl Cancer Inst 102:680–691, 2010
6. Sickles EA, Wolverton DE, Dee KE: Performance parameters for screening and diagnostic mammography: specialist and general radiologists. Radiology 224:861–869, 2002
7. Buist DS, Anderson ML, Haneuse SJ, et al: Influence of annual interpretive volume on screening mammography performance in the United States. Radiology 259:72–84, 2011
8. Fenton JJ, Wheeler J, Carney PA, et al: Reality check: perceived versus actual performance of community mammographers. Am J Rotentgenol 187:42–46, 2006
9. Warner E, Plewes DB, Hill KA, et al: Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA 292:1317–1325, 2004
10. Hubbard RA, Kerlikowske K, Flowers CI, et al: Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. Ann Intern Med 155:481–492, 2011
11. Burrell HC, Pinder SE, Wilson AR, et al: The positive predictive value of mammographic signs: a review of 425 non-palpable breast lesions. Clin Radiol 51:277–281, 1996
12. Nishikawa RM: Current status and future directions of computer-aided diagnosis in mammography. Comput Med Imaging Graph 31:224–235, 2007
13. Freer TM, Ulissey MJ: Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology 220:781–786, 2001
14. Brem RF, Baum J, Lechner M, et al: Improvement in sensitivity of screening mammography with computer-aided detection: a multi-institutional trial. Am J Roentgenol 181:687–693, 2003
15. 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 96:185–190, 2004
16. Fenton JJ, Abraham L, Taplin SH, et al: Effectiveness of computer-aided detection in community mammography practice. J Natl Cancer Inst 103:1152–1161, 2011
17. Zheng B, Chough D, Ronald P, et al: Actual versus intended use of CAD systems in the clinical environment. Proc SPIE 6146:9–14, 2006
18. Birdwell RL, Ikeda DM, O'Shaughnessy KF, Sickles EA: Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection. Radiology 219:192–202, 2001
19. Khoo LA, Taylor P, Given-Wilson RM: Computer-aided detection in the United Kingdom National Breast Screening Programme: prospective study. Radiology 237:444–449, 2005
20. Ko JM, Nicholas MJ, Mendel JB, Slanetz PJ: Prospective assessment of computer-aided detection in interpretation of screening mammograms. Am J Roentgenol 187:1483–1491, 2006
21. Alberdi E, Povyakalo A, Strigini L, Ayton P: Effect of incorrect computer-aided detection (CAD) output on human decision-making in mammography. Acad Radiol 11:909–918, 2004
22. Giger ML, Huo Z, Vyborny CJ, et al: Intelligent CAD workstation for breast imaging using similarity to known lesions and multiple visual prompt aides. Proc SPIE 4684:768–773, 2002
23. El-Kwae E, Xu H, Kabuka MR: Content-based retrieval in picture archiving and communication systems. J Digit Imag 13:70–81, 2000
24. Muller H, Rosset A, Garcia A, et al: Benefits of content-based visual data access in radiology. RadioGraphics 25:849–858, 2005
25. El-Napa I, Yang Y, Galatsanos NP, Nishikawa RM, Wernick MN: A similarity learning approach to content-based image retrieval: application to digital mammography. IEEE Trans Med Imag 23:1233–1244, 2004
26. Alto H, Rangayyan RM, Desautels JE: Content-based retrieval and analysis of mammographic masses. J Electron Imag 14:023016, 2005
27. Kinoshita SK, de Azevedo-Marques PM, Pereira RR, Rodrigues J, Rangayyan R: Content-based retrieval of mammograms using visual features related to breast density patterns. J Digit Imag 20:172–190, 2007
28. Zheng B, Mello-Thoms C, Wang X, Abrams GS, et al: Interactive computer aided diagnosis of breast masses: computerized selection of visually similar image sets from a reference library. Acad Radiol 14:917–927, 2007
29. Mazurowski MA, Habas PA, Zurada JM, Tourassi GD: Decision optimization of case-based computer-aided decision systems using genetic algorithm with application to mammography. Phys Med Biol 53:895–908, 2008
30. Muramatsu C, Li Q, Schmidt RA, et al: Determination of subjective similarity for pairs of masses and pairs of clustered microcalcifications on mammograms: comparison of similarity ranking scores and absolute similarity ratings. Med Phys 34:2890–2895, 2007
31. Gur D, Stalder JS, Hardesty LA, Zheng B, Sumkin JH: Computer-aided detection performance in mammographic examination of masses: assessment. Radiology 223:418–423, 2004

32. Zheng B, Lu A, Hardesty LA, et al: A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment. Med Phys 33:111–117, 2006
33. Park SC, Wang XH, Zheng B: Assessment of performance improvement in content-based medical image retrieval schemes using fractal dimension. Acad Radiol 16:1171–1178, 2009
34. Park SC, Pu J, Zheng B: Improving performance of computer-aided detection scheme by combining results from two machine learning classifiers. Acad Radiol 16:266–274, 2009
35. Wang XH, Park SC, Zheng B: Improving performance of content-based image retrieval schemes in searching for similar breast mass regions: an assessment. Phys Med Biol 54:949–961, 2009
36. Mitchell TM: Machine learning. WCR/McGraw-Hill, Boston, 1997
37. Park SC, Sukthankar R, Mummert L, Satyanarayanan M, Zheng B: Optimization of reference library used in content-based medical image retrieval scheme. Med Phys 34:4331–4339, 2007
38. Yang L, Jin R, Mummert L, Sukthankar R, Goode A, Zheng B, Satyanarayanan M: A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. IEEE Trans on PAMI 32:30–44, 2010
39. Wang XH, Park SC, Zheng B: Assessment of performance and reliability of computer-aided detection scheme using content-based image retrieval approach and limited reference database. J Digit Imag 24:352–359, 2011