Developing a new case based computer-aided detection scheme and an adaptive cueing method to improve performance in detecting mammographic lesions

Maxine Tan[1,2,4], **Faranak Aghaei**[2], **Yunzhi Wang**[2], and **Bin Zheng**[2,3]

[1]Electrical and Computer Systems Engineering, School of Engineering, Monash University Malaysia, 47500 Bandar Sunway, Malaysia

[2]School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019

[3]Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15213

Abstract

The purpose of this study is to evaluate a new method to improve performance of computer-aided detection (CAD) schemes of screening mammograms with two approaches. In the first approach, we developed a new case based CAD scheme using a set of optimally selected global mammographic density, texture, spiculation, and structural similarity features computed from all four full-field digital mammography (FFDM) images of the craniocaudal (CC) and mediolateral oblique (MLO) views by using a modified fast and accurate sequential floating forward selection feature selection algorithm. Selected features were then applied to a "scoring fusion" artificial neural network (ANN) classification scheme to produce a final case based risk score. In the second approach, we combined the case based risk score with the conventional lesion based scores of a conventional lesion based CAD scheme using a new adaptive cueing method that is integrated with the case based risk scores. We evaluated our methods using a ten-fold cross-validation scheme on 924 cases (476 cancer and 448 recalled or negative), whereby each case had all four images from the CC and MLO views. The area under the receiver operating characteristic curve was AUC = 0.793 ± 0.015 and the odds ratio monotonically increased from 1 to 37.21 as CAD-generated case based detection scores increased. Using the new adaptive cueing method, the region based and case based sensitivities of the conventional CAD scheme at a false positive rate of 0.71 per image increased by 2.4% and 0.8%, respectively. The study demonstrated that supplementary information can be derived by computing global mammographic density image features to improve CAD-cueing performance on the suspicious mammographic lesions.

Keywords

Computer-aided detection (CAD) of mammograms; Full-field digital mammography; Adaptive cueing; Improvement of screening mammography efficacy; Four-image feature analysis

---

[4]Corresponding author, Maxine.Tan@monash.edu.

1. Introduction

Breast cancer screening is widely considered as an effective approach to detect breast cancer at its early stages and to reduce the mortality rate of cancer patients. Among all available cancer screening imaging modalities, mammography is the only imaging modality that is accepted to conduct population based breast cancer screening to date (Smith *et al.*, 2015). However, due to the overlapping dense fibroglandular tissue in two-dimensional (2D) projection images and the large variability of breast lesions, reading and interpreting mammograms is difficult. Radiologists may miss or overlook early cancers (Birdwell *et al.*, 2001) and also generate high false positive recall rates (Hubbard *et al.*, 2011). Although previous study has well demonstrated that image double reading could significantly improve the performance of screening mammography (Thurfjell *et al.*, 1994), it is not a practical and viable option due to the shortage of radiologists in most countries around the world. Thus, in the last few decades, many Computer-Aided Detection (CAD) schemes of breast mammograms have been proposed and developed as a "second reader" in efforts to assist radiologists in reading/interpreting mammograms (Birdwell *et al.*, 2001). However, despite their relatively-high lesion detection sensitivities, the current commercialized CAD schemes generate substantially high false positive (FP) detection rates and also have high correlations in positive lesion detection with radiologists (Gur *et al.*, 2004a). Thus, studies have shown that the success of using current CAD to improve the detection of malignant masses has been less than overwhelming (Gur *et al.*, 2004b), and many radiologists generally ignore CAD-cued suspected mass regions in clinical practice due to their low confidence in the CAD-cued results (Zheng *et al.*, 2006a). The task of continuously exploring new approaches to develop and use CAD is thus ongoing/still needed (Nishikawa and Gur, 2014).

Recently, we proposed a different approach to improve the efficacy of CAD for mammography by developing a case based CAD scheme (Tan *et al.*, 2014; Tan *et al.*, 2015b). Namely, by computing global mammographic density based image features from all four craniocaudal (CC) and mediolateral oblique (MLO) view mammograms, we trained a classifier to analyze the bilateral global mammographic image features and their differences to generate a likelihood score of the case in question being positive for cancer. Thus, unlike the current lesion based CAD schemes that generate many FP detections, our new case based CAD scheme outputs only one likelihood score for each case in question of being positive or negative for cancer. We hypothesize that since radiologists read four-view mammograms simultaneously and the bilateral asymmetry of mammographic density or tissue patterns is the first important sign for detecting suspicious lesions, developing a case based CAD scheme to determine the cancer risk based on the quantitative assessment of bilateral mammographic tissue asymmetry may play a useful role to warn radiologists to pay more attention to reading the suspicious signs/lesions on these cases.

In the current lesion based CAD schemes, the local image based features (e.g., shape, texture, and spiculation features) were only computed on the segmented lesion region of interest (ROI). Thus, in this study, we hypothesize that the performance of the lesion based CAD schemes may be improved by combining supplementary information derived from our global bilateral mammographic density feature analysis based CAD scheme (Tan *et al.*, 2015b; Tan *et al.*, 2014). This was observed in our preliminary studies on a smaller dataset

(Tan *et al.*, 2016a). In this study, we also extend the image feature set of our case based CAD scheme to include many more diverse features that had performed well in our recent breast cancer risk prediction scheme (Tan *et al.*, 2016b). The purpose of this study is to analyze whether fusion of the case based CAD risk score (Tan *et al.*, 2016b; Häberle *et al.*, 2012) with lesion based CAD scores can help improve the performance of lesion based CAD cueing used in current clinical practice to assist radiologists in reading mammograms.

2. Materials and methods

2.1. A full-field digital mammography (FFDM) image dataset with mass-like abnormalities

Under an institutional review board (IRB)-approved image data collection protocol, we have been collecting a large and diverse FFDM image dataset from our clinical database. An IRB-certified research staff randomly selected screening mammography cases based on the screening outcome (i.e., positive or negative for cancer) recorded in the existing clinical database without viewing the mammograms. All the FFDM examinations were acquired using Hologic Selenia FFDM systems. The fully-anonymized FFDM images and corresponding clinical information were transferred and stored in a research dataset for our studies after a de-identification process.

We assembled a dataset of 924 cases including: (1) 454 verified cancer cases, i.e. masses that were detected during mammographic screening examination and that were confirmed by diagnostic work-up (Group 1); (2) 22 interval cancer cases consisting of masses that were detected in the interval between two screening mammographic examinations (Group 2); (3) 5 patients who had been recalled for diagnostic work-up of high-risk pre-cancer cases (e.g., lobular carcinoma in situ) with recommended surgical excision of lesions (Group 3); (4) 88 patients recalled for diagnostic work-up with masses that were ultimately determined to be benign (Group 4); (5) 355 screening cases that were rated as negative, i.e. not recalled during the screening examinations (Group 5).

All 924 cases had all four FFDM images acquired from CC and MLO views of the left and right breasts. Thus, our FFDM dataset consisted of 3,696 images altogether. After assembling the FFDM dataset, a three-step process consisting of a review, verification and final confirmation of identifying/marking the center-of-mass as well as the encapsulating margin of the mass in question were performed by the radiologists. These steps were performed using information from de-identified source documents provided by research staff (honest brokers). Table 1 summarizes the case distribution as well as the distribution of marked mass regions in each different group/category of cases in the FFDM dataset. Altogether, 1,275 mass ROIs were marked including 963 ROIs associated with cancer and 312 associated with benign and high-risk abnormalities.

2.2. The lesion based computer-aided detection (CAD) scheme

The detailed description of our lesion based CAD scheme has been reported in our previous publications (Zheng *et al.*, 2012a; Zheng *et al.*, 1995). The CAD scheme had performed comparably with two leading commercial CAD schemes on a large independent clinical database (Gur *et al.*, 2004a). A brief description of the scheme is provided in this section,

and the reader is referred to the references (Zheng *et al.*, 1995; Zheng *et al.*, 2012a; Gur *et al.*, 2004a) for a detailed description.

The CAD scheme consists of three main image processing stages. In the first stage, we used a Difference-of-Gaussian (DoG) filter as a blob detector to identify suspicious mass regions. This stage typically detects between 10 and 50 suspicious mass regions depending on the density of the breast and patterns of overlapping fibro-glandular tissue (FGT). In the second CAD stage, we applied a multilayer topographic region growth algorithm to segment the mass candidate regions using the results of the first stage as the initial seed point. After the second stage, the number of suspicious mass candidates typically reduces to less than five per image. In the third stage of the CAD scheme, we computed 14 shape, texture and gray level based features on each segmented mass region. We then applied the features to a multi-feature based artificial neural network (ANN) classifier to generate a detection score indicating the likelihood of the segmented mass region in question being positive for cancer. Finally, we applied an operating threshold to the detection scores of the ANN classifier; only the suspicious regions with detection scores that were greater than the threshold were cued on the images, whereas the other regions were discarded. Figure 1 displays two examples of applying our CAD scheme on two cancer cases in our image dataset.

2.3. A case based CAD scheme incorporating new bilateral asymmetry based features

We recently developed a case based CAD scheme that computed global mammographic density based features from four-view (CC and MLO) mammograms (Tan *et al.*, 2015b; Tan *et al.*, 2014), and fused the features using a novel "scoring fusion" classifier. In the initial schemes, we had computed the simpler gray level and density based features, including mean, standard deviation, skewness, and run length statistics (RLS) features (Galloway, 1975). However, our recent studies (Tan *et al.*, 2015a; Tan *et al.*, 2016b) showed that the more complex texture and structural similarity based features including Weber Local Descriptor (WLD) (Chen *et al.*, 2010) and Structural SIMilarity (SSIM) index (Wang *et al.*, 2004) features were very useful for predicting short-term breast cancer risk occurrence in sequential FFDM images. Thus, in this study, we comprehensively extended our image feature set to compute numerous texture, structural similarity, and mammographic density based features in order to improve the performance of our case based CAD scheme. We computed altogether 158 features that can be divided into four general subgroups: (1) structural similarity features, (2) WLD and Gabor directional similarity features, (3) gray level co-occurrence matrix (GLCM) and RLS features, and (4) other gray level magnitude and texture based features. These features have been explained in detail in previous publications (Tan *et al.*, 2016b; Tan *et al.*, 2015a; Chen *et al.*, 2010; Casti *et al.*, 2015; Wang *et al.*, 2004; Soh and Tsatsoulis, 1999), and we will just provide their brief description in this section.

The first feature subgroup consists of the *structural similarity* features. We first computed the SSIM index metric proposed by Wang et al. (Wang *et al.*, 2004). Given $\mathbf{x} = \{ x_i | i = 1, \ldots, M \}$ and $\mathbf{y} = \{ y_i | i = 1, \ldots, M \}$ are two nonnegative image signals that have been aligned with each other, the SSIM index is defined as:

$$\mathrm{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x{}^2 + \mu_y{}^2 + C_1)(\sigma_x{}^2 + \sigma_y{}^2 + C_2)} \quad (1)$$

whereby $\sigma_{xy} = 1/M \sum_{i=1}^{M}(x_i - \mu_x)(y_i - \mu_y)$, $\mu_x = 1/M \sum_{i=1}^{M} x_i$, $\sigma_x^2 = 1/M \sum_{i=1}^{M}(x_i - \mu_x)^2$. C_1 and C_2 are two small positive constants. Casti et al. (Casti *et al.*, 2015) extended the SSIM index metric to a Correlation-Based SSIM or CB-SSIM index to estimate the structural similarity between two *different-sized* regions. Given $\mathbf{x}_R$ and $\mathbf{y}_L$, a pair of right and left rectangular regions of size $A \times B$ and $P \times Q$ pixels, respectively, the CB-SSIM index is defined as:

$$\mathrm{CB-SSIM}(\mathbf{x}_R, \mathbf{y}_L) = \frac{(2\mu_R\mu_L + K_1)\{2\max[\mathrm{corr}(\mathbf{x}_R, \mathbf{y}_L)]\}}{(\mu_R{}^2 + \mu_L{}^2 + K_1)\{\max[\mathrm{corr}(\mathbf{x}_R, \mathbf{x}_R)] + \max[\mathrm{corr}(\mathbf{y}_L, \mathbf{y}_L)]\}} \quad (2)$$

whereby μ_R and μ_L are the mean values of the right and left breast regions, respectively. *corr*$(\mathbf{x}_R, \mathbf{y}_L)$ is the two-dimensional (2D) cross-correlation between the 2 regions:

$$\mathrm{corr}(\mathbf{x}_R, \mathbf{y}_L) = \sum_{a=0}^{A-1}\sum_{b=0}^{B-1}\{[\mathbf{x}_R(a, b) - \mu_R][\mathbf{y}_L(a - p, b - q) - \mu_L]\} \quad (3)$$

whereby $-P + 1 \quad p \quad A - 1$ and $-Q + 1 \quad q \quad B - 1$. *corr*$(\mathbf{x}_R, \mathbf{x}_R)$ and *corr*$(\mathbf{y}_L, \mathbf{y}_L)$ are two auto-correlation functions of $\mathbf{x}_R$ and $\mathbf{y}_L$, respectively. K_1 is a small positive constant set to 0.01 (Wang *et al.*, 2004), (Casti *et al.*, 2015). As SSIM is highly sensitive to small geometric distortions, such as small rotations, scale, and translations, a Complex Wavelet SSIM index (CW-SSIM) was proposed (Sampat *et al.*, 2009) and defined as:

$$\mathrm{CB-SSIM}(\mathbf{c}_x, \mathbf{c}_y) = \frac{2\left|\sum_{i=1}^{N} c_{x,i} c_{y,i}^*\right| + K}{\sum_{i=1}^{N}|c_{x,i}|^2 + \sum_{i=1}^{N}|c_{y,i}|^2 + K} \quad (4)$$

whereby $\mathbf{c}_x = \{c_{x,i} | i = 1, \dots, N\}$ and $\mathbf{c}_y = \{c_{y,i} | i = 1, \dots, N\}$ are two sets of coefficients extracted at the same spatial location in the same wavelet subbands of the two images being compared in the complex wavelet transform domain, and c^* denotes the complex conjugate of c. K is a small positive constant. Similar to what was performed in the spatial domain, Casti et al. (Casti *et al.*, 2015) defined a new correlation-based complex wavelet SIMilarity (CB-CW-SSIM) index defined as:

$$\mathrm{CB-CW-SSIM}(\mathbf{c}_R, \mathbf{c}_L) = \frac{2\max_{p,q}\left\{\left|\sum_{i,j}[\mathbf{c}_R(i,j)\mathbf{c}_L^*(i-p,j-q)]\right|\right\}}{\sum_{i,j}|\mathbf{c}_R(i,j)|^2 + \sum_{s,t}|\mathbf{c}_L(s,t)|^2} \tag{5}$$

whereby $\mathbf{c}_R$ and $\mathbf{c}_L$ are the complex wavelet coefficients obtained by decomposing regions $\mathbf{x}_R$ and $\mathbf{y}_L$, respectively. In this study, we replicated all the parameters recommended in Refs. (Casti *et al.*, 2015; Wang *et al.*, 2004; Sampat *et al.*, 2009). As we computed all the structural similarity features on the whole and dense breast regions extracted using the method proposed in our previous publication (Chang *et al.*, 2002; Tan *et al.*, 2013), we computed altogether 8 structural similarity features for each case.

The second feature group consists of the *WLD and Gabor directional similarity features*. Inspired by Weber's Law, WLD features are simple yet robust texture features that have outperformed the Gabor (Marcelja, 1980), scale-invariant feature transform (SIFT) (Lowe, 2004), and local binary pattern (LBP) (Ojala *et al.*, 2002) features in various pattern detection and recognition tasks (Chen *et al.*, 2010). The two main components of WLD are the differential excitation and orientation filter responses. Gabor filter features have been widely used for face recognition (Haghighat *et al.*, 2013) and detection of architectural distortion in prior mammograms (Rangayyan *et al.*, 2010). In this study, we computed the SSIM, CB-SSIM, CW-SSIM, and CB-CW-SSIM structural similarity features on the WLD differential excitation and gradient orientation responses and on the Gabor magnitude and phase responses of the image central regions. Thus, altogether 40 directional similarity features were computed to analyze the bilateral differences between the left and right image feature responses. Figure 2 displays an example of a positive (cancer) case in which a mass (blue arrow) was detected and later confirmed by pathology as an invasive ductal carcinoma (IDC). The Gabor magnitude response, WLD differential excitation and gradient orientation components of the abnormal (right) breast (Figs. 2b-d) have structural dissimilarity in the region around the mass (red circle) compared to the normal (left) breast (Figs. 2e-h), which demonstrates that the Gabor and WLD features can identify changes that make the breast susceptible to cancer development.

In the third feature group, we computed the *GLCM and RLS features* to analyze the gray level distributions and patterns within the images. We computed the following RLS features on the whole and dense breast regions: short and long run emphasis, run length non-uniformity, low and high gray level run emphasis, short run low and high gray level emphasis, long run low and high gray level emphasis, gray level non-uniformity, and run percentage. We also computed the following GLCM based features: contrast, energy, homogeneity defined by Soh and Tsatsoulis (Soh and Tsatsoulis, 1999), homogeneity defined in Matlab®, inverse difference normalized, inverse difference moment normalized (Clausi, 2002), maximum probability, correlation defined in Matlab®, and correlation defined in (Haralick *et al.*, 1973). Thus, we computed a total of 80 RLS and GLCM features on the whole and dense breast regions.

In the fourth feature group, we computed 30 *other gray level magnitude and texture* based features described in (Gierach *et al.*, 2014; Häberle *et al.*, 2012). These features included 5 moments based features, 3 percentage density (*PD*) measures similarly defined as *PD* defined by the Cumulus software (Byng *et al.*, 1994; Tan *et al.*, 2014), features based on the directional gradients and gradient direction computed using a Sobel operator, as well as other regional and gray level based features.

As the feature dimensionality (158 features) is very high compared to the dataset size, we need to reduce the number of features due to the "curse of dimensionality" (Powell, 2007). Thus, we applied a fast and accurate modified sequential floating forward selection (SFFS) method by exploiting the information loss of the Mahalanobis distance in high dimensions, as proposed by Ververidis and Kotropoulos (Ververidis and Kotropoulos, 2008). In order to prevent or minimize overfitting on the training set, the feature insertion and exclusion procedure in this modified SFFS has been implemented with an early stopping criterion to limit the number of feature insertion and exclusion steps (Ververidis and Kotropoulos, 2009). In addition, we also used a repeated ten-fold cross-validation to estimate the correct classification rate (CCR) of the feature selection results. By taking these steps, we could select a small sets of the relevant optimal features, while minimize the risk of classifier "overfitting" during the classification stage. Finally, the feature selection was only applied on the training set of each cycle, which was kept completely independent and separated from the evaluation set.

The texture and density of the overlapping breast fibro-glandular tissue (FGT) patterns in the CC and MLO views are often different (see Figure 1) due to the different projection directions. Thus, the bilateral mammographic texture/density asymmetric patterns computed from the two (CC and MLO) views can be different (Zheng *et al.*, 2006b; Wang *et al.*, 2011). In order to fuse the image feature differences and supplementary information computed from both views, we first extracted the maximum feature computed between left and right breasts of each individual view, and then applied the features to a "scoring fusion" artificial neural network (ANN) classification scheme described in detail in our previous publications (Tan *et al.*, 2014; Tan *et al.*, 2015a). The "scoring fusion" classifier generates a final case based classification score, which can be used as a standalone case based CAD scheme, or can be incorporated with the lesion based CAD scores using a new case based adaptive cueing method that will be explained in the next section.

2.4. New adaptive cueing of lesion based CAD scores using a case based CAD probability score

To assess the supplementary information provided by the global feature extraction based method, we combined the "scoring fusion" ANN-generated classification scores of our case based CAD scheme with our lesion based CAD-generated scores using a case based adaptive cueing method. We had previously analyzed this method for cuing subtle masses that were only visible on one view in the "prior" images (Wang *et al.*, 2012). In brief, we desire to adaptively adjust the original lesion based CAD-generated detection scores (S_{org}) of a detected suspicious mass region based on the computed case based score (S_{case}) of the case associated with each detected region. In order to do this, we can project each original

CAD score (S_{org}) into a new cueing reference line using the following projection equation to compute a new CAD cueing score (S_{new}):

$$S_{\text{new}} = S_{\text{org}} + [S_{\text{case}} - S_{\text{org}} \times \tan(\alpha)] \times \sin(\alpha) \times \cos(\alpha) \quad (6)$$

Whereby α is an angle between the projection line and the horizontal axis (i.e., $S_{case} = 0$); $\tan(\alpha)$ is the slope of the projection line to the horizontal axis. Based on equation (6), when $\alpha = 0$, $S_{new} = S_{org}$.

In order to explain how our new case based adaptive cueing approach works, Figure 3 displays two examples of suspicious mass regions (marked R1 and R2). The original lesion based CAD scheme generated detection scores, S_{org} for the two regions, R1 and R2 as 0.51 and 0.56, respectively. Using a lesion-based CAD operating threshold of $T = 0.55$ as used in our original scheme (Gur *et al.*, 2004a), for the projection line slope (i.e., the dashed line) shown in the figure, region R2 is marked (cued) and region R1 is discarded (not cued). However, as region R1 has a high corresponding case based CAD score, when projecting this region unto a new reference line (i.e., the dashed line in Figure 3), its new cueing score computed using equation (6) is $S_{new} = 0.61$. Region R2 was originally cued as $S_{org} = 0.56$. After applying the adaptive case based cueing method, $S_{new} = 0.47$. Thus, using the case based adaptive cueing approach, the originally uncued lower score (region R1) will be cued, whereas the originally cued higher score (region R2) will be discarded.

2.5. Data analysis and performance assessment and comparison

We trained and evaluated our scheme using a ten-fold cross-validation method. In this method, the sum of positive (cancer) cases and negative (cancer-free) cases in our dataset were randomly divided into 10 exclusive partitions. Nine partitions were used to train the "scoring fusion" ANN classifier in each validation cycle using the bilateral image features computed from the CC and MLO view images, respectively. The trained classifier was then applied to the cases in the remaining evaluation partition. For case based adaptive cueing, the result of the trained classifier was also used to adaptively cue or adjust the lesion based CAD scores in the evaluation partition. We iteratively repeated this procedure 10 times using the 10 different combinations of partitions. Thus, each of the positive and negative cases in our dataset was evaluated once for the case based CAD scheme, and with and without adaptive cueing for the lesion based CAD scheme.

In order to assess the performance of our new case based CAD scheme to identify FFDM examinations with high risk of being positive (i.e., cancer), we used several performance assessment indices. We first computed the AUC including the mean and standard deviation of AUC values over the ten folds of the cross-validation experiments for the two individual (CC and MLO) views, as well as the "scoring fusion" (combined) result. The curves were computed using a ROC curve fitting program that applies an expanded binormal model and the maximum likelihood estimation method (ROCKIT http://www.-radiology.uchicago.edu/krl/, University of Chicago, 1998). We then sorted the classification scores of all evaluation cases (including both cancer and cancer-free cases) in ascending order and selected five

threshold values to segment all cases into five subgroups/bins with an approximately equal number of cases within each subgroup. We computed the adjusted odds ratios (ORs), which are defined as a measure of association between an exposure (in our study, bilateral asymmetry between the mammographic image features of left and right breasts) and an outcome (breast cancer occurrence) (Szumilas, 2010), for all subgroups using a multivariate statistical model, and analyzed an OR increasing trend using a publically-available statistical software package, *R* (https://www.r-project.org/). In addition, we also computed the classification accuracy, positive predictive value (PPV), negative predictive value (NPV), and generated a confusion matrix table using a threshold at the midpoint of the classification scores.

To assess the performance of our lesion based CAD scheme, we analyzed all the recorded CAD scores and computed the free-response receiver operating characteristic (FROC) curves with and without applying the new case based adaptive cueing method. We reported both region (ROI) based and case based FROC performance. For the case based analysis, a mass region was counted as "detected" if it was cued by CAD on either one or both views of the same examination. For the ROI based analysis, each mass region was independently counted. Then, using an operating threshold at the same level as previously implemented for evaluation (Gur *et al.*, 2004a), we compared the actual CAD-cueing performance including sensitivity and FP rate per image. All evaluation results were tabulated and compared.

3. Results

Figure 4 displays and compares three ROC curves of the case based CAD scheme optimized based on the image features computed from the individual CC and MLO views, and as a combined result using the "scoring fusion" method of both views. Table 2 displays the corresponding AUC results of the three ROC curves plotted in Figure 4. Using DeLong's test (DeLong *et al.*, 1988) for two correlated ROC curves, the AUC results of the CC and MLO based classifiers were not significantly different from each other at the 5% significance level ($p > 0.94$). However, the AUC result of the "scoring fusion" classifier was significantly different from that of the two individual classifiers ($p < 0.003$), which demonstrates that combining image features from both views significantly improved performance. Furthermore, we also observed that the new combined result of 0.793 ± 0.015 was a significant improvement over the results of our initial case based CAD scheme of AUC = 0.707 ± 0.031 in Ref. (Tan *et al.*, 2015b). This demonstrates that the new structural similarity, texture, and gray level magnitude based features described in Section 2.3 significantly improved the performance of our initial case based CAD scheme (Tan *et al.*, 2015b). The new CAD scheme even outperformed our preliminary scheme that had been trained with the addition of three epidemiology based risk factors of age, family breast cancer history and subjectively-rated mammographic density (BIRADS), i.e., AUC = 0.779 ± 0.025 in the previous study (Tan *et al.*, 2015b). The high performance of the new scheme, which was entirely trained on image features was likely due to the efficacy of the new computed features including structural similarity that can better detect bilateral asymmetry between both breasts, and also better capture the textural information of the breast compared to the more simplistic gray level statistical based features that were implemented in the initial scheme (Tan *et al.*, 2015b) – see Figure 2.

Table 3 summarizes the ORs and corresponding 95% confidence intervals (CIs) computed for the five subgroups (bins) of FFDM examinations. An increasing trend was observed in the case based CAD generated classification scores from subgroups 1 to 5. Namely, the computed ORs monotonically increased from 1.0 to 37.21 in subgroups 1 to 5 using the cases in subgroup 1 as a baseline. We obtained the result that the slope of a regression trend using the data pairs between the classification scores and the adjusted ORs was significantly different from zero ($p < 0.01$). This demonstrates a positive association of classification scores generated by the case based CAD scheme and an increasing risk probability trend of the FFDM examinations of interest being positive. Similar to the AUC performance, the computed ORs for our new case based CAD scheme are much higher than that of our previous scheme (Tan *et al.*, 2015b), using only image features (1 to 7.31) and including epidemiology features (1 to 31.55).

Table 4 displays a confusion matrix when applying a threshold at the midpoint (i.e., 0.5) of the CAD-generated classification scores of all evaluation cases. The FFDM cases with CAD-generated classification scores greater than 0.5 were assigned to the positive ("high risk") case group. Otherwise, the cases were assigned to the negative ("low risk") case group. Using this criterion, the case based CAD scheme correctly classified 72.1% (666 of 924) of cases in the evaluation subsets. The classification accuracy was slightly lower in the negative case group, which was 71.4% (320 of 448), than in the positive case group, which was 72.7% (346 of 476). The computed PPV was 73.0% (346 of 474) and the NPV was 71.1% (320 of 450).

Table 5 displays the result of increasing the projection line angle/slope to modify the cueing weights on cases with a higher score generated by our case based CAD scheme. From Table 5, we observe that by increasing the projection line angle at the same FP rate of 0.89 per image, the ROI based sensitivity gradually increases from 68.8% up to a maximum value of 71.2% at $a = \pi / 19$, followed by a general decreasing trend. Figure 5 displays the ROI and case based FROC curves obtained with and without applying case based adaptive cueing to the original lesion based scores (S_{org}) at $a = \pi / 19$ (note that we used $a = \pi / 19$ as it produced the highest ROI based sensitivity among all the results in Table 5). Tables 6 and 7 tabulate and compare the corresponding ROI and case based FROC sensitivity results, respectively. The results show that better sensitivity results are obtained by incorporating the global feature or case based CAD scores using the adaptive cueing method at lower FP rates (i.e., between 0.5 to 1.1 FP/image). From the tables, we observe that improvements of sensitivities of up to 2.4% and 0.8% were obtained for the ROI and case based results, respectively, using the case based adaptive cueing method. Using DeLong's test (DeLong *et al.*, 1988), the original sensitivity and the adaptive cueing sensitivity results were statistically significantly different from each other at the 5% significance level ($p < 0.05$).

As the previous results were obtained by deriving the optimal value of angle, a retrospectively on the evaluation set, we also repeated the experiments and report the results of optimizing a for each individual training partition independently. By optimizing a on the training partitions, the results of varying a on each individual training partition were very close for $a = \pi / 10$, $\pi / 16$, $\pi / 19$, $\pi / 22$, and $\pi / 28$, with just very slight differences in the results obtained at the same false positive per image rate (or non-lesion localization

fraction). However, we found that $a = \pi / 10$ produced a slightly better result than all the other values of a for all ten training partitions. The poorest result was obtained for $a = \pi / 4$, whose sensitivities were considerably lower than the corresponding sensitivities of the other values of a, which is consistent with the results obtained in Table 5. Figure 6 displays the ROI and case based FROC curves obtained with and without applying case based adaptive cueing to the original lesion based scores (S_{org}) using a optimized on each training partition individually.

4. Discussion

As current CAD of mammograms are generally regarded to be disappointing in terms of adding value to improve screening mammography efficacy, researchers have suggested that the task of exploring new approaches to use and develop CAD schemes is still required (Nishikawa and Gur, 2014). Thus, several researchers have proposed different/alternative CAD approaches in efforts to improve the efficacy of CAD (Samulski *et al.*, 2010; Moin *et al.*, 2011). This study examined two different and new approaches to improve performance and/or clinical utility of CAD for screening mammography.

In the first approach, we examined a case based CAD scheme that extracted many image features from all four images of each patient, applied a fast and accurate modified SFFS feature selection algorithm (Ververidis and Kotropoulos, 2008) individually to the two views, and combined the classification scores of both views using a "scoring fusion" (Tan *et al.*, 2014) ANN classifier. By expanding the image feature set with the inclusion of many more texture, spiculation, mammographic density, and structural similarity or bilateral asymmetry based features (i.e., 158 features, altogether), the performance of our preliminary case based CAD scheme (Tan *et al.*, 2015b) significantly improved. The performance improvement was consistently observed using different performance measurement metrics including AUC and ORs.

The advantages of our case based CAD scheme compared to the conventional lesion based schemes are two-fold. Firstly, the case based cueing can cue/provide "warning signs" to attract radiologists' attention to take a closer look at the cases with the "high-risk" classification scores. Due to the low cancer prevalence rate (i.e., < 1%) in the population based screening environment and studies also show that a high percentage of missed/overlooked breast cancers were visually detectable in the retrospective reviews (Birdwell *et al.*, 2001), this could be one of the main reasons that cause radiologists to overlook the subtle cancer cases from the overwhelmingly negative ones. Thus, the case based CAD scheme will be useful to cue the "warning signs" on the cases with high risk of being positive for cancer. The "warning signs" will attract radiologists' attention to pay more careful attention to the subtle lesions or abnormalities in these cases.

The second advantage of the case based CAD scheme is that unlike the lesion based scheme, which segments and pinpoints the locations of many candidate detections within an image, the case based scheme generates only one cancer likelihood score per case. Previous studies have shown that the lesion based schemes not only generate many FP detections that are distracting to radiologists, but also have high correlation in positive lesion detection with

radiologists (Gur *et al.*, 2004a). In contrast, our case based CAD scheme estimates the bilateral asymmetry between left and right breasts based on the computation of global mammographic density, texture, and spiculation based features, which is a completely new approach different from the current lesion based schemes. Furthermore, unlike the lesion based schemes that typically cue more FPs on the images with denser breast tissue (Ho and Lam, 2003), since our case based CAD scheme estimates bilateral asymmetry between image features extracted from the left and right breasts, it is not correlated to the mammographic density assessed from only one image (Zheng *et al.*, 2012b; Tan *et al.*, 2015a).

In the second approach, we analyzed a new case based adaptive cueing method to adjust the cancer likelihood/probability scores of our in-house lesion based CAD scheme (Zheng *et al.*, 2012a). As the two different types of lesion and case based CAD schemes are developed based on very different concepts and are unlikely to be highly correlated, we incorporated the classification scores generated by the case based scheme into the lesion based scheme using an adaptive cueing method that was examined previously to increase mass cueing performance on two views (Wang *et al.*, 2012). Although the case based adaptive cueing method improved the performance of the lesion based scheme slightly, i.e., 2.4% and 0.8% sensitivity increases were observed for ROI and case based performance, respectively, the performance increases demonstrated the potential of fusing two different classification scores that contribute supplementary information using a simple fusion method. Furthermore, the sensitivity increases were observed at lower FP rates (1.1 FP per image), which are the preferred operating thresholds (in order to generate low FP detections or recall rates) in present CAD schemes used in the clinical practice.

A possible avenue of future work would be to investigate other more complex fusion methods, such as an ANN "scoring fusion" method (Tan *et al.*, 2014) or by using fixed weights (Wang *et al.*, 2011). In this study, we had assumed the existence of a linear relationship, i.e. we used a linear projection line to adjust (raise/decrease) the lesions' cueing scores. In reality, a non-linear relationship between the case based and lesion based classification scores, which can only be derived by implementing more complex fusion methods might yield a better classification performance. However, as a preliminary study and a new concept that to our knowledge is being analyzed for the first time, the initial sensitivity improvements indicate that this is an important new approach and/or a paradigm changing step in the direction of improving the performance of CAD schemes for detecting soft tissue abnormalities (namely, masses).

Another unique characteristic of this study is that we analyzed our scheme on a relatively large and diverse FFDM database of 924 cases, whereby each case had all four images from the CC and MLO views i.e., 3696 images altogether. Compared to the previous studies that had also proposed new CAD approaches i.e., Refs. (Samulski *et al.*, 2010) (120 cases) and (Moin *et al.*, 2011) (200 cases), this study was conducted on a much more large/diverse dataset. This also increases the reliability of assessing our scheme's performance.

In summary, we examined two new concepts/approaches of developing CAD schemes for screening mammograms. The first approach demonstrated the potential usefulness/utility of

a case based CAD scheme that generates a case based risk/probability score based on the detection and analysis of bilateral mammographic density and asymmetry (structural similarity) global image features. The second approach demonstrated that potential improvements can be obtained by fusing the case based CAD scores derived from global image features with the conventional lesion based CAD scores derived from local (i.e., ROI) extracted features. Despite the encouraging results, this study also has a number of limitations that needs to be addressed in our future work. First, the case based CAD scheme was trained entirely on image based features. In our previous study (Tan *et al.*, 2015b), we observed that adding the epidemiology based risk factors of age, family breast cancer history and subjectively-rated mammographic density (BIRADS) considerably improved our scheme's performance. Second, as this is a retrospective study with an enriched number of positive cases in the image dataset, it does not represent the cancer prevalence ratio of screening examinations in clinical practice. Thus, the performance of our two schemes require proper validation in future prospective studies. Third, this is just a preliminary technology development study. In order to analyze the effectiveness of our scheme in radiologists' routine clinical workflow, the performance of our scheme has to be validated in future observer performance studies. Thus, in this study, we assessed two different approaches to CAD schemes for mammograms that need to be examined further before they can be utilized by radiologists in clinical practice.

Acknowledgments

References

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. 2001; 219:192–202. [PubMed: 11274556]

Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. Phys Med Biol. 1994; 39:1629–38. [PubMed: 15551535]

Casti P, Mencattini A, Salmeri M, Rangayyan RM. Analysis of structural similarity in mammograms for detection of bilateral asymmetry. IEEE Trans Med Imaging. 2015; 34:662–71. [PubMed: 25361502]

Chang YH, Wang XH, Hardesty LA, Chang TS, Poller WR, Good WF, Gur D. Computerized assessment of tissue composition on digitized mammograms. Acad Radiol. 2002; 9:899–905. [PubMed: 12186438]

Chen J, Shan S, He C, Zhao G, Pietikainen M, Chen X, Gao W. WLD: A Robust Local Image Descriptor. IEEE Trans Pattern Anal Mach Intell. 2010; 32:1705–20. [PubMed: 20634562]

Clausi DA. An analysis of co-occurrence texture statistics as a function of grey level quantization. Can J Remote Sens. 2002; 28:45–62.

DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 11:837–45.

Galloway M. Texture analysis using gray level run lengths. Comput Graph Image Process. 1975; 4:172–9.

Gierach GL, Li H, Loud JT, Greene MH, Chow CK, Lan L, Prindiville SA, Eng-Wong J, Soballe PW, Giambartolomei C, Mai PL, Galbo CE, Nichols K, Calzone KA, Olopade OI, Gail MH, Giger ML.

Relationships between computer-extracted mammographic texture pattern features and BRCA1/2 mutation status: a cross-sectional study. Breast Cancer Res. 2014; 16:424. [PubMed: 25159706]

Gur D, Stalder JS, Hardesty LA, Zheng B, Sumkin JH, Chough DM, Shindel BE, Rockette HE. Computer-aided detection performance in mammographic examination of masses: assessment. Radiology. 2004a; 233:418–23. [PubMed: 15358846]

Gur D, Sumkin JH, Rockette HE, Ganott M, Hakim C, Hardesty L, Poller WR, Shah R, Wallace L. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. J Natl Cancer Inst. 2004b; 96:185–90. [PubMed: 14759985]

Häberle L, Wagner F, Fasching P, Jud S, Heusinger K, Loehberg C, Hein A, Bayer C, Hack C, Lux M, Binder K, Elter M, Münzenmayer C, Schulz-Wendtland R, Meier-Meitinger M, Adamietz B, Uder M, Beckmann M, Wittenberg T. Characterizing mammographic images by using generic texture features. Breast Cancer Res. 2012; 14:1–12.

Haghighat, M.; Zonouz, S.; Abdel-Mottaleb, M. Computer Analysis of Images and Patterns. Wilson, R., et al., editors. Heidelberg: Springer Berlin; 2013. p. 440-8.

Haralick RM, Shanmugam K, Dinstein I. Texture features for image classification. IEEE Trans Syst Man, Cybern. 1973; 3:610–21.

Ho WT, Lam PW. Clinical performance of computer-assisted detection (CAD system in detecting carcinoma in breasts of different densities. Clin Radiol. 2003; 58:133–6. [PubMed: 12623042]

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. 2011; 155:481–92. [PubMed: 22007042]

Lowe DG. Distinctive Image Features from Scale-Invariant Keypoints. Int J Comput Vision. 2004; 60:91–110.

Marcelja S. Mathematical description of the responses of simple cortical cells. J Opt Soc Am. 1980; 70:1297–300. [PubMed: 7463179]

Moin P, Deshpande R, Sayre J, Messer E, Gupte S, Romsdahl H, Hasegawa A, Liu BJ. An observer study for a computer-aided reading protocol (CARP) in the screening environment for digital mammography. Acad Radiol. 2011; 18:1420–9. [PubMed: 21971259]

Nishikawa RM, Gur D. CADe for Early Detection of Breast Cancer—Current Status and Why We Need to Continue to Explore New Approaches. Academic Radiology. 2014; 21:1320–1. [PubMed: 25086951]

Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell. 2002; 24:971–87.

Powell, WB. Approximate Dynamic Programming: Solving the Curses of Dimensionality. 1st. Hoboken, NJ: Wiley-Interscience; 2007.

Rangayyan RM, Banik S, Desautels JEL. Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer. J Digital Imaging. 2010; 23:611–31.

Sampat MP, Wang Z, Gupta S, Bovik AC, Markey MK. Complex wavelet structural similarity: a new image similarity index. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society. 2009; 18:2385–401. [PubMed: 19556195]

Samulski M, Hupse R, Boetes C, Mus RD, den Heeten GJ, Karssemeijer N. Using computer-aided detection in mammography as a decision support. Eur Radiol. 2010; 20:2323–30. [PubMed: 20532890]

Smith RA, Manassaram-Baptiste D, Brooks D, Doroshenk M, Fedewa S, Saslow D, Brawley OW, Wender R. Cancer screening in the United States, 2015: A review of current American Cancer Society guidelines and current issues in cancer screening. CA: A Cancer Journal for Clinicians. 2015; 65:30–54. [PubMed: 25581023]

Soh LK, Tsatsoulis C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans Geosci Remote Sens. 1999; 37:780–95.

Szumilas M. Explaining Odds Ratios. Journal of the Canadian Academy of Child and Adolescent Psychiatry. 2010; 19:227–9. [PubMed: 20842279]

Tan M, Aghaei F, Wang Y, Qian W, Zheng B. Improving the performance of lesion-based computer-aided detection schemes of breast masses using a case-based adaptive cueing method. Proc SPIE Medical Imaging: Computer-Aided Diagnosis. 2016a; 9785 97851V-V-7.

Tan M, Pu J, Cheng S, Liu H, Zheng B. Assessment of a Four-View Mammographic Image Feature Based Fusion Model to Predict Near-Term Breast Cancer Risk. Annals Of Biomedical Engineering. 2015a; 43:2416–28. [PubMed: 25851469]

Tan M, Pu J, Zheng B. Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme. Phys Med Biol. 2014; 59:4357–73. [PubMed: 25029964]

Tan M, Qian W, Pu J, Liu H, Zheng B. A new approach to develop computer-aided detection schemes of digital mammograms. Phys Med Biol. 2015b; 60:4413–27. [PubMed: 25984710]

Tan M, Zheng B, Leader JK, Gur D. Association between Changes in Mammographic Image Features and Risk for Near-term Breast Cancer Development. IEEE Trans Med Imaging. 2016b; 35:1719–28. [PubMed: 26886970]

Tan M, Zheng B, Ramalingam P, Gur D. Prediction of near-term breast cancer risk based on bilateral mammographic feature asymmetry. Acad Radiol. 2013; 20:1542–50. [PubMed: 24200481]

Thurfjell EL, Lernevall KA, Taube AA. Benefit of independent double reading in a population-based mammography screening program. Radiology. 1994; 191:241–4. [PubMed: 8134580]

Ververidis D, Kotropoulos C. Fast and accurate sequential floating forward feature selection with the Bayes classifier applied to speech emotion recognition. Signal Process. 2008; 88:2956–70.

Ververidis D, Kotropoulos C. Information loss of the mahalanobis distance in high dimensions: application to feature selection. IEEE Trans Pattern Anal Mach Intell. 2009; 31:2275–81. [PubMed: 19834146]

Wang X, Lederman D, Tan J, Wang XH, Zheng B. Computerized prediction of risk for developing breast cancer based on bilateral mammographic breast tissue asymmetry. Med Eng Phys. 2011; 33:934–42. [PubMed: 21482168]

Wang X, Li L, Xu W, Liu W, Lederman D, Zheng B. Improving the performance of computer-aided detection of subtle breast masses using an adaptive cueing method. Physics In Medicine And Biology. 2012; 57:561–75. [PubMed: 22218075]

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing : a publication of the IEEE Signal Processing Society. 2004; 13:600–12. [PubMed: 15376593]

Zheng B, Chang YH, Gur D. Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. Acad Radiol. 1995; 2:959–66. [PubMed: 9419667]

Zheng B, Chough D, Ronald P, Cohen C, Hakim CM, Abrams G, Ganott MA, Wallace L, Shah R, Sumkin JH, Gur D. Actual versus intended use of CAD systems in the clinical environment. Proc SPIE. 2006a; 6146 614602--6.

Zheng B, Leader JK, Abrams GS, Lu AH, Wallace LP, Maitz GS, Gur D. Multiview-based computer-aided detection scheme for breast masses. Med Phys. 2006b; 33:3135–43. [PubMed: 17022205]

Zheng B, Sumkin JH, Zuley ML, Lederman D, Wang X, Gur D. Computer-aided detection of breast masses depicted on full-field digital mammograms: a performance assessment. Br J Radiol. 2012a; 85:e153–e61. [PubMed: 21343322]

Zheng B, Sumkin JH, Zuley ML, Wang X, Klym AH, Gur D. Bilateral mammographic density asymmetry and breast cancer risk: a preliminary assessment. Eur J Radiol. 2012b; 81:3222–8. [PubMed: 22579527]
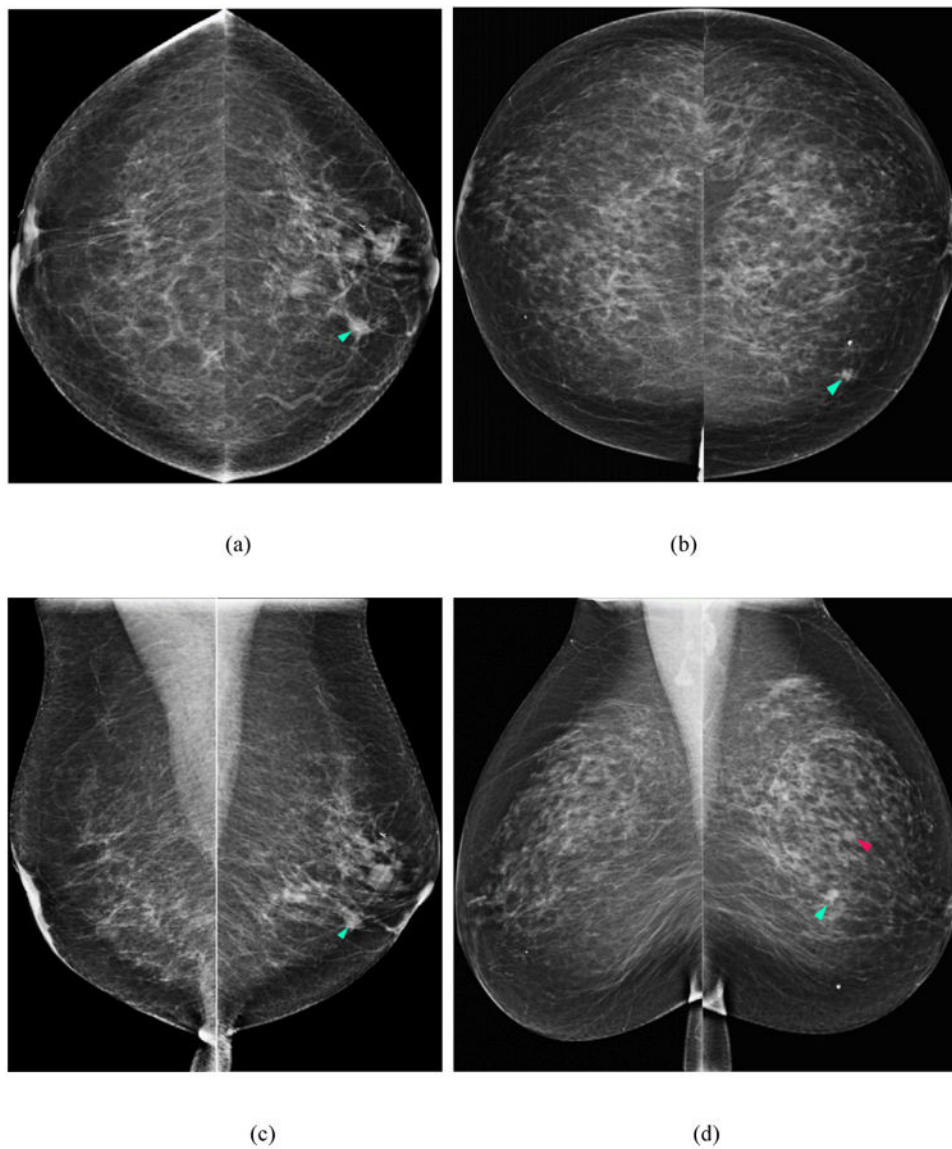
Figure 1.
Two examples of applying our lesion based computer-aided detection (CAD) scheme on two cancer cases in our image dataset. Figures 1 (a) and (c) display the first example with a cancerous mass detected by the scheme (i.e., indicated with light blue arrow) in both the CC and MLO images of the left breast. Figures 1 (b) and (d) display the results obtained on the second example – a cancerous mass was marked with a FP detection (magenta arrow).
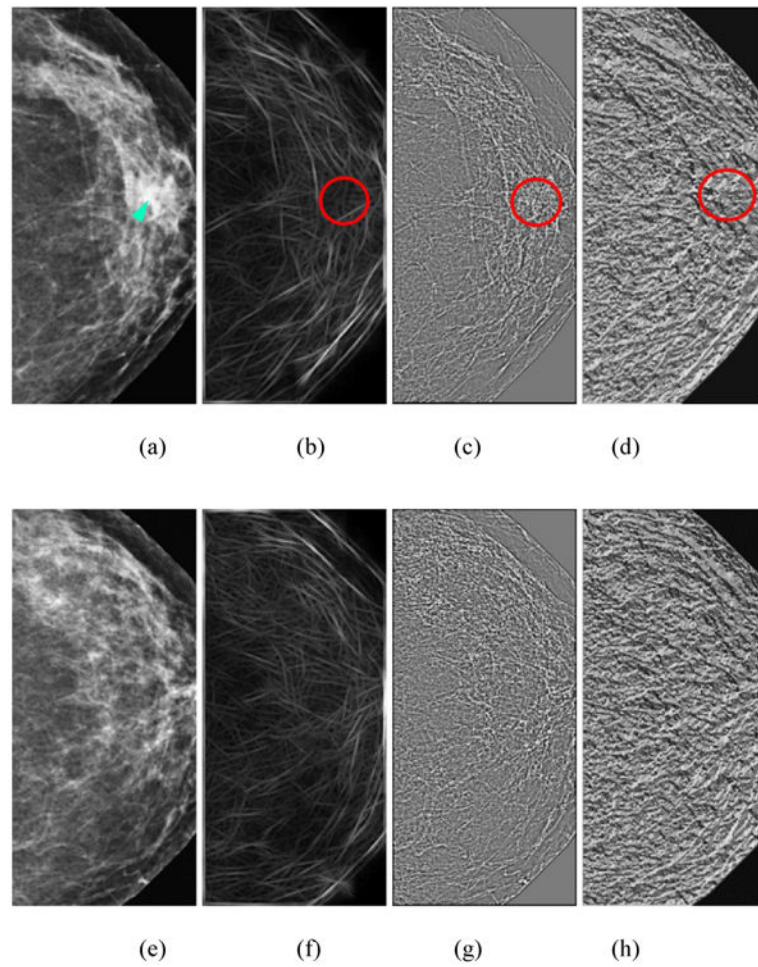
Figure 2.
Example of a positive (cancer) case in which a mass (blue arrow) was detected on the CC view FFDM of the right breast (a); the Gabor magnitude response (b); WLD differential excitation (c); and gradient orientation (d) images computed on the central region of the right breast (red circle indicates approximate mass location). Corresponding images of the left (normal) breast (e)-(h).
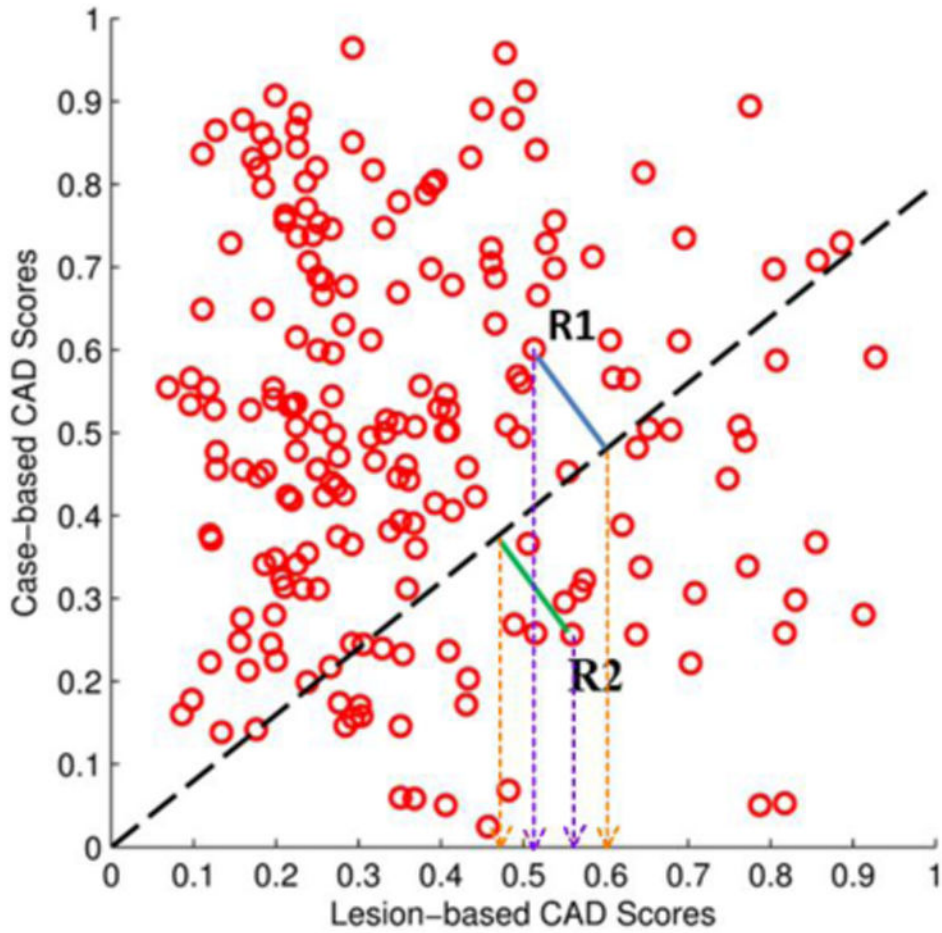
Figure 3.
An illustration of using the case based adaptive cueing method to adaptively change the CAD-generated lesion based detection scores by incorporating the case based CAD scores, and projecting the original lesion based scores unto a new scoring reference line depicted by the dashed black line in the figure. Two purple (inner) vertical dashed lines indicate the original lesion based CAD scheme generated detection scores, S_{org} for two regions, R1 and R2, while two orange (outer) vertical dashed lines indicate the new detection scores after applying the adaptive cueing method on R1 and R2, respectively.
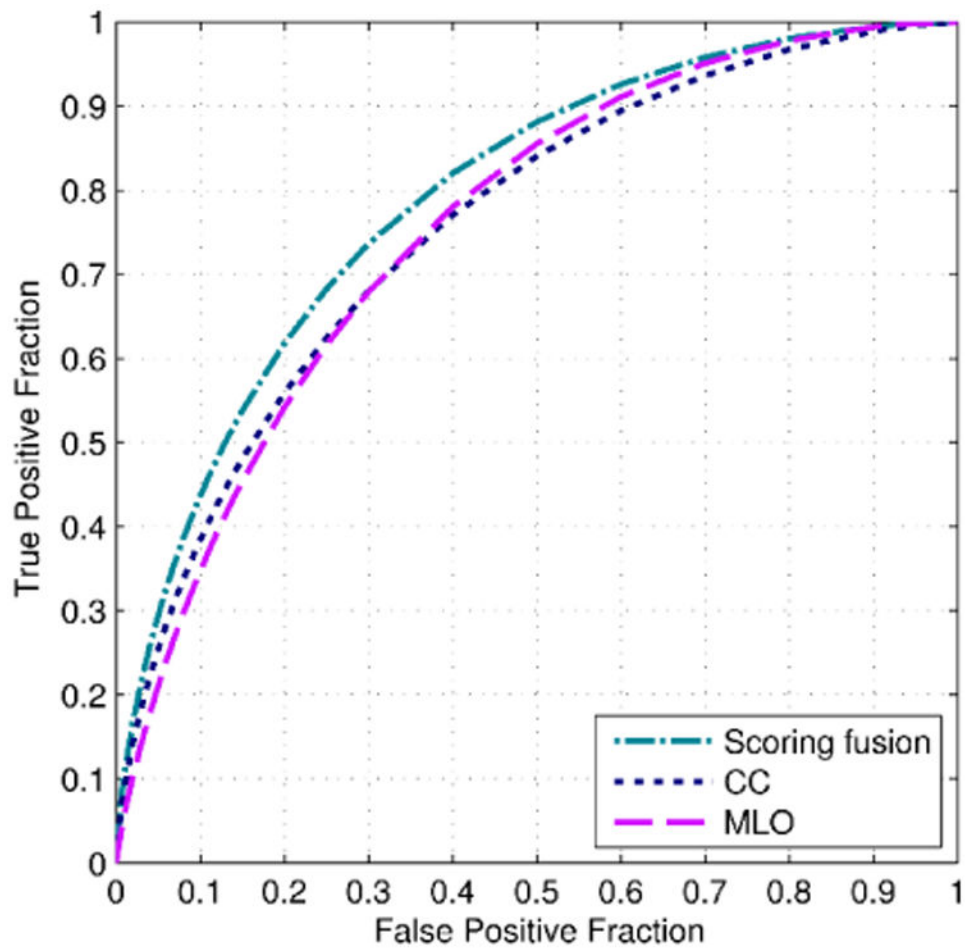
Figure 4.
Comparison of three receiver operating characteristic (ROC) curves of our case based CAD scheme optimized on new features computed from individual CC and MLO views, and combined using a "scoring fusion" scheme of both views.
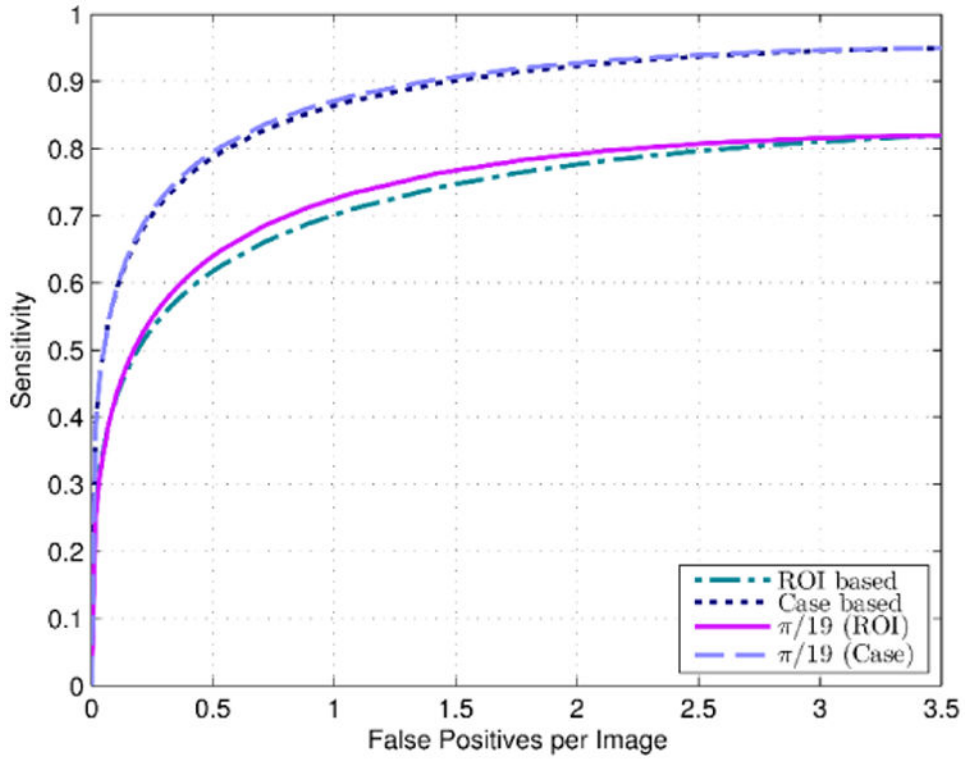
Figure 5.
ROI based and case based FROC curves obtained by applying our original lesion based CAD scheme in Ref. (Zheng *et al.*, 2012a) to our evaluation dataset (924 cases), as well as the ROI and case based FROC curves obtained after applying the adaptive cueing method using the new case based CAD scheme scores at projection line angle, $a = \pi / 19$.
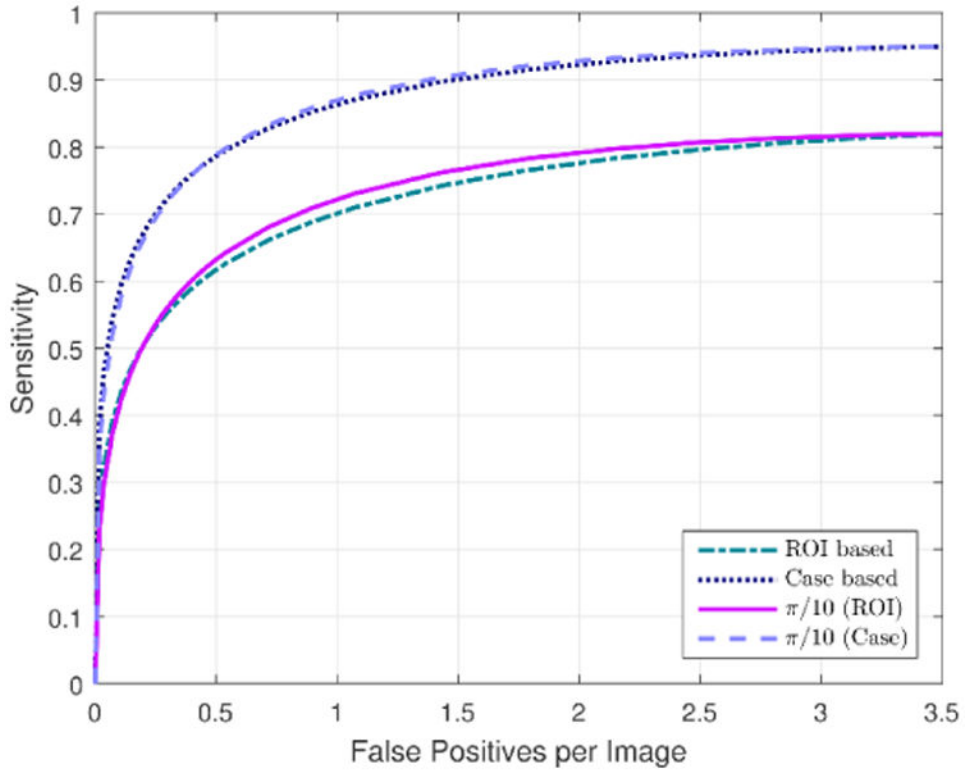
Figure 6.
ROI and case based FROC curves obtained after applying the adaptive cueing method using the new case based CAD scheme scores with projection line angle, a optimized on the training partitions of the cross-validation scheme. The curves are compared with the ROI and case based FROC curves obtained by applying our original lesion based CAD scheme in Ref. (Zheng *et al.*, 2012a).

Table 1

Case distribution and distribution of radiologist-marked mass-like abnormalities of our image dataset of 924 full-field digital mammography (FFDM) cases.

	Number of cases	Total number of radiologist-marked mass-like abnormalities
All selected cases	924	1275
Cancer cases	454	918
Interval cancer cases	22	45
Benign cases	88	302
High-risk cases (with surgical excision)	5	10
Negative cases	355	0

Table 2

Average AUC results and corresponding standard deviation intervals of three ANN classifiers of the case based CAD scheme trained with new features computed on the (1) CC and (2) MLO views, and (3) features combined using a "scoring fusion" method of both views.

No.	Trained ANN classifier	AUC
1	CC view features	0.758±0.016
2	MLO view features	0.758±0.016
3	Optimized "scoring fusion" method of both views	0.793±0.015

Table 3

Summary of the adjusted odds ratios (ORs) and 95% confidence intervals (CIs) at five subgroups of the probability scores generated by the "scoring fusion" classifier of our case based CAD scheme.

Subgroup	Number of cases (Positive – Negative)	Adjusted Odds Ratio (OR)	95% Confidence Interval (CI)
1	25 – 160	1.00	Baseline
2	63 – 122	3.30	[1.97, 5.56]
3	103 – 82	8.04	[4.82, 13.41]
4	128 – 57	14.37	[8.51, 24.28]
5	157 – 27	37.21	[20.69, 66.93]

Table 4

A confusion matrix obtained when applying a threshold at the midpoint of the case based CAD-generated probability scores.

Actual ↓	Negative cases	Positive cases
Negative cases	320	128
Positive (cancer) cases	130	346

Table 5

CAD adaptive cueing region (or ROI) and case based performance comparison by changing the scoring projection line slopes at a FP rate of 0.89 per image.

Projection line angle, α	Adaptive cueing ROI based sensitivity at projection line angle, α (%)	Adaptive cueing case based sensitivity at projection line angle, α (%)
0.0	68.8	85.2
$\pi/4$	64.5	78.1
$\pi/10$	70.9	85.8
$\pi/16$	71.1	86.2
$\pi/19$	71.2	86.0
$\pi/22$	70.9	86.2
$\pi/28$	70.9	85.7

Table 6

Corresponding sensitivity and FP rate results and increases in sensitivities using the adaptively cued scores for the ROI based FROC curves of Figure 5.

FP rate	Original sensitivity (%)	Adaptive cueing sensitivity at $\alpha = \pi/19$ (%)	Increase in sensitivity (%)
0.25	53.2	54.7	1.5
0.50	61.7	64.0	2.3
0.71	66.1	68.5	2.4
0.89	68.8	71.2	2.4
1.1	71.0	73.3	2.3
1.4	74.2	76.3	2.1
1.8	76.5	78.3	1.8

Table 7

Corresponding sensitivity and FP rate results and increases in sensitivities using the adaptively cued scores for the case based FROC curves of Figure 5.

FP rate	Original sensitivity (%)	Adaptive cueing sensitivity at $a = \pi/19$ (%)	Increase in sensitivity (%)
0.25	70.1	70.7	0.6
0.50	78.7	79.5	0.8
0.71	82.8	83.6	0.8
0.89	85.2	86.0	0.8
1.1	87.1	87.8	0.7
1.4	89.7	90.4	0.7
1.8	91.5	92.0	0.5