HHS Public Access

Author manuscript

Phys Med Biol. Author manuscript; available in PMC 2019 May 15.

Published in final edited form as:

Phys Med Biol. ; 63(10): 105005. doi:10.1088/1361-6560/aabefe.

Applying a New Computer-aided Detection Scheme Generated Imaging Marker to Predict Short-term Breast Cancer Risk

Seyedehnafiseh Mirniaharikandehei[1], **Alan B. Hollingsworth**[2], **Bhavika Patel**[3], **Morteza Heidari**[1], **Hong Liu**[1], and **Bin Zheng**[1]

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

[2]Department of Surgery, Mercy Health Center, Oklahoma City, OK 73120, USA

[3]Department of Radiology, Mayo Clinic, Phoenix, AZ 85259, USA

Abstract

This study aims to investigate the feasibility of identifying a new quantitative imaging marker based on false-positives generated by computer-aided detection (CAD) scheme to help predict short-term breast cancer risk. An image dataset including four view mammograms acquired from 1,044 women was retrospectively assembled. All mammograms were originally interpreted as negative by radiologists. In the next subsequent mammography screening, 402 women were diagnosed with breast cancer and 642 remained negative. An existing CAD scheme was applied "as is" to process each image. From CAD-generated results, 4 detection features including the total number of (1) initial detection seeds and (2) the final detected false-positive regions, (3) average and (4) sum of detection scores, were computed from each image. Then, by combining the features computed from two bilateral images of left and right breasts from either craniocaudal or mediolateral oblique view, two logistic regression models were trained and tested using a leave-one-case-out cross-validation method to predict the likelihood of each testing case being positive in the next subsequent screening. The new prediction model yielded the maximum prediction accuracy with an area under a ROC curve of AUC=0.65±0.017 and the maximum adjusted odds ratio of 4.49 with a 95% confidence interval of [2.95, 6.83]. The results also showed an increasing trend in the adjusted odds ratio and risk prediction scores (p<0.01). Thus, this study demonstrated that CAD-generated false-positives might include valuable information, which needs to be further explored for identifying and/or developing more effective imaging markers for predicting short-term breast cancer risk.

Index Terms

Computer-aided detection (CAD); Quantitative imaging marker; Breast cancer screening; Prediction of short-term cancer risk; False-positive detection; Mammography screening

I. INTRODUCTION

Mammography is the most commonly used imaging modality in the population-based breast cancer screening to date. However, the efficacy of screening mammography is controversial (Berlin *et al* 2010) due to the relatively lower detection sensitivity (in particular among

women with dense breasts or younger than 50 years old) (Carney *et al* 2003) and higher false-positive recall rates (Hubbard *et al* 2011) with the potential long-term psychosocial consequences (Brodersen *et al* 2013). In order to improve the efficacy of breast cancer screening, developing a new and more effective personalized breast cancer screening paradigm has been recently attracting extensive research interest (Brawley *et al* 2012). A prerequisite for realizing an optimal personalized screening is to identify more effective breast cancer risk factors and/or develop more accurate risk prediction models to stratify women into two groups with a higher and lower risk of having or developing breast cancer in a short-term (i.e., < 1 to 3 years). As a result, using these risk factors or prediction models may assist clinicians and/or the individual patient to better decide whether she should currently have more frequent screening (e.g., annually) or if she can be screened at longer interval until her short-term risk significantly increases in future reassessments.

Although many epidemiological studies based on breast cancer risk prediction models have been previously developed and applied to identify high-risk women (Amir *et al* 2010) the models have not provided a reliable way to determine who should or should not be screened in the short-term to improve the efficacy of breast cancer screening. To illustrate, these models primarily predict long-term risk of subgroups of women compared to the general population, resulting in low discriminatory power. (Gail *et al* 2010). As a result, identifying new breast cancer risk factors and developing new breast cancer risk prediction models remain crucial and challenging tasks (Hollingsworth *et al* 2014), which still attracts great research interest and effort. In this research field, many researchers believe that mammograms contain important phenotype markers that can be quantified and used to help improve prediction of breast cancer risk (Wei *et al* 2011). In our previous studies (Zheng *et al* 2012b, Tan *et al* 2016), a new quantitative imaging marker based on the computed bilateral asymmetry of mammographic density features between the left and right breasts was identified. The study demonstrated that higher discriminatory power or an increasing trend in predicting short-term risk or the likelihood of women having or developing mammography-detectable cancers on the next subsequent mammographic screening.

During our studies of exploring new quantitative imaging markers, we recently recognized a potential new approach. In order to assist radiologists reading and interpreting mammograms, computer-aided detection (CAD) schemes of mammograms have been installed and used as “the second reader” in a large number of breast imaging facilities or clinics since early 2000 (Astley *et al* 2004). Whether or not using CAD can help increase the accuracy of radiologists in breast cancer detection is also controversial, primarily due to the high false-positive detection rates generated by CAD schemes of mammograms (Nishikawa *et al* 2014). However, whether we can use the quantitative image feature analysis and detection scores on the CAD-generated false-positive regions to help predict short-term breast cancer risk has not been previously investigated. Thus, we proposed a hypothesis. Since previous studies have demonstrated that CAD enabled to detect more early abnormalities that were missed or overlooked by radiologists and later became image-detectable cancer (Birdwell *et al* 2001, Zheng *et al* 2003), the CAD-generated false-positives on the negative images may not be totally harmful or useless. In fact, the CAD-generated false-positives might contain valuable information as quantitative imaging markers to help predict short-term breast cancer risk. The objective of this study is to test our hypothesis

using a relatively large and diverse image dataset of 1,044 negative mammography screening cases. After utilizing a CAD scheme to process these negative images, we applied a machine learning method to develop a new quantitative imaging marker for predicting the risk of having mammography-detectable cancer at the next subsequent screening (12 to 18 months later).

II. MATERIALS AND METHOD

In this study, we used full-field digital mammography (FFDM) images selected from the preexisting image dataset in our laboratory, which has been retrospectively collected and assembled in our previous studies to develop CAD schemes of mammograms (Zheng *et al* 2010, Zheng *et al* 2012a) and quantitative imaging markers for prediction of breast cancer risk (Zheng *et al* 2012b, Tan *et al* 2016). In brief, by excluding the interval cancer cases and the cases without four images of both craniocaudal (CC) and mediolateral oblique (MLO) views of the left and right breasts, the dataset used in this study includes FFDM images acquired from 1,044 women who participated in routine annual mammography screenings. Each case had at least two or more subsequent FFDM screenings of 4 view images of both breasts. The latest screening is named “current” screening, which can be either positive or negative. All “prior” screenings are negative as reported by radiologists in the original image reading and interpretation, which may include a small fraction of “false-negative” cases in which the “early suspicious tumors” may be considered detectable in the retrospective review (Zheng *et al* 2010). However, since these “suspicious tumors” were missed or overlooked by the radiologists in the original screening, the prior images of these cases are classified as negative images in the clinical database. Thus, the negative images used in this study are determined by the clinical record generated from the real mammography screening.

From this dataset, we selected images acquired from the first “prior” mammography screening, which was taken 12 to 18 months prior to the “current” mammography screening. In the “current” screening, 402 cases were positive in which cancer was detected from mammograms and verified by biopsy, while the rest of 642 cases remained negative (cancer-free) in the “current” screenings. Thus, these 1,044 “prior” negative screenings were divided into two groups or classes of the high and low-risk cases. Table 1 summarizes the basic characteristics of the cases involved in our dataset.

Next, as shown in the flowchart of Figure 1, we took the following steps to process each image, compute CAD-generated features, and build a machine learning classifier to predict cancer risk. First, an existing CAD scheme of mammograms (Zheng *et al* 1995) was used. The previous study showed that performance of this CAD scheme was quite comparable to two leading commercialized CAD schemes in detecting malignant breast lesions using an independent testing image dataset to all 3 CAD schemes under comparison (Gur *et al* 2004). In brief, CAD scheme applies three steps to detect suspicious mass regions depicted in an image. First, a Gaussian bandpass filter is applied to detect initially suspicious seeds (i.e., typically 10 to 50 per image depending on the complexity of breast tissue structure). Second, an adaptive multi-layer topographic region growing algorithm is applied to segment each suspicious region. Based on the set of region growing criteria, this step typically results in

discarding more than 50% of suspicious regions initially detected in step one. Third, a multi-feature based artificial neural network (ANN) is applied to process each remaining suspicious region and generate a detection score, which indicates the likelihood of the detected region associating with a malignant lesion. In this study, we applied this CAD scheme of mammograms "***as is***" to process each FFDM image of all study cases in our dataset.

Second, we summarized CAD-generated detection results and scores on all detected suspicious regions on each image. First, four features were computed, which are (1) the number of initial suspicious seeds detected by the first step of CAD scheme, (2) the number of final suspicious mass regions detected by the second step of the CAD scheme, (3) total (summation) score of all detected suspicious mass regions, and (4) average score of all detected suspicious mass regions. Second, we computed bilateral summation between two left and right (CC or MLO view) images. In this way, we generated a total of 8 features, namely 4 for the CC view and 4 for the MLO view images. Third, we computed the mean value (μ) and standard deviation (σ) of each feature computed from all 1044 testing cases. Each feature was then normalized to a range from 0 to 1 (within $\mu \pm 2\sigma$). All outliers were assigned to either 0 or 1 depending on whether their actual value are smaller than zero or greater than one.

Third, since a mammogram is a two-dimensional projection image overlapped with fibro-glandular tissues along the projection direction, mammographic tissue patterns or image features computed from CC and MLO view of one breast may vary significantly. Thus, we separately built two classifiers that combine the four bilateral summation features computed from either CC or MLO view images to stratify the testing cases into the high and low-risk classes of having mammography-detectable cancers in subsequent mammography screening. Despite the fact that many different machine learning classifiers can be used for this purpose, we selected a multinomial logistic regression model based classifier because it is inherently simple, low variance, fast in training and has a lower probability in overfitting.

Fourth, in order to build the classifier, we used a Weka data mining and machine learning software platform (Witten *et al* 2011), which has been successfully applied and tested in a number of our previous studies in developing multi-feature fusion based machine learning classifiers to predict cancer risk and prognosis (Emaminejad *et al* 2016, Danal *et al* 2017, Heidari *et al* 2018). In order to minimize the case selection or partition bias, we applied a leave-one-case-out (LOCO) based cross-validation method to train and test the classifier (Li 2006) in which a classifier was trained using 1,043 cases and tested using one remaining case. Thus, through 1,044 training and testing iterations for each classifier, each case (either two bilateral CC view or MLO view images) had a CAD-generated independent classification score ranging from 0 to 1. The higher score indicates the higher risk or likelihood of the study case having or developing mammography-detectable cancer in a short-term, defined as occurring 12 to 16 months later.

Last, we conducted the following data analysis tasks to assess the performance of using the new risk prediction models for case stratification. We used a receiver operating characteristic (ROC) based data analysis method. A maximum likelihood-based ROC curve fitting

program (ROCKIT, http://xray.bsd.uchicago.edu/krl/roc_soft.htm) was used to generate ROC curve and compute the area under the curve (AUC). We also applied an operating threshold (T = 0.5) on the classification scores to divide the cases into two groups and generated the corresponding confusion matrix. We then computed odds ratio (OR) from the confusion matrix. The adjusted odds ratios and their increasing trend were also computed and analyzed using a statistical software package (R version 2.1.1, http://www.r-project.org). The data analysis results were tabulated and compared.

III. RESULTS

Figure 2 shows an example of applying our CAD scheme to process 4 bilateral CC and MLO view images of a “prior” mammography screening case. Table 2 summarizes the computed 4 features and detection scores for all suspicious regions detected by the CAD scheme on 4 view images of this testing case. In this case, CAD initially detected 81 (ranging from 14 to 23) seeds for the suspicious lesions in the first step. After applying the second step of region growing algorithms, the numbers of the suspicious mass reduced to 32 suspicious regions detected in 4 images. In the third step of the CAD scheme, an artificial neural network based classifier generated a detection score for each of 32 regions, which indicates the likelihood of the region associated with a positive mass. Thus, unlike a conventional CAD cueing method used in clinical practice, which only cues the regions with detection scores greater than a predetermined threshold, all 32 suspicious regions are cued in Figure 2 and all detection scores reported in Table 2 were extracted and used to build the prediction model. All 32 CAD-generated cueing markers were discarded as false-positives and the case was classified as the negative case. In the subsequent mammography screening, a malignant mass-type lesion was detected by the radiologists as shown in Figure 3. Comparing images in Figures 2 and 3, we observed that the lesion detected in “current” images are not “visible” in the “prior” images. However, CAD has two cueing markers with detection scores of 0.56 and 0.65 (as shown in Table 2) on the left MLO view “prior” images (Figure 2), which seem to match with the region that has a malignant mass detected in the “current” left MLO image (Figure 3).

Two logistic regression model based classifiers trained using CC and MLO view images yielded areas under ROC curves, AUC=0.652±0.017 and AUC= 0.652±0.017 to predict cancer risk, respectively. The results show that using CAD-generated detection results or features computed from bilateral MLO view images yielded higher prediction performance than using bilateral CC view images ($p < 0.05$). Figure 4 shows the ROC curve of applying the logistic regression model based classifier to MLO view images.

Tables 3 shows a confusion matrix of using CAD-generated detection scores on the bilateral MLO view images, which was generated by applying an operation threshold of T=0.5 on the risk prediction score. Using this operation score, the risk prediction sensitivity is 27.6% and the specificity is 88.0%. Table 4 summarizes the corresponding computed odds ratio and risk ratio with their 95% confidence intervals. Table 5 reports the adjusted odds ratios, which were computed after applying a set of thresholds to automatically divide 1,044 cases into 5 subgroups with an approximately equal case number. The risk prediction scores gradually increase from subgroup 1 to 5. By using subgroup 1 as a reference (baseline) with 208 cases,

the adjusted odds ratios and 95% confidence intervals for subgroups 2 to 5 with 209 cases were computed. The maximum adjusted odds in subgroup 5's ratio is 4.49 with a 95% confidence interval of [2.95, 6.83]. Using a regression analysis, the slope of the regression line generated from the adjusted odds ratios is significantly different from the zero slope ($p < 0.01$), which indicates an increasing trend of the adjusted odds ratios with an increase of risk prediction scores.

IV. DISCUSSION

The impact of false-positives on the subsequent risk of breast cancer has been previously investigated. For example, one study reported that women undergoing false-positive mammography at the first screening were less likely to participate in subsequent screenings, yet were more likely to develop interval cancers or cancers at subsequent screening (McCann *et al* 2002). Another study reported that women having false-positives involving a fine-needle aspiration cytology or a biopsy had a higher breast cancer detection risk than those involving additional imaging procedures alone in subsequent screening participants over a 17-year period. The odds ratios ranged from 1.81 to 2.69 (Casells *et al* 2013). This study provided new evidence and experimental data regarding the possible association between false-positives and risk of cancer detection in the subsequent screenings. Unlike previous studies of using qualitative assessment, our approach aims to explore, identify and/or develop a new quantitative imaging marker based on CAD-generated false-positive detection results to help predict short-term breast cancer risk.

Since CAD schemes of mammograms are currently available and used in clinical practice, extracting a new quantitative imaging marker from CAD findings in the "current" negative images is a more efficient and cost-effective approach to help predict short-term risk of breast cancer detection in the subsequent screenings, which does not require additional imaging or genomics tests. Since the dataset used in this study was originally assembled for developing CAD schemes, it is not an age-matched dataset (as shown in Table 1) and thus includes a large fraction of difficult negative cases of younger women with dense breasts (Zheng *et al* 2010, Zheng *et al* 2012b). Thus, applying CAD scheme to this dataset tends to produce more false-positive detections as indicated in the previous study (Gur *et al* 2004). However, despite using this relatively difficult dataset, the result of the risk prediction in this study is encouraging, which yielded an odds ratio of 2.80 (as shown in Table 4) or the maximum adjusted odds ratio of 4.49 (as shown in Table 5). The odds ratio is higher than or quite comparable to many existing risk factors reported in previous breast cancer risk prediction studies (Amir *et al* 2010, McCann *et al* 2002, Casells *et al* 2013). Thus, this study identified a potentially new clinical application for CAD schemes of mammograms. Also, the study indicated that further exploration is worth to optimize this new CAD-based imaging marker in future studies.

In this study, we also had several specific observations. First, although CC and MLO view mammograms provide complementary information and reading mammograms of both views can help detect more cancers and reduce false-positive recalls (Giess *et al* 2014), CAD results on two bilateral CC and MLO view images were highly correlated. In this dataset, CAD results on MLO images yielded higher risk prediction performance. AUC value for

using CAD-generated results on MLO view images was 0.652±0.017 with 95% CI of [0.617, 0.686] while using CAD results on CC view images, AUC value was 0.586±0.018 with 95% CI of [0.550, 0.621]. The computed correlation coefficients are 0.75 for all classification scores of 1,044 cases (or 0.76 and 0.73 for 402 high risk and 642 low-risk cases, respectively). Thus, due to the higher correlation coefficient of CAD-generated detection results between using bilateral CC and MLO view images, using the conventional fusion methods (Wang *et al* 2011, Tan *et al* 2015) was unable to further increase AUC value beyond that yielded using MLO view images. To clarify, using an average fusion method yielded AUC=0.631± 0.018.

Furthermore, previous studies have reported that in the retrospective review of the prior images of the positive cases, a large fraction of subtle or occult “early tumors” can be detected by radiologists. Thus, in developing the conventional CAD schemes of mammograms, these missed or overlooked “early tumors” by the radiologists in the real screening environment are typically considered as “false negative” cases and selected as positive training cases in order to increase CAD sensitivity to detect more subtle tumors. In this study, we also conducted a test by removing 53 cases in which the “early masses” were considered “visible” or detectable in our previous retrospective review (Zheng *et al* 2010). After removing these cases, the computed AUC value remained quite constant with only a slight change (or increase) from 0.652±0.017 to 0.659±0.018. The results suggest that unlike the task to develop the conventional CAD schemes of mammograms, for this new specific task of predicting risk of having mammography-detectable cancer in the subsequent (or annual) screening, the negative case group should include all cases that were determined as negative in the real mammography screening environment (include those potential “false negative” cases). More accurately predicting the risk of cancer detection in the next annual screening of these “false negative” cases can help prevent further delay of cancer detection.

Although this is a unique study, which has demonstrated the feasibility of a new concept with promising experimental results, the study also has a number of limitations. First, although we used a relatively large dataset involving 1,044 mammography screening cases, it has very high cancer prevalence ratio (402 vs. 642), which does not represent the cancer prevalence ratio in the real mammography screening environment. Instead, this is just a laboratory-based retrospective study. The potential clinical utility of this approach and/or new imaging marker need to be validated in future prospective studies. Second, this study only focuses on predicting the risk of having or developing mammography-detectable cancer in the subsequent screening. Although detection or distinction between invasive (or aggressive) and non-invasive cancer is an important clinical issue and research topic in reducing over-diagnosis and over-treatment, whether we can develop a new clinical marker or model based on CAD of mammograms to help solve this issue has not been investigated in this study. Third, we only extracted 4 simple features from the CAD-generated detection results and selected a simple multinomial logistic regression model. In order to further improve the performance of risk prediction, more studies are needed to continue exploring, identifying and/or developing more effective imaging markers from the CAD-generated detection results and their impact on future breast cancer risk.

V. CONCLUSION

We investigated and demonstrated the feasibility of extracting a new quantitative mammographic imaging marker from the existing CAD-generated false-positives on the negative mammograms to help predict the risk of cancer detection in the next subsequent mammography screening. The discriminatory power of this new imaging marker or risk prediction model is higher than or quite comparable to many other breast cancer risk factors reported in the literature and/or used in many epidemiology studies based on breast cancer risk prediction models. In addition, since commercialized CAD schemes of mammograms are currently available and used in clinical practice, computing and utilizing this new imaging marker, if successful in the future validation of the prospective studies using the large and diverse databases, is a cost-effective approach to help improve the efficacy of breast cancer screening using mammography.

Acknowledgments

This work is supported in part by Grants R01 CA160205 and R01 CA197150 from the National Cancer Institute, National Institutes of Health, USA. The authors would also like to acknowledge the support from the Peggy and Charles Stephenson Cancer Center, University of Oklahoma, USA.

VII. REFERENCES

Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. JNCI: Journal of the National Cancer Institute. 2010; 102:680–691. [PubMed: 20427433]

Astley SM, Gilbert FJ. Computer-aided detection in mammography. Clinical Radiology. 2004; 59:390–399. [PubMed: 15081844]

Berlin L, Hall FM. More mammography muddle: emotions, politics, science, costs, and polarization. Radiology. 2010; 255:311–316. [PubMed: 20413746]

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]

Brawley OW. Risk-based mammography screening: an effort to maximize the benefits and minimize the harms. Annals of internal medicine. 2012; 156:662–663. [PubMed: 22547477]

Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. The Annals of Family Medicine. 2013; 11:106–115. [PubMed: 23508596]

Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, Cutter G. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Annals of internal medicine. 2003; 138:168–175. [PubMed: 12558355]

Castells X, Roman M, Romero A, Blanch J, Zubizarreta R, Ascunce N. Cumulative False Positive Risk Group. Breast cancer detection risk in screening mammography after a false-positive result. Cancer epidemiology. 2013; 37:85–90. [PubMed: 23142338]

Danala G, Thai T, Gunderson CC, Moxley KM, Moore K, Mannel RS, Qiu Y. Applying quantitative CT image feature analysis to predict the response of ovarian cancer patients to chemotherapy. Academic radiology. 2017; 24:1233–1239. [PubMed: 28554551]

Emaminejad N, Qian W, Guan Y, Tan M, Qiu Y, Liu H, Zheng B. Fusion of quantitative image and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients. IEEE Transactions on Biomedical Engineering. 2016; 63:1034–1043. [PubMed: 26390440]

Gail, MH., Mai, PL. Comparing breast cancer risk assessment models. 2010.

Giess CS, Frost EP, Birdwell RL. Interpreting one-view mammographic findings: minimizing callbacks while maximizing cancer detection. RadioGraphics. 2014; 34:928–940. [PubMed: 25019432]

Gur D, Stalder J, Hardesty LA, Zheng B, Sumkin JH, Chough D, Rockette HE. CAD performance on sequentially ascertained mammographic examinations of masses: an assessment. Radiology. 2004; 233:418–423. [PubMed: 15358846]

Heidari M, Khuzani A, Hollingsworth AB, Danala G, Mirniaharikandehei S, Qiu Y, Liu H, Zheng B. Prediction of breast cancer risk using a machine learning approach embedded with a locality preserving projection algorithm. Physics in Medicine and Biology. 2018; 63:035020. [PubMed: 29239858]

Hollingsworth AB, Stough RG. An Alternative Approach to Selecting Patients for High-risk Screening with Breast MRI. The breast journal. 2014; 20:192–197. [PubMed: 24387050]

Hubbard RA, Kerlikowske K, Flowers CI, Yankaskas BC, Zhu W, Miglioretti DL. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. Annals of internal medicine. 2011; 155:481–492. [PubMed: 22007042]

Li Q. Reduction of bias and variance for evaluation of computer-aided diagnostic schemes. Medical physics. 2006; 33:868–875. [PubMed: 16696462]

McCann J, Stockton D, Godward S. Impact of false-positive mammography on subsequent screening attendance and risk of cancer. Breast Cancer Research. 2002; 4:R11. [PubMed: 12223128]

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; 21:1320–1321. [PubMed: 25086951]

Tan M, Qian W, Pu J, Liu H, Zheng B. A new approach to develop computer-aided detection schemes of digital mammograms. Physics in Medicine and Biology. 2015; 60:4413–4427. [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 transactions on medical imaging. 2016; 35:1719–1728. [PubMed: 26886970]

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. Medical Engineering, and Physics. 2011; 33:934–942. [PubMed: 21482168]

Wei J, Chan HP, Wu YT, Zhou C, Helvie MA, Tsodikov A, Sahiner B. Association of computerized mammographic parenchymal pattern measure with breast cancer risk: a pilot case-control study. Radiology. 2011; 260:42–49. [PubMed: 21406634]

Witten, IH., Frank, E., Hall, MA., Pal, CJ. Data Mining: Practical machine learning tools and techniques. Elsevier; Amsterdam: 2016. http://www.cs.waikato.ac.nz/ml/weka/

Zheng B, Chang YH, Gur D. Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. Academic radiology. 1995; 2:959–966. [PubMed: 9419667]

Zheng B, Good WF, Armfield DR, Cohen C, Hertzberg T, Sumkin JH, Gur D. Performance change of mammographic CAD schemes optimized with most-recent and prior image databases. Academic radiology. 2003; 10:283–288. [PubMed: 12643555]

Zheng B, Wang X, Lederman D, Tan J, Gur D. Computer-aided detection: the effect of training databases on detection of subtle breast masses. Academic radiology. 2010; 17:1401–1408. [PubMed: 20650667]

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. The British journal of radiology. 2012a; 85:e153–e161. [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. European journal of radiology. 2012b; 81:3222–3228. [PubMed: 22579527]

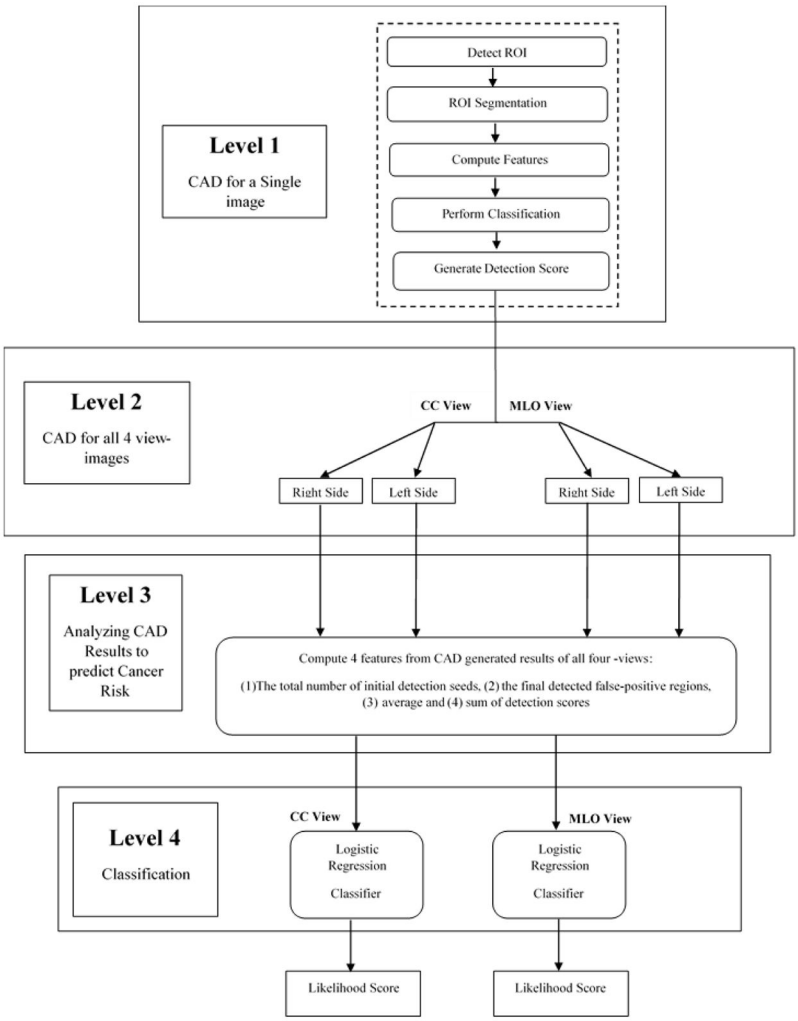

Figure 1.

Flowchart showing the steps of applying the proposed image processing and risk prediction scheme.

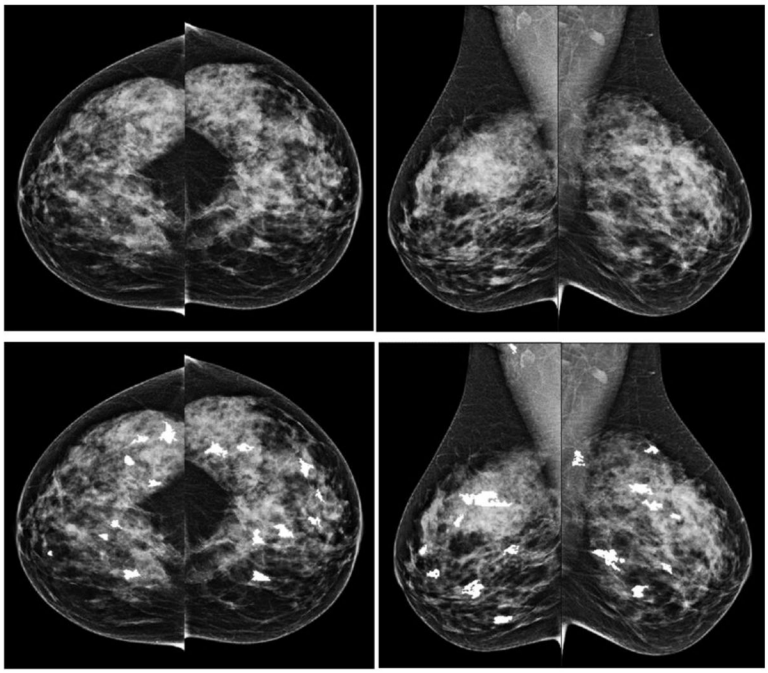

Figure 2.

An example showing CAD processing results in one testing case of 4 bilateral CC and MLO view images. The top row shows 4 original images and the bottom row shows images marked with CAD-detected suspicious mass regions. The CAD-generated detection scores are listed in Table 2.

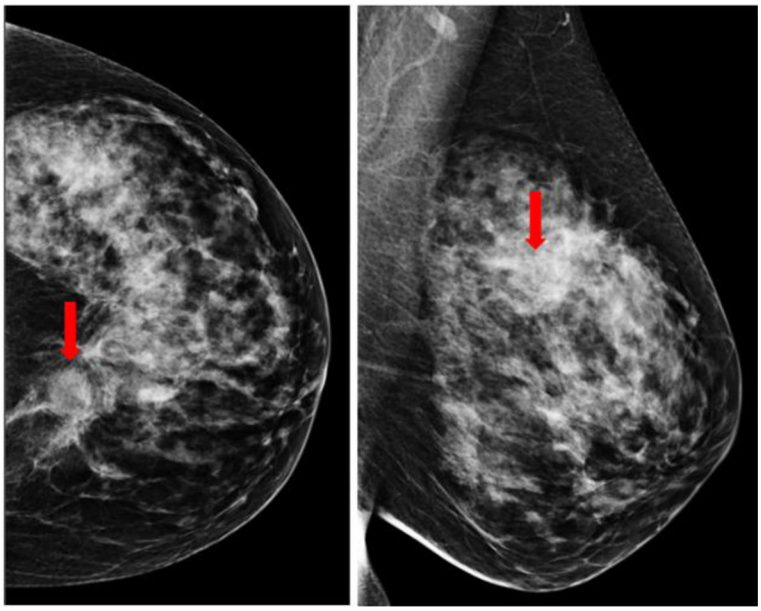

Figure 3.
A malignant lesion was detected in the CC and MLO views of the left mammogram (as pointed by the arrow) in the next subsequent (“current”) screening of the same case as shown in Figure2.

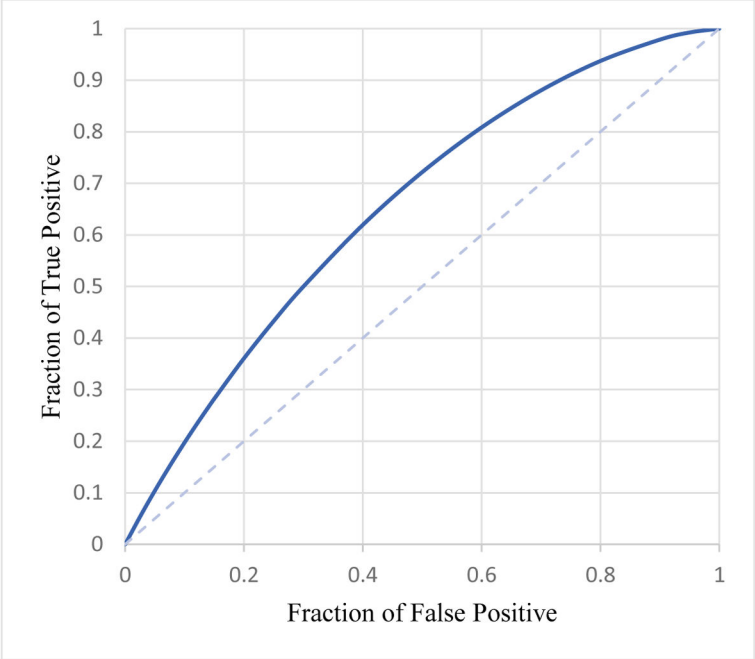

Figure 4.
ROC curve for MLO view with AUC value of 0.652±0.017.

Table 1

Distributions of women's age and subjectively rated breast density (BIRADS), and family history of breast cancer in the two groups of high and low-risk cases.

Risk Factor	Category	High-Risk Cases	Low-Risk Cases
Total Cases		402	642
Age (years old)	< 45	29 (7%)	169 (26%)
	45 – 55	93 (23%)	251 (39%)
	55 – 65	148 (37%)	146 (23%)
	> 65	132 (33%)	76 (12%)
	Mean ± SD	60.88±11.10	52.48±10.52
	Median	60	50
Density BIRADS	Almost all fatty tissue	22 (5%)	40 (6%)
	Scattered fibroglandular densities	153 (38%)	248 (39%)
	Heterogeneously dense	217 (54%)	328 (51%)
	Extremely dense	10 (2%)	26 (4%)
Family History	No family history was known	149 (37%)	285 (44%)
	Cancers in the 1st-degree relatives	48 (12%)	50 (9%)
	Cancers in the 2nd-degree relatives	59 (15%)	83 (13%)
	Cancers in the 3rd-degree relatives	3 (1%)	22 (3%)
	other	143 (36%)	202 (31%)

Table 2

Summary of CAD-detection result and computed 4 features from one example of mammography screening case.

Image View	CC (left)	CC (right)	MLO (left)	MLO (right)
Detected regions	8	8	8	8
Average score	0.47	0.38	0.48	0.38
Total score	3.74	3.01	3.82	3.05
Initial suspicious seeds	23	14	22	22
Detailed detected regions and their scores	1) 0.44	1) 0.41	1) 0.27	1) 0.34
	2) 0.34	2) 0.39	2) 0.31	2) 0.31
	3) 0.50	3) 0.66	3) 0.56	3) 0.54
	4) 0.46	4) 0.27	4) 0.65	4) 0.19
	5) 0.32	5) 0.34	5) 0.38	5) 0.41
	6) 0.51	6) 0.31	6) 0.55	6) 0.26
	7) 0.52	7) 0.28	7) 0.40	7) 0.44
	8) 0.66	8) 0.37	8) 0.70	8) 0.58

Table 3

Confusion matrix of using CAD generated detection results on bilateral MLO view images.

	Actual Positive	**Actual Negative**
Predicted Positive	**111**	77
Predicted Negative	291	**565**

Table 4

Odds Ratio and Risk Ratio of using CAD generated detection results on bilateral MLO view images.

Significance Level	95%
Odds Ratio	2.03 < **2.80** < 3.87
Critical Odds Ratio (COR)	1.11
Risk Ratio	1.77 < 2.30 < 2.99

Table 5

Adjusted Odds Ratio and 95% confidence interval using risk prediction scores computed from bilateral MLO view images.

Group Number	High-risk cases	Low-risk cases	Adjusted Odds Ratio	95% Confidence interval
1	51	157	1	N/A
2	55	154	1.09	0.71 – 1.71
3	80	129	1.91	1.25 – 2.91
4	92	117	2.42	1.59 – 3.68
5	124	85	**4.49**	2.95 – 6.83