Bond University

Research Repository



Presenting a simplified assistant tool for breast cancer diagnosis in mammography to radiologists

Zhang, Ping; Doust, Jenny; Kumar, Kuldeep

Published in:
Medical Biometrics - Second International Conference, ICMB 2010, Proceedings

DOI:
10.1007/978-3-642-13923-9_39

Licence:
Other

Link to output in Bond University research repository.

Recommended citation(APA):
Zhang, P., Doust, J., & Kumar, K. (2010). Presenting a simplified assistant tool for breast cancer diagnosis in mammography to radiologists. In *Medical Biometrics - Second International Conference, ICMB 2010, Proceedings* (Vol. 6165 LNCS, pp. 363-372). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 6165 LNCS). https://doi.org/10.1007/978-3-642-13923-9_39

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

Download date: 29 Mar 2024

Presenting a Simplified Assistant Tool for Breast Cancer Diagnosis in Mammography to Radiologists

Ping Zhang[1], Jenny Doust[1], Kuldeep Kumar[2],

[1]Faculty of Health Sciences and Medicine, Bond University
Gold Coast, Australia
[2]Faculty of Business, Technology and Sustainable Development, Bond University
Gold Coast, Australia
{pzhang, jdoust, kkumar}@bond.edu.au

Abstract. This paper proposes a method to simplify a computational model from logistic regression for clinical use without computer. The model was built using human interpreted featrues including some BI-RADS standardized features for diagnosing the malignant masses. It was compared with the diagnosis using only assessment categorization from BI-RADS. The research aims at assisting radiologists to diagnose the malignancy of breast cancer in a way without using automated computer aided diagnosis system.

Keywords: breast cancer diagnosis, mammography, logistic regression, simplified model

1 Introduction

Mammography is considered the most reliable method of early detection. However, in the earliest stage the visual clues are subtle and varied in appearance, making diagnosis difficult, challenging even for specialists. The benefits of early detection generate a powerful motivation for researchers to develop automated detection and diagnosis systems to assist specialists and radiologists.

A historical review of computer-aided diagnosis (CAD) has been given by Doi [1], indicated that CAD has become a part of the routine clinical work for detection of breast cancer on mammograms at many screening sites and hospitals in the United States. The requirement of high sensitivity and high specificity is difficult to meet for researchers who are developing computer algorithms for detection of abnormalities on radiologic images. It would not be possible for most advanced countries to employ the current computer results for automated computer to make the diagnosis decision [1].

A number of investigators have reported positive findings on the usefulness of CAD in detecting various lesions, including clustered microcalcifications [2] and masses [3] in mammograms, based on the evaluation with ROC (receiver operating characteristic [4])

Recently, the effectiveness of computer assisted detection systems have been assessed for clinical use. In both the studies of Taplin et al [5] and Ko et al [6], CADs were used as the assistants to expert radiologists. Despite the differences between the

studies, they both indicated the possible benefit and some of the issues with the clinical use of current CAD technology. Taplin et al used the ImageChecker M2 1000 system (version 2.2, R2 Technology) [7, 8] as the assistant to the radiologists. They compared the reading from an expert radiologist with or without CAD following the BI_RADS[1] criteria. The conclusion drawn from their study was that the CAD did not affect overall sensitivity, but its effect differed for visible masses that were marked by CAD compared with those were not marked by CAD. The study also showed that the CAD had a greater effect on both specificity and sensitivity among radiologists who interpret more than 50 mammograms per week. Ko et al [6] prospectively assessed the clinical usefulness of CAD in the interpretation of screening mammograms. The CAD used in their study was iCAD MammoReader [9]. This study showed that use of CAD can increase cancer detection rate by at least 4.7% and sensitivity by at least 4% with significant increased recall rates and not significant effect on positive predictive value for biopsy.

The effect of CAD on double reading of paired screen-film and full field digital screening mammograms was evaluated by Skaane et al [10]. The CAD (Image Checker, version 8.0 R2 Technology) showed the potential to increase the cancer detection rate for both FFDM (full field digital mammography) and screen-film mammography in breast cancer screen performed with independent double reading. Georgian-Smith et al [11] compared the CAD with a blinded human second reader for detection of additional breast cancer not seen by a primary radiologist. The research aimed to compare the practice of a human second reader with a CAD reader for the reduction of the number of false-negative cases resulting from review by a primary radiologist. A conclusion from their small sample sized investigation was that a human second reader or the use of a CAD system can increase the cancer detection rate, but no statistical significant difference between the two.

Measuring actual improvement of cancer detection is not a simple task and the results of these studies are dependent on the performance of the CAD itself, the population of cases to which it is applied, and the reviewers who use it [6]. CAD has been used as a complementary tool to mammography, prompting the reader to consider lesions on the mammogram that may represent cancer. However emerging evidence and improved CAD technology are likely to help define its role in breast screening [12].

As we know, the commercialized CADs have not been world widely accepted yet, for example, there have not CAD used in Australia. However, radiologists would like the ability of CAD to change the contrast of a mammogram [13]. In this situation, we would like to present a simple model to assist radiologists for their diagnosis without having to use computer.

---

[1] ACR stands for the American College of Radiology. BI-RADS stands for Breast Imaging Reporting and Data System, which contains a guide to standardized mammographic reporting including a breast-imaging lexicon of terminology, a report organization and assessment structure and coding system.

2 Methodology

Various classification techniques have been applied for breast abnormality classification including neural network, KNN (K-neariest neighbour), logistic regression, and support vector machines. In our previous study [14, 15], we also tested various classification techniques including discriminant analysis, logistic regression and neural networks, and compared the different set of features for breast cancer diagnosis in mammography. Our study showed that logistic regression performs very well for breast abnormality classification in mammography, though the theory is not complicated.

2.1 Logistic regression

Logistic regression is useful for situations in which we want to predict the presence or absence of a characteristic or outcome based on values of a set of predictor variables. It is similar to the linear regression model but is suited to models where the dependent variable is dichotomous [16]. It is widely used in medical research since many studies involve two-category response variables [17].

Logistic regression model for a binary dependent variable can be written as

$$E(y) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k)}$$

where in this research, y=1, if the patient has malignant tumor; y=0, if the patient has benign tumor.

E(y)= P=probability of the breast area is malignant

x1, x2….. xk are quantitative or qualitative independent variables.

2.2 Features for classification

Researchers have extracted different features from the mammograms for cancer diagnosis, such as region-based features [18], shape-based features [19], image structure [20], texture based features [21] and position related features [22]. In our previous research, we extracted 25 features including computer extracted features and human interpreted features [14, 15]. The human extracted features are the features interpreted from the radiologists including the ones interpreted following the BI-RADS lexicon. The computer extracted features are the commonly used which are statistically calculated based on the grey-level value distribution from the mammograms. These features were used to build different classification models for breast cancer diagnosis. Our research result showed that human extracted features contributed to the diagnosis more significantly than the computer extracted features[23].

The BI-RADS was developed to standardize mammographic reporting, to improve communication, to reduce confusion regarding mammographic findings, to aid

research, and to facilitate outcomes monitoring [24, 25]. The assessment categorization from the BI-RADS has been recommended as a predictor of malignancy [26]. Lo et al [27, 28] developed the computer-aided diagnostic techniques that took consideration the description of lesion morphology according to BI-RADS. The models were reported that performed at a level comparable to or better than the overall performance of the expert mammographers who originally interpreted the mammograms.

In this paper, we present a simplified model derived from a logistic regression equation built on human extracted features to assist radiologists for breast cancer diagnosis.

3 Materials

Digital Database for Screening Mammography (DDSM) [29] from University of South Florida was used for this research. In DDSM, the outlines for the suspicious regions are derived from markings made on the film by at least two experienced radiologists. DDSM provides some information related to the marked masses using the ARC BI-RADS definition. Such as assessment, density, mass shape and mass margin are all described using the standard terminology in BI-RADS. Patient age and some others information involved in the digital mammogram, such as subtlety, are also provided in related files. These information are considered that can be important for distinguishing the malignant and benign masses or calcifications

A total of 200 mass suspicious areas were extracted from DDSM for the experiments. It includes 100 malignant case and 100 benign ones. The mammograms used in this study are all scanned on the HOWTEK scanner at the full 43.5 micron per pixel spatial resolution.

Seven human interpreted features related to mass suspicious areas are patient age, breast density, mass shape, mass margin, assessment, subtlety, and calcification association. Except “calcification association” (Calc-asso) was created by us in our research, others are obtained from DDSM. Calc-asso describes how the mass is related to a calcification, for example, some of the masses are also marked as calcifications. In this research, we simply categorize calc-asso as ‘yes’ or ‘no’ for having the association or not having the association. Some of BI-RADS features are described with text in the database. We assign them to the different categories with numbers. For example, “1” represents “round” and “2” is used to represent “oval” for the shape feature. These variables are treated as categorical (nominal) variable for building the models.

4 Implementation

4.1 Cross validation for comparison of the models

The assessment categorization from the BI-RADS as a predictor of malignancy was tested by Orel et al [26]. In their research, 1312 localization (mass areas) were

investigated. The PPV (positive predictive value) was 0% for assessment category 2 (refer to BI-RADS), 2% for category 3, 30% for category 4 and 97% for category 5. In our research, we tested logistic regression model built with all 7 human extracted features. This model is compared with the model built with only 6 features (out of 7 features excluding "assessment"), and with using only assessment categorization from the BI-RADS for prediction of malignancy.

In this study, the logistic regression models were constructed based on the set of features with iteratively reweighted least squares (IWLS) method using R (version 10.1.0), a free software environment for statistical computing and graphics [30]. 5-fold cross validation was conducted to compare the performance between the models using 7 features, 6 features or assessment categorization alone. Table 1 shows the validation results from the models using different sets of features. In 5 splits of the datasets, every split used the different 40 cases out of the whole 200 cases as testing and the rest for training.

Table 1. Areas under the ROC from different Models

Dataset (split)	7-features*		6-features**		"assessment" only	
	training	testing	training	testing	training	testing
1	0.989	0.893	0.983	0.886	0.841	0.863
2	0.972	0.998	0.969	0.988	0.835	0.880
3	0.972	0.988	0.971	0.973	0.844	0.850
4	0.977	0.970	0.973	0.968	0.848	0.830
5	0.972	0.991	0.969	0.981	0.845	0.843
Mean	0.977	0.968	0.973	0.959	0.843	0.853
Standard deviation	0.007	0.043	0.006	0.042	0.005	0.019

*7-features: age, density, shape, margin, subtlety, calcification association and assessment;
**6-features: age, density, shape, margin, subtlety and calcification association

4.2 Models to be Recommended to Radiologists

From table 1, we can see that with 7 features combining the 6 human extracted features and the assessment categorization from the BI-RADS predict better than using only 6 human extracted features or using only the assessment categorization from BI-RADS ($P<0.05$, from t-tests). We used the whole set of 200 cases to build a final logistic regression model with these 7 features. It is shown below:

$$log(p/(1-p)) = -41.9 + C1*age + C2*density + C3*mass_shape + C4*margin + C5*subtlety + C6*Cal_asso + C7*assessment$$

p is the predicted probability of the mass being malignant. *C1* to *C7* are the coefficient values for the different variables in the model. They take different values when the variables are in different categories. See table 2 for the coefficient values in the final logistic regression model. The coefficients of the independent variables were simplified to an integer weight in order to be easily managed by radiologists for assisting the diagnosis. For example, a 55 years old patient has the mammogram showing a mass in "round" shape, obscured margin, subtlety 1, is seen as also calcifications and is assessed as category 3. The total diagnostic weight will be 2+2+0+1+15+0=19.

Table 2. Logistic Regression Model Built with the Whole Dataset

Variable	Category	Coefficient (C1-C7)	Simplified Weight
Age	Over 50	2.41	2
Density	1	1.99	2
	2	1.97	2
	3	1.54	2
Mass shape	Round	1.68	2
	Lobulated	1.44	1
	Irregular	1.80	2
Mass margin	Circumscribed	16.00	16
	Microlobulated	17.97	18
	Ill_defined	17.23	17
	Spiculated	22.34	22
Subtlety	1	1.25	1
	2	1.36	1
	3	0.98	1
Cal-asso	Yes	15.33	15
Assessment	4	18.20	18
	5	20.19	20

Note: the variable values are interpreted following the BI-RADS standard or recoded in DDSM database.
Assessment categories from BI-RADS: 1- negative, 2- benign, 3- probably benign, 4- suspicious abnormality, 5- highly suggestive malignancy.

5 Results and Discussion

Figure 1 shows the ROC curves produced from the simplified model, and is compared with logistic model with all 7 features and with assessment categorization alone for

diagnosis. We can see the area under the simplified model is obviously higher than that from only assessment categorization (0.961 vs 0.843).

Table 3 shows the diagnostic performance from the simplified model built with our 200 cases and the result from using only assessment value for diagnosis. It can be used as the second diagnosis reference on top of the diagnosis with assessment categorization alone. For instance in our above example, though the assessment category is 3 which represent only 13% specificity for malignancy, the model from the logistic regression shows the high probability (specificity 100% with total weight score 19) of it being malignant.

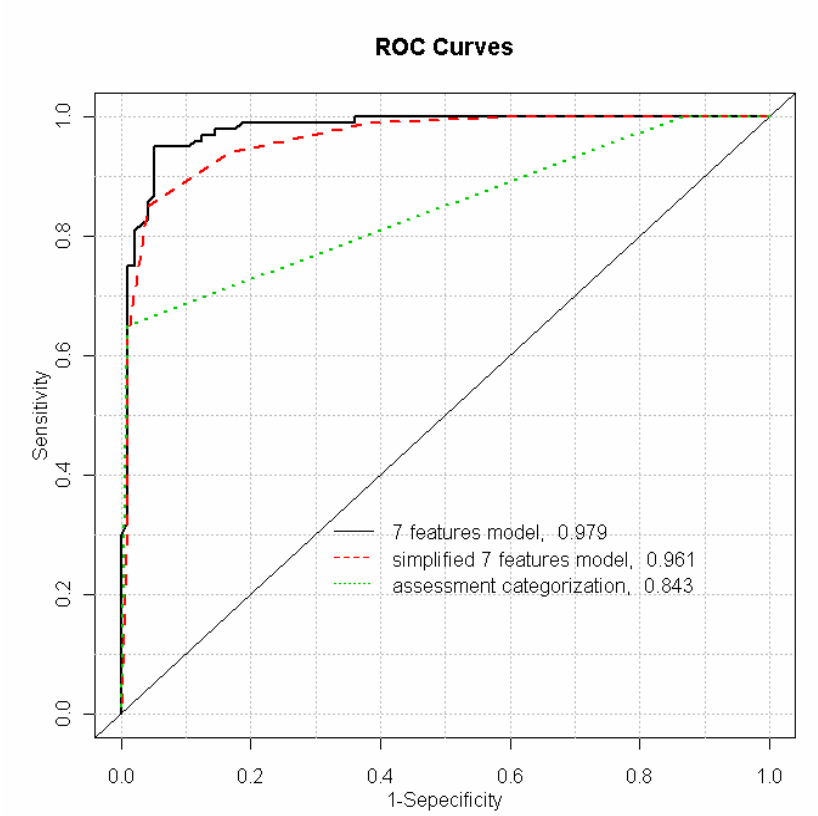


Figure 1: ROC and the areas under the curves from the different models

Table 3. Performace of Logistic Regression Model Built with the Whole Dataset and using Assessment Categorization Alone

	Total Weight score	Sensitivity	Specificity	Youden index
Final Diagnostic Model Result	9	1.00	0.02	0.02
	10	1.00	0.09	0.09
	11	1.00	0.24	0.24
	12	1.00	0.39	0.39
	13	0.99	0.62	0.61
	14	0.94	0.83	0.77
	15	0.85	0.96	0.81
	16	0.63	0.98	0.61
	17	0.43	0.99	0.42
	18	0.30	0.99	0.29
	19	0.00	1.00	0.00
Diagnosis with only assessment value	2	1.00	0.00	0.00
	3	1.00	0.13	0.13
	4	0.65	0.99	0.64
	5	0.00	1.00	0.00

From the results we can see, especially when the assessment category falls to 3 (probably benign), the proposed model will give a reference for the diagnosis. Need to mention, the model proposed in the paper is just a preliminary model built based on only 200 masses which do not cover all the possible features of the mammograms. More variety of the dataset will be used in our future research and aim to build a clinical evaluated model for practical use.

References

1. Doi, K.: Computer-Aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential. Computerized Medical Imaging and Graphics. 31 (4-5), 198-211 (2007).
2. Chan, H.P., Doi, K., Vyborny, C.J., Schmidt, R.A., Metz, C.E., Lam, K.L., Ogura, T., Wu, Y. and MacMahon, H.: Improvement in Radiologists' Detection of Clustered Microcalcifications on Mammograms: The Potential of Computer-aided Diagnosis. Investigative Radiology. 25, 1102-1110. (1990).
3. Moberg, K., Bjurstam, N., Wilczek, B., Rostgard, L., Egge, E. and Muren, C.: Computer Assisted Detection of Interval Breast Cancers. European Journal of Radiology. 39, 104-110 (2001).
4. Zweig, M.H. and Campbell, G.: Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine. Clin. Chem. 39(4), 561-577 (1993).

5. Taplin, S.H., Rutter, C.M. and Lehman, C.D.: Testing the Effect of Computer-Assisted Detection on Interpretive Performance in Screening Mammography. Am. J. Roentgenol. 187(6), 1475-1482 (2006).
6. Ko, J.M., Nicholas, M.J., Mendel, J.B. and Slanetz, P.J.: Prospective Assessment of Computer-Aided Detection in Interpretation of Screening Mammography. Am. J. Roentgenol. 187(6), 1483-1491 (2006).
7. Wang, S.-p.L.A., CA), *Computer-aided Dagnosis System and Method.* Patent Application No: 6434262, 2002, United States.
8. Wang, S.-p.B.L., Los Altos, CA, 94022), *Computer-aided Diagnosis Method and System.* Patent Application No: 6266435, 2001, United States.
9. Rogers, S.K.B., OH, US), *Use of Computer-aided Detection System Outputs in Clinical Practice.* Patent Application No: 6970587, 2005, ICAD, Inc. (Nashua, NH, US), United States.
10. Skaane, P., Kshirsagar, A., Stapleton, S., Young, K. and Castellino, R.A.: Effect of Computer-Aided Detection on Independent Double Reading of Paired Screen-Film and Full-Field Digital Screening Mammograms. Am. J. Roentgenol. 188(2), 377-384 (2007).
11. Georgian-Smith, D., Moore, R.H., Halpern, E., Yeh, E.D., Rafferty, E.A., D'Alessandro, H.A., Staffa, M., Hall, D.A., McCarthy, K.A. and Kopans, D.B.: Blinded Comparison of Computer-Aided Detection with Human Second Reading in Screening Mammography. Am. J. Roentgenol. 189(5), 1135-1141 (2007).
12. Houssami, N., Lord, S.J. and Ciatto, S.: Breast Cancer Screening: Emerging Role of New Imaging Techniques as Adjuncts to Mammography. eMJA The Medical Journal of Australia. 190(9), 493-497 (2009).
13. Rangayyan, R.M., Ayres, F.J. and Leo Desautels, J.E.: A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute. 344(3-4), 312-348 (2007).
14. Zhang, P., Kumar, K. and Verma, B.: A Hybrid Classifier for Mass Classification with Different Kinds of Features in Mammography. In: Wang, L., Jin, Y. (Eds.) FSKD 2005. Lecture Notes of Artificial Intelligence, vol 3614, pp. 316-319. Springer, Heidelberg (2005).
15. Zhang, P., Verma, B. and Kumar, K.: Neural vs. Statistical Classifier in Conjunction with Genetic Algorithm Based Feature Selection. Pattern Recognition Letters. 26, 909-919 (2005).
16. Norusis, M.J.: SPSS Advanced Statistics User's Guide. SPSS Inc., Chicago, (1990)
17. Everitt, B.S.: Statistical Methods for Medical Investigations (2nd ed.). Edword Arnold, London (1994).
18. Woods, K.S., Doss, C., Bowyer, K., Solka, J., Priebe, C. and Kegelmeyer, P.: Comparative Evaluation of Pattern Recognition Techniques for Detection of Microcalcifications in Mammography. International Journal of Pattern Recognition and Artificial Intelligence. 7(6), 80-85 (1993).
19. Arbach, L., Bennett, D.L., Reinhardt, J.M. and Fallouh, G.: Classificaiton of Mammographic Masses: Comparison between Backpropagation Neural Network (BNN) and Human Readers. In: Sonka, M., Fitzpatrick, M.J. (Eds.) Medical Imaging: Image Processing. Proceedings of SPIE, vol 5032, pp. 810-818, SPIE Press (2003).
20. Chitre, Y. and Dhawan, A.P.: ANN Based Classification of Mammographic Microcalcifications Using Image Structure Features. State of the Art in Digital Mammographic Image Analysis. 7(6),167-97 (1994).
21. Varela, C., Karssemeijer, N. and Tahoces, P.G.: Classification of Breast Tumors on Digital Mammograms Using Laws' Texture Features. In: Niessen, W.J, Viergever, M.A. (Eds.) Medical Image Computing and Computer-Assisted Intervention-MICCAI 2001: 4th International Conference. LNCS, vol 2208, pp. 1391-1392. Springer, Heidelberg (2001).

22. Antonie, M.L., Zaiane, O.R. and Coman, A.: Application of data mining techniques for medical image classification. In: The Second International Workshop on Multimedia Data Mining (MDM/KDD'2001), in conjunction with ACM SIGKDD Conference, pp. 94-101 (2001).
23. Zhang, P. and Kumar, K.: Analyzing Feature Significance from Various Systems for Mass Diagnosis. In: International Conference on Computational Intelligence for Modeling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06). pp. 141, DOI: 10.1109/CIMCA.2006.46 (2006)
24. Amerian College of Radiology: BI-RADS:American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) (3rd ed.). American College of Radiology, Reston (1998).
25. Eberl, M.M., Fox, C.H., Edge, S.B., Carter, C.A. and Mahoney, M.C.: BI-RADS Classification for Management of Abnormal Mammograms. The Journal of the American Board of Family Medicine (JABFM). 19, 161-164 (2005).
26. Orel, S.G., Kay, N., Reynolds, C. and Sullivan, D.C.: BI-RADS Categorization as a Predictor of Malignancy. Radiology. 211, 845-850 (1999).
27. Baker, J.A., Kornguth, P.J., Lo, J.Y., Williford, M.E. and Floyd, C.E.J.: Breast Cancer: Prediction with Artificial Neural Network Nased on BI-RADS Standardized Lexicon. Radiology. 196, 817-822 (1995).
28. Lo, J.Y., Markey, M.K., Baker, J.A. and Floyd, C.E.J.: Cross-institution Evaluation of BI-RADS Predictive Model for Mammographic Diagnosis of Breast Cancer. AJR. 178, 457-463 (2002).
29. Heath, M., Bowyer, K.W. and Kopans, D.: Current Status of the Digital Database for Screening Mammography. In: Digital Mammography, pp. 457-460, Kluwer Academic (1998).
30. Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2008).

View publication stats