A Hybrid Similarity Measure for Mammography CAD using CBIR Approach

Changwang Liu[1], LiFang[2], Ziju Peng[3], Heng Yang4[*], Jinxin Wan[5], Yihua Lan[3]

1 School of Software
Nanyang Normal University
Nanyang 473061, Henan, China
2 Physics and Electronic Engineering College
Nanyang Normal University
Nanyang 473061, Henan, China
3 School of Computer and Information Technology
Nanyang Normal University
Nanyang 473061, Henan, China
4 School of Education Science
Nanyang Normal University
Nanyang 473061, Henan, China
5. Medical Imaging Department
Lianyungang second people's hospital
Lianyungang 222005, Jiangsu, China

***Abstract* — Breast cancer is one of the most devastating and deadly cancer among women. Early detecting and diagnosing can reduce mortality. Mammography screening is recognized as the most effective method for early detection. However, it is an error-prone task to interpret mammogram. Computer-aided detecting and diagnosis (CADx) systems have been investigated to assist radiologists to interpret mammographic images. In a new type of content-based image retrieval (CBIR) CAD systems, similarity, between a being detected suspicious region and regions of interest (ROI) which are beforehand included in a images database, is most important for the performance of a CAD system. According to the similarity measure, there are two types of CBIR CAD, multi-feature based and information template based. In this paper, a hybrid similarity measure method using the above two measures is proposed. This new hybrid method taken the advantages of two methods into account and avoid their disadvantages. A lot of experiments were carried out. CADs' performance was evaluated using a leave-one-out sampling method and Receiver Operating Characteristics analysis. The experimental results show that the proposed hybrid method has better performance by compare with the two single types of method.**

Keywords - Breast cancer, Mammography, mass, computer-aided detection, diagnosis system

I. INTRODUCTION

For the time being, breast cancer has become the most common type of deadly disease for middle-aged or older women in the whole world [1-4]. From the statistics, there are 62,570 cases of breast carcinoma in situ were expected to be newly diagnosed in the United States in 2014. The overall estimate showed that breast cancer alone is expected to account for 29% (232,670) of all new cancers and it is one of the 3 most commonly diagnosed types of cancer (i.e., breast cancer, lung and bronchus) among women [1]. As in most other countries, breast cancer is also now the most frequently diagnosed cancer in Chinese women. Cases in China account for 12.2% of all newly diagnosed breast cancers and 9.6% of all deaths from breast cancer worldwide [4]. Scientific evidence has shown that early detection and diagnosis of breast cancer is a key factor to reduce mortality and improve the patients' quality of life [5]. At present, a lot of techniques, such as but not limit to tomography/computed tomography (PET/CT) [6], infrared ray imaging [7], ultrasound imaging [8], magnetic resonance imaging (MRI) [9] and mammography screening [10], and so on, have been investigated and developed. Mammography screening remains the most effective method for the detection of early stage breast cancer. Due to the complexity of the screening images, it is a very error-prone, time-consuming task to read and interpret mammograms. Many factors, such as experience or fatigue and subjective judgment, and so on, can influence the radiologists' diagnosis accuracy. A possible approach to improve the diagnosis performance of radiologist is double reading [11-12]. However, double reading is also a very costly method especially in the case of increasing large number of mammograms. Computer-aided detection and diagnosis (CAD) systems have been developed and tested as "a second reader" to distinguish between benign and malignant lesions in mammograms.

II. RELATED WORK

A number of mammographic CAD systems have been developed to aid radiologist in detecting abnormalities by researchers for the last three decades. Previous studies showed that radiologists can detect more micro-calcification

clusters by using CAD schemes. However, it is not successful in detecting cancers associated with mass-like abnormalities when using conventional CAD schemes. Some studies even showed that using CAD actually reduced the performance of radiologist due to the relatively higher false positive detection rates in the clinical practice [13-14]. Hence, many radiologists have low confidence in the CAD-cued mass-like suspicious lesions which lead to their always ignore CAD-cued results [15]. Besides the relatively low performance of CAD schemes in mass detection, there another major limitation in conventional CAD is that those CAD systems are usually difficult to explain the reasoning of the CAD decision-making (the "black-box" type approach) [16]. A new type of CAD schemes using content-based image retrieval (CBIR) approaches have been attracted wide research interest to improve the performance of mammographic CAD in detecting mass-like malignant regions.

Different from those conventional CAD systems using "black-box" type approach, in a CBIR CAD scheme, once a suspicious breast region depicted mass is queried, CBIR system searches for those most similar regions of interest (ROIs) in a constructed reference database, and then computes a likelihood score of being a malignant mass or benign mass by using a CBIR approach on the basis of the above results. At last, the set of most similar ROIs and the CAD-computed score are displayed to radiologists together [17]. Observer performance study indicated those CBIR CADs can improve the radiologists' performance and increase their confidence in their decision making [18].

In term of the similarity measure type used, CBIR CAD schemes are mainly categorized into two types: multi feature-based k-nearest neighbor (KNN) algorithm and pixel value-based template matching methods [17]. Tourassi Georgia D et al. developed a knowledge-based scheme based on a mutual information-based (MI) template matching method for the detection of masses on digitized screening mammograms [19]. In their study, they utilized MI as the similarity metric to determine if a query ROI depicts a mass. A database of 1465 ROIs, including 809 ROIs with confirmed masses and 656 normal ROIs, was employed to evaluate the method. From their report, their CAD scheme achieved performance as high as Az=0.87 0.01. Zheng et al. proposed a CBIR CAD by using a multi-image feature-based KNN similarity searching method [20]. They converted each ROI (including the queried ROI and ROIs in reference database) as a point located in a multi dimensional feature space domain by using a segmentation algorithm and feature exaction method. Then their CAD searches for the points that have the smallest distance to the being queried point (i.e., searches for the ROIs that most similar to the being queried ROI). Xiao hui Wang et al. assessed three methods (i.e., multi feature based KNN method, MI and Pearson's correlation) usually used in CBIR schemes [17]. They established a reference database involving 3000 ROIs and randomly selected 400 ROIs from it to form a testing dataset. The experimental results showed that the Az values of use multi feature KNN method, MI and Pearson's correlation were 0.893±0.009, 0.606±0.021 and 0.699±0.026 respectively. If using ROIs with the size adaptively adjusted according to the actual mass size, the Az values of the last two method can increase to 0.724±0.017 and 0.787±0.016. The study indicated that the CAD used multi feature based KNN method achieved best performance by compare with other two methods due to the diversity of mammographic images.

In this study, we proposed a hybrid similarity measure method which based on the multi feature based method and a template matching method based on information theory (i.e., MI). To test and evaluate the proposed method, we conducted a set of experiments, and then compared it with two single similarity measure methods.

III. MATRIALS AND METHODS

A. Data set

At present, UCSF/LLNL(University of California, San Francisco and Lawrence Livermore National Laboratory, UCSF/LLNF) database, MIAS (Mammographic Image Analysis Society, MIAS) database, and DDSM(Digital Database for Screening Mammography, DDSM) are three commonly used publicly available databases. In this study, all of the mammograms were taken from DDSM. The pixel gray level of the original mammograms is 12 bits (i.e., 4096 gray levels). if we process the original images directly, it is time consuming. We first compressed the range of a pixel gray level from 12 to 8 bits to reduce computational complexity by using an uncompress software provided by the Michael Heath Computer Vision Laboratory at the University of South Florida. Every mammogram depicting mass in DDSM has an overlay file which records abundant information for each mass, such as mass boundaries, pathologic classification and so on. To achieve a fixed size (i.e., 125×125 pixels) ROI for a region depicting mass, we first determined a pixel as the center of an ROI by subjectively estimating, and then we extracted the ROI with 125×125 pixels. Hence, we extracted 426 true-positive ROIs. We used an Artificial Neural Networks (ANN)-CAD to extract 426 negative ROIs. Due to the diversification of breast' normal tissue, each negative ROI selected in our study actually contain a false-positive mass. Hence, 852 ROIs included 426 true-positive ROIs and 426 negative ROIs to form the reference database.

B. *The framework of CBIR CAD scheme*

Figure 1 shows the overview of the CBIR CAD scheme used in this study. This scheme followed a general CBIR framework. The query processes can be described as follows. Once an ROI is provided for querying, the scheme compares the being queried ROI and each ROI to search for the K most similar ROIs in reference ROIs database by using similarity measures (e.g., MI or multi feature-based method). Then, on the basis of the above results, the scheme calculates a decision index (i.e., likelihood score of the being queried ROI depicted a positive mass) using a decision algorithm. At last, the scheme returns the results including the K most similar ROIs and the likelihood score to user.

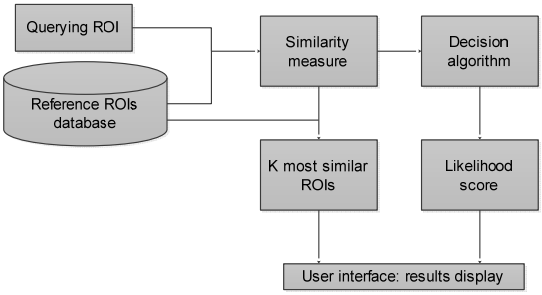

Figure 1. Retrieval framework of the CBIR CAD scheme

The decision algorithm used in this study is described as following formula (1).

$$Score(Q) = \frac{\sum_{i=1, Y_i \in Mass}^{K} (Sim(Q, R_i) \times (K + 1 - O(R_i)))}{\sum_{i=1}^{K} (Sim(Q, R_i) \times (K + 1 - O(R_i)))} \quad (1)$$

where Q represents the querying ROI, R represents one of the ROIs in reference database. K is the number of the most similar ROIs. $Score(Q)$ means likelihood score between Q and $R_i (i \in [1, K])$. $Sim(Q, R_i)$ is the similarity between Q and R_i. $O(R_i)$ is the sort number of each similar ROI in K most similar ROIs by the similarity value.

C. MI based similarity method

MI, a derivation from information theory, is widely used in registering multi-modal images. MI can be described general interdependence between two random variables. We can extend this concept to images easily. Assumed that Q and R are two images, let Q is the queried ROI, R is one of the most similar ROI to Q, the mutual information of Q and R, $MI(Q,R)$ can be computed as follows

$$MI(Q,R) = \sum_x \sum_y P_{QR}(x,y) \log_2 \frac{P_{QR}(x,y)}{P_Q(x) P_R(y)} \quad (2)$$

where $P_{QR}(x,y)$ is the joint probability density function of the two images Q and R, $P_Q(x)$ and $P_R(y)$ are the marginal probability density functions respectively.

The meaning of MI between two images can be described as follows, if two images are similar, those pixels with as certain intensity value in one image should correspond to a more clustered distribution of intensity values in another image. Hence, we can utilize MI as a similarity metric to measure the similar extent between two images. In our CBIR CAD system, the value of MI increase when one of the reference ROI R_i depict more similar structures to the querying ROI Q. The more the R and Q are similar, the more information R provides for Q. In experiments in this article, we used a histogram approach to compute probability density function.

D. Multi feature based similarity method

In a multi feature KNN-based CBIR CAD system, all ROIs are treated as points located in a multi dimensional feature space by using image features. To achieve images features accurately, it is an important step to segment the suspicious region and all ROIs in reference image database. To improve accuracy and robustness of the ROI segmentation, researchers have been investigated a number of segmentation algorithms, including but not limit to region growth algorithm, dynamic programming algorithm and active contour models, and so on. Enmin Song et al. proposed a hybrid segmentation of mass in mammograms using template matching and dynamic programming. Their method achieved good performance by comparing with a lot of segmentation algorithms. In this study, we implemented a vision of it and use it to segmentation all of the ROIs in reference image database and the queried suspicious region.

We extracted 49 features based on the segmentation result of ROI. There are three types of feature categories, 15 edge based features, 19 gray level based features and 15 texture based features. In order to represent image content more powerful, we extracted features as many as possible. However, those discriminating or redundant features were extracted usually with those features have powerful discrimination simultaneous, we should perform feature selection to improve the performance of classification. Genetic algorithm (GA) is a commonly used feature selection method. We employed GA to selection a relative satisfactory feature subset in this article.

In our study, we used Euclidean distance to measure the distance between a querying ROI Q and each of reference ROI R_i, thereby, the similarity between Q and R_i can be defined as following equation 3:

$$Sim(Q, Y_j) = \frac{1}{\sqrt{\sum_{r=1}^{n} (f_r(Q) - f_r(R_j))^2}} \quad (3)$$

where $f_r(Q)$ is the value of the rth ($r \in [1,n]$) feature in the selected feature subset. From equation (3), it can be seen that the larger the $Sim(Q, Y_j)$, the more the two ROIs are alike.

E. Hybrid method

At above two sections, we simply introduced two similarity measures which belonged to two types of categories. To hybridize the two methods, there is a problem should be resolved, that is, which method used firstly in CAD scheme? In general, radiologists have not patience or a lot of time to wait for the retrieval results. To compute MI between two ROIs, the computation process should be performed at the radiologists' query process. Hence, the

computation of MI will spend too much time. In a multi feature based KNN based CAD, the features value of all ROIs in reference database can be pre-calculated. Once a queried ROI is provided, the CAD scheme can calculate similarity between it with reference ROIs fast, and then returns the results in time.

Based on the above analysis, multi feature KNN method was considered to use first in the hybrid method. The complete process of hybrid method is described as follows.

Step 1: achieving an initial most similar ROIs subset.

In this step, we used multi feature based KNN algorithm to search for the initial K most similar reference ROIs for the querying ROI.

Step 2: Calculating the MIs

In this step, CAD scheme calculated MIs between the querying ROI and ROIs in the initial most similar ROIs subset.

Step 3: Resorting the order of ROIs in initial subset

In this step, the original order of ROIs in initial subset were resorted by the MIs in Step 2.

Step 4: calculating decision scores

In this step, the CBIR CAD scheme calculate a decision score for queried ROI.

IV. Experimental Results and Discussions

To test and evaluate the proposed hybrid method and other two methods (i.e., MI and multi feature based KNN) in CBIR CAD system, we conducted a lot of experiments. The performance of CAD schemes using different similarity methods were evaluated using a leave-one-out sampling scheme. Under this experimental scheme, to the give reference ROI database including 852 ROIs, every one of 852 ROIs was singly used once as a queried ROI, and the other remaining 851 ROIs were severed as reference images to form the reference database. CBIR CAD system searched for K most similar ROIs in those 851 ROIs for the queried ROI using the three different similarity measures respectively. Then the CAD systems used decision algorithm described in formula (1) to calculate the likelihood score for each ROI. On the basis of those scores for 852 ROIs, receiver operator characteristic (ROC) data-fitting and analysis program ROCKIT was applied to compute areas under curves (Az) and 95% confidence intervals to assess the performance of CBIR CAD systems using different similarity measures.

The Az values of three CBIR CAD schemes using different methods (i.e., Multi feature based KNN-CAD, MI-CAD and the proposed hybrid method-CAD) are detailed in Table 1.

TABLE I The Az values of three CAD schemes using different methods

CBIR CAD using different similarity mehtods	Az Value	Std. Err.
Multi feature based KNN-CAD	0.8326	0.010
MI-CAD	0.7904	0.015
Hybrid method-CAD	0.8497	0.012

Figure 2 shows the ROC curves of three different CAD schemes.

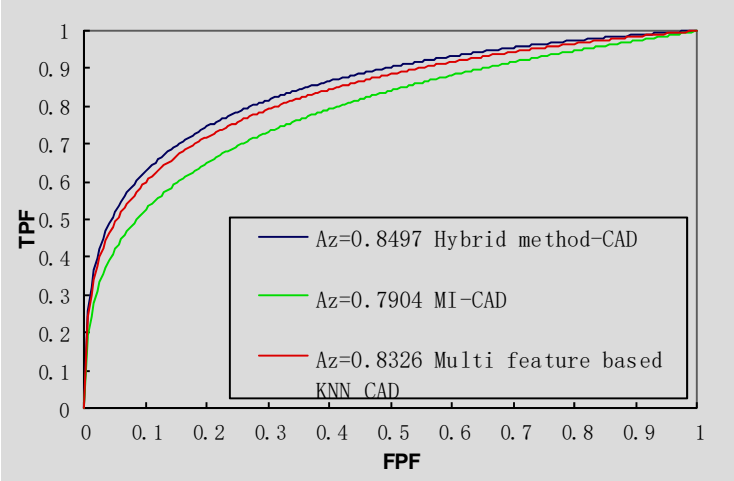

Figure 2. ROC curves of three different CAD schemes.

From Table 1 and Figure 2, the Az values of three CBIR CAD schemes using different methods are 0.8326 ± 0.010, 0.7904 ± 0.015, 0.8497 ± 0.012. The CBIR CAD using proposed hybrid similarity method achieved the best performance. From the experimental results, we also can draw a conclusion that the performance of CBIR CAD using multi feature based KNN is better than the CAD using MI similarity method. We know in feature based CAD, the CAD' performance is heavily depended on accurate segmentation results. Although a number of researchers have worked hard to develop accurate and robust segmentation algorithms, it is remains a difficult task to achieve

satisfactory segmentation results now due to the lesion boundaries are usually overlapping or touching normal structures, obscured or irregular or low contrast. MI, as a template matching algorithm, can avoid to the issue of suspicious ROI segmentation. However, from the definition of MI it can be seen that MI methods is based on pixel value distribution, which made the MI be more affected by the orientation of mass or background regions, or the actual size of mass [17]. The CAD using hybrid method can take advantages of the both multi feature KNN algorithms and MI. experimental results show it achieve the best Az value by compare with the other two different approaches.

V. CONCULSIONS

In the last recent years, it is a very active research topic that developed CAD using CBIR approach. A very important factor which influenced the classification performance in those CBIR CAD systems is similarity measures. In this study, we proposed a hybrid similarity measure in searching those most similar breast mass regions in a large scale reference mammographic images database for CBIR CAD. This method can take advantages of both multi-feature similarity measure and template matching similarity measure. Experimental results show that the Az value yielded by the hybrid method is large then those values by the other two single methods.

ACKNOWLEDGMENT

This study was supported in part by the National Natural Science Foundation of China (Grant No. 61401242), China Postdoctoral Science Foundation(Grant No.2015M572143), the Research Foundation for Advanced Talents of Nanyang Normal University (ZX2014058), Technology Research and Development Program of Lianyungang (Grant No.SH1223), the Research plan for Basic and frontier Research of Henan (Grant No. 142300410044), and the Key Programs of Education Department of Henan Province (Grant No. 14A520057).

REFERENCES

[1] R. Siegel, J. Ma, Z. Zou and A. Jemal, "Cancer statistics," CA: a cancer journal for clinicians, vol.64, pp.9-29,2014.

[2] B. M. Dish, P. M. Dish, S. M. Dish, P. M. Dish, P. M. Dish, V. M. Dish, V. S. Dish, G. S. Dish, S. Calbeans and E. Calbeans, "Cancer survivors---United States," Morbidity and Mortality Weekly Report (MMWR), vol.60,pp.269-272, 2011.

[3] N. M. A. El-Maqsoud and D. M. A. El-Rehim, "Clinicopathologic implications of EpCAM and Sox2 expression in breast cancer," Clinical breast cancer, vol.14, pp. e1-e9, 2014.

[4] L. Fan, K. Strasser-Weippl, J.-J. Li, J. St Louis, D. M. Finkelstein, K.-D. Yu, W.-Q. Chen, Z.-M. Shao and P. E. Goss, "Breast cancer in China," The lancet oncology, vol.15,pp. e279-e289, 2014.

[5] J. Y. Kim, J. Woo, S. S. Lee, H. W. Kim, D. Khang and H. D. Rim, "Psychosocial Factors Predicting Delayed Diagnosis of Breast Cancer: The Role of Marital Relationship Functioning," Korean Journal of Psychosomatic Medicine, vol.22, pp.13-22, 2014.

[6] D. Groheux, E. Hindié, S. Giacchetti, A.-S. Hamy, F. Berger, P. Merlet, A. de Roquancourt, P. de Cremoux, M. Marty and M. Hatt, "Early assessment with 18 F-fluorodeoxyglucose positron emission tomography/computed tomography can help predict the outcome of neoadjuvant chemotherapy in triple negative breast cancer," European Journal of Cancer, vol.50,pp.1864-1871, 2014.

[7] K. Michaelsen, V. Krishnaswamy, B. W. Pogue, K. Brooks, K. Defreitas, I. Shaw, S. P. Poplack and K. D. Paulsen, "Characterization of materials for optimal near-infrared and x-ray imaging of the breast," Biomedical optics express, vol.3, pp.2078-2086, 2012.

[8] W. A. Berg, Z. Zhang, D. Lehrer, R. A. Jong, E. D. Pisano, R. G. Barr, M. Böhm-Vélez, M. C. Mahoney, W. P. Evans and L. H. Larsen, "Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk," Jama, vol.307, pp.1394-1404, 2012.

[9] N. Houssami, R. Turner and M. Morrow, "Preoperative magnetic resonance imaging in breast cancer: meta-analysis of surgical outcomes," Annals of surgery, vol.257, pp.249-255, 2013.

[10] P. Filev, L. Hadjiiski, B. Sahiner, H.-P. Chan and M. A. Helvie, "Comparison of similarity measures for the task of template matching of masses on serial mammograms," Medical physics, vol.32, pp.515-529, 2005.

[11] Waldmann, S. Kapsimalakou, A. Katalinic, I. Grande-Nagel, B. M. Stoeckelhuber, D. Fischer, J. Barkhausen and F. M. Vogt, "Benefits of the quality assured double and arbitration reading of mammograms in the early diagnosis of breast cancer in symptomatic women," European radiology, vol.22, pp.1014-1022, 2012.

[12] D. Georgian-Smith, R. H. Moore, E. Halpern, E. D. Yeh, E. A. Rafferty, H. A. D'Alessandro, M. Staffa, D. A. Hall, K. A. McCarthy and D. B. Kopans, "Blinded comparison of computer-aided detection with human second reading in screening mammography," American Journal of Roentgenology, vol.189, pp.1135-1141, 2007.

[13] X. Wang, L. Li, W. Xu, W. Liu, D. Lederman and B. Zheng, "Improving the performance of computer-aided detection of subtle breast masses using an adaptive cueing method," Physics in medicine and biology, vol.57, pp. 561, 2012.

[14] J. Gilbert, S. M. Astley, M. A. McGee, M. G. Gillan, C. R. Boggis, P. M. Griffiths and S. W. Duffy, "Single Reading with Computer-aided Detection and Double Reading of Screening Mammograms in the United Kingdom National Breast Screening Program, " Radiology, vol.241,pp.47-53, 2006.

[15] B. Zheng, J. K. Leader, G. S. Abrams, A. H. Lu, L. P. Wallace, G. S. Maitz and D. Gur, "Multiview-based computer-aided detection scheme for breast masses," Medical physics, vol.33, pp.3135-3143, 2006.

[16] B. Zheng, "Computer-aided diagnosis in mammography using content-based image retrieval approaches: current status and future perspectives," Algorithms, vol.2,pp.828-849, 2009.

[17] X.-H. Wang, S. C. Park and B. Zheng, "Improving performance of content-based image retrieval schemes in searching for similar breast mass regions: an assessment," Physics in medicine and biology, vol.54, pp. 949-961, 2009.

[18] D. Tourassi, R. Vargas-Voracek, D. M. Catarious Jr and C. E. Floyd Jr, "Computer-assisted detection of mammographic masses: A template matching scheme based on mutual information," Medical Physics, vol.30, pp. 2123-2130, 2003.

[19] Zheng B, Lu A, Hardesty L A, et al. "A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment," Medical physics, vol.33, pp.111-117, 2006.

[20] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," Evolutionary Computation, IEEE Transactions on, vol. 6, pp.182-197, 2002.