Variations on breast density and subtlety of the findings require different computational intelligence pipelines for the diagnosis of clustered microcalcifications

Ioannis I. Andreadis, George M. Spyrou, Panos A. Ligomenides and Konstantina S. Nikita

***Abstract*—In this work, we study the factors that influence the efficacy of a proposed Computer Aided Diagnosis (CAD_x) framework for the diagnosis of clustered microcalcifications (MCs) using a large dataset of mammograms containing cases of varying breast density and findings' subtlety. The reported results indicate that the proposed framework performs towards the right direction, as it appears high classification performance (A_z=0.909) for specific subsets of cases, while outperforming at the same time the performance of the radiologists who evaluated the same cases. The effect of the initial enhancement of mammograms in the CAD_x pipeline is then investigated, by applying three different image enhancement techniques on several subsets of mammograms. We observed that for the considered subsets of dense mammograms, a wavelet-based enhancement algorithm outperformed the rest and provided superior classification performance (A_z=0.849). We indicate therefore that the density of the breast determines the need of different computational algorithms for the analysis of a mammogram and as a result the a priori knowledge of this factor may be exploited for the optimization of the diagnostic process.**

I. Introduction

Microcalcifications are tiny deposits of calcium which can be located anywhere in breast tissue [1]. Mammography is considered nowadays the most effective screening tool for the examination of breast MCs [2]. However, specific inherent limitations of the method led to the development of Computer Aided Diagnosis systems (CAD_e for the detection of suspicious lesions and CAD_x for the diagnosis) [3]. While there are many studies reported in the literature concerning CAD_x algorithms for the specific task [4], there are still many challenges and limitations that have to be faced.

First of all, each CAD_x system follows a specific pipeline which mainly includes: pre-processing of the image, segmentation of microcalcifications, feature extraction, feature selection and finally classification. The specific pipeline is followed when evaluating a new mammogram and remains steady independently of the properties of the considered mammogram. Our proposal is that the CAD_x framework should be flexible depending on the image properties of the evaluated mammogram. In a previous work [5], we have demonstrated that the density of the breast plays a crucial role in the proper selection of image features used for the analysis of a cluster. This fact may imply that different tissue types may require different computational intelligence pipelines for the CAD_x analysis.

Secondly, there are many studies where the proposed CAD_x pipeline differs slightly from the one discussed above. For example, few studies consider as a prerequisite step the image enhancement preprocessing of the mammogram, whose role is to improve the contrast features of a Region of Interest (ROI) in a mammogram. There are many proposed image enhancement algorithms reported in the literature [6], based mainly on histogram equalization, unsharp masking, linear stretching and wavelet decomposition. Papadopoulos et al. [7] have compared several image enhancement algorithms to study their effect on the computer aided detection (CAD_e) performance of cluster of MCs. Sivaramakrishna et al. [8] compared also various enhancement methodologies, using as evaluation procedure a radiologist's score. However, it seems that there is no work comparing different image enhancement methods to study their effect on a CAD_x validation tool.

In this study, we investigate the performance of a proposed CAD_x framework [5] for the diagnosis of clusters of MCs on a large dataset of mammograms, while we further examine the effect of an image enhancement stage using the performance of the CAD_x system as a validation tool. To the best of our knowledge, the dataset used in the specific study is the largest dataset that has been used in CAD_x diagnosis for cluster of MCs, as we used almost all the cases that are available by the Digital Database of Screening Mammography (DDSM) [9]. We exploit the specific fact to work in two main axes: firstly, we apply the proposed CAD_x framework on all the available cases in order to study the effect of the breast density and the subtlety of findings on the performance of the CAD_x system. Secondly, based on the aforementioned factors, we work on well defined subsets of specific image properties in order to study the contribution of an enhancement preprocessing stage to the correct categorization of clusters of MCs. Three different enhancement algorithms are considered and their efficacy is

Manuscript received September 20, 2013.

I. I. Andreadis is with the the School of Electrical and Computer Engineering, National Technical University of Athens, 15573, Athens, Greece (e-mail: iandr@ biosim.ntua.gr).

G. M. Spyrou is with the Biomedical Research Foundation Academy of Athens, Athens, Greece (e-mail: gspyrou@bioacademy.gr).

P. A. Ligomenides is with the Academy of Athens, Athens, Greece (pligo@academyofathens.gr).

K. S. Nikita is with the School of Electrical and Computer Engineering, National Technical University of Athens, 15573, Athens, Greece (corresponding author phone: +30-2107722968; fax: +30210 7722320; e-mail: knikita@ece.ntua.gr).

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

compared based on the classification results achieved by the CAD_x system. We investigate whether the a priori knowledge of the breast density or the subtlety of findings may require different enhancement algorithms for the optimization of the diagnostic process.

II. Methods and Materials

A. Data Collection

We used cases provided by the DDSM database which is the largest database currently available. We extracted all ROIs from the DDSM which included clusters of MCs. We excluded only cases with unproven pathology result and a few cases whose ROI included almost the whole breast area. We ended with a dataset of 1715 ROIs of both views. For each case in the DDSM database, the density of the breast, the subtlety of the cluster of MCs and the BI-RADS assessment had been specified by an experienced radiologist and are provided in the annotation files accompanying each case's mammograms. The specific dataset of the 1715 mammograms will be divided on subsets following the density of the breast and the subtlety of the clusters, in order to investigate the effect of these factors on the efficacy of a proposed CAD_x framework discussed in the next section.

B. CAD_x Framework

Each CAD_x system includes specific steps such as the segmentation of microcalcifications, the feature extraction process and the final classification of a considered case [3]. The framework investigated in the current study has been presented in a previous work [5]. A segmentation algorithm is used to isolate the detected MCs from the surrounding breast tissue. Afterwards, feature extraction methodologies are applied to extract image features concerning the shape and the brightness of each MC in the cluster, the distribution of the MCs in the cluster and the morphology of the whole cluster and finally first-order and second-order statistics to describe the texture of the whole ROI. Finally, a classification scheme is required to discriminate between benign and malignant cases. We used a Support Vector Machines classifier using a Gaussian RBF kernel. The leave-one-out (LOO) method was applied for performance evaluation, while 10-fold cross validation has been performed for the parameterization of the SVM algorithm by adjusting the regularization parameter C and the parameter g [10]. Prior to the classification phase, a feature selection process is applied to find a satisfactory feature subset, eliminating features that may be irrelevant to the classification task. To this end, the Sequential Forward Selection (SFS) method has been applied to choose an appropriate features subset. The evaluation of the CAD_x performance is based on ROC analysis.

C. Image Enhancement Algorithms

We investigated three enhancement algorithms. The first two are well-known algorithms derived from conventional image analysis, while the third is based on wavelet decomposition of the image. Each algorithm is applied prior to the segmentation task of MCs. Few details are reported below for each algorithm:

1) Contrast Limited Adaptive Histogram Equalization (CLAHE): It is a well-known algorithm which is highly used for image enhancement [7]. The algorithm is based on dividing the initial image in sub-images and performing histogram equalization on each of them. Two parameters have to be adjusted, the size of each sub-image that was selected equal to 10×10 and the clip threshold for contrast expansion which was defined equal to 0.1. The specific values have been proven to achieve the best image quality after proper subjective tests.

2) Local Range Modification (LRM): The LRM algorithm implements a linear stretching formula to generate an enhanced image transforming the pixel values of the original image [7]. The method is based on the double processing of the initial image by dividing it in sub-blocks and estimating the maximum and minimum value per block. The specific values are used to define the parameters of a linear equation which is subsequently used for the transformation of each pixel's value. The size of the block has been defined equal to 51×51.

3) Wavelet-Based Algorithm (WB): We applied an algorithm based on the wavelet decomposition of the image which proposes the use of a multiscale contrast measure in the wavelet domain to enhance directly the contrast of the initial image [11]. A local contrast is estimated for each direction (vertical, horizontal and diagonal) and therefore the enhancement of the contrast can be performed in different scales of the image. The method is simple enough as only one parameter has to be defined (λ). For the current study, the parameter λ has been defined equal to 4.9, as we observed that the best image quality has been achieved for the specific value.

III. Results

A. Effect of breast density and subtlety of findings

Each cluster in the database is characterized by a value of subtlety (1 to 5), where smaller values imply increasing grade of difficulty for cluster's analysis. It is expectable that the performance of the CAD_x framework should be highly influenced from the specific value. We divided the initial dataset of 1715 cases on five subsets, depending on the subtlety value of each cluster. We applied then the proposed CAD_x framework discussed in section II, paragraph B and we performed ROC analysis, recording the values of accuracy (ACC), sensitivity (SN), specificity (SP) and Area Under Curve (A_z). For reasons of completeness and comparison, we estimated the A_z value for the radiologists who assessed the same cases, following the method proposed in [12] using the BI-RADS assessment [13]. In table 1, we present the number of cases for each subset (#Cases, inside the brackets we fill the number of benign and malignant cases respectively) and the results obtained for the CAD_x system and the radiologists (RD).

TABLE I
CAD$_X$ AND RADIOLOGISTS PERFORMANCE DEPENDING ON CLUSTER'S SUBTLETY

Subtlety	***#Cases***	***CADx***				***RD***
		ACC	***SN***	***SP***	***A_z***	***A_z***
1	170 (83/87)	0.529	0.54	0.518	0.556	0.538
2	390 (213/177)	0.61	0.542	0.667	0.654	0.47
3	527 (279/248)	0.66	0.597	0.716	0.686	0.618
4	371 (194/177)	0.698	0.65	0.742	0.756	0.699
5	257 (113/144)	0.852	0.882	0.814	0.909	0.932

We repeated the same pipeline by dividing this time the initial subset of 1715 cases based on the breast density of each case. According to the BIRADS standard [13], there are four ratings of the breast based on its density: (1) entirely fat, (2) scattered fibrogranular densities, (3) heterogeneously dense and (4) extremely dense. As in the case of the subtlety value, we present in table 2 the performance measures for both the CAD$_x$ framework and the radiologists who assessed the same subsets.

TABLE II
CAD$_X$ AND RADIOLOGISTS PERFORMANCE DEPENDING ON BREAST DENSITY

Density	***#Cases***	***CADx***				***RD***
		ACC	***SN***	***SP***	***A_z***	***A_z***
1	120 (69/51)	0.767	0.686	0.826	0.823	0.762
2	533 (269/264)	0.689	0.655	0.721	0.759	0.654
3	579 (288/291)	0.636	0.632	0.639	0.695	0.682
4	483 (256/227)	0.592	0.524	0.652	0.624	0.619

B. *Effect of Image Enhancement*

The second step of our computational experiments is to evaluate the effect of an image enhancement processing stage on the diagnosis of clusters of MCs using the CAD$_x$ system as a validation tool. Based on the results of the previous section, the algorithms have to be compared on cases of different density and subtlety values. In order to reduce the computational cost and experimental time, we did not use all the available 1715 cases. Instead, we selected randomly only 1000 cases dividing them on the following subsets, namely: (S1) cases of high density (density 3, 4) and low subtlety (subtlety≤3), (S2) cases of high density and high subtlety (subtlety>3), (S3) cases of low density (density 1, 2) and low subtlety and finally (S4) cases of low density and high subtlety. All subsets include 250 cases (125 benign and 125 malignant) and obviously no case is included in more than one subset. The considered subsets are enough large to enable us to extract results and conclusions with high generalization ability. We added into the CAD$_x$ flowchart prior to the segmentation of MCs each one of the three image enhancement algorithms considered. We applied the rest pipeline on all the four subsets and we recorded the results which are presented in table 3. The results concern all the considered subsets of mammograms and the three enhancement algorithms (CLAHE, LRM, WB). We also add the performance achieved by the CAD$_x$ system without applying any enhancement algorithm (NO_ENHANCE). We highlight in bold the best performance for each subset, based on the greatest A_z value.

TABLE III
CAD$_X$ PERFORMANCE FOR EACH DIFFERENT SUBSET OF CASES AND DIFFERENT IMAGE ENHANCEMENT ALGORITHM

SUBSET	***METHOD***	***ACC***	***SN***	***SP***	***A_z***
S1	CLAHE	0.622	0.644	0.607	0.673
	LRM	0.592	0.668	0.528	0.642
	WB	**0.663**	**0.682**	**0.643**	**0.735**
	NO_ENHANCE	0.61	0.647	0.589	0.677
S2	CLAHE	0.765	0.728	0.81	0.839
	LRM	0.753	0.729	0.778	0.809
	WB	**0.771**	**0.769**	**0.774**	**0.849**
	NO_ENHANCE	0.757	0.731	0.785	0.806
S3	CLAHE	0.69	0.672	0.713	0.756
	LRM	0.744	0.724	0.766	0.808
	WB	0.711	0.656	0.766	0.77
	NO_ENHANCE	**0.731**	**0.71**	**0.755**	**0.816**
S4	CLAHE	0.763	0.663	0.865	0.836
	LRM	**0.786**	**0.686**	**0.885**	**0.893**
	WB	0.805	0.756	0.858	0.84
	NO_ENHANCE	0.784	0.74	0.829	0.87

IV. DISCUSSION

As far as the results based on cluster's subtlety are concerned, we observe that the CAD$_x$ system performs indeed towards the right direction. As expected, the results are poor for difficult cases (subtlety 1), but the performance is increased as the subtlety value increases. This fact indicates that the proposed algorithms appear a steady behavior and do not classify the cases randomly. It is noticeable that for the majority of the subsets, the CAD$_x$ framework presents superior A_z value to that achieved by the radiologists. Observing the results of table 2, we may proceed on corresponding conclusions for the density rating. The classification results decrease as the rating of the breast density increases. It is obvious that dense tissues harden the analysis of a cluster of MCs. Similar conclusions may be extracted for the case of radiologists, as their performance also decreases when assessing subsets of denser tissues. It is also noteworthy that the A_z value achieved by CAD$_x$

outperforms the corresponding value of radiologists in all the considered subsets of cases.

The general conclusion that we may extract based on the specific results is that the CAD_x system works towards the right direction. Its performance is affected by the same factors that affect the radiologists in daily clinical practice. The fact that the performance of the CAD_x system is almost always superior to the performance of the radiologists who assessed the same cases reveal the potential of the proposed algorithms to support radiologists' diagnostic process.

Since the breast density and findings' subtlety influence strongly the performance of CAD_x methodologies, we studied the effect of enhancement algorithms on the classification performance using various subsets, as these factors may determine different computational intelligence pipelines for the analysis of a mammogram. Studying Table III, we may notice that image enhancement does not always lead to improvement of classification results. For example, we observe that in the case of subset S3 the greater A_z value is achieved when no enhancement algorithm is applied. Similarly, concerning the subset S4, the performance of the unenhanced pipeline is close enough to that of the LRM algorithm which is the best for the specific subset. In other words, it seems that in cases containing fatty tissues (low density tissues, subsets S3 and S4) the image enhancement stage cannot improve the classification results, as the LRM algorithm is the only method able to provide results close to those obtained without enhancement. However, we can observe that in both subsets S1 and S2 (high density cases) the greatest performance is achieved when using the WB algorithm. This observation indicates the potential of the specific algorithm to enhance the classification when cases with dense tissues are under consideration. On the contrary, it seems that in the case of fatty tissues (subset S3 and S4) the algorithm is not beneficial for the classification performance. The CLAHE algorithm has similar behavior, but presents comparable or even inferior performance to that of the unenhanced pipeline.

The majority of the CAD_x studies found in the literature use different sets of mammograms for the evaluation of their algorithms, while there is usually no information on the composition of the dataset (whether there are cases of varying breast density and subtlety). As a result, a straightforward comparison is not feasible. In general, high A_z values are reported (A_z>0.8) for small datasets (cases<150), while in this study we achieve similar performance for larger subsets. An overview of the ROC performance of representative studies may be found in [4], [5].

V. Conclusions

The aim of this paper was to study the factors that may affect the performance of a CAD_x framework for the diagnosis of clusters of MCs and investigate whether alternative pipelines may improve its performance. We demonstrated that the most important factor which influenced the CAD_x performance is the composition of the evaluated dataset, due to the fact that inherent properties of a mammogram such as the density of the breast and the subtlety of the findings play crucial role to the classification performance achieved. The specific conclusions led us to study the effect of a preprocessing stage which aims to the improvement of the contrast in the ROI. Several enhancement algorithms were compared on subsets of varying breast density and findings' subtlety and the obtained results revealed that different pipelines should be followed depending on these factors. The wavelet-based enhancement algorithm seems capable to improve the classification results in cases of dense tissues. On the contrary, for breast tissues of low density, the LRM algorithm may provide comparable results to those achieved without enhancement. The conclusions extracted in this study may form a proper baseline to work towards refining the diagnostic process of the CAD_x framework, by adopting new pipelines depending on the properties of each mammogram.

References

[1] M. Lanyi, " Microcalcifications in the breast—a blessing or a curse? A critical review," *Diagn. Imaging Clin. Med*., vol. 54, pp. 126–145, 1985.

[2] H. H. Ng and M. Muttarak, "Advances in mammography have improved early detection of breast cancer," *J. HK Coll. Radiol.*, vol. 6, no. 3, pp. 126–131, 2003.

[3] M. Giger, H. Chan and J. Boone "Anniversary paper: history and status of CAD and quantitative image analysis: the role of medical physics and AAPM," *Med. Phys.*, vol. 35, pp. 5799–5820, 2008.

[4] M. Elter and A. Horsh, "CADx of mammographic masses and clustered microcalcifications: a review," *Med. Phys.*, vol. 36, pp. 2052–2068, 2009.

[5] I. Andreadis, G. Spyrou and K. Nikita, "A comparative study of image features for classification of breast microcalcifications," *IOP Measurement Science and Technology*, vol. 22, no. 11, 2011.

[6] J. Tang, R. Rangayyan, J. Xu, I. El Naqa and Y. Yang, "Computer-Aided Detection and Diagnosis of Breast Cancer With mammography: Recent Advances", *IEEE Trans. Inform. Tech. Biomed*., vol. 13, no. 2, pp. 236-251, 2009.

[7] A. Papadopoulos, D. Fotiadis and L. Costaridou, "Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques," *Comput. Biol. Med.*, vol. 38, pp. 1044-1055, 2008.

[8] R. Sivaramakrishna, N.A. Obuchowski, W.A. Chilcote, G. Cardenosa and K.A. Powell, "Comparing the performance of mammographic enhancement algorithms—A preference study,", *Amer. J. Roentgenol.*, vol. 175, no. 1,pp. 45–51, 2000.

[9] M. Heath, K. Bowyer, D. Kopand, R. Moore and W. Kegelmeyer. "The Digital Database for Screening Mammography,"in 2001 Proc. of the 5th IWDM, Yaffe M. Medical Physics Publishing, pp. 212-218.

[10] C.J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discovery*, vol. 2, pp. 121–167, 1998.

[11] J. Tang, X. Liu and Q. Sun, "A Direct Image Contrast Enhancement Algorithm in the Wavelet Domain for Screening Mammograms," *IEEE Journal of Selected topics in Signal Processing*, vol. 3, no. 1, pp.74-80, 2009.

[12] A. Karahaliou, S. Skiadopoulos, I. Boniatis, P. Sakellaropoulos, E. Likaki, G. Panayiotakis and L. Costaridou, " Texture analysis of tissue surrounding microcalcifications on mammograms for breast cancer diagnosis," *Br. J. Radiol.*, vol. 80, pp. 648-656, 2007.

[13] American College of Radiology (ACR), *ACR Breast Imaging Reporting and Data System, Breast Imaging Atlas*, 4th Edition, Reston, VA. USA, 2003.