Breast cancer computer-aided diagnosis systems: analysis of breast tissues in mammograms and ultrasound images

Mohamed Abdel-Nasser *

Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili
Tarragona, Spain
mohamed.abdelnasser@urv.cat

1 Introduction

Breast cancer attacks women in their 40s. Based on a statistic in the European Union, breast cancer is the leading cause of cancer death in 2014 [5]. Early detection through screening with computer aided diagnosis (CAD) systems can help to reduce the fatalities. The causes of breast cancer are still unknown; however, there are several factors that can indicate the risk of breast cancer such as *age*, *family profile*, *genetics* and *breast density*. Mammographies are considered the most effective tool for early detection of breast cancer. Figure 1 shows an example for breast screening using a mammography and the appearance of breast cancer in the mammogram.

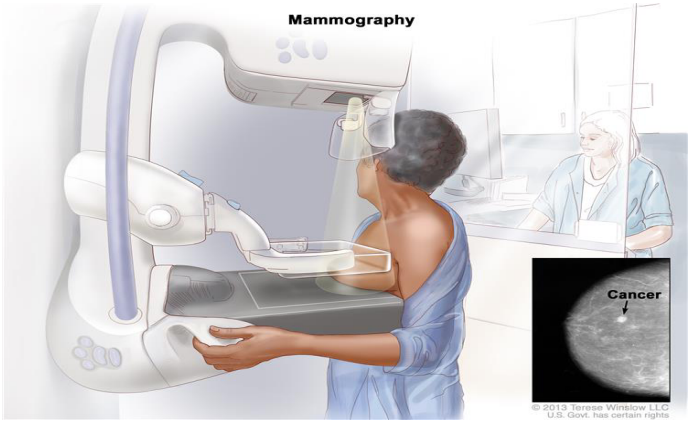

Fig. 1: Breast screening through mammography. Reprinted from: http://www.cancer.gov/types/breast/patient/breast-treatment-pdq

* PhD advisors: Dr. Antonio Moreno and Dr. Domenec Puig.

In a mammography, each breast is compressed using compression plates, and then X-rays are used to take images of breast tissue. A study in [4] showed that younger women usually have *denser breasts* than older women. *Dense* breasts have more glandular and fibrous tissues, and they appear white in the mammogram. Thus, they hide masses, which also usually appear white in mammograms. On the contrary, *fatty* tissues appear grey in mammograms.
Breast ultrasound (BUS) images are superior to mammograms in their ability to detect abnormalities in dense breasts. BUS is considered a complementary tool to mammograms in breast cancer detection. They cannot replace a mammogram for breast screening, but they can provide more information to physicians. Figure 2 shows breast tissues in a BUS image and a mammogram for the same breast.

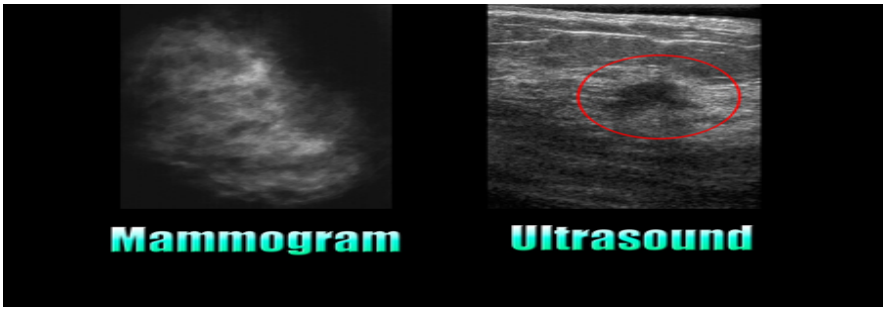

Fig. 2: A mammogram and BUS image for the same breast. Reprinted from: *http://www.thedoctorstv.com/articles/1442-procedures-you-need-to-know*

CAD systems are typically used to analyse mammograms in screening. They use several *machine learning*, *computer vision* and *image processing* techniques. A breast cancer CAD system consists of three main stages: *segmentation* of a region of interest (ROI) from the image, *feature extraction* from the ROI, and *classification*. The literature shows no consensus on the optimal feature set for breast tissue characterization. A poor description of the breast tissues leads to a high number of false positives (ROIs interpreted by a CAD system as abnormal cases when they are actually normal).
In our study, we focus on the feature extraction sub-task of a breast cancer CAD system, where we analyse breast cancer in both mammograms and BUS images. Although several feature extraction methods have been proposed for mammogram and BUS images analysis, improving the classification results remains a challenge. Texture analysis methods constitute one of the options for improving the performance of classification algorithms used in breast CAD systems. To that purpose, we analyse the performance of various texture analysis methods for breast mass classification aiming at reducing the number of *false positives*. Moreover, we propose novel descriptors.

2 Texture analysis for breast mass classification

2.1 Materials

In our experiments, we use several publicly available mammography databases such as mini-MIAS, DDSM and INbreast [6]. To analyse BUS images, we use a BUS database that was collected in UDIAT Diagnostic Centre of Sabadell (Spain).

2.2 Methods

We analyse the performance of several texture analysis methods for breast mass classification in mammograms and BUS images. We have selected widely used texture analysis methods such as local binary pattern (LBP), local directional number (LDN), histogram of oriented gradients (HOG), Gabor filters (GF) and Haralick's features (HAR). In addition, we used several classification methods such as k-nearest neighbour, linear discriminant analysis, linear support vector machines (LSVM), non-linear support vector machines (NLSVM), and random forests (RF). The performance of the CAD systems was measured in terms of the area under the curve (AUC) of the receiver operating curve (ROC), the sensitivity and the specificity. Our study is organized as follows.

1- Study of the performance of various texture analysis methods with breast mass classification in mammograms.
2- Study the effect of breast density on texture analysis for mass classification in mammograms.
3- Propose a fuzzy logic-based texture analysis method for breast mass detection in BUS images.

3 Results

In [1,2], we performed a twofold analysis. Firstly, we analysed the performance of several texture methods individually. As shown, LDN improved the results of well-known texture analysis methods like LBP, HOG, HAR or GF. It achieved a sensitivity of 84.0% with the KNN and a specificity of 99.0% with the LSVM. Secondly, we evaluated two feature combination techniques: the majority output of the individual classifiers and building new models on the concatenation of features provided by different texture methods. Among all combinations, $LDN + LBP$ gave the smallest percentage of false positives. It achieved a sensitivity of 92.0% with the NLSVM and a specificity of 96.5% with the RF.

In [3], we proposed the fuzzy local directional pattern (FLDP) for breast tissue characterization. It describes each pixel in a given ROI by its edge responses and makes use of fuzzy membership functions. The rationale behind

the use of fuzzy logic is to compensate the uncertainty of the visual appearance of breast tissues due to noise, breast density and the variation in breast compressions. FLDP properly discriminates between mass and normal tissues in both dense and fatty breasts. We showed that the use of FLDP improved the classification results of breast tissues in BUS images when compared to some of the state-of-the-art descriptors such as LBP, HOG, LDN, HAR and GF. It achieved an AUC of 0.87 with the LSVM and 0.914 with the NLSVM.

4 Future work

The future work includes a study for breast tumour changes through several motion analysis methods. We will explore the use of the anatomical information in the breast region (nipple position, pectoral muscle, fatty and dense regions) to tune the results of motion analysis methods.

Acknowledgement. Mohamed Abdel-Nasser was awarded a Ph.D. grant from URV in 2013 (Martí Franquès program). Our work was partly supported by the Spanish Government through project TIN2012-37171-C02-02.

References

[1] Mohamed Abdel-Nasser, Antonio Moreno, and Domenec Puig. Towards cost reduction of breast cancer diagnosis using mammography texture analysis. *Journal of Experimental & Theoretical Artificial Intelligence*, 0(0):1–18, 2015.

[2] Mohamed Abdel-Nasser, Domenec Puig, and Antonio Moreno. Improvement of mass detection in breast X-ray images using texture analysis methods. In *Artificial Intelligence Research and Development: Recent Advances and Applications*, volume 269, pages 159–168. IOS Press, 2014.

[3] Mohamed Abdel-Nasser, Domenec Puig, Antonio Moreno, Adel Saleh, Joan Marti, Luis Martin, and Anna Magarolas. Breast tissue characterization in x-ray and ultrasound images using fuzzy local directional patterns and support vector machines. In *Proc. of the International Conference on Computer Vision Theory and Applications (VISAPP 2015)*, volume 1, pages 387–394, 2015.

[4] Karla Kerlikowske, Deborah Grady, John Barclay, Edward A Sickles, and Virginia Ernster. Effect of age, breast density, and family history on the sensitivity of first screening mammography. *Journal of the American Medical Association*, 276(1):33–38, 1996.

[5] M. Malvezzi, P. Bertuccio, F. Levi, C. La Vecchia, and E. Negri. European cancer mortality predictions for the year 2014. *Annals of Oncology*, 2014.

[6] Inês C Moreira, Igor Amaral, Inês Domingues, António Cardoso, Maria João Cardoso, and Jaime S Cardoso. Inbreast: toward a full-field digital mammographic database. *Academic radiology*, 19(2):236–248, 2012.