COMPLIANCE OF PUBLICLY AVAILABLE MAMMOGRAPHIC DATABASES WITH ESTABLISHED CASE SELECTION AND ANNOTATION REQUIREMENTS

Inês C. Moreira[1,2,3,4], Gustavo Bacelar-Silva[1,2,5] and Pedro Pereira Rodrigues[1,5]

[1]*Faculty of Medicine of the University of Porto, Al. Prof. Hernâni Monteiro, Porto, Portugal*
[2]*Faculty of Sciences of the University of Porto, Rua do Campo Alegre, Porto, Portugal*
[3]*Superior School of Health Technology of Porto, Rua Valente Perfeito, Vila Nova de Gaia, Portugal*
[4]*INESC Porto, Faculty of Engineering of University of Porto, Rua Dr. Roberto Frias, Porto, Portugal*
[5]*CINTESIS, Al. Prof. Hernâni Monteiro, Porto, Portugal*

Keywords: Mammographic database, CAD, Computer-aided detection, Computer-aided diagnosis.

Abstract: Mammographic databases play an important role in the development of algorithms aiming to improve Computer-Aided Detection and Diagnosis systems (CAD). However, these often do not take into consideration all the requirements needed for a proper study, previously discussed at the Biomedical Image Processing Meeting in 1993. Case selection and annotation requirements are the most commonly referenced in literature, when describing a database used for the development of such algorithms. This work aims to assess the compliance and suitability of case selection and annotation requirements in the publicly available mammographic databases for development and optimization of CADs. A literature review has been made, applying proper selection criteria related to the research question. In the literature, we found citations to 3 publicly available mammographic databases and ten having restricted access. Through the analysis of the results attained, we noticed that none of the two requirements previously described is on its way to be fully complied in mammographic databases. We can conclude that researchers need a database that fulfils all the mentioned requirements in order to develop efficacious and effective CAD systems. We also believe that the requirements, discussed in 1993, need to be reviewed and updated. New paradigms and ideas to increase algorithms' performance are needed in order to improve CAD schemes.

1 INTRODUCTION

Breast cancer related mortality incidence reaches 1500 women every year in Portugal, whereas in the Europe Union breast cancer is responsible for one in every six deaths from cancers in women (Eurostat, 2009). The earlier detection of breast cancer through mammographic screening is strongly recommended by all medical community, in order to decrease its associated mortality rate (WHO, 2009).

The common findings that can be found on mammography are masses, calcifications, architectural distortion of the breast tissue, and asymmetric densities when comparing the two breasts (ACR, 2003). In order to standardize the Terminology of the mammographic report, of the assessment of the findings and of the action to be taken, has been proposed by the American College of Radiology (ACR), the Breast Imaging Reporting and Data System (BI-RADS) scale (ACR, 2003). Other important characteristic referred by the ACR is the breast composition tissue, related to the breast density shown in X-Ray (ACR, 2003).

Computer-Aided Detection and Diagnosis (CAD) systems have been developed in the past two decades to assist the radiologist, in order to provide a second opinion (Bin Zheng et al., 2003). In order to increase the efficiency and obtain greater sensitivity/specificity from these systems, researchers have been developed algorithms for detection and segmentation of abnormalities. To proper develop their techniques, researchers need a large number of mammograms to test and tune their algorithms to recognize signs of abnormalities (Nishikawa, 1998). Thus, mammographic databases play an important role in the development of algorithms aiming to detect and diagnose lesions. They are also important because they are used to test

algorithms and CAD schemes, and allow the comparison of results from different studies (e.g. Jiang et al., 2008).

According to Nishikawa (1997), mammographic databases should take into consideration the following requirements. **Case Selection**: the database should include various cases with images with none and all types of findings, and also all types of breast density. Normal images with structures that may be misleading (e.g. superimposed tissue that looks like a mass) are important in order to make the classifiers more robust. Also, the cases should be collected by a specialist experienced in mammography, and each case should contain the four standard views, unless it is a case from a patient with one breast only. It is considered that for each 100 cases, approximately 200 images should contain a lesion. **Ground Truth**: Biopsy proof for all cases should be available, and for cases in which a biopsy is not recommended, the mammography should have the same BI-RADS for at least three years. Annotations should include the "ground truth" concerning the degree of malignancy, the location and the boundary of the lesion and this outline should be performed by a specialist. **Associated Information: c**linical history (e.g. age, family history, and previous biopsies) can be useful to improve the performance of CADs. **Requirements of the digitizer**: This is still a point of controversy, but one common approach is to digitalize at a very small pixel size, for example, at 25 microns. **Organization of Database**: A specific file format for digital mammograms does not exist. Medical images are usually saved in the DICOM (Digital Imaging and Communications in Medicine) format that gathers not only the image but also some related metadata. A division of the images on training and test sets should also be suggested, in order to have comparable sets and different methods can be compared. **Distribution of Database**: The database should be available, preferentially over the World Wide Web. Continuous user support is also indispensable.

There are several image databases, both public and restricted to individual groups, which are used by researchers in the breast cancer area. However, these often do not take into consideration all the requirements needed for a study (Oliver et al., 2010).

Nishikawa (1998) made a review about mammographic databases for teaching and research purposes. However, the article was written in 1998 and we consider that there is a need to research and find out if there are new databases. Thus, mentioning some aspects that are not specified in

that work will bring some advantage for this area, like the case selection and annotation type of lesion used in each of them. These requirements are the most commonly referenced in literature, when describing the database used for the development of the algorithm under study.

This work aims to assess the compliance and suitability of case selection and annotation requirements in the publicly available mammographic databases for development and optimization of CADs.

2 MATERIAL AND METHODS

A bibliographic search in three digital libraries – Pubmed, ISI WEB of Knowledge and SCOPUS has been made between November 2010 and January 2011. Our inclusion criteria considered: 1) mammographic databases description; 2) work related with mammographic databases, such as algorithms for CAD detection and diagnosis of abnormalities; and the exclusion criteria were: 1) not English written; 2) studies considering other modalities, as ultrasound or magnetic resonance; 3) other related work, like content-retrieval based issues or search function systems. Relevance and suitability of the papers on the subject under study were assessed using and the abstract and full paper analysis, applying the inclusion and exclusion criteria. Some articles, whose access was restricted and for which contact with the author could not be established, were also excluded.

Of the 32 selected articles, 13 mammographic databases were described, from which 3 are available databases and 10 have restricted access. 19 papers are related to algorithms' development where the authors used one or more of these databases.

3 THE MAMMOGRAPHIC IMAGE ANALYSIS SOCIETY DIGITAL MAMMOGRAM DATABASE

The MIAS database (Suckling, 1994) is the oldest one and it is widely used in literature, although it is no longer supported.

In (Rangayyan et al., 2000) the authors noticed that there is a big amount of benign findings, in relation to the malign ones. Due to the increasing usage of the ACR standard, in (Oliver, Lladó, et al.,

2010) it was decided to classify the set of mammograms according to that reference.

MIAS annotations are considered to be not sufficient for some studies (e.g. Oliver et al., 2010), where all circumscribed and spiculated lesions had to be manually segmented. Another drawback is the resolution to which it has been digitalized, which makes it not suitable for experiments on detection of micro-calcifications (Rojas Dominguez & Nandi, 2007). Llobet (2005) considered that in case of calcifications, the ground truth region contains much healthier tissue than affected tissue, and justifies that with the shape of calcifications, which are small lesions spreaded into a wide area and the annotation includes all that. For this reason, calcifications were not included in his study.

4 THE DIGITAL DATABASE FOR SCREENING MAMMOGRAPHY

The DDSM database (Heath et al., 1998) is the most used, but is no longer supported.

Annotations include pixel level boundary of the findings. There are several papers reporting satisfactory results using this annotation (e.g. D. Wang et al., 2009). However, as noted in some studies (e.g. Enmin Song et al., 2010), they are not adequate for the validation of segmentation algorithms because the precision is not good enough.

5 THE BANCOWEB LAPIMO DATABASE

This database (Matheus and Schiabel, 2010) is a more recent database which is supported and users can contribute to the database.

It has annotations in only some images, as a Region of Interest (ROI), but all have textual description of finding. We didn't find any work related to this database, due being a recent project.

A summary of these databases concerning case selection and annotation can be found in Table 1.

6 OTHER DATABASES

Ten other databases were found. However, given that they are not publicly available, they were not considered (the complete list of references is omitted due to space requirements, but can be made available upon request).

Table 1: Available database's case selection and annotation requirements.

	MIAS	DDSM	BancoWeb
Number of Cases	161	2620	320
Views	MLO	MLO and CC	MLO, CC and other
Number of Images	322	10480	1400
Breast Density	YES (not ACR)	YES (ACR)	YES (not ACR)
Lesion Type	All kind (concentration of spiculated masses)	All kind	All kind
Breakdown of images	204 normal 66 benign 52 malign	2780 normal 4044 benign 3656 malign	294 normal 994 benign 112 malign
Ground Truth	Centre and radius of a circle around the interest area	Pixel level boundary of the findings	ROI is available in a few images only
BI-RADS	NO	YES	YES
Biopsy Proven	YES	YES	YES

7 DISCUSSION

None of the two requirements previously described is on its way to be fully complied in mammographic databases. Concerning case selection, there's a need to review the common thought that "more is better". Zheng et al. (2010) claim that, in the development of CAD systems, including difficult cases leads to better results than simply increasing the size of the database with easy masses. We believe that having a small set of well chosen cases is better than to have a large database filled with redundant cases. Regarding ground truth, we found some incoherence in literature. Annotation is considered to be a subjective, tedious, and extremely time consuming task (Nishikawa, 1998), and it has to be performed by specialists, what can be extremely costly and difficult to find. That is probably the main reason why the currently available databases do not have accurate contours. The importance of having accurate annotations depends on the work at hands. Detection algorithms, for instance, may not need exact contours, while segmentation algorithms have

to be validated by comparing automatic contours with highly detailed manual ones. Nevertheless, we believe that a public database whose objective is to be used in works with several different purposes, should have as accurate ground truth as possible.

Notwithstanding the importance of the digitalized databases, technological advances in image acquisition devices for Radiology led to the development of the Full Field Digital Mammography (FFDM), where the digitalization-related loss of information is absent. Thus, the development of new databases that cover such technological advances is a crucial step to develop future CADs. Besides case selection and annotation requirements, there are some authors (e.g. Oliver, Freixenet et al., 2010) who referred that this issue must also be taken into account when developing new algorithms for CAD improvement. As noted in this review, agreeing with previews works (Oliver et al., 2010), there is no publicly available database made with digital mammograms, all the images are digitized.

We can conclude that researchers need a database that fulfils all the mentioned requirements in order to develop CAD systems. Having in attention the actual state of the art on the breast cancer research, databases with great variability of cases, accurate annotations FFDM images are the natural step in the evolution of mammographic databases.

The requirements discussed at the Biomedical Image Processing Meeting in 1993 need to be reviewed and updated, as new paradigms and ideas to increase algorithms performance are needed in order to improve CAD schemes.

REFERENCES

American College of Radiology, 2003. American College of Radiology Breast Imaging and Data System (BI-RADS) 4th ed.

Eurostat, 2009. Health Statistics Atlas on Mortality in the European Union.

Heath, M. et al., 1998. Current status of the Digital Database for Screening Mammography. In *Digital Mammography*. p. 457–460.

Jiang, L. et al., 2008. Automated Detection of Breast Mass Spiculation Levels and Evaluation of Scheme Performance. *Academic Radiology*, 15(12), p.1534-1544.

Llobet, R., Paredes, R. and Pérez-Cortés, J.C., 2005. Comparison of Feature Extraction Methods for Breast Cancer Detection. In J. S. Marques, N. Pérez de la Blanca, & P. Pina, orgs. *Pattern Recognition and*

Image Analysis. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, p. 495-502.

Matheus, B. R. N. and Schiabel, H., 2010. Online Mammographic Images Database for Development and Comparison of CAD Schemes. *Journal of Digital Imaging*.

Nishikawa, R. M., 1997. Development of a Common Database for Digital Mammography Research.

Nishikawa, R. M., 1998. Mammographic databases. *Breast Disease*, 10(3-4), p.137-150.

Oliver, A., Freixenet, J., et al., 2010. A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, 14, p.87-110.

Oliver, A., Lladó, X., et al., 2010. A Statistical Approach for Breast Density Segmentation. *Journal of Digital Imaging*, 23, p.527-537.

Rangayyan, R. M., Mudigonda, N. and Desautels, J., 2000. Boundary modelling and shape analysis methods for classification of mammographic masses. *Medical and Biological Engineering and Computing*, 38(5), p.487–496.

Rojas Dominguez, A. and Nandi, A., 2007. Detection of masses in mammograms using enhanced multilevel-thresholding segmentation and region selection based on rank. In Proceedings of the 5th IASTED International Conference on Biomedical Engineering, BioMED 2007. p. 370-375.

Song, Enmin et al., 2010. Hybrid Segmentation of Mass in Mammograms Using Template Matching and Dynamic Programming. *Academic Radiology*, 17(11), p.1414-1424.

Suckling, J., 1994. The Mammographic Image Analysis Society Digital Mammogram Database. In *Exerpta Medica. International Congress Series 1069*. York, England, p. 375–378.

Wang, D., Shi, L. and Ann Heng, P., 2009. Automatic detection of breast cancers in mammograms using structured support vector machines. *Neurocomputing*, 72(13-15), p.3296-3302.

World Health Organization, 2009. Fact sheet Nº 297: *Cancer*.

Zheng, Bin et al., 2003. Mammography with Computer-Aided Detection: Reproducibility Assessment - Initial Experience. *Radiology*, 228, p.58-62.

Zheng, Bin et al., 2010. Computer-Aided Detection: The Effect of Training Databases on Detection of Subtle Breast Masses. *Academic radiology*, 17(11), p.1401-1408.