Identification of benign breasts during mammogram screening

Inês Domingues[1]
https://sites.google.com/site/indezitindezit/
Jaime S. Cardoso[1]
http://www.inescporto.pt/~jsc/
Pedro Cardoso[2]
http://www.eacampos.pt/

[1] INESC TEC (formerly INESC Porto) and Faculdade de Engenharia, Universidade do Porto Porto, Portugal

[2] Emílio de Azevedo Campos, S.A. Portugal

Abstract

A "pre-CAD" system that aims to detect normal mammograms instead of abnormal ones is being designed. This pre-CAD system will work as a "first look" by screening-out normal mammograms, letting the experts focus on "harder" cases. The method consists of two blocks. The first block detects dense breasts and automatically flag them for expert review. Non-dense breasts are sent to the second block where benign cases are identified. Tests made on the INBreast database automatize a significant number of cases, without missclassifing a single malign case.

1 Introduction

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among females, accounting for 23% of the total cancer cases and 14% of the cancer deaths [5]. It is estimated that, in the United States, 226,870 women will be diagnosed with and 39,510 women will die of cancer of the breast in 2012 [1]. At present there is no known method to prevent breast cancer but early detection and diagnosis increase the chance of cure. Therefore, screening is recommended by the medical community. Screening can, however, be a very tedious and tiring task for the specialist. We believe that easy cases can be automatically detected, alleviating the human effort and giving the specialist more time to carefully evaluate more ambiguous cases.

2 State of the art

The first work published in this field was by Sahiner *et al* in 1996 [9]. The authors investigated the classification of regions of interest (ROI's) on mammograms as either mass or normal tissue using a convolution neural network (CNN). The input images to the CNN were obtained from the ROI's using two techniques. The first technique employed averaging and subsampling. The second technique employed texture feature extraction methods applied to small subregions inside the ROI. Features computed over different subregions were arranged as texture images, which were subsequently used as CNN inputs. A data set consisting of 168 ROIs containing biopsy-proven masses and 504 ROI's containing normal breast tissue was extracted from 168 mammograms by radiologists experienced in mammography. With the best combination of CNN architecture and texture feature parameters, the area under the test ROC curve reached 0.87, which corresponded to a true-positive ratio of 90% at a false positive ratio of 31%.

Babbs and Delp team have published in this field from 1998 to 2007. In their most recent work, a Support Vector Machine (SVM) based method is introduced [2]. Crossed-distribution feature pairs are identified and mapped into new features that can be separated by a zero-hyperplane of the new axis. The probability density functions of the features of normal and abnormal mammograms are then sampled and the local probability difference functions are estimated to enhance the features. From 1,000 ground-truth-known mammograms, 250 normal and 250 abnormal cases, including spiculated lesions, circumscribed masses or microcalcifications, are used for training. The classification results tested with another 250 normal and 250 abnormal sets show testing performances with 90% sensitivity and 89% specificity.

The other group with strong contributions in the field is led by Elshinawy. They noticed that, in general, there are two major types of mammograms according to their tissue type: fatty and dense mammograms. A "pre-CAD" system was designed for detecting only normal mammograms [3]. Mammograms are first separated into two different categories according to their tissue type, and each category is studied individually. A total of 13 features extracted from the Gray Level Co-occurrence Matrices (GLCM). One-class and two-class SVMs were compared. A majority voting approach based on combining the outcome of the classifier of all of the blocks that correspond to the Mediolateral oblique (MLO) and Craniocaudal (CC) views of the same breast was then used. Results showed that separating the mammograms into two disjoint categories reduced the false negative rate in each of the fatty and dense mammograms while keeping the false positive rate as low as possible.

3 Materials and Methods

In this section both the database and the classification methodology used in this work are detailed.

3.1 Database

INBreast [6] is a recently proposed database that, due to its contour quality, we believe will become a benchmark in the field. It consists of 205 breasts with MLO and CC views (making a total of 410 images). Images area ranges from 8 to 14×10^6 pixels^2 (average area of 11×10^6 pixels^2) with an average height of 2.9×10^3 and an average width of 3.6×10^3. Besides contour information, INBreast also provides ground truth information on breast density (classified according to the ACR standard) and malignancy (in the BI-RADS standard).

Breast density classes were binarized by considering ACR I and II as non-dense and ACR III and IV as dense. In this way, 283 images are non-dense, while the remaining 127 are considered dense.

Malignancy was also binarized by joining BI-RADS classes 1, 2 and 3 (by assuming they correspond to benign findings) and similarly classes 4, 5 and 6 (assuming they are malign). In total, 310 images were considered benign and 100 malign.

3.2 Benign breasts identification

Each breast (MLO and CC views) is fed into system whose flow chart can be seen in Figure 1.

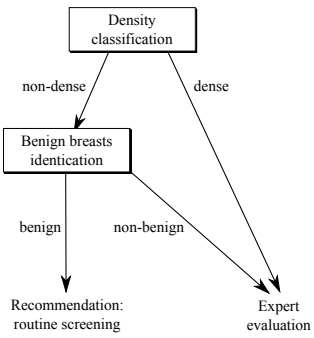


Figure 1: General block diagram.

The density classification block only uses the CC views and outputs a binary result. As density is strongly associated with breast cancer and decreases the sensitivity of automatic systems [10], breasts classified as dense are directly forwarded to a specialist. Otherwise, the breast goes thought the Benign breasts identification block. In this stage, both views are analysed and again a binary output determines if the breast is benign or suspicious. If bening, the general, country specific, routine screening is advised, however if suspicious, no automatic decision is made and the breast image is passed to a specialist. The next sections detail the classification blocks.

3.2.1 Density classification

In this block, only the CC view is used. Intensities are first normalized by histogram stretching to the range $[0,1]$. Then, a histogram analysis (by comparing the average intensity level of the right half with the left half) is made to mirror, if necessary, the image so that the nipple is pointing to the right side. The breast region is detected by using the Otsu threshold [8]. Some features are then extracted from the breast region. After some studies, it was observed that the most relevant characteristics are the 32 bin histogram (from which only 8 bins were kept, namely bins 8, 9, 10, 11, 12, 16, 17, and 24) and a SVD feature as described in [7], making a total of 9 features. These features are normalized to have zero mean and unit variance. A SVM classifier is then created with an RBF kernel and a grid search is performed to select the optimal values for C and γ.

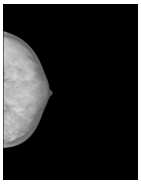
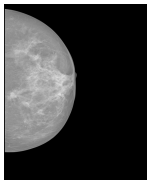
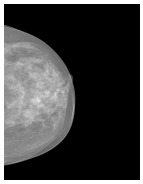
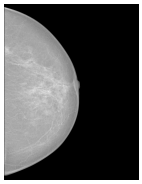
3.2.2 Benign breasts identification

Non-dense breasts are passed to the benign breasts identification block. Pre-processing steps (histogram stretching, mirroring and breast detection) are performed as described above. The image is then divided into blocks of 512×512 (around 40 blocks for each image) and, for each block (with at least 40% of breast tissue), 15 features are extracted. GLCM matrices with distance 1 and angles $0°$, $45°$, $90°$ and $135°$ are constructed. From each one of the 4 matrices, 5 properties are extracted: Contrast, Correlation, Energy, Entropy and Homogeneity. The mean and standard deviation of these 5 properties are then calculated. The last feature is derived from the Local Binary Pattern (LBP) image by computing its contrast value. The average values of each one of the above features over all the image blocks is taken as the feature vector. The final breast feature vector of length 30 corresponds to the concatenation of the features for the two views. Again, the features are normalized and a SVM classifier is created as before.

4 Results

The INBreast database was divided into train and test sets (75%/25%). The density block classifier presents an accuracy of 50% and False Negative rate $FN = 6\%$. Some results can be seen in Table 1.

Table 1: Density classification results

correctly classified		incorrectly classified	
dense	non-dense	dense	non-dense
			

The benign breasts identification classifier was trained with 46 breasts (26 benign and 23 malign) classified as non-dense by the previous block. The performance on the test set of 69 breasts is: True Positive rate $TP = 12\%$, True Negative rate $TN = 28\%$, False Positive rate $FP = 61\%$ and False Negative rate $FN = 0\%$. Some results can be seen in Figures 2, 3 and 4.

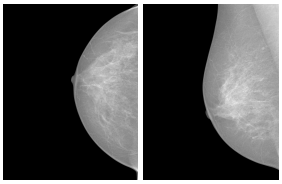


Figure 2: Benign breast recommended for routine screening.

Note that INBreast is a diagnosis database whose cases were purposely selected to include hard to analyse images. In a real screening scenario, most of the mammograms ($> 90\%$) that are interpreted by radiologists are normal cases, and only very few (0.58%) cases are cancerous [4]. We thus believe that in a real scenario, the workload removed from the specialist will be even greater than the one here presented.

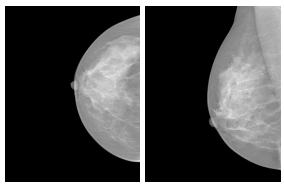


Figure 3: Benign breast sent to be reviewed by an expert.

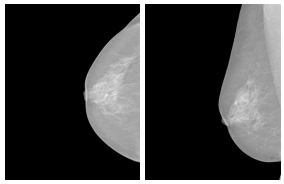


Figure 4: Malign breast sent to be reviewed by an expert.

5 Conclusion

We have presented a fully automatic benign breast identification system. To the best of our knowledge this is the first work that uses an automatic density classification block to filter images before passing them to the Benign breasts identification classifier. Future work consists of improving both classifiers and in testing the system in a real screening environment.

6 Acknowledgment

This work is financed by the ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT - Fundação para a Ciência e Tecnologia (Portuguese Foundation for Science and Technology) within projects FCOMP-01-0124-FEDER-022701, PTDC/SAU-ENB/114951/2009 and SFRH/BD/70713/2010.

References

[1] Cancer facts & figures. *American Cancer Society*, 2012.

[2] W. Chiracharit, Y. Sun, P. Kumhom, K. Chamnongthai, C. F. Babbs, and E. J. Delp. Normal mammogram detection based on local probability difference transforms and support vector machines. *Trans on Information and Systems*, (1):258–270, 2007.

[3] M. Elshinawy, A. Badawy, W. Abdelmageed, and M. Chouikha. Effect of breast density in selecting features for normal mammogram detection. In *ISBI*, pages 141–147, 2011.

[4] M. Y. Elshinawy, A. H. A. Badawyy, W. W. Abdelmageed, and M. F. Chouikha. Normal mammogram detection using density information and texture features. In *SiiM*, 2011.

[5] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman. Global cancer statistics. *CA: A cancer journal for clinicians*, 61: 69–90, 2011.

[6] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso. INbreast: toward a full-field digital mammographic database. *Academic Radiology*, 19(2):236–248, 2012.

[7] J. Oliviera, A. Albuquerque, and T. M. Deserno. Content-based image retrieval applied to BI-RADS tissue classification in screening mammography. *World Journal of Radiology*, 3(1):24–31, 2011.

[8] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[9] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE Trans on Medical Imaging*, 15(5):598–610, 1996.

[10] M. Tortajada, A. Oliver, R. Marti, M. Vilagran, S. Ganau, L. Tortajada, M. Sentis, and J. Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. *Breast Imaging*, pages 561–568, 2012.