

Classifying clusters of microcalcifications in digitized mammograms by bayesian network

ANA CLAUDIA PATROCINIO¹, HOMERO SCHIABEL¹, ROSELI AP. FRANCELIN ROMERO²

¹Departamento de Engenharia Elétrica - EESC - Universidade de São Paulo

²Instituto de Ciências Matemáticas e de Computação

Av. Dr. Carlos Botelho, 1465 - CEP 135630-970 - Cx. Postal 359 - São Carlos – SP
acpatroc@sel.eesc.sc.usp.br, homero@sel.eesc.sc.usp.br, rafrance@icmc.sc.usp.br

Abstract. The purpose of this work is the evaluation and analysis of bayesian network models in order to classify clusters of microcalcification to supply a second opinion to the specialists in the detection of breast diseases by mammography. From regions of interest (ROI) containing clusters of microcalcifications, features based on the moments theory were extracted and used as the input to a bayesian network. By using a free software of bayesian network models construction, some tests were performed in order to build the classifier. The first results of validation have shown 83.17% of correct results.

1 Introduction

Computer-aided diagnosis (CAD) schemes have been developed aimed to help the early detection of breast cancer [1]. As part of a CAD scheme this work presents an investigation of the influences of features extracted from clusters of microcalcifications as well as from medical report information by using bayesian networks generated models.

Bayesian networks are statistics techniques which provide explanation about the inferences and influences among features and classes of a determined problem [2]. Therefore, the technique investigation will aid in obtaining more detailed information to the diagnosis in a CAD scheme.

2 Methodology

Digitized mammographic images were obtained from an image database which is being constructed by our research group, all images with histologic reports - 60 ROIs were segmented by a previous developed procedure [3]. Features to the classifier were calculated using the segmented images, by considering only one cluster by ROI as shown in Figure 1.

- **Features Extraction:** detailed image analysis, pixel to pixel; in this step shape descriptors were extracted, by using mathematical descriptors, as moment (Hu Invariant Moments, second and third order moments and radius gyration [4]); irregularity measure; compactness; area and perimeter, according to the literature [4].
- **Bayesian Network:** by the software *Bayesware* (<http://bayesware.com/>), different bayesian network classifier models could be generated [2], using the extracted features mentioned above in order to verify their behavior and probabilistic influences.

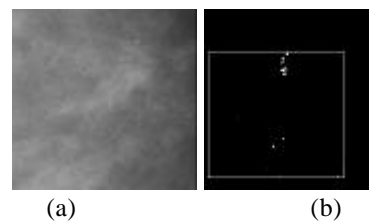


Figure 1: (a) original ROI (b) segmented ROI

3 Results

The results of generated nets models validation correspond to an average of 10 tests made with 6 different database sub-groups. They are divided according to the “cross-validation” default used by the *BayesWare*. With all the features extracted (22), 66.33% of correct results in suspect and non-suspect cases were obtained. The best results were obtained with a net generated with 11 selected features and the average was 83.17% of correct results. Although the database was composed of only 60 cases, the results indicate good performance in the validation test.

References

- [1] Doi, K., et al. “Recent Progress in Development of Computer-Aided Diagnosis (CAD) Schemes in Radiology”. *Medical Image Technology*, v.13, n.6, p. 822-35, Nov. 1995.
- [2] Pearl, J. - *Probabilistic Reasoning in Intelligent Systems*. San Francisco, CA: Morgan Kaufmann, 1988.
- [3] Schiabel, H., et al., “A computerized scheme for detection of clusters of microcalcifications by mammograms image processing” - *World Congress On Medical Physics And Biomedical Engineering*, Nice (France), pp.705, 1997.
- [4] Pratt, W. K. *Digital Image Processing*. A Wiley-Interscience Publication, 1991