

Classification of benign and malignant masses in breast mammograms

A. Šerifović-Trbalić*, A. Trbalić**, D. Demirović*, N. Prljača* and P.C. Cattin***

* Faculty of Electrical Engineering, University of Tuzla, Tuzla, Bosnia and Herzegovina

** Drvodom d.o.o, Tuzla, Bosnia and Herzegovina

*** Medical Image Analysis Center (MIAC), University of Basel, Basel, Switzerland

Introduction

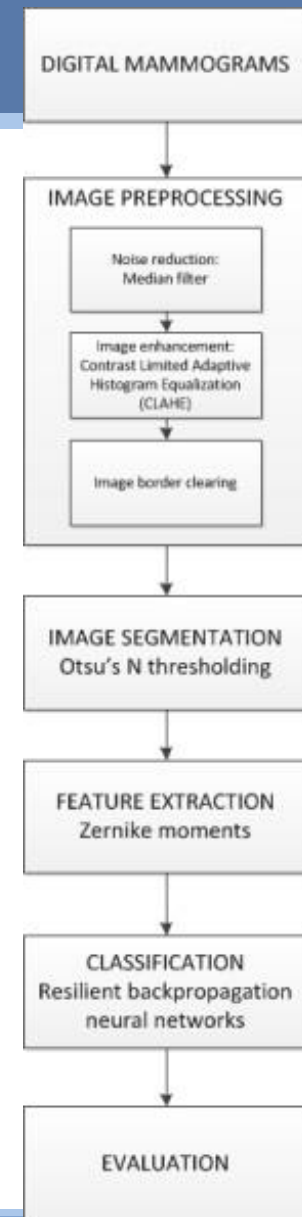
- Breast cancer represents the major cause of death by cancer among women.
- Most breast cancer tumors at earlier stages are benign and not yet malignant and life-threatening.
- The early detection of the breast cancer masses can increase the survival rate for patients.
- Mammography has been one of the most reliable methods for early detection and diagnosis of the breast cancer, typically through detection of characteristic masses and/or microcalcification.

Introduction

- An accurate and efficient computer-aided mammography diagnosis system plays an important role as a second opinion to assist radiologists.
- Finding an accurate and robust computer-aided diagnosis system for the classification of the abnormalities in the mammograms still remains a challenge.
- In the past two decades, a lot of research has been invested in computer-aided analysis and diagnosis of breast mammograms.

Materials and methods

- Proposed method consists of four major stages:
 - Image preprocessing - used to suppress noise and improve the contrast of the image;
 - Image segmentation - used to identify the masses in mammograms, i.e. to detect the Regions of Interest (ROI);
 - Feature extraction;
 - Classifications of masses.

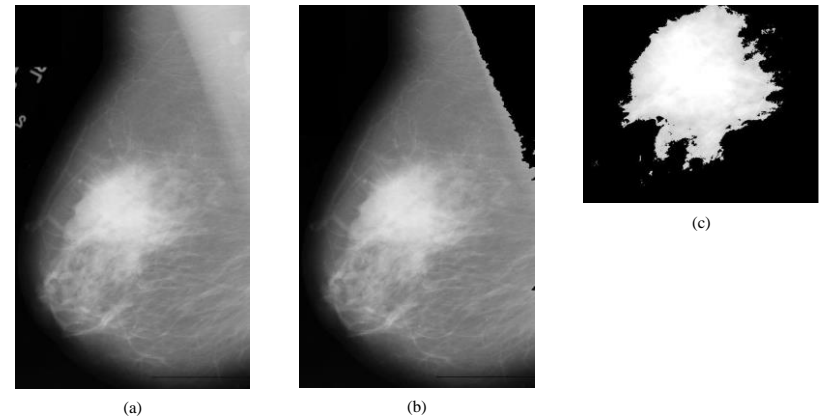


Materials and methods – image preprocessing

- It is used to improve the image quality.
- This stage involves:
 - Image filtering using a two-dimensional (2D) median filter with a neighborhood to reduce a digitization noises.
 - Image enhancement using *contrast-limited adaptive histogram equalization (CLAHE)*.
 - The automatic thresholding and image border cleaning to suppress the light structures connected to the image border and thus to identify the foreground in the image.

Materials and methods – image segmentation

- A multiclass Otsu's thresholding has been used to find the Region of Interest (ROI).
- It is:
 - a nonparametric and unsupervised method of automatic threshold selection for the image segmentation;
 - a simple procedure that utilizes the zeroth- and first-order moments of the graylevel histograms and it maximizes the between class variances, while minimizing the within class variances.



Different segmentation steps. (a) Original Image. (b) Image after marking foreground. (c) Obtained ROI.

Materials and methods – feature extraction

- Zernike moments represent a set of descriptors obtained using complex kernel functions based on Zernike polynomials.
- The computation of Zernike moments from an input image includes three steps:
 - computation of radial polynomials,
 - computation of Zernike basis functions,
 - computation of Zernike moments by projecting the image onto the Zernike basis functions.
- A group of 32 low-order moments with different orders and iterations which satisfy the following conditions have been used:

$$Features = \{Z_{n,m}\} \forall \begin{cases} 3 \leq n \leq 10 \\ |m| \leq n \\ n - |m| = 2k \\ k \in \mathbb{N} \end{cases}$$

Materials and methods – classification

- Classification is conducted using resilient backpropagation neural networks.
- These networks have following characteristics:
 - During training, the error must be propagated from the output layer back to hidden layer in order to perform the learning of the input-to-hidden weights;
 - Only the sign of the derivative is used to determine the direction of the weight update, while the size of the weight change is determined by a separate update parameter;
 - They are generally much faster than the standard steepest descent algorithm and it requires only a modest increase in memory consumptions

Experiments and results

- The digital mammograms used for experimentation were taken from the Mammography Image Analysis Society (MIAS) database.
- The input images were split in three sets intended for training, validation and testing of the neural network.
- The number of hidden layers and their nodes are being changed until the best network topology, which yields the best accuracy, is found.
- The final neural network consists of:
 - one input (32 nodes, sigmoid function),
 - one hidden (32 nodes, sigmoid function) and
 - one output layer (linear activation function).

Experiments and results

- Through visual inspection of the segmented masses, it was noticed that:
 - the best result is for round-shape masses that have well-defined edges,
 - most false detections occurred for:
 - non-circular cases with calcifications, which consist of very small spots in mammograms,
 - and speculated masses, which are not segmented quite well.

Experiments and results

- As benchmark functions of proposed system the following terms were used:

- the area under ROC (Receiver operating characteristic) curve;

- $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$

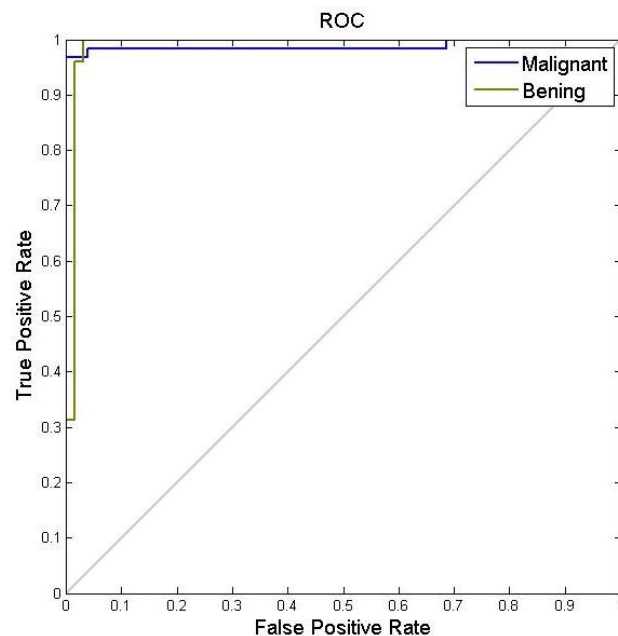
- $Sensitivity = TPR = \frac{TP}{TP + FN}$

- $Specificity = \frac{TN}{TN + FP}$

where FP, FN, TP and TN denote false-positive, false-negative, true-positive and true-negative answers, respectively.

Experiments and results

The system yielded a malignant mass versus benign mass classification accuracy represented as an area under ROC curve of 0.8920.



Comparison with state-of-the-art methods based on sensitivity, specificity and accuracy

Author's name	Sensitivity	Specificity	Accuracy (%)
Menut et al. [4]	-	---	76.0
Hussain [9]	-	---	99.0
Gorgel et al. [16]	0.9470	0.7140	84.8
Islam et al. [15]	0.9091	0.8371	---
Proposed method	0.9216	0.9608	90.4

Conclusion

- A novel fully autonomous computer-aided diagnosis system has been presented for mass detection and classification in breast mammography images.
- Input ROIs are obtained from the digital mammograms using image preprocessing and Otsu's N thresholding.
- After the first stage, a group of 32 Zernike moments with different orders and iterations are extracted from the ROIs.
- These moments have been applied to the resilient backpropagation neural network classifier.
- Experimental results show that the proposed technique is able to classify efficiently malignant and benign mass types of abnormalities in digital mammograms.

References

- O. Menut, R. Rangayyan, and J. Desautels, “Classification of breast tumors via parabolic modeling of their contours,” in Proceedings 1997 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Victoria, BC, Canada, pp. 1002 – 1005, 1997
- M. Hussain, “False Positive Reduction using Gabor Feature Subset Selection”, 2013 International Conference on Information Science and Application (ICISA), 2013
- P. Gorgel, A. Sertbas, N. Kilic, O.N. Ucan and O. Osman, “Mammographic mass classification using wavelet based support vector machine”. J. Elect. Elect. Eng., 9: 867-875, 2009
- M.J. Islam, M. Ahmad and M.A. Sid-Ahmed, “An efficient automatic mass classification method in digitized mammograms using artificial neural network”. Int. J. Artif. Intell. Applic., 1: 6-6, 2010
- A. Tahmasbi, F. Saki, S. B. Shokouhi, “Classification of benign and malignant masses based on Zernike moments,” J. Computers in Biology and Medicine, vol. 41, no. 8, pp. 726-735, 2011