Computer aided monitoring of breast abnormalities in X-ray mammograms

SELVAN, Arul <http://orcid.org/0000-0001-9222-5538>, SAATCHI, Reza <http://orcid.org/0000-0002-2266-0187> and FERRIS, Christine

Computer Aided Monitoring of Breast Abnormalities in X-ray Mammograms

Arul N. Selvan[1]
http://www.shu.ac.uk/research/meri/mmvl/research/projects/medical/index.html
Reza Saatchi[1]
Christine M. Ferris[2]

[1] Materials & Engineering Research Inst.
Sheffield Hallam University,
Sheffield. UK.

[2] Faculty of Health & Wellbeing
Sheffield Hallam University,
Sheffield. UK.

Abstract

X-ray mammography is regarded as the most effective tool for the detection and diagnosis of breast cancer, but the interpretation of mammograms is a difficult and error-prone task. Computer-aided detection (CADe) systems address the problem that radiologists often miss signs of cancers that are retrospectively visible in mammograms. Furthermore, computer-aided diagnosis (CADx) systems assist the radiologist in the classification of mammographic lesions as benign or malignant [1].

This paper details a novel alternative system namely computer-aided monitoring (CAM) system. The designed CAM system can be used to objectively measure the properties of a suspected abnormal area in a mammogram. Thus it can be used to assist the clinician to objectively monitor the abnormality. For instance its response to treatment and consequently its prognosis. The designed CAM system is implemented using the Hierarchical Clustering based Segmentation (HCS) [2] [3] [4] process. Brief description of the implementation of this CAM system is as follows : Using the approximate location and size of the abnormality, obtained from the user, the HCS process automatically identifies the more appropriate boundaries of the different regions within a region of interest (ROI), centred at the approximate location. From the set of, HCS process segmented, regions the user identifies the regions which most likely represent the abnormality and the healthy areas. Subsequently the CAM system compares the characteristics of the user identified abnormal region with that of the healthy region; to differentiate malignant from benign abnormality. In processing sixteen mammograms from mini-MIAS [5], the designed CAM system demonstrated a success rate of 100% in differentiating malignant from benign abnormalities.

1 Introduction

The introduction of systems for automated reading in mammography has been proposed to improve the sensitivity [computer-aided detection (CADe) systems] and, more recently, the specificity [computer-aided diagnosis (CADx) systems] of the test. Only CADe systems have been approved by the U.S. Food and Drug Administration (FDA) [6]. The first CADe tools were approved by the U.S. FDA for clinical use in 1998. Several commercial and non-commercial CAD systems have since become available [7].

CAD systems are trained on a database of mammograms of selected patients before they are used in clinical practice [6]. The method for assessing CAD performance during the

Doi:06/06/2011 #46#

training period is to plot the free-response receiver operating characteristic (FROC). This represents the sensitivity of the system in detecting cancer as a function of the averaged number of false-positive. The best performing CAD systems will have a low false-positive rate as well as high sensitivity [7].

During actual usage CAD system use thresholding algorithms to identify as many true signals and as few false signals as possible. Signal data are separated from the background in a process known as segmentation. The signals are then subjected to a probabilistic analysis to assess the likelihood that the structure on the image contains malignancy-induced abnormalities [7]. The final result is a CAD prompt, if the probability of cancer being present is sufficiently high [8].

The exact impact of CAD on radiologists' sensitivity has not been determined. One retrospective study of screening mammograms, involving two different CAD systems, showed that CAD improved the detection of cancers that had been overlooked by a single radiologist. This reduction in false negatives led to an increase in sensitivity from 71.2% for the single radiologist to 84.8% and 80.3% for the two CAD systems [9]. Another study, focusing on the detection of previously missed cancers, demonstrated a sensitivity of 51.5% for CAD, 62.5% for the radiologist, and 86.2% for the radiologist and CAD combined [10].

The major limitations of the current CAD(e/x) systems are two fold, firstly the way the systems are trained and secondly the way the systems are used. The limitation of the training process is that the training samples might have had features associated with symptomatic lesions. In the actual usage environment the abnormalities might have less obvious mammographic features than symptomatic lesions. Consequently this might lead to false-negative rate of up to 25% [6]. The limitation during the usage is that the thresholding, learnt by the CADe system during the training, may not be appropriate for the actual image data under consideration. Consequently most of the CAD prompts are false-positive calls which leads to needless breast biopsies [8].

In the HCS based CAM system, designed in this study, the above issues were addressed as follows :

- The designed CAM system does not need any prior training.
- The designed system does not use a fixed thresholding to differentiate the abnormalities from the healthy tissue, rather it uses an adaptive measure adapted to the actual mammogram data being analysed.

The rest of the paper is organized as follows. Firstly the operation of HCS process is outlined briefly. Secondly the methodology adopted to implement the designed CAM system, using the HCS process, is discussed. Then the performance of the designed CAM system in differentiating benign from malignant abnormalities is discussed. Finally the possibilities of using the designed CAM systems to aid radiologists is discussed.

2 Hierarchical Clustering based Segmentation

The HCS process partitions an image into its constituent regions for a hierarchical levels of allowable dissimilarity between the different regions. As the allowable dissimilarity is incremented; at any particular level in the hierarchy, the HCS process clusters together all the pixels and/or regions that have dissimilarity among them less than or equal to the dissimilarity allowed for that level. At each level the HCS process yields an optimized segmentation output related to the dissimilarity allowed for that level. The algorithmic diagram, shown in Figure 1, illustrates the overall operation of HCS [3].

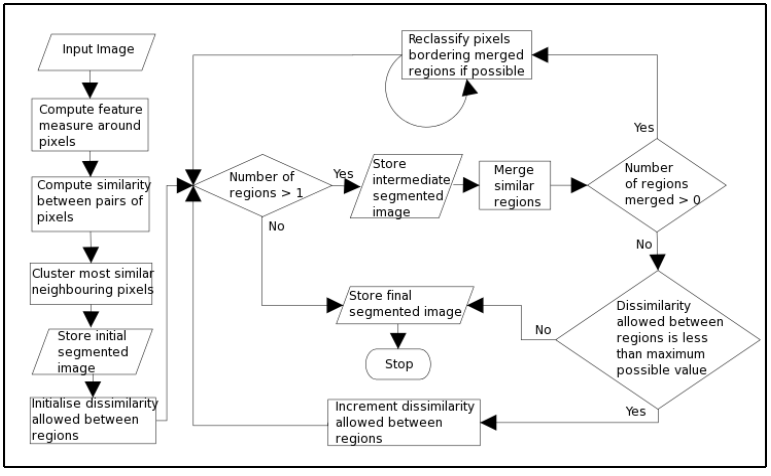


Figure 1: Flow chart illustrating the working of the HCS process [3].

3 Design and Implementation of the CAM System

To start with it was hypothesised that in the X-ray mammogram, the benign abnormalities will have image properties closer to the healthy tissue and malignant abnormalities will be markedly dissimilar from the surrounding healthy tissue. Hence the measured dissimilarity, between the region comprising the abnormality and the region comprising the healthy tissue, will indicate whether the abnormality is benign or not. To measure the dissimilarity between the abnormality and the healthy region the basic requirements are :

- appropriate delineation of the boundary of the abnormality
- identifying the region of the image comprising the healthy part.

The process, how the above requirements are addressed, is explained below by an example

Figure 2 shows a X-ray mammogram (mdb102) of a dense glandular breast having a malignant asymmetry class of abnormality. Making use of the information provided in the mini-MIAS database the approximate boundary of the abnormality (Green circle Figure 2 a), was located. The HCS process was applied within a ROI centred on the abnormality. Inspecting the HCS process output the user selected the region corresponding to the abnormality (Red Figures 2 b, c and d). The area within the approximate circular boundary, other than the abnormality, was selected as healthy, (Green Figure 2 d). Inspecting the HCS process output the user also selected a location, within the abnormal area, which was considered as the core of the abnormality (Yellow Figure 2 d and e).

To estimate the dissimilarity between the abnormal and the healthy regions, the HCS process was applied only to the pixel locations within the abnormality (Red Figure 2 d and e) and the healthy (Green Figure 2 d and e) areas of the image. As the HCS process goes about merging similar regions within the abnormality and the healthy areas, the maximum average dissimilarity, measured between the cluster having the user tagged location (within the abnormality) and the clusters within the healthy region, is estimated. The heuristic used for differentiating malignant from benign abnormality is, if the value of the above estimated measure is less than fifty percent then the abnormality is benign else malignant.

Graph shown in Figure 3 (a) demonstrates how the above measure and the criteria is able to classify the abnormality, under consideration, as malignant. Similarly the graph shown in Figure 3 (b) demonstrates how a benign abnormality is correctly classified.

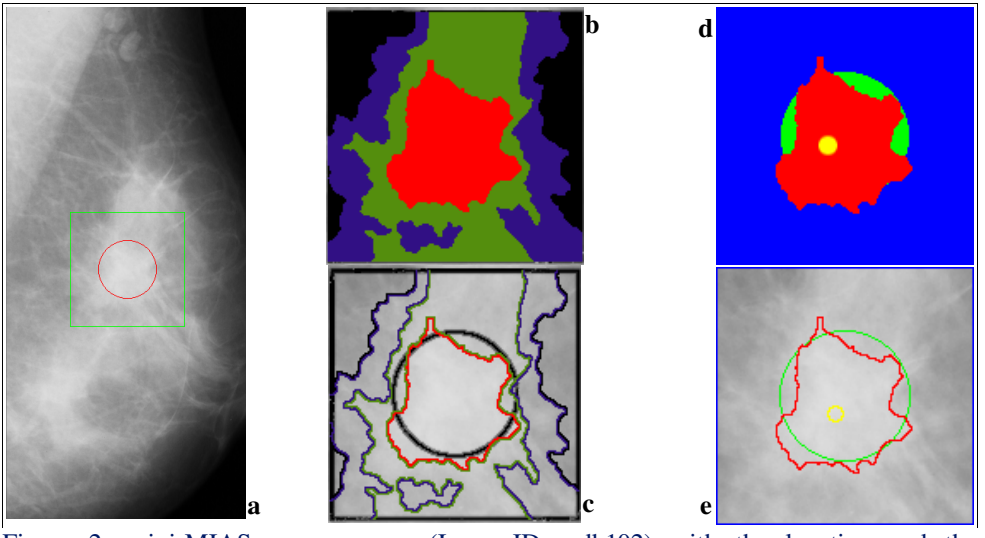
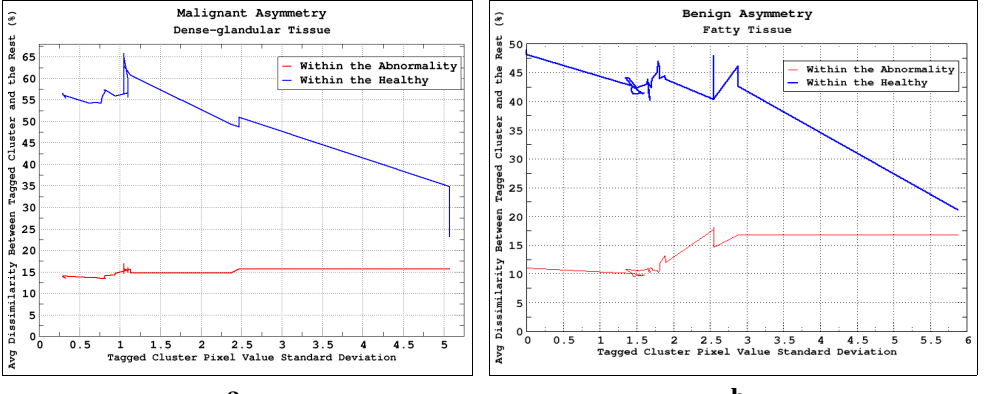


Figure 2: mini-MIAS mammogram (Image-ID mdb102) with the location and the approximate boundary of the abnormality circled in Red by the user (a).HCS process intermediate segmentation of four regions and their boundaries (b and c). Regions, and their boundaries, identified by the user as healthy (Green) and abnormal (Red) (d and e).



a b

Figure 3: For mini-MIAS Mammogram images mdb102 and mdb097 graphs (a and b) showing the dissimilarity between the cluster having the user tagged location and other clusters belonging to the abnormality area and the healthy area.

4. Performance of the Designed CAM System

For sixteen of the mini-MIAS [5] cases the designed system successfully differentiated benign (B) from malignant (M) abnormalities. Table 1 lists the quantitative measures used.

5. Conclusion

The designed CAM system, which does not need any prior training, can help the clinician to visualise and quantitatively measure dissimilarities between healthy and abnormal areas in X-ray mammograms. Work is in progress to evaluate how the designed CAM system could augment the diagnostic capabilities of clinicians.

Table 1 : Overall performance of the CAM system.

Image ID	Breast Tissue Type	Class of Abnormality	Severity of Abnormality	Dissimilarity Measure	CAM based Classification
mdb002	Fatty Glandular	Circumscribed	Benign	16.08%	B (True -ve)
mdb005	Fatty	Circumscribed	Benign	30.50%	B (True -ve)
mdb012	Fatty	Circumscribed	Benign	40.50%	B (True -ve)
mdb028	Fatty	Circumscribed	Malignant	81.96%	M (True +ve)
mdb075	Fatty	Asymmetry	Malignant	52.81%	M (True +ve)
mdb083	Fatty Glandular	Asymmetry	Benign	43.77 %	B (True -ve)
mdb090	Fatty Glandular	Asymmetry	Malignant	67.37 %	M (True +ve)
mdb092	Fatty	Asymmetry	Malignant	53.86%	M (True +ve)
mdb095	Fatty	Asymmetry	Malignant	61.48 %	M (True +ve)
mdb097	Fatty	Asymmetry	Benign	48.79 %	B (True -ve)
mdb099	Dense Glandular	Asymmetry	Benign	48.18 %	B (True -ve)
mdb102	Dense Glandular	Asymmetry	Malignant	65.72 %	M (True +ve)
mdb104	Dense Glandular	Asymmetry	Benign	32.43%	B (True -ve)
mdb110	Dense Glandular	Asymmetry	Malignant	77.88%	M (True +ve)
mdb141	Fatty	Circumscribed	Malignant	82.64%	M (True +ve)
mdb270	Fatty Glandular	Circumscribed	Malignant	53.23%	M (True +ve)

References

[1] A. L. Elter M, Horsch A., "CADx of mammographic masses and clustered microcalcifications: a review" *.Medical Physics* 2009 Jun;36(6):2052-68. 2009..

[2] Selvan, A.N.; Saatchi, R.; Amavasai, B.P.; Travis, J.R; "A Dissimilarity Visualisation System for CT: Pilot Study" T*he International Conference on Computer as a Tool. EUROCON* 2005. Volume 2, Pages:1007 - 1010, 2005.

[3] Selvan, A. N. "Highlighting Dissimilarity in Medical Images Using Hierarchical Clustering Based Segmentation (HCS)" *M. Phil. dissertation,* Faculty of Arts Computing Engineering and Sciences Sheffield Hallam Univ., Sheffield, UK, 2007.

[4] Selvan, A.N.; Saatchi, R.; Ferris, C.M.; "Improving medical image perception by hierarchical clustering based segmentation" *9th International Conference on Information Technology and Application in Biomedicine,* ITAB 2009. Pages:1 - 6, 2009.

[5] Suckling, J. Parker, J. Dance, DR. Astley, S. Hutt, I. Boggis, C. Ricketts, I. Stamatakis, E. Cerneaz, N. Kok, SL. Taylor, P. Betal, D. and Savage, J. "The Mammographic Image Analysis Society Digital Mammogram Database" *Exerpta Medica. International Congress Series* 1069 pp375-378, 1994.

[6] Bazzocchi M, Mazzarella F, Del Frate C, Girometti R, Zuiani C."CAD systems for mammography: a real opportunity? A review of the literature". *Radiology Medicine*;112(3):329-353. 2007

[7] Nishikawa RM. "Current status and future directions of computer-aided diagnosis in mammography". *Computer Med Imaging Graph*;31(4-5):224-235. 2007

[8] Rossand G., Lorenzo C., Massimo B. and Chiara Z., "CAD comes under scrutiny in breast screening debate" *Diagnostic Imaging Europe*. Vol. 26 No. 7 October 29, 2010.

[9] Ciatto S, Ambrogetti D, Bonardi R, Brancato B, Catarzi S, Risso G, Rosselli Del Turco M. "Comparison of two commercial systems for computer-assisted detection (CAD) as an aid to interpreting screening mammograms." *Radiology Medicine*;107(5-6):480-488. 2004.

[10] Ciatto S, Ambrogetti D, Collini G, Cruciani A, Ercolini E, Risso G, Rosselli Del Turco M. "Computer-aided detection (CAD) of cancers detected on double reading by one reader only." Breast;15(4):528-532. 2006