Detection of Masses in Mammograms Using Texture Features

Keir Bovis and Sameer Singh
PANN Research, Department of Computer Science, University of Exeter, Exeter, UK

Abstract

The aim of this study is to detect masses in mammograms on the basis of textural features. Suspicious regions are identified following the bilateral image subtraction of left and right breast image pairs. The study uses the nipple as a common rotational point thereby facilitating an alignment with the highest correlation prior to subtraction. Within this study, 144 breast images from the MIAS database are considered. Five co-occurrence matrices are constructed at four different distances for each suspicious region. Twelve texture features defined by Haralick et. al. [5], angular second moment, correlation, contrast, entropy, inverse difference moment, sum average, sum entropy, sum variance, difference entropy, difference variance and two information measures of correlation. Two further features defined by Chan et. al [2], inertia and difference average, are also computed giving a total of fourteen texture measures. Following classification of six principal components calculated for the extracted features using an ANN (Artificial Neural Network) and 10-fold cross-validation, an average recognition rate of 77% was achieved. Using Receiver Operating Characteristic (ROC) analysis, the overall sensitivity of the technique measured by the value of Az, was found to be 0.74.

1. Introduction

X-ray mammography is the most common technique used by radiologists in the screening and diagnosis of breast cancer in women. Although it is seen as the best examination technique for the early detection of breast cancer reducing mortality rates by up to 25%, their interpretation requires skill and experience by a trained radiologist. The aim of this study is to analyse digitised mammograms by applying computer image processing techniques to enhance x-ray images and then subsequently extract features from suspicious regions characterising the underlying texture of the breast regions. These features can then be passed to a classifier for discrimination for different regions of interest to test whether they are masses or non-masses.

In order to identify suspicious regions, the clinical observation of focal asymmetric densities within the internal structure of the breast is implemented [7]. Implementing this method requires left and right breast images to be aligned around a common reference point, then subtracted bilaterally (see Figure 1). On the basis of several past studies, we follow the fundamental assumption that asymmetries represent a region of interest that could be a mass.

Several previous studies have applied the bilateral subtraction technique to mammographic image pairs. Yin et al.[13] investigated mammographic asymmetries for the identification of mass lesions. In their research the authors compared the performance of a non-linear bilateral subtraction technique with grey level thresholding. For bilateral subtraction, left and right breast images were aligned, thresholded and following subtraction subsequently thresholded again. For the comparative method, the authors implemented a method of local grey scale thresholding a single image divided into 100 x 100 pixel blocks. For bilateral subtraction using ROC analysis the performance of the technique, measured by the area under the ROC curve Az, was found at Az=0.530 and for local grey level thresholding the Az=0.385. Ideally, the area under the curve Az should be as close to 1 as possible for an accurate technique. Mendez et al. [9], investigated a computerised method to automatically detect malignant masses on digital mammograms based on bilateral subtraction to identify asymmetries between left and right images. The nipple was detected to align the images, and alignment of image pairs achieved by translation using the nipple location and rotation against a correlation coefficient. To reduce false positives, texture tests were applied to each suspicious region, coarseness and contrast. Using ROC analysis, the value of Az=0.667 was achieved. In their study, Chan et. al. [3], studied the effectiveness of using texture features derived from Spatial Grey Level Dependency (SGLD) matrices for classification of masses and normal breast tissue on mammograms. All images were digitised at 4,096 grey levels and a region of interest confirmed by biopsy comprising 256 x 256 pixels. The following texture features were extracted from the SGLD matrices: correlation, entropy, angle of second moment, inertia, inverse difference moment, sum average, sum entropy,

0-7695-0750-6/00 $10.00 © 2000 IEEE

difference entropy. The accuracy was evaluated using ROC analysis and the maximum value of the area under the ROC curve obtained was $Az = 0.823$.

2. Methodology

The MiniMIAS database [11] containing left and right breast images for a total of 161 patients is used in this study. All images containing spiculated/circumscribed masses and a selection of normal types are considered. Supplied ground truth data for each image includes tissue type and regions of interest encompassing the abnormalities given in terms of a centroid and region radius. All images are digitised at a resolution of 1024 X 1024 pixels and at 8-bit grey scale level. Forty images contain abnormalities (circumscribed masses n=21, spiculated masses n=19) and 104 images are classed as normal. In the analysis of results within the study, we use the following definitions:

True Positive (TP): lesions called cancer and prove to be cancer
False Positive (FP): lesions called cancer that prove to be benign
False Nagative (FN): lesions that are called negative or benign and prove to be cancer
True Negative (TN): lesions that are called negative and prove to be negative

On the basis of this terminology, we can evaluate the performance of our technique by calculating True Positive Fraction (TPF) and False Positive Fraction (FPF). These are defined as:

$$TPF = \frac{TP}{TP+FN} \qquad FPF = \frac{FP}{FP+TN}$$

The study involved five different phases:

(i) Location of a Common Reference Point

In order to align left and right breast image pairs prior to bilateral subtraction a common reference, the spatial position of the nipple, is located.

(ii) Alignment and Bilateral Subtraction of Left and Right Breast Images

Once the nipple has been located in each breast image of a matching pair, the observed image is translated such that the nipple locations of both breast images are aligned. Taking different rotations of the observed image and determining a correlation measure against the reference image, the left original breast image, the best alignment can be determined. Initially the observed image is subjected to large degrees of rotation, following which the best range is located and the process repeated with smaller incremental rotations until the best correlation is found. Using aligned left and right breast image pairs, two images are generated by bilaterally subtracting one image from the other. One is a positive image and details differences that occur in the left breast image and not in the right, and a negative image that details differences that occur in the right breast image and not in the left.

(iii) Reduction of False Positives

Each difference image generated from bilateral subtraction, following contrast enhancement using a histogram stretching technique [12], will contain regions that are true positives, and those that are false positives. Difference images are segmented using a region splitting technique, the complete set of regions being modelled using a quad-tree data structure representation [6]. Once regions have been identified, the aim is to subsequently remove as many false positive regions on the basis of region characteristics size, shape [10], difference in homogeneity [12] and entropy [4].

(iv) Feature Extraction

Using the quad-tree region model generated in the previous phase for all remaining suspicious regions within a difference image, five co-occurrence matrices are constructed in four different spatial orientations, horizontal, left diagonal, vertical and right diagonal, (0°, 45°, 90° and 135°, respectively). A fifth matrix is also constructed as the mean of the first four. Each co-occurrence matrix reflects the joint probability of a pixel pair at a given orientation and distance. From each of the five normalised co-occurrence matrix, the 14 textures features are extracted giving a total of 70 texture features per region. Each region is labelled mass or non-mass using the supplied MIAS ground-truth data. These 70 features are extracted from the co-occurrence matrices constructed at four pixel distances (d=1, 3, 6, 9) giving four separate feature vectors.

(v) Classification

In order to determine the discriminating effectiveness of texture features extracted from co-occurrence matrices constructed at different distances, a classification is initially performed using each normalised feature vector and linear discriminant analysis. The feature vector giving the best true positive fraction was selected for subsequent classification using an Artificial Neural Network (ANN) developed using the Stuttgart Neural Network Software

(SNNS) package. For classification using an ANN, the selected feature vector is normalised and principal component analysis performed to reduce the dimensionality of the data. Subsequently, samples for classes are interleaved and training/test files can be generated using a 10-fold cross validation method [1]. In each fold, 90% of data is used for training the classifier and the remaining 10% is used for testing. The process continues 10 times taking different training and test partitions, each time with a different training and test set. Each of the six principal components is mapped onto an individual input node and linked by a varying number of hidden nodes to two output nodes (mass, non-mass). For training, a back-propagation with momentum learning function is used (learning rate η= 0.01, momentum μ= 0.9). In order to prevent over-generalisation by controlling the complexity of the ANN, the technique of structural stabilisation [1] is used. Using this technique, the TPF for a range of ANNs with differing number of hidden nodes is compared. The ANN model giving the highest TPF is selected. ROC analysis is finally performed [8] on the test results from each fold and the value of Az computed.

3. Results

Result in Table 1 show the TPF and FPF obtained from performing linear discriminant analysis on the all features extracted from co-occurrence matrices constructed at four different distances. Cross-validation is done for all cases in the analysis. From the 144 input images, the dataset comprised 820 non-masses and 48 actual masses.

The co-occurrence matrices constructed using a distance *(d=3)* gave the largest TPF value thus being the most sensitive to detecting true positives. This distance represents a trade-off between the construction of the complete joint probability histogram for each region, needed to describe its texture, limited by a region's size, coupled with a minimum value of d required to characterise the underlying texture associated with pixel pairs. For each test set tested on a trained ANN using 10-fold cross-validation, the associated feature vector (where d=3) is subsequently used for classification. Table 2 shows the ANN structure giving the recognition rate together with the TPF, FPF and the value of Az for the associated computed ROC curve. Note that in the generation of fold-10, as indicated by a asterisk in Table 2, only non-masses were presented within the test set, thus no TPF or FPF score is given.

4. Conclusions

A key aim of this study was the implementation of a technique that could automatically detect suspicious regions. The bilateral subtraction technique of left and right breast image pairs is seen as an effective method of identifying such suspicious regions based on focal asymmetric densities. The reduction of false-positive regions based on size, shape and uniformity measures ensured that the ultimate FPF could be kept low without affecting the overall sensitivity.

The use of statistical measures of texture as features, their subsequent extraction and classification, can be seen from the results in Table 2 to be highly discriminant in differentiating between masses and non-masses with a very high TPF with a low TPF. In comparison with previous studies the overall performance of the technique as indicated by the value of Az=0.74, which improves on similar previous studies where the detection masses has been the objective. Future work will attempt to improve on these results and include different texture and shape measures applied to other mammography databases.

5. References

[1] C. Bishop, 1998, Neural Networks for Pattern Recognition, Clarendon Press.

[2] H. Chan, B. Sahiner, N. Petrick, M. Helvie, K.L. Lam, D. Adler, M. Goodsitt, 1997, Computerised Classification of Malignant and Benign Microcalcifications on Mammograms: Texture Analysis Using and Artificial Neural Network, *Physical Medical Biology*, vol. 42, pp. 549 - 567.

[3] H. Chan, D. Wei, N. Helvie, B. Sahiner, D. Adler, M. Goodsitt, N. Petrick, 1995, Computer-aided Classification of Mammographic Masses and Normal Tissue: Linear Discriminant Analysis in Texture Feature Space, *Physical Medical Biology*, vol. 40, pp. 857 - 876.

[4] R. Gonzalez, R. Woods, 1993, Digital Image Processing, Addison Wesley.

[5] R. Haralick, K. Shanmugam, I. Dinstein, 1973, Textural Features for Image Classification, *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610 - 621.

[6] A. Jain, 1995, Fundamentals of Digital Image Processing, Prentice-Hall of India.

[7] D. Kopans, 1998, Breast Imaging, Lippincott-Raven.

[8] D. McNicol, 1972, A Primer of Signal Detection Theory, Australasian Publishing Company.

[9] A. Mendez, P. Tahoces, M. Lado, M. Souto, J. Vidal, 1998, Computer-aided Diagnosis: Automatic Detection of Malignant Masses in Digitised Mammograms, *Medical Physics*, vol. 25, no. 6, pp 957-964.

[10] N. Petrick, H. Chan, D. Wei, B. Sahiner, M. Helvie, D. Adler, 1996, Automated Detection of Breast Masses on Mammograms Using Adaptive Contrast Enhancement and Texture Classification, *Medical Physics*, vol. 23, no. 10, pp.

1685-1696.
[11] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal and J. Savage. "The mammographic images analysis society digital mammogram database." Exerpta Medica. International Congress Series 1069: 375-378 (1994). (mias@sv1.smb.man.ac.uk).
[12] S. Umbaugh, 1998, Computer Vision and Image Processing: A Practical Approach Using CVIPtools, Prentice Hall.
[13] F. Yin, M. Giger, C. Vyborny, K. Doi, R. Schmidt, 1993, Comparison of Bilateral-Subtraction and Single-Image Processing Techniques in the Computerised Detection of Mammographic Masses, *Investigative Radiology*, vol. 28, no. 6, pp. 473 - 481.

Table 1. Results from discriminant analysis

Distance (d)	*TPF*	*FPF*
1	0.563	0.267
3	0.625	0.266
6	0.458	0.223
9	0.583	0.253

Table 2. Results from each test fold using 10-fold cross-validation

Fold No.	*Hidden Nodes*	*Recog. Rate*	*TPF*	*FPF*	*Az*
1	1	0.70	0.80	0.31	0.743
2	5	0.78	0.40	0.19	0.617
3	3	0.70	0.60	0.30	0.601
4	3	0.77	0.80	0.23	0.789
5	1	0.75	1.00	0.27	0.836
6	2	0.68	0.60	0.31	0.674
7	4	0.67	0.80	0.34	0.736
8	4	0.82	0.60	0.16	0.737
9	5	0.83	1.00	0.18	0.898
10 (*)	1	1.00	-	-	-
Mean		**0.77**	**0.73**	**0.25**	**0.74**

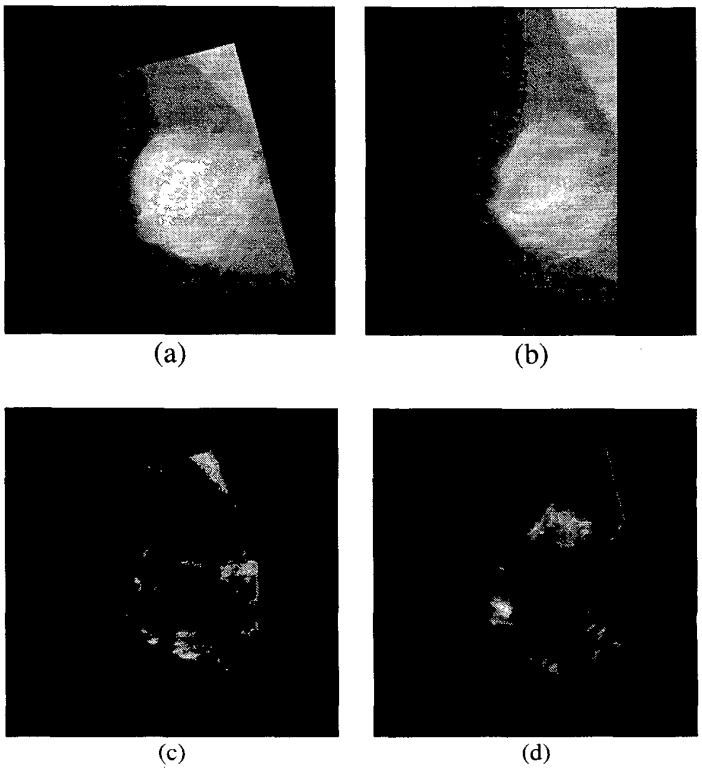

Figure 1. (a) is left breast and (b) is the aligned and rotated right breast image for a patient with benign circumscribed masses. The bottom images represent differences between (a) and (b). Image (c) is the negative difference image and (d) is the positive difference image.