Structural image texture and early detection of breast cancer

Shijia Lu and Murk J. Bottema
School of Informatics and Engineering
Flinders University
PO Box 2100, Adelaide SA 5001, Australia
and
Cooperative Research Centre for Sensor Signal and Information Processing
SPRI Building, Mawson Lakes Blvd, Mawson Lakes, SA 4095, Australia
lu@infoeng.flinders.edu.au
murkb@infoeng.flinders.edu.au

Abstract

Structural texture measures are used to address three aspects of early detection of breast cancer in screening mammograms: detection of microcalcification, detection and classification of clustered microcalcification as benign or malignant, and the detection of invasive lobular carcinoma. The use of structural texture features replaces the task of initial detection of complex and poorly modelled image structures such as masses or clustered microcalcifications by initial detection of primitive image structures.

Receiver operating characteristic (ROC) analysis yields high performance scores for detection of microcalcification ($A_z = 0.945$) and for classification of clusters as benign or malignant ($A_z = 0.858$). In the case of invasive lobular carcinoma, cancer was detected in half of the images with no false positive detections. In these cases, no evidence of cancer could be found by visual inspection of the mammogram, thus demonstrating the potential of structural texture measures to encompass diagnostically useful information not accessible to the human expert observer.

1. Introduction

In the last fifteen years, a plethora of papers have appeared in the literature on the use of computer algorithms for improving early detection of breast cancer in screening mammograms. A number of review papers have also appeared [2],[4], [12]. Currently, nearly all screening mammograms are film mammograms and, in order to implement computer algorithms, images must be digitised. Although commercial systems for computer assisted reading exist that accept film mammograms as input, wide spread use of computer aided screening awaits the impending switch to direct digital acquisition. The technology for direct digital mammography is just emerging and thus the motivation to further improve algorithms for detection is greater than ever before.

Algorithms reported in the literature for detection of breast cancer are designed to search for signs of cancer such as masses, clustered microcalcifications, and stellate patterns. These are the same signs of cancer used by radiologists to evaluate screening mammograms. Detection is difficult because the anomalies due to cancer are complex and are only subtly different from patterns arising from normal tissue. To overcome this problem, many algorithms work in two stages. In the first stage, regions of interest (ROI) are identified that resemble one of the signs of cancer and in the second stage, RIO are analysed more carefully to separate cancer from normal tissue. In the first stage, many false positive detections are accepted with the understanding that the second stage will separate these from true cancer. In the second stage, additional features are measured on the ROI. Some of these features reflect the experience of radiologists, and some, including image texture features, are not based on known manifestations of cancer in mammograms. Texture measures have been shown to contribute positively to discrimination between cancerous and normal tissue [6],[8],[9], [11],[12].

Although many good results have been reported, there are inherent weaknesses in this general approach.

1. Initial detection is focussed on complex signs of cancer.
2. Training studies for developing detection algorithms require accurate information regarding the location of anomalies (truthing). This information is often not available. In the case of detecting clustered microcal-

cifications, for example, nearly all algorithms initially detect individual calcifications. To train such an algorithm, the location in the image of every calcification must be known. Many calcifications are "obvious" but the greatest interest is in subtle calcifications. The error rate in assigning these is large and leaves open the possibility that information, not apparent to visual inspection, is omitted in the analysis.

3. Texture analysis is applied only to ROI, leaving open the possibility that better initial detection could be accomplished if texture information is used throughout the process. Measuring texture features over the entire image for the purpose of initial detection is often not feasible due to computational load.

1.1 Structural texture features

Structural texture refers to statistical distributions of image structures such as lines, edges, or bumps [5]. In contrast, statistical texture refers to statistical distributions of individual pixel values and includes measurements of the local mean, variance, higher order moments, run length statistics, and co-occurrence matrices.

Structural texture analysis provides a means to address the difficulties listed above.

1. Initial detection is focussed on simple structural primitives rather than complex manifestations of cancer.
2. Since statistics are computed over regions, only the disease state of regions (or the entire breast) is required for training.
3. Statistics are computed over image structures rather than pixels. The number of structures is usually at least two orders of magnitude smaller than the number of pixels in the image.

The structures considered in this paper are domains associated with local image intensity extrema. Their use is illustrated by three tasks in automatic detection of breast cancer: the detection of individual microcalcifications, detection and classification of clustered microcalcification, and the detection of invasive lobular carcinoma.

1.2 Clustered microcalcifications

Clustered microcalcifications appear in mammograms as groups of bright dots. Clustered microcalcifications are an important sign of cancer in mammograms, but their detection is not, in itself, a great achievement. The reason is that a very large percentage of normal breasts also show some signs of calcification. Although there are several forms of clusters associated with cancer and several forms associated with benign processes, there are some general differences between benign clusters and ones associated with cancer. Individual calcifications in clusters associated with cancer tend to be more irregular in shape and more varied in terms of contrast and size than those associated with normal tissue. Also, some benign clusters, particularly those associated with calcified blood vessels, can be distinguished based on the shape of the cluster itself [10].

In training algorithms for detecting clustered microcalcifications or for classifying clusters as benign or malignant, it is not possible to know with certainty, exactly which bright spots represent true calcification and which are manifestations of normal tissue such as crossing filaments. From histopathology reports, it is possible to know, with a great deal of certainty, the disease state of the tissue associated with the cluster. This provides motivation for using structural texture measures.

Our implementation of structural texture involves computing three features for every local image intensity maximum in the image: the radius of the bump associated with the local maximum, the height of the bump above background, and a measure of the symmetry of the bump. Detection of clusters is based on local distributions of these three features. In this way, the issue of identifying individual calcifications is bypassed. If there are very small calcifications or other disturbances in the tissue that are associated with cancer but cannot be discerned visually, this information is potentially incorporated into the detection process.

1.3 Invasive lobular carcinoma

The use of structural texture, and in particular the focus on local image intensity bumps, for detecting microcalcifications is natural given the visual appearance of microcalcifications in mammograms and guidance provided by the criteria used by radiologists during visual assessment. In contrast, there is no basis for using these structural textures detecting invasive lobular carcinoma, except by default.

Invasive lobular carcinoma is a form of breast cancer that is not visible at screening much more often than other forms of breast cancer. This form of cancer is not usually associated with the appearance of clustered microcalcifications and in many cases the growth pattern of the tumour is such that masses often are not seen in the mammogram.

Hence, there is no guide by which to select image features that are likely to lead to useful detection algorithms, except that visually obvious features need not be considered. This leaves spatially small and low contrast features as the only likely candidates. As a first attempt, image intensity minima were selected as candidate structure features.

2 Methods and materials

2.1 Data

Images used for these studies were digitised at 50 μ m spatial resolution and 12 bit depth using a Lumisys Lumiscan 150 laser digitiser. A number of image pre-processing steps were used to reduce the non-linear response of the film and digitiser system, to identify the breast region in the image, and to remove patient information, and subtract the background. These steps have been described elsewhere [3].

Images were included as examples of cancer only if the presence of cancer was confirmed by the pathology report and images were included as examples of normals only if no evidence of cancer was found in three years.

2.2 Feature extraction

In each image, local maxima (in the case of detecting microcalcification) and local minima (in the case of detecting invasive lobular carcinoma) were found by comparing pixel values with those of its 8-connected neighbours. For each local extrema point, p, the average pixel value on each concentric ring about p was found. In the case of image maxima, the average ring values form an initially decreasing function of the ring radius. In the case of image minima, the ring values form an initially increasing function. The radius at which these functions ceased to be strictly decreasing or increasing was taken to be the radius, R, associated with the local extremum. The region bounded by the ring of radius R was used to compute the net height, H, the volume, V, and the background level, B, associated with local extremum. The symmetry, S was taken to be the l^2 difference between the image values on the disk of radius R centred at p and the function obtained by revolving the average ring values about p. This latter function can be viewed as the ideally symmetric surface having the same height radius, volume, and average profile as the image intensity surface centred at p. Hence small values of S correspond to symmetric image features.

3 Experimental studies

3.1 Study 1. Detecting individual calcifications

Although the benefit of using structural texture features lies in detecting clusters without initial detection of individual microcalcifications, a study was conducted to determine the suitability of using the structural texture features described above to separate known microcalcifications from other image anomalies [3].

A total 107 individual microcalcifications were marked by a radiologist with experience in mammography. Using the features R, S, V, and H, and linear discriminant analysis, the detection rate was 90 percent at the operating point of 10 percent false positive rate. The area under the receiver operating characteristic (ROC) curve was $A_z = 0.945$.

This study does not confirm that structural texture features surpass other methods for detecting microcalcification. In fact, techniques for detecting microcalcifications are sufficiently mature, that there is very little room for improvement. At the recent International Workshop on Digital Mammography (IWDM, Bremen, June 2002) a panel of radiologists with experience in using commercial systems for automatic detection of microcalcifications agreed that current systems essentially do not miss clusters of diagnostic interest. However, many clusters not associated with caner are also detected by these systems. Hence, work on algorithms for classifying clusters as benign or malignant is still important.

This study does confirm that, in principle, structural texture features identify information relevant to distinguishing microcalcifications from other anomalies.

3.2 Study 2. Identification and classification of clusters

This study comprised 85 images, including 40 images containing clusters of microcalcification associated with benign clusters and 45 images containing clusters associated with cancer. The structural texture features R, S, V, and H were used for initial detection of clusters [7].

The algorithm identified many clusters not marked by the radiologist and assigned separate small clusters within general regions marked by the radiologist as a single cluster. This result cannot be used to measure the accuracy of the algorithm for identifying clusters for two reasons. First, radiologists usually note only the general region of the breast where a cluster or clusters occur and are not in the habit of delineating the spatial extent of individual clusters in detail since, in visual assessment of the mammogram, there is no benefit in doing so. Once a single cluster or region is found in an image that warrants calling the woman back for further tests, other regions are often not be noted even if further evidence of cancer is visible. Second, it is possible that clusters of subtle calcification are present that are not detectable by visual inspection.

The goal of the study was to assign each image as containing benign clusters only or as containing at least one malignant cluster. Accordingly, the performance of the entire system was measured in terms of the correct assignment of images as containing a malignant cluster or not. This process involved the initial detection of clusters as described above plus the extraction of statistical texture features from

these regions. In this particular study, classification of the image as cancer or non-cancer was based on the statistical texture features measured on the clusters rather than the structural texture features used for initial detection of the clusters [7].

The classification of test data resulted in an area under the ROC curve of $A_z = 0.858$ if examples of ductal carcinoma in situ (DCIS) comedo type were excluded and $A_z = 0.701$ for all types together. The reason for the discrepancy is that DCIS comedo type calcification is not well characterised by the structural texture features considered here. This type of calcification forms large, linear, and branching structures (the calcium fills ducts and acquires their shape) rather than round spots. The failure of this algorithm to detect these clusters is not important since they are easily discovered by visual assessment of the mammogram or by a separate algorithm designed specifically for detecting DCIS comedo calcification [3].

3.3 Study 3. Detection of invasive lobular carcinoma

A preliminary study was conducted to test the feasibility of detecting invasive lobular carcinoma in screening mammograms [1]. For this project, 24 images with invasive lobular carcinoma were obtained from 12 women. In each of these cases, the screening mammogram had been judged to show no evidence of cancer, but cancer was discovered by other methods within 28 months. The sizes of the tumours ranged from 15 mm to 100 mm in diameter. The mean diameter was 35 mm. As part of normal policy, these mammographically occult cases were reviewed by radiologists with expertise in mammography. Each image was judged to show no evidence of cancer in retrospect. In addition, 24 normal images were included in the study.

In the case of detection and classification of microcalcifications, classification was based directly on the statistics of R, H, V, and S values in local neighbourhoods. In the case of invasive lobular carcinoma, the locations of the tumours in the training data were not well known and so it was necessary to classify entire images as containing evidence of cancer or not. The volume feature, V, was not used in this study, but the local background statistic, B, was used.

First, attention was restricted to two regions within the $RHSB$ feature space. Five images with cancer and five normal images were selected randomly. The H, R, S, B values for all local minima in these images were plotted as points in the four-dimensional feature space. Visual inspection showed that, the great majority of points from the two classes of images formed an indistinguishable glob in the feature space. This was expected since large portions of the images with cancer correspond to normal tissue, and noise characteristics of cancer tissue and normal tissue are the same. However, one portion of the feature space was occupied nearly entirely by points from cancer images only. This portion of the feature space was divided into two sets.

$$\Omega_1 = \{H > 19, R = 1, S \leq 150, B > 2100\}$$
$$\Omega_2 = \{H > 38, R = 2, S \leq 200, B > 2100\}$$

Second, two image features were defined by

$$N_1 = |\Omega_1|/N \qquad \text{and} \qquad N_2 = |\Omega_2|/N,$$

where N is the total number of local minima in the image. Normalisation by N was used to compensate for breast size. Classification was based on these two image features.

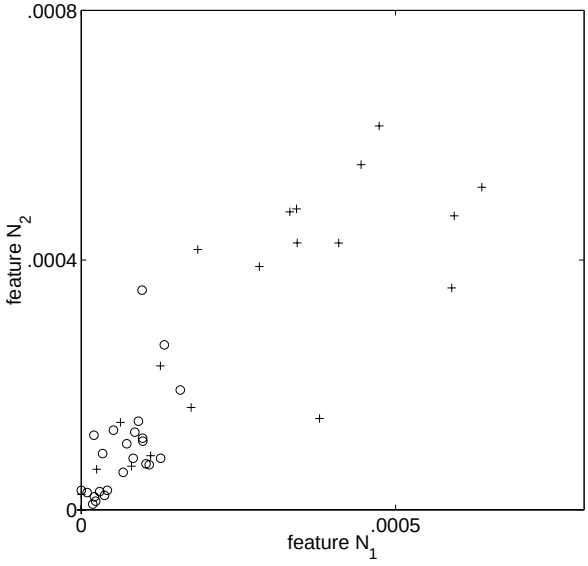

Figure 1. Scatter plot for detection of invasive lobular carcinoma. "o" - normal images, "+" - invasive lobular carcinoma. Approximately half of the carcinoma images cannot be distinguished from the normal images, but the other half are well separated from the normals.

An algorithm for detecting invasive lobular carcinoma must produce very few false positive detections. This is because the radiologist cannot verify a detection by visual inspection of the mammogram and the decision to call the woman back for further tests must be based on the algorithm's finding alone. For this reason, the results for detection of invasive lobular carcinoma should be reported in terms of the number of detections at the operating point of zero false positives rather than in terms of A_z scores.

In this study 50 percent of the images with invasive lobular carcinoma were correctly detected with no false positive reports.

4 Discussion

Computer algorithms will not replace human interpretation of screening mammograms in the foreseeable future. Current commercial systems for computer assisted screening mammography focus on bringing suspicious regions of the mammogram to the attention of the radiologist and serve mostly to catch evidence of cancer that is missed by the radiologist due to inattention. These systems only detect evidence of cancer that could have been detected by the radiologist.

It is reasonable to conjecture that more information regarding the presence of cancer is present in mammograms than can be extracted by visual interpretation. A viable goal is to develop algorithms that complement and extend the information available to the radiologist rather than mimic this information.

Structural texture measures provide a class of features that are not based on the experience of radiologists but are well suited for embodying information of diagnostic interest. They offer a computational advantages since image segmentation is focused on simple structures, local image extrema in this study, instead of complex structures such as masses, stellate patterns, or clustered microcalcification.

The results on detection and classification of clustered microcalcification indicate that simple structural texture features are useful in identifying known signs of cancer in mammograms. Unfortunately, it has not been possible to compare our classification performance with other studies since there are no public databases available that are both of sufficient resolution and contain a repertoire of both benign and malignant clusters.

Invasive lobular carcinoma represents approximately 8 percent of breast cancer. At screening, roughly half of these cases can be detected by visual inspection. Thus 4 percent of breast cancer cases are not be detected at screening because they are examples of occult invasive lobular carcinoma. In our study, half of these occult cases were detected indicating a potential improved detection rate for breast cancer of 2 percent.

The structure features used in the detection of invasive lobular carcinoma were not motivated by models or by the experience of radiologists. The only guidance was that these carcinomas could not be seen in the mammograms by radiologists and thus large or high contrast features could be neglected. Local extrema were chosen mostly because algorithms for characterising these structures were available from previous studies. The results obtained were surprisingly good considering that these features were largely unmotivated and were the first ones tested for this purpose. Unless these choices were very fortuitous, this implies that other structure features may well supply equal or greater discriminatory power for detecting invasive lobular carcinoma.

The study on detection of invasive lobular carcinoma was based on only 48 images and should be regarded as a preliminary study.

References

[1] M. J. Bottema and J. P. Slavotinek. Automatic detection of invasive lobular carcinoma in screening mammograms. submitted.

[2] M. J. Bottema and J. P. Slavotinek. Computer aided screening mammography. In B. Pham, editor, *New Approaches in Medical Image Analysis*, volume 3747, pages 177–190. SPIE, 1999.

[3] M. J. Bottema and J. P. Slavotinek. Detection of microcalcifications associated with cancer. In M. J. Yaffe, editor, *IWDM 2000 5 th International Workshop on Digital Mammography*, pages 149–153, 2001.

[4] H.-P. Chan, N. Petrick, and B. Sahiner. Computer-aided breast cancer diagnosis. In A. Jain, A. Jain, S. Jain, and L. Jain, editors, *Artificial Intelligence techniques in breast cancer diagnosis and prognosis*, pages 179–264. World Scientific, 2000.

[5] R. C. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16(2):131–164, 1989.

[6] H. D. Li, M. Kallergi, et al. Markov random field for tumor detection in digital mammography. *IEEE Transactions on Medical Imaging*, 14(3):565–576, 1995.

[7] S. Lu. Texture analysis for classification of microcalcification as benign or malignant in screening mammograms. PhD thesis, in preperation.

[8] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels. Gradient and texture analysis for the classification of mammographic masses. *IEEE Transactions on Medical Imaging*, 19(10):1032–1043, 2000.

[9] B. Sahiner, H.-P. Chan, N. Petrick, M. A. Halvie, and M. M. Goodsitt. Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis. *Medical Physics*, 25(4):516–526, 1998.

[10] L. Tabar and P. B. Dean. *Teaching Atlas of Mammography*. Thieme, New York, 1985.

[11] G. M. te Brake, N. Karssemeijer, and J. H. C. L. Hendriks. An automatic method to discriminate malignant masses from normal tissue in digital mammograms. *Physics in Medicine and Biology*, 45:2843–2857, 2000.

[12] C. J. Vyborny, M. L. Giger, and R. M. Nishikawa. Computer-aided detection and diagnosis of breast cancer. *Radiology Clinics of North America*, 38(4):725–740, 2000.