A Neural Network Detector for Clinically Normal Mammograms with Multiscale Feature Enhancement

RAQUEL GÓMEZ-SÁNCHEZ, DIEGO ANDINA
Dep. Señales, Sistemas y Radiocomunicaciones-E.T.S.I. Telecomunicación
Universidad Politécnica de Madrid
Ciudad Universitaria s/n, 28040 Madrid
SPAIN

Abstract: A computer-aided diagnosis (CAD) scheme to detect clinically normal mammograms is presented. The objective is to develope a method for the automatic recognition of normal mammograms with a very low probability of classifiying abnormal images as normal. Since the radiologist spends an enormous amount of time investigating images lacking any abnormality, and the vast majority of mammograms are clinically normal, the method would potentially save valuable time.

The method is divided in two steps. The first step accomplishes a feature enhancement to improve the visualization of any abnormal lesion. Multiscale wavelet representations of the mammographic images are used for this task. A multiscale representation of a mammogram provides a set of images, each of them giving a different scale view of the initial mammogram. Abnormal mammographic lesions appear enhanced in one of more than one of the images, that means, their intensity profiles appear emphasized. These lesions appear as singularities easier to discriminate than in the original image.

The second step obtains a set of parameters that compiles the information from the first step. Parameter examples are mean pixel intensity values and standard deviation of the pixel intensity values in the regions of interest. This set constitutes the input to a detector. A neural network optimized in the Neyman-Pearson sense is used as a detector. For training the neural network, a data base with normal and abnormal mammograms (with biopsy proven abnormal lesions) is used. IMACS/IEEE CSCC'99 Proceedings, Pages:2051-2055

Key-Words: - wavelets, multiscale representations, neural networks, mammograms, microcalcifications

1 Introduction

Screen/film mammography has been widely recognized as being the only effective imaging modality for the early detection of breast cancer in asymptomatic women. Screening asymptomatic women using screen/film mammography has been shown to significantly reduce breast cancer mortality. Major advances in screen/film mammography have occurred over the past decade which have resulted in significant improvements in image resolution and film contrast. Despite these advances, however, screen/film mammography remains a diagnostic imaging modality where image interpretation is difficult.

An early sign of disease in 30-50% of mammographically detected cases is the appearance of clusters of fine, granular microcalcifications [1] whose individual grains typically range in size from 0.05-1 mm in diameter. Individual microcalcifications are difficult to detect because of variations in their shape and size and because they are embedded in and camouflaged by varying densities of parenchymal tissue structures. Computer-aided diagnosis (CAD) schemes using digital image processing techniques have the goal of improving the detection performance of secreening mammography.

An important branch of CAD methods for feature enhancement in mammography employs wavelet transforms [2]-[5]. Our proposal for detecting microcalcifications in mammograms is to perform a multiscale edge detection [6] and analyse the properties of the edges that have been detected. The information obtained from the above process can be packed in a set of parameters, which constitutes the input to a detector. A neural network will perform the detection task. This neural network must have been previously trained with a sufficient amount of examples corresponding to proven normal and abnormal mammograms regions.

Since the vast majority of mammograms are normal, our interest will be focused in detecting clinically normal mammograms. The final objective

is to save the valuable time that radiologists spend analysing normal mammograms. Therefore the most important factor is to achieve a Pfa (false alarm probability) for the neural detector zero or close to zero, that is, abnormal regions must not be detected as normal. This optimization of the Pfa leads to a decreasing of the Pd (detection probability). This means that an amount of normal mammograms could not be detected and, then, considered abnormal. But this has less impact, since the abnormal mammograms will be further analysed by radiologists.

The paper is organized as follows. Section 2 focuses in the analysis of the mammographic data. Section 3 reviews the wavelet transform properties and relates multiscale edge detection to the wavelet transform. Section 4 proposes a clinically normal mammograms detection scheme employing a neural network with the wavelet preprocessing introduced in Section 3, including the obtained results. We finalize with discussion and conclusions in Section 5.

2. Characteristics of mammographic images

Although mammography currently is the best method for the detection of breast cancer, between 10% to 30% of women who have breast cancer and undergo mammography have negative mammograms. In approximately two thirds of these false-negative mammograms, the radiologist failed to detect the cancer that was evident retrospectively. The missed detections may be owing to poor image quality, eye fatigue or oversight by the radiologist. Double reading has been suggested, with the number of lesions found increasing by 15%. Thus, the goal of CAD research is to develop computer methods as aids to the radiologists, in order to increase diagnostic accuracy in mammography screening programs.

Microcalcifications are often a presenting sign among early breast cancers. On screening studies, 90% of all cases of nonpalpable ductal carcinoma in situ (DCIS) [8] and 70% of all cases of minimal carcinoma (infiltrating cancer smaller than 0.5 cm and all DCIS) were seen on the basis of microcalcifications alone.

The search of microcalcifications lends itself to computer detection methods because of their high clinical relevance and the lack of normal structures that have the same appearance. Individual microcalcifications appear as small (typically 0.05-1 mm) particulate objects of variable shape (from glandular to rod-shaped) and fairly uniform optical density. A typical example of microcalcification is presented in Figure 1. Although microcalcifications vary in outline and degree of elongation, the average form is roughly circular, with a tapered cross-sectional profile. Microcalcifications often appear in clusters.

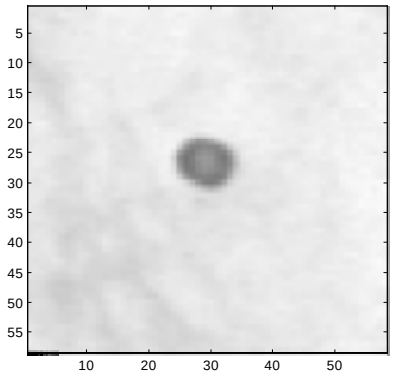

Figure 1: Example of circular microcalcification in a region of 1cmx1cm.

The visibility of microcalcifications is often degraded by the high frequency texture of the breast tissue. Special attention must be paid in not to confuse singularities owing to breast tissue with those owing to microcalcifications.

A different type of lesions in mammograms are mass lesions. The computerized detection of mass lesions is different from that of microcalcifications because of some mass lesions and some normal parenchymal tissues mimic each other, thus making the interpretation difficult for both the human and the computer. In this work we are not involved with this type of lesions.

Once a lesion is detected, benignancy or malignancy must be determined. We are concerned only in detection of microcalcifications, without further analysis.

3 Wavelet representations of images and multiscale edge detection

Let $\psi^1(x,y)$ and $\psi^2(x,y) \in L^2(\mathbf{R}^2)$ two bidimensional wavelet functions and a integer j. We denote that:

$$\psi_{2^j}{}^1(x,y) = \frac{1}{2^{2j}}\psi^1\left(\frac{x}{2^j},\frac{y}{2^j}\right) \qquad (2.1)$$
$$\psi_{2^j}{}^2(x,y) = \frac{1}{2^{2j}}\psi^2\left(\frac{x}{2^j},\frac{y}{2^j}\right)$$

The wavelet transform of a function $f(x,y) \in L^2(\mathbf{R}^2)$ at the scale 2^j has two components defined by the following convolutions:

$$W_{2^j}{}^1 f(x,y) = f * \psi_{2^j}{}^1(x,y)$$
$$W_{2^j}{}^2 f(x,y) = f * \psi_{2^j}{}^2(x,y) \quad (2.2)$$

We refer to the 2-D dyadic wavelet transform of $f(x,y)$ as the set of functions:

$$Wf = \left(W_{2^j}{}^1 f(x,y),\ W_{2^j}{}^2 f(x,y)\right)_{j\in\mathbf{Z}} \quad (2.3)$$

We use the term 2-D smoothing function to describe any function $\theta(x,y)$ whose integral over x and y is equal to 1 and converges to 0 at infinity. The image $f(x,y)$ is smoothed at different scales s by a convolution with $\theta_s(x,y) = (1/s^2)\theta(x/s, y/s)$. If we define:

$$\psi^1(x,y) = \frac{\partial\theta(x,y)}{\partial x}$$
$$\psi^2(x,y) = \frac{\partial\theta(x,y)}{\partial y} \quad (2.4)$$

it can be easily proved [6] that the wavelet transform can be rewritten

$$\begin{bmatrix} W_{2^j}{}^1 f(x,y) \\ W_{2^j}{}^2 f(x,y) \end{bmatrix} = 2^j \begin{bmatrix} \frac{\partial}{\partial x}(f * \theta_{2^j})(x,y) \\ \frac{\partial}{\partial y}(f * \theta_{2^j})(x,y) \end{bmatrix} = \quad (2.5)$$
$$= 2^j \vec{\nabla}(f * \theta_{2^j})(x,y)$$

The two components of the wavelet transform are proportional to the two components of the gradient vector of the function f smoothed at the scale 2^j. The first component measures how sharp $f(x,y)$ smoothed at a scale 2^j varies along horizontal directions, while the second component measures the variation along vertical directions. At each scale 2^j, the modulus of the gradient vector is proportional to:

$$M_{2^j} f(x,y) = \sqrt{\left|W_{2^j}{}^1 f(x,y)\right|^2 + \left|W_{2^j}{}^2 f(x,y)\right|^2} \quad (2.6)$$

The angle of the gradient vector with the horizontal direction is given by:

$$A_{2^j} f(x,y) = tan^{-1}\left(\frac{W_{2^j}{}^2 f(x,y)}{W_{2^j}{}^1 f(x,y)}\right) \quad (2.7)$$

The sharp variation points of $f * \theta_s(x,y)$ are the points (x,y) where the modulus $M_s f(x,y)$ has a local maximum in the direction of the gradient given by $A_s f(x,y)$. We record the position of each of this modulus maxima as well as the values of the modulus $M_s f(x,y)$ and the angle $A_s f(x,y)$ at the corresponding locations. At fine scales, there are many local maxima created by the image noise, but at this locations, the modulus value has a small amplitude. We are interested in edge points whose modulus is larger than a given threshold at all scales. At coarse scales, the modulus maxima have different positions than at fine scales. This is due to the smoothing of the image by the function $\theta_s(x,y)$

Sharp variations of 2-D signals are often not isolated but belong to curves in the image plane. Along these curves, the image intensity can be singular in one direction while varying smoothly in the perpendicular direction. It is well known that such curves are more meaningful than edge points by themselves beacuse they generally are the boundaries of the image structures. For discrete images, we reorganize the maxima representation into chains of local maxima to recover these edge curves. Then, we can characterize the properties of edges from the modulus maxima evolution across scales.

At the scale s, the wavelet modulus maxima detect the sharp variation points of $f * \theta_s(x,y)$. Some of these modulus maxima define smooth curves in the image plane along which the profile of the image intensity varies smoothly. At any point along the maxima curve,
the gradient of $f * \theta_s(x,y)$ is perpendicular to the tangent of the edge curve. We thus chain two adjacent local maxima if their respective position is perpendicular to the direction indicated by the angle $A_s(x,y)$. Since we want to recover edge curves along which the image profile varies smoothly, we only chain together maxima points where the modulus $M_s f(x,y)$ has close values. This chaining procedure defines an image representation that is a set of maxima chains.

4. Microcalcifications detection method and results

In this section we introduce a CAD method to detect the presence of clustered microcalcifications in mammograms. First we perform a multiscale feature enhancement and secondly we propose a detector adapted to the preprocessed images.

4.1 Feature enhancement

We perform a feature enhancement with a multiscale analysis. The objective in this step is to remove as much as possible information which is not relevant. We refer to intensity variations owing to parenchymal tissue structures and film noise. The first step is to obtain a dyadic wavelet transform of the full breast area with mother wavelets defined as in (2.4). Once obtained we can compute the modulus and the angle of the gradient vector (2.5) for each scale 2^j, as well as the sharp variation points inside the breast area.

The majority of these sharp variation points do not correspond to lesions in the mammogram, thus we can eliminate them. The removing criterion is to

establish a threshold value, we keep the singularities whose gradient modulus values overcome this threshold. As a first threshold value we select a half of all maxima values.

At this point we have reduced the information content of the mammogram to a set of singularity points. The characteristics of this set determine the normality or abnormality of the mammogram.

Figures 2 and 3 show a region of a mammogram before and after the multiscale feature enhancement. As it is shown, most of the noise and tissue variations are removed in the processed image, remaining the three present microcalcifications.

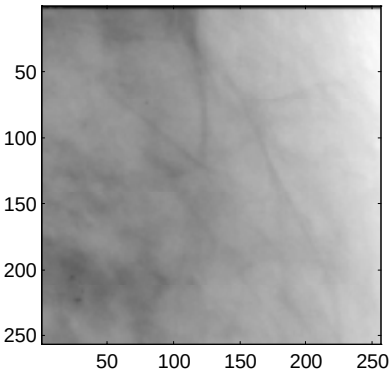

Figure 2: Mammographic area with three proven microcalcifications.

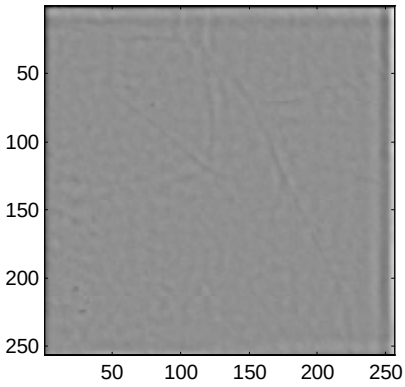

Figure 3: The same mammographic area after the multiscale enhancement process.

2.2 Detection

We have selected an artificial neural network to perform the detection task. Artificial neural networks constitute a nonalgorithmic approach to information processing. These neural networks, which are capable of processing a large amount of information simultaneously, address problems not by means of prespecified algorithms but rather by learning from examples that are presented repeatedly. The popularity of neural networks is due primarily to their apparent ability to make decisions and draw conclusions when presented with complex, noisy or partial information and to adapt their behaviour to the nature of the training data. In medical imaging, artificial neural networks have been applied to a variety of data-classification and pattern recognition tasks, such as the differential diagnosis of interstitial diseases, and have been shown to provide a potentially powerful classification tool [10].

We propose a three-layered feedforward neural network with a backpropagation algorithm for the interpretation of the mammographic features. The enhanced image is divided in regions of size 8x8 pixels, corresponding to regions of 1.5mmx1.5mm. Both the mean intensity value and the standard desviation will be higher for regions containing a microcalcification, as can be observed in Figure 4. This figure shows the microcalcification in figure 1 after enhancement. In consequence, these two parameters are selected as input to the neural network.

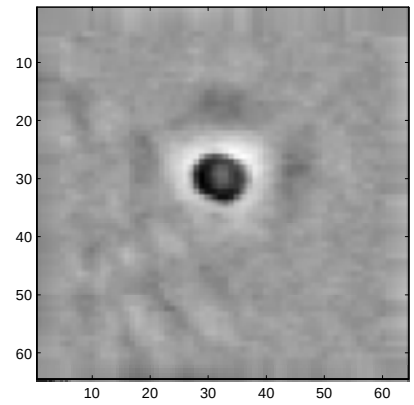

Figure 4: Microcalcification in Figure 1 after enhancement.

We have trained a three layer neural network with two inputs and one output by means of the cross-validation method, achieving a Pfa value of 0.001, with a Pd of normal mammograms of 50%. This early result predicts the success of the wavelet-based tools and motivates further investigations.

5. Conclusions

The proposed method emphasizes the ability of wavelets to perform feature extraction. In applications where the amount of information to manage is large they offer a potential tool that have been proved in several fields.

One of these fields is the detection and characterization of singularities in images. When the

images are constituted by mammograms, we can use this potential for detecting microcalcifications, due to the characteristics of this type of lesions. In this proposal we have combined the properties of wavelet processing with characteristics of mammographic lesions to improve their visualization. The emphasized charactetristics combined with a neural network detector lead to a detection method of normal mammograms with very low probabilities of classify abnormal mammograms as normal.

References:

[1] B. S. Monseses, "Evaluation of breast microcalcifications", *Radiologic Clinics of North America*, Vol. 33, No. 6, pp. 1109-1121, Nov 1995.

[2] R. N. Strickland, H. I. Hahn, "Wavelet transform for detecting microcalcifications in mammograms", *IEEE Trans. on Medical Imaging*, Vol. 15, No. 2, April 1996, pp. 218-229.

[3] A. Laine, S. Schuler, J. Fan, W. Huda, "Mammographic feature enhancement by multiscale analysis", *IEEE Trans. on Medical Imaging*, Vol. 13, No. 4, April 1994, pp. 725-740.

[4] W. Qian, L. P. Clarke, M. Kallergi, R. Clark, "Tree-structured nonlinear filters in digital mammography", *IEEE Trans. on Medical Imaging*, Vol. 13, No. 1, March 1994, pp. 25-36.

[5] J. Jeine, S. R. Deans, "Multirresolution statistical analysis of high-resolution digital mammograms", *IEEE Trans. on Medical Imaging*, Vol. 16, No. 5, October 1997, pp. 503-515.

[6] S. Mallat, S. Zhong, "Characterization of signals from multiscale edge", *IEEE Trans.on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 7, July 1992, pp. 710-732.

[7] Y. Meyer, *Ondelettes et Operateurs*, New York: Herman, 1990.

[8] S. A. Feig, G. S. Shaber, A. Patchefsky, "Analysis of clinically occult and mammographically occult breast tumors" AJR 1997, Vol. 128, pp. 403-408.

[9] S. A. Feig, "Mammographic evaluation of calcifications", *RSNA Categorical Course in Breast Imaging*, 1995, pp. 93-105.

[10] Y. Mu, M. L. Giger, K. Doi, "Artificial Neural Networks in Mammography: Application to Decision Making in the Diagnosis of Breast Cancer", *Radiology*, April 1993, 187:81-87.