ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 195-200

Automatic detection of clustered microcalcifications in digital mammograms using an SVM classifier

Armando Bazzani †◊, Alessandro Bevilacqua ‡◊, Dante Bollini †◊,
Rosa Brancaccio †◊, Renato Campanini †◊, Nico Lanconelli ◊,
Alessandro Riccardi †◊, Davide Romani ◊ and Gianluca Zamboni †

† Department of Physics, University of Bologna, Italy
‡ Department of Electronics, Computer Science and Systems, University of Bologna, Italy
◊ National Institute for Nuclear Physics, Bologna, Italy

Abstract. In this paper we investigate the performance of a Computer Aided Diagnosis (CAD) system for the detection of clustered microcalcifications in mammograms. Our detection algorithm consists on the combination of two different methods. The first one, based on difference-image techniques and gaussianity statistical tests, finds out the most obvious signals. The second one is able to discover more subtle microcalcifications by exploiting a multiresolution analysis by means of the wavelet transform. In the false-positive reduction step we separate false signals from microcalcifications by means of an SVM classifier.
Our algorithm yields a sensitivity of 94.6% with 0.6 false positive cluster per image on the 40 images of the Nijmegen database.

1. Introduction

Breast cancer is the most common form of cancer among women. The presence of microcalcifications in breast tissues is one of the main features considered by radiologists for its diagnosis. Several techniques developed for the automated detection of microcalcifications can mainly be grouped in three different categories: multiresolution analyses [1,2], filtering methods [3] and statistical methods [4,5]. By comparing the different methods it turns out that some microcalcifications are detected by one method but missed by others.
In this paper we propose an approach based on the combination of different detection methods in order to get optimal performances. Yoshida *et al* pointed out that the simultaneous use of two or more techniques might improve the results of an optimized single method [6]. The basic idea of our method is to combine a multiresolution analysis based on wavelet transform with a difference-image method and a gaussianity statistical test and to perform a logical OR operation on the detected microcalcifications before clustering.
In the false-positive reduction (*fpr*) step we try to separate false signals from microcalcifications by using a classifier based on a Support Vector Machine (SVM).
The detection scheme has been tested on 40 digitized mammograms coming from Nijmegen Hospital.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 195-200

2. Methods

2.1. Detection scheme

Microcalcifications are very small spots that are relatively bright compared with the surrounding normal tissue. Typically they are between 0.1 mm and 1 mm in size and are of particular clinical significance when found in clusters of five or more in a 1 cm^2 area. Most of the clusters consist of at least one evident microcalcification and other more hidden signals. Our approach includes two different methods: the first one (coarse) is able to detect the most obvious signals and uses difference-image techniques and gaussianity tests, while the second one (fine), based on multiresolution analyses, discovers more subtle microcalcifications. Signals coming out from these methods are combined through a logical OR operation and then clusterized to give the final result.

2.2. Coarse method

In this part of the algorithm we remove structured image background by means of a difference-image technique. The scheme of the coarse method is shown in figure 1.

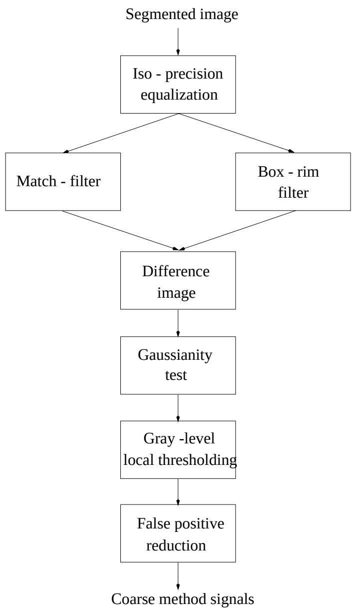

Figure 1. Scheme of the coarse method.

First of all we perform an iso-precision noise equalization as described in [4]. The equalized image is passed through two different filters: a 3x3 match-filter that produces a signal-enhanced image and a 7x7 box-rim filter that gives a signal-suppressed image. By subtracting the suppressed image from the enhanced one, we obtain a difference-image, which contains noise and signals resembling

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 195-200

microcalcifications. According to experimental evidences we assume that the remaining noise is gaussian, since we have reduced the structured noise in the previous steps. We then employ a gaussianity test on the difference image in order to choose ROI's that include interesting signals. Since the difference-image contains only gaussian noise and signals with a high contrast we should have a deviation from gaussianity in regions including microcalcifications. Here we perform the gray-level thresholding: the central pixel of the considered 51x51 window of the difference-image is retained only if its gray level is greater than the mean pixel value multiplied by a preselected k multiple of standard deviations σ.
The next step is a false-positive reduction (*fpr*) phase based on a local edge-gradient analysis [7]. We consider five features (*area*, *average pixel value*, *edge gradient*, *degree of linearity*, *average local gradient*) to separate microcalcifications from false signals. These features are the inputs of a SVM classifier described in subsection 2.4.
Signals survived to the *fpr* stage will join others coming from the fine method described in the next subsection.

2.3. Fine method

In this part of the detection scheme we try to discover more subtle microcalcifications. In figure 2 is depicted the scheme of the fine algorithm. Microcalcifications are characterized by well-defined range size and high local contrast, so we find out signals having these features. We split the algorithm into two independent sections.

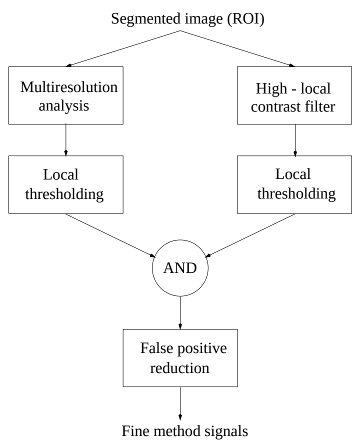

Figure 2. Scheme of the fine method.

The first one detects signals having size smaller than 1 mm by means of a multiresolution analysis based on wavelet transform. To extract interesting signals we perform a local thresholding in 40x40 pixel size windows. Assuming a gaussian distribution for the noise we fit with a parabola the gray level histogram of the window in semi-logarithmic scale. Then we retain pixels having a gray level greater than the intersection of the parabola with the x axis.

In the second section signals having a high local contrast are enhanced by using a difference image technique. We subtract a suppressed image obtained by a 9x9 moving average filter from an enhanced image coming from a 3x3 match-filter. We carry out the same local thresholding on the difference image followed by the morphological opening.
After that a logical AND operation is accomplished on signals extracted by the two sections of the fine method.
To split false signals from microcalcifications a *fpr* phase similar to the one used in the coarse method is performed.
Finally, microcalcifications which have passed the *fpr* step are joined with others coming from the coarse method through a logical OR operator.

2.4. SVM classifier

False-positive reduction step is a two classes pattern recognition problem: we must distinguish true microcalcifications from false signals. Support Vector Machines (SVMs) are learning machines used in pattern recognition and regression estimation problems [8,9]. They grow up from Statistical Learning Theory (SLT) [8], which gives some useful bounds on the generalization capacity of machines for learning tasks. The SVM algorithm constructs a separating hypersurface in the input space. Its way to do that is:

a) mapping the input space into a high dimensional features space through some non linear mapping chosen a priori (*kernel)*;
b) constructing in this features space the Maximal Margin Hyperplane.

Hyperplanes are defined by $\mathbf{w}\cdot\mathbf{x}+b=0$: when training data $(\mathbf{x}_i, y_i), i=1,...,l$ are separated by this hyperplane it happens that $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1$, where $y_i=\pm 1$ are the labels. It can be shown that the margin is $2/\|\mathbf{w}\|$, so finding the hyperplane which separates data with maximal margin is equal to:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 \\ \text{with } y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1 \end{cases} \tag{1}$$

In order to allow for misclassication errors, constraints are relaxed to $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1-\xi_i$, $\xi_i\geq 0$. (1) becomes then:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 + C\cdot\sum\xi_i \\ \text{with } y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1-\xi_i \end{cases} \tag{2}$$

The dual formulation of (2) reduces to:

$$\begin{cases} \text{maximize} \sum\alpha_{\mathrm{i}} - \frac{1}{2}\sum\alpha_i\alpha_j(\mathbf{x}_i\cdot\mathbf{x}_j)y_i y_j \\ \text{with} \sum\alpha_{\mathrm{i}} y_i = 0, 0\leq\alpha_i\leq C \end{cases} \tag{3}$$

This formulation, where example vectors are present only in dot products, makes quite simple the execution of point a), because of a theorem by Mercer [8]. It gives an easy way to compute dot products in features space, where vectors of input space are non-linearly mapped by a function $\phi(\mathbf{x})$. By using a suitable function K such that $\phi(\mathbf{x}_i)\cdot\phi(\mathbf{x}_j)=K(\mathbf{x}_i,\mathbf{x}_j)$ we do not need to calculate each singular mapping $\phi(\mathbf{x})$. For our purposes we need positive misclassified examples to outweigh negative ones: in order to do this it is necessary to modify the primal in the following way:

$$\begin{cases} \text{minimize } \|\mathbf{w}\|^2/2 + C^+\cdot\sum \xi_i^+ + C^-\cdot\sum \xi_i^- \\ \text{with } (\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1-\xi_i^+,\ (\mathbf{w}\cdot\mathbf{x}_i+b)\leq -1+\xi_i^- \end{cases} \tag{4}$$

where C^+ and C^- give different costs to false-positive and false-negative errors.
We have chosen a polynomial *kernel*: $(\mathbf{x}\cdot\mathbf{y}+1)^d$, with degree ranging from 2 to 7. For each value of d we have constructed a Free Response Operating Characteristics (FROC) curve changing the ratio C^+/C^-. It turned out that this ratio is much more important for fine tuning than the choice of d or of the absolute value of C^-. We train the machine with *fpr* features of signals coming from 20 images of the Nijmegen database.

3. Results

We have compared the SVM classifier with a Multi-Layer Perceptron (MLP) one. The MLP has one hidden layer and we used a Resilient Back-propagation (Rprop) training algorithm. The performance of our detection algorithm is shown in figure 3 and figure 4. In figure 3 there are depicted two FROC curves regarding the whole database (*test* + *training*): one related to the detection method with MLP and one with the SVM classifier. In figure 4 there are depicted the curves relative to *training* and *test* with the SVM classifier. On the 40 images we yield a sensitivity of 94.6% with 0.6 false per image. These results are comparable to others obtained on the same database [4,10,11].

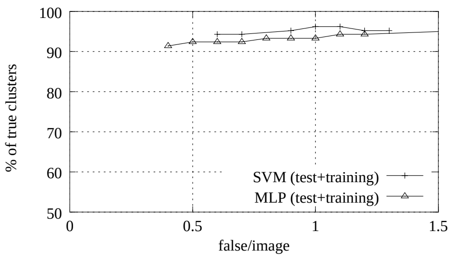

Figure 3. FROC of the detection system on the Nijmegen database with MLP classifier and with SVM classifier.

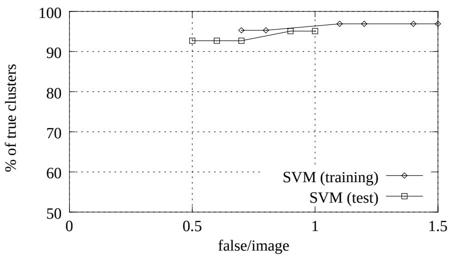

Figure 4. FROC of the detection for *training* and *test* with SVM classifier.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 195-200

This work is supported by Italian National Institute for Nuclear Physics (CALMA project). Images were provided by courtesy of the National Expert and Training Center for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, the Netherlands.

References

1. H. Yoshida, K. Doi, R. M. Nishikawa: Automatic detection of clustered microcalcifications in digital mammograms using wavelet transform techniques. Proc. SPIE **2167**, 868-86 (1994).
2. M. J. Lado, P. G. Tahoces, A. J. Mendez, M. Souto, J. J.Vidal: A wavelet based algorithm for detecting clustered microcalcifications in digital mammograms. Med. Phys. **26**, 1294-305 (1999).
3. H. P. Chan, K. Doi, S. Galhotra, C. J. Vyborny, H. MacMahon, P. M. Jokich: Image feature analysis and Computer Aided Diagnosis in digital radiography: automated detection of microcalcifications in mammography. Med. Phys. **14**, 538-48 (1987).
4. N. Karssemeijer: Adaptive noise equalization and recognition of micro-calcification clusters in mammograms. International Journal of Pattern Recognition and Artificial Intelligence **7**, 1357-76 (1993).
5. M. N. Gürcan, Y. Yardimci, A. E. Cetin: Microcalcifications detection using adaptive filtering and gaussianity tests. Digital Mammography: Nijmegen, Kluwer Academic Pub, 157-64 (1998).
6. H. Yoshida, K. Doi, R. M. Nishikawa, M. L. Giger, R. A. Schmidt: An improved CAD scheme using wavelet transform for detection of clustered microcalcifications in digital mammograms. Acad. Radiol. **3**, 621-7 (1996).
7. T. Ema, K. Doi, R. M. Nishikawa, Y. Jiang, J. Papaioannou: Image feature analysis and Computer Aided Diagnosis in digital radiography: reduction of false-positive clustered microcalcifications using local edge-gradient analysis. Med. Phys. **22**, 161-9 (1995).
8. V. Vapnik: The nature of statistical learning theory. Springer Verlag (1995).
9. T. Joachims: Making large-scale SVM learning practical. Advances in kernel methods- Support Vector Learning. B. Schölkopf, C. Burges, A. Smola, MIT-press (1999).
10. W. Veldkamp, N. Karssemeijer: Improved correction for signal dependent noise applied to automatic detection of microcalcifications. Digital Mammography: Nijmegen, Kluwer Academic Pub, 169-76 (1998).
11. S. Brown, R. Li, L. Brandt, L. Wilson, G. Kossof, M. Kossof: Development of a multi-feature CAD system for mammography. Digital Mammography: Nijmegen, Kluwer Academic Pub, 189-96 (1998).