Detection of Microcalcifications Clusters in Mammograms through TS-MRF Segmentation and SVM-based Classification

C. D'Elia°, C. Marrocco°, M. Molinara°, G. Poggi*, G. Scarpa*, F. Tortorella°

(°) *Dipartimento di Automazione, Elettromagnetismo, Ingegneria dell'Informazione e Matematica Industriale*
Università degli Studi di Cassino
via G. di Biasio, 43 03043 Cassino - Italy
E-mail: {delia,marrocco,molinara,tortorella}@unicas.it

(*) Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni
Università degli Studi di Napoli "Federico II"
via Claudio, 21 80125 Napoli - Italy
E-mail: {poggi,giscarpa}@unina.it

Abstract

At present, mammography is the only not invasive diagnostic technique allowing the diagnosis of a breast cancer at a very early stage. A visual clue of such disease particularly significant is the presence of clusters of microcalcifications. Reliable methods for an automatic detection of such clusters are very difficult to accomplish because of the small size of the microcalcifications and of the poor quality of the digital mammograms. A method designed for this task is described. The mammograms are firstly segmented by means of the Tree Structured Markov random field algorithm which extracts the elementary homogeneous regions of interest on the image. Such regions are then submitted to a further analysis (based both on heuristic rules and Support Vector classification) in order to reduce the false positives. The approach has been successfully tested on a standard database of 40 mammographic images, publicly available.

1. Introduction

Mammography is a radiological screening technique which makes it possible to detect lesions in the breast using low doses of radiation. At present, it represents the only not invasive diagnostic technique allowing the diagnosis of a breast cancer at a very early stage, when it is still possible to successfully attack the disease with a suitable therapy. For this reason, programs of wide mass screening via mammography for the female population at risk have been carried out in many countries.

A visual clue of breast cancer particularly meaningful is the presence of clusters of microcalcifications. Microcalcifications are tiny granule-like deposits of calcium that appear on the mammogram as small bright spots. Their size ranges from about 0.1 mm to 0.7 mm, while their shape is sometimes irregular. Isolated microcalcifications are not, in most cases, clinically significant. However, the low quality of mammograms and the intrinsic difficulty in detecting likely cancer signs make the analysis particularly fatiguing, especially in a mass screening where a high number of mammograms must be examined by a radiologist in a day. In this case, a computer aided analysis could be very useful to the radiologist both for prompting suspect cases and for helping in the diagnostic decision as a "second reading". The goal is twofold: to improve both the *sensitivity* of the diagnosis, i.e. the accuracy in recognizing all the actual clusters and its *specificity*, i.e. the ability to avoid erroneous detections.

In the recent past, many approaches have been proposed for the automatic detection of clusters of microcalcifications, based on wavelets, Gaussian filtering, artificial neural networks, texture analysis, mathematical morphology and fuzzy logic.

In this paper we present a novel method for detecting clustered microcalcifications on digital mammograms. The first step is a segmentation of the mammographic image by means of a Markov random field (MRF) model. The MRF approach to the segmentation gives the possibility of including an a priori knowledge on the segmentation results x. Indeed, we can consider x as a realization of a random field X whose density function $p(x)$ models our a priori knowledge. A MRF-based approach has been previously used for detecting clustered microcalcifications [1].

Two critical points of such approach are the computational burden and the sensitivity of the results to the model parameters. For these reasons, we employ a tree-structured MRF-based segmentation [2], which address these topics and allows us to obtain a segmentation process fast and quite spatially adaptive since all field parameters are estimated locally.

The regions obtained by the segmentation are successively examined in order to keep only the actual microcalcifications. Such analysis is accomplished by means of a two-stage, coarse-fine classification: the first

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

stage takes into account some constraints on the geometry and on the size of the regions. The second stage is performed by means of a Support Vector Machine [3] on more refined features.

The approach has been experimented with a standard database of mammograms, obtaining encouraging results which confirmed its effectiveness.

2. The Image Segmentation

Mammogram segmentation can be easily formulated as a MAP estimation problem. Suppose each pixel of the image S belongs to one of C different classes, and let x_s in $\{1,...,C\}$ indicate the class of pixel s. Then $x = \{x_s, s \in S\}$ is the segmentation of the image S in C classes. Of course, x is unknown, and must be recovered from the observable data $y = \{y_s, s \in S\}$ where y_s is the value of the pixel s in the original mammogram. If we model all quantities as random fields, i.e. if we assume that x and y are particular realizations of two random fields X and Y, a natural way to carry out the segmentation process is to select x as the realization $\hat{x}$ with the largest conditional probability given the data y, namely:

$$\hat{x} = \arg\max_x p(x|y) = \arg\max_x p(y|x)p(x) \qquad (1)$$

The last equality can be written since the prior probability $p(y)$ does not affect the result.

The image data are modeled as conditionally independent Gaussian, given the class, namely $p(y|x) = \prod_{s \in S} p(y_s|x_s)$, with $p(y_s|x_s) \sim N(\mu_{x_s}, \sigma_{x_s})$.

As for the field of classes, it is convenient to model it as a MRF. Indeed, this is a reasonably simple, yet general, model which keeps into account the spatial dependencies in the image through the conditional probability that a pixel belongs to a given class given the classes of its neighbours. As a result, X has Gibbs distribution

$$\Pr(X = x) = \frac{1}{Z} \exp\left(\frac{1}{T} \sum_{c \in C} V_c(x, \beta) \right) \qquad (2)$$

where Z is a normalizing constant, and $V_c(x,\beta)$ are potential functions, defined on suitable cliques c of the image, and depending on some hyperparameters.

Given this model, the segmentation problem amounts to maximizing the function $p(y|x)p(x)$ over x, where all the quantities C, $\mu_1,\ldots, \mu_C$, $\sigma_1,\ldots, \sigma_C$ and β are in general unknown and must be estimated from the data. Due to the inherent complexity of this problem, in practical applications one must resort to heuristics that reduce the search complexity, and accept suboptimal solutions.

To drastically reduce the search complexity, we adopt a tree-structured MRF model [2], where the full segmentation is obtained through a sequence of binary segmentations. More precisely, the whole image is associated to the root node $t = 1$ of a tree T, and is segmented in two regions using a binary MRF model. The two new regions, associated with the children of the root, $t = 2$ and $t = 3$, can be likewise segmented by means of newly defined local binary MRF, and the growth of the tree continues until a suitable stopping condition is met. Therefore, each node t of the tree is associated with a region of the image S_t, a field of observables Y_t with realization y_t, a binary MRF X_t with realization x_t, and a set of parameters $\{\mu_t, \sigma_t, \beta_t\}$. The leaves of the tree partition the image in C disjoint regions, i.e. provide the desired segmentation. In [2] it is shown that the growth of the tree can be based exclusively on local decisions. In fact, a split gain G_t is associated with each leave t,

$$G_t = \frac{p(y_t|x_t)p(x_t)}{p(y_t|t)} \qquad (3)$$

defined as the likelihood ratio between the two hypotheses of splitting the region in two (according to the realization x_t of the local binary MRF) or leaving it unaltered. If the split gain is greater than 1, the region S_t is better described by a two-class field rather than by a uniform field. When all the leaves have split gain less than 1 the tree stops growing. The use of binary fields only, together with the locality of the splitting (the segmentation of a region does not depend on other regions) leads to a significant reduction of the computational complexity with respect to the case where a flat C-class MRF is used.

3. The Region Classification

After the segmentation step, the mammogram is subdivided in a huge number of homogeneous elementary regions. In order to identify the actual microcalcifications a classification phase is needed. To this aim, for each region both geometrical and textural features are extracted according to the characteristics of the microcalcifications. Features considered are:

Area: number of pixels in a single region.

Perimeter: number of pixels on the contour of a region.

Compactness: Area/(Perimeter)^2.

X Axis: maximum width along the horizontal axis.

Y Axis: maximum height along the vertical axis.

Elongation: (X Axis)/(Y Axis).

Mean: the value μ_i of the i-th class to which the region belongs. This feature is provided directly by the segmentation process.

Variance: the value σ_i of the i-th class to which the region belongs. This feature is provided directly by the segmentation process.

Gradient on the original data: minimum, maximum and average value of the gradient strength on the region

contour pixels. Such feature is strictly related to the single region, but it is really significant only if the original data are sharp enough on the border of the region.

Gradient on the mean: minimum, maximum and average value of the gradient evaluated as in the previous feature but in this case the value of the pixels is equal to the mean μ_i of the class which the region belongs to. In this way we obtain a feature related both to the class and to the region, which is significant even if the border of the region is not sharp.

The analysis of the regions is performed by means of a two-stage, coarse-fine classification. Such scheme allows us to discard rapidly the regions less likely to be microcalcifications and to employ more sophisticated classifier and features only on the most difficult cases, thus improving the effectiveness of the whole classification process. Through the first stage all the regions with size, shape or brightness strongly different from those typical for a microcalcification are eliminated. In particular, a first criterion adopted in this phase discards all the regions with horizontal or vertical size greater than 10 pixels. In this way, we eliminate all the regions greater than 1.0 mm in at least one dimension: such regions cannot be microcalcifications because their size is excessive. In a similar way, also the regions with *gradient on the mean* less than or equal to zero are discarded. These regions, in fact, have a grey level very similar to their neighbourhood or even darker: this excludes the presence of a microcalcification, which should appear as a bright spot.

After the first stage, we keep only the regions with characteristics acceptable for a microcalcification, but these still contain a large number of erroneous detections. In order to distinguish the actual microcalcifications, a further, more refined classification system is needed which, on the basis of the features previously described, is able to correctly discriminate between true and false detections. To this aim, Support Vector Machines represent one of the best two-class classifiers now available. SVM map the set of samples from the R^n input space to a high-dimensional feature space F and search the hyperplane separating the two classes with a maximum margin. To this aim, the learning algorithm of SVM considers a training data set containing, say, m samples u_i, each described by an n-dimensional vector; the samples are assigned to corresponding labels $l_i = \pm 1$, where the sign of the label indicates the class. The decision for a new sample z to be classified is based on the function:

$$f(z) = \frac{1}{2}\sum_{i=1}^{m} \alpha_i l_i K(z, u_i) + b \qquad (4)$$

where the parameters α_i and b are evaluated during the learning phase and $K(u,v)$ is a particular function, called *kernel*, which evaluates the scalar product between the projections of the two n-dimensional vectors u and v in the feature space F. It is worth noting that $f(z)$ provides the signed distance of the point z from the optimal separating hyperplane; hence the decision is made according to the sign of $f(z)$.

4. Experimental Results

The system has been tested on a standard database, publicly available on Internet, provided by courtesy of the National Expert and Training Centre for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, the Netherlands. It contains 40 digitized mammographic images composed of both oblique and craniocaudal views from 21 patients. Each mammogram has one or more clusters of microcalcifications marked by radiologists. The total number of clusters is 105, 76 of which are malignant and 29 benign. All images have a size of 2048x2048 pixels and use 12 bits per pixel for the gray levels. A preprocessing phase was applied to convert them into an 8 bit/pixel format by using the adaptive noise equalization described in [1]. The images obtained were firstly segmented using the TS-MRF algorithm with a number of classes equal to 80, thus achieving approximately 20,000 regions for each image. After the segmentation the coarse classification was applied, reducing in this way from 20,000 to 700, on average, the number of regions given in input to the SVM. The classifiers used in our experiments have been implemented by means of SVM^{light} tool [4], available at http://svmlight.joachims.org.

The low number of clusters in the database made very difficult the training phase of the SVM. For this reason, we adopted a *leave one out* approach for our experiments. According to this procedure we accomplished 40 different training procedures, one for each image, each time using a different training set. To build it we assumed as positive samples all the microcalcifications we detected with the segmentation (according to the ground truth of the database) in 39 images and, as negative samples, a number of false regions, equal to the number of positives, chosen randomly among all the negatives that we had in input to the SVM. Then, for each training set we performed a test on the image we did not consider before. For our experiments we used the polynomial kernel $K(u,v) = (\langle u \cdot v \rangle + 1)^2$; moreover, all features given in input to the SVM were previously rescaled so as to have zero mean and unit standard deviation.

The SVM performed very well on the true microcalcifications with a drop of 98% on the number of false regions. Nevertheless we still had some false regions per image. However, what is of medical interest in the breast cancer is the detection of microcalcifications clusters. To this aim, we adopted the criterion used in [1],

that is to consider the presence of a cluster if there are three or more microcalcifications within a region of 1 cm^2. So, in order to detect microcalcification clusters, we did not consider isolated regions that are usually produced by scratches or emulsion errors.

Groups of detected regions consistent with the ground truth in the database were considered true positive (TP) clusters, otherwise they were counted as false positive (FP) clusters. We have compared our experimental results with those reported in [1], [5] and [6] since in these papers the same set of mammograms was used. In particular, [1] also employs an MRF model, while both [5] and [6] are based on a scale space approach; the latter also introduces a fuzzy logic based enhancement of the mammogram as a preprocessing step. Table 1 shows the results obtained.

Table 1. The results obtained on the database.

Img	N	True Positives				False Positives			
		TMS	*KAR*	*CHE*	*NET*	*TMS*	*KAR*	*CHE*	*NET*
c01c	3	3	2	3	3	0	1	0	0
c01o	3	3	2	3	3	0	0	0	0
c02c	2	2	2	2	2	0	1	0	0
c02o	1	1	1	1	1	1	2	0	3
c03c	2	1	1	2	1	0	0	0	0
c03o	1	1	1	1	1	0	1	0	0
c04c	2	2	2	2	2	0	0	0	1
c04o	2	2	2	2	2	0	0	2	0
c05c	1	1	1	1	1	0	1	0	1
c05o	2	2	2	2	1	0	0	0	0
c06c	3	2	2	2	3	0	0	2	1
c06o	2	2	2	1	2	0	3	0	0
c07c	1	1	1	1	1	0	0	0	2
c07o	1	1	1	1	1	0	1	0	1
c08c	4	3	3	3	3	0	0	0	0
c08o	6	2	4	4	6	0	1	4	0
c09c	1	1	1	1	1	0	1	0	0
c09o	2	1	2	2	1	1	1	0	1
c10c	1	0	1	1	1	0	1	1	0
c11c	1	1	1	1	1	2	8	0	4
c11o	1	1	1	1	1	0	2	0	1
c12c	15	11	9	13	13	2	2	3	1
c12o	13	9	11	12	11	2	2	0	1
c13c	1	1	1	1	1	0	0	0	1
c13o	1	1	1	1	1	0	0	0	0
c14c	2	2	2	2	2	0	0	0	0
c14o	2	1	1	2	1	0	0	0	1
c15c	1	1	1	1	1	1	1	0	0
c15o	1	1	1	1	1	1	1	0	1
c16c	1	1	1	1	1	0	0	0	0
c16o	1	1	1	1	1	0	1	0	0
c17c	9	7	9	7	8	0	1	1	2
c17o	5	3	4	5	5	2	2	1	1
c18c	2	2	2	2	2	0	1	0	1
c18e	1	1	1	1	1	0	0	0	0
c18o	1	1	1	1	1	0	4	0	1
c19c	2	2	2	2	2	0	0	0	1
c19o	3	3	3	3	1	0	0	0	0
c20c	1	1	1	1	1	0	0	0	8
c21o	1	1	1	1	1	0	3	0	5
Total	105	83	88	95	93	12	42	14	39

The first two columns contain the label of the mammogram in the database and the respective number of clusters; the successive columns show the results obtained by the four methods compared in terms of false positives and true positives. We have denoted with *TMS* the method here proposed, while *KAR*, *CHE* and *NET* designate the results obtained in [1], [6] and [5] respectively. We can observe that the proposed method does not reach the same results in the detection of TP clusters but gives the best results in terms of false positives. In particular, the comparison with the other MRF-based method shows how our method provides results slightly worse in terms of true positives, but significantly better for the false positives. Also the comparison with [5] shows the same behavior, while [6] exhibits a higher number of true positives and an amount of false positives near to what we obtain with our method. It is worth noting, however, that in [6] the tests have been performed in selected areas containing all the clusters of the image, while in all our tests the whole mammograms have been used. Such situation is more realistic, even though less favorable for the MRF-based segmentation.

On account of these first encouraging results, future work will be aimed to improve the classification phase by including new features and by introducing a more suitable cluster validation method based also on topological criteria.

References

[1] N. Karssemeijer, "Adaptive Noise Equalization and Recognition of Microcalcification Clusters in Mammograms", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, 1993, pp. 1357-1376.

[2] C. D'Elia, G. Poggi and G. Scarpa, "A Tree-Structured Markov Random Field Model for Bayesian Image Segmentation", *IEEE Trans. on Image Processing* , vol. 12, no. 10, 2003, pp. 1259-1273.

[3] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*, Cambridge Univ. Press, , 2000.

[4] T. Joachims, "Making large-scale SVM learning practical", in B. Schölkopf, C.J.C. Burges, A.J. Smola, eds., *Advances in Kernel Methods*, MIT Press, 1999, pp. 169-184.

[5] T. Netsch and H. Peitgen, "Scale-Space Signatures for the Detection of Clustered Microcalcifications in Digital Mammograms", *IEEE Trans. on Medical Imaging*, vol. 18, no. 9, 1999, pp. 774-786.

[6] H.D. Cheng, J. Wang and X. Shi, "Microcalcification Detection Using Fuzzy Logic and Scale Space Approach", *Pattern Recognition*, vol. 37, 2004, pp. 363-375.