A SUPERVISED MICRO-CALCIFICATION DETECTION APPROACH IN DIGITISED MAMMOGRAMS

Albert Torrent, Arnau Oliver, Xavier Lladó, Robert Martí & Jordi Freixenet

Dept. of Computer Architecture and Technology
University of Girona
17071, Girona, Spain
{atorrent,aoliver,llado,marly,jordif}@eia.udg.edu

ABSTRACT

We present in this paper a supervised approach for automatic detection of micro-calcifications. The system is based on learning the different morphology of the micro-calcifications using local features, which are extracted using a bank of filters. Afterwards, this set of features is used to train a pixel-based boosting classifier which at each round automatically selects the most salient one. Therefore, when a new mammogram is tested only the salient features are computed and used to classify each pixel of the mammogram as being part of a micro-calcification or actually being normal tissue. The experimental results shows the validity of our approach. Moreover, the robustness of our method is also demonstrated using a digitised database for the learning process and a different one for the testing, providing satisfactory results.

Index Terms— Biomedical image processing, Computer aided detection, Mammography

1. INTRODUCTION

Breast cancer continues to be a significant health problem in the world. It constitutes the most common cancer among women in the European Union [1], and it is estimated that between one in eight and one in twelve women will develop breast cancer during their lifetime [2]. Mammography is the most effective and reliable method for an early detection of breast cancer which is fundamental for improving prognosis. Mammographic images are characterised by high spatial resolution allowing the detection of subtle scale signs such as micro-calcifications and masses. In this work, we focused on the detection of micro-calcifications, which are tiny granular deposits of calcium that generally appear in a mammogram as small bright spots within an inhomogeneous background. Fig. 1 shows two mammograms with (easy-to-find) micro-calcifications (extracted from the MIAS database [3]).

In the past and recent years there has been a considerable interest in developing methods for automatic detection of micro-calcifications in mammograms [4, 5]. However, none of them has emerged as a standard algorithm. We present here a new approach for micro-calcifications detection based on extracting local features for describing and learning the micro-calcifications morphology. Once the system is trained, new mammograms are classified pixel-by-pixel and suspicious areas are detected. The reliability of our approach is proved in the experimental section, using different databases for training and testing the system.

The rest of this paper is structured as follows. Section 2 describes the proposed approach. Section 3 describes the testing methodology, the parameter optimisation, and the obtained results. Finally, the paper ends with the conclusions and further work.

2. MICRO-CALCIFICATIONS DETECTION

The presented approach for micro-calcification detection is based on the work of Murphy et al. [6] for object detection using local features and boosting classifier. As is shown in Fig. 2, the approach is divided in three parts. Firstly, we create a visual word dictionary, which is composed by convolving patches containing a micro-calcification with a bank of filters. Secondly, the training data is found by convolving positive samples (patches containing a micro-calcification) and negative samples (patches of other tissues) with the words of the dictionary defined as the duple patch-filter. Finally, new mammograms are classified pixel-by-pixel using the classifier. Hence, the detection problem is translated to a pixel-based classification approach.

In the following subsections we describe in more detail the three parts of our approach.

2.1. Building the dictionary

The first task of the system consists in building the feature dictionary. This dictionary is similar to an atlas, since it con-
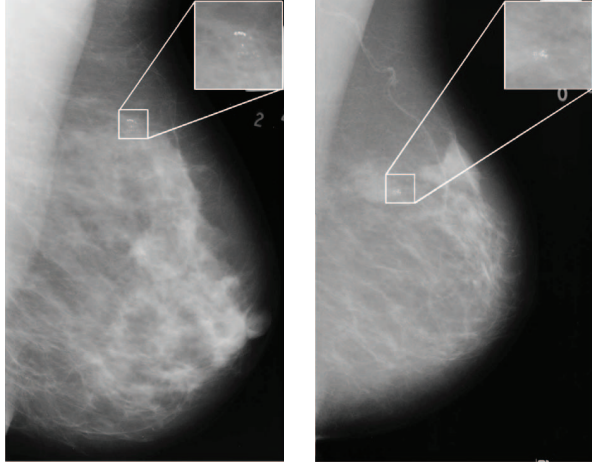
Fig. 1. Two mammograms containing micro-calcifications. Both examples were selected for a good visualisation of the problem, although in general micro-calcifications are more subtle and difficult to appreciate, even for experts radiologists.

tains samples (patches) of micro-calcifications. However, it also contains the convolution of these patches with a bank of filters, including a delta function (which gives the own patch as a result), 4 Gaussian derivatives, a Laplacian filter, a corner detector, and 2 Sobel filters. Hence, the dictionary contains grey-level and gradient information of the micro-calcifications and their neighbourhood. Therefore, each dictionary word w_{ij} can be understood as the duple (p_i, f_j), where p_i represents the patch and f_j the filter.

2.2. Training step

Once the dictionary is built, the words are used to extract the mammographic features that will be used for the micro-calcifications detection. For the training step, we need a database of patches containing instances of both patches with micro-calcifications (positive training examples) and patches from the rest of the mammographic tissues (negative training examples). Note that in the original approach [6], each pixel of the training images was used as the centre of a positive or a negative patch but in our case, this implies a high computational cost due to the large size of the mammograms (the small size of micro-calcifications prevents subsampling the mammograms, which are for example 4320×2600 pixels). Hence, for reducing this cost, we only select some points in each training image. In particular, we select the centre of the micro-calcifications (positive training examples) and some random locations of the background containing examples of different tissues.

The feature extraction of each training image patch consists in two operations. Firstly, the patch is convolved with all the bank filters, and secondly, the normalised cross-correlation with all the words is computed. Mathematically,

both operations can be summarised as:

$$v = (I * f_j) \otimes w_{ij} \qquad (1)$$

where I is the training image patch convolved ($*$) with the filter f_j and cross-correlated ($\otimes$) with the word w_{ij} (note that $w_{ij} = p_i * f_j$). The resulting value v represents the similarity of the training patch and the dictionary word. Therefore, for each training image patch, a vector of features v is constructed by cross-correlating all the dictionary words w_{ij} with the convolution of the patch itself with filter f_j. Notice here the necessity of keeping the filter as well as the patch in the dictionary word.

At this point, the positive and negative training examples have been characterised. Therefore, this data can be used to train a classifier. In this work, we have used the Gentleboost algorithm [7]. Boosting algorithms are based on the simple idea that the sum of weak classifiers can produce a strong classifier. In the Gentleboost algorithm, the weak classifiers (h_t) are simple regression stumps with one of the features, so at each round t the feature with less error is selected. The weak classifier used is:

$$h_t(x) = a\delta(x_i > th) + b \qquad (2)$$

where th is a threshold that determines if pattern x belongs to the object class, x_i is the i'th dimension of x, and a and b are parameters selected to minimise the error of the classifier (a is the regression slope and b the offset):

$$e = \sum(z(y - (a\delta(x_i > th) + b))^2) \qquad (3)$$

At each round the training data weights (z) are updated, increasing in the following round the possibility of classifying correctly the previous incorrectly classified points. In the GentleBoost algorithm the data weights are updated using:

$$z_{t+1} = z_t e^{y \cdot h_t(x)} \qquad (4)$$

Hence, when testing a new data, the final (strong) classifier is computed using the weak classifier created at each round of the boosting. Therefore, the testing data is classified according to the sign of the sum of weak classifiers:

$$H(x) = \sum h(x) \qquad (5)$$

The absolute value of $H(x)$ shows the confidence of the classified data.

2.3. Testing step

Once the classifier is built, the system is ready for the testing step, where the strong pixel-based classifier $H(x)$ is applied to new images in order to evaluate the micro-calcifications detection. Notice that the result of our approach after evaluating a mammogram is a probability image, where high values represents more confidence to be a micro-calcification.
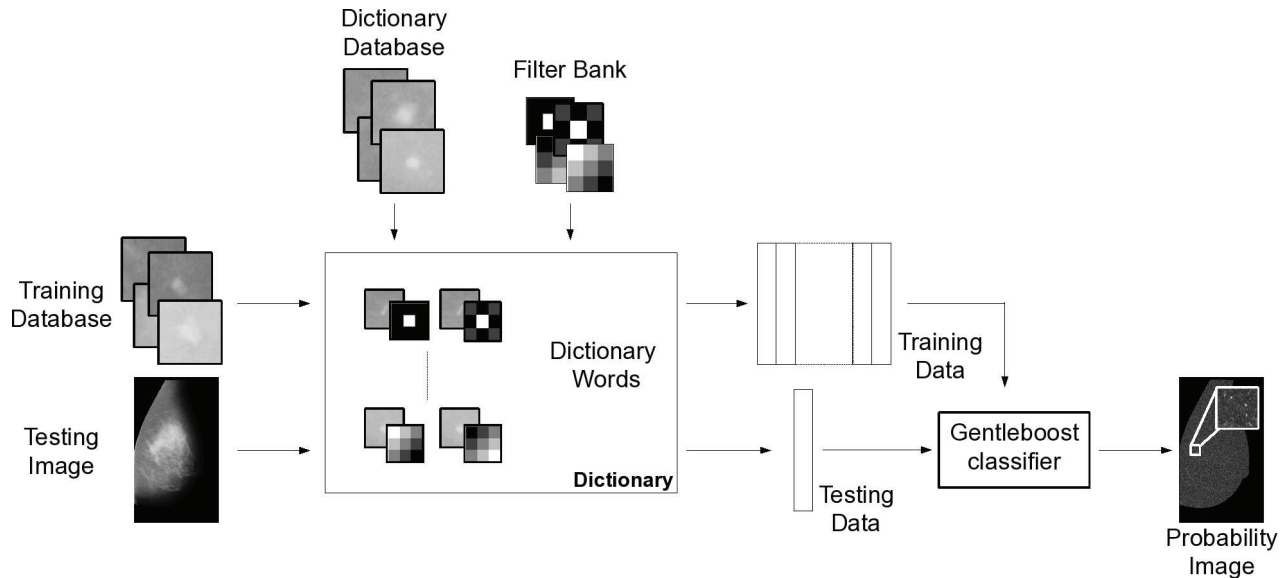
4346

Fig. 2. Schematic representation of our approach. Note that the result of testing a new image is a probability image, where suspicious regions are highlighted over the rest of the image.

3. RESULTS

The experimental results were performed using two different subsets of mammograms. The first subset of 112 mammograms was extracted from the MIAS database, and contained all the mammograms with micro-calcifications (22 in total) and a set of 90 normal mammograms. This subset was used for training and testing the algorithm. In order to train the classifier with positive examples, an expert accurately marked among 5 and 15 micro-calcifications in each mammogram containing micro-calcifications, while the negative examples were obtained from the rest of tissues using around 20 marks in all mammograms. The second subset of mammograms was extracted from the DDSM database, containing a total of 1141 mammograms, 386 containing micro-calcifications and 755 being normal ones. From this dataset we did not have accurate individual micro-calcifications annotations, and therefore we could not use this database for the training step. Hence, this dataset was only used for testing purposes.

The evaluation of the results was done using ROC analysis [8]. In this analysis, a graphical curve represents the true positive rate (number of detected mammograms with micro-calcifications divided by the total number of mammograms with micro-calcifications) as a function of the false positives rate (number of normal mammograms incorrectly detected as containing micro-calcifications divided by the total number of normal mammograms). Moreover, the percentage value under the curve (A_z) is an indication for the overall performance of the observer, and is typically used to analyse the performance of the algorithms.

We optimised the parameters involved in our detection ap-

proach using the first subset of mammograms and a 10-folder cross-validation methodology. The 112 images were divided into 10 different groups, from where 1 was used to create the dictionary, 8 groups were used to train the boosting classifier, and the remaining group was used for testing. This procedure was repeated until all image groups were used for testing. Notice that using this methodology each image appears in the test set only once. The best results were achieved using around 500 visual words for describing the different micro-calcifications morphology, providing the ROC curve shown in Fig. 3 ($A_z = 0.88$). The computational time of the whole process dramatically increased when increasing the number of patches and words used for building the dictionary, and the empirical values used here provided a good trade-off between performance and feature vector length.

With the same set of parameters and the same training dataset (1 group was used to create the dictionary while the remaining 9 were used to train the classifier) we also tested the second subset of mammograms. The obtained ROC curve is shown in Fig. 3, where we obtained $A_z = 0.70$. Note that worse results were obtained compared with the first dataset. This is due to the fact that we are training the classifier with mammograms obtained from a different database. Moreover, note that we are testing a large dataset of mammograms using a small training set, which probably do not generalise all the micro-calcifications morphology (note that we are testing 386 mammograms containing micro-calcifications with a training set containing just 22 mammograms with micro-calcifications). However, taking both limitations into account, we consider the obtained results as promising.

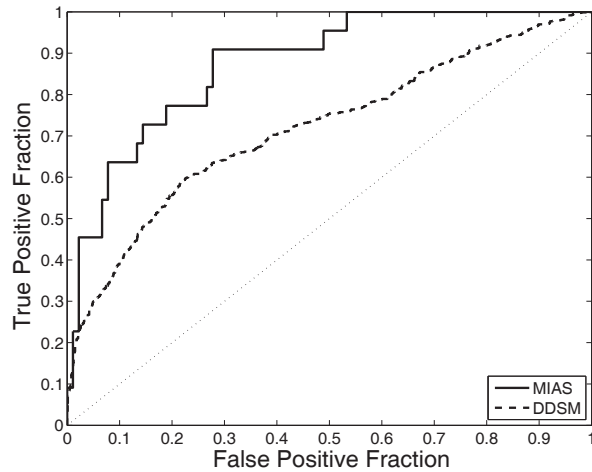Finally, we compare in Table 1 the obtained results with

Fig. 3. Obtained ROC curves when testing both the MIAS (solid line) and DDSM (dashed line) datasets. The dotted line indicates random guess.

those obtained by current state-of-the-art approaches. These approaches are usually based on two steps. Firstly the detection of suspicious regions is performed, usually tuning the algorithm parameters in order to detect as many suspicious regions as possible (i.e., detecting the largest number of micro-calcifications but also increasing the probabilities to detect normal tissue as being a micro-calcification). Secondly, a false positive reduction algorithm is applied in order to find those detected suspicious regions being in fact normal tissue. Comparing the obtained results using the MIAS database note that they are inline with the ones reported by current approaches. However, we want to stress here that our approach do not include a false positive reduction step, and hence our results are obtained directly from the detection step.

4. CONCLUSION

We have presented a new approach for micro-calcifications detection based on extracting local features for characterising the morphology of the micro-calcifications. The proposed boosting approach allows the selection of the most salient features at each round, reducing the computational time of the testing step. The performed experiments have shown the validity of our proposal, even when training with mammograms from a different database than the testing one.

Further work is directed in two directions. Firstly, we would like to integrate a false positive reduction step into the boosting algorithm to improve the results, and secondly, we want to test our approach using a digital database. We expect that the use of this technology will also improve the results, since it is well known that digital mammography improves the contrast between the different internal structures [9].

Table 1. Comparison of the obtained experimental results with state-of-the-art algorithms for micro-calcification detection in digitised mammograms.

Authors	Cases	Results (A_z)
Chang et al. (2008) [10]	194	0.90
Nunes et al. (2007) [11]	121	0.93
Papadopoulos et al. (2008) [12]	60	0.92
Our approach – MIAS	**112**	**0.88**
Our approach – DDSM	**1141**	**0.70**

5. REFERENCES

[1] Eurostat, *Health statistics Atlas on mortality in the European Union*, Office for Official Publications of the European Union, 2002.

[2] F. Bray, P. McCarron, et al., "The changing global patterns of female breast cancer incidence and mortality," *Breast Cancer Research*, vol. 6, no. 6, pp. 229–239, 2004.

[3] J. Suckling, J. Parker, et al., "The Mammographic Image Analysis Society digital mammogram database," in *Int. Work. Dig. Mammography*, 1994, pp. 211–221.

[4] H. D. Cheng, X. Cai, et al., "Computer-aided detection and classification of micro-calcifications in mammograms: a survey," *Pattern Recogn.*, vol. 36, no. 12, pp. 2967–2991, 2003.

[5] R. M. Rangayyan, F. J. Ayres, et al., "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *J. Frankl. Inst.*, vol. 344, no. 3–4, pp. 312–348, 2007.

[6] K. Murphy, A. Torralba, et al., "Object detection and localization using local and global features," in *Sicily W. on Object Recognition*, 2005.

[7] J. Friedman, T. Hastie, et al., "Additive logistic regression: a statistical view of boosting," *Ann. Stat.*, vol. 38, no. 2, pp. 337–374, 2000.

[8] C. E. Metz, "Evaluation of digital mammography by ROC analysis," in *Int. Work. Dig. Mammography*, 1996, pp. 61–68.

[9] U. Fischer, F. Baum, et al., "Comparative study in patients with micro-calcifications: full-field digital mammography vs screen-film mammography," *Epidemiol. Rev.*, vol. 11, no. 12, pp. 2679–2683, 2002.

[10] T. T. Chang, J. Feng, et al., "Clustered micro-calcification detection based on a multiple kernel support vector machine with grouped features," in *IAPR Int. Conf. Pattern Recogn.*, 2008, pp. 1–4.

[11] F. L. S. Nunes, H. Schiabel, et al., "Contrast enhancement in dense breast images to aid clustered micro-calcifications detection," *J. Digit. Imaging*, vol. 1, no. 20, pp. 53–66, 2007.

[12] A. Papadopoulos, D. I. Fotiadis, et al., "Improvement of micro-calcification cluster detection in mammography utilizing image enhancement techniques," *Comput. Biol. Med.*, vol. 10, no. 38, pp. 1045–1055, 2008.