MASS DETECTION IN MAMMOGRAMS COMBINING TWO METHODS

N. Tóth, G. Takács, B. Pataki

Department of Measurement and Information Systems,
Budapest University of Technology and Economics,
Magyar tudósok körútja 2, H-1117 Budapest, Hungary

ntoth@mit.bme.hu

Abstract: Breast cancer is one of the most common forms of cancer among women. Currently mammography (X-ray examination of the breast) is the most efficient method for early detection. Two independently developed methods for detecting cancer-indicating masses in mammograms are presented in this paper. The first method is based on pixel intensities, the second one on texture features. The possible combinations of the two different approaches are investigated to achieve better mass detection rate with less false warnings. The composite system was tested with 523 mammographic cases, each containing 4 images.

Introduction

Breast cancer is one of the most common forms of cancer among women. Every 12th woman suffers from this disease at least once in her lifetime [1]. This means more than 300 000 diagnosed new cases per year in Europe [2]. Since the cause of the disease is unknown, early detection is very important, for which currently mammography (X-ray examination of the breast) is the most efficient method. If breast cancer is detected early, the five-year survival rate exceeds 95% [1].

The evaluation of images taken at mammographic screening examinations needs a large amount of human resource and money. Therefore computer-aided diagnosis (CAD) for mammography is an active area of research (e.g. [3], [4], [5]). The main goals of a CAD system are to (i) increase the accuracy of examination by aiming radiologist's attention to suspicious cases and to (ii) decrease the cost by filtering out normal cases.

In a mammographic session usually two images are taken of both breasts. Craniocaudal (CC) is a top view, mediolateral (ML) is roughly a side view image of the breast., Radiologists typically notice suspicious-looking structures in one view and then verify their suspicion by checking the corresponding area of the other view of the same breast.

The most important mammographic symptoms of breast cancer can be divided into two main classes:

- Microcalcification: a group of small white calcium spots.
- Mass: usually approximately round object brighter than its surrounding tissue.

Not all microcalcifications and masses are cancerous, both can be benign, too. The two main classes can be divided into subclasses, for example the ACR (American College of Radiology) BI-RADS recommendation [6] defines 9 mass and 13 micro-calcification types. Combined mass-microcalcification lesions are also possible.

Figures 1-4 show some typical forms of microcalcifications and masses appearing in mammograms:

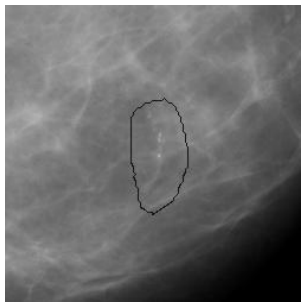


Figure 1: A typcal benign microcalcification

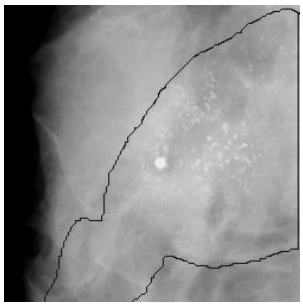


Figure 2: A typical malig-nant microcalcification

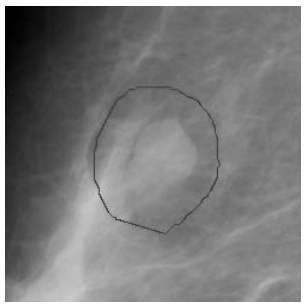


Figure 3: A typical benign mass

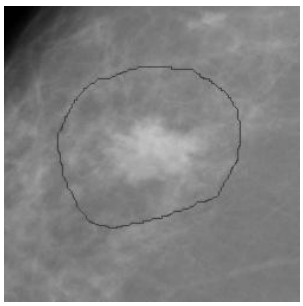


Figure 4: A typical malig-nant mass

In our work we examined the mass detection subproblem of breast cancer detection by independently developing a texture-based and an intensity-based mass detection algorithm and then combining the two approaches.

Intensity-Based Algorithm

The first mass detection algorithm is denoted as intensity-based, because it works directly with pixel intensities of the input image and uses only simple, low-level features (e.g. average intensity of some pixels). The advantage of such an algorithm is obviously computational efficiency, because it does not need any transformation or complex feature extraction step.

ISSN: 1727-1983

The size of a mammographic mass can vary in a wide range (~5 mm – ~50 mm). The way of handling this size variability is an interesting question of computerized mass detection. The intensity based mass detector defines two size classes ("small" and "large") and uses different methods for each. Most real-life mammographic masses belong to the "small" class so the small mass detector is the critical part of the system. Masses belonging to the "large" class are rare and usually easy to detect. Therefore the large mass detector is a less-frequently-used, simpler, but necessary part.

Detection of Small Masses

For the detection of small masses (smaller than 20 mm in diameter) a variant of the AFUM mass detection algorithm [7] was applied. At each pixel position (x, y) the minimal intensity value at distance r_1 from location (x, y) is computed (m_1), then the fraction of pixels at distance r_2 from (x, y) that have lower intensity value than the m_1 is measured. This fraction under the minimum (FUM) calculation is done over many scales using a range of r_1 and r_2 values and the average of those calculations yields the average FUM (AFUM) value.

This AFUM algorithm variant slightly differs from the original one, because in the original algorithm the minimal intensity at distance *less than or equal* r_1 from (x, y) is compared to intensity values at distance r_2 from (x, y). In real mammograms some masses contain small dark dots inside. The original AFUM algorithm prohibits this case while the proposed variant tolerates it in some degree.

If $r_1 = R_{min}$, $R_{min} + 1$, $R_{min} + 2$, ..., R_{max} and $r_2 = r_1 + D$ then AFUM value calculation can be written as:

$$\frac{1}{R_{max} - R_{min} + 1} \cdot \sum_{r_1 = R_{min}}^{R_{max}} FUM(r_1, r_1 + D) \qquad (1)$$

Values R_{min}, R_{max} and D are fixed a priori choices based on the problem definition.

A fast small mass detector can be obtained by running the AFUM algorithm for each non-background pixel of a mammographic image. The result of the AFUM algorithm is an output (filtered) image having the AFUM values of the original pixels at each position. The filtered image is thresholded and continuous regions are identified by a region-filling algorithm. Regions with a too high perimeter-area ratio are excluded from further examinations. The location of the maximal AFUM value is computed for each region, and an "energy" value is assigned for each maximum location based on the AFUM value of that position and its neighbouring pixels. A structure is accepted as a mass if this energy is higher than a limit. (The energy limit parameter can be used to adjust the sensitivity of the small mass detector.) Finally the locations of the highest N energy maxima are returned (Figure 5).

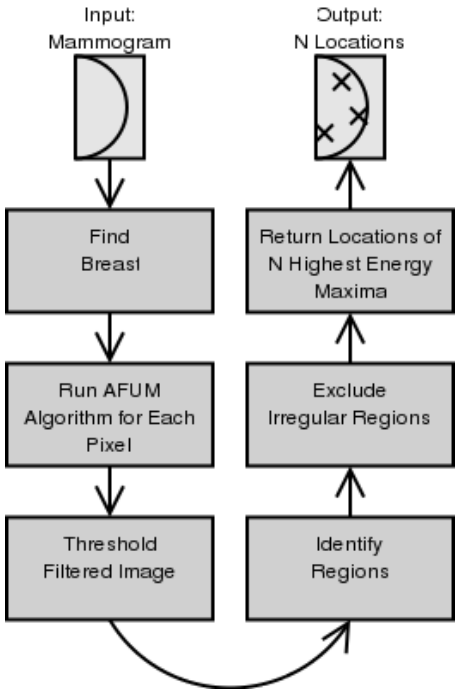


Figure 5: Block diagram of the small mass detector

Detection of Large Masses

The a priori parameters of the AFUM algorithm (R_{min}, R_{max} and D) could not be set to deal with arbitrary mass size. At the resolution of 400 microns $R_{min} = 0$, $R_{max} = 6$ and $D = 12$ seemed to be a good choice but worked well only for masses smaller than 20 mm in diameter. For the detection of larger masses a different algorithm was developed:

At each pixel position 8 lines are started from the center (vertically, horizontally and diagonally) and a mass boundary point is assessed for each direction (Figure 6).

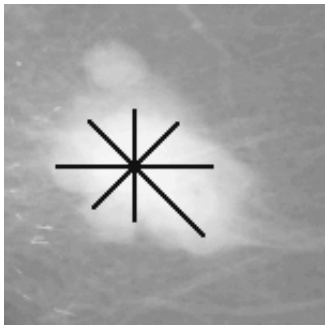


Figure 6: Example result of mass boundary assessment

Then a "conspicuousness" value can be obtained from line lengths (l_i), average intensity along the lines (*Brightness*) and average contrast at the end of the lines (*Contrast*) with the following heuristic formula:

$$Conspicuousness = (\min_i l_i) \cdot Brightness \cdot Regularity \cdot Contrast \qquad (2)$$

ISSN: 1727-1983

$$Regularity = (\min_i l_i) \big/ (\max_i l_i) \quad (3)$$

Since the detection of large masses is usually easier than the detection of small ones, the large mass detector returns only the location of the highest conspicuousness value if processing a whole mammogram. Obviously a conspicuousness limit can be applied to control the sensitivity of the large mass detector (Figure 7).

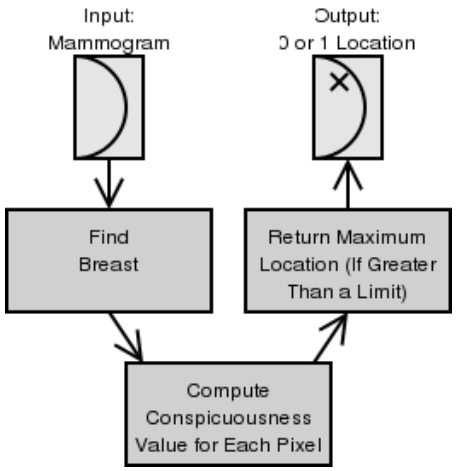


Figure 7: Block diagram of the large mass detector

The intensity-based mass detector

The combination of the previous two methods yields the intensity-based mass detector that can handle arbitrary mass size. Another advantage of the intensity-based mass detector is that it is fast enough to process each pixels of a mammogram (at 400 microns resolution).

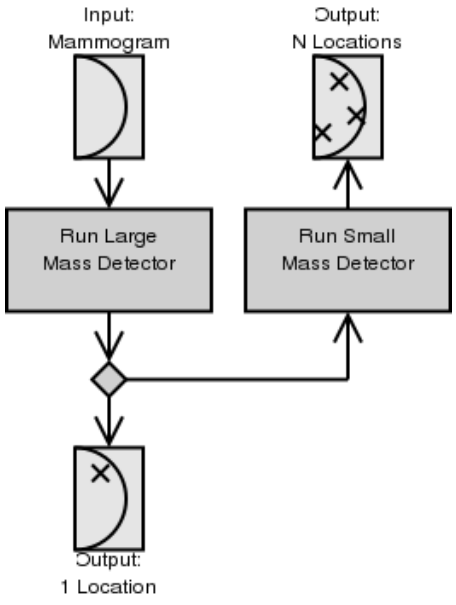


Figure 8: Block diagram of the intensity-based mass detector

The combination of the two parts was implemented in the following way (Figure 8):

i. Run the large mass detector for the input image.
ii. If the conspicuousness value of the returned location is greater than a given value, stop processing and return this location as the result of the intensity-based mass detector.
iii. If the conspicuousness value of the returned location is smaller than a given value, run the small mass detector for the input image.
iv. Return the small mass detector's output locations in addition to the large mass detector's output location as the result of the intensity-based mass detector.

Texture-Based Algorithm

The second mass detection algorithm is denoted as texture-based, because it works with texture features extracted from pixel intensities around a pixel's neighbourhood in the input image. This algorithm is computationally more expensive than the intensity-based algorithm, because of the derivation of texture features. These are high level features, such as texture variation or coarseness. This segmentation method gives information about the location of the most conspicuous mass tissues in the breast and some information of the shape as well . This texture-based localization of masses in the mammograms consists of two main parts: "coarse" and "fine" segmentation.

"Coarse" segmentation

This step calculates various texture parameters in a particular sample window to create the feature vector. This window (in which the texture features are calculated) is sliding through the entire breast image. The feature vector contains 17 texture parameters. Four of them are based on the histogram [8], four of them on the co-occurrence matrix [8], four of them on the gray level run length [9] and the last five on the gray level differences histogram [9] (Table 1).

Table 1: Used texture features and their references.

Feature Group	Feature	Reference
Histogram features	Mean Variance Skewness Kurtosis	[8]
Co-occurence matrix features	Varinace Contrast Homogenity Correlation	[8]
Gray level run length features	Short-run emphasis Long-run emphasis Gary-level distribution Run-length distribution	[9]
Gray level differences features	Mean Variance Contrast Skewness Kurtosis	[9]

ISSN: 1727-1983

The feature vector is classified by a set of decision trees [10]. Each one of the individual decision tree classifiers is generated from a training image. The set of the decision trees are generated from a set of training images using a form of the bagging [10] algorithm (Figure 9). The difference to the original bagging algorithm is that here each traning image is used only once to grow one decision tree. These training images – which are extracted from mammograms previously evaluated by physicians – contain a mass and some surrounding background tissue. From each training image the corresponding decision tree is generated using the CART (Classification and Regression Trees) [11] algorithm. The feature vectors – extracted at sample window positions – are used as training data for the decision trees. The tree growing process chooses the optimal test based on the Gini diversity index. Tests are made on the feature vector parameters. Each tree node uses 1 of 1 test (only one parameter is tested in a node). The optimal tree size is estimated using 10-fold cross-validation, minimal cost-complexity pruning and 1-SE rule [11] to prevent overfitting. Different trees – that are generated from different training images – will choose different texture features from the set to classify the image regions according to their training sample.

After the training phase the new mammogram is segmented using all the decision trees (Figure 9). The output is a voteboard where each image segment can have a vote value between zero and the number of the classifying trees (each tree votes '1' if it classifies the actual image segment as part of a lesion, '0' otherwise). The higher the vote number the greater the probability that the actual image region is a mass. Adaptive thresholding is applied to this voteboard to find the most suspicious regions. This step will binarize the voteboard, marking the locations where the vote value is greater than its surrounding's. This segmentation method gives a rough representation of the shape and localization of the most conspicuous mass tissues in the breast .

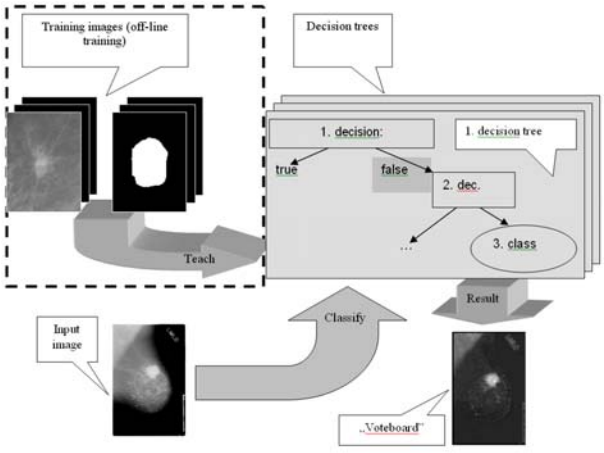


Figure 9: The schematic of the "coarse" segmentation classifier system.

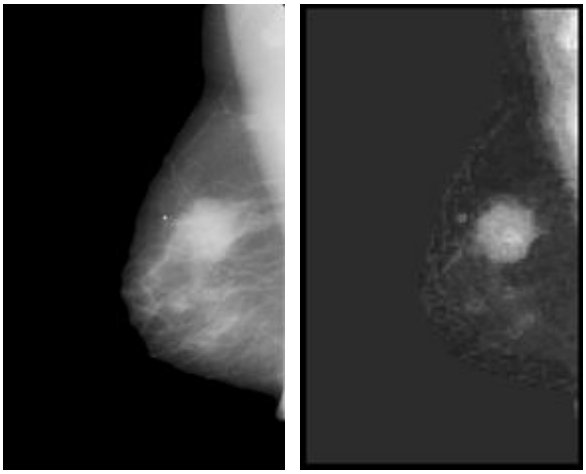
"Fine" segmentation

This forthcoming algorithm uses a Markov random field [12] [13] to improve the preliminary result provided by the "coarse" segmentation. This method models the segmented image as a Markov Random Field, where pixel class probabilities are defined by their neighbors' classes and intensities in the original image. The model chosen for the segmented image is the MultiLevel Logistic Field (MLL) [12]. It's advantage is simplicity, and the power to encode spatial patterns. First, using this model, at every spatial location *s* in the segmented image a potential value (*V*) is defined. This value is lower when the pixels in the neigborhood of *s* have the same class label as the pixel at *s*. Next, in this model the original image is derived from the segmented image with additive Gaussian noise [15]. As a result each pixel is more likely to have a class of its neighbors having intensities close to itself. The optimal segmentation is defined using the MAP (Maximum A Posteriori) principle [12]. Pixel class probabilities defined by MLL (MultiLevel Logistic Field) :

$$p(x_s = k \mid y_s) = e^{\{-\frac{[y_s - m_s(k)]^2}{2\sigma^2} - V(x_s)\}} \qquad (4)$$

where y_s is the pixel intensity at location *s*, x_s is the class of the pixel at location *s*, $m_s(k)$ is the average intensity of pixels around location *s* in a particular window (w_s) which have class label *k*.

An image with 1024x1024 pixels and 2 pixel classes has $2^{1048676}$ possible segmentation results. Finding the optimal one is prohibitively expensive in computation time. For this reason we use the ICM (Iterated Conditional Modes) algorithm [12] [13] to approximate the MAP (Maximum A Posteriori) segmentation of the image in several steps. Each step uses the result of the previous step as the input segmentation. In every step each pixel class is replaced by the most probable one defined by Eq. (4). This step creates a mask that contains the mass candidates with better size and shape representation (Figure 10).



ISSN: 1727-1983

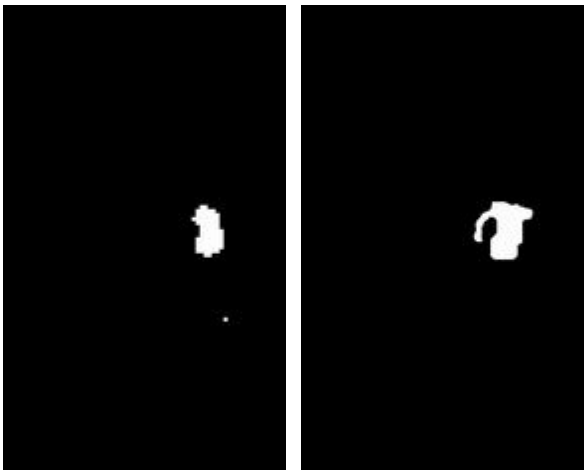


Figure 10: Original image, “voteboard”, “coarse” segmentation, “fine” segmentation.

Combination Possibilities of the methods

Each of the methods presented in this paper creates an output mask, covering the suspected mass locations. These locations are also marked with a conspicousness number. The combined result of the two algorithms is a mask where certain patches are kept from each algorithm.

Several combination possibilities were investigated. 131 images with known masses in each were selected for combination tests. The mass candidates that were marked by both algorithms are marked in the output image as well. Unfortunately if only those common masses are kept, the true positive ratio of the combined algorithm barely reaches 50% (Fig. 11).

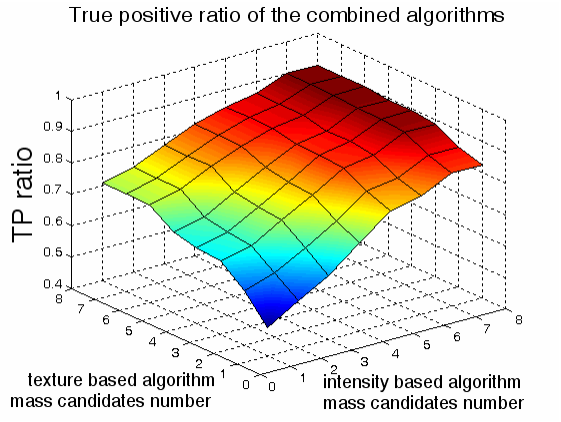


Figure 11a: Surface plot of the combined true positive ratio on test data set.

When a certain number of masses are kept – according to their conspicousness number –, plus the common ones, the true positive ratio increases (Fig. 11). In the optimal combination 6 mass candidates from the intensity based algorithm and 2 from the texture based algorithm are selected. Using this combination true positive ratio increases to 85% on the test data set. The search for the optimal combination was done on image level. As a result here true positive ratio is calculated for images. Later for validation purposes case level was used instead (see results and conclusion) as recommended by physicians. (A case contain 4 images of the two breasts.)

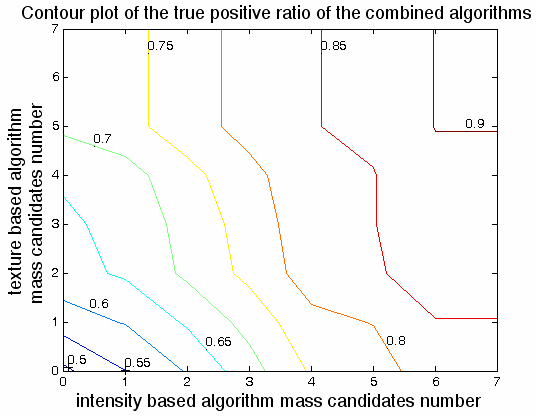


Figure 11b: Contour plot of the combined true positive ratio on test data set.

Further increasing the true positive ratio is prohibitively expensive in the means of false positive markers / image (Fig. 12).

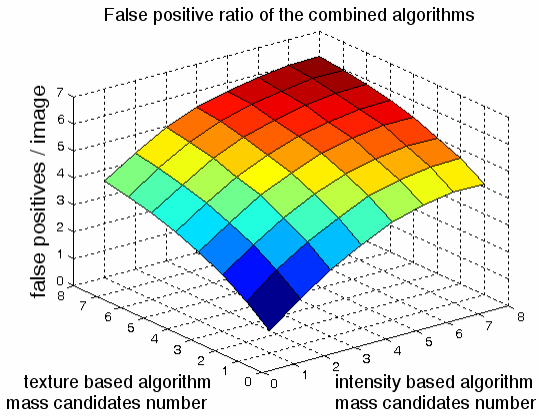


Figure 12a: Surface plot of the combined algorithm's false positive markers / image on test data set.

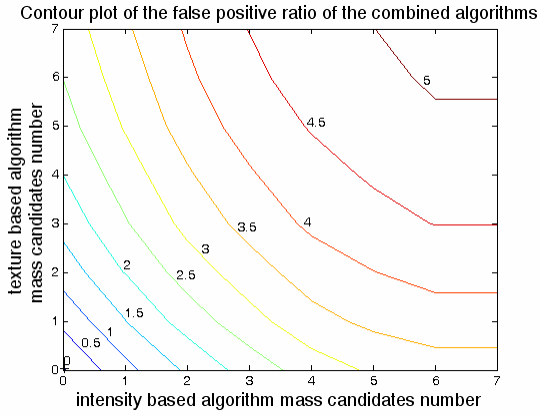


Figure 12b: Contour plot of the combined algorithm's false positive markers / image on test data set.

ISSN: 1727-1983

Results and conclusion

The combined algorithm was evaluated on 523 mammographic cases (523 x 4 = 2092 images) of the DDSM database [4]. 349 cases contained malignant, 8 cases benign masses and 166 cases contained no masses. A malignant case was counted as recognized if one of the marker centers returned by the combined mass detector was inside the radiologist-drawn boundary of the malignant mass in any image of the case. A marker was counted as a false positive if its center was not inside the radiologist-drawn boundary of any (malignant or benign) mass. (See Table 2)

Table 2: Results, MCRR = Malignant case recognition rate, FPM / image = False positive markers / image

	MCRR	FPM / image
Relative intensity based algorithm (setting 1)	92.8%	3.6
Relative intensity based algorithm (setting 2)	93.9%	4.5
Texture-based algorithm	70.0%	3.0
Combined algorithm	*95.1%*	*4.3*

According to an exhaustive study [14], the chance of detection of a tumor by a radiologist ranges from 98% to 48%, according to the breast density. In our test we used cases from all density categories at roughly equal number. Considering this, our results are promising.

Results show that the intensity based algorithm is more effective in mass localization, although there are a minor number of cases where texture-based information is required to identify the mass.

Acknowledgments

This work is sponsored by Research and Development Secretariat of the Hungarian Ministry of Education under contract IKTA 102/2001 and by the Hungarian Fund for Scientific Research (OTKA) under contract T046771.

References

[1] HIGHNAM, R. and BRADY, M. (1999, Eds): 'Mammographic Image Analysis', (Kluwer Academic Publishers)

[2] FERLAY, J., BRAY, F., PISANI, P. and PARKIN, D.M. (2001): GLOBOCAN 2000: 'Cancer Incidence, Mortality and Prevalence Worldwide, Version 1.0', (IARC Cancer Base No. 5., Lyon, IARC Press)

[3] LEE, S., CHUNG, P., CHANG, C., LO, C., LEE, T., HSU, G. and YANG, C. (2003): 'Classification of Clustered Microcalcifications Using a Shape Cognitron Neural Network', *Neural Networks*, **16**, pp. 121-132.

[4] CAMPANINI, R., DONGIOVANNI, D., IAMPIERI, E., LANCONELLI, N., MASOTTI, M., PALERMO, G., RICCARDI, A. and ROFFILLI, M. (2004): 'A novel featureless approach to mass detection in digital mammograms based on Support Vector Machines', *Physics in Medicine and Biology*, **49**, pp. 961-976

[5] ALTRICHTER, M., LUDÁNYI, Z., HORVÁTH, G. (2005): 'Joint Analysis of Multiple Mammographic Views in CAD Systems for Breast Cancer Detection', Accepted at SCIA 2005 – 14th Scandinavian Conference on Image Analysis, Joensuu, Finland, 2005

[6] AMERICAN COLLEGE OF RADIOLOGY (1998): 'Illustrated Breast Imaging Reporting and Data System (BI-RADS)', (3rd edition, American College of Radiology, Reston, VA)

[7] HEATH, M.D., BOWYER, K.W. (2000): 'Mass Detection by Relative Image Intensity', Proc. of IWDM 2000 – Fifth International Workshop on Digital Mammography, Toronto, Canada, 2000, pp. 219-255

[8] J. IIVARINEN (1998), 'Texture Segmentation and Shape Classification with Histogram Techniques and Self-Organizing Maps', *Acta Polytechnica Scandinavica*, No 95, TTA, Helsinki

[9] I. PITAS (2000), 'Digital Image Processing and Algorithms and Applications', (John Wiley & Sons, New York)

[10] RICHARD O. DUDA, PETER E. HART, DAVID G. STORK (2001), 'Pattern Classification 2nd ed'., (John Wiley and Sons, New York)

[11] L. BREIMAN, J. H. FRIEDMAN, R. A. OLSHEN, C. J. STONE (1984), 'Classification And Regression Trees', (Chapman & Hall)

[12] H. D. LI, M. KALLERGHI, L. P. CLARKE, V. K. JAIN (1995), 'Markov Random Field for Tumor Detection in Digital Mammography', IEEE Transactions on Medical Imaging, vol. 14, no. 3, pp. 565-576, September 1995.

[13] S. Z. LI (1995), 'Markov Random Field Modeling in Computer Vision', (SpringerVerlag, New York)

[14] THOMAS M. KOLB, MD, JACOB LICHY, MD AND JEFFREY H. NEWHOUSE, MD, 'Comparison of the Performance of Screening Mammography, Physical Examination, and Breast US and Evaluation of Factors that Influence Them: An Analysis of 27,825 Patient Evaluations', URL: http://radiology.rsnajnls.org/cgi/content/full/225/1/165?gca=225%2F1%2F165&sendit=Get+All+Checked+Abstract%28s%29&

ISSN: 1727-1983