# Automatic Labelling and BI-RADS Characterisation of Mammogram Densities

K. Marias[1], M. G. Linguraru[3], M.A. Gonzalez Ballester[4], S. Petroudi[2], M. Tsiknakis[1] and Sir M. Brady[2]

[1]Institute of Computer Science, FORTH ICS– Heraklion, Greece, [2]Wolfson Medical imaging Laboratory, University of Oxford, UK, [3]Biorobotics Laboratory, Harvard University, Cambridge MA, USA, [4]Medical Image Analysis Group, MEM-ISTB, University of Bern, Bern, Switzerland

*Abstract*—**Intelligent management of medical data is an important field of research in clinical information and decision support systems. Such systems are finding increasing use in the management of patients known to have, or suspected of having, breast cancer. Different types of breast-tissue patterns convey semantic information which is reported by the radiologist when reading mammograms. In this paper, a novel method is presented for the automatic labelling and characterisation of mammographic densities. The presented method is first concerned with the identification of the prominent structures in each mammogram. Subsequently, 'dense tissue' is labelled in a mammogram data set, and BI-RADS classification is performed based on a 2D pdf that is contracted from a "ground truth" data set as well as a shape analysis framework. The presented method can be used in large-scale epidemiological studies which involve mammographic measurements of tissue-pattern, especially since breast-tissue density has been linked to an increased risk of breast cancer.**

*Keywords*—**Image Analysis and Classification; Intelligent Decision Support Systems.**

## I. Introduction

Mammogram images are hard to interpret because of the complex tissue morphology of the breast and the numbers of imaging parameters that affect mammogram acquisition. Moreover, mammogram acquisition involves breast tissue compression, which results in a non-rigid deformation of breast-tissue structures. To date, most researchers have focused on developing holistic approaches to mass and microcalcification detection aiming at being invariant to the mammographic appearance of the breast tissue. However, little work has been done on the identification and characterisation of significant tissue categories prior to mass or calcification detection. Classification of breast density is important for epidemiological studies investigating the relationship between mammogram density and the occurrence of cancer. In addition, mammographic tissue appearance has been correlated with breast cancer risk [1], while change of appearance following Hormone Replacement Therapy (HRT) can also be a sign of cancer.

For these reasons, there is increasing interest in using measurements of mammographic density patterns in computer-aided detection. The most widely accepted pattern coding and characterisation system is the Breast Imaging Reporting and Data system (BI-RADS) suggested by the American College of Radiology [2]. The pattern of breast tissue density is an important characteristic and is routinely reported by physicians around the world, especially those following the ACR BI-RADS reporting system [2]. The work presented in this paper aims to automatically identify and characterise dense groups of tissue in mammograms and subsequently classify dense tissue according to the BI-RADS criteria.

Previous work on mammogram tissue identification has been aimed primarily at mass-detection. Li et al. [3], present a 3-stage method for determining the number of 'dense' tissue regions by optimising a generalised Gaussian mixture intensity model. This allows the identification of a number of 'suspicious' densities in each mammogram. However, the method includes a scale-dependent pre-processing step (structural elements of morphological operations) while results for medio-lateral mammograms, where the pectoral muscle makes segmentation difficult, are not presented. Similarly, in [4] a wavelet enhancement pre-processing step is used but the final 'dense' tissue extraction is parameter dependent. Finally, in [5], both a dynamic contour model and a region growing method are used to detect single densities that are suspicious for cancer.

Figure 1 is a schematic of the proposed method for mammogram tissue labelling and classification. First, each mammogram is labelled in 2 different classes; dense (ducts, lobules and connective tissue) and fat (fatty tissue including the edge along the breast outline and small vessels). The labelling methodology is presented in section II, and can be considered as a generalised tool for tissue labelling in mammograms. Subsequently, the tissue labelling method is applied to BI-RADS classification (section III). By using the 'dense' labelled regions, a set of mammograms is classified based on a Parzen PDF model and shape analysis.

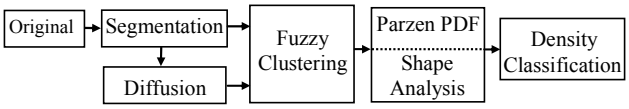



Fig. 1. A schematic of the proposed method for tissue labelling and BI-RADS classification.

## II. Methodology

In this section we propose a method for labelling dense breast tissue. This task can be considered as an important generic pre-processing step, since mammograms can be decomposed into several consistent categories enabling more specific image analysis (e.g. direct matching of corresponding tissue classes for temporal or bilateral mammogram registration and examination only of the dense part of a mammogram, for mass detection). In order to

identify and label important structures in mammograms, it is necessary to segment the pectoral muscle, background and film edge from the mammogram (section II.*A*). This is necessary, since the pectoral muscle can have similar intensities or textures to the denser mammographic structures. Subsequently, each segmented mammogram is filtered using an adaptive anisotropic diffusion-based filter in order to remove small curvilinear structures and to homogenise the important groups of tissue prior to labelling, while preserving their edges (section II.*B*). This step enhances dense regions, reduces structural noise and, unlike morphological enhancement [3], is independent of image size. Finally, a fuzzy-clustering classification scheme is employed to identify the important tissue groups (section II.*C*). These steps are now presented in more detail:

### *A. Segmentation*

The method used for segmenting the breast background and film edges was developed by Highnam and Brady [6] and was chosen because it is fast and reliable, especially compared to histogram-based methods. The method is based on a combination of the linear Hough transform, image gradient operators and morphology.

In order to also remove the pectoral muscle, each pixel intensity $I$(x, y) is transformed according to (1):

$$I(x,y) = 4\,I(x,y)\,\frac{(r-x)^2 + (c-y)^2}{r^2 + c^2} \qquad (1)$$

where r ,c are the image dimensions.

Subsequently, the image is classified in two classes using fuzzy clustering which is also used to label the 'important' tissue classes (i.e. 'dense' and 'fat') and is explained in section II.*C*. A typical pectoral muscle segmentation result is shown in Fig. 2b. Unlike other reported approaches, the method does not assume linearity of the pectoral muscle line.

### *B. Diffusion*

In order to disregard the less important features of the mammogram, images are processed using an adaptive anisotropic diffusion-based filter [7]. This edge-preserving filter removes the less significant structures from the image, while making important mammographic entities (e.g. density and fat) easier to differentiate. The parameters of the filter are computed from a statistical analysis of the image gradient (2-3) and the mammogram is anisotropically blurred (4). The statistical analysis is based on an adaptive Gaussian derivative filter:

$$g_i = M_i - \frac{1}{N}\sum_{j\in\delta_i} M_j \qquad (2)$$

where

$$M = |K_\sigma'(I)|, \qquad K_\sigma(I) = \frac{1}{2\pi\sigma^2} * \exp\left(-\frac{|I|^2}{2\sigma^2}\right) \qquad (3)$$

$I$ represents the intensities in a mammogram, N is the number of pixels in the neighbourhood $\delta_i$ of $i$, $g_i$ a measure of local contrast and $K_\sigma$' the Gaussian derivative. The parameter $\lambda$ that is used in anisotropic filtering is related to the image contrast and results from the estimation of $g_i$. Subsequently, it is used in the computation of the eigenvalues used by the diffusion tensor of the anisotropic diffusion:

$$\lambda_1 = \begin{cases} 1 & |\nabla u_\sigma| = 0, \\ 1 - \exp\left(\dfrac{-1}{(|\nabla u_\sigma|/\lambda)}\right) & |\nabla u_\sigma| > 0 \end{cases} \qquad \lambda_2 = 1 \qquad (4)$$

where $u_\sigma$ is the gradient of the smoothed image u.

Having the values of $\lambda_1$ and $\lambda_2$ and their associated eigenvectors (one perpendicular to and one along the isophote), the diffusion tensor can be computed as a symmetric matrix. Subsequently, the flux of diffusion can be estimated [8]. A diffused mammogram is shown in Fig. 2c.

### *C. Fuzzy Clustering*

In this final step, both original intensities and diffused intensities are used to characterise each pixel. From an image segmentation standpoint, this can be considered as a trade-off between accuracy (original intensities) and smoothness (diffused intensities). Each non-zero pixel $i$ is represented by the following vector:

$$V_i = \left[I^n(i),\, a \cdot I^n_{Diffused}(i)\right], \forall\, i > 0 \qquad (5)$$

where $I_{Diffused}$ is the diffused version of the original image $I$, $\alpha$ is the smoothness trade-off (set to 0.7) and $n$ is an adaptive parameter that aims to characterise the dense-tissue content of each mammogram. The larger its value, the more 'brighter' regions are favoured and 'darker' suppressed. Our algorithm estimates the maximum intensity $I_{\max}$ that avoids "outliers" such as noise or microcalcifications (e.g. by adding a constraint that the $I_{\max}$ frequency is larger than 50). This way, the useful histogram extent of each mammogram is estimated, avoiding small densities corresponding to segmentation residuals or microcalcifications. Subsequently, $n$ is calculated by (5):

$$n = 1.5 \cdot (256 - I_{\max}) \qquad (6)$$

This results in $n$ being larger for the less dense mammograms, thus enhancing the 'contrast' in their feature vectors. Having calculated the feature vectors, fuzzy clustering assigns different degrees of membership to each point. The membership of a point is thus shared among various clusters. In order to compute the fuzzy partition of the data, the following cost-function is minimized iteratively:

$$J_q(U, C) = \sum_{j=1}^{M}\sum_{i=1}^{K} (u_{ij})^q \cdot d^2(V_j, C_i) \qquad (7)$$

where $q$ is the weighting exponent for $u_{ij}$ and controls the "fuzziness" of the resulting clusters, $V$ the feature vector, $C_i$ the center of each cluster, $u_{ij}$ the degree of membership of $V_j$ in the $i^{\text{th}}$ cluster, $M$ the number of non-zero pixels and $K$

the pre-defined number of desired clusters. $C_i$ and $u_{ij}$ are calculated iteratively according to (8):

$$u_{ij} = \left[d^2(V_j, C_i)\right]^{\frac{-1}{(q-1)}} \Big/ \sum_{k=1}^{K} \left[d^2(V_j, C_k)\right]^{\frac{-1}{(q-1)}},$$
$$C_i = \frac{\sum_{j=1}^{M} (u_{ij})^q V_j}{\sum_{j=1}^{M} (u_{ij})^q} \quad (8)$$

For mammogram tissue classification, $K$ is initially set to 3, provided that the pectoral muscle, film edge and background have been segmented correctly. The 'denser' class (highest grey-level statistics) is an approximation to the mammographic density, while the 'fat' (lowest grey-levels) to the fatty edge (a homogenous, low-intensity region near the breast edge). The 'intermediate' class corresponds to the fatty background of the mammogram. In order to refine the final result, the 'intermediate' region is classified again into 2 sub-regions. The denser sub-region is merged with the 'dense' class unless its statistics are 'closer' to the 'fat' one. This results in a refinement of the 'dense' region estimation. In addition, to increase robustness, an additional optimisation step is included; the clustering is repeated if the 'dense' region is <3% of the total mammogram area. Thus, the classification of small yet very bright image regions as the 'entire' breast density is avoided (e.g. macrocalcifications can be significantly brighter than the rest of the tissue). Fig. 2 illustrates the labelling process; the original mammogram (2a) is gradually segmented (2b), diffused (2c) and labelled (2d). The brighter class corresponds to 'dense' fibroglandular tissue and is used to characterise mammographic densities in 146 mammograms based on the BI-RADS criteria for breast tissue patterns. The results are presented in the next section, where a 2D Parzen PDF is constructed from a "ground truth" dataset and used in conjunction with a shape analysis framework for automated tissue-pattern classification.

## III. Results

To date, BIRADS represents the most widely accepted reporting systems for mammographic findings in the US and Europe and is therefore considered a milestone for developing automated methods for characterising mammogram findings. The American College of Radiologists suggests that breast composition should be reported in all patients using the BI-RADS [2] classification: I describes an almost entirely fat breast, II breasts with scattered fibroglandular densities, III heterogeneously dense breasts and IV extremely dense breasts.

Previous work on breast-density classification has mainly used global texture measures to differentiate the four classes. The novelty of the presented approach lies in the fact that semantic mammogram information is extracted before classification, enabling the design of intuitive classification algorithms based on the definition of BI-RADS criteria and not on the assumption that all categories have consistently difference texture, an assumption that doesn't generally hold. Therefore, by using the tissue labelling method described in section II, it is possible to perform BI-RADS classification based on the morphology and intensity of only the dense 'labelled' regions of each mammogram.

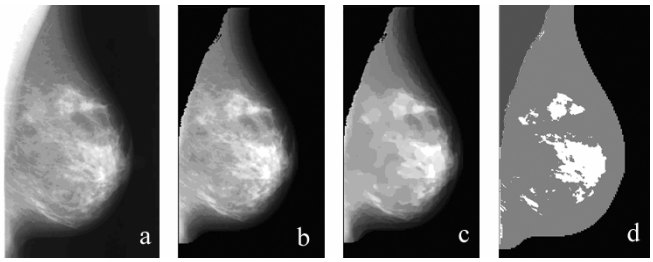


Fig. 2. a: the original mammogram, b: segmented mammogram (section A), c: diffused mammogram (section B), d: the labelled one (section C).

To describe a previously labelled density, both area and intensity information is included, since several mammographic densities may have the same spatial extent, but different intensity profile, reflecting actual tissue density differences that result from variable X-ray absorption. Previous results in this field [9], indicate that 'estimated' density is a good measure to separate 'high' (III and IV) from 'low' (I and II) density mammograms, but not I from II and III from IV, due to a partial overlap in density and texture measures. Therefore, the first aim is to separate 'high' from 'low' density classes and subsequently to try to further separate each one of these sub-categories. This is done by first calculating a 2D-PDF, and then employing a shape analysis framework. The results obtained in each one of these steps are reported in the following subsections:

### *A. Parzen Estimation of 'High' and 'Low' Density PDF*

'High' and 'low' density PDFs are estimated from 65 mammograms with identical BI-RADS descriptions amongst 3 different observers in [9] that define the 'ground truth' dataset. This is done to minimise inter-observer variability. For each mammogram the % area (labelled dense area divided by total area excluding pectoral muscle and background) and % density (defined as the sum of dense area intensities divided by the sum of all intensities excluding pectoral muscle) is calculated from the dense class (using the method in Section II). By using these 2D measurements of the "ground truth" mammograms we compute a 2D PDF of high' (III, IV) and 'low' density (I, II) mammogram patterns, using Parzen window estimation:

$$p(x, y) = \frac{(2\pi)^{-d/2} h^{-d} \sum_{i=1}^{M} e^{-\frac{1}{2h^2}\left((x-x_i)^2+(y-y_i)^2\right)}}{M} \quad (9)$$

where $d$ is the data dimension (2 in this case), $h$ the width of a Parzen window, and ($x_i$, $y_i$) the area and density measures of all 'high' and 'low' density mammograms of the 'ground truth' dataset. Using this dataset, the 'high' and 'low' density 2D PDFs are estimated, as is shown in Fig. 3a. It is noticeable that each PDF comprises two sub-distributions that partially overlap. This confirms that it is hard to separate within the 'high' and 'low' density classes.

Subsequently, a test set of 146 mammograms [9] is classified into 'high' and 'low' based on the Parzen probabilities. The measurements for computing the probabilities are shown in Fig. 3b and the separation accuracy between 'high' and 'low' was found to be 92%. To further separate III to IV a shape analysis scheme is adopted (section III.B). However, this analysis proved to be useful only for the 'high' density class and the classification accuracy for further separating I from II, based only on the measurements of Fig. 3 (right), was 65%.

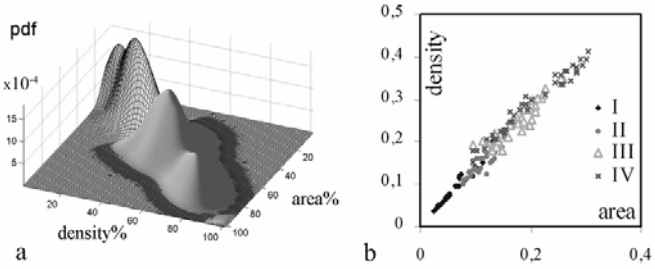



Fig. 3. Characterising breast densities. a: The parzen pdfs for 'high' (interpolated surface) and 'low' (meshed surface) density classes based on the 'standard' dataset. b: The area, density measures for the test dataset.

### B. *Shape Analysis*

According to the BI-RADS criteria, dense tissue in category III mammograms is more scattered than in category IV. Translating this criterion into an image analysis algorithm based on texture is rather difficult due to the large variability in anatomical morphology of tissue, breast compression and imaging parameters. However, once dense tissue is extracted from each mammogram (section II), the 'dense' categories can be further classified by directly describing the shape's solidity (Solidity = Area/Convex Area) and Euler number (the total number of objects in the image minus the total number of holes in those objects). Fig. 4a shows a number of extracted densities from the 'dense' category of the tested set. It is noticeable that the densities of category IV are more compact than those of III, as expected. As shown in Fig. 4b, the separation accuracy based on shape description was 86%, while it would be less than 50% based only on the density-area measurements.

Solidity and Euler number provide global shape information. More detailed insight into shape differences can be obtained by measuring the 'jaggedness' of the extracted object's boundary, which seems to be of importance to separate categories III and IV (Fig. 4a). This can be measured by computing the fractal dimension [10]. The correlation of the fractal dimension and high breast tissue density has been reported previously [11-12], but, to our knowledge, this is the first time it has been used to differentiate between breast tissue classes according to established criteria such as BI-RADS.

In Fig. 5, the fractal dimension measure, computed using a box-counting algorithm [13], is shown for the same image dataset as above. In addition, the fractal dimension measurements were normally distributed according to the D'Agostino Pearson normality test [14] at a significance level of 0.05. Based on this observation, the probability error of "overlap" was 0.18 (in the ideal case this should be equal to 0 i.e. no overlap between the 2 distributions). From a purely statistical perspective the significance probability that the populations generating III and IV are identical was less than $10^{-10}$ using both Gaussian (t-test) and non-Gaussian (Wilcoxon rank sum test) statistical tests.

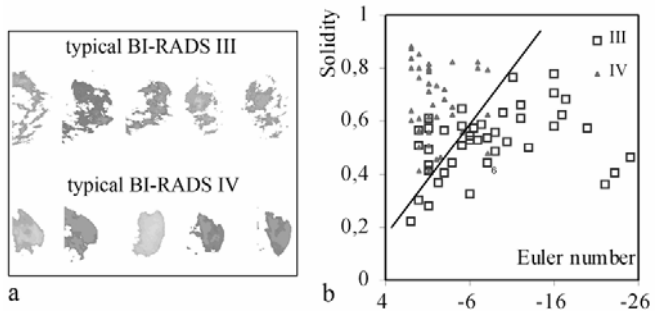



Fig. 4. a: typical 'dense' labelled regions of BI-RADS III (first row) and IV (second row) mammograms. b: shape analysis of the 'high' density mammograms calculated in III.A.

## IV. Discussion

A fully automatic method to label mammogram tissue groups was presented. The main advantage of this method is that it is non-parametric and is based on extracting semantic density information rather than using non-intuitive global textural measures. The method was tested for BI-RADS classification of tissue patterns and the results were satisfactory, especially for discriminating the 'high' density categories III and IV, which are of greater clinical importance, since women with dense breast run an increased risk of breast cancer [1].

The labelling method (section II) has shown consistent results in hundreds of mammograms used to date. In Fig. 4a, the last 2 mammograms in each row are temporal, acquired with different compression and imaging parameters. It is noticeable, that their 'dense' labelled classes show consistent shape. The method has the potential to be used as a pre-processing tool in CAD systems for mammography. Computational time and false positives could be reduced by only considering the labelled 'dense' tissue regions as mass candidates. Moreover, tissue-labelling information can be pre-computed and stored in large DICOM mammogram databases. Last, it can be used as an alternative to the 'interactive thresholding' technique that has been broadly used for defining the mammoghaphic densities. Although this method has been used in epidemiological studies [1,11,15], it is important to mention that density is globally segmented and characterised. In contrast, the presented method decomposes the mammogram in significant categories while a statistical and shape-based framework can

further define dense tissue patterns according to widely acceptable criteria (BI-RADS).

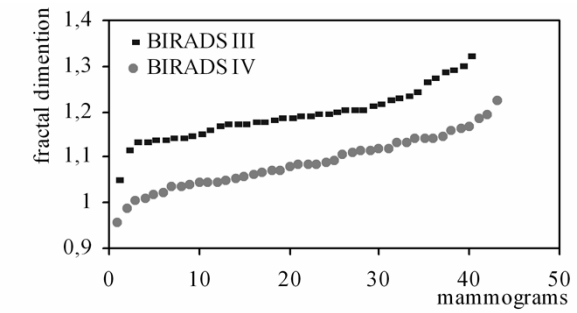



Fig. 5. The fractal dimension of the mammograms corresponding to the BIRADS III and IV categories in ascending order.

It should be noted that the presented method does not completely solve the problem of BI-RADS classification which suffers from both significant interpretation variability and difficulties in translating ontological criteria to computer measures. The long-term scientific vision of this work is the development of software assisted reporting systems for mammography aiming to reduce reading/reporting time and improve interoperability and exchangeability of heterogeneous and distributed mammographic data by enabling data descriptions of high semantic content.

## REFERENCES


[1] N.F. Boyd, G.A. Lockwood, J.W. Byng, D.L. Tritchler, M.J. Yaffe, "Mammographic densities and breast cancer risk," *Cancer Epidimiol Biomarkers*, vol. 7, pp. 1133-1144, 1998.

[2] American College of Radiology, *Illustrated breast imaging reporting and data system (BI-RADSTM)*. ACR, 1998.

[3] H. Li, Y. Wang, K.J.R. Liu, S.C.B. Lo and M.T. Freedman, "Computerized Radiographic Mass Detection – Part I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation," *IEEE Transaction on Medical Imaging*, vol. 20, pp. 289-301, 2001.

[4] J.T. Neyhart, M. Kirlakovsky, K.M. Coleman, R. Polikar, M. Tseng, S.A. Mandayam, "Automated segmentation and quantitative characterization of radiodense tissue in digitized mammograms," *Review of Progress in Quantitative Nondestructive Evaluation (QNDE 2001)*, 21B , pp. 1866-1879, Published by the American Institute of Physics, 2002.

[5] G. M. te Brake and N. Karssemeijer, "Segmentation of suspicious densities in digital mammograms," *Medical Physics*, vol. 28, pp. 259-266, 2001.

[6] R.P. Highnam and J.M.Brady. *Mammographic Image Analysis*. Kluwer Academic Publishers, 1999.

[7] M.G. Linguraru, J.M. Brady, M. Yam, "M.G. Linguraru, J.M. Brady, M. Yam, "Filtering hint Images for the Detection of Microcalcifications,," in *Proc. Medical Image Computing and Computer-Assited Intervention*, Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg New York 2001, pp. 629-636.

[8] J. Weickert, *Anisotropic Diffusion in Image Processing*. B.G. Teubner, Stuttgart, 1998.

[9] K. Marias, S. Petroudi, R .English, R. Adams and J.M.Brady, "Subjective and computer-based characterisation of mammographic patterns," in *Proc. IWDM*, Lecture Notes in Computer Science, Springer Verlag Berlin Heidelberg, 2002, pp. 552-557.

[10] B.B. Mandelbrot, *The Fractal Geometry of Nature*. W.H. Freeman, New York, 1983.

[11] J.W. Byng, M.J. Yaffe, G.A. Lockwood, L.E. Little, D.L. Tritchler, and N.F. Boyd, "Automated analysis of mammographic densities and breast carcinoma risk," *Cancer*, vol. 80(1), pp. 66-74, Jul 1997.

[12] V. Velanovich, "Fractal analysis of mammographic lesions: a feasibility study quantifying the difference between benign and malignant masses," *Am J Med Sci.*, vol. 311(5), pp. 211-214, May 1996.

[13] K. Falconer, *Fractal Geometry - Mathematical Foundations and Applications*. John Wiley & Sons, Chichester, 1990.

[14] R.B. D'Agostino, A. Belanger, and R.B. D'Agostino Jr., "A suggestion for powerful and informative tests of normality," *The American Statistician*, vol. 44(4), pp. 316-321, November 1990.

[15] J.W. Byng, M.J. Yaffe, R.A. Jong, R.S. Shumak, G.A. Lockwood, D.L. Tritchler, N.F. Boyd, "Analysis of mammographic density and breast cancer risk from digitized mammograms," *Radiographics*, vol. 18(6), pp.1587-98 Nov-Dec 1998.