Automated Mammographic Risk Classification Based on Breast Density Estimation

Zhili Chen[1,*], Arnau Oliver[2], Erika Denton[3], and Reyer Zwiggelaar[1]

[1] Department of Computer Science,
Aberystwyth University, Aberystwyth, SY23 3DB, UK
{zzc09,rrz}@aber.ac.uk
[2] Department of Computer Architecture and Technology,
University of Girona, 17071 Girona, Spain
aoliver@eia.udg.edu
[3] Department of Radiology,
Norfolk and Norwich University Hospital, Norwich, NR4 7UY, UK
erika.denton@nnuh.nhs.uk

Abstract. This paper presents a method for automated mammographic risk classification based on breast density estimation in mammograms. The overall profile of breast tissue density is represented using a topographic map, which is a hierarchical representation, obtained from the upper level sets of an image. A *shape* tree is constructed to describe the topological and geometrical structure of the *shapes* (i.e. connected components) within the topographic map. Two properties, *saliency* and *independency*, are defined to detect *shapes* of interest (i.e. dense regions) based on the *shape* tree. A density map is further generated focusing on dense regions, which provides a quantitative description of breast density. Finally, mammographic risk classification is performed based on the breast density measures derived from the density map. The validity of this method is evaluated using the full MIAS database and a large dataset taken from the DDSM database. A high agreement with expert radiologists is indicated according to the BIRADS density classification. The obtained classification accuracies are 76.01% and 81.22%, respectively.

1 Introduction

Numerous studies have indicated that breast density is a strong risk factor for developing breast cancer [1–6], and therefore quantitative estimation of breast density is of great importance for mammographic risk classification. A variety of semi-automated and fully-automated approaches have been developed to quantify breast density. The most commonly used method is interactive thresholding (known as Cumulus), where the greylevel threshold is manually tuned by observers to segment dense tissue [1]. Petroudi and Brady [2] implemented breast density segmentation using textons. Oliver *et al.* [3] used the fuzzy c-means

---

clustering to segment the breast into fatty and dense tissue, and features were extracted from both tissue classes for the subsequent breast density classification. Tzikopoulos *et al.* [4] employed the minimum cross entropy thresholding to detect dense tissue. He *et al.* [5] developed a method for segmenting mammographic images into Tabár's four mammographic building blocks and the relative proportions of the four building blocks were used for mammographic risk classification. Kallenberg *et al.* [6] proposed a segmentation method where different approaches to breast density segmentation in the literature were integrated.

We propose a new method to classify mammographic images into different risk categories based on quantitative estimation of breast density. Dense tissue regions are detected using the topographic map of the breast region. Quantitative measures of breast density are derived from the resulting density map and used for mammographic risk classification. A topographic approach was also used by Hong *et al.* [7] for a different purpose of detecting the breast boundary, the pectoral muscle and candidate masses. In addition, most previous methods need a learning procedure (e.g. [2, 3, 5]) or use a high dimensional feature vector for segmenting/classifying breast density (e.g. [3, 4, 6]), while the proposed method is straightforward and uses only few features for classification.

2   Topographic Representation

A topographic map is a morphological and multiscale decomposition of an image relying on the connected components of the level sets. In this work, the topographic map is built by the upper level sets at a sequence of intensity levels over the full intensity range of the image, which is defined as $TM(I) = \{U_{l_i}(I) \mid l_i \in L, i = 1, 2, \cdots, N\}$, where $U_{l_i}(I)$ represents the upper level set of the image I at an intensity level of l_i, L denotes the full intensity range, and N is the number of intensity levels for generating the upper level sets. The connected components (8-connectivity) of the topographic map are referred to as the *shapes* in which the holes have been filled (see [8] for details). The upper level sets constitute a decreasing family corresponding to an increasing intensity level. Therefore, the topographic map has a hierarchical structure from large to small scales (the scale here corresponds to the area of the *shapes*), and the number of intensity levels N behaves as a resolution factor in the sense that a large N can capture a high-resolution topographic map. In addition, the topographic map is contrast invariant to any increasing contrast change due to its hierarchical property.

The topological and geometrical structure of the *shapes* in the topographic map can be described by a *shape* tree, which is constructed based on an inclusion relationship between the *shapes*. The root node represents the whole topographic map and each node descended from the root corresponds to a *shape* in the map. If *shape* S_i spatially encloses *shape* S_j, S_i is the *parent* node of S_j, and S_j is the *child* node of S_i. If one node has multiple immediate *child* nodes, it is called a *branching* node. The immediate *child* node of a *branching* node is called a *base* node. If one node has no *child* node, it is called a *terminal* node. An M-generation *ancestor family* of a *shape* S, denoted by $\mathcal{A}^M(S)$, is defined as

$\mathcal{A}^M(S) = \{p^m(S), \mid m = 1, 2, \cdots, M\}$, where $p^m(S)$ is the mth *parent* node of S. An M-generation *descendent family* of S, denoted by $\mathcal{D}^M(S)$, is defined as $\mathcal{D}^M(S) = \{c^m(S), \mid m = 1, 2, \cdots, M\}$, where $c^m(S)$ is the mth *child* node of S.

3   Breast Density Estimation

An example mammogram is shown in Fig. 1(a). As a pre-processing step, the breast region is segmented using the approach in [9] as shown in Fig. 1(b) and the level lines (i.e. *shape* contours) are shown in Fig. 1(c) (32 uniformly spaced intensity levels are used). The corresponding topographic map is shown in Fig. 1(d) (the *shapes* smaller than 300 pixels are removed to eliminate sensitivity to noise and small intensity irregularities). All *shapes* contained in the topographic map are shown in Fig. 1(e). It is shown that the *shape* contours closely capture the boundaries of regions with various densities. On the one hand, a nesting pattern is formed near the boundary of a distinctively dense region where the *shape* contours are densely nested. On the other hand, a significant transition happens between the contours of two adjacent *shapes* if one of which is the outermost boundary of a dense region. Dense tissue regions are regarded as *salient* or *independent shapes* in the topographic map. Specifically, a *salient shape* is a *shape* which indicates the local maximum intensity value with respect to the surrounding background; and an *independent shape* corresponds to a *shape* of which the contour has a significant transition from that of the *parent shape*.

Dense tissue regions are considered as *salient shapes*, when regions have higher intensity values than the surroundings. The *saliency* of a *shape* can be evaluated based on the *shape* tree. In each branch of the tree, the nodes closer to the *terminal* node have a higher *saliency* score compared to those further away, as the *terminal* node indicates the maximum intensity value among all the nodes in this branch. Therefore, the *terminal* nodes are rated as the most *salient shapes*. However, the *base* node corresponding to each *terminal* node is finally selected, as it is the outermost *shape* enclosing a dense region. Dense tissue regions are detected from the finest scale to coarser scales by iteratively tracing the *base* nodes from multiscale *shape* trees, obtained by gradually removing small *shapes* (*shapes* smaller than $300 \times i$ pixels are removed at the ith iteration). The iterative process terminates when a trunk is formed. Figure 1(f) concisely illustrates this iterative process covering the core *shape* trees (*shapes* $S_3, S_4, \cdots, S_{22}$ located in the trunk are not displayed). It is shown that the tree is cut at iterations 2, 6 and 9, and the trunk is formed at the 9th iteration. Finally, the nodes S_{24}, S_{25}, S_{29} and S_{30} (indicated by red solid squares in Fig. 1(e)) are selected as dense tissue regions as shown in Fig. 1(g).

Dense tissue regions are regarded as *independent shapes* as a sharp shape transition happens between their *ancestor* and *descendent families*. The *independency* of the *shape* S is defined as:

$$Ind(S) = \frac{1}{Area(S)} \cdot \frac{\langle Area(\mathcal{A}^M(S)) \rangle - Area(S)}{Area(S) - \langle Area(\mathcal{D}^M(S)) \rangle}$$
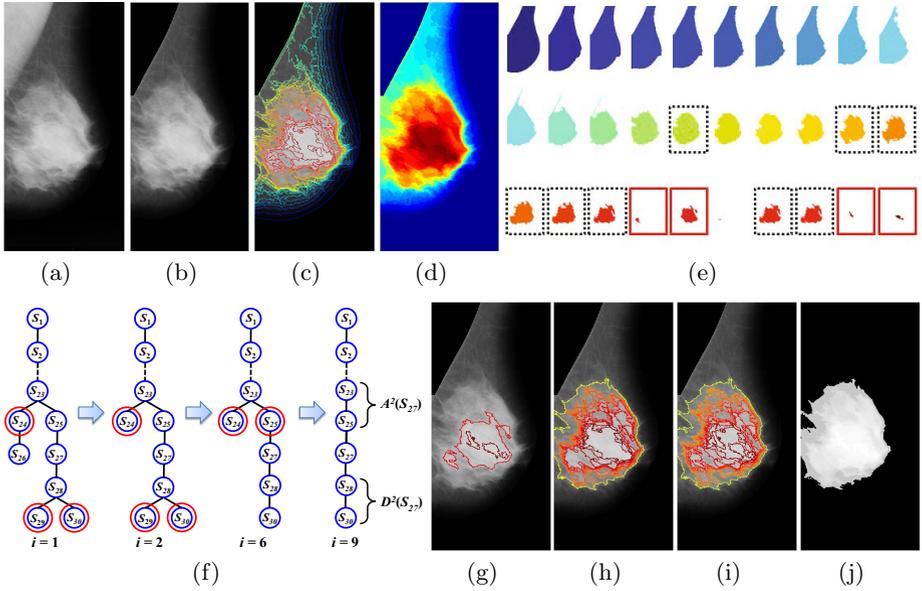
Fig. 1. (a) Original mammogram; (b) breast region segmentation; (c) level lines (*shape* contours); (d) topographic map; (e) all *shapes* contained in (d) (S_1 - S_{30}); (f) multiscale *shape* trees in the iterative process indicating the change in the tree layout (nodes marked with red circles are the selected *shapes* at that iteration); (g) dense tissue regions extracted based on *saliency*; (h) dense tissue regions extracted based on *independency*; (i) all extracted dense tissue regions (after removing false positives); and (j) density map (the contrast has been normalised for better visualisation).

where $Area(S)$ is the area of the *shape* S, $\langle Area(\mathcal{A}^M(S)) \rangle$ and $\langle Area(\mathcal{D}^M(S)) \rangle$ are the mean areas of the *ancestor* and *descendent families* of S, which are computed by $\frac{1}{M}\sum_{m=1}^{M} Area(p^m(S))$ and $\frac{1}{M}\sum_{m=1}^{M} Area(c^m(S))$, respectively. $Ind(S)$ is normalised by $Area(S)$ to avoid bias caused by the size of the *shape*. It is noted that a large *independency* value indicates the *shape* has significant shape change with respect to its *ancestor family* but insignificant shape change with respect to its *descendent family*, which indicates a high probability of it being the outermost *shape* enclosing a dense region. Thus, dense tissue regions are detected from the resulting trunk of the *shape* tree by setting a threshold on $Ind(S)$. The *shapes* with *independency* values larger than the threshold are selected. For an illustration of the detection of subtle region boundaries, $M = 1$ and a threshold of 8.6×10^{-6} are used here (the threshold of *independency* was automatically determined for each mammographic image when generating the experimental results presented in the following section, but small variations provided similar results). As a result, S_{15}, S_{19}, S_{20}, S_{21}, S_{22}, S_{23}, S_{25}, S_{27} and S_{28} (indicated by black dotted squares in Fig. 1(e)) are selected as dense tissue regions as shown in Fig. 1(h).

To remove false positive regions, the 2nd order central moments of the resulting dense tissue regions are computed. Two descriptors, *elongation* and *compactness*, are computed for each region. In our work, the regions of interest are compact, approximately circular or elliptical regions, and therefore regions with a large *elongation* value (> 5.5) or a small *compactness* value (< 0.04) are discarded. The two thresholds were set by experiments mainly for the purpose of removing the elongated *shapes* along the straight boundary of the breast region due to the incomplete removal of the pectoral muscle. The final detection result is shown in Fig. 1(i).

In order to quantify breast density for mammographic risk classification, a density map is created for mammographic images based on the detected dense tissue regions (see Fig. 1(j) for an example), which is defined as:

$$DM(x,y) = \begin{cases} 0 & (x,y) \notin Regions \\ \frac{1}{|R_{min}|} \sum_{(x,y) \in R_{min}} I(x,y) & (x,y) \in Regions \end{cases}$$

where $DM(x,y)$ is the intensity value of pixel (x,y) in the density map, *Regions* denotes the union of all the dense tissue regions, R_{min} represents the smallest region containing pixel (x,y) in *Regions*, and $|R_{min}|$ calculates the number of pixels within R_{min}. Two quantitative measures of breast density are derived from the density map for mammographic risk classification, which are referred to as *dense area* and *average density*, respectively. The *dense area* corresponds to the overall area of the dense regions in the breast, which is defined as the number of the non-zero pixels in the density map, calculated by $density\ area = |Regions|$. The *average density* corresponds to the average density of the dense regions, which is defined as the average intensity value of the non-zero pixels in the density map, computed by $average\ density = \frac{1}{|Regions|} \sum_{(x,y) \in Regions} DM(x,y)$.

4   Results

To evaluate the performance of the proposed method for mammographic risk classification, it has been applied on two well-known databases. The first one is the full Mammographic Image Analysis Society (MIAS) database [10], which contains 322 mammograms from 161 women. Bilateral (left and right) MLO mammograms were taken for each woman. Three expert radiologists were involved to classify 321 valid mammograms (mdb295ll was excluded for historical reasons) into four classes according to the Breast Imaging Reporting and Data System (BIRADS) density classification [11]. The consensus between the three individual ratings was considered as the final class label which was used as the classification ground truth in this work. The second one is a larger dataset taken from the Digital Database for Screening Mammography (DDSM) database [12], which contains 1662 mammograms from 831 women. Four mammograms were taken for each woman, including both MLO and CC views of each breast. The right MLO and CC mammograms of each woman were extracted for the subsequent experiments. In contrast to the MIAS database, the BIRADS density rating was provided in the DDSM database, which was used as the ground truth.
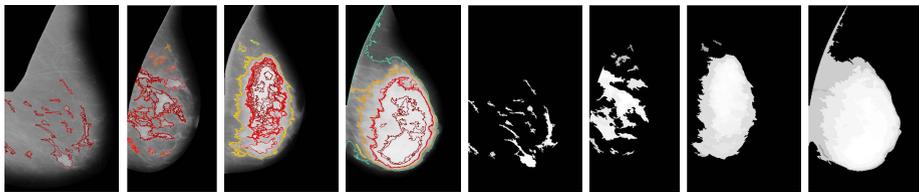
Fig. 2. Detected dense tissue regions and density maps (the contrast has been normalised for better visualisation) of example mammograms. From left to right, the mammograms range from BIRADS I to BIRADS IV.

For each mammogram, eighty-five uniformly spaced intensity levels ($N = 85$) were used to generate a high-resolution topographic map. The *shape* tree was first constructed at the finest scale, containing all the *shapes* larger than 300 pixels. Then, the *shape* tree was simplified to coarser scales by iteratively removing small *shapes* and as such only the *shapes* larger than $300 \times i$ pixels were retained at the ith iteration. On the basis of this, dense tissue regions were iteratively detected by searching for the *base* nodes. A three-generation *family* ($M = 3$) was used to detect dense tissue regions from the remaining trunk afterwards. Fig. 2 shows the resulting dense tissue regions and the corresponding density maps of example mammograms, which indicated realistic segmentation of dense regions. As shown in Fig. 2, the overall area of dense regions and the average intensity of the density map both increase from BIRADS I to BIRADS IV, which is consistent with the increasing breast density and mammographic risk.

For mammographic risk classification, the two quantitative measures, *dense area* and *average density*, derived from the density map were used as features. The value of *dense area* was normalised by dividing the area of the breast region to avoid the bias induced by the size of individual breasts and as such the normalised *dense area* represented the relative proportion of dense tissue to the whole breast. On the other hand, the value of *average density* was normalised by dividing the maximum intensity value of 255 within the full greylevel scale (8 bits). The purpose of using a uniform value to normalise *average density* instead of a varied value depending on mammograms was to preserve the original correlation of intensity information between mammograms. A k-Nearest Neighbours (kNN) based classifier was used for classification, which was initially based on a simple majority vote unless multiple classes indicated the same number of training samples among the k nearest neighbours (i.e. a tie occurs in the decision making) in which case a distance weighted approach was applied to resolve the tie. The Euclidean distance was used as the similarity measure. The leave-one-woman-out methodology was used for cross-validation. For the MIAS database, when classifying one MLO mammogram of one woman, the opposite-side mammogram was excluded from the training set. For the DDSM database, we classified mammograms of the MLO view and the CC view separately. Thus, the leave-one-woman-out methodology was regarded as the leave-one-image-out cross-validation for mammograms of a single view.

Table 1. Confusion matrices for the MIAS database and the DDSM database

(a) MIAS (CA = 76.01%)

	BIRADS	I	II	III	IV	CA
Truth	I	72	13	2	0	83%
	II	18	73	12	0	71%
	III	0	19	70	5	74%
	IV	0	0	8	29	78%

(b) MLO DDSM (CA = 75.21%)

	BIRADS	I	II	III	IV	CA
Truth	I	80	26	0	0	75%
	II	33	266	35	2	79%
	III	0	40	186	29	73%
	IV	0	4	37	93	69%

(c) CC DDSM (CA = 74.61%)

	BIRADS	I	II	III	IV	CA
Truth	I	72	34	0	0	68%
	II	29	280	27	0	83%
	III	0	54	180	21	71%
	IV	0	5	41	88	66%

(d) MLO/CC DDSM (CA = 81.22%)

	BIRADS	I	II	III	IV	CA
Truth	I	93	13	0	0	88%
	II	27	286	23	0	85%
	III	0	41	199	15	78%
	IV	0	5	32	97	72%

The confusion matrices for the two databases can be found in Table 1. The overall classification accuracy (CA) was 76.01% for the MIAS database, while for the DDSM database, the CAs were 75.21% and 74.61% for the MLO view and the CC view, respectively. The CA for the DDSM database of the individual view was slightly worse than that for the MIAS database. This might be due to the fact that the DDSM database used in the experiments contained more mammograms than the MIAS database and the mammograms belonging to the same class showed higher variance than the MIAS database. However, the obtained results for the DDSM database still indicated the robustness of our method on a large dataset. In addition, we investigated the capability of combining the MLO and CC views of each woman for mammographic risk classification. The two features extracted from each individual view were concatenated into a single feature vector for each woman. The combination of the two views provided a good improvement as shown in Table 1(d). An overall CA of 81.22% was obtained and the CAs for the four BIRADS classes were all improved.

It is shown that good classification results have been obtained compared to the closely related work where the same database and the same BIRADS classification ground truth were used for classification, which enabled a direct comparison. In [3], the achieved best CAs were up to 86% and 77% for the MIAS and DDSM databases, respectively. However, it should be noted that the sophisticated classification using more features and more advanced classifiers is outside the scope of this paper, and will be investigated in future work. In [5], the proportions of different breast tissue classes were directly used for classification and the CA was 70% for the MIAS database. In addition, it should be noted that the proposed method is distinct from a greylevel thresholding method. We have compared the obtained results with a simple segmentation method (Otsu thresholding). The resulting CA of Otsu was 48.29% for the MIAS database by using the area and average intensity of the segmented regions for classification.

5   Conclusions

We have presented a method for automated mammographic risk classification, which is based on quantitative estimation of breast density. The topographic map was exploited to represent a series of regions having a range of densities. A novel way was proposed to define dense tissue regions in the topographic map, in terms of *saliency* and *independency*. The resulting density maps have indicated realistic description of breast density. High classification accuracies have been obtained for the MIAS and DDSM databases according to the four BIRADS categories, which indicates the promising capability of the proposed method. In addition, the proposed breast density measures have been demonstrated to have the potential for quantifying breast density and tracking the quantitative change in breast density over time for mammographic risk assessment.

References

1. Byng, J.W., et al.: The Quantitative Analysis of Mammographic Densities. Physics in Medicine and Biology 39, 1629–1638 (1994)
2. Petroudi, S., Brady, M.: Breast Density Segmentation Using Texture. In: Astley, S.M., Brady, M., Rose, C., Zwiggelaar, R. (eds.) IWDM 2006. LNCS, vol. 4046, pp. 609–615. Springer, Heidelberg (2006)
3. Oliver, A., et al.: A Novel Breast Tissue Density Classification Methodology. IEEE Transactions on Information Technology in Biomedicine 12(1), 55–65 (2008)
4. Tzikopoulos, S.D., et al.: A Fully Automated Scheme for Mammographic Segmentation and Classification Based on Breast Density and Symmetry. Computer Methods and Programs in Biomedicine 102(1), 47–63 (2011)
5. He, W., et al.: Mammographic Image Segmentation and Risk Classification Based on Mammographic Parenchymal Patterns and Geometric Moments. Biomedical Signal Processing and Control 6(3), 321–329 (2011)
6. Kallenberg, M.G.J., et al.: Automatic Breast Density Segmentation: An Integration of Different Approaches. Physics in Medicine and Biology 56(9), 2715–2729 (2011)
7. Hong, B.W., Sohn, B.S.: Segmentation of Regions of Interest in Mammograms in a Topographic Approach. IEEE Transactions on Information Technology in Biomedicine 14(1), 129–139 (2010)
8. Xia, G.S., et al.: Shape-based Invariant Texture Indexing. International Journal of Computer Vision 88(3), 382–403 (2010)
9. Chen, Z., Zwiggelaar, R.: Segmentation of the Breast Region with Pectoral Muscle Removal in Mammograms. In: Medical Image Understanding and Analysis, pp. 71–75 (2010)
10. Suckling, J., et al.: The Mammographic Image Analysis Society Digital Mammogram Database. Excerpta Medica. International Congress Series 1069, 375–378 (1994)
11. American College of Radiology: Illustrated Breast Imaging Reporting and Data System BIRADS, 3rd edn. American College of Radiology, Philadelphia (1998)
12. Heath, M., et al.: The Digital Database for Screening Mammography. In: International Workshop on Digital Mammography, pp. 212–218 (2000)