Breast lesion Segmentation in Ultra-Sound images based on super-pixels and high-level descriptors

Joan Massich[a] and Guillaume Lemaître[a,b] and Joan Martí[b] and Fabrice Mériaudeau[a]

[a]LE2I-UMR CNRS 6306, Université de Bourgogne, 12 rue de la Fonderie, 71200 Le Creusot, France;

[b]ViCOROB, Universitat de Girona, Campus Montilivi, Edifici P4, 17071 Girona, Spain

ABSTRACT

This document shows the desired format and appearance of a manuscript prepared for the Proceedings of the SPIE. It contains general formatting instructions and hints about how to use LaTeX. The LaTeX source file that produced this document, `article.tex` (Version 3.3), provides a template, used in conjunction with `spie.cls` (Version 3.3).

Keywords: Breast Ultra-Sound, Machine-Learning based Segmentation

1. INTRODUCTION

Breast cancer is the second most common cancer (1.4 million cases per year, 10.9% of diagnosed cancers) after lung cancer, followed by colorectal, stomach, prostate and liver cancers.[1] In terms of mortality, breast cancer is the fifth most common cause of cancer death. However, it place as the leading cause of cancer death among females both in western countries and in economically developing countries.[2]

Medical imaging plays an important role in breast cancer mortality reduction, contributing to its early detection through screening, diagnosis, image-guided biopsy, treatment follow-up and suchlike procedures.[3] Although Digital Mammography (DM) remains the reference imaging modality, US imaging has proven to be a successful adjunct image modality for breast cancer screening,[3,4] specially as a consequence of the discriminative capabilities that US offers for differentiating between solid lesions that are benign or malignant[5] so that the amount of unnecessary biopsies, which is estimated to be between $65 \sim 85\%$ of the prescribed biopsies,[6] can be reduced[7] in replacing them by short-term US screening follow-up.[8]

Figure . . . shows . . . what doctors look for. The image needs to be compacted and would be further used to justify the features used.

Analysing figre . . . it can be observed that most of the markers depend on the lesion delineation. Therefore in order to develop releable Computer Aided Diagnose (CAD) systems accurate segmentations to properly delineate the lesions are needed. This article presents a segmentation technique based on classifying superpixels based on their appearance.

2. SEGMENTATION METHODOLOGY DESCRIPTION

Optimization methodologies offer a standardized manner to approach segmentation by minimizing an application-driven cost function.[9] Figure 2 illustrates a generic representation of the segmentation strategy here adopted to delineate breast tissues or lesions in US images. The overall segmentation can be seen as a three steps strategy: (1) a mapping or encoding of the image into a discrete set of elements $\mathcal{S}$, (2) the optimization stage which is formulated as *metric labelling* problem, and (3) re-mapping or re-coding of the labels obtained from the previous stage to produce the final delineation.
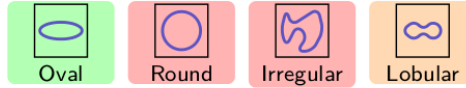
In order to formulate the segmentation like a metric labelling problem, the image is conceived as a discrete set of elements $\mathcal{S}$ that need to be labelled using a label l from the labelling set $\mathcal{L}$ (i.e. $l \in \{\text{lesion}, \overline{\text{lesion}}\}$ or $l \in \{\text{lungs}, \text{fat}, \cdots, \text{lesion}\}$). Let $\mathcal{W}$ be all the possible labelling configurations of the set $\mathcal{S}$ given $\mathcal{L}$; and, let

---

- BKGD Echotexture : adipose, fibro-glandular, heterogeneous
- Mass shape : Oval | Round | Irregular | Lobular
- Mass orientation : Parallel | Non-parallel
- Mass margin : Circumscribed | Indistinct | Angular | Microlobulated | Spiculated
- Lesion boundary : Abrupt Interface | Echogenic halo
- Echo pattern : Anechoic | Hyperechoic | Complex | Isoechoic | Hypoechoic
- Posterior acoustic pattern : Shadowing | Combined | Enhancement | No pattern
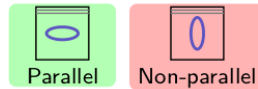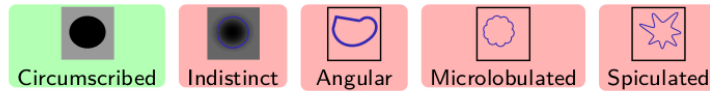
[0] benign, malignant and undetermined founds obtained from a compound of medical works[1].

[1] Sughra Raza et al. "US of Breast Masses Categorized as BI-RADS 3, 4, and 5: Pictorial Review of Factors Influencing Clinical Management". In: *Radiographics* 30.5 (2010), pp. 1199–1213

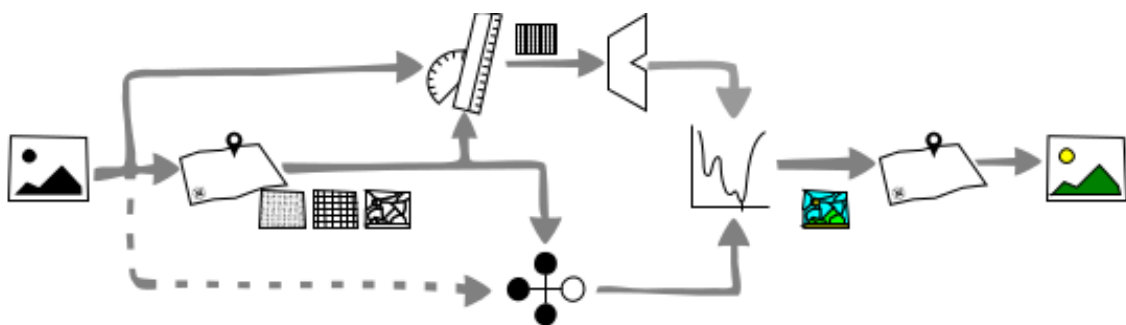Figure 1: Visual compilation of breast lesions characteristics

Figure 2: Conceptual block representation of the segmentation methodology

(a) Problem definition (b) Data term

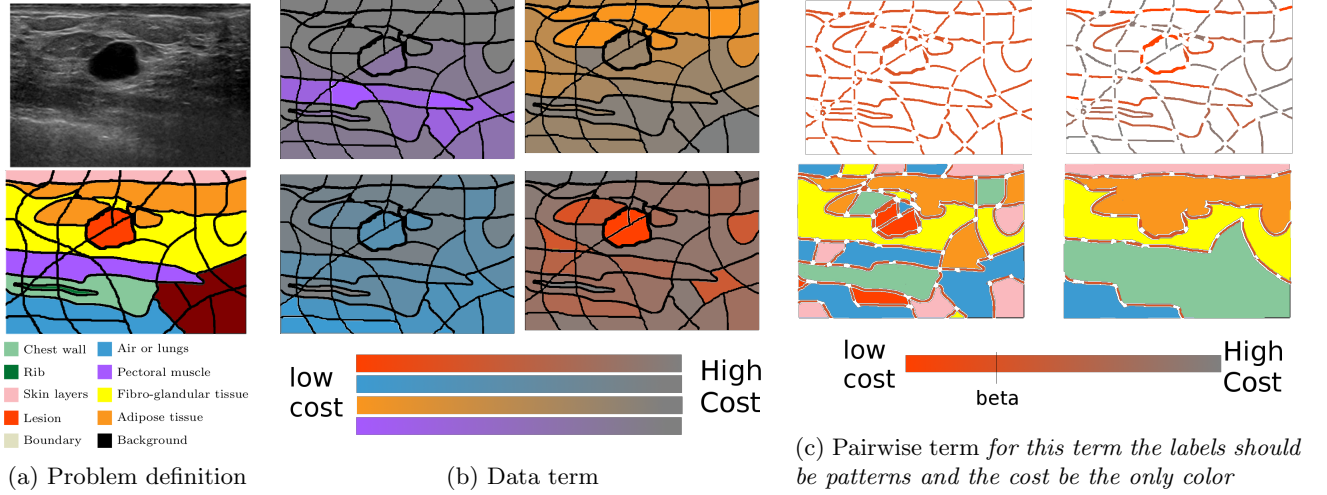(c) Pairwise term *for this term the labels should be patterns and the cost be the only color*

Figure 3: Methodology terms interpretation

$U(\cdot)$ be a cost function encoding how good is a labelling configuration $\omega \in \mathcal{W}$ based on the appearance of the elements in $\mathcal{S}$, their inner relation and some designing constrains. Then, the desired segmentation $\hat{\omega}$ corresponds to the labelling configuration that minimize this cost function, as described in **??**.

$$\hat{\omega} = \arg\min_{\omega} U(\omega) \tag{1}$$

Selecting the appropriated strategy to minimize the cost function $U(\omega)$ is part of the designing process since depending on the nature of $U(\cdot)$ and $\mathcal{W}$, not all the minimizing strategies are suitable or desirable.

?? shows the details of the cost function which combines two independent cost. Both costs are shaped by $\mathcal{S}$, evaluated in $\mathcal{W}$ and need to be simultaneously minimized as a whole. The former term $D_s(\omega_s)$, is referred to as the *data* term, while the latter, $\sum_{r \in \mathcal{N}_s} V_{s,r}(\omega_s, \omega_r)$, is indistinctly referred to as the *pairwise* or *smoothing* term.

$$U(\omega) = \sum_{s \in s} D_s(\omega_s) + \sum_{s} \sum_{r \in \mathcal{N}_s} V_{s,r}(\omega_s, \omega_r) \tag{2}$$

Figure 3 takes the study case of delineating breast tissues in US images, to relate **??** and fig. 2, and offer an interpretation their terms and elements.

Despite the fact that $\mathcal{S}$ can be any discrete set representing the image (i.e. pixels, overlapping or non overlapping windows, etc.), for this application, $\mathcal{S}$ is the super-pixels representation of the image.[10] Figure 3a shows a Breast Ultra-Sound (BUS) image example and a its associated super-pixels representation $\mathcal{S}$ coloured according to the image's Ground Truth (GT). Bear in mind, that the ultimate goal given an unseen BUS image is to represent the image as a set of super-pixels and infer the appropriated labelling for each of them.

2.1 Data term

Designing $D(\cdot)$ that accomplish the desired behaviour by defining an obscure heuristic, is rather complicated. Therefore, an easier and cleaner approach is to take advantage of Machine Learning (ML) techniques. The idea is to generate image or data model for each class, based on training samples, and let $D(\cdot)$ be a distance or goodness measure reflecting how likely is for the site s to belong to class l. Figure 2 shows how this is incorporated to the segmentation framework here proposed. Each site s is treated as a sample and the features to describe the site are extracted from the original image. For the work here reported, a Support Vector Machine (SVM) classifier is used to determine the data model during the training stage and during testing stage $D_s(\omega_s = l)$ corresponds to the distance between the testing sample and the vector supporting the data model associated to l.

Notice that defining $D(\cdot)$ in this manner allows for many designing choices such as: which features to use, how is the class model created: which classifier, which training policy; or, how is defined the relation between the testing sample and the model. Further discussion regarding the feature choices can be found in **??**, whereas other designing choices regarding ML are out of the scope for this work.

2.2 Pairwise or smoothing term

The pairwise term represents the cost of the assignation ω_s taking into account the labels of its neighbour sites, ω_r, $r \in \mathcal{N}_s$. This term models a Markov Random Fields (MRFs) or a Conditional Random Fields (CRFs). The typical form of this term, given in **??**, is called homogenization which acts as a regularization factor favouring configurations that have coherent labelling.

$$V_{s,r}(\omega_s, \omega_r) = \begin{cases} \beta, & \text{if } \omega_s \neq \omega_r \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Figure 3c offers a visual interpretation of this cost. If the resulting segmentation associated to the current labelling configuration ω has a boundary segment, this boundary brings a penalization β to the total cost $U(\omega)$. In this manner the regularization term can be seen as a post-processing or deionising stage since some sites will flip their labelling if the cost of producing and edge is larger than the cost of adopting the neighbour's label.

More sophisticated smoothing terms where boundaries contribute have different or variable costs (see fig. 3c) are also possible by taking into account not only relations in $\mathcal{S}$ of but also image information (see fig. 2). Further details can be found in **??**.

2.3 Searching the best labelling configuration

Once defined $U(\omega)$ so that the cost for a particular labelling configuration ω can be computed, the problem of finding $\hat{\omega}$ corresponding to the global minimum of the space $\mathcal{W}$ of all possible labelling configurations needs to be faced.

This problem falls into the category of NP-hard problems. The dimension of the solution space can be expressed as $||\mathcal{W}|| = ||\mathcal{L}||^{||\mathcal{S}||}$. This means that in order to perform an exhaustive search for a toy example of 20 sites and two possible labels, the cost function needs to be calculated way more than a million times. More over, due to limitations in building $U(\cdot)$ such as noise, training policies, etc. there are no guarantees that the global minimum $\hat{\omega}$ corresponds to the true labelling.

Nevertheless, there is a large body of literature proposing methodologies to find suboptimal solutions to the problem trading-off between time of convergence and accuracy of the solution reached. Szeliski et al.[11] conducted an exhaustive review in terms of solution quality and runtime of the most common energy minimization algorithms used in Computer Vision (CV), such as Iterated Conditional Modes (ICM), Simulate Anealing (SA) or Graph-Cuts (GC).

The minimization strategy used for this work is GC. A technique initially introduced to solve CV applications by Boykov et al.[12] that rapidly become the minimization technique of choice for CV problems. GC can only be applied in situations where the pairwise term is designed to favour labelling configurations which are coherent, and to penalize configurations which neighbouring sites do not share labels. The fact that GC can be applied to our cost function, allows to rapidly find a strong local minima guaranteeing that no other minimum with with lower energy can be found.[13]

2.4 Similitude with other optimization techniques

It is worth to mention here, that this pairwise term links this segmentation strategy to the family of segmentation methodologies based on optimization using Active Contour Model (ACM), such as levelsets. On its basic form, the family of ACM segmentation defines some forces to be applied to an initial contour and this contour evolves by minimizing its length while constrained by the forces properly designed for the task in hand.

3. FEATURE DESCRIPTION

Features description

ACKNOWLEDGMENTS

REFERENCES

1. J. Ferlay, H.-R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *International Journal of Cancer* **127**(12), pp. 2893–2917, 2010.

2. A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA: A Cancer Journal for Clinicians* **61**(2), pp. 69–90, 2011.

3. R. A. Smith, D. Saslow, K. A. Sawyer, W. Burke, M. E. Costanza, W. Evans, R. S. Foster, E. Hendrick, H. J. Eyre, and S. Sener, "American cancer society guidelines for breast cancer screening: update 2003," *CA: a cancer journal for clinicians* **53**(3), pp. 141–169, 2003.

4. W. A. Berg, L. Gutierrez, M. S. NessAiver, W. B. Carter, M. Bhargavan, R. S. Lewis, and O. B. Ioffe, "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer," *Radiology* **233**(3), pp. 830–849, 2004.

5. A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions," *Radiology* **196**(1), pp. 123–34, 1995.

6. Y. Yuan, M. L. Giger, H. Li, N. Bhooshan, and C. A. Sennett, "Multimodality computer-aided breast cancer diagnosis with ffdm and dce-mri.," *Academic radiology* **17**(9), p. 1158, 2010.

7. S. Ciatto, M. Rosselli del Turco, S. Catarzi, D. Morrone, *et al.*, "The contribution of ultrasonography to the differential diagnosis of breast cancer.," *Neoplasma* **41**(6), p. 341, 1994.

8. P. B. Gordon and S. L. Goldenberg, "Malignant breast masses detected only by ultrasound. A retrospective review," *Cancer* **76**(4), pp. 626–630, 1995.

9. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision* **72**(2), pp. 195–215, 2007.

10. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," 2012.

11. R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for markov random fields with smoothness-based priors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **30**(6), pp. 1068–1080, 2008.

12. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **23**(11), pp. 1222–1239, 2001.

13. A. Delong, A. Osokin, H. N. Isack, and Y. Boykov, "Fast approximate energy minimization with label costs," *International Journal of Computer Vision* **96**(1), pp. 1–27, 2012.