Breast Ultra-Sound image segmentation: an optimization approach 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

Breast cancer is the second most common cancer and the leading cause of cancer death among women. Medical imaging has become an indispensable tool for its diagnosis and follow up. During the last decade, the medical community has promoted to incorporate Ultra-Sound (US) screening as part of the standard routine. The main reason for using Ultra-Sound (US) imaging is its capability for differencing between benign and malignant masses, when compared to other imaging techniques. The increasing usage of Ultra-Sound (US) imaging encourages the development of Computer Aided Diagnose (CAD) systems applied to BUS images.

However in order to produce proper diagnosis using computer systems, accurate delineations of the lesions and structures of the breast are essential.

This article proposes a highly modular and flexible framework for segmenting lesions and tissues present in BUS images. The proposal takes advantage of optimization strategy using super-pixels and high-level descriptors, which are analogous to the visual cues used by radiologists for diagnosis. Qualitative and quantitative results are provided stating a performance within the range of the state-of-the-art.

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 is ranked as the leading cause of cancer death among females in both western countries and 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 for breast cancer screening, Ultra-Sound (US) imaging has proven to be a successful adjunct image modality.[3,4] The main advantage of US imaging, opposed to other image modalities, is its discriminative power for visually differentiate solid lesions as benign and malignant.[5] US screening contributes to reduce the amount of unnecessary biopsies,[6] which is estimated to be between $65 \sim 85\%$ of the prescribed biopsies,[7] in favour of a less traumatic short-term US screening follow-up.[8] For all these reasons, there is a growing interest in the medical community to incorporate US screening as part of the standard procedure,[9] which encourages the development of Computer Aided Diagnose (CAD) systems using US to be applied to breast cancer diagnosis.

The American College of Radiology (ACR), in order to provide a common ground for radiologists when assessing BUS images, compiled and proposed the Breast Imaging-Reporting and Data System (BI-RADS) lexicon for BUS images.[9] A lexicon is a standardized set of markers to describe the visual cues found in BUS images that are recommended to be analysed when performing image based diagnosis. This lexicon, proposed by the ACR, can be found in this document at section 2.1.1, where visual cues of BUS images and breast structures are discussed to define feature descriptors. While visual cues are discussed further in this document, it is worth to mention here that the US BI-RADS lexicon is designed to be used by expert radiologists to characterize the

Further author information: (Send correspondence to J.M.)
J.M.: E-mail: joan.massich@u-bourgogne.fr

lesions and produce a diagnosis based on the lexicon description of the lesions. This implies that radiologists, during the visual assessment of the images, locate the lesions and determine their extension prior to utilize the lexicon. Obviously, this is an intrinsic the process carried out by the trained radiologists when visually reading the images and there is no need for explicit delineation of the lesions. However, developing accurate segmentation methodologies breast lesions and structures is crucial for developing CAD systems that can take advantage of the already existing tools for characterizing the lesions.

This article proposes a highly modular and flexible framework for segmenting lesions and tissues present in BUS images. The proposal takes advantage of an energy-based strategy to perform segmentations based on discrete optimization using super-pixels and a set of novel features analogous to the elements encoded by the US BI-RADS lexicon.[9]

2. SEGMENTATION METHODOLOGY DESCRIPTION

Optimization methodologies offer a standardized manner to approach segmentation by minimizing an application-driven cost function.[10] Figure 1 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-step 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 the labels obtained from the previous stage to produce the final delineation.

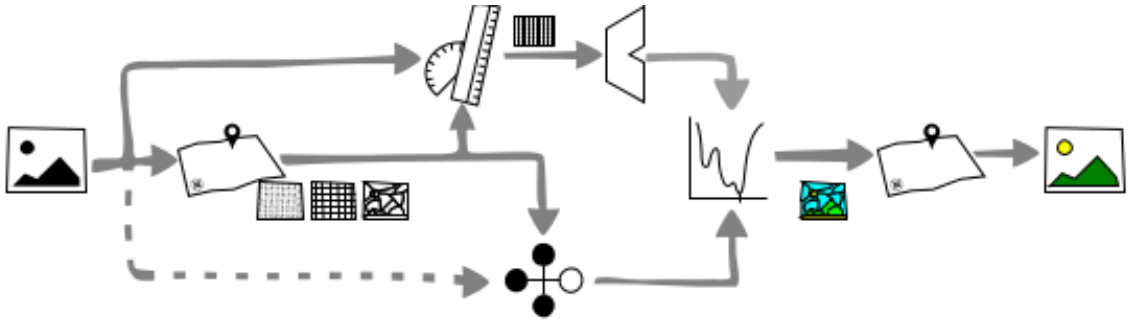


Figure 1: Conceptual block representation of the segmentation methodology

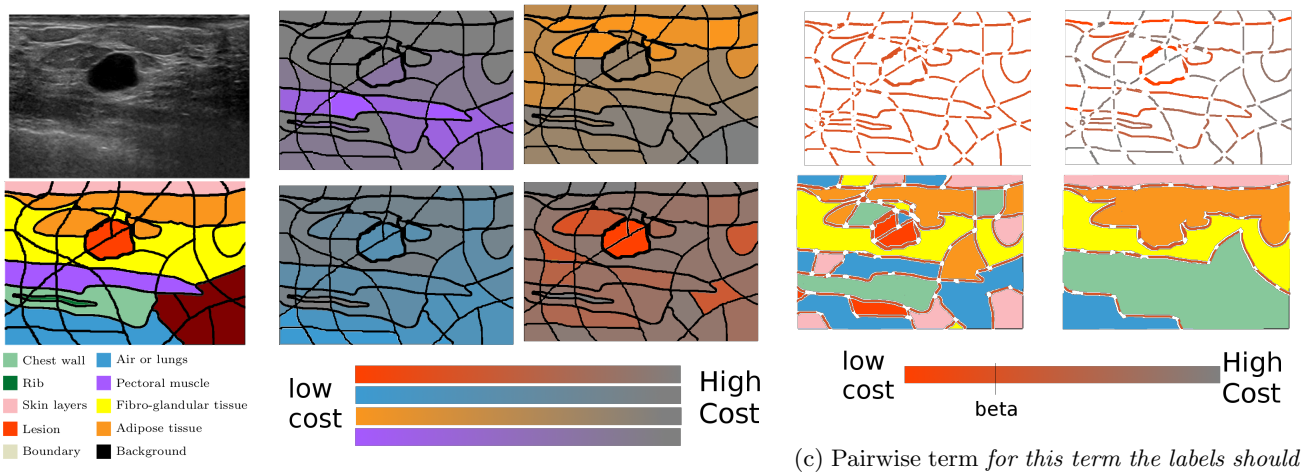
In order to formulate the segmentation like a metric labelling problem, the image is conceived as a discrete set of elements $\mathcal{S}$ that need to be labelled using a label l from the labelling set $\mathcal{L}$ (i.e. $l \in \{\text{lesion}, \overline{\text{lesion}}\}$ or $l \in \{\text{lungs}, \text{fat}, \cdots, \text{lesion}\}$). Let $\mathcal{W}$ be all the possible labelling configurations of the set $\mathcal{S}$, given $\mathcal{L}$. Let $U(\cdot)$ be a cost function encoding the goodness of the labelling configuration $\omega \in \mathcal{W}$ based on the appearance of the elements in $\mathcal{S}$, their inner relation and some designing constraints. Then, the desired segmentation $\hat{\omega}$ corresponds to the labelling configuration that minimize this cost function, as described in eq. (1).

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

The nature of $U(\cdot)$ and $\mathcal{W}$ determines which minimization strategies are applicable (or more desirable) in order to find $\hat{\omega}$; and which minimization strategies are not.

This goodness measure $U(\cdot)$ must be defined to take into account the appearance of the target region, its relation with other regions and other designing constrains. Equation (2) describes this cost function as the combination of two independent costs that need to be simultaneously minimized as a whole. The left hand side of the expression integrates the so called *data* term, while the right hand side integrates the *pairwise* term, which is also referred as the *smoothing* term. Both terms are shaped by $\mathcal{S}$ and evaluated in the labelling space $\mathcal{W}$.

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



(a) Problem definition (b) Data term (c) Pairwise term *for this term the labels should be patterns and the cost be the only colour*

Figure 2: Methodology terms interpretation

Figure 2 illustrates the terms and elements relating the framework's outline, presented in fig. 1, with its formulation in eq. (2). The problem of delineating the tissues present in BUS image has been used here for this illustrative purposes.

In general, $\mathcal{S}$ can be any discrete set representing the image (i.e. pixels, overlapping or non overlapping windows, etc.). For this work $\mathcal{S}$ is chosen to be a super-pixels representation of the image. Super-pixels can be seen as the output of a over-segmentation process or as a set of pixel collections that are contiguous and coherent with respect to some metric. Either way super-pixels are no overlapped irregular groups of similar connected pixels.[11] Figure 2a shows a BUS image example and a its associated super-pixels representation $\mathcal{S}$ coloured according to the image's Ground Truth (GT).

For the rest of this work, $\mathcal{S}$ is considered to be the super-pixels resulting from over-segmentation of the image using Quick-shift (see[12] for details).

Bear in mind that given an unseen BUS image, the ultimate goal is to represent the image as a set of super-pixels and infer the appropriated labelling for each of them. To do so using the strategy here proposed, it requires to define: a data term, a pairwise term, and a proper minimization methodology.

2.1 Data term

Given a label configuration $\omega \in \mathcal{W}$, the data term penalizes the labelling of a particular image element or site $(\omega_s = l)$ based on the data associated to s. In this manner $D_s(\omega_s = l_{\checkmark}) << D_s(\omega_s = l_{\times})$. fig. 2b shows some labelling configurations ω' where all the sites share the same label, $\omega' \in \{\omega_s = l,\ \forall s \in \mathcal{S}\}$ to illustrate the effect (or behaviour) of this data term. On it, the labelling preference of each site based on its appearance can be easily observed when comparing this labelling configurations.

Designing an obscure heuristic to comply with the desired behaviour of $D(\cdot)$ out of the box, is rather a complicated task. Therefore, an easier and cleaner approach is to take advantage of Machine Learning (ML) techniques to design this data cost in a systematic manner based on a training stage. The idea is to generate a data model for each class from training samples, and let $D(\cdot)$ be a distance or goodness measure reflecting the likelihood for s to to belong to class l. Defining the data term in this manner allows for great flexibility while offering a systematic approach towards its design. Using ML to define $D(\cdot)$ allows for customizing the features to represent the data, allows for customizing the construction of the model where several classifiers and training techniques can be applied; or even to include some arbitrary constrains.

Despite regarding the construction $D(\cdot)$ are out of the scope of this report, the rest of this section 2.1 summarizes the process. For further details the reader is referred to Massich et al.[12] The usage of ML as part of the proposed framework to determine $D(\cdot)$ is represented at the upper side of the diagram in fig. 1, where two

main blocks allow for design: (a) the features to represent the samples, and (b) the tools to encode $D(\cdot)$ based on the features and the training.

2.1.1 BUS features to build the data term

Figure 3b illustrates the BI-RADS lexicon: a standardized description of the visual cues found in BUS images, widely used by radiologists to produce a diagnosis based on the image readings.

Figure 3 relates the BI-RADS lexicon and the features designed to build $D(\cdot)$.

The following features are extracted to describe the sample s:

Appearance Based on the multi-labelled GT, a Median Absolute Deviation (MAD) histogram model for every tissue label is build. The Appearance feature is computed as the Quadratic-Chi (QC) distance between histogram of s and the models.

Atlas Based on the multi-labelled GT an atlas is build to encode the label likelihood based on the location of s.

Brightness Takes an intensity descriptor of s (*i.e:* mean, median, mode) and compares it with some intensity markers of the set $\mathcal{S}$ such as the minimum intensity value, the maximum, its mean, etc.

Self-Invariant Feature Transform (SIFT)-Back-of-Features (BoF) s is represented as the occurrences of a SIFT dictionary of 36 words.[13]

In order to incorporate multi-resolution, each super-pixel is group with its adjacent super-pixels such that $s' = \{s \cup \mathcal{N}_s\}$, the features are recalculated using s' and concatenated to the original feature descriptor. This operation can be repeated several times.

Mass shape, *Mass orientation*, and *Mass margin* are in not encoded in $D(\cdot)$ since this visual cues do not belong to a single super-pixel but to the group of contiguous super-pixels sharing the same label. Otherwise lesions and super-pixels should be elements of same order. This is not the case since the bottom-line is to label small regions to from the segmentation. For similar reasons, the *Lesion boundary* cue has been left out of the data term. In order to take the echogenic halo into account as a super-pixel appearance, those should be small enough to be contained within the halo limiting the high-level region description allowed by larger super-pixels.

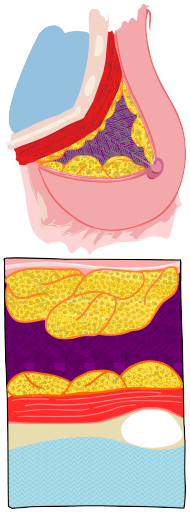
2.2 Pairwise or smoothing term

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 eq. (3), is called homogenization which acts as a regularization factor favouring configurations that have coherent labelling.

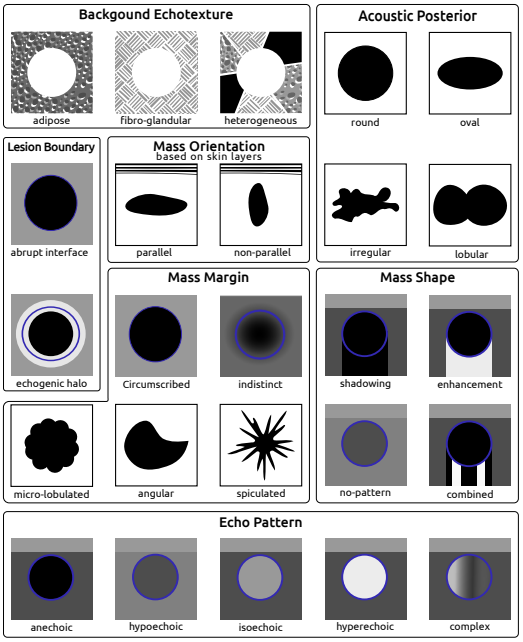
$$V_{s,r}(\omega_s, \omega_r) = \begin{cases} \beta, & \text{if } \omega_s \neq \omega_r \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Figure 2c 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 denoising 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 have different penalization based not only on site relations in $\mathcal{S}$ but also based on image information would (see fig. 2c) be found in the final version of the manuscript.



(a) Breast structure



(b) Breast lesion characteristics in US screening influencing clinical management[14]

	Background	Echo-Texture	Posterior	L.Boundary	M.Orient.	M.Margin	M.Shape	E.Pattern
Appearance		x						x
Atlas			x					x
Brightness			x					x
SIFT-BoF		x						

Figure 3: Visual reference of breast structures and visual cues used for standard BUS image assessment and diagnosis.

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. 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.[15] 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. This technique was initially introduced to solve CV applications by Boykov et al.[16] Soon after its introduction, it become the minimization technique of choice for CV problems. Since, when GC is applicable, it allows to rapidly find a strong local minima guaranteeing that no other minimum with with lower energy can be found.[17] GC is applicable if, and only if, the pairwise term favours coherent labelling configurations and penalizes labelling configurations where neighbours labels differs. Such is our case, given eq. (3).

3. RESULTS

A lack of public data to perform fair comparison between methodologies using a common benchmark is a recurrent problem in medical imaging. BUS imaging is no exception.[18] Therefore our framework can only be compared for the lesions segmentation case, and only against the results reported in the bibliography.

Figure 4 compares our segmentation strategy against the state-of-the-art methodologies assuming the following limitations: (a) the other methodologies' performance has been collected from the literature. (b) Since all the segmentation results are reported using different metrics, those have been translated to Area Overlap (AOV) as a common evaluation metric. (c) Evaluation datasets not only differ in image acquisition but also their sizes suffer a large variation.

Each radius (a .. p) represents a methodology from the literature. Those methodologies have been grouped in terms of ML, Active Contour Model (ACM), other methodologies, and combinations of those three classes. The figure reports the size of the dataset the authors have used for evaluation and also the AOV reported. Highlighted in blue there is also represented an experiment conduced by Pons et al.[19] where 50 BUS images with a single lesion were all delineated by 5 experts in order to study inter- and intra-observer variability of GT annotation. The experiment reported an AOV rate between 0.8 and 0.852 for the 6 actors, when counting the original GT accompanying the images.

Our segmentation results are represented as a black circle showing that those are within the state-of-the-art. A more meticulous analysis of the results is present in the complete version of the manuscript.

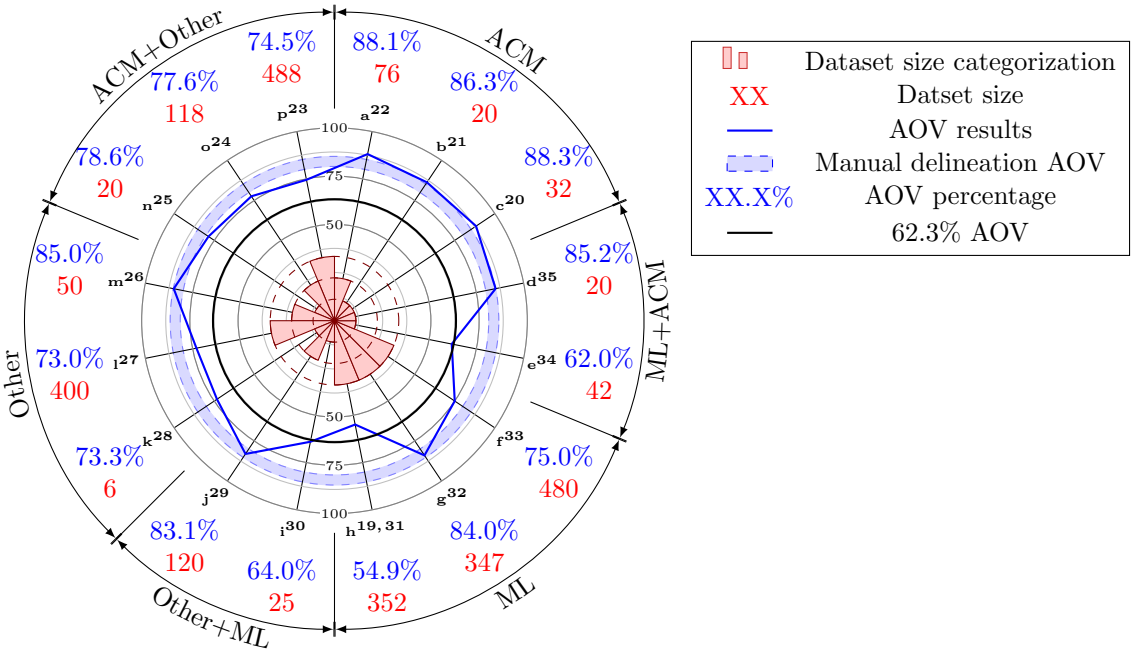


Figure 4: Quantitative AOV results

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. 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.
7. 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.
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. E. Mendelson, J. Baum, B. WA, *et al.*, *BI-RADS: Ultrasound, 1st edition in: D'Orsi CJ, Mendelson EB, Ikeda DM, et al: Breast Imaging Reporting and Data System: ACR BIRADS – Breast Imaging Atlas*, American College of Radiology, 2003.
10. 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.
11. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," 2012.
12. J. Massich i Vall *et al.*, "Deformable object segmentation in ultra-sound images," 2013.
13. J. Massich, F. Meriaudeau, M. Sentís, S. Ganau, E. Pérez, D. Puig, R. Martí, A. Oliver, and J. Martí, "Sift texture description for understanding breast ultrasound images," in *Breast Imaging*, H. Fujita, T. Hara, and C. Muramatsu, eds., *Lecture Notes in Computer Science* **8539**, pp. 681–688, Springer International Publishing, 2014.
14. S. Raza, A. L. Goldkamp, S. A. Chikarmane, and R. L. B irdwell, "US of breast masses categorized as BI-RADS 3, 4, and 5: Pictor ial review of factors influencing clinical management," *Radiographics* **30**(5), pp. 1199–1213, 2010.
15. 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.
16. 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.
17. 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.
18. H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognition* **43**(1), pp. 299–317, 2009.
19. G. Pons, J. Martí, R. Martí, S. Ganau, J. Vilanova, and J. Noble, "Evaluating lesion segmentation in breast ultrasound images related to lesion typology," *Journal of Ultrasound in Medicine* , 2013.
20. M. Alemán-Flores, L. Álvarez, and V. Caselles, "Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation," *J Math Imaging Vis* **28**(1), pp. 81–97, 2007.
21. L. Gao, X. Liu, and W. Chen, "Phase- and GVF-Based level set segmentation of ultrasonic breast tumors," *Journal of Applied Mathematics* **2012**, pp. 1–22, 2012.
22. B. Liu, H. D. Cheng, J. Huang, J. Tian, X. Tang, and J. Liu, "Probability density difference-based active contour for ultrasound image segmentation," *Pattern Recognition* , 2010.
23. J. Cui, B. Sahiner, H.-P. Chan, A. Nees, C. Paramagul, L. M. Hadjiiski, C. Zhou, and J. Shi, "A new automated method for the segmentation and characterization of breast masses on ultrasound images," *Medical Physics* **36**(5), p. 1553, 2009.
24. Y.-L. Huang, Y.-R. Jiang, D.-R. Chen, and W. K. Moon, "Level set contouring for breast tumor in sonography," *Journal of digital imaging* **20**(3), pp. 238–247, 2007.
25. Y.-L. Huang and D.-R. Chen, "Automatic contouring for breast tumors in 2-D sonography," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005*, pp. 3225–3228, IEEE, 2006.

26. W. Gómez, L. Leija, A. V. Alvarenga, A. F. C. Infantosi, and W. C. A. Pereira, "Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation," *Medical Physics* **37**(1), p. 82, 2010.
27. K. Horsch, M. L. Giger, L. Venta, and C. Vyborny, "Automatic segmentation of breast lesions on ultrasound," *Medical Physics* , 2001.
28. C. Yeh, Y. Chen, W. Fan, and Y. Liao, "A disk expansion segmentation method for ultrasonic breast lesions," *Pattern Recognition* , 2009.
29. J. Shan, H. D. Cheng, and Y. Wang, "Completely automated segmentation approach for breast ultrasound images using multiple-domain features," *Ultrasound in Medicine & Biology* **38**(2), pp. 262–275, 2012.
30. J. Massich, F. Meriaudeau, E. Pérez, R. Martí, A. Oliver, and J. Martí, "Lesion segmentation in breast sonography," *Digital Mammography* , pp. 39–45, 2010.
31. G. Xiao, M. Brady, J. A. Noble, and Y. Zhang, "Segmentation of ultrasound B-mode images with intensityinhomogeneity correction," *IEEE Transactions on medical imaging* **21**(1), pp. 48–57, 2002.
32. J. Zhang, S. K. Zhou, S. Brunke, C. Lowery, and D. Comaniciu, "Database-guided breast tumor detection and segmentation in 2D ultrasound images," in *SPIE Medical Imaging*, **7624**, pp. 762405–762405, International Society for Optics and Photonics, 2010.
33. Z. Hao, Q. Wang, Y. K. Seong, J.-H. Lee, H. Ren, and J.-y. Kim, "Combining CRF and multi-hypothesis detection for accurate lesion segmentation in breast sonograms," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pp. 504–511, Springer, 2012.
34. A. Madabhushi and D. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Transactions on medical imaging* , 2003.
35. Q.-H. Huang, S.-Y. Lee, L.-Z. Liu, M.-H. Lu, L.-W. Jin, and A.-H. Li, "A robust graph-based segmentation method for breast tumors in ultrasound images," *Ultrasonics* **52**(2), pp. 266–275, 2012.