the leading cause of cancer deaths among females in both western and economically developing countries [**?**].

Medical imaging contributes to its early detection through screening programs, non-invasive diagnosis, follow-up, and similar procedures. Despite Breast Ultra-Sound (BUS) imaging not being the imaging modality of reference for breast cancer screening [**?**], Ultra-Sound (US) imaging has more discriminative power when compared with other image modalities to visually differentiate benign from malignant solid lesions [**?**]. In this manner, US screening is estimated to be able to reduce between $65 \sim 85\%$ of unnecessary biopsies [**?**], in favour of a less traumatic short-term screening follow-up using BUS images. As the standard for assessing these BUS images, the American College of Radiology (ACR) proposes the Breast Imaging-Reporting and Data System (BI-RADS) lexicon for BUS images [**?**]. This US BI-RADS lexicon is a set of standard markers that characterizes the lesions encoding the visual cues found in BUS images and facilitates their analysis. Further details regarding the US BI-RADS lexicon descriptors proposed by the ACR, can be found in Sect. **??**, where visual cues of BUS images and breast structures are discussed to define feature descriptors.

The incorporation of US in screening policies and the emergence of clinical standards to assess image like the US BI-RADS lexicon, encourage the development of Computer Aided Diagnosis (CAD) systems using US to be applied to breast cancer diagnosis. However, this clinical assessment using lexicon is not directly applicable to CAD systems. Shortcomings like the location and explicit delineation of the lesions need to be addressed, since those tasks are intrinsically carried out by the radiologists during their visual assessment of the images to infer the lexicon representation of the lesions. Therefore, developing accurate segmentation methodologies for breast lesions and structures is crucial to take advantage of this already validated clinical tools.

This paper 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 segmentation based on discrete optimization using super-pixels and a set of novel features analogous to the elements encoded by the US BI-RADS lexicon [**?**].

2 Description of the segmentation methodology

Optimization methodologies offer a standardized manner to approach segmentation by minimizing an application-driven cost function [**?**]. Figure **??** illustrates a generic representation of the segmentation strategy, concrete examples of its terms, applied to BUS, can be found in **??**. The overall segmentation can be seen as a three-steps strategy: (1) a mapping of the image into a discrete set of elements $\mathcal{S}$, (2) the optimization stage which is formulated as a *metric labelling* problem, and (3) a re-mapping the labels obtained from the previous stage to produce the final delineation.

To formulate segmentation as 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

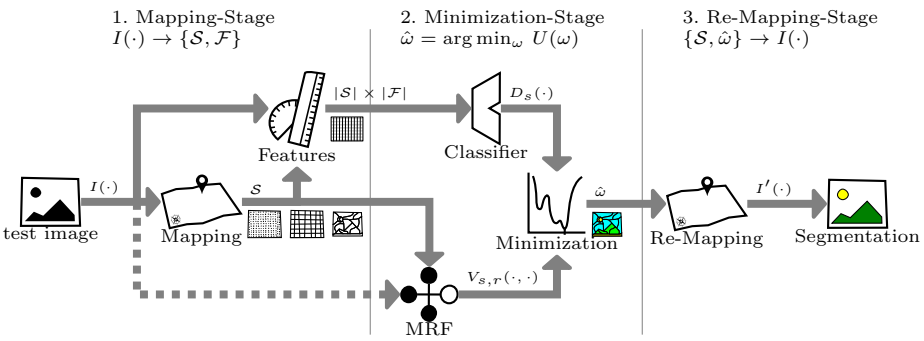

Fig. 1: Conceptual block representation of the segmentation methodology.

from the labelling set $\mathcal{L}$. 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 minimizes this cost function, as described in Eq. (**??**).

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

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 constraints. Equation (**??**) describes this cost function as the combination of two independent costs that need to be simultaneously minimized as a whole.

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

Where, 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 to as the *smoothing* term. Both terms are shaped by $\mathcal{S}$ and evaluated in the labelling space $\mathcal{W}$. In our quest to optimize the cost function $U(\cdot)$, it is required to define a representation for the set $\mathcal{S}$, a data term $D(\cdot)$, a pairwise term $V(\cdot)$, and a proper minimization methodology.

The set $\mathcal{S}$ can be, in general, any discrete set representing the image (i.e. pixels, overlapping or non overlapping windows, super-pixels, etc.).

The data term $D(\cdot)$, given a label configuration $\omega \in \mathcal{W}$, penalizes the labelling of a particular image element or site ($\omega_s = l$) based on the data associated to s. In this manner, $D_s(\omega_s = l_{\checkmark}) << D_s(\omega_s = l_{\times})$. Figure **??**b illustrates the data cost associated to some arbitrary labelling configurations to clarify the desired effect (or behaviour) of this data term. Designing an obscure heuristic to comply with the desired behaviour of $D(\cdot)$ out of the box, is rather a complicated task.

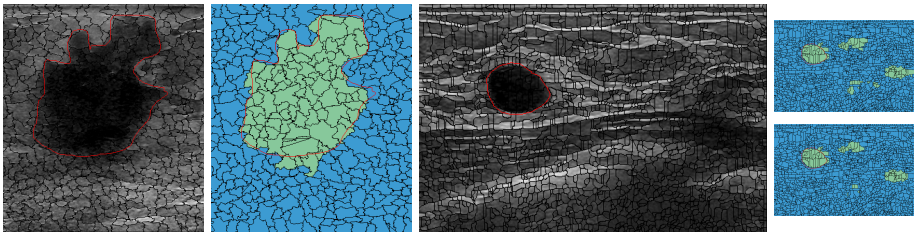

Fig. 4: Qualitative results. (a) Example 1: orignal image, super-pixels' delineations and GT. (b) Differences between GT and the delination resulting from super-pixels' boundary. (c) Ex. 2. (d) weak $V(\cdot,\cdot)$ (e) strong $V(\cdot,\cdot)$

Brightness Intensity descriptors are computed based on statistics of s (*i.e:* mean, median, mode) and are compared 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 described as a histogram of visual words based on SIFT [**?**]. The dictionary is built with 36 words.

The relationships between the lexicon and the descriptors described previously are depicted in Table **??**. More precisely, we highlight the corresponding elements of the lexicon which is encoded by each feature. A choice regarding the encoding of the data term $D(\cdot)$ has to be made by using a ML classifier. Support Vector Machines (SVM) classifier with Radial Basis Function (RBF) kernel is selected to determine the data model during the training stage. The pairwise term in our framework was defined as in Eq. (**??**). The optimization method used as a solver to minimize our cost function $U(\cdot)$ is Graph-Cuts (GC). GC, where appropriately applied, allows to rapidly find a strong local minima guaranteeing that no other minimum with lower energies can be found [**?**]. GC is applicable if, and only if, the pairwise term favours coherent labelling configurations and penalizes labelling configurations where neighbours' labels differ such as in our case, given by Eq. (**??**).

Table 1: Design choices summary

$\mathcal{S}$	Quick-Shift super-pixels
	Background Echotexture: encoded in Appearance and SIFT-BoW
$D(\cdot)$	Echo Pattern: encoded in Appearance, Atlas and Brightness
	Acoustic Posterior: encoded in Atlas and Brightness
$V(\cdot,\cdot)$	homogeneity as Eq. (**??**)
$\arg\min U(\cdot)$	Graph-Cuts

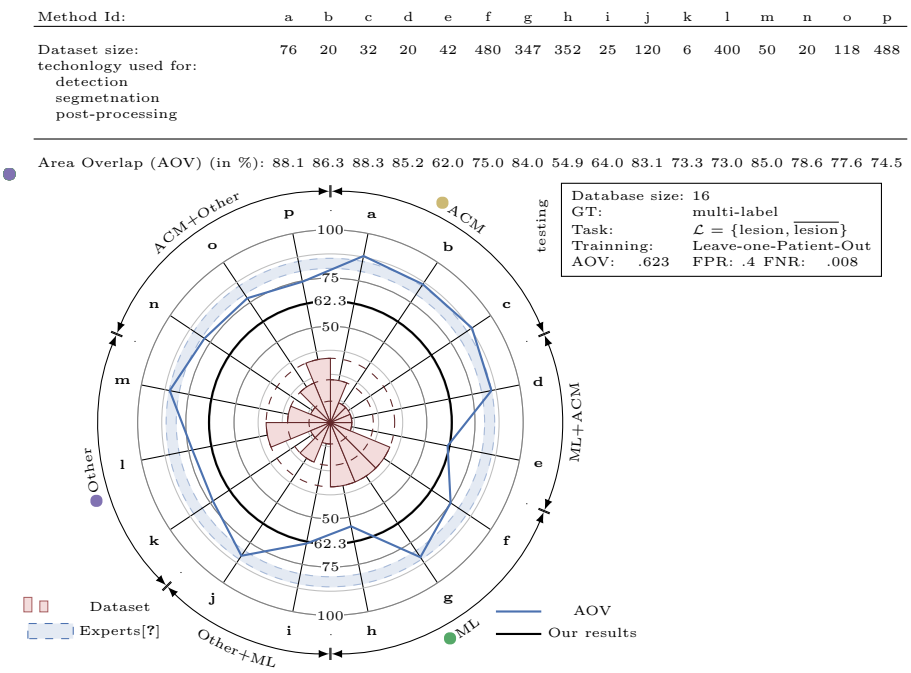

Fig. 5: Quantitative results compilation and comparison

4 Method evaluation and comparison

The proposed methodology is evaluated using a dataset of 16 BUS images presenting a single lesion of variable extension. The size of the lesions ranges from under 1/100 to over 1/5 of the image size. The dataset is composed of cysts, Fibro-Adenomas (FAs), Ductal Inflating Carcinomas (DICs) and Inflating Lobular Carcinomas (ILCs). Every image has accompanying multi-label GT delineating all the depicted structures. This dataset is now publicly available at `http://visor.udg.edu/dataset/#breast`

The lack of publicly available data and source code, limits the comparison between the different methods. For this study, the results published by the other authors have been collected and expressed in terms of AOV in order to share a common metric. Further details can be found in [**?**] and summarized in Fig. **??**.

Figure **??** is divided into three main parts: (i) a table on the top summarizes the core stages of each study framework, (ii) a legend box on the right side informs our testing setup and, (iii) a comparison of the different metrics in a radial manner. An extra element is also represented in this radial representation: a blue swatch delimited by two blue dashed lines. The boundaries of this swatch correspond to the performance of some expert radiologists based on an inter- and intra-observer experiments carried out by Pons et al. [**?**]. It is interesting

to note that some methodologies outperform this swatch. A publicly available dataset should allow a better comparison in that regard.

The results point out the inherent capabilities of ML to cope with data scalability and variability, induce its usage in conjunction with larger datasets. Whereas, Active Contour Model (ACM) methodologies show its effectiveness to model the boundary in a natural manner.

For our proposed framework, the performance in terms of AOV lies within the state-of-the-art despite its final delineation limited by the capacity of the super-pixels to snap the desired boundary. Figure **??** shows some qualitative results where there are limitations of labeling super-pixels when compared with hand-drawn GT. Figure **??** also illustrates the influence of the pair-wise term.

5 Conclusions

This work presents a segmentation strategy to delineate lesions in BUS images using an optimization framework that takes advantage of all the facilities available when using ML techniques. Despite the limitation that the final segmentation is subject to the super-pixels' boundaries, the AOV results reported here are similar to those reported by other methodologies in the literature. A higher AOV result can be achieved by refining the delineation resulting from our proposed framework by post-processing it with an ACM. In this manner, the contour constraints could be applied to achieve a more natural delineation.

References

1. Achanta, R., et al.: SLIC superpixels compared to state-of-the-art superpixel methods (2012)
2. Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. International journal of computer vision 72(2) (2007)
3. Delong, A., Osokin, A., Isack, H.N., Boykov, Y.: Fast approximate energy minimization with label costs. International Journal of Computer Vision 96(1), 1–27 (2012)
4. Jemal, A., et al.: Global cancer statistics. CA: A Cancer Journal for Clinicians 61 (2011)
5. Massich, J.: Deformable object segmentation in ultra-sound images. Ph.D. thesis
6. Massich, J., et al.: Sift texture description for understanding breast ultrasound images. In: Breast Imaging, Lecture Notes in Computer Science, vol. 8539, pp. 681–688. Springer International Publishing (2014)
7. Mendelson, E., Baum, J., WA, B., 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)
8. Pons, G., Martí, J., Martí, R., Ganau, S., Vilanova, J., Noble, J.: Evaluating lesion segmentation in breast ultrasound images related to lesion typology. Journal of Ultrasound in Medicine (2013)

9. Smith, R.A., et al.: American cancer society guidelines for breast cancer screening: update 2003. CA: a cancer journal for clinicians 53(3), 141–169 (2003)
10. Stavros, A.T., Thickman, D., Rapp, C.L., Dennis, M.A., Parker, S.H., Sisney, G.A.: Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions. Radiology 196(1), 123–34 (1995)
11. Yuan, Y., Giger, M.L., Li, H., Bhooshan, N., Sennett, C.A.: Multimodality computer-aided breast cancer diagnosis with ffdm and dce-mri. Academic radiology 17(9), 1158 (2010)