

# Automatic Nonlinear Filtering and Segmentation for Breast Ultrasound Images

Mohamed Elawady<sup>1(✉)</sup>, Ibrahim Sadek<sup>2</sup>, Abd El Rahman Shabayek<sup>3</sup>,  
Gerard Pons<sup>4</sup>, and Sergi Ganau<sup>5</sup>

<sup>1</sup> Universite Jean Monnet, CNRS, UMR 5516,  
Laboratoire Hubert Curien, 42000 Saint-Etienne, France  
[mohamed.elawady@univ-st-etienne.fr](mailto:mohamed.elawady@univ-st-etienne.fr)

<sup>2</sup> Image and Pervasive Access Lab, CNRS UMI 2955, Singapore, Singapore

<sup>3</sup> Department of Computer Science, Faculty of Computers and Informatics,  
Suez Canal University, Ismailia, Egypt

<sup>4</sup> Department of Computer Architecture and Technology,  
University of Girona, Girona, Spain

<sup>5</sup> Radiology Department, UDIAT Centre Diagnostic, Sabadell, Spain

**Abstract.** Breast cancer is one of the leading causes of cancer death among women worldwide. The proposed approach comprises three steps as follows. Firstly, the image is preprocessed to remove speckle noise while preserving important features of the image. Three methods are investigated, i.e., Frost Filter, Detail Preserving Anisotropic Diffusion, and Probabilistic Patch-Based Filter. Secondly, Normalized Cut or Quick Shift is used to provide an initial segmentation map for breast lesions. Thirdly, a postprocessing step is proposed to select the correct region from a set of candidate regions. This approach is implemented on a dataset containing 20 B-mode ultrasound images, acquired from UDIAT Diagnostic Center of Sabadell, Spain. The overall system performance is determined against the ground truth images. The best system performance is achieved through the following combinations: Frost Filter with Quick Shift, Detail Preserving Anisotropic Diffusion with Normalized Cut and Probabilistic Patch-Based with Normalized Cut.

**Keywords:** Breast cancer · Lesion segmentation · Ultrasound imaging · Speckle noise removal · Nonlinear filtering

## 1 Introduction

In 2012, breast cancer is ranked as the second most frequent cancer among women worldwide. The breast cancer incidence and mortality have been raised by more than 20 % and 14 % respectively, since the 2008 estimates [1, 2]. Early detection and diagnosis can significantly increase the breast cancer survival. Consequently, there is an insisting demand for developing accurate and affordable computer aided diagnosis (CADx) systems. Mammography is the most widely used imaging modality for predicting breast cancer in women. However, it introduces some limitations such as radiation risk, risk of false alarm, over-diagnosis or

over-treatment, and limited sensitivity [3]. Fortunately, the introduction of ultrasound (US) medical imaging modality has helped to reduce these side effects. In general, breast ultrasound (BUS) imaging has advantages in terms of safety, cost, sensitivity, and accuracy over the conventional mammography. Although its own advantages, it requires professional radiologists. Therefore, radiologists are in need of CADx systems to help detect and analyze breast cancer. In the previous work, most of CADx systems consist of three main steps, i.e., preprocessing, feature extraction, classification [4, 5].

Liu et al. [6] performed anisotropic diffusion filtering to eliminate speckle noise besides unsharp masking for edge enhancement. The filtered and enhanced image is fed into a Normalized Cut (NC) to get multiple small regions. Subsequently, alongside regions are merged into several bigger regions using region merging. Finally, potential lesions are extracted by morphological operations. Quan et al. [7] proposed to use region based NC instead of pixel based NC. In the first stage, a sigmoid filtering is applied to reinforce the differences between the ROI and the background, afterward the filtered image is divided into a group of over-segmented regions by a linear iterative clustering algorithm, where the over-segmented regions are treated as nodes rather than pixels. In the final step, the NC is applied to merge the over-segmented regions and to segment the ROI. A semi-automatic approach is introduced by Zhou et al. [8] to segment lesions in BUS images, Gaussian filter and histogram equalization are utilized to smooth and to enhance image contrast. Then, a pyramid mean shift filtering is applied to improve the homogeneity of the enhanced image. The final segmentation step is achieved, through NC and morphological operations.

This work contributes with (1) a comparative study on the most common preprocessing nonlinear techniques [9], and (2) a lesion isolation algorithm by means of Quick Shift. The paper is organized as follows. In Sect. 2, the fully automatic lesion extraction algorithm is introduced; in Sect. 3, the experimental results for proposed methods are shown and discussed. Finally, Sect. 4 concludes the work.

## 2 Framework

The ultrasound image contrast between the abnormality and the surrounding breast tissue is insufficient for direct lesion detection. This inherent difficulty in detection requires a considerable amount of processing to isolate candidate tumor regions. The process followed in this work can be summarized as follows: (a) the input images are preprocessed in Subsect. 2.1 to reduce the existing noise and to improve the contrast between the lesion and its surrounding, (b) the preprocessed images are segmented in Subsect. 2.2 to seek candidate lesions, and finally (c) the segmentation results are postprocessed in Subsect. 2.3 to remove any extra noisy regions. The suggested segmentation framework is presented in Fig. 1 and detailed in the following subsections.

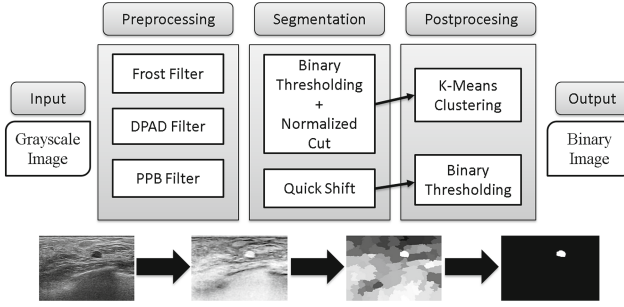


Fig. 1. The proposed framework used for lesion segmentation in BUS images.

### 2.1 Preprocessing

Noise reduction is a typical preprocessing step in BUS images to improve the results of later processing. Ultrasound imaging system is a coherent imaging system that produces images which suffer from speckle noise, primarily due to the interference of the returning wave at the transducer aperture [10], that degrades its quality. Several methods are used to eliminate speckle noise according to different mathematical models of the speckle phenomenon.

**Frost Filter (FR).** Frost et al. [11] proposed a local adaptive filter designed to despeckle images based on the local statistics. The noisy image is modeled as follows:

$$f(x, y) = [g(x, y) \cdot n(x, y)] * h(x, y), \tag{1}$$

where  $f(x, y)$ ,  $g(x, y)$ ,  $n(x, y)$ , and  $h(x, y)$  denote noisy image, ideal noise-free image, noise, and impulse response in spatial coordinate  $(x, y)$ , respectively. It relies upon three fundamental assumptions in its mathematical model: (1) Speckle noise is in direct proportion to the local grey level in any area, (2) The signal and the noise are statistically independent of each other, and (3) The sample mean and variance of a single pixel are equal to the mean and variance of the local area that is centered on that pixel.

**Detail Preserving Anisotropic Diffusion (DPAD).** Anisotropic diffusion technique was firstly introduced by Perona and Malik [12]. Speckle Reducing Anisotropic Diffusion (SRAD) filter was then proposed by Yu and Acton [13] to remove the speckle noise without removing significant parts of the image content, typically edges, lines or other details that are important for the interpretation of the image. Aja-Fernández and Alberola-López [14] has improved the speckle statistical properties estimation of [13]. The model equation is as follows:

$$I_p^{t+\Delta t} = I_p^t + \frac{\Delta t}{|\eta_p|} \text{div}[c(C_{p,t}) \nabla I_p^t] \tag{2}$$

where  $I_p^t$  indicates the discrete image at  $p$  coordinate position.  $t$  represents the time step, and  $|\eta_p|$  is the number of pixels in the window.  $\nabla I_p^t$  indicates the gradient value, and  $C_{p,t}$  represents the ratio between the local standard deviation and the local mean.

**Probabilistic Patch-Based (PPB) Filter.** Deledalle et al. [15] recently proposed a nonlocal means filter that performs a weighted average of the values of similar patches. These weights can be iteratively calculated based on both the similarity of noisy patches and their previous estimated similarity. The estimate value  $\hat{f}_s^i(c)$  of the center pixel  $c$  at the patch  $s$  at  $i$ th iteration can be obtained by performing a weighted average  $w$  of all the pixels in the image,

$$\hat{f}_s^i(c) = \sum_t w(s,t) f_t(c), \quad (3)$$

where  $f_t(c)$  is the center pixel at the patch  $t$ . The iteration is repeated until there is no more change between two consecutive estimates.

## 2.2 Segmentation

Image segmentation is the process of partitioning the image into multiple segments (e.g. set of pixels or superpixels). The goal of segmentation is to isolate the lesion(s) from its surrounding pixels to put it into a more meaningful representation for doctors where normal and suspected regions are identified. In this work, Quick Shift is proposed as a robust fast segmentation step with significant effect in the breast cancer detection framework. The significance of our results is shown by comparing it to Normalized Cut which has been frequently used for the same purpose [6, 16].

**Quick Shift (QS).** Quick Shift is based on an approximation of kernelized Mean-Shift. It is a local mode-seeking algorithm and is applied to the 5D space consisting of color information and image location. QS computes a hierarchical segmentation on multiple scales simultaneously and iteratively forms a tree of links to the nearest neighbor [17].

For each pixel  $(x, y)$ , QS regards  $(x, y, I(x, y))$  as a sample from a  $d+2$  dimensional vector space. It then calculates the Parzen density estimate  $P(x, y, I(x, y))$  with a Gaussian kernel [18]. Then QS constructs a tree connecting each image pixel  $(x, y)$  to its nearest neighbor  $(x', y')$  which has greater density value where  $(x', y') > (x, y) \Leftrightarrow P(x', y', I(x', y')) > P(x, y, I(x, y))$ . Each pixel is connected to its closest higher density pixel.

**Normalized Cut (NC).** Normalized Cut is a graph partitioning problem based on a global criterion for segmentation that measures both the total similarity

within the groups and the total dissimilarity between the different groups. It is based on a generalized eigenvalue problem used to optimize the NC criterion [19]. Gao et al. [16] proposed using NC after a boundary-detection function which combines texture and intensity information. Their algorithm defines a homogeneous patch for each pixel using the boundary map from the boundary-detection function.

### 2.3 Postprocessing

Lesion selection is a postprocessing step to find a correct lesion per image. The main reason of this step is choosing only one non-boundary region with the highest contrast value and the largest area among other candidate regions in global neighborhood. The eigenvectors of Normalized Cut segmentation are applied to k-means clustering method [20] to find lesion candidates across these different eigenvectors, while the Quick Shift easily selects the correct lesion by applying the gray-scale output to empirical binary thresholding.

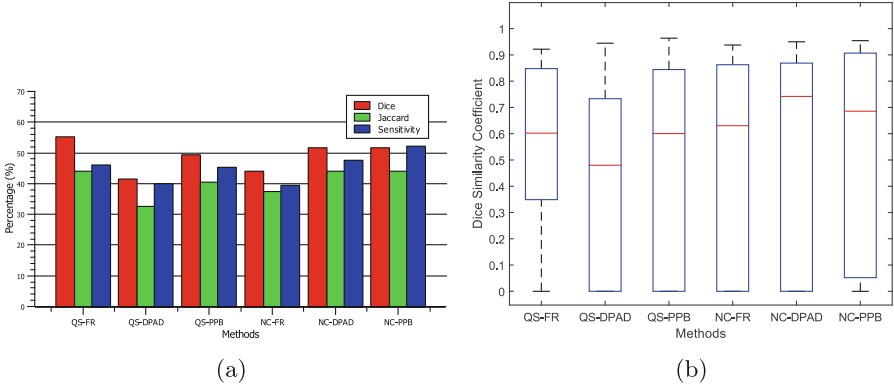
## 3 Experiments and Discussion

A set of 20 breast B-mode ultrasound images [21] has been collected from different patients through UDIAT Diagnostic Center of Sabadell (Spain) with a Siemens ACUSON Sequoia C512 system 17L5 HD linear array transducer (8.5 MHz). Ground truth assessments are provided by an experienced radiologist to delineate any kind of lesion inside the images. For the proposed preprocessing algorithms, default parameter values of [FR, DPAD, PPB] are stably assigned, as stated in the very recent survey by Zhang et al. [9] to compare despeckle filters for breast ultrasound images. For the proposed segmentation methods, source codes of [QS<sup>1</sup>, NC<sup>2</sup>] are publicly available. The parameters were empirically selected from half of the dataset. The second half was tested with these parameters. The reported results show that there is no need to further tune any of them and they can be directly applied to any new image. Values of the QS parameters (color/spatial ratio, kernel size, and maximum distance between pixels) are 0.8, 5 and 20 respectively. The input image of NC is empirically binarized after the preprocessing process, and its parameter (number of segments) is set to 4.

In order to evaluate the proposed methods quantitatively, three well-known statistical measures are used: Dice similarity coefficient, Jaccard similarity index (aka Area Overlap), and Sensitivity (aka True Positive Rate or Recall). Previous works [4–8] described different segmentation methods to discriminate bright lesions within BUS images. It should be noted that a fair comparison between methods is hard to make since results are based on different datasets. The proposed methods are implemented using MATLAB (R2014b, MathWorks Inc., MA) on a windows-based PC platform (Intel core i7-3630QM, 2.4 GHz and 8 GB

<sup>1</sup> <http://www.vlfeat.org/overview/quickshift.html>.

<sup>2</sup> <http://www.timotheecour.com/software/ncut/ncut.html>.

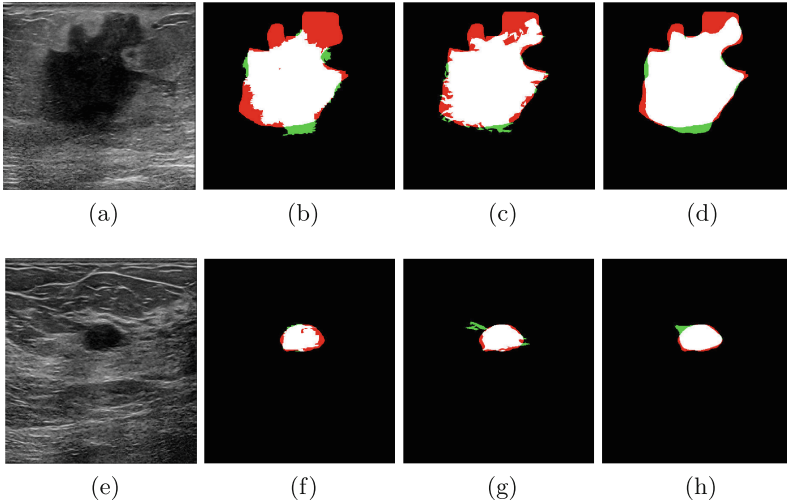


**Fig. 2.** Performance results across all proposed methods (segmentation [QS: Quick Shift and NC: Normalized Cut] and preprocessing [FR: Frost Filter, DPAD: Detail Preserving Anisotropic Diffusion and PPB: Probabilistic Patch-Based]): (a) Statistical metrics of Dice, Jaccard and Sensitivity measures calculated in average across all dataset images. (b) Box plot of Dice similarity coefficient. (Color figure online)

RAM). For  $360 \times 528$  image, the run time of Quick Shift method (7.92 s) is  $8\times$  faster than Normalized Cut method (65.51 s). Figure 2(a) shows analytical results of the proposed methods (segmentation, preprocessing). The following methods (Quick Shift with Frost Filter [QS-FR], Normalized Cut with Detail Preserving Anisotropic Diffusion [NC-DPAD] and Normalized Cut with Probabilistic Patch-Based [NC-PPB]) achieve best results among variant metrics (especially in Jaccard similarity coefficient) with slightly difference in comparison. Dice similarity coefficient shows superior achievement of [QS-FR] against other best candidates [NC-DPAD, NC-PPB], while the Sensitivity of NC-PPB is the best among others.

Figure 2(b) describes box plot of Dice Jaccard similarity coefficient for all proposed methods. [QS-FR] is less distributed among other methods (especially compared to the other candidates: [NC-DPAD, NC-PPB]), although the later methods achieve the best median results respectively. The failure cases exist in all methods due to the intensity similarity of surrounding tissues around the target lesion, leading to incorrect segmentation behavior without any prior knowledge while executing the preprocessing phase.

Figure 3 displays qualitative results of corrected lesion extraction for the best proposed methods [QS-FR, NC-DPAD, NC-PPB], such that similar acceptable outcomes are computed leading to decide the best method to be [QS-FR] as shown in the previously stated quantitative results. To sum up, QS needs a non-complex despeckle filter (FR) with efficient computation perspective to get a proper segmentation result.



**Fig. 3.** Results of some successful lesion extraction for best proposed candidates. First column represents some of the input images. Second, third and fourth columns show the output results of [QS-FR, NC-DPAD, NC-PPB] respectively, in which white color is true segmented lesion, green color is false positive, red color is false negative and black color is true negative. (Color figure online)

## 4 Conclusion

In this paper, an automatic approach is proposed to detect breast lesions in ultrasound images. Three different filtering methods are analyzed for speckle noise reduction: FR, DPAD and PPB. NC and QS are used for breast lesions segmentation followed by a postprocessing step to select the correct candidate region from the output of segmentation step. The quantitative results are computed as average across all images such that our best performance is conducted through FR with QS, DPAD with NC and PPB with NC. The first combination is superior in terms of computational complexity, thereby it is a more preferable choice in real time applications. In the future, we would like to increase the dataset size and to use superpixel segmentation approaches to obtain more accurate and robust results.

## References

1. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., Bray, F.: Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136**(5), E359–E386 (2015)
2. Bray, F., Ren, J.S., Masuyer, E., Ferlay, J.: Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int. J. Cancer* **132**(5), 1133–1145 (2013)

3. Heywang-Köbrunner, S.H., Hacker, A., Sedlacek, S.: Advantages and disadvantages of mammography screening. *Breast Care* **6**(3), 199–207 (2011)
4. Cheng, H., Shan, J., Ju, W., Guo, Y., Zhang, L.: Automated breast cancer detection and classification using ultrasound images: a survey. *Pattern Recognit.* **43**(1), 299–317 (2010)
5. Noble, J.A., Boukerroui, D.: Ultrasound image segmentation: a survey. *IEEE Trans. Med. Imaging* **25**(8), 987–1010 (2006)
6. Liu, X., Huo, Z., Zhang, J.: Automated segmentation of breast lesions in ultrasound images. In: *27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS 2005*, pp. 7433–7435. IEEE (2006)
7. Quan, L., Zhang, D., Yang, Y., Liu, Y., Qin, Q.: Segmentation of tumor ultrasound image via region-based ncut method. *Wuhan Univ. J. Nat. Sci.* **18**(4), 313–318 (2013)
8. Zhou, Z., Wu, W., Wu, S., Tsui, P.H., Lin, C.C., Zhang, L., Wang, T.: Semi-automatic breast ultrasound image segmentation based on mean shift and graph cuts. *Ultrason. Imaging* **36**(4), 256–276 (2014)
9. Zhang, J., Wang, C., Cheng, Y.: Comparison of despeckle filters for breast ultrasound images. *Circuits Syst. Signal Process.* **34**(1), 185–208 (2015)
10. Forouzanfar, M., Abrishami-Moghaddam, H.: Ultrasound Speckle Reduction in the Complex Wavelet Domain. In: *Principles of Waveform Diversity and Design*. SciTech Publishing an imprint of the IET, pp. 558–577 (2010)
11. Frost, V.S., Stiles, J.A., Shanmugan, K., Holtzman, J.: A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Pattern Anal. Mach. Intell. PAMI* **4**(2), 157–166 (1982)
12. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639 (1990)
13. Yu, Y., Acton, S.: Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.* **11**, 1260–1270 (2002)
14. Aja-Fernández, S., Alberola-López, C.: On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering. *IEEE Trans. Image Process.* **15**(9), 2694–2701 (2006)
15. Deledalle, C.A., Denis, L., Tupin, F.: Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. *IEEE Trans. Image Process.* **18**(12), 2661–2672 (2009)
16. Gao, L., Yang, W., Liao, Z., Liu, X., Feng, Q., Chen, W.: Segmentation of ultrasonic breast tumors based on homogeneous patch. *Med. Phys.* **39**(6), 3299–3318 (2012)
17. Vedaldi, A., Soatto, S.: Quick shift and kernel methods for mode seeking. In: *European Conference on Computer Vision* (2008)
18. Parzen, E.: On estimation of a probability density function and mode. *Ann. Math. Stat.* **33**(3), 1065–1076 (1962)
19. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888–905 (2000)
20. Vedaldi, A., Fulkerson, B.: VLFeat: an open and portable library of computer vision algorithms. In: *Proceedings of the International Conference on Multimedia*, pp. 1469–1472. ACM (2010)
21. Pons, G., Martí, J., Martí, R., Ganau, S., Vilanova, J.C., Noble, J.A.: Evaluating lesion segmentation on breast sonography as related to lesion type. *J. Ultrasound Med.* **32**(9), 1659–1670 (2013)