Tumour Ellipsification in Ultrasound Images for Treatment Prediction in Breast Cancer

Mehrdad J. Gangeh, *Senior Member, IEEE,* Hamid R. Tizhoosh,
Kan Wu, Dun Huang, Hadi Tadayyon, and Gregory J. Czarnota

Abstract— Recent advances in using quantitative ultrasound (QUS) methods have provided a promising framework to non-invasively and inexpensively monitor or predict the effectiveness of therapeutic cancer responses. One of the earliest steps in using QUS methods is contouring a region of interest (ROI) inside the tumour in ultrasound B-mode images. While manual segmentation is a very time-consuming and tedious task for human experts, auto-contouring is also an extremely difficult task for computers due to the poor quality of ultrasound B-mode images. However, for the purpose of cancer response prediction, a rough boundary of the tumour as an ROI is only needed. In this research, a semi-automated tumour localization approach is proposed for ROI estimation in ultrasound B-mode images acquired from patients with locally advanced breast cancer (LABC). The proposed approach comprised several modules, including 1) feature extraction using keypoint descriptors, 2) augmenting the feature descriptors with the distance of the keypoints to the user-input pixel as the centre of the tumour, 3) supervised learning using a support vector machine (SVM) to classify keypoints as "tumour" or "non-tumour", and 4) computation of an ellipse as an outline of the ROI representing the tumour. Experiments with 33 B-mode images from 10 LABC patients yielded promising results with an accuracy of 76.7% based on the Dice coefficient performance measure. The results demonstrated that the proposed method can potentially be used as the first stage in a computer-assisted cancer response prediction system for semi-automated contouring of breast tumours.

I. IDEA AND MOTIVATION

Precision medicine is an emerging technique that tailors medical treatment depending on the characteristics of each patient [1]. For example, in cancer therapy, the main goal of precision medicine is to decide on the most effective cancer therapy for a patient based on his/her response to treatment administrations. To this end, the first step in personalized cancer therapy is to either predict or provide a means to assess the individual responses to treatment early during the course of a therapy. There are several functional imaging modalities such as magnetic resonance imaging (MRI), diffuse optical spectroscopy (DOS), and positron emission tomography (PET) that can provide imaging at a microscopic level to detect cell death [2]. Two main limitations of these imaging technologies include: the requirements for a large capital investment and an external agent. The latter is also expensive, and may cause some side effects and/or allergic reactions [3]. In contrast, quantitative ultrasound (QUS) methods provide a portable, non-expensive, and non-invasive means for a rapid acquisition of functional images that can be used for an early assessment of cancer therapeutic effects [3]. Moreover, in QUS methods, the endogenous contrast – generated by the very process of cell death – is employed in treatment assessment, which alleviates the requirement for injecting external agents.

The applications of QUS methods have recently been extended from cancer response monitoring [4] to cancer response prediction [5], and tissue characterization/visualization using 3-D automated breast ultrasound (ABUS) scanners [6]. The first major step in the implementation of each of these applications is to contour a region of interest (ROI) inside the tumour in frames with identifiable tumour areas. This step is currently performed manually as there is no automated software to segment an ROI in ultrasound B-mode images. The manual segmentation of tumours, however, is a very time-consuming task as hundreds of frames should be contoured in a study due to the availability of multiple frames for each subject (an animal in preclinical and a patient in clinical studies). With the availability of 3-D scanners such as ABUS technologies, the problem will be even more severe as tens of frames should be contoured in each patient. Therefore, designing an automated segmentation method can save a significant amount of experts' time and efforts.

In this study, a semi-automated supervised tumour localization method was proposed for ROI estimation in B-mode images acquired from patients with locally advanced breast cancer (LABC).

II. BACKGROUND

There is an extensive literature on image segmentation. Several segmentation techniques have been tried for ultrasound images [7]. In a more recent literature, classification techniques applied on some expressive features have been more frequently employed to extract segments from images. These classification techniques include simple methods like *k*-means and fuzzy *c*-means (FCM) and more sophisticated algorithms such as support vector machines (SVM). These

M. J. Gangeh, H. Tadayyon, and G. J. Czarnota are with the Departments of Medical Biophysics, and Radiation Oncology, University of Toronto, Toronto, ON, M5G 2M9 Canada; also with the Departments of Radiation Oncology, and Imaging Research - Physical Sciences, Sunnybrook Health Sciences Centre, Toronto, ON, M4N 3M5 Canada (e-mails: mehrdad.gangeh@utoronto.ca, {hadi.tadayyon,gregory.czarnota}@sunnybrook.ca).

H. R. Tizhoosh, K. Wu, and D. Huang are with KIMIA Lab, University of Waterloo, Waterloo, ON, N2L 3G1 Canada. (e-mails: {tizhoosh,k8wu,dun.huang}@uwaterloo.ca).

classifiers group pixels into segments to enable subsequent measurements and recognition. Most of these methods, when applied to digital images, rely on the availability of invariant features. Among feature extraction methods in the literature, those which are based on invariant features extracted from some *keypoints* demonstrated success in segmentation tasks. Keypoints represent pixel positions in an image surrounded (within certain vicinity) with significant information, e.g., textures, corners, edges, etc [8]. For instance, FAST (features from accelerated segment test) [9], BRISK (binary robust invariant scalable keypoints) [10], and SURF (speeded up robust features) [11] are among the most commonly used keypoint-based feature descriptors providing inputs to a classifier.

SURF is a feature detection method, which is partly inspired by the scale-invariant feature transform (SIFT) descriptor, but is generally faster than SIFT [11]. It uses an integer approximation of the determinant of the Hessian blob detector to detect keypoints. Its feature descriptor is based on the sum of the Haar wavelet response around the point of interest. The three main modules of the algorithm include: interest point detection, local neighborhood descriptor, and matching.

In BRISK detector, the scale of each keypoint is estimated in the continuous scale-space [10]. The BRISK descriptors are composed of binary vectors by concatenating the results of brightness comparisons. The key concept of the BRISK descriptor is based on using a specific pattern to sample the neighborhood of all keypoints.

FAST is a corner detection method [9]. The advantage of using FAST corner detector is its high computational efficiency and a better performance when subsequent machine-learning methods such as a supervised classification are applied [12]. It uses a 16-pixel circle to classify a candidate point as a corner or no corner, where for each pixel in the circle, an integer number from 1 to 16 will be labeled clockwise.

III. PROPOSED APPROACH

The first step in tumour response prediction is contouring (segmenting) the tumours. Subsequently, texture analysis will be performed on the QUS parametric maps computed within the contoured tumours (tumour cores), and their margins (including surrounding healthy tissue) to predict patient's response [5]. In the aforementioned analysis steps, however, a rough localization of the tumour is of more interest than its exact segmentation [5]. Considering this, and also the fact that ultrasound images are known to be very difficult for segmentation, a semi-automated approach was proposed in this study to obtain an estimation of the tumour location as described in the next paragraphs.

Algorithm 1 describes the proposed approach.

Patient Data – The study involved 10 LABC patients with tumour sizes between 5 and 15 cm. The data acquisition was performed in accordance with the clinical research ethics approved by Sunnybrook Health Sciences Centre. A biopsy was used as the gold standard test to confirm all

---

Algorithm 1 Proposed Approach

1: —— Preprocessing and Feature Extraction ——
2: Read number of training images N
3: Calculate the average width $\bar{w}$ and height $\bar{h}$ of all segments
4: **while** $i < N$ **do**
5:      Read current image $\mathbf{I}_i$ and its ground truth image $\mathbf{G}_i$

6:      Pre-process $\mathbf{I}_i$ (contrast adjustment and filtering)
7:      Extract n keypoints with m descriptors $\mathbf{F}_i(n, m)$
8:      Acquire a point (x_c^i, y_c^i) close to the centre of $\mathbf{G}_i$
9:      **while** $j < n$ **do**
10:          Calculate the weighted distance d between (x_c^i, y_c^i) and the coordinates of keypoint n_j:
11:          $d_j = \sqrt{(x_j - x_c^i)^2 + [\frac{\bar{w}}{\bar{h}}(y_j - y_c^i)]^2}$
12:          Add distance to the descriptors $\mathbf{F}_i(n_j, m+1) = d_j$
13:      **end while**
14:      Save $\mathbf{F}_i(n, m'), (m' = m + 1)$
15: **end while**
16: —— Testing ——
17: **while** $i < N$ **do**
18:      Read the current image I_i
19:      Use remaining images to train a classifier with features $\mathbf{F}(n, m')$ and target E
20:      Classify pixels of I_i using the trained classifier
21:      Fit an ellipse E_i into the tumour pixels $(x_1^i, y_1^i), (x_2^i, y_2^i), \ldots$ provided by classifier
22:      Read $\mathbf{G}_i$
23:      Calculate the accuracy $D_i =$Dice$(\mathbf{G}_i, E_i)$
24: **end while**

---

cancer cases. In order to measure the size of tumours, all patients were imaged using MRI. Also, all patients were imaged using ultrasound before the start of neoadjuvant chemotherapy ("pre-treatment"). The acquisition of US data was performed using a Sonix RP ultrasound system. The system was equipped with an L14-5/60 linear transducer (centre frequency at $\sim$7 MHz). Depending on tumour size and location, the transducer was focused at the midline of the tumour with a maximum depth of 4-6 cm. The number of scans obtained from each tumour depended on it size. This resulted in three to five scan planes from each tumour, with a scan plane separation of about 1 cm.

Pre-Processing – In order to increase the quality and contrast of B-mode images, before computation of feature descriptors, each image was pre-processed by applying the fuzzy histogram hyperbolization [13]. Subsequently, a 3×3 median filter was applied to each image.

Feature Extraction – Features were extracted from the keypoints identified on the pre-processed images for submission to a classifier. Three major feature extraction algorithms explained in Section II, including SURF, FAST, and BRISK were adapted in this stage.

Extended Features Descriptors – After applying feature extraction methods, a feature descriptor of size m for each

keypoint was used to characterize an image. However, the preliminary results using these features revealed that there was no strong correlation between the tumour and the feature points to obtain an acceptable result (the initial performance using only extracted features resulted in relatively low performance, i.e., less than 60% accuracy for segmentation of tumours). Although the manual contouring of multiple scans is a very time-consuming and tedious task, requiring the clinician to provide the rough location of the tumour centre (e.g., via a mouse click) does not pose an unacceptable burden on him/her. Assuming that the tumour centre (x_c, y_c) is available, the distance d of each keypoint from (x_c, y_c) can be added to the descriptor to augment the feature space (Algorithm 1, lines 9-12). This, as our experiments demonstrated later, considerably increased the accuracy of contouring tumours.

Classification – A support vector machine (SVM) classifier with a radial basis function (RBF) kernel was used in a supervised paradigm to classify the keypoints in each image as "tumour" or "non-tumour" points based on the extended feature descriptors **F** extracted from the keypoints (Algorithm 1, line 14). In addition, k-means and FCM were also used for the classification of the keypoints in the same way for the purpose of comparison with the SVM performance.

Ellipsification – After classification, all keypoints would be classified as "tumour" or "non-tumour" pixels. Suppose that there are n points which are defined as tumour points in one image. For the purpose of ellipsification of a tumour area, at first, the positions of all the n points were found in the vertical direction and the leftmost position x_l and the rightmost position x_r were identified. In the horizontal direction, the highest position y_h and the lowest position y_l were determined. These four points are the four vertices of an ellipse, which was used as the boundary of the ROI representing the tumour.

Accuracy Measurement – Dice coefficient was measured for calculating the accuracy of classification results. The Dice coefficient is given by

$$D = \frac{2|E \cup G|}{|E| + |G|} \tag{1}$$

where E is the fitted ellipse and G is the ground-truth. It can be interpreted as a measure of overlapping between the estimated and correct classes [14].

Validation – Leave-one-out validation was performed to evaluate the performance of the designed ellipsification system on the 33 available B-mode images. The algorithm was developed using Matlab (R2011a, Mathworks, USA) on a 64-bit Intel Core i5-4200U CPU @ 2.30 GHz processor equipped with 8 GB of memory and Windows 8.1.

IV. RESULTS

Fig. 1 depicts the result of pre-processing on one typical breast ultrasound image using fuzzy histogram heprbolization. A comparison between the pre-processed (right) and the
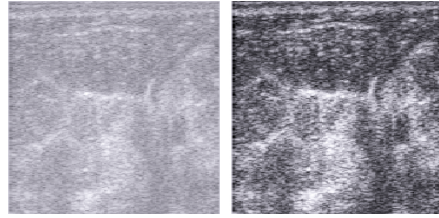

Fig. 1. A sample breast ultrasound image (left) and its enhanced pre-processed version after contrast adjustment using fuzzy histogram hyperbolization (right).
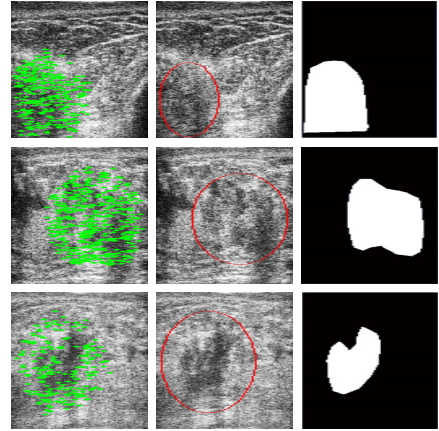

Fig. 2. Representative results for good, average, and bad ellipsification performances: 92.66% (*top*), 77.23% (*middle*), and 46.94% (*bottom*). Figure shows the computed keypoints (*left*), estimated ellipses (*middle*), and ground-truth contours (*right*) on representative B-mode images.

original (left) images shows the improvement in the quality of the B-mode image after pre-processing.

Fig. 2 shows representative good, average, and bad ellipsification results. As can be observed from this figure, the accuracy can significantly vary. Images with an irregular or small-size shape may not be classified correctly (accuracy as low as 46.9% in Fig. 2-bottom row). In contrast, for images with a normal size and shape, the accuracy can be quite high (e.g., 92.7% in Fig. 2-top row).

Table I provides the results of ellipsification using the Dice measure on all 33 B-mode images for various feature descriptors and classifiers used in the experiments. The highest average accuracy is 76.7% with a standard deviation of 15.2% using the SURF features and SVM classifier. In terms of computational cost, the SVM took longer time than other methods (approximately 10 minutes for each image, or 5 hours in total), while the other two classification methods were relatively fast. The results of classification can be improved by increasing the number of images as using only 33 scans in the experiments restricted the learning capability of the proposed ellipsification system.

Fig. 3 shows the distribution of the accuracy of ellipsification in respect to the number of generated keypoints for all the images used in the experiments (the results are only shown for the SVM classifier). BRISK, with an average accuracy of 75.63%, produced 979 ± 332 keypoints. FAST, with an average accuracy of 73.66%, resulted in

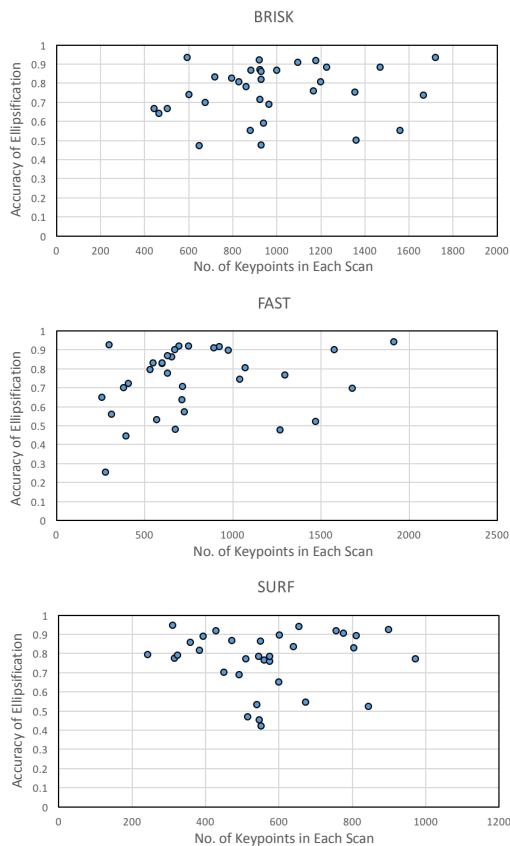| Features	Classifier	$D \pm \sigma$
BRISK	SVM	0.7563 ± 0.1387
FAST	SVM	0.7366 ± 0.1746
SURF	SVM	0.7665 ± 0.1515
BRISK	k-Means	0.7427 ± 0.1575
FAST	k-Means	0.7376 ± 0.1602
SURF	k-Means	0.7507 ± 0.1671
BRISK	FCM	0.7503 ± 0.1507
FAST	FCM	0.7379 ± 0.1601
SURF	FCM	0.7510 ± 0.1670

Fig. 3. The distribution of the accuracy of ellipsification versus the number of keypoints generated for all 33 B-mode images in the experiments. The slight advantage of SURF in our results (Table I) can be verified visually; compared to the BRISK and FAST, SURF appeared to result in higher accuracies more consistently, and within a narrower interval of features.

791 ± 420 keypoints. And SURF, with an average accuracy of 76.65%, generated 565 ± 179 keypoints. Therefore, the SURF achieved a higher performance using a smaller number of keypoints, which reduced the computational cost of subsequent classification step. Moreover, the SURF was more consistent in selecting the number of keypoints from B-mode images, as can be judged from the smaller standard deviation (179) achieved compared to the other two methods.

V. DISCUSSION AND CONCLUSION

In this study, a semi-automated localization method was proposed to mark a tumour with an ellipse based on a user single-click input, keypoint descriptors, and the SVM classification. The results demonstrated that the SURF features extended with the proposed weighted distances from the tumour's centre pixel (identified by the expert), along with the SVM classifier, achieved promising results. The original breast cancer ultrasound images exhibited very poor quality. Employing the central tumour point provided a higher correlation between the tumour area and the extracted features in noisy ultrasound images. Nonetheless, the segmentation of these images remains a challenging task as also reported by other researchers. For example, a recently published paper reported only between 50% and 60% accuracy in the segmentation of B-mode US images of breast [15].

In future work, a larger image dataset will be used for the training process. This will increase the chance of including examples of the images of different sizes and shapes, which will help to fully automate the approach.

REFERENCES

[1] F. S. Collins and H. Varmus, "A new initiative on precision medicine," *The New England Journal of Medicine*, vol. 372, no. 9, pp. 793–795, 2015.

[2] K. Brindle, "New approaches for imaging tumour responses to treatment," *Nature Reviews Cancer*, vol. 8, no. 2, pp. 94–107, 2008.

[3] M. L. Oelze and J. Mamou, "Review of quantitative ultrasound: Envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound," *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 2, pp. 336–351, Feb. 2016.

[4] M. J. Gangeh, H. Tadayyon, L. Sannachi, A. Sadeghi-Naini, W. T. Tran, and G. J. Czarnota, "Computer aided theragnosis using quantitative ultrasound spectroscopy and maximum mean discrepancy in locally advanced breast cancer," *IEEE Trans. on Medical Imaging*, vol. 35, no. 3, pp. 778–790, 2016.

[5] H. Tadayyon, L. Sannachi, M. J. Gangeh, A. Sadeghi-Naini, M. Trudeau, and G. J. Czarnota, "Early prediction of breast tumour response to chemotherapy using multiparametric quantitative ultrasound," in *40th International Symposium on Ultrasonic Imaging and Tissue Characterization (UITC)*, 2015.

[6] M. J. Gangeh, A. Raheem, H. Tadayyon, S. Liu, F. Hadizad, and G. J. Czarnota, "Breast tumour visualization using 3-D quantitative ultrasound methods," in *SPIE Medical Imaging*, 2016.

[7] S. Sridevi and M. Sundaresan, "Survey of image segmentation algorithms on ultrasound medical images," in *International Conference on Pattern Recognition, Informatics and Mobile Engineering*, 2013, pp. 215–220.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[9] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *10th IEEE International Conference on Computer Vision (ICCV)*, Oct. 2005, vol. 2, pp. 1508–1515.

[10] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *13th International Conference on Computer Vision (ICCV)*, Nov. 2011, pp. 2548–2555.

[11] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[12] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[13] H. R. Tizhoosh and M. Fochem, "Image enhancement with fuzzy histogram hyperbolization," *Proceedings of EUFIT*, pp. 1695–1698, 1995.

[14] V. Labatut and H. Cherifi, "Accuracy measures for the comparison of classifiers," in *arXiv preprint*, 2012, vol. arXiv:1207.3790.

[15] M. Elawady, I. Sadek, A. E. R. Shabayek, G. Pons, and S. Ganau, "Automatic nonlinear filtering and segmentation for breast ultrasound images," in *13th International Conference on Image Analysis and Recognition*, Berlin, Heidelberg, 2016, pp. 206–213, Springer-Verlag.