

# Database-Guided Breast Tumor Detection and Segmentation in 2D Ultrasound Images

Jingdan Zhang<sup>a</sup>, Shaohua Kevin Zhou<sup>a</sup>,  
Shelby Brunke<sup>b</sup>, Carol Lowery<sup>\*b</sup>, and Dorin Comaniciu<sup>a</sup>

<sup>a</sup>Robust Analysis and Content Retrieval Program, Siemens Corporate Research, Princeton, NJ, USA;

<sup>b</sup>Ultrasound Division, Siemens Medical Solution, Issaquah, WA, USA

## ABSTRACT

Ultrasonography is a valuable technique for diagnosing breast cancer. Computer-aided tumor detection and segmentation in ultrasound images can reduce labor cost and streamline clinic workflows. In this paper, we propose a fully automatic system to detect and segment breast tumors in 2D ultrasound images. Our system, based on database-guided techniques, learns the knowledge of breast tumor appearance exemplified by expert annotations. For tumor detection, we train a classifier to discriminate between tumors and their background. For tumor segmentation, we propose a discriminative graph cut approach, where both the data fidelity and compatibility functions are learned discriminatively. The performance of the proposed algorithms is demonstrated on a large set of 347 images, achieving a mean contour-to-contour error of 3.75 pixels with about 4.33 seconds.

**Keywords:** breast tumor detection and segmentation, database-guided methods, discriminative graph cut

## 1. INTRODUCTION

Ultrasonography is widely used in medical imaging due to its low cost, portability, and low risk to patients. In diagnosing breast cancers, it is a useful complementary technique to mammography.<sup>1</sup> Not only can ultrasonography be used as an early screening tool for breast cancer, it also shows unique value to distinguish benign tumors from malignant ones.<sup>2,3</sup> When diagnosing breast tumors based on ultrasound images, the shape information of tumors is important for physicians to make decisions.<sup>4,5</sup> As a result, tumor segmentation in ultrasound images is a routine clinical practice. Manually segmenting tumors is tedious and time-consuming. Computer-aided tumor detection and segmentation is valuable to reduce cost and streamline clinic workflows.

Automatic tumor detection and segmentation are challenging due to large shape and appearance variations of tumors, ultrasound speckle noises, signal dropouts, etc. Fig. 1 shows some tumors captured in typical ultrasound breast images. Several approaches have been proposed in the literature for tumor segmentation. In,<sup>6</sup> the Active Contour Model (ACM) is used to segment solid nodules in ultrasound images. In,<sup>7</sup> ACM is combined with the watershed transform to achieve better segmentation results. Other general segmentation techniques can also be applied to segment tumors. In,<sup>8</sup> the random walker algorithm<sup>9</sup> is used to segment tumors in CT images. In,<sup>10</sup> the graph cut algorithm is invoked to segment anatomic structures in medical images. A drawback of these algorithms is that they rely on empirical prior knowledge of objects' appearance, such as "boundary smoothness", "similar intensity or texture", and "high gradient boundary". These assumptions can be easily violated in ultrasound images, causing poor segmentation performance. Also these approaches require initializing the location of target objects manually.

Database-guided approaches<sup>11–15</sup> have been proposed to detect and segment objects automatically and more accurately by training discriminative models, typically classifiers, based on large medical databases. In detection, classifiers are trained to discriminate between the object of interest and background.<sup>11,12,14</sup> In segmentation, either boundary classifiers<sup>12,15</sup> or foreground/background classifiers<sup>13</sup> are plugged into the traditional segmentation algorithms to improve their performance.

In this paper, we propose a database-guided approach to **automatically** segmenting tumors in 2D ultrasound breast images. Segmentation is formulated as a two-step problem. The first step of our approach is tumor

---

\* This research was done when Carol Lowery was an employee at Siemens.

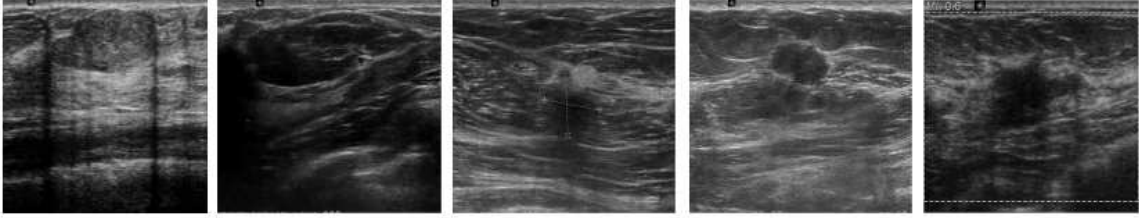


Figure 1. Tumors in 2D ultrasound breast images.

detection, where the location of a tumor in a given image is determined. The second step is tumor segmentation, where the tumor is delineated based on the detection result. Both steps use classifiers trained from a large tumor database. In segmentation, because of the complexity of tumor appearance and image noise, neither a boundary classifiers nor a foreground/background classifier alone can provide sufficient information to guide the segmentation algorithm. We propose a new segmentation algorithm called *discriminative graph cut* to learn a boundary classifiers *offline* and a foreground/background classifiers *online*. The information from these two classifiers is combined using a Markov Random Field (MRF), which can be solved efficiently by the graph cut algorithm. We demonstrate the effectiveness of the algorithm on a large data set of 347 images, which outperforms the baseline graph cut by a large margin. To the best of our knowledge, this is the first in the literature to attempt automatic segmentation of breast image in 2D ultrasonography using such a large data set.

## 2. TUMOR DETECTION

The goal of tumor detection is to locate tumors in images. Since the shape of a tumor is irregular, there is no common shape for tumors. For this reason, we represent a tumor location by its tight bounding box  $\lambda = [t_x, t_y, w, h, \theta] \in \mathbf{R}^5$ , where  $[t_x, t_y]$  is the position,  $[w, h]$  is the size, and  $\theta$  is the orientation of the tumor as specified by its long principal axis.

We use a learning based approach<sup>11</sup> to detect tumors. The basic idea is to formulate the detection as a two-class classification problem. In offline training, we first collect positive and negative image patches from the training images. Each image patch  $I_\lambda$  is determined by a bounding box  $\lambda$ . The positive image patches are determined by the bounding boxes of tumors, computed from expert annotations. The negative patches are randomly cropped from background. Fig. 2(a) shows some examples of image patches. Then we train a classifier to discriminate between the positive and negative patches. In testing, given a new image, tumors are detected by exhaustively checking all possible bounding boxes within a search range.

We use the Probabilistic Boosting Tree (PBT)<sup>16</sup> to train the classifier, which involves the recursive construction of a tree, where each of its node is an Adaboost classifier.<sup>17</sup> Other classifiers<sup>18</sup> can be used too. A PBT classifier computes the probability  $P(y = 1|I_\lambda)$  that the image patch  $I_\lambda$  contains a target object. We use this probability as a measurement of detection confidence. We will also use PBT classifiers in tumor segmentation, which will be discussed in the next section. In PBT computation, Haar-like features are used to extract image information. These features can be efficiently evaluated by using integral images.<sup>18</sup>

## 3. TUMOR SEGMENTATION USING DISCRIMINATIVE GRAPH CUT

### 3.1 Markov Random Field (MRF)

Segmentation can be viewed as a labeling problem, where each pixel is labeled as either foreground or background. Because there are spatial correlations among neighboring pixels in an image, MRFs are commonly used to describe these spatial correlations. Based on the MRF theory, we can segment an image by labeling its pixels to minimize an energy function:

$$E(\{l\}) = \sum_u D(l_u) + \alpha \sum_{(u,v)} V(l_u, l_v), \quad (1)$$

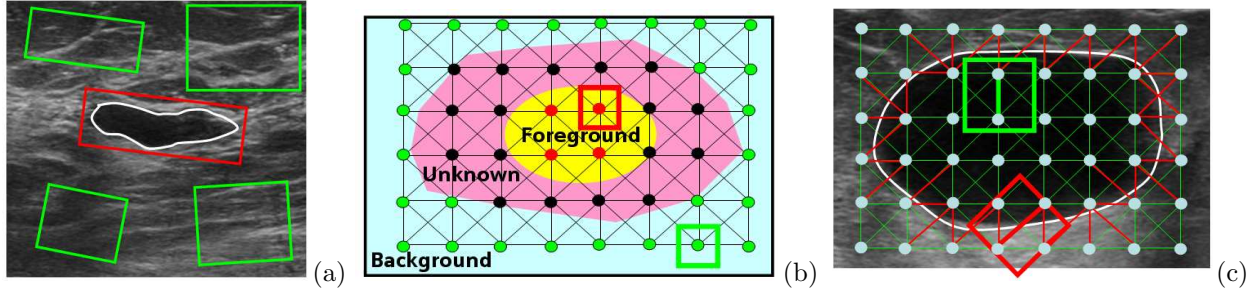


Figure 2. (a) The positive (red) and negative (green) image patches collected from an image for training the tumor detector. The white contour is ground truth. (b) Based on a given detection location, an image patch containing a tumor can be temporarily partitioned to three regions: foreground, background, and an unknown region. The red dots in foreground are positives and green in background are negatives in online learning. (c) An eight-connected grid used by MRF and illustration of sampling edges to train a classifier as the data compatibility function. The red edges are positive samples and the green edges are negatives.

where  $E$  is the energy,  $u$  is a pixel,  $l_u \in \{0, 1\}$  is the label at pixel  $u$  ( $l_u = 0$  means  $u$  is a background pixel and  $l_u = 1$  means  $u$  is foreground),  $D(l_u)$  is a *data fidelity function* measuring the cost of assigning the label  $l_u$  to the pixel,  $(u, v)$  is an edge in an eight-connected image grid shown in 2(b),  $V(l_u, l_v)$  is a *data compatibility function* describing the cost of labeling two neighboring pixels  $u$  and  $v$  with labels  $l_u$  and  $l_v$  respectively, and  $\alpha$  is a coefficient determining the relative weight between  $D(l_u)$  and  $V(l_u, l_v)$ . Energy function (1) can be minimized efficiently using the graph cut segmentation algorithm,<sup>10</sup> where the problem is formulated as a max-flow problem on a graph, and a global minimum solution is the min-cut between the source and the sink nodes in the graph.

The segmentation accuracy depends on the choices of data fidelity and compatibility functions. In the baseline graph cut segmentation,<sup>10</sup>  $D(l_u)$  is learned from data using generative models, and  $V(l_u, l_v)$  is defined to penalize the energy function when two pixels with similar intensities are labeled differently. These definitions have limits when applying them to segment tumors in ultrasound images. As seen in Fig. 1, the images are noisy and the foreground and background appearance of some tumors are quite similar. In such a condition, discriminative models usually outperforms generative models in terms of discriminating between the foreground and background region of a tumor. Also, compared with the edges caused by fat tissues, shadows, and noise, the boundaries of tumors in ultrasound images have distinctive appearance, which is an important cue for segmenting tumors. An empirically defined function  $V(l_u, l_v)$  has difficulty in encoding this data-specific information.

A variety of approaches have been proposed to learn  $D(l_u)$  and  $V(l_u, l_v)$  more effectively from the training images. In,<sup>13</sup>  $D(l_u)$  is learned using a discriminative model, but  $V(l_u, l_v)$  is still defined empirically. Conditional Random Fields (CRFs)<sup>19,20</sup> are introduced to learn both  $D(l_u)$  and  $V(l_u, l_v)$  from the training data using the maximum likelihood principle. However, CRFs have a computational cost exponential to the number of data points, due to the joint learning of  $D(l_u)$  and  $V(l_u, l_v)$ . To decrease the learning cost, Decoupled Conditional Random Fields (DCRFs)<sup>21</sup> are proposed to approximate CRFs by learning  $D(l_u)$  and  $V(l_u, l_v)$  separately. Experiment results in<sup>21</sup> show that DCRFs have comparable performance comparing with other CRF variants, but it is much faster to train. In this paper, we also learn  $D(l_u)$  and  $V(l_u, l_v)$  separately using discriminative models. The main difference between<sup>21</sup> and our approach are two fold: (i) We train  $D(l_u)$  and  $V(l_u, l_v)$  using boosting while SVM is used in.<sup>21</sup> Since boosting also is connected with feature selection,<sup>18</sup> during testing it is fast to evaluate the energy function. (ii) We learn the data fidelity function  $D(l_u)$  in an *online* fashion. This is possible because we perform automatic detection before segmentation.

### 3.2 Online learned data fidelity function

The data fidelity function  $D(l_u)$  is defined as:

$$\begin{cases} D(l_u = 1) = -\ln P(y = 1|I_u) \\ D(l_u = 0) = -\ln(1 - P(y = 1|I_u)) \end{cases} \quad (2)$$

where  $I_u$  is a small image patch around a pixel  $u$ , and  $P(y = 1|I_u)$  is the probability of  $u$  being a foreground tumor pixel.  $P(y = 1|I_u)$  can be learned offline from the training images like the approach in.<sup>13</sup> However,

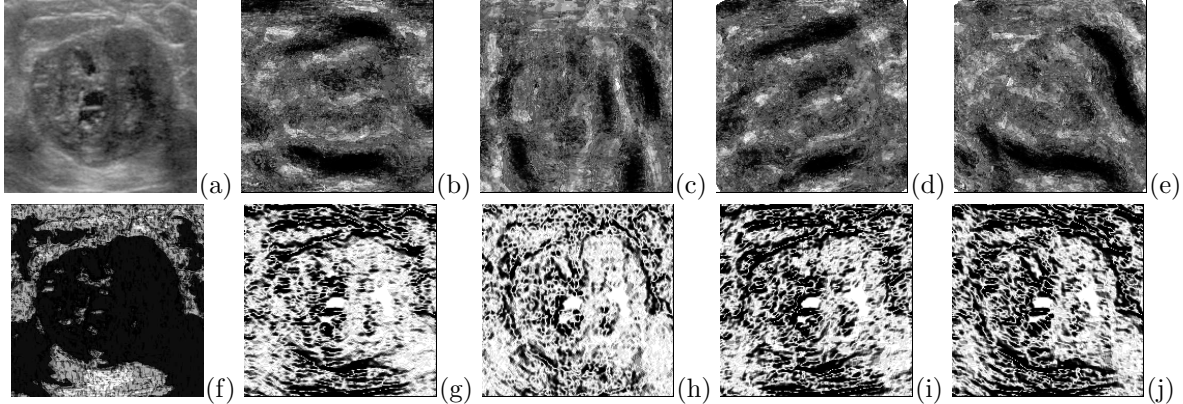


Figure 3. A tumor image and its associated cost functions: (a) original image; (b,c,d,e) the learned data compatibility function in horizontal, vertical, diagonal, and off-diagonal directions; (f) the online learned data fidelity function; and (g,h,i,j) the baseline data compatibility function<sup>10</sup> in horizontal, vertical, diagonal, and off-diagonal directions.

because of large foreground/background appearance variations of tumors in ultrasound images, as shown in Fig. 1, it is more effective to learn  $P(y = 1|I_u)$  online based on the detection results. After the rough location of a tumor is estimated, we can temporarily partition the image region around the tumor to three parts: foreground, background, and unknown. Fig. 2(b) shows the partition. The pixels in the foreground region have high confidence of being included in the tumor and those in the background region of being outside of the tumor. In training, we sample small image patches in the foreground region as positives and those in the background region as negatives. A PBT classifier is trained online using the sampled image patches. Fig. 3(a) and (f) show a tumor image and its associated  $D(l_u = 1)$ , which is learned online.

### 3.3 Offline learned data compatibility function

The data compatibility function is defined as:

$$\begin{cases} V(l_u \neq l_v) = -\ln P(y = 1|I_{uv}) \\ V(l_u = l_v) = -\ln(1 - P(y = 1|I_{uv})) \end{cases} \quad (3)$$

where  $I_{uv}$  is a small image patch around the edge  $[u, v]$ , and  $P(y = 1|I_{uv})$  is the probability that there is a tumor boundary passing between pixels  $u$  and  $v$ . We learn  $P(y = 1|I_{uv})$  offline from the training images. In training, we first determine a grid, on which the graph cut algorithm will be applied. For each training image with its annotation, edges on the grid are sampled randomly. Edges crossing tumor boundaries are positive samples and edges completely inside or outside tumors are negatives. An example of sampling is shown in Fig. 2(c). Again, a PBT classifier is trained to learn tumor boundary information. Fig. 3 compares the learned and the baseline data compatibility functions in horizontal, vertical, diagonal, and off-diagonal directions. The learned function highlights the tumor boundary more distinctively as evidenced by the dark (low energy) regions around it.

The sizes of some tumors are too small to fully contain the image patches for computing  $D(l_u)$  and  $V(l_u, l_v)$ . In order to solve this problem, we normalize the size of the image patch  $I_\lambda$  containing a tumor to a predefined size and conduct tumor segmentation in this normalized images. We learn  $D(l_u)$  and  $V(l_u, l_v)$  on a predefined coarse grid in training, and apply the discriminative graph cut algorithm on the grid with same resolution in testing. After the coarse segmentation, we refine the segmentation results by applying the baseline graph cut algorithm in the region around the coarse segmentation result.

## 4. EXPERIMENT RESULTS

In this section, the system performance is demonstrated using a large database. We collected 347 BMode ultrasound breast images from different patients. The tumors in these images are annotated by experts using contours defined by control points. Most images contain only one tumor. If there are multiple tumors in an

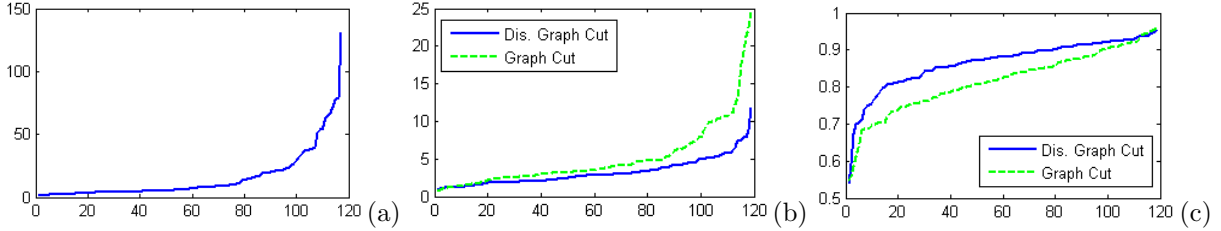


Figure 4. Sorted errors of the experiment results: (a) detection errors, (b) contour-to-contour distances, and (c) area overlap ratios.

(a) Detection	mean $E_{BB}$ (pix.)	90% mean $E_{BB}$ (pix.)	med. $E_{BB}$ (pix.)		time(sec.)
	$15.47 \pm 20.49$	$9.70 \pm 8.81$	6.62		3.53
(b) Semi-Auto Seg.	mean $E_{CC}$ (pix.)	med. $E_{CC}$ (pix.)	mean $E_{OR}$	med $E_{OR}$	time(sec.)
graph cut	$5.07 \pm 4.37$	3.57	$0.82 \pm 0.09$	0.83	0.37
discriminative graph cut	$3.21 \pm 1.82$	2.83	$0.86 \pm 0.07$	0.88	0.74
(c) Auto Seg.	mean $E_{CC}$ (pix.)	med. $E_{CC}$ (pix.)	mean $E_{OR}$	med $E_{OR}$	time(sec.)
graph cut	$4.84 \pm 3.36$	3.45	$0.81 \pm 0.09$	0.82	0.38
discriminative graph cut	$3.75 \pm 2.85$	2.86	$0.84 \pm 0.10$	0.87	0.80

Table 1. The benchmarks of the detection and segmentation algorithms. For each measurement, we reported its mean, standard deviation, and median.

image, the experts annotate the most obvious one. We used 228 images for training and the remaining 119 images for testing.

In testing, we first tested the performance of tumor detection. The average detection time is 3.53 seconds on a regular PC (3.2GHz CPU, 2GB RAM). For a given image, we only detected one tumor by selecting a location that has the highest aggregated probability. The first column in Fig. 5 shows some detection results. A result is represented as a bounding box containing the detected tumor. The detection error is defined as the average Euclidean distance between the corners of the detected box and the ground truth box. Table 1(a) shows the mean and median of the detection errors. The mean error is much larger than the median error because there are outliers in detection; hence we also computed the mean of 90% of test results by excluding 10% outliers. Fig. 4(a) show the sorted detection errors.

We then tested the performance of tumor segmentation. We compared two algorithms: the baseline graph cut algorithm<sup>10</sup> and our discriminative graph cut. The parameters of the baseline algorithm was tuned based on the training data to achieve good performance. We first tested these algorithms in a semi-automatic scenario, where an initial location of a tumor is determined by the ground truth annotation. We used two error measurements: average contour-to-contour distance  $E_{CC}$  and overlapping ratio  $E_{OR} = \frac{A_S \cap A_G}{A_S \cup A_G}$ , where  $A_S$  is the segmented area and  $A_G$  is the ground truth area. Table 1(b) shows the semi-automatic segmentation benchmarks. We computed both the mean and median of the two measurements. Fig. 4(b) and (c) are plots of the sorted errors, where the points on the curve with a same horizontal position do not necessarily correspond to a same test case. The discriminative graph cut outperforms the baseline graph cut, which is 58% above in term of the mean contour-to-contour distance. The discriminative graph cut uses slightly more time, due to extra cost in computing the cost functions.

Finally, we tested the full pipeline, in which the locations of tumors are determined by the detection module. Again the two segmentation algorithms are used for comparison. It is evident that using the automatically detected bounding box might introduce large segmentation error when the detector fails to find the rough location of a tumor. In order to compute a meaningful segmentation benchmark, we excluded 10% detection outliers. Table 1(c) shows the benchmarks. The discriminative graph cut still has a better performance: the baseline graph cut yields the mean contour-to-contour distance 29% more than the discriminative graph cut. The second and third column in Fig. 5 show some fully automatic segmentation results.

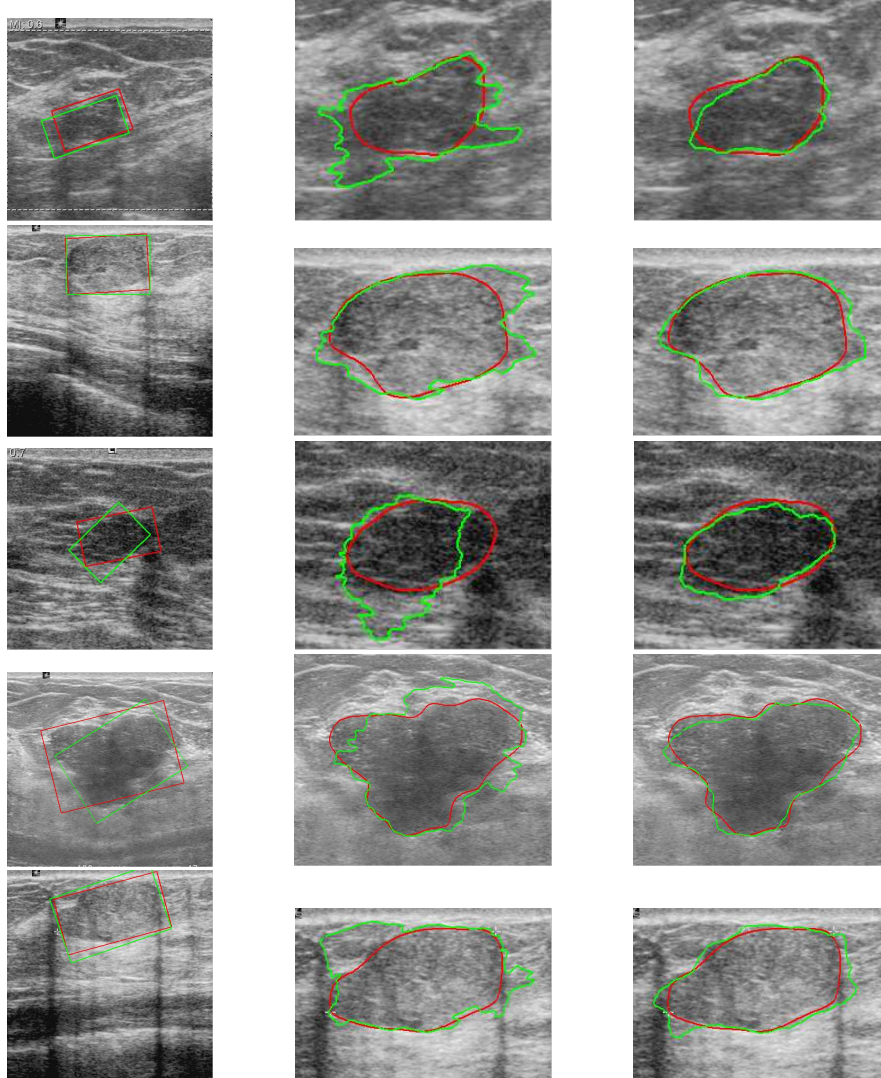


Figure 5. Several fully automatic segmentation examples on which our algorithm works better than the the baseline graph cut. The first column is the detected result. The second is the result of the baseline graph cut. The third is the result of our algorithm. The ground truth is plotted as red and the detection and segmentation as green.

## 5. CONCLUSION

In this paper, we have proposed a fully automatic system to detect and segment breast tumors in 2D ultrasonography. For tumor detection, we have applied a classification approach to discriminate between tumors and their background. For tumor segmentation, we have proposed a *discriminative graph cut* algorithm, where the data fidelity function is online learned and the data compatibility function is offline learned, both discriminatively. We have demonstrated the performance of the proposed algorithms on a large image database of breast tumors. We have shown the proposed learning based segmentation algorithm outperforms the baseline graph cut by a large margin.

## REFERENCES

- [1] Bosch, A., Kessels, A., Beets, G., Vrancken, K., Borstlap, A., von Meyenfildt, M., and van Engelshoven, J., “Interexamination variation of whole breast ultrasound,” *British Journal of Radiology* **76**, 328–331 (2003).

- [2] Shankar, P., Reid, J., Ortega, H., Piccoli, C., and Goldberg, B., "Use of non-rayleigh statistics for identification of tumors in ultrasonic B-scans of the breast," *IEEE Trans. Medical Imaging* **12(4)**, 687–692 (1993).
- [3] Stavros, A., Thickman, D., Rapp, C., Dennis, M., Parker, S., and Sisney, G., "Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions," *Radiology* **196**, 122–134 (1995).
- [4] Garra, B., Krasner, B., Horii, S., Ascher, S., Mun, S., and Zeman, R., "Improving the distinction between benign and malignant breast lesions: The value of sonographic texture analysis," *J. Ultrasound Imaging* **15**, 267–285 (1993).
- [5] Chen, C., Chou, Y., Han, K., Hung, G., Tiu, C., Chiou, H., and Chiou, S., "Breast lesions on sonograms: Computer-aided diagnosis with nearly setting-independent features and artificial neural networks," *Radiology* **226**, 504–514 (2003).
- [6] Aleman, M., Aleman, P., Alvarez, L., Esteban, M., Fuentes, R., and Montesdeoca, J., "Computerized ultrasound characterization of breast tumors," *Computer Assisted Radiology and Surgery (Elsevier International Congress)* **1281**, 1063–1068 (2005).
- [7] Huang, Y. and Chen, D., "Automatic contouring for breast tumors in 2-D sonography," in [*Proc. IEEE Engineering in Medicine and Biology 27th Annual Conference*], (2005).
- [8] Jolly, M. and Grady, L., "3D general segmentation in CT," in [*Proc. IEEE Int'l Sym. Biomedical Imaging*], (2008).
- [9] Grady, L., "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.* **28(11)**, 1768–1783 (2006).
- [10] Boykov, Y. and Jolly, M., "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," in [*Proc. Int'l Conf. Computer Vision*], (2001).
- [11] Georgescu, B., Zhou, X. S., Comaniciu, D., and Gupta, A., "Database-guided segmentation of anatomical structures with complex appearance," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], (2005).
- [12] Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., and Comaniciu, D., "Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features," in [*Proc. Int'l Conf. Computer Vision*], (2007).
- [13] Wels, M., Carneiro, G., Aplas, A., Huber, M., Hornegger, J., and Comaniciu, D., "A discriminative model-constrained graph cuts approach to fully automated pediatric brain tumor segmentation in 3D MRI," in [*Proc. Int'l Conf. Medical Image Computing and Computer Assisted Intervention*], (2008).
- [14] Zhang, J., Zhou, S., McMillan, L., and Comaniciu, D., "Joint real-time object detection and pose estimation using probabilistic boosting network," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], (2007).
- [15] Zhang, J., Zhou, S., Comaniciu, D., and McMillan, L., "Discriminative learning for deformable shape segmentation: A comparative study," in [*Proc. European Conf. Computer Vision*], (2008).
- [16] Tu, Z., "Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering," in [*Proc. Int'l Conf. Computer Vision*], (2005).
- [17] Freund, Y. and Schapire, R., "A decision-theoretic generalization of online learning and an application to boosting," *J. Computer and System Sciences* **55(1)**, 119–139 (1997).
- [18] Viola, P. and Jones, M., "Rapid object detection using a boosted cascade of simple features," in [*Proc. IEEE Conf. Computer Vision and Pattern Recognition*], (2001).
- [19] Lafferty, J., McCallum, A., and Pereira, F., "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in [*Proc. International Conference on Machine Learning*], (2001).
- [20] Kumar, S. and Hebert, M., "Discriminative fields for modeling spatial dependencies in natural images," in [*Advances in Neural Information Processing Systems*], (2003).
- [21] Lee, C., Greiner, R., and Zaïane, O., "Efficient spatial classification using decoupled conditional random fields," in [*Proc. Knowledge Discovery in Databases*], (2006).