






Anatomical Structure Segmentation in Ultrasound Volumes Using Cross Frame Belief Propagating Iterative Random Walks

Debarghya China , Graduate Student Member, IEEE, Alfredo Illanes , Prabal Poudel ,
Michael Friebe , Senior Member, IEEE, Pabitra Mitra, Senior Member, IEEE,
and Debdoot Sheet , Member, IEEE

Abstract—Ultrasound (US) is widely used as a low-cost alternative to computed tomography or magnetic resonance and primarily for preliminary imaging. Since speckle intensity in US images is inherently stochastic, readers are often challenged in their ability to identify the pathological regions in a volume of a large number of images. This paper introduces a generalized approach for volumetric segmentation of structures in US images and volumes. We employ an iterative random walks (IRW) solver, a random forest learning model, and a gradient vector flow (GVF) based interframe belief propagation technique for achieving cross-frame volumetric segmentation. At the start, a weak estimate of the tissue structure is obtained using estimates of parameters of a statistical mechanics model of US tissue interaction. Ensemble learning of these parameters further using a random forest is used to initialize the segmentation pipeline. IRW is used for correcting the contour in various steps of the algorithm. Subsequently, a GVF-based interframe belief propagation is applied to adjacent frames based on the initialization of contour using information in the current frame to segment the complete volume by frame-wise processing. We have experimentally evaluated our approach using two different datasets. Intravascular ultrasound (IVUS) segmentation was evaluated using 10 pullbacks acquired at 20 MHz and thyroid US segmentation is evaluated on 16 volumes acquired at 11–16 MHz. Our approach obtains a Jaccard score of 0.937 ± 0.022 for IVUS segmentation and 0.908 ± 0.028 for thyroid segmentation while processing each frame in 1.15 ± 0.05 s for the IVUS and in 1.23 ± 0.27 s for thyroid segmentation without the need of any computing accelerators such as GPUs.

Index Terms—Belief propagation, volumetric analysis, thyroid segmentation, iterative random walks, random forests, intravascular ultrasound segmentation.

I. INTRODUCTION

ULTRASOUND imaging has been employed for clinical use for half a century, and its salient advantages including non-ionizing nature of radiation, real time imaging capability are some of the factors which make it popular over other modalities like X-rays, computed tomography (CT) or magnetic resonance (MRI). Some of the commonly practiced clinical applications includes intravascular ultrasound (IVUS), thyroid, abdominal, cardiac, breast imaging, bone sonometry, prostate, fetal, ophthalmic imaging [1]. Current advances in signal acquisition and tracking systems, enable volumetric ultrasound acquisition using approaches such as 2D+t, 3D, and 3D+t [2]. The primary limitation in 2D ultrasound due to its planar appearance of structures, may lead to unavoidable occlusion of important 3D attributes characteristic of pathologies leading to inter-observer variability in reporting. Volumetric ultrasound can directly be acquired using a 3D imaging probe or by using a tracking device integrated with a 2D probe.

Ultrasound imaging primarily relies on sensing of backscattered echoes to image an object submerged within another medium that conceals it from direct external visibility [3]. A pulse of acoustic wave at a fixed frequency typically more than human audible range, is transmitted from the transducer and travels through the tissue. During this process, its energy is partially absorbed, attenuated and backscattered [4]. The backscattered signal moves back to the transducer and contributes to the formation of the US signal. Backscattering is contributed by changes of acoustic index while traversing across different tissues. The received pulse is envelope detected and a log-compression for formation of ultrasound signal in B-mode per scan-line that is subsequently stacked to form an image and displayed on a monitor as a gray-scale image.

Ultrasound image segmentation is complicated and a challenging task due to the presence of certain characteristic artifacts viz. speckles, shadows, and signal dropout as well as due to the orientation based variability that affects signal acquisition [1]. In this paper, primarily address the task of segmenting 3D US volume using frame-wise 2D segmentation approach following an iterative convergence approach per-frame and an transfer of belief on contour location for solving volumetric segmentation.

Manuscript received February 17, 2018; revised July 4, 2018; accepted July 28, 2018. Date of publication August 15, 2018; date of current version May 6, 2019. (Corresponding author: Debarghya China.)

D. China, P. Mitra, and D. Sheet are with the Indian Institute of Technology Kharagpur, Kharagpur, 721302, India (e-mail: debarghyachina@iitkgp.ac.in; pabitra@cse.iitkgp.ac.in; debdoot@ee.iitkgp.ac.in).

A. Illanes, P. Poudel, and M. Friebe are with the Otto-von-Guericke University Magdeburg, Magdeburg 39106, Germany (e-mail: alfredo.illanes@ovgu.de; prabal.poudel@ovgu.de; michael.friebe@ovgu.de).

Digital Object Identifier 10.1109/JBHI.2018.2864896

The brief prior art for both 2D and 3D segmentation in general and specifically for IVUS and thyroid are presented in Sec. II. Sec. III presents the problem statement and details the algorithm. The experiment description and results for both applications are stated in Sec. IV. The discussion of various characteristics of our approach as validated experimentally in Sec. V and finally we conclude our approach in Sec. VI.

II. PRIOR ART

In this section we present a comparison of prior art related to US image and volumetric segmentation in general and also specifically look into thyroid and IVUS segmentation problem.

1) *Segmentation on 2D Ultrasound Images*: Approaches include genetic algorithm based optimization of active contour, maximum likelihood based region segmentation, adaptive region growing based segmentation, texture-based split-and-merge techniques, knowledge-based approach, watershed, and learning based approaches [1].

a) *IVUS*: Prior art predominantly make use of active surface segmentation, Hopfield neural networks, fast-marching method, implicit mode of active contour and with anisotropic contour closing or explicit form of active contour model using snakes, edge tracking and gradient-based techniques, shape-driven method and optimal graph search that have been summarized by Noble *et al.* [1]. Papadogiorgaki *et al.* [5] proposed a method for intensity based segmentation of lumen and media-adventitia boundaries which use radial basis function for refining initially recognized contours. Ciompi *et al.* [6] introduced a multi-class classifier using pixel-wise and contextual features to segment lumen from media. The final segmentation is achieved using the coarse definition of media and adventitia tissues. Balocco *et al.* [7] had proposed a standardized framework for evaluation of IVUS segmentation at the 2011 MICCAI workshop, where a total of eight algorithms including four automatic and four semi-automatic ones were evaluated. In an earlier work [8] we had proposed a graph theoretic random walk method for lumen wall segmentation. Subsequently also on ultrasound backscattering physics based learning model for segmentation of lumen and external elastic laminae were proposed in [9], [10].

The limitations in prior art is their primary dependence on shape of the edges and the region information of contours which is weakly discriminable in speckle images. Also most algorithms are device frequency specific and not generally semi-automated or fully automated whether cumulatively they are challenged with inability to handle guide-wire shadows and pitch shaft discontinuity arising due to helical scan.

b) *Thyroid*: Keramidas *et al.* [11] had designed a neither for thyroid gland segmentation incorporating boundary detection along with local binary patterns based textures analysis. Maroulis *et al.* [12], [13] developed algorithms based on level-sets and active contours incorporating variable background modeling, active contours and joint echogenicity-texture. These models primarily reduce the effect of intensity inhomogeneity. Feedforward neural network [14] have also been proposed for segmentation of thyroid region. Wunderling *et al.* had proposed approaches based on level set, graph cut, and feature

classification for thyroid segmentation [15], and Narayan *et al.* [2] have introduced a multiorgan segmentation for thyroid gland, carotid artery, muscles and the trachea employing spatial location based search techniques.

The segmentation algorithms predominantly rely on using pattern of the region of interest and its structure are challenged when quality of the image degrades. Moreover they fail to work when small sized objects appear in frames.

2) *Segmentation of Structures in 3D Ultrasound Volumes*: Segmentation using 3D active surfaces operating on region-based external forces, neural network based approaches for detection of boundary and elastic surface model fitting are some of the popularly used approaches [1]. One common denominator is the predominance in use of speckle reduction filters like anisotropic diffusion to enable better localization of the boundary around high-intensity gradient. Active models viz. deformable active surface, intensity gradient based active surfaces and structural models have been used to segment lesions in breast tissue [1].

a) *IVUS*: A semi-automated technique using discrete dynamic contour model and active surface segmentation has been proposed in [16] to segment lumen and adventitial border across a stack of IVUS frames in a pullback. In another approach [17], a fast active surface method using neighborhood-search technique was used. Cardinal *et al.* [18] had designed a 3D fast-marching method based on the gray level distribution modeled using mixture of Rayleigh on the whole IVUS pullback. A graph-based approach had been proposed for lumen and external elastic laminae segmentation [19] using gated IVUS image sequences. Mendizabal-Ruiz *et al.* [20] proposed an approach based on minimization of a probabilistic cost function where the likelihoods were acquired from a support vector machine. Zakeri *et al.* has proposed a four-fold method [21], where an initial contour drawn from the classification in the sparse representation framework is used to initialize on active surface model.

These methods are highly specific to the application in hand and cannot be extend beyond IVUS for generalizable. Furthermore most of them are also US frequency specific and required modifications to model parameter for extension to different frequencies.

b) *Thyroid*: A semi-automated approach with the help of active geodesic contour had been proposed in [22] to classify and analyze the thyroid. This framework had been used for volumetric quantification, and the algorithm had been extended to segmenting other parenchymatous organ as well. Chang *et al.* [23] had proposed a full solution to estimate the thyroid volume, as well as classify the blocks within thyroid.

The primary challenge posted by these methods are image size and size of the organ as visible on the image such that smaller sized contours predict high amount of false areas and process is also susceptible to the imaging frequency.

The existing limitations of these US segmentation methods in respect of being entirely application oriented, and highly dependent on the shape of the region of interest while being fully focused on a particular frequency of which imaging on they are designed. Furthermore they do not exploit the reduction in compute complexity which comes with using the inter-frame

dependency arises in a 3D US volume acquired. The motivation of this paper is to address these limitations using a fully automated, robust and low time complexity algorithm. Our contributions has two distinct characteristics. First, we employ the apriori information based on physics of the US imaging process using its statistical mechanics to obtain a rough estimate of the tissues. Second, we exploit the redundancy in compute arising due to the inter-frame dependency across adjacent 2D frames constituting a 3D volume which we also term as belief propagation.

III. MATHEMATICAL MODEL FOR CONTINUOUS ULTRASOUND FRAME SEGMENTATION

A basic observation on an ultrasonic acoustic pulse passing through tissues is that one part of the signal energy is backscattered, while another part is attenuated, and the remaining energy of the signal is absorbed by the tissues traverses through it. Imaging relies on the use of these returned echo signals. The leading cause of backscattering is the presence of scatterers in the tissue. Their nature varies with the tissue type and that leads to variation in statistical behavior of the envelope of the detected ultrasonic echo signal. The scatterers' contribution to an echo can be treated as a random walk due to its presence at arbitrary locations within the resolution limits of the backscattered ultrasonic echo pulse. The primary challenge associated with segmenting structures in ultrasonic images is the stochastic nature of speckles in images. Here we handle this challenges by relying on the statistical nature of envelope of the received ultrasonic echo. Here we employ (i) a machine learning based model to weakly predict the type of the tissue based on parametric models of ultrasound echo in order to provide an initial seed point to the iterative random walks solver, and subsequently (ii) calculate the gradient vector flow between neighboring frames for belief-propagation based segmentation solution across neighboring frames for volumetric analysis, and (iii) an iterative random walks solver for computing the segmentation across the serial frames in an ultrasound volume. The mathematical problem statement is defined as below.

Let us consider an US frame \mathcal{I} , where $i(\mathbf{x})$ is the intensity at location \mathbf{x} and k is the number of unique contours split that \mathcal{I} into $k + 1$ disjoint set as $\mathcal{I}_{L_1}, \mathcal{I}_{L_2}, \dots, \mathcal{I}_{L_{k+1}}$ such that $\mathcal{I}_{L_{k_1}} \cap \mathcal{I}_{L_{k_2}} = \emptyset$ and $\forall L_1 \neq L_2, (k_1, k_2) \in \{1, 2, \dots, k + 1\}^2$ and $\mathcal{I}_{L_1} \cup \mathcal{I}_{L_2} \cup \dots \cup \mathcal{I}_{L_{k+1}} = \mathcal{I}$. The image \mathcal{I} can further also be represented as a graph \mathcal{G} where the nodes of the graph can be represented as $n \in \mathcal{I}$ and the weights of edges connecting the nodes of graph \mathcal{G} are modeled based on the physics of the acoustic energy propagation and attenuation within highly scattering biological tissues. The probability of each node $n \in \mathcal{G}$ to belong to the regions in $\{\mathcal{I}_{L_1}, \mathcal{I}_{L_2}, \dots, \mathcal{I}_{L_{k+1}}\}$ can be obtained using the random walks approach for the image segmentation [24]. The class posterior probability at $\mathbf{x} \in \mathcal{I}$ is the probability of the correspond node $n \in \mathcal{G}$. A pixel at location \mathbf{x} is labeled as $y = \arg \max\{p(L_1|\mathbf{x}, \mathcal{I}), p(L_2|\mathbf{x}, \mathcal{I}), \dots, p(L_{k+1}|\mathbf{x}, \mathcal{I})\}$.

A set of seeds \mathcal{S}_M constitutes some of the marked nodes of graph \mathcal{G} such that $\mathcal{S}_M \subseteq \{(\mathcal{S}_M \in \mathcal{I}_{L_1}) \cup (\mathcal{S}_M \in \mathcal{I}_{L_2}) \cup$

$\dots \cup (\mathcal{S}_M \in \mathcal{I}_{L_{k+1}})\}$ and $(\mathcal{S}_M \in \mathcal{I}_{L_1}) \cap (\mathcal{S}_M \in \mathcal{I}_{L_2}) \cap \dots \cap (\mathcal{S}_M \in \mathcal{I}_{L_{k+1}}) = \emptyset$ is defined for initialization of the random walker. An ultrasonic pulse backscattering physics based model is used to provide these initial seeds. Thus on solution, a class specific posterior probability is assigned to the unmarked nodes $\mathcal{S}_U = \mathcal{G} - \mathcal{S}_M$ of the graph to obtain the segmented anatomical structure such that $\mathcal{G} \subseteq \{\mathcal{S}_M \cup \mathcal{S}_U\}$ and $\mathcal{S}_M \cap \mathcal{S}_U = \emptyset$. The detail information flow of our proposed algorithm is presented in Fig. 1. Subsequently, we discuss the various stages of the algorithm and their significance to the solution.

A. Statistical Mechanics of Ultrasound Backscattering

Let us consider the value of the signal received at a specific location as r and let the tissue type be ω , then the conditional likelihood $p(r|\omega)$ can be written as a function of the following tissue specific effects on ultrasonic signals:

- a) likelihood of the received ultrasonic signal: $f_1(r; \phi_1|\omega)$;
- b) reliability of the received ultrasonic signal: $f_2(r; \phi_2|\omega)$

where ϕ_1 and ϕ_2 are the set of parameters defining $f_1(\cdot)$ and $f_2(\cdot)$ respectively which are dependent on the tissue specific properties. The likelihood $p(r; \dots|\omega)$ can be parametrically represented taking into consideration these factors such that

$$p(r; \dots|\omega) \propto \underbrace{\{f_1(r; \phi_1|\omega)\}}_{\text{backscattering stats.}}, \underbrace{\{f_2(r; \phi_2|\omega)\}}_{\text{confidence}} \quad (1)$$

The received US signal r is Nakagami distributed [25] such that $p(r|\omega) \propto \mathcal{N}(r|m, \Omega)$ [4], [10]. Since in a B-mode image, the image intensity i is a log-compressed version of the signal r , the intensity $i \in I$ is accordingly Fisher-Tippett distributed [4], [10] such that $p(i|\omega) \propto F(i|\sigma)$ and

$$F(i|\sigma) \propto \exp([2i - \ln(2\sigma^2)] - \exp[2i - \ln(2\sigma^2)]) \quad (2)$$

where σ is the standard deviation of the intensity.

The parameters of i and σ are estimated through a nonlinear multiscale estimation. According to our proposition [26], these parameters are estimated at different scales $\tau = (\tau_{trans}, \tau_{axial})$ where τ_{trans} represents the span along the the number of neighboring scan lines and τ_{axial} is the number of samples along each scan line.

The factor of reliability of an ultrasonic signal measured in terms of confidence in $f_2(\cdot)$ is estimated using the method of random walks [10], [27]. The confidence of the ultrasonic signal is estimated as the probability of a random walker starting at a node of the graph equivalent of the US image located on the scan-line to reach the origin of each scan line where the virtual transducer element is placed akin to the physical location of the sensor element on the transducer. Thus the signal confidence is represented as

$$p(r; \dots|\omega) \propto f_2(r; \phi_2|\omega) \quad (3)$$

where $\phi_2(\cdot)$ is the ultrasonic signal confidence associated with the received echo r that was backscattered by a tissue type ω . This set of ultrasonic backscattering modeling statistical features enable on to learn a Random forest classifier [28] to classify different tissue layer from the US images.

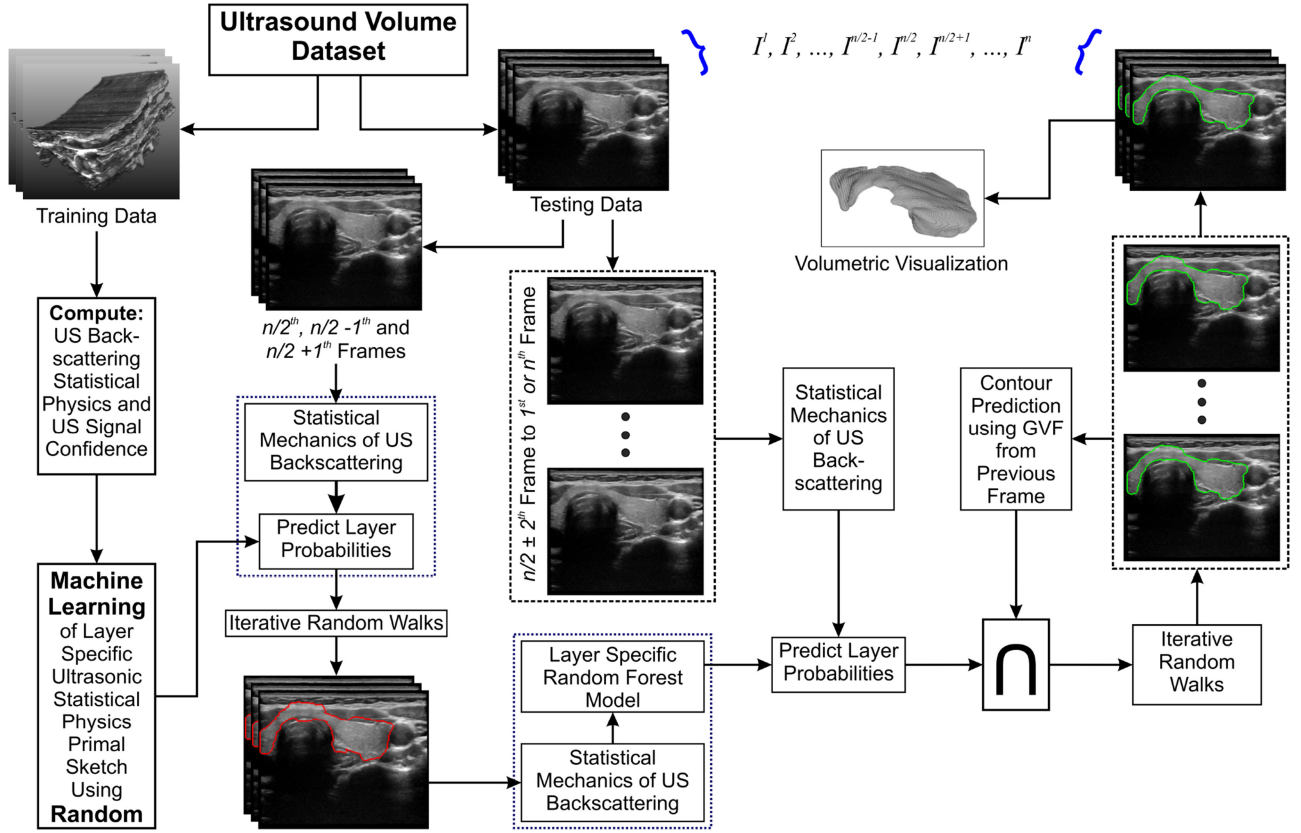


Fig. 1. Segmentation framework of the proposed approach for ultrasound (US) volume. The US dataset has m number of volumes (V_1, V_2, \dots, V_m) where $m - 1$ number of volume has been used for Random Forest (RF) training of initial segmentation and one volume is used for testing where the testing volume is consecutive of n number of individual frames i.e. $I^1, I^2, \dots, I^{n/2-1}, I^{n/2}, I^{n/2+1}, \dots, I^n$. Iterative random walks (IRW) is used for final segmentation in each different stages. Gradient vector flow (GVF) for each frame has been calculated from previous two frames. Finally the segmented volume of the US is visualized in volumetric visualization block.

B. Transfer Learning of Ultrasound Backscattering Statistics

A non-parametric machine learning framework using random forest [4] has been employed for the purpose of learning apriori information on speckles to seed. The parameters of $f_1(r; \dots | \omega)$ and $f_2(r; \dots | \omega)$ obtained earlier (Sec. III-A) constitutes the feature space to be modeled for prediction of the tissue-specific posterior probability. The prediction model of a random forest can be represented as

$$p(\omega | \Theta; r) = H(\omega | \Theta; r) \quad (4)$$

where $H(\omega | \Theta; r)$ is the learn RF model that is formally defined as a classifier consisting of a collection of tree-structured decision makers $\{h(\omega | \Theta, \phi_z), z = 1, \dots\}$, where $\{\phi_z\}$ are independent and identically distributed random vectors representing sample features, $\{\phi_z\} \subseteq \Theta$ and each tree $h(\omega | \Theta, \phi_z)$ casts a unit vote for the most popular class ω at input Θ [28]. During prediction, the vote casted by the forest is the class specific mean response of each of the trees such that $p(\omega | \Theta; i) = E[h(\dots, \phi_z)]$. The initial segmentation done using this random forest model is used to initialize seeds for each tissue type $\omega \in \{L_1, L_2, \dots, L_{k+1}\}$.

C. Belief Propagation Across Neighboring Frames Using Gradient Vector Flow

Let us consider that a US volume consists of n number of frames organized as an ordered set $\{I^1, I^2, \dots, I^{n/2-1}, I^{n/2}, I^{n/2+1}, \dots, I^n\}$. The process starts with segmenting the $\{I^{n/2-1}, I^{n/2}, I^{n/2+1}\}$ frames in the volume using the method in Sec. III-B. The segments are used to initialize the seeds for the random walks based fine segmentation detailed subsequently in Sec. III-D. Those three segmented frames are then used to further train a volume specific design random forest model presented in Sec. III-B. This volume specific learnt model helps to precisely predict the posterior probability of the contours and the background tissue in the current volume taking into account and learning patient specificity of ultrasonic tissue interaction with the contour in $\{I^{n/2}, I^{n/2+1}\}$, GVF is employed to estimate the contour in $I^{n/2+2}$ and this serves to initialize seeds for segmentation of the $I^{n/2+2}$ frame using IRW presented in Sec. III-D. This process repeats the other way to predict contour on $I^{n/2-2}$ using GVF over $\{I^{n/2-1}, I^{n/2}\}$. The final segmentation is done using intersection of the segmentation results obtained from the learnt model and the GVF for every frame process continues till it reaches the end frames on the volume. Since seed initialization with the RF on individual frames is not required, computation time is considerably reduced, and the method serves to

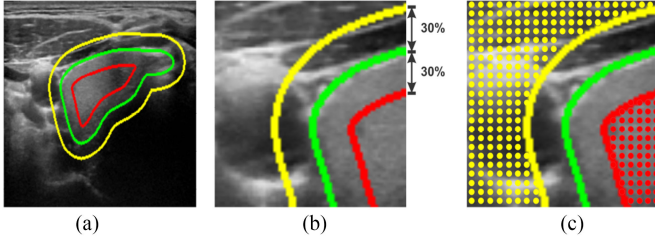


Fig. 2. Illustration of the seed selection from foreground and background for solving IRW. (a) Initial contour (green) with the seeds for background region (yellow) and foreground region (red). (b) Foreground (red) and background (yellow) region has been selected by a morphological regularization operation on the initial estimated region (green). (c) Neighborhood around the foreground (red dots) and background (yellow dots) that are selected as seeds for different regions.

accurately segment accounting for tissue specific variations on an auto-updating mechanism.

D. Solution to Iterative Random Walks for Final Segmentation

A predominant observation with RW is its mild variation in segmentation due to updates on seeds, and we follow this to a convergence using IRW. We iteratively employ random walk solver [24]. Here, we denote the probability of a random walker starting at a node v_q to reach a seeded point belonging to tissue type $\omega \in \{L_1, L_2, \dots, L_{k+1}\}$ as x_q^ω such that $\sum_\omega x_q^\omega = 1$.

Therefore, for label ω , the solution probabilities of a random walker could be generated at a node $q \in \mathcal{G}$ where the \mathcal{G} is the graphical representation of the image I and has been solved accordingly [24].

$$p(L_Z|x, I) = x_q^Z \forall \{q \in \mathcal{G} \Leftrightarrow x \in I\}, Z = 1, 2, \dots, k+1 \quad (5)$$

The initialization of the contour for random walks is obtained from the intersection between the prediction of the RF and GVF. The foreground and background seeds in the image were designed using morphological operation on the initial contour. First erosion operation is performed and the resultant region is used for foreground seeding while the inverse of that dilated region is used for background seeding. The structuring element in these morphological operations is chosen depending on the size of the initial contour and typically sized at 30% of the radius of the initial estimated region of the structure. This morphology based regularization is illustrated in Fig. 2. In case of iterative random walks (IRW), the solution of the current step acts as the initial contour for the subsequent step until convergence is reached. This iterative process converges when the change in the posterior probability obtained in the current stage is smaller than a pre-defined tolerance (δ). The stepwise method for Iterative Random Walks is shown in Algorithm 1.

IV. EXPERIMENTS AND RESULTS

The segmentation results at different stages of the algorithm is illustrated in Fig. 1 for an example of thyroid segmentation. The first step in our work is to segment the three middle frames of the volume. We have employed RF and IRW to solve this

TABLE I
PERFORMANCE MEASURE WITH THE IVUS AND THYROID DATASET IN THE 3D CONTEXT

Dataset	SE	SP	DSC	PAD
Phantom	0.992 ± 0.003	0.956 ± 0.031	0.950 ± 0.040	0.001 ± 0.001
IVUS	0.990 ± 0.003	0.944 ± 0.049	0.961 ± 0.023	0.002 ± 0.001
Thyroid	0.989 ± 0.006	0.938 ± 0.051	0.889 ± 0.043	0.002 ± 0.001

Algorithm 1: Iterative random walks (IRW) algorithm.

Input: Complete layer specific set of marked seeds $S_M \in \{S_{M1}, S_{M2}, \dots, S_{Mk}\}$ with k number of layers and all the unmarked seeds S_U where present contour $\mathcal{G}_P = S_M \cup S_U$

Input: Stopping criteria δ

Output: Optimal marking the unmarked seeds (\mathcal{G}_P)

for $i \leftarrow 1$ **to** n **do**

Compute layer specific prior probability $p(L_Z|x, I)$ from (5);

Generate new contour \mathcal{G}_N ;

Count change of pixel $\mathcal{C}_P = \text{count}(\mathcal{G}_P - \mathcal{G}_N)$;

If $\mathcal{C}_P > \delta$ **then**

$\mathcal{G}_P = \mathcal{G}_N$;

Reinitialize S_M and S_U ;

else

Return \mathcal{G}_P ;

end

end

problem. A RF model is learned (Sec. III-B) by training over a feature set (Sec. III-A) over 500 random samples, selected for each class in the training images and has been tested on these three middle frame in the test volume, which acts as initial segmentation of the layers for IRW (Sec. III-D). In the next step, a volume specific RF model is designed using the information in these three segmented frames. The model predicts the layers in the other frames of the particular volume. Subsequently a frame specific GVF is obtain for a frame using the previous two neighboring frames (Sec. III-C). Finally, we have taken the intersection of the two results from both RF and GVF model, which has acts as initial seeds for segmentation using IRW on a given frame. The algorithm has been experimentally evaluated on kidney segmentation using a multi-modal abdominal phantom (CIRS abdomen phantom),¹ where four 3D US were acquired using a tomographic ultrasound device (Piur Imaging, Germany).² The results of the segmentation are shown in Table I together with the ground truth results are visually illustrated in Fig. 3. Beyond the phantom based validation we also employ it for IVUS and thyroid segmentation using real patient data.

¹<http://www.cirsinc.com/products/all/65/triple-modality-3d-abdominal-phantom/>

²<https://www.piurimaging.com/>

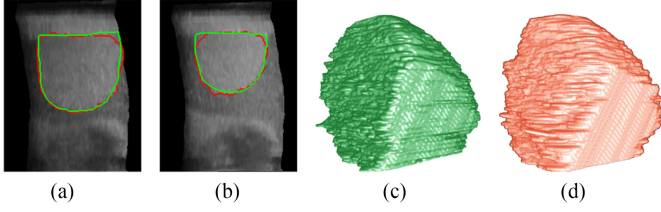


Fig. 3. Contour segmentation for kidney phantom where Fig. 3(a) and Fig. 3(b) are two frames from different volumes. Fig. 3(c) is 3D visualization of the segmented phantom kidney and Fig. 3(d) is the ground truth volume of the same. GREEN - ground truth data and RED - result of our proposed method.

A. Intravascular Ultrasound Segmentation

1) **Data Description:** The IVUS data used in this experiment is from the Lumen + External Elastic Laminae border detection of the IVUS Challenge dataset³ [7]. It consists of 10 pullbacks each from a different patient acquired at 20 MHz. The data acquisition technique is detailed in [7]. All the frames are provided in DICOM format as grayscale images. The manually labeled ground truth of the lumen and external elastic laminae are also provided with this set of data. Each frame has a size of 384×384 pixels.

2) **Experiments and Results:** At the time of random forest learning (Sec. III-B), a D (number of features) dimensional ordered vector Θ representing multiscale estimated Fisher-Tippett statistics parameter and ultrasonic signal confidence is computed and represented as $\{(\Theta; r) \forall r \in G\}$, where G is the IVUS image represented in polar domain. In this experiments we have the tissue specific labels $y \in Y$ with $Y = \{lumen, media, externa\}$ corresponding to the ultrasound echo measurements at grid points $r \in G$. Statistics of backscattered ultrasonic B-mode signals parameters are estimated at scales $\{(3 \times 3), (5 \times 3), (7 \times 3), \dots, (30 \times 3)\}$ and this range of values are chosen according to [4]. This code is implemented on Python.⁴ The random forest⁵ classifier is learnt using hyper-parameter detailed in Table III. This trained model is finally tested on pre-selected test images (not used during training). This experiment is performed using a 10-fold cross-validation technique. The data was taken from 10 patient cumulatively yielding 2,175 images used in this experiment and at each fold data from 9 patients were used for training and the remaining patient data was used for testing.

Evaluation for segmentation of IVUS is performed using three measures (i) Jaccard index (JCC), (ii) Percentage of Area Difference (PAD) measure, and (iii) Hausdorff Distance (HD) following the approach presented in [7]. We compare our method with eight prior art i.e. shape-driven segmentation method [7] (P1), geodesic active contour [7] (P2), fast-marching method based on gray level distributions [7] (P3), graph search method [7] (P4), multi-scale stacked sequential learning [7] (P5), holistic approach [7] (P6), support vector machine with a radial basis

TABLE II
PERFORMANCE EVALUATION METRICS OF RESULTS OBTAINED WITH DATASET AND COMPARISON WITH PRIOR ART. EEL DENOTES External Elastic Laminae

Methods	JCC		HD		PAD		DC	
	Lumen	EEL	Lumen	EEL	Lumen	EEL	Lumen	EEL
P1 [7]	0.81 ± 0.12	0.76 ± 0.13	0.47 ± 0.39	0.64 ± 0.48	0.14 ± 0.13	0.21 ± 0.16	—	—
P2 [7]	0.83 ± 0.08	—	0.51 ± 0.25	—	0.14 ± 0.12	—	—	—
P3 [7]	0.88 ± 0.05	0.91 ± 0.04	0.34 ± 0.14	0.31 ± 0.12	0.06 ± 0.05	0.05 ± 0.04	—	—
P4 [7]	0.77 ± 0.09	0.74 ± 0.17	0.47 ± 0.22	0.76 ± 0.48	0.15 ± 0.12	0.23 ± 0.19	—	—
P5 [7]	0.79 ± 0.08	—	0.46 ± 0.30	—	0.16 ± 0.09	—	—	—
P6 [7]	—	0.84 ± 0.10	—	0.57 ± 0.39	—	0.12 ± 0.12	—	—
P7 [7]	0.84 ± 0.08	—	0.38 ± 0.26	—	0.11 ± 0.12	—	—	—
P8 [7]	0.81 ± 0.09	0.79 ± 0.11	0.42 ± 0.22	0.60 ± 0.28	0.11 ± 0.11	0.19 ± 0.19	—	—
Intra-obs [7]	0.88 ± 0.05	0.92 ± 0.03	0.28 ± 0.13	0.24 ± 0.12	0.11 ± 0.08	0.06 ± 0.04	—	—
Inter-obs [7]	0.93 ± 0.05	0.95 ± 0.03	0.17 ± 0.13	0.14 ± 0.10	0.04 ± 0.06	0.03 ± 0.03	—	—
Intermediate	0.911 ± 0.045	0.917 ± 0.049	0.239 ± 0.134	0.227 ± 0.116	0.092 ± 0.084	0.059 ± 0.050	0.876 ± 0.051	0.903 ± 0.073
Proposed	0.927 ± 0.017	0.947 ± 0.027	0.168 ± 0.081	0.173 ± 0.114	0.002 ± 0.001	0.002 ± 0.001	0.909 ± 0.025	0.922 ± 0.035

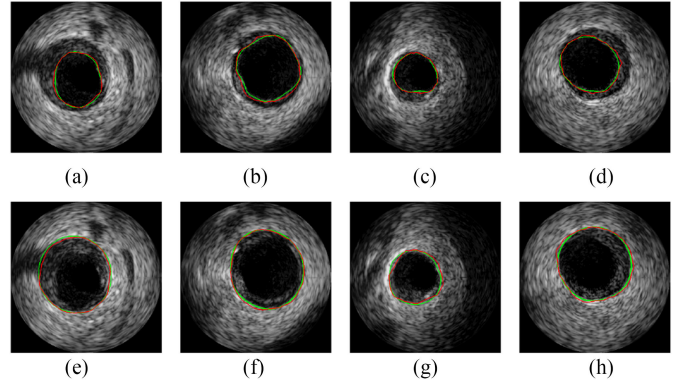


Fig. 4. 4(a), 4(b), 4(c), and 4(d) are four segmented lumen contour and 4(e), 4(f), 4(g), and 4(h) are four segmented external elastic laminae (EEL) contour from four different IVUS pullback. Here, GREEN – ground truth data and RED – result of our proposed method.

function kernel [7] (P7), and angiocard [7] (P8) respectively which has been reported by Balocco *et al.* [7]. The results are quantitatively detailed in Table II.

B. Thyroid Segmentation

1) **Data Description:** A total of sixteen healthy human subjects were imaged using a Logiq E9 US device (General Electric, USA) with a 11–16 MHz probe and equipped with an electromagnetic (EM) tracking system to form the free hand US volume dataset. All the volume are provided in DICOM format as RGB images. The datasets are available online⁶ [15].

2) **Experiment Implementation and Results:** The experimental setup for thyroid segmentation is similar to IVUS as presented earlier. We have used tissue-specific labels $y \in Y$ and $Y = \{thyroid, background\}$ corresponding to the ultrasound echo measurements. The ultrasonic signal confidence is estimated using the parameters shown in Table III. The RF parameters (Table III) are similar to as used for earlier use case of

³<http://www.cvc.uab.es/IVUSchallenge2011/dataset.html>

⁴http://scikit-image.org/docs/dev/auto_examples/segmentation/plot_random_walker_segmentation.html

⁵<http://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

⁶<http://opencaas.webarchiv.kit.edu/?q=node/29>

TABLE III
HYPER-PARAMETERS FOR BOTH THE EXPERIMENTS

Hyper-parameters	IVUS	Thyroid	Explanation
$nClasses$	3	2	Number of classes
α	2	2	It affects the likelihood of vertical random walks
β	1250	5500	It effects on the robustness and accuracy of the segmentation
γ	0.05	0.05	It penalizes horizontal and diagonal random walks in the graph
$nTrees$	50	50	Number of trees
\mathcal{D}	17	17	Number of features
$minLeaf$	50	50	Number of leaves for each trees
$treeDepth$	∞	∞	Level of the tree
$splitObj$	GDI	GDI	Gini's diversity index

TABLE IV
QUANTITATIVE ANALYSIS WITH THE THYROID DATASET
AND COMPARISON WITH PREVIOUS APPROACHES

Methods	SE	SP	DSC	PPV	JCC	HD	PAD
Narayan [2]	0.955 ± 0.024	0.889 ± 0.065	0.839 ± 0.047	0.806 ± 0.086	—	—	—
JCR [2]	0.564 ± 0.070	0.926 ± 0.070	0.479 ± 0.066	0.326 ± 0.057	—	—	—
Chang [23]	0.874 ± 0.117	0.560 ± 0.324	0.512 ± 0.288	0.531 ± 0.347	—	—	—
Garg [14]	0.473 ± 0.182	0.864 ± 0.224	0.400 ± 0.143	0.265 ± 0.111	—	—	—
Intermediate	0.937 ± 0.028	0.892 ± 0.068	0.828 ± 0.079	0.795 ± 0.157	0.889 ± 0.054	0.538 ± 0.488	0.011 ± 0.015
Proposed	0.989 ± 0.006	0.923 ± 0.059	0.854 ± 0.066	0.807 ± 0.110	0.908 ± 0.028	0.488 ± 0.430	0.002 ± 0.001

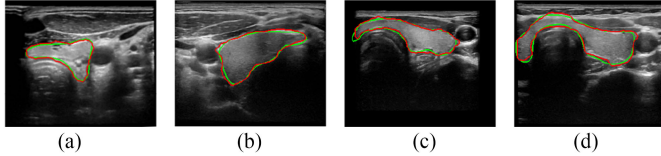


Fig. 5. Contour segmentation for thyroid where 5(a), 5(b), 5(c), and 5(d) are four frames from four different volumes. GREEN – ground truth data and RED – result of our proposed method.

IVUS pullback segmentation (Sec. IV-A2). This experiment is performed using a 10-fold cross-validation technique. The data was obtained from 16 patients for this experiment and at each fold 14–15 volumes were used for training and the remaining volumes were used for testing.

The segmentation performance of our algorithm applied to thyroid US has been evaluated in two different ways following the approach in [2]. The proposed algorithm is qualitatively assessed using (i) Probabilistic rand index (PRI), (ii) Global consistency error (GCE), (iii) Variation of information (VOI) and (iv) Boundary distance error (BE). Also measures quantitatively based on the overlapping area are used, i.e. (i) Sensitivity (SE), (ii) Specificity (SP), (iii) Dice coefficient (DSC) and (iv) Positive predictive value (PPV). We have compared our approach with four different algorithms i.e. echogenicity-based quantization [2], joint classification-regression [2], RBF neural network [23], and feedforward neural network [14] which has been reported by Narayan *et al.* [2], (Table IV). The only limitation trying to compare commonality is that the dataset has

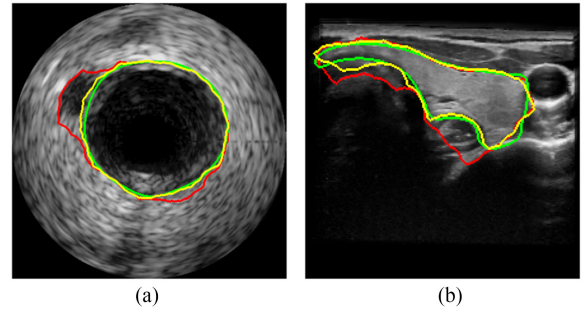


Fig. 6. Contour segmentation in presence of artifacts for IVUS in 6(a) and thyroid in 6(b) where GREEN – ground truth of the contour, RED – result after applying only RF and YELLOW - result after applying IRW on the result of RF.

used for this comparison is not same as the dataset used in prior art [2].

V. DISCUSSION

A. Initializing Seeds of Random Walks Through Learning of Statistical Mechanics of Ultrasound

A key feature of this algorithm is its ability to predict the tissue layers in US images by learning of parametric model of speckle statistics. This prediction facilitates a reasonable initial estimate for RW seed initialization. The RF model is trained using 500 samples from each tissue type per image. Except for a few limitations like guidewire artifact, necrotic core, muscles, trachea, and carotid, this algorithm yields perfect accurate results for segmenting the lumen, external elastic laminae as well as the thyroid. While the RF-based contour initialization on each frame acts as a good starting point, it has limitations in not being able to correctly steer the contour across neighboring frames in the presence of necrotic core as is evident in the red contour on Fig 6(a) and trachea as is evident in the red contour on the Fig 6(b). To overcome this limitation, IRW has been employed iteratively in our formulation termed IRW as can be seen in the yellow contour in Fig 6.

B. Iterative Random Walker for Correcting Contour in Presence of Initialization Error

Due to the false prediction in the presence of the different artifacts, we have employed a subsequent refining stage using the RW for final segmentation using the seeds initialized by the RF model. In this experiment, 0.1% change of the contour is set as the convergence criteria for IRW. Fig 6 illustrates that IRW is well suited to correct the final contour. In Table II, the method *Intermediate* shows performance of the RF initialized IRW based segmentation obtained using method described in Sec. III-B and Sec. III-D, which is comparably better than prior art [7] and the segmented contours are visually illustrated marked in yellow contours in Fig. 6.

This approach while improving on the results, has the drawback when applied to a whole pullback for segmentation. When the seeds initialized for the contour are erroneous, the

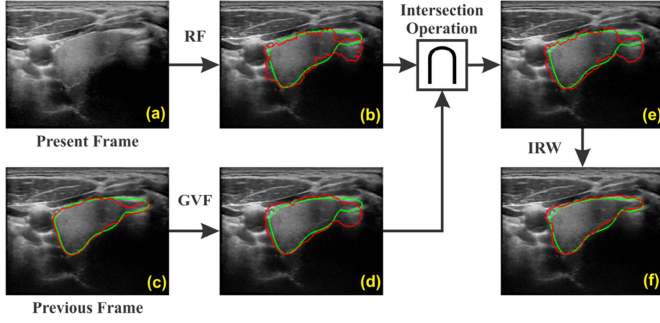


Fig. 7. The graphical framework for the minimization of leaking and drifting from the RF model and GVF respectively.

computation time increases. This problem is accentuated for small size contour, i.e., it is difficult for the RF model to predict small area accurately, which sometimes leads to inaccurate segmentation results. In order to overcome this problem, we have suggested the use of belief propagation across neighboring frames as a possible solution, and its benefits are discussed subsequently.

C. Belief Propagation for Cross Frame Segmentation

The solution of the problem for segmentation on the whole volume using a frame wise approach has been discussed in Sec. V-B. We have employed belief propagation, where the predicted contour in the current frame is propagated to the next frame as an initial estimate of the contour, and then an IRW initialized with these seeds corrects the initial contour. The drawback of this method is that the error if accrued intermediate is propagated till the last frame. This helps in putting the belief propagation of the previous result with in a bound. In this method, we again employ another RF model, which is learnt to be the volume specific using the set of three segmented middle frames. This RF model predicts roughly the tissue presence for a frame, and for the same frame, the initial contour comes from the results of the previous frames as well. The intersection between these two contours is then used for the initial contour for IRW. If the RF model is predicted wrongly, the propagated results can also be used to correct the contour and vice versa. To get more accurate belief propagated contour, we introduce a gradient vector flow (GVF) over contour drifts across neighboring frames instead of conveying the exact result from the neighboring frame. The GVF is calculated using the last two frames cumulative neighbors (Sec. III-C). The propagation error has decreased through this combined method. The graphical framework for the minimization of leaking has been illustrated in Fig. 7. Finally, all the segmented frames have been stacked into a volume and evaluated in comparison with the ground truth volume shown in Table I and visualized using MeVisLab, as shown in Fig. 8.

In this experiment we have obtained an average JCC value of 0.927 ± 0.017 and 0.947 ± 0.027 , HD of 0.168 ± 0.081 and 0.173 ± 0.114 , and PAD value of 0.002 ± 0.001 and 0.002 ± 0.001 for the segmentation of lumen and external elastic luminae

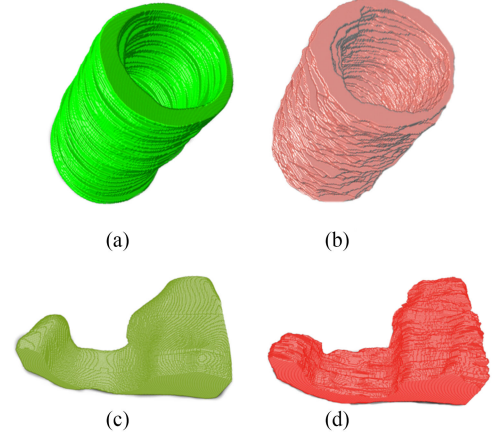


Fig. 8. 8(a), and 8(b) are the media volume of IVUS where 8(a) is the ground truth and 8(b) is the segmented media volume. 8(c), and 8(d) are the thyroid volume where 8(c) is the ground truth and 8(d) is the segmented thyroid volume.

TABLE V
PERFORMANCE MEASURE WITH THE THYROID DATASET AND COMPARISON WITH PREVIOUS APPROACH IN TERMS OF PRI, GCE, VOI, AND BE

Methods	PRI	GCE	VOI	BE
Narayan [2]	0.844 ± 0.037	0.192 ± 0.044	1.169 ± 0.206	12.221 ± 2.788
Proposed	0.972 ± 0.011	0.024 ± 0.009	0.172 ± 0.054	6.660 ± 1.962

boarder respectively in IVUS. Table II, shows the performance of our method obtained from Sec. III-C, which are better than as reported in the prior art [7]. Fig. 4(a)–4(d) and Fig. 4(e)–4(h) presents the visualization of the segmented contours for lumen and external elastic luminae from four different IVUS pullbacks. The visual representation of the segmented media for whole pullback is shown in Fig. 8(b), which is visibly similar to the ground truth Fig. 8(a).

In the experiment of thyroid contour segmentation, the performance of the proposed method has been evaluated in two different ways. In Table IV and Table V, quantitative and qualitative results are summarized for our algorithm in comparison with approaches presented in prior art. In both cases it is clear that our approach outperforms the other compared algorithms [2]. Fig. 5(a)–5(d) presents the visualization of segmented thyroid contours from four different thyroid volume and volumetric visual representation of thyroid has shown in Fig. 8(d) which is similar to the ground truth volume in Fig. 8(c) for thyroid segmentation.

D. Execution Time

The algorithm was implemented in Python 2.7 and executed on a PC with Intel Core i5 CPU operating at 3.20 GHz, 24.0 GB of RAM, and Ubuntu 14.04 LTS operating system. Per frame processing time was under 1.15 ± 0.05 s and 1.23 ± 0.27 s respectively for IVUS and thyroid segmentation respectively, without the need of deployment on any computing accelerators like GPUs. In order to process an IVUS volume of

$384 \times 384 \times 100$ it takes around 115 s and a thyroid US volume of size $372 \times 252 \times 100$ it takes about 123 s.

VI. CONCLUSION

The approach to layer characterization and subsequent boundary detection using the ultrasonic backscattered signals operating on the freehand ultrasound images is a crucial task. In this paper, we have presented a computationally efficient and robust algorithm to automatically segment anatomical structures and without need for any user-interaction for initializing contours in US image sequences. The algorithm addresses the problem of freehand US volume segmentation using the formulation of a machine learning based model with the help of mathematical models of US backscattering statistical mechanics. Further to this with use of belief propagation across neighboring frames of the volume. The method shows high accuracy while consuming less time and yielding robust contours. Those abilities are in fact a significant advantage over existing tissue classifiers that are not adaptive to patient level variations. The proposed method can directly be leveraged to facilitate online learning or relearning on new cases for improved performance of segmentation without any manual interaction for seeding. The algorithm also demonstrates (1) reliable US layer segmentation from different tissue present in the frame and independent of US application, (2) interframe contour segmentation consistency in the volume, (3) independence to the size and shape of anatomical structure for detection of the contour, (4) ability to segment in the presence of various artifacts in the frame and (5) accurate prediction using a dominant learning model (RF) and less time complexity of per frame processing. These attributes make the method unique and better performing in comparison with state-of-art of US image and volume segmentation.

REFERENCES

- [1] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.
- [2] N. S. Narayan, P. Marziliano, J. Kanagalingam, and C. G. Hobbs, "Speckle patch similarity for echogenicity-based multiorgan segmentation in ultrasound images of the thyroid gland," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 172–183, Jan. 2017.
- [3] B. Saleh, *Introduction to Subsurface Imaging*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [4] D. Sheet *et al.*, "Joint learning of ultrasonic backscattering statistical physics and signal confidence primal for characterizing atherosclerotic plaques using intravascular ultrasound," *Med. Image Anal.*, vol. 18, no. 1, pp. 103–117, 2014.
- [5] M. Papadogiorgaki, V. Mezaris, Y. S. Chatzizisis, G. D. Giannoglou, and I. Kompatsiaris, "Automated IVUS contour detection using intensity features and radial basis function approximation," in *Proc. 20th IEEE Int. Symp. Comput.-Based Med. Syst.*, 2007, pp. 183–188.
- [6] F. Ciompi *et al.*, "Holimab: A holistic approach for media-adventitia border detection in intravascular ultrasound," *Med. Image Anal.*, vol. 16, no. 6, pp. 1085–1100, 2012.
- [7] S. Balocco *et al.*, "Standardized evaluation methodology and reference database for evaluating IVUS image segmentation," *Comput. Med. Imag. Graph.*, vol. 38, no. 2, pp. 70–90, 2014.
- [8] D. China, M. K. Nag, K. Mandana, A. K. Sadhu, P. Mitra, and C. Chakraborty, "Automated in vivo delineation of lumen wall using intravascular ultrasound imaging," in *Proc. 38th Annu. Int. Conf., IEEE Eng., Med., Biol. Soc.*, 2016, pp. 4125–4128.
- [9] D. China, P. Mitra, and D. Sheet, "Segmentation of lumen and external elastic laminae in intravascular ultrasound images using ultrasonic backscattering physics initialized multiscale random walks," in *Proc. Int. Conf. Comput. Vis., Graph., Image Process.*, 2016, pp. 393–403.
- [10] D. China, P. Mitra, and D. Sheet, "On the fly segmentation of intravascular ultrasound images powered by learning of backscattering physics," in *Classification in BioApps*. New York, NY, USA: Springer, 2018, pp. 351–380.
- [11] E. G. Keramidas, D. K. Iakovidis, D. Maroulis, and S. Karkanis, "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection," in *Proc. Int. Conf. Image Anal. Recognit.*, 2007, pp. 1052–1060.
- [12] D. E. Maroulis, M. A. Savelonas, D. K. Iakovidis, S. A. Karkanis, and N. Dimitropoulos, "Variable background and texture for delineation of computer-aided delineation of nodules in thyroid ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 5, pp. 537–543, Sep. 2007.
- [13] M. A. Savelonas, D. K. Iakovidis, I. Legakis, and D. Maroulis, "Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 4, pp. 519–527, Jul. 2009.
- [14] H. Garg and A. Jindal, "Segmentation of thyroid gland in ultrasound image using neural network," in *Proc. Int. Conf. Comput. Commun. Netw. Technol.*, 2013, pp. 1–5.
- [15] T. Wunderling, B. Golla, P. Poudel, C. Arens, M. Friebe, and C. Hansen, "Comparison of thyroid segmentation techniques for 3d ultrasound," *Med. Imag. Image Process.*, vol. 10133, 2017, Art. no. 1013317.
- [16] R. Shekhar, R. Cothren, D. Vince, S. Chandra, J. Thomas, and J. Cornhill, "Three-dimensional segmentation of luminal and adventitial borders in serial intravascular ultrasound images," *Comput. Med. Imag. Graph.*, vol. 23, no. 6, pp. 299–309, 1999.
- [17] J. D. Klingensmith, R. Shekhar, and D. G. Vince, "Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images," *IEEE Trans. Med. Imag.*, vol. 19, no. 10, pp. 996–1011, Oct. 2000.
- [18] M.-H. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, É. Therasse, and G. Cloutier, "Intravascular ultrasound image segmentation: A three-dimensional fast-marching method based on gray level distributions," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 590–601, May 2006.
- [19] S. Sun, M. Sonka, and R. R. Beichel, "Graph-based IVUS segmentation with efficient computer-aided refinement," *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1536–1549, Aug. 2013.
- [20] E. G. Mendizabal-Ruiz, M. Rivera, and I. A. Kakadiaris, "Segmentation of the luminal border in intravascular ultrasound b-mode images using a probabilistic approach," *Med. Image Anal.*, vol. 17, no. 6, pp. 649–670, 2013.
- [21] F. S. Zakeri, S. K. Setarehdan, and S. Norouzi, "Automatic media-adventitia IVUS image segmentation based on sparse representation framework and dynamic directional active contour model," *Comput. Biol. Med.*, vol. 89, pp. 561–572, 2017.
- [22] E. N. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger, and T. Kuwert, "Quantification of thyroid volume using 3-d ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 27, no. 4, pp. 457–466, Apr. 2008.
- [23] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, and S.-R. Shih, "Thyroid segmentation and volume estimation in ultrasound images," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1348–1357, Jun. 2010.
- [24] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.
- [25] J. A. Noble, "Ultrasound image segmentation and tissue characterization," in *Proc. Inst. Mech. Eng., H, J. Eng. Med.*, 2010, vol. 224, no. 2, pp. 307–316.
- [26] P. K. Sen and J. M. Singer, *Large Sample Methods in Statistics*. London, U.K.: Chapman & Hall, 1993.
- [27] A. Karamalis, W. Wein, T. Klein, and N. Navab, "Ultrasound confidence maps using random walks," *Med. Image Anal.*, vol. 16, no. 6, pp. 1101–1112, 2012.
- [28] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.