

A topo-graph model for indistinct target boundary definition from anatomical images

Hui Cui¹, Xiuying Wang^{1*}, Jianlong Zhou², Guanzhong Gong³, Stefan Eberl^{1,4}, Yong Yin³, Lisheng Wang⁵, Dagan Feng^{1,6} and Michael Fulham^{4,7}

¹ School of Information Technologies, The University of Sydney, Sydney, Australia

² Data61, CISRO, Sydney, Australia

³ Department of Radiation Oncology, Shandong Tumor Hospital, Jinan, China

⁴ Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, Sydney, Australia

⁵ Institute of Image Processing and Pattern Recognition, Department of Automation, Shanghai Jiao Tong University, Shanghai, China

⁶ Med-X Research Institute, Shanghai Jiao Tong University, Shanghai, China

⁷ Sydney Medical School, The University of Sydney, Sydney, Australia

Abstract

Background and Objective: It can be challenging to delineate the target object in anatomical imaging when the object boundaries are difficult to discern due to the low contrast or overlapping intensity distributions from adjacent tissues. **Methods:** We propose a topo-graph model to address this issue. The first step is to extract a topographic representation that reflects multiple levels of topographic information in an input image. We then define two types of node connections - nesting branches (NBs) and geodesic edges (GEs). NBs connect nodes corresponding to initial topographic regions and GEs link the nodes at a detailed level. The weights for NBs are defined to measure the similarity of regional appearance, and weights for GEs are defined with geodesic and local constraints. NBs contribute to the separation of topographic regions and the GEs assist the delineation of uncertain boundaries. Final segmentation is achieved by calculating the relevance of the unlabeled nodes to the labels by the optimization of a graph-based energy function. We test our model on 47 low contrast CT studies of patients with non-small cell lung cancer (NSCLC), 10 contrast-enhanced CT liver cases and 50 breast and abdominal ultrasound images. The validation criteria are the Dice's similarity coefficient and the Hausdorff distance. **Results:** Student's t-test show that our model outperformed the graph models with pixel-only, pixel and regional, neighboring and radial connections (p-values < 0.05). **Conclusions:** Our findings show that the topographic representation and topo-graph model provides improved delineation and separation of objects from adjacent tissues compared to the tested models.

Keywords: Topology, Graph, Segmentation

1. Introduction

The accurate segmentation of a target object from adjacent tissues is a fundamental component of medical image processing. It is critical for accurate diagnosis and has major ramifications for patient management as surgical, radiological intervention and radiotherapy approaches often depend on the relationship of the target to its surrounding structures. Image segmentation is also typically performed in feature analysis of tissues for disease classification. Automated target object segmentation and boundary delineation algorithms have attracted intensive research interest because of their efficiency and objectivity when compared to manual approaches. An ongoing challenge for anatomical image segmentation is the delineation of blurred or indistinct object boundaries. Ideally, objects of interest belonging to the foreground would be dissimilar in appearance or have disparate features to the background. In reality, objects of interest and background regions can have overlapping intensity distributions or low contrast on anatomical images. Under these circumstances, it is difficult to delineate the boundaries of the target object, especially when it abuts adjacent tissues that have a similar

appearance. For example, a lung tumor on chest CT may be difficult to separate from soft tissues of the chest wall or associated collapse and consolidation.

Target object segmentation refers to segmentation where users provide inference for objects of interest or background, by using labels or scribbles. Segmentation is then achieved by assigning labels to unknown/unlabeled image regions according to their relevance to the labels. Active contour [1] models have been investigated by a number of researchers for computer-aided anatomical image segmentation. Active contours have obtained promising performance when there was contrast variability and non-uniform illumination [2]. Edge-based geometric models [3] were among the early level set formulations that minimized an objective function of the contour's geodesic length [4]. For blurry or weak boundaries, region-based active contour models [5] were proposed to control the motion of level set functions by using a region descriptor [6]. The region descriptor can be obtained by unsupervised or supervised statistical models such as clustering models and Gaussian models [4] [7].

Graph-based methods have shown great value in target tumor or organ boundary delineation from medical images [8]. For

this work, we focused on graph-based algorithms and designed a new graph model to delineate indistinct boundaries in anatomical images.

Graph theory has a long history in mathematics and computer science [9] and graph models have been used in image processing, for semi-supervised clustering, user interactive segmentation and saliency detection [8]. Fundamental graph theory approaches such as graph cut (GC) [10] and the random walk algorithm (RW) [11] have been widely used [12-14]. The RW [11] can capture the local intensity changes and solve ‘weak boundary’ problems for various organs from different image modalities [11, 15, 16]. Graph models have also been used in the co-delineation of the target object boundary, for instance a RW based model [14] and a combined GC and Markov Random Field (MRF) model [17, 18] have been applied to regulate and penalize the energy functions of PET and CT intensity distributions for co-segmentation. In all these aforementioned graph models, the nodes represent 4 or 8 connected adjacent/neighbors pixels in two dimensional (2D) images. In 3D volumes, there can be 6, 18 or 26 adjacent connections. The weights of the graph edges are calculated to reflect the changes in the neighboring feature spaces of intensity or gradient magnitude. When the target object’s appearance is inhomogeneous, the local pixel level intensity similarities are insufficient to achieve good segmentation results [19-22]. Hence, additional approaches are required to improve segmentation accuracy and these approaches include the construction of specialized graph models and the incorporation of prior knowledge to assist the local pixel level information.

Specialized graph models vary in the definition of nodes. In some graph models, regions are used as graph nodes [23-25] where the images are pre-processed and partitioned into a number of irregular regions in an unsupervised manner such as mean shift, quick shift etc. Region-level nodes represent more informative image features and textures than pixel-level nodes. These specialized graph models have thus improved segmentation results over textured images [23]. Models with region-level information also have the advantage of propagating local grouping cues to broader image ranges while minimizing the influence of frequent local intensity changes or noise, but the segmentation results are largely dependent on the initial region partition results, as perfect pre-partitions cannot always be guaranteed [25].

The construction of graph models also vary in node connections / edge definitions and these can be categorized as geometrical and topological connections. In geometrical connections, the edges are constructed based on the spatial locations of nodes. For instance a radial connection is proposed in the graph model [24] where each node is connected to its neighboring nodes as well as the nodes sharing the common boundaries with these neighboring nodes. Other edge definitions include full connection. When the number of nodes is limited, the nodes can be fully connected to achieve the whole image information propagation [23]. Full connection, however, decreases computation efficiency with increasing node numbers and we previously reported that full regional connections might produce misleading grouping information

and result in leakage or under-segmentation of inhomogeneous objects [26]. Topological connections and topological graphs provide abstracted and structured data representation that can be analyzed. These abstractions and representations are used in volume rendering and scalar field visualization [27]. Topological graph models include an ‘extreme graph’, which is an abstraction of the gradient flow of an image [28] and contour tree that represents how the level sets merge and split to form individual components [29]. The construction of topology graphs can be complicated in medical image processing. For instance, since PET images are noisy, the conventional contour tree construction methods may produce giant size tree structures making data analysis and visualization impractical [21, 30]. Hence, research has focused on the simplification of topological representation [27, 30, 31]. Recently, we have explored region of interest (ROI) topology extraction from PET images and exploited the topological connections to graph based segmentations from PET-CT images [21, 22, 32].

As mentioned above, incorporating prior knowledge or prior models is another way to alter conventional graph based segmentation that affects the edge weights calculations. Yang et al. [20] proposed a segmentation model where the prior distributions were estimated by Gaussian mixture models (GMM) and the weights were calculated with the estimated prior models instead of intensity values. If there are overlapping intensity distributions between foreground and background regions, however, the prior models may not be able to solve the separation or weak boundary problems. [21] Kim et al. [23] reported a boundary based weighting function, which measured the gradient magnitude along the edges, to capture the object boundary when the foreground and background had similar intensity distributions. This approach, however, may result in over-segmentation because of the frequent changes of gradient magnitudes of textured objects. Grady et al. further suggested a combined intensity and boundary weighting function where it was assumed that the object was more likely to be homogeneous than the background, but this is generally not the case with medical imaging [33].

In contrast to Euclidean distance, the geodesic distance can capture the geometric structure of the data and has been used in various image analysis and computer vision tasks. For instance, it was used as a shape and surface descriptor for object matching [34, 35] and classification [36]. Liu et al. [37] used geodesic distance to capture the depth information in an image for low resolution image up-sampling. When segmenting images to superpixels, Wang et al. [38] proposed a structure-sensitive superpixel generation algorithm that incorporated intensity information and the geodesic distance. Their results showed improved performance when using the geodesic distance information.

1.1. Our contributions

The unique contributions of our model include the introduction of a new topographic image representation to extract regional relationships in a multiple-level manner, in particular, for adjacent tissues that have similar intensity

distributions. The image topography provides regional information (referred to as topoRegion) and more importantly, the topological relations of multilevel topoRegions inherited from the concept of contour tree. The number of topoRegions is automatically determined by the density estimation and the partition of the feature space where local entropy is employed for boundary definition. Further, we use a new weighted graph model where two types of weighted edges are constructed to associate the pairwise similarities and link multilevel topoRegions according to the topographic relations. Finally, given the pre-defined labels for the target object, the relevance of individual unlabeled nodes to the given seeds/labels was calculated by the optimization of a quadratic cost function. The function is composed of inter-smoothness and intra-correlation terms to associate the label information and the embedded affinities in the graph edges.

The paper is structured as follows: Section 2 introduces the graph model and experimental dataset; in Section 3 we present the experimental evaluation results and discussion, the concluding statements are in Section 4.

2. Methods: Topo-Graph Model

2.1. Related work: Graph-based image segmentation

Given a set of foreground F and background B labels $L = \{F, B\}$ in an image, the target object segmentation can be achieved by ranking the unlabelled pixels according to their relevance to the given labelled pixels / seeds as a query by solving a graph model.

A graph model $G = (V, E, \mathbf{W})$ is constructed with a graph node $v_i \in V$ representing an image pixel or a region, a graph edge $e_{ij} \in E$ connecting nodes v_i, v_j , a weight $w_{ij} \in \mathbf{W}_{|V| \times |V|}$ denoting the similarities between connected nodes. As discussed by Couprie et al. [39], a graph-based optimization framework can be generalized with a unary term that formulates the nodes pairwise and a binary term that penalizes the nodes and edges as:

$$\Psi(\mathbf{f}) = \sum_{e_{ij} \in E} w_{ij} \|f_i - f_j\|^2 + \sum_{e_{ij} \in E} \sum_j w_{ij} \|f_i - f_i^*\|^2, \quad \begin{cases} f_i^* = 1, v_i \in V_F \\ f_i^* = 0, v_i \in V_B \end{cases} \quad (1)$$

where $\mathbf{f} = [f_{iL_k}]_{|V| \times 2}$ and f_{iL_k} is the relevance of an image pixel/graph node to the $F(B)$ labels. $V_F (V_B)$ denotes graph nodes initialized with $F(B)$ labels. By optimizing Eq.1, the final segmentation is obtained by ranking the remaining unlabeled nodes $V_U = V \setminus (V_F \cup V_B)$ according to their relevance to the labeled nodes.

Most previous graph models have nodes that are defined as pixels or super-pixels and edges that are designed to connect adjacent or neighboring nodes. These models may not work when the target object boundaries are difficult to discern due to the low contrast or overlapping intensity distributions from adjacent tissues.

2.2. Hypothesis and framework

Our hypothesis is that embedding topographic regional information in a graph model will assist the object separation and boundary delineation especially when the target object is in close proximity to other regions with overlapping intensity distributions in anatomical images. The overview of the proposed Topo-Graph model is given in Fig. 1. Given input images and user inference of the target object to be segmented, a topographic representation is extracted and indexed by constructing initial and fine-grained topology trees. Then, the graph model is defined with nesting branches connecting nodes corresponding to initial level topoRegions and geodesic edges linking the fine level topoRegions. Lastly, the segmentation is achieved by estimating the relevance of the unlabeled nodes to the labeled nodes by graph regulation.

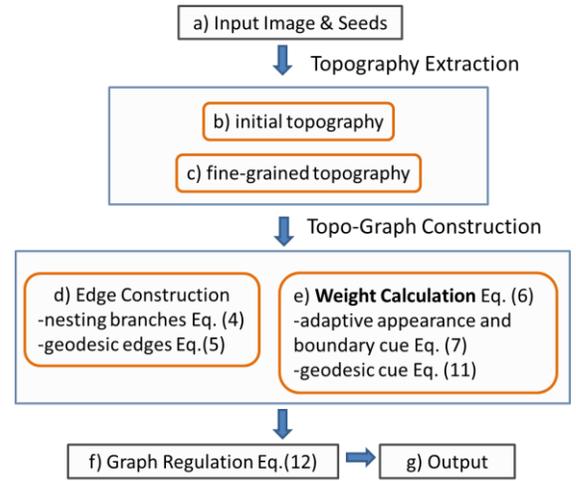


Fig.1. Overview of the proposed Topo-Graph model. For a) an input image, d) the edges in topo-graph are constructed with nesting branches connecting topoRegions corresponding to (b) initial topology tree and geodesic edges linking topoRegions at detailed level based on c) fine-grained topology. e) The graph weighting function associates adaptive appearance and boundary cues and geodesic distance information. The final segmentation is achieved by f) graph regulation.

2.3. Multi-level Topographic Representation

We define the topographic representation as a visual index of image regions inspired by the concept of contour tree. The contour tree is an abstract description of the image and reflects how the iso-contours merge and split to form individual components [40]. In the contour tree structure, the list of nodes corresponds to the critical points that are the local extremes and saddle points in the image. The arcs represent a set of regions where each of the regions is composed of the set of iso-contours between the critical points. Conventional contour tree generation methods, such as merging the split and joint trees [41], can be unmanageable and impractical for clinical image analysis because of noise and artifacts in real-world data. Thus, further simplification procedures are normally performed to optimize the tree size so that it is small enough for user interaction while maintaining the essential structure of the data.

Nevertheless, in this work, we propose an alternative approach to initially identify a set of regions and assign them the topological relation based on the concept of a contour tree. Then these initial regions are locally refined to find detailed topologies.

2.3.1 Initial topography extraction

The initial topography is designed to reflect the landscape of the whole image domain. To obtain the initial level topoRegions, we partition the image by firstly estimating the density distribution of feature space and secondly clustering the feature space according to optimal threshold levels.

1) Feature space

Local entropy of the input image is exploited as the feature space. The areas with high intensity changing frequency such as region boundaries appear as ‘hot’ regions in the feature map. The entropy map is then normalized to $[0, 1]$. The density distribution of the normalized feature space X is estimated by diffusion-based Gaussian kernel density estimator [42] as

$$KDE^{diff}(X, Q_i; t) = \sum_{\sigma=-\infty}^{\infty} (\eta(X, 2\sigma + Q_i; t) + \eta(X, 2\sigma - Q_i; t)), X \in [0, 1] \quad (2)$$

where $Q_{i \in \{1, \dots, M\}}$ is one of M independent realizations; η denotes the Gaussian probability density at scale t defined as

$$\eta(X, Q_i; t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(X-Q_i)^2}{2t}} \quad (3)$$

2) Feature space clustering

Given Eq.(2) representing the feature space distribution, the next step is to obtain the cluster numbers and optimal thresholding levels. An affinity propagation (AP) based clustering method [43] is performed for image partition. This method was proposed to identify spatially diffuse and multifocal radiotracer uptake in PET images [43] which have similar appearance with our local entropy feature map. Unlike conventional thresholding methods or clustering methods, the initial topoRegions are generated based on the automatically estimated thresholding levels and cluster numbers without user intervention.

2.3.2. Fine-grained topography extraction

The initially extracted topoRegions provide general and global information of the input image. To obtain more detailed information of the target object, fine level topography is extracted from the target object (ROI) by performing local refinement until a stopping criterion is satisfied (as shown in Fig. 2). The theoretical basis for the stopping criterion is that: in a contour tree structure, the topology changes at the saddle point where two or more existing components are joined into a new component, or an existing component is split into two or more components. [27]

For each of the initially obtained saddle points, a *local refinement* is performed by a sweep through changing iso-values of this saddle point’s neighboring initial topoRegion contours. The sweeping is stopped when the following *stopping criterion* is satisfied: when locally increasing or decreasing the iso-values, the iso-contours are about to join or split at the

saddle point.

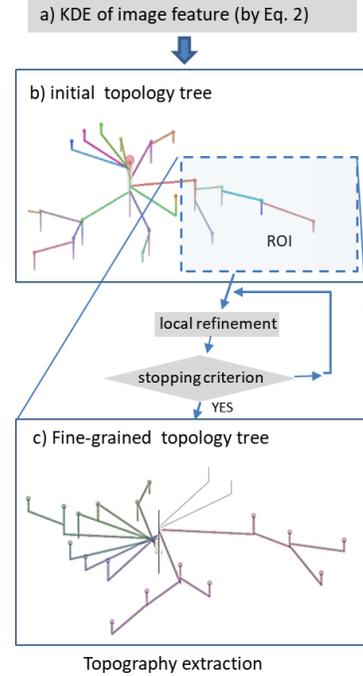


Fig. 1. Fine-grained topography extraction. Given a) the density distribution of the feature space estimated by *KDE*, b) initial topology tree is constructed and c) fine-grained topology is extracted by focusing on the ROI until the depth stopping criterion is satisfied.

By going through all the saddle points within the ROI, a set of fine-grained topoRegions are obtained based on the updated iso-values. When compared to using super-pixels as sub-regions in other graph models, the main advantage of extracting topoRegions is its automatic partition procedure while super-pixels require prior case-by-case parameter settings. Moreover, the initial and fine-grained topoRegions, which are composed of sets of iso-contours, inherit the topological information from the contour tree.

2.3.3. Topographic relations definition

Two topographic relations derived from a contour tree are defined to represent regional topology. If the outer boundary of a region r_a is a subset of the inner boundaries of r_b , then r_a and r_b are defined as having a nesting/inclusion relation, i.e. $r_a \subset r_b$. If two regions r_a, r_b are both included in r_c , $r_a \subset r_c, r_b \subset r_c, r_a \cap r_b = \emptyset$, then r_a, r_b are considered as having an exclusion relation.

2.4. Weighted Graph Model Construction

Given an image I and a set of topoRegions $R = R_0 \cup R_1$ where R_0 (R_1) denotes the initial level (fine-grained level) topoRegions, we construct a weighted undirected graph to associate the initial-to-fine level image topography.

Table 1: Symbols, abbreviations and descriptions used for Topo-Graph model			
Symbol	Description	Symbol	Description
I	Input image	G	Graph
V	Graph nodes	E	Graph edge
\mathbf{W}	Edge weights	Ψ	Graph optimization function
\mathbf{D}	Degree matrix	\mathbf{T}	Transition matrix
\mathbf{f}	Relevance vector	\mathbf{f}^R	Relevance scores of topoRegions
\mathbf{f}^I	Relevance scores of pixels	L	Label set
F	Foreground labels	B	Background labels
V_F	Nodes with foreground labels	V_B	Nodes with background labels
R	topoRegions set	R_0	Initial level topoRegion
N_{R_0}	Number of initial level topoRegions	R_1	Fine-grained topoRegion
E_{NB}	Nesting branches connecting initial level topoRegions	E_{GE}	Geodesic edge connecting region and pixel
φ_{ij}	Pairwise feature distances of the nodes connected by nesting branches	ϕ_{ij}	Pairwise feature distances of regions and pixels connected by geodesic edges
GM	Gradient magnitude	EMD	Earth Mover's Distance
d_G	Geodesic distance	KDE	Diffusion-based Gaussian kernel density estimator
Ψ_{inter}	Inter energy term	Ψ_{intra}	Intra energy term
Θ	Fitting constraint	λ	Label confidence parameter
μ_1	region level information confidence factor	μ_2	pixel level information confidence factor

2.4.1. Graph nodes and edges

Each graph node $v_i \in V$ corresponds to a topoRegion $r_i \in R$. According to the level of topoRegions and their topographic relations, two types of graph edges $E = (E_{NB}, E_{GE})$ are defined where E_{NB} denotes nesting branches and E_{GE} is the geodesic edge set. Nesting branches are defined to assist the long-range propagation of label information; the geodesic edges are designed to assist the local splitting of topoRegions and proper grouping of uncertain regions.

1) Nesting Branch

Given two nesting topoRegions $r_m, r_n \in R_0$ as nodes, a nesting branch (as shown in Fig. 3 (b)) is defined as:

$$\exists r_m \rightarrow v_i, r_n \rightarrow v_j, r_m \subset r_n : e_{ij} \in E_{NB} \quad (4)$$

where $r_m \rightarrow v_i$ denotes a node v_i corresponding to a topoRegion r_m ; $r_m \subset r_n$ denotes that a topoRegion r_m is included/nested in r_n .

2) Geodesic Edge

Given an initial level topoRegion $r_m \in R_0$ and a pixel I_k inside fine level topoRegion $r_n \in R_1$ as nodes, geodesic edges (Fig. 3 (c)) are defined as:

$$\exists r_m \rightarrow v_i, I_k \rightarrow v_j, I_k \in r_n, r_n \subset r_m : e_{ij} \in E_{GE} \quad (5)$$

where $I_k \in r_n, r_n \subset r_m$ denotes that node v_j is located inside a fine level topoRegion r_n which has a nesting relation with r_m . In terms of a tree structure, each topoRegion can be considered as connected to the regions of nesting relations with its ‘parent’ (as shown in Fig. 3 (c)).

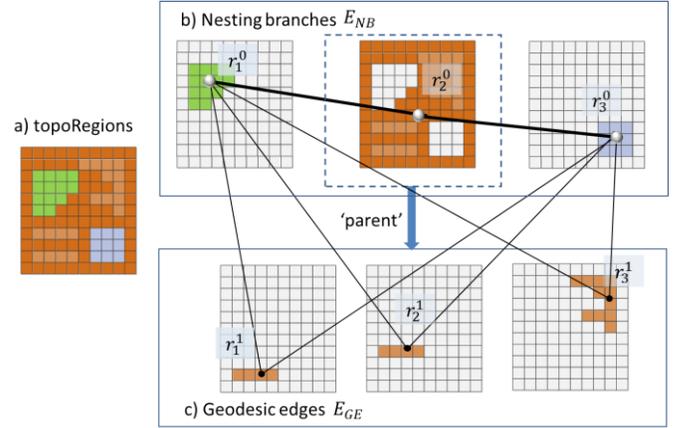


Fig. 2. Illustration of nesting branches and geodesic edges. (a) Each square represents a pixel index with different colors indicating various topoRegions. (b) Nesting initial topoRegions are connected by nesting branches. (c) Geodesic edges link fine-grained topoRegions and their nesting ‘parent’ initial level topoRegion.

2.4.2. Weighting functions

The edge weights reflecting the pairwise feature distances / affinities between topographically connected nodes are defined as:

$$w_{ij} = \begin{cases} \exp(-\beta \cdot \varphi_{ij}), & \text{if } e_{ij} \in E_{NB} \\ \exp(-\beta \cdot \phi_{ij}), & \text{if } e_{ij} \in E_{GE} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where β is free parameter and set to 60 by default; φ_{ij} and ϕ_{ij} denote the pairwise feature distances of the nodes connected by nesting branches and geodesic edges respectively.

The feature distance φ_{ij} between nodes of initial level topoRegions is defined with an adaptive appearance and boundary cue. The prior appearance cue is particularly useful when the background and foreground/object have dissimilar appearances. When the background has a similar appearance to the foreground, the boundary cue provides complementary information and contributes to the edge disconnection. The feature distance φ_{ij} is calculated as:

$$\varphi_{ij} = \alpha (g(v_i) - g(v_j))^2 + (1 - \alpha) \cdot \max_{i \in \bar{j}} \|GM_i\|^2 \quad (7)$$

where α is a factor balancing the appearance and boundary

terms defined by Eq. 8; GM_i denotes the gradient magnitude on the common boundary \bar{ij} of nodes v_i, v_j ; $g(v_i)$ is calculated as the probability that a node v_i fits the foreground model by:

$$g(v_i) = \frac{g(v_i | F)}{g(v_i | F) + g(v_i | B)} \quad (8)$$

where $g(v_i | F)$ and $g(v_i | B)$ denote the foreground and background likelihood respectively. α is determined adaptively with respect to the distance between $g(v_i | F)$ and $g(v_i | B)$ as [20]:

$$\alpha = \frac{1}{N_{R_0}} \sum_{i=1}^{N_{R_0}} \left| \frac{g(v_i | F) - g(v_i | B)}{g(v_i | F) + g(v_i | B)} \right| \quad (9)$$

where N_{R_0} denotes the number of initial topoRegions. According to factor α , the weight ϕ_{ij} is adaptively tuned to depend more on the appearance cue when the foreground and background have disparate appearances. When the appearance dissimilarity is indistinct, the value of α decreases and the boundary cue contributes more to the weight calculation. Given the overlapping intensity distributions of foreground and background objects, features such as the average intensity values would be insufficient to distinguish the target object and background tissues. Therefore, $g(v_i | F)$ ($g(v_i | B)$) is calculated as the Earth Mover's Distance (EMD) [44] between node v_i and foreground (background) seeds. The foreground likelihood $g(v_i | F)$ is determined by

$$g(v_i | F) = \sum_{k=1}^{|F|} EMD(v_i, v_k \in V_F) \quad (10)$$

where $EMD(v_i, v_k)$ denotes the EMD between nodes v_i, v_k .

$g(v_i | B)$ is defined as $g(v_i | B) = \sum_{k=1}^{|B|} EMD(v_i, v_k \in V_B)$ in a similar way.

The feature distance ϕ_{ij} between the nodes connected by geodesic edges is defined to be structure-sensitive. The geodesic distance is used for calculation as it can enhance and capture the thin boundary information [37, 45] when compared with Euclidean distance as shown in Fig. 4.

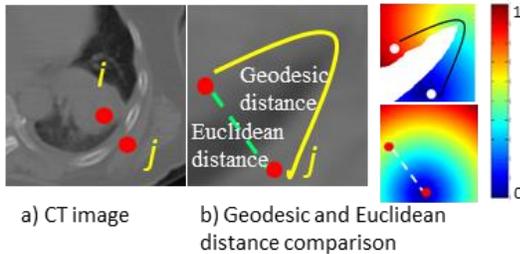


Fig. 4. Comparison of Euclidean and geodesic distances from points i to j . The geodesic distance (b) captures the thin boundary information giving a much larger distance than Euclidean distance (c) between points i and j .

ϕ_{ij} is calculated as:

$$\phi_{ij} = d_G(v_i, r_j(c)), r_j \in R_0 \quad (11)$$

where $d_G(x, y)$ denotes the geodesic distance between two points x, y and is calculated using an efficient method in [46]; $r_j(c)$ denotes the center of region r_j .

2.5. Energy Function and Graph Regulation

To solve the proposed graph model, the ranking score \mathbf{f} is obtained by solving the following function as:

$$\Psi(\mathbf{f}) = \Psi_{inter}(\mathbf{f}) + \Psi_{intra}(\mathbf{f}) + \lambda \Theta(\mathbf{f}, L) \quad (12)$$

where Ψ_{inter} correlates the inter affinities between connected nodes; Ψ_{intra} correlates the intra affinities between a topoRegion and a pixel, Θ is a constraint associating label information. The parameter λ reflects how much the segmentation results would like to trust the initial labels. $\mathbf{f} = [\mathbf{f}^R, \mathbf{f}^I] = [f_{iL_k}]_{|V| \times 2}$ is a vector with \mathbf{f}^R denoting the relevance between topoRegions, \mathbf{f}^I denoting the relevance between pixels, and $L_{k=1}$ for F , $L_{k=2}$ for B . Ψ_{inter} is defined as:

$$\Psi_{inter}(\mathbf{f}) = \sum_{e_{ij} \in E_{GE}} w_{ij} \|f_i - f_j\|^2 + \sum_{e_{ij} \in E_{NB}} w_{ij} \|f_i - f_j\|^2 \quad (13)$$

and Ψ_{intra} is defined as

$$\Psi_{intra}(\mathbf{f}) = \sum_{e_{ij} \in E_{GE}} \mu_1 \sum_j w_{ij} \|f_i - \bar{f}_i\|^2 + \sum_{e_{ij} \in E_{NB}} \mu_2 \sum_j w_{ij} \|f_i - \bar{f}_i\|^2 \quad (14)$$

where μ_1 is to control the confidence of using pixel level similarities to classify regions; μ_2 is to control the confidence of using region level similarities to classify pixels, \bar{f}_i is the average relevance score of all the nodes connected with v_i . Θ associates labels and is defined as

$$\Theta(\mathbf{f}, L) = \sum_{e_{ij} \in E_{GE}} \sum_j w_{ij} \|f_i - f_j^*\|^2 + \sum_{e_{ij} \in E_{NB}} \sum_j w_{ij} \|f_i - f_j^*\|^2, \begin{cases} f_i^* = 1, \text{ if } v_i \in V_F \\ f_i^* = 0, \text{ if } v_i \in V_B \end{cases} \quad (15)$$

By differentiating Eq.13 with respect to $\mathbf{f}^R, \mathbf{f}^I$ respectively as below:

$$\frac{\partial \mathbf{f}}{\partial \mathbf{f}^R} = \mathbf{f}^R - \mathbf{T}^{NB} \mathbf{f}^R + \Lambda(\mathbf{f}^R - \mathbf{f}^{*R}) + \mu_1(\mathbf{f}^R - \mathbf{T}^{GE} \mathbf{f}^I) \quad (16)$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{f}^I} = \mathbf{f}^I - \mathbf{T}^I \mathbf{f}^I + \Lambda(\mathbf{f}^I - \mathbf{f}^{*I}) + \mu_2(\mathbf{f}^I - \mathbf{T}^{GE} \mathbf{f}^R) \quad (17)$$

where $\mathbf{T} = \mathbf{D}^{-1} \mathbf{W}$ is the transition matrix; $\mathbf{D} = \text{diag}(\sum_j w_{ij})$ is the degree matrix; \mathbf{T}^{NB} is the transition matrix along nesting branches; \mathbf{T}^{GE} is the transition matrix along geodesic edges, \mathbf{T}^I is the matrix between pixels. By setting the derivative of Eq.16 and Eq.17 to 0, the two equations can be jointly transformed to

$$(\mathbf{I} - (\mathbf{I} - \Omega)\mathbf{I}) \mathbf{f} = \Omega \mathbf{f}^* \quad (18)$$

where $\Omega = \text{diag}(\Lambda / ((1 + \mu_1)\mathbf{I} + \Lambda), \Lambda / ((1 + \mu_2)\mathbf{I} + \Lambda))$,

$$\mathbf{\Pi} = \begin{bmatrix} \mathbf{T}^I / (1 + \mu_1) & \mu_1 \mathbf{T}^{GE} / (1 + \mu_1) \\ \mu_2 \mathbf{T}^{GE} / (1 + \mu_2) & \mathbf{T}^{NB} / (1 + \mu_2) \end{bmatrix}. \text{ The final decision to}$$

assign a label L_1 or L_2 to pixel i is determined by

$$l_i = \arg \max_{L_k} f_{il_k}^I \quad (19)$$

2.6. Image datasets

The proposed model was applied to lung tumor segmentation from low contrast CT images, liver segmentation from enhanced CT images, and target object segmentation from breast and abdominal ultrasound images. The CT images were used to validate the performance of target object separation when the foreground and background objects share similar intensity distributions and the boundaries are not easily discernible. The ultrasound images were used to evaluate the ability of identifying uncertain/blurred boundaries with speckle noise.

2.6.1. Clinical low contrast NSCLC CT studies

We studied 47 non-small cell lung cancer CT studies including 27 patients from Shandong Cancer Hospital (SCH), China and 20 patients from Royal Prince Alfred (RPA) Hospital, Australia. The CT scans were done as part of a PET-CT study using ^{18}F FDG. The SCH patients were scanned on a Discovery LS PET-CT system (GE Healthcare, Milwaukee, WI, USA). The CT images were reconstructed using a matrix of 512×512 pixels. The CT voxel size was $1.17 \text{ mm} \times 1.17 \text{ mm} \times 5 \text{ mm}$. The manual tumor delineations were performed by one experienced radiation oncologist on CT with PET images as reference. The 20 RPA scans were carried out on a Biograph TrueV 64 slice PET-CT scanner (Siemens Medical Solutions, Hoffman Estates, IL, USA) where the CT data were reconstructed using a matrix of 512×512 pixels with voxel size of $0.98 \text{ mm} \times 0.98 \text{ mm} \times 3 \text{ mm}$. The manual tumor delineations were performed by one senior clinical expert. The manual delineations were used as the reference, ground truth (GT), for segmentation accuracy comparison.

2.6.2. Public liver dataset from high contrast CT

These images were obtained from the 3D Image Reconstruction for Comparison of Algorithm Database¹ (3D-IRCADb). There were 10 enhanced CT images of liver with no more than two tumors. The CT images were reconstructed with matrix of 512×512 pixels and inter-slice distances varying from 1.0 to 2.4 mm. The manual segmentations were done by multiple experienced radiologists.

2.6.3. Public breast and abdominal ultrasound images

We obtained fifty breast and abdominal ultrasound images [33]² including, scan-converted, monochromatic, B-mode and elastography ultrasound acquisitions. The segmentation results and 14 sets of manual delineations were also provided to compare accuracy. In our experiments, the manual segmentation results named by “subject 8” and “subject 10”

were used for segmentation accuracy evaluation because they were listed for qualitative and quantitative analysis [33]. These two manual delineations are referred to as GT-1 and GT-2.

2.7. Evaluation methods

To assess the accuracy of the proposed method, we calculated the spatial overlap and shape dissimilarity between the segmentation results and manual delineations by Dice’s Similarity Coefficient (DSC) and the Hausdorff distance (HD). DSC was defined as:

$$DSC(Vol_1, Vol_2) = \frac{2|Vol_1 \cap Vol_2|}{|Vol_1| + |Vol_2|} \quad (20)$$

where Vol_1 is the segmented volume, and Vol_2 is the GT volume. A higher DSC value indicates greater spatial overlap between the segmentation and ground truth.

HD was defined as:

$$HD(Vol_1, Vol_2) = \max\{\sup_{i \in Suf_1} \inf_{j \in Suf_2} d_{Eu}(i, j), \sup_{j \in Suf_2} \inf_{i \in Suf_1} d_{Eu}(i, j)\} \quad (21)$$

where Suf_1 and Suf_2 denote the boundary of the segmented volume and the GT volume, sup represents the supremum and inf the infimum; d_{Eu} is the Euclidean distance between point i and j . For our cases, the HD measurement can be implemented by using the maximum and minimum surface distances between Vol_1 and Vol_2 to replace supremum and infimum. [47] A low HD value indicates high segmentation accuracy.

A paired t-test with two-tailed distribution was performed to evaluate any statistical significance of performance improvement.

2.8. Comparison methods

We compared our proposed model, using the 3 datasets outlined above, to 5 other methods: 1) GC [48]; 2) RW [11], 3) graph model [24] where each superpixel / node is connected to its neighboring nodes as well as the nodes sharing the common boundaries with these neighboring nodes (referred to as radial connection (RSP)); 4) graph model with multi-level superpixel and pixel connection (NHLIS) [49]. For the NSCLC datasets we also compared our model to our previous topology polymorphism graph model (P-Graph) [21]. P-Graph was only used for the NSCLC datasets because the P-Graph model requires PET and CT images together for segmentation. For the ultrasound datasets, the segmentation from a box algorithm (S-Box) [33] was used for comparison where the authors provided the segmentation results. We refer to our proposed model as PM from here.

2.9. Implementation, initialization and parameter settings

Our algorithm was implemented with MATLAB R2017a on a PC with 3.50GHz Intel(R) Core(TM) i7-4770K CPU and 16.0GB memory, running a 64-bit Windows operating system. The model was implemented on 2D slices. The graph model can, however, be extended to a 3D graph by changing the nodes

¹ 3D-IRCADb data <http://www.ircad.fr/research/3d-ircadb-01/>

² The ultrasound data and segmentation results are from http://cns.bu.edu/~lgrady/box_study/box_study.html

and edges from a neighboring 4-connected 2D lattice to a 3D lattice with 6, 18 or 26 adjacent connections. The number of nodes and edges in the graph, however, will be increased correspondingly and hence it will require substantially larger memory resources and a longer computation time.

The AP based method for feature space clustering was implemented using the published code [43]. The graph and the ranking function were constructed and solved using the Graph Analysis Toolbox³. RSP and NHLIS were implemented using the published code by the authors. The geodesic distance was calculated by fast marching toolbox⁴. EMD is implemented using fast EMD⁵.

A “one-touch” user-input is required to segment the target object. The ROI is obtained by a rectangle box which is drawn outside the object with the user-input “one-touch” foreground seed as the centroid. We set the size of background box as 40×40 for the NSCLC cases as this size allowed all the lung tumors to be enclosed. We set the background box as 300×300 for the liver volume and the foreground seeds with a rectangle of 15×20. For the ultrasound datasets, the boundaries of the images as provided in the dataset comprised the background boxes [33].

For the free parameters in this paper, λ needs to be assigned a high value (10^5) to associate prior knowledge and ensure that the segmentation results would not change from the foreground and background seeds. μ_1, μ_2 were set as 0.002 and 0.2 which are the same as NHLIS for comparison. In the weighting function, β was set as 60 which is generally used by graph models.

3. Results and discussion

3.1. Comparison and evaluation of NSCLC CT images

The segmentation accuracy for the lung tumor cases is shown in Tables 2 and 3. Our PM achieved the best results followed by P-graph and RSP; GC and RW with sole pixel level information had lower segmentation accuracy than the graph models using regional or combined regional and pixel level information.

	Mean±SD	p-value
RW	0.677±0.156	1.07E-08
GC	0.603±0.142	1.11E-08
NHLIS	0.635±0.110	2.99E-09
RSP	0.708±0.097	3.04E-09
P-Graph	0.842±0.051	0.00954
PM	0.878±0.046	-

	Mean±SD	p-value
RW	13.212±11.352	8.7E-05
GC	16.102±10.213	5.4E-06
NHLIS	17.523±10.265	8.74E-07
RSP	12.154±6.564	3.49E-07
P-Graph	6.341±5.231	0.00521
PM	5.741±3.226	-

The segmentation result of a case with indistinct tumor boundaries is shown in Fig. 5. In this example, the tumor was located in the left lower lobe of the lung, adjacent to the pericardium and descending thoracic aorta and abutting the posteromedial pleura. Note that there is similar intensity in the tumor, the pericardium, heart, aorta, posteromedial pleura and the chest wall and the tumor boundaries are difficult to discern on CT. Our PM gave the best tumor delineation from the surrounding structures with a DSC of 0.886 and HD of 3.153 (mm). The second best segmentation was achieved by P-Graph with a DSC of 0.865 and HD of 5.652 (mm). The GC, RSP and NHLIS methods all showed leakage into the heart. RW, RSP and NHLIS all included non-tumor regions.

3.2. Comparison and evaluation of liver CT images

The segmentation results for the ten liver cases are shown in Tables 4 and 5. Our PM obtained consistently better spatial volume overlap similarity. As shown in the two cases (3Dircadb03 and 3Dircadb06) in Fig. 6, although RSP and NHLIS failed to delineate the weak boundary between the liver and the chest wall, they were able to propagate long range foreground and background label information. Considering only pixel level information, RW and GC failed to delineate the complete liver volume. Our PM was able to delineate the whole object and capture detailed boundary information.

	Mean±SD	p-value
RW	0.768±0.072	2.25E-09
GC	0.677±0.108	3.08E-09
NHLIS	0.746±0.068	6.11E-07
RSP	0.743±0.075	5.95E-08
PM	0.871±0.034	-

	Mean±SD	p-value
RW	18.49±6.230	5.13E-06
GC	20.23±6.320	6.09E-06
NHLIS	14.63±5.478	4.21E-06
RSP	13.268±4.404	6.83E-05
PM	6.989±2.033	-

3.3. Comparison and evaluation of ultrasound images

Across the 50 images, the object in Case 32 had two separate components and the segmentation result of S-box had only one

³ Grady L The graph analysis toolbox: image processing on arbitrary graphs 2003. Boston University, Boston, MA, Tech. Rep. TR-03-021

⁴ Fast marching toolbox: <http://au.mathworks.com/matlabcentral/fileexchange/6110-toolbox-fast-marching>

⁵ <http://www.ariel.ac.il/sites/ofirpele/FastEMD/code/>

component, and so this case was excluded from the analysis. For the remaining 49 cases, our PM obtained better segmentation results than S-Box as shown in Tables 6 and 7. Student's t-test performed between our results and S-box showed a statistical difference (p -values = $1.17E-05$ and $4.49E-06$ with respect to GT-1 and GT-2).

	GT-1		GT-2	
	Mean± SD	p-value	Mean± SD	p-value
RW	0.533±0.401	4.58E-06	0.546±0.400	6.54E-07
RSP	0.392±0.341	6.87E-08	0.384±0.344	4.96E-07
NHLIS	0.684±0.292	5.35E-07	0.664±0.295	3.47E-06
S-Box	0.866±0.065	1.17E-05	0.859±0.070	4.49E-06
PM	0.891±0.080	-	0.888±0.084	-

*Note: Case 32 was removed because it contained two objects

The case-by-case segmentation results of PM and S-Box with respect to DSC of the 49 cases are shown in Fig. 8. The lowest segmentation accuracy was seen for Case 37. As shown for Case 37 in Fig. 7, RW failed to delineate the whole target object due to high frequency intensity changes inside the object. The results of RSP and NHLIS were sensitive to the intensity distributions and variations of the target object.

	GT-1		GT-2	
	Mean± SD	p-value	Mean± SD	p-value
RW	18.256±12.653	5.77E-06	19.578±12.765	4.96E-07
RSP	21.563±15.648	6.59E-07	23.564±17.689	5.37E-07
NHLIS	12.563±10.247	7.35E-07	15.365±11.547	5.03E-07
S-Box	7.812±4.231	6.05E-06	6.548±3.546	7.25E-06
PM	5.311±3.022	-	4.882±2.865	-

*Note: Case 32 was removed because it contained two objects

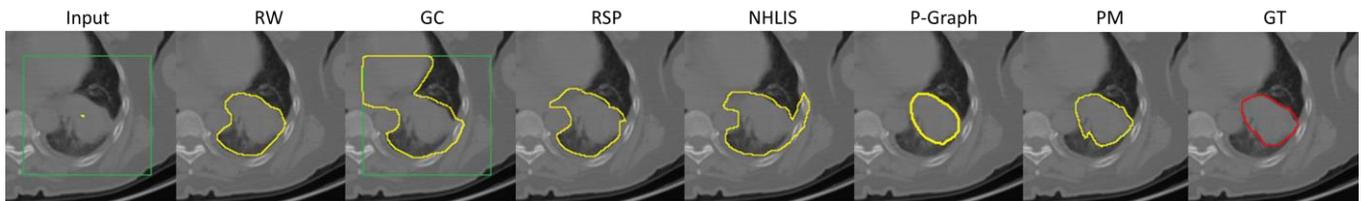


Fig. 5. Cropped transaxial CT images with lung windows show tumor delineation results of a NSCLC CT case; segmentation results are shown in yellow and the ground truth is shown in red.

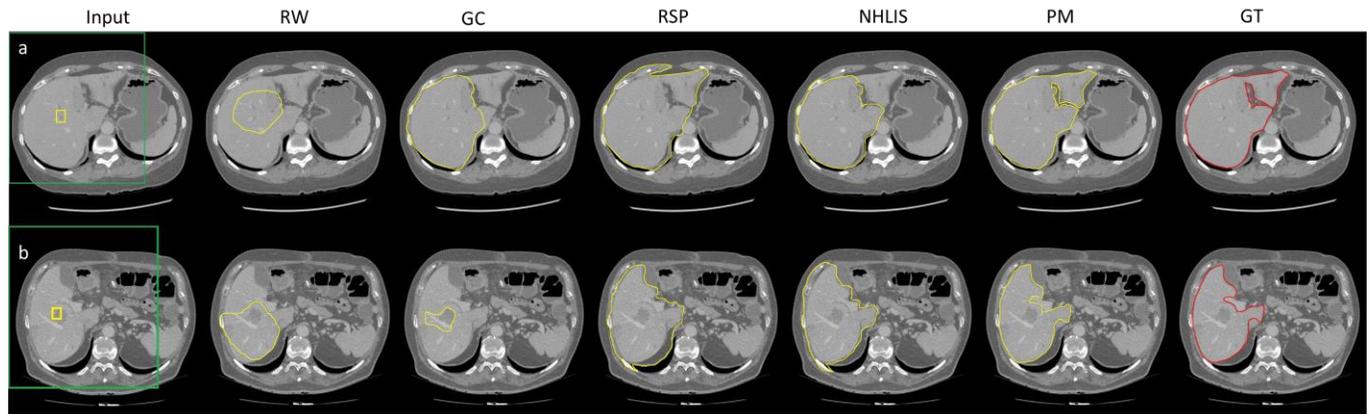


Fig. 6. Segmentation results of two liver CT cases (3Dircadb03 and 3Dircadb06); segmentation results are shown in yellow and the ground truth is shown in red.

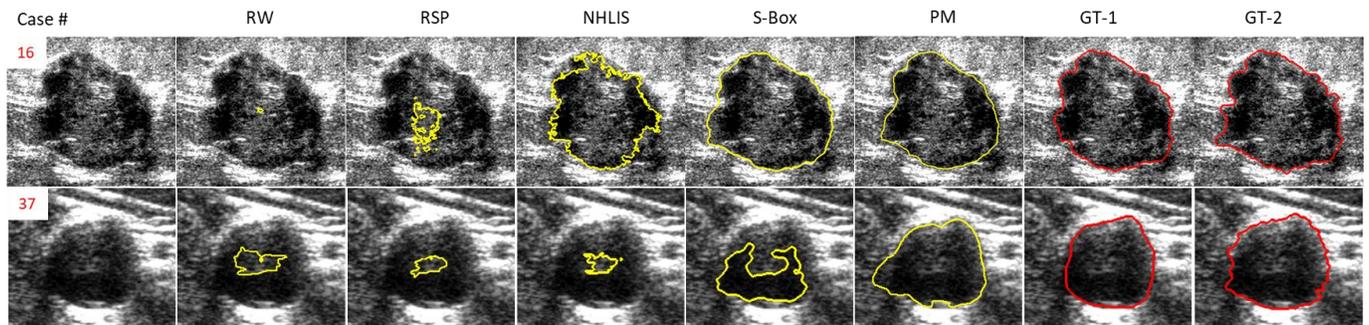


Fig. 7. Segmentation results of Case 16 and 37; the segmentation results are shown in yellow and the manual delineations of GT-1 and GT-2 are shown in red.

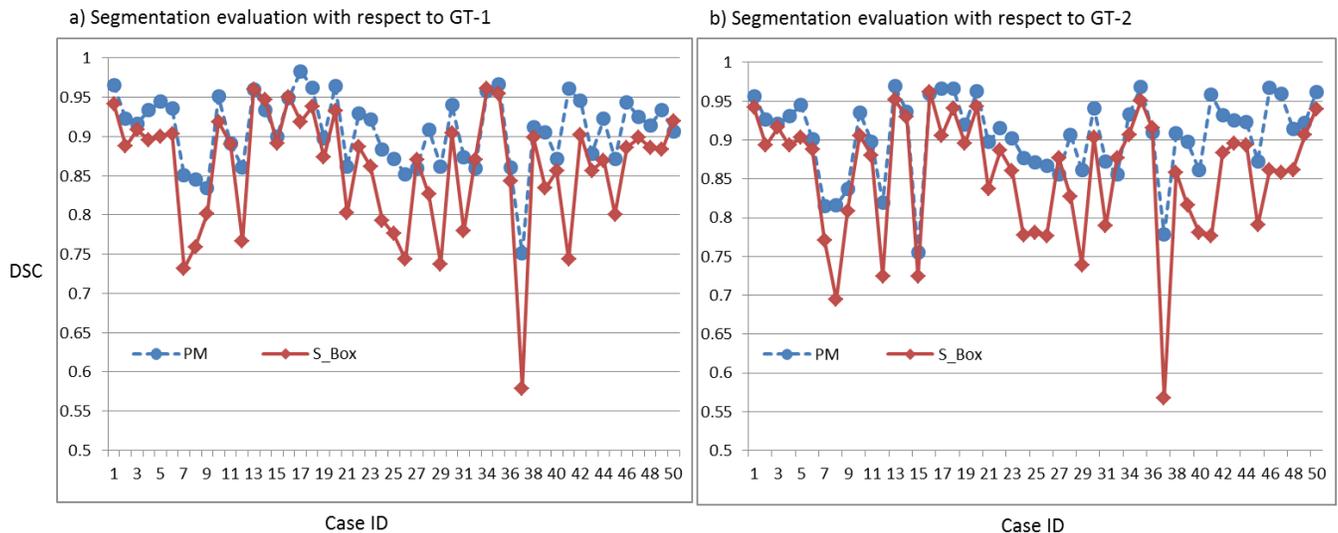


Fig. 8. Case-by-case DSC comparisons between PM and S-Box with respect to GT-1 (a) and GT-2 (b).

3.4. Performance evaluation on parameter settings

3.4.1. Confidence parameters

λ , μ_1 and μ_2 are all confidence parameters for initialization, regional and pixel levels, respectively. We firstly investigated the sensitivity to the initialization confidence parameter λ . The segmentation results with different λ values are plotted in Fig. 9. As shown in Fig. 9, when $\lambda \geq 10^4$, the segmentation accuracy was almost the same and the

differences were not appreciable. Not surprisingly, the segmentation accuracy decreased with decreasing λ when $\lambda \leq 10^4$. This is because a smaller λ value indicates less confidence on the pre-defined seeds with foreground and background labels. Therefore, a smaller λ value may result in a final segmentation where each seed is less likely to be assigned to its initial label. Based on these experimental results, λ can be fixed at 10^5 to generate robust segmentations.

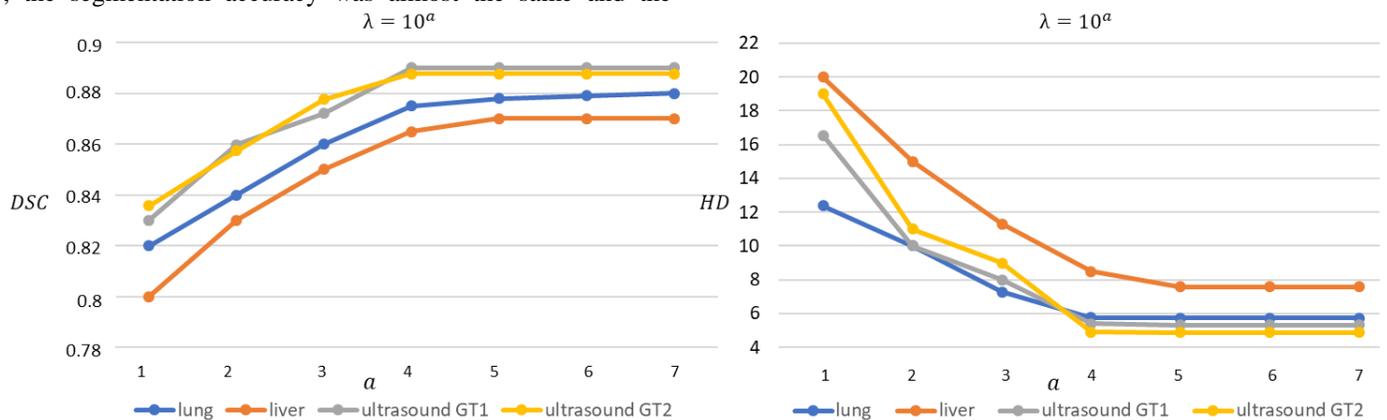


Fig. 9. Effects of λ on segmentation accuracy.

To evaluate the sensitivity to regional and pixel level confidence parameters μ_1, μ_2 , we fixed λ at 10^5 and tested our model with different μ_1 and μ_2 values ($2 \times 10^{-6} \leq \mu_1 \leq 2 \times 10^{-2}$, $2 \times 10^{-5} \leq \mu_2 \leq 2 \times 10^{-1}$). As shown by the segmentation results in Fig. 10, there was no significant difference when $\mu_1 \geq 2 \times 10^{-5}$. The segmentation results were less sensitive to μ_2 when compared with μ_1 . Generally, when we decreased the value of μ_1 and increased the value of μ_2 , the segmentation accuracy decreased. Given a fixed μ_2 value,

higher μ_1 values yield better segmentation results. This finding also emphasizes the contributions of regional information in object separation, especially for the noisy images in the ultrasound cases and for large objects such as the liver.

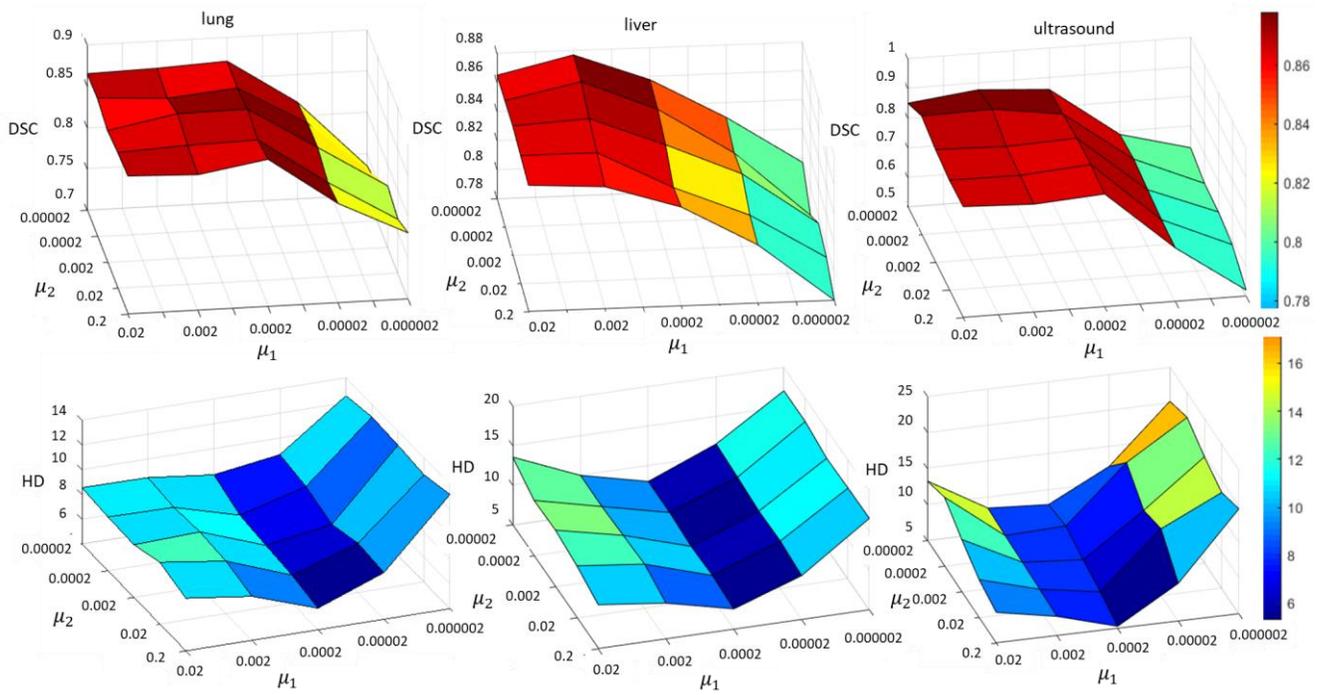


Fig. 10. Effects of μ_1, μ_2 on the segmetnation accuracy.

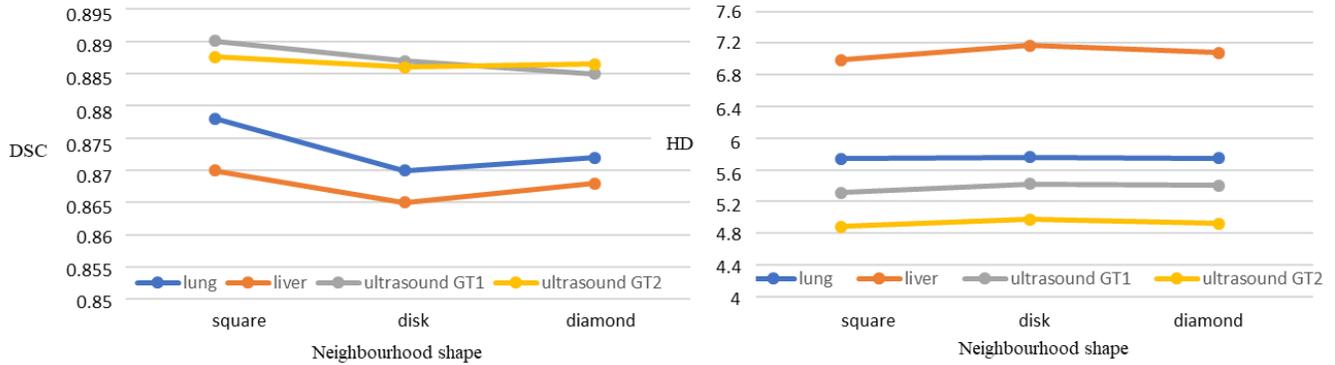


Fig. 11. Effects of neighbourhood shape on segmentation accuracy.

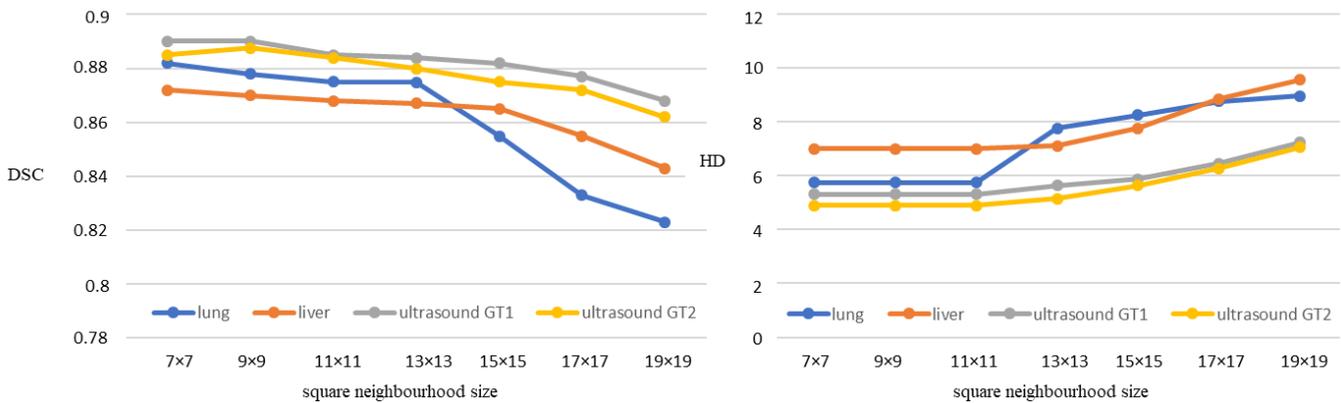


Fig. 12. Effects of neighbourhood size on segmentation accuracy

3.4.2. Neighborhood in feature space calculation

When calculating entropy feature space, the neighborhood was defined as a 9×9 square by default. The sensitivity to varying neighborhood shape (square, disk and diamond) and

size is shown in Fig. 11 and Fig. 12. To evaluate the sensitivity of neighborhood shape selection, the sizes of square and diamond were set as 9×9 and the radius of the disk was set as 4. In Fig. 11, the shape selection did not have significant impact on the segmentation results (p-value

<0.05). The variations of segmentation accuracy were within 0.01 in terms of DSC and 0.02 mm in terms of HD.

Given a square neighborhood, the segmentation results with respect to various neighborhood sizes were plotted in Fig. 12. The increment of neighborhood size had the smallest impact on the segmentation of ultrasound images. For the lung tumor cases, when the neighborhood size was greater than 13×13 , the difference was statistically significant ($p > 0.05$). For the liver cases, when the neighborhood size was set $> 15 \times 15$, the segmentation accuracy variations were statistically significant ($p > 0.05$). In summary, the selection of large neighborhood sizes may impact the segmentation results of objects with varying sizes. From our results, the neighborhood size can be set as 9×9 by default and this is widely accepted in entropy calculations.

3.4.3. Initialization

In regard to the initialization, we tested the segmentation results on the lung tumor and liver cases using different background box sizes. For lung tumor segmentation, our segmentation results were not significantly different ($p > 0.05$) when the size was set as 40×40 or 50×50 . Bigger background boxes may fail in the tumor separation when the tumor is attached to the heart. For liver cases, larger foreground box sizes resulted in improved segmentation results for all the methods due to the increased foreground confident pixels/regions.

3.5. Execution time and memory consumption

We measured the execution time and memory consumption of our PM and the five comparative graph models. We calculated the execution, the total amount of memory allocated within a model and any functions it called on (allocated memory), and the maximum amount of memory in use at any one time during the execution of a model (peak memory) for each slice. The average results over all the cases are given in Table 8. For all 6 graph models, the largest portions of the execution time and memory use were required for the weight calculations. As shown in Table 8, not surprisingly, graphs with both regional and pixel level nodes required more memory space and longer execution times. NHLIS had the longest execution time and largest memory consumption because of the full pixel level and regional level nodes connections. Our model required larger memory space and a longer execution time than RW, GC and P-Graph because of the incorporation of regional level nodes.

	RW	GC	RSP	NHLIS	P-Graph	PM
execution time	0.24	0.57	1.62	2.70*	1.12	1.33
allocated memory	24.63	46.83	102.65	197.60	106.03	111.76
peak memory	8.22	10.24	32.84	32.84	16.42	20.05

* excluding the superpixel generation time

3.6. Discussion

Our PM delineated the target object boundaries when the object and background shared similar or overlapping intensity distributions and where the boundaries were uncertain or not easily discernible. We explain this finding by the topographic nesting branches providing our model with the ability to separate adjacent different structures with similar intensities. For instance, in the case shown in Fig. 5, the topoRegions indicating tumor and the heart have an exclusive relation. Thus there is no direct connection between the two regions *a* and *c* in our graph model (as shown in Fig. 13 (c)). In comparison, the tumor and heart regions *a* and *c* are directly connected according to the adjacent or radial connection defined in RSP and NHLIS (as shown in Fig. 13 (a) and (b)). Therefore, when the intensity distributions of the target object and background are similar, it is difficult for these graph models to achieve the separation or identify the boundary.

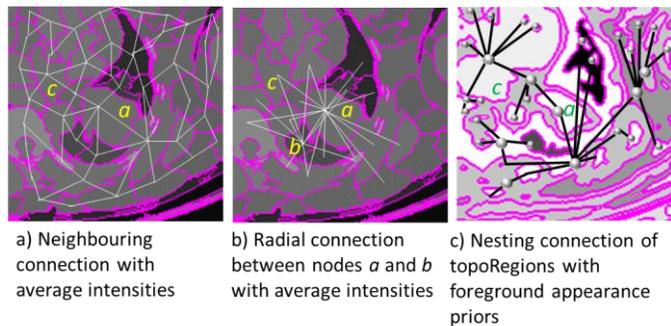


Fig. 13. Cropped transaxial slices with three types of edges. (a) and (b) show the neighboring and radial connections (white lines) of oversegmentations generated by mean-shift algorithm. c) The proposed nesting branches (black lines) connecting initial topoRegions.

The extracted initial topography from the whole image domain is essential and important in our PM due to: a) Noise and artefacts in medical images, the direct adoption of a conventional contour tree, even within the ROI, may generate giant tree structures with redundant information that makes data analysis impractical [30]. The proposed initial topography extraction method optimized the tree size for data understanding while maintaining the essential image structure. b) The initial topography reflects the landscape of the whole image domain. The derived nesting branches enable

long-range propagation and association of label information at the global level within the whole image that serves as a global guidance to assist the separation and merging of regions/pixels at fine levels. These contributions are shown in the regional/pixel level confidence parameter settings where in Fig. 10, given a fixed pixel level confidence, the greater regional confidence, the better the segmentation results. Moreover, those models that focused solely on fine level information without global association and propagation of label information, may fail to delineate the whole target object especially for noisy ultrasound images (as shown by Fig. 7) and where there are large objects such as the liver (as shown by Fig. 6).

Our PM obtained complete target object segmentation using the adaptive appearance and boundary similarity estimation in the weights calculation of the nesting branches. Both NHLIS and RSP, incorporate the regional information in edge weights by calculating the average intensity values in each region. This regional information provided separation cues in general images. For medical images when the object and background have similar or overlapping intensity distributions, the average intensity values from each representative region do not allow these models to achieve target object separation (see regions *a* and *c* in Fig. 13(a) and (b)). Further, the geodesic edges in our model made it possible to obtain the final boundary delineation with detailed structural information. The other graph models, without the geodesic constraint, failed to identify detailed boundary information (see the lung tumor and liver cases in Fig. 5 and 6). In our PM, the edge weight was moderated by the geodesic distances and it takes more energy for the foreground label information to reach background when crossing the gap than in Euclidean space. This also explains why the graph models such as NHLIS were unable to obtain accurate boundary delineation although they incorporated the combined regional and pixel level information as well.

4. Conclusions

We present a new graph model for target object segmentation and boundary delineation in anatomical images. Our model incorporates topographic relations of multilevel topoRegions by constructing nesting branches and geodesic edges. The evaluation we conducted on cases of lung tumors in low contrast chest CT volumes, liver CT cases and ultrasound images show that our graph model improved segmentation accuracy, in particular, for cases where there are overlapping intensity distributions and uncertain boundaries.

Conflict of interest

The authors do not have any financial and personal relationships with other people or organizations that could

inappropriately influence (bias) their work.

References

- [1] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61-79, 1997.
- [2] B. R. Lee, A. B. Hamza, and H. Krim, "An active contour model for image segmentation: A variational perspective," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002, vol. 2, pp. 1585-1588.
- [3] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische Mathematik*, journal article vol. 66, no. 1, pp. 1-31, December 01 1993.
- [4] C. Li, X. Wang *et al.*, "A Likelihood and Local Constraint Level Set Model for Liver Tumor Segmentation from CT Volumes," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2967-2977, 2013.
- [5] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001.
- [6] C. Li, X. Wang *et al.*, "Robust Model for Segmenting Images With/Without Intensity Inhomogeneities," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3296-3309, 2013.
- [7] L. A. Vese and T. F. Chan, "A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model," *International Journal of Computer Vision*, journal article vol. 50, no. 3, pp. 271-293, December 01 2002.
- [8] B. Peng, L. Zhang, and D. Zhang, "A survey of graph theoretical approaches to image segmentation," *Pattern Recognition*, vol. 46, no. 3, pp. 1020-1038, 2013.
- [9] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: theory and its application to image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1101-1113, 1993.
- [10] J. Carreira and C. Sminchisescu, "Automatic object segmentation using constrained parametric min-cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, pp. 1312-1328, 2012.
- [11] L. Grady, "Random walks for image segmentation," *IEEE Trans Pattern Anal Mach Intell*, vol. 28, no. 11, pp. 1768-1783, 2006.
- [12] Q. Song, J. Bai *et al.*, "Optimal co-segmentation of tumor in PET-CT images with context information," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1685-1697, 2013.
- [13] D. Han, J. Bayouth *et al.*, "Globally optimal tumor segmentation in PET-CT images: a graph-based co-segmentation method," *Information processing in medical imaging*, vol. 22, pp. 245-56, 2011.
- [14] U. Bagci, J. K. Udupa *et al.*, "Joint segmentation of anatomical and functional images: Applications in quantification of lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT images," *Medical Image Analysis*, vol. 17, no. 8, pp. 929-945, 2013.
- [15] L. Grady, T. Schiwietz, S. Aharon, and M. Westermann, "Random walks for interactive organ segmentation in two and three dimensions: Implementation and validation," in *Medical Image Computing and Computer-Assisted Intervention*, 2005, vol. 3750, pp. 773-780.
- [16] M. Chen, E. Helm, N. Joshi, and S. M. Brady, "Random walk-based automated segmentation for the prognosis of malignant pleural mesothelioma," in *International Symposium on Biomedical Imaging*, 2011, pp. 1978-1981.
- [17] X. Chen, J. K. Udupa *et al.*, "Medical image segmentation by combining graph cuts and oriented active appearance models," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2035-2046, 2012.
- [18] D. Han, J. Bayouth *et al.*, "Globally optimal tumor segmentation in PET-CT images: A graph-based co-segmentation method," in *Information Processing in Medical Imaging*, 2011, pp. 245-256.

- [19] T. Kim, K. Lee, and S. Lee, "Generative image segmentation using random walks with restart," in *European Conference on Computer Vision*, 2008, pp. 264-275.
- [20] W. Yang, J. Cai, J. Zheng, and J. Luo, "User-friendly interactive image segmentation through unified combinatorial user inputs," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2470-2479, 2010.
- [21] H. Cui, X. Wang *et al.*, "Topology polymorphism graph for lung tumor segmentation in PET-CT images," *Physics in Medicine Biology*, vol. 60, no. 12, pp. 4893-4914, 2015.
- [22] H. Cui, X. Wang *et al.*, "Primary lung tumor segmentation from PET-CT volumes with spatial-topological constraint," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 1, pp. 19-29, 2016.
- [23] T. Kim, K. Lee, and S. Lee, "Learning full pairwise affinities for spectral segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 7, pp. 1690-1703, 2013.
- [24] C. Yang, L. Zhang *et al.*, "Saliency detection via graph-based manifold ranking," in *Computer Vision and Pattern Recognition*, 2013, pp. 3166-3173.
- [25] J. Wang, H. Jiang *et al.*, "Regularized tree partitioning and its application to unsupervised image segmentation," *IEEE Transactions on Image Processing*, 2014.
- [26] H. Cui, X. Wang *et al.*, "Topology constraint graph-based model for non-small-cell lung tumor segmentation from PET volumes," in *IEEE International Symposium on Biomedical Imaging*, 2014, pp. 1243-1246.
- [27] C. Correa, P. Lindstrom, and P.-T. Bremer, "Topological spines: A structure-preserving visual representation of scalar fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 1842-1851, 2011.
- [28] C. Correa and P. Lindstrom, "Towards robust topology of sparsely sampled data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 1852-1861, 2011.
- [29] J. Zhou, *Gaining Insights Into Volumetric Data Visualization (A Semi-Automatic Transfer Function Generation Approach Using Contour Tree Analyses)*. Germany: LAP LAMBERT Academic Publishing, 2012.
- [30] J. Zhou, C. Xiao, and M. Takatsuka, "A multi-dimensional importance metric for contour tree simplification," *Journal of Visualization*, vol. 16, no. 4, pp. 341-349, 2013.
- [31] D. Aydogan and J. Hyttinen, "Contour tree connectivity of binary images from algebraic graph theory," in *IEEE International Conference on Image Processing*, 2013, pp. 3054-3058.
- [32] H. Cui, X. Wang *et al.*, "Improved segmentation accuracy for thoracic PET-CT in patients with NSCLC using a multi-graph model (MGM)," *Journal of Nuclear Medicine*, vol. 56, no. supplement 3, p. 2527, 2015.
- [33] L. Grady, M. Jolly, and A. Seitz, "Segmentation from a Box," in *International Conference on Computer Vision*, 2011, pp. 1-8.
- [34] A. B. Hamza and H. Krim, "Geodesic matching of triangulated surfaces," *IEEE Transactions on Image Processing*, vol. 15, no. 8, pp. 2249-2258, 2006.
- [35] A. B. Hamza and H. Krim, "Probabilistic shape descriptor for triangulated surfaces," in *IEEE International Conference on Image Processing 2005*, 2005, vol. 1, pp. I-1041-4.
- [36] D. Aouada, S. Feng, and H. Krim, "Statistical Analysis of the Global Geodesic Function for 3D Object Classification," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007, pp. 645-648.
- [37] M. Liu, O. Tuzel, and Y. Taguchi, "Joint geodesic upsampling of depth images," presented at the Computer Vision and Pattern Recognition, 2013.
- [38] P. Wang, G. Zeng *et al.*, "Structure-Sensitive Superpixels via Geodesic Distance," *International Journal of Computer Vision*, vol. 103, no. 1, pp. 1-21, 2013.
- [39] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power Watershed: A Unifying Graph-Based Optimization Framework," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 7, pp. 1384-1399, 2011.
- [40] J. Zhou, X. Wang *et al.*, "Topology-aware illumination design for volume rendering," *BMC Bioinformatics*, vol. 17, no. 1, pp. 309-326, 2016.
- [41] H. Carr, J. Snoeyink, and M. van de Panne, "Flexible isosurfaces: simplifying and displaying scalar topology using the contour tree," *Computational Geometry*, vol. 43, no. 1, pp. 42-58, 2010.
- [42] G. J. Botev, Z. and Kroese.D, "Kernel density estimation via diffusion," *The Annals of Statistics*, vol. 38, no. 5, pp. 2916-2957, 2010.
- [43] B. Foster, U. Bagci *et al.*, "Segmentation of PET Images for computer-aided Functional quantification of tuberculosis in small animal models," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 711-724, 2014.
- [44] O. Pele and M. Werman, "Fast and robust earth mover's distances," presented at the International Conference on Computer Vision, 2009.
- [45] B. Price, B. Morse, and S. Cohen, "Geodesic graph cut for interactive image segmentation," presented at the Computer Vision and Pattern Recognition, 2010.
- [46] B. Xue and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," in *International Conference on Computer Vision*, 2007, pp. 1-8.
- [47] B. H. Menze, A. Jakab *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024, 2015.
- [48] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images " in *International Conference on Computer Vision*, 2001, pp. 105-112.
- [49] T. H. Kim, K. M. Lee, and S. U. Lee, "Nonparametric higher-order learning for interactive segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 3201-3208.