Automatic Pathological Lung Segmentation in Low-dose CT Image using Eigenspace Sparse Shape Composition

Geng Chen, Dehui Xiang, Bin Zhang, Haihong Tian, Xiaoling Yang, Fei Shi, Weifang Zhu, Bei Tian, Xinjian Chen

***Abstract*—Segmentation of lungs with severe pathology is a nontrivial problem in clinical application. Due to complex structures, pathological changes, individual differences and low image quality, accurate lung segmentation in clinical 3D CT images is still a challenging task. To overcome these problems, a novel dictionary-based approach is introduced to automatically segment pathological lungs in 3D low-dose CT images. Sparse shape composition is integrated with eigenvector space shape prior model, called eigenspace sparse shape composition, to reduce local shape reconstruction error caused by weak and misleading appearance prior information. To initialize the shape model, a landmark recognition method based on discriminative appearance dictionary is introduced to handle lesions and local details. Furthermore, a new vertex search strategy based on gradient vector flow field is also proposed to drive shape deformation to target boundary. The proposed algorithm is tested on 78 3D low-dose CT images with lung tumors. Compared to state-of-the-art methods, the proposed approach can robustly and accurately detect pathological lung surface.**

***Index Terms*—Pathological lung segmentation, eigenspace sparse shape composition, gradient vector flow, discriminative appearance dictionary.**

I. Introduction

Computed tomography (CT) is a diagnostic imaging technique widely used for lung diseases, and especially low-dose CT is commonly used for lung tumor analysis. As an important preprocessing step in automatically analyzing lung, automatic lung segmentation has received considerable attention from researchers [1]. Accurate and automatic lung segmentation can save physicians' efforts to annotate lung anatomy since it is tedious and time-consuming to label each voxel in huge amount of slices. In addition, lung segmentation can help to improve the accuracy of lung tumor segmentation and lung nodule detection by 17% [2]. However, it is a nontrivial task to accurately and automatically segment lungs in CT images, especially in the low-dose CT images. Pathological lung segmentation is still a challenging task in medical image processing field, as shown in Fig.1. First, the internal structures of lungs are complex and difficult to recognize. Lungs are part of the lower respiratory and divided into different lobes. Pulmonary arteries and veins spread throughout the lungs. Second, there are often large pathological lesions. Large tumors lead to large difference in image intensity values, and great changes to structures and shapes. Third, the anatomy of lungs varies largely from different healthy individuals, both in shape and size. Forth, image quality is often low. The image contrast to background is often low due to the low-dose radiation and pathological lesions such as tumors. The boundaries between the tumors and neighboring organs or structures are also not clear. The movement artifacts produced by breaths are obvious in 3D low-dose CT images.

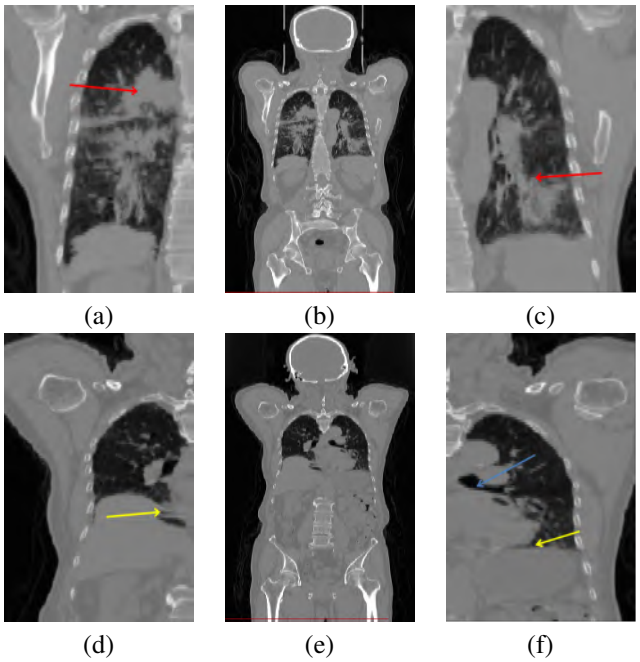


Fig. 1. Illustration of the challenges in pathological lung segmentation. (b) and (e) A low-dose CT image; (a)(d) the locally enlarged right lungs; (c)(f) the locally enlarged left lungs. Red arrows indicates large tumors. Yellow arrows indicates the fuzzy boundary, and blue arrows denotes the trachea with the similar appearance to the lung.

This work has been supported in part by the National Basic Research Program of China (973 Program) under Grant 2014CB748600, and in part by the National Natural Science Foundation of China (NSFC) under Grant 61622114, 61771326, 81371629, 61401293, 61401294, 81401451, 81401472, and in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant SJCX17_0650. Geng Chen, Dehui Xiang and Bin Zhang contributed to this work equally. Corresponding authors: Dehui Xiang, Xinjian Chen, Bei Tian.

Geng Chen, Dehui Xiang, Haihong Tian, Xiaoling Yang, Fei Shi, Weifang Zhu and Xinjian Chen are with the School of Electronics and Information Engineering, State Key Laboratory of Radiation Medicine and Protection, Soochow University, Jiangsu 215006, China (email: xiangdehui@suda.edu.cn).

Bin Zhang is with the first affiliated hospital of Soochow University, Suzhou, China.

Bei Tian is with the Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University. No.1 Dongjiaominxiang Street, Dongcheng District, Beijing, 100730, China.

In this paper, a novel shape prior model based segmentation framework with appearance priors is proposed to address the aforementioned challenges in the segmentation module of the surgery planning. The proposed framework consists of two parts: training and testing. The proposed segmentation framework is successfully applied to segment 3D pathological lungs in 78 low-dose clinical CT scans with leave-one-out strategy. The main contributions of the proposed method are as follows:

1) Inspired by [3]–[5], a new sparse shape composition (ESSC) is proposed to reconstruct 3D surface of pathological lungs. In [3], [5], overall shape was reconstructed from the aligned shape dictionary and an error variable was introduced to handle single point location errors. Compared to traditional sparse shape composition (SSC), eigenvectors obtained by principal component analysis (PCA) are integrated into the objective function of SSC. All training shapes are registered to the average shape of the point distribution model (PDM) using similarity transformation to construct shape dictionary for sparse shape composition in the eigenspace. This process refines the rough surface obtained by shape deformation, and avoids the loss of details due to the anatomy variations of individual lungs and large lesions. This strategy reduces the loss of local detail in the traditional SSC in [3]–[5] when the shape is constrained to be deformed to the target boundary.

2) A new hybrid search strategy is developed to deform shape based on normal direction and gradient vector flow (GVF) field [6]. In [3], [5], a learning-based method was employed for landmark detection. There may be gross errors or point missing from the detection. Worse still, the initial shape was only deformed following the image gradient information. However, there is no point missing in our landmark detection by using our ESSC model in the process of shape deformation. Our deformation model was based on a hybrid search strategy including normal-based, GVF-based in the deformation stage. On the one hand, normal-based deformation [7] constrained with our ESSC model can rapidly transform the three-dimensional mesh into the capture range of GVF. On the other hand, GVF-based deformation also constrained with our ESSC model allows mesh vertices to smoothly move towards target boundary even with large concavity. In addition, neighboring vertices are also considered to increase the accuracy and robustness of our approach.

3) A new vertex deformation algorithm is designed by using the GVF field and a distance function in the iterative optimization process. As the deformed mesh moves toward target boundary, the gradient vector field should be almost parallel with the normal of vertices of deformed mesh at the boundary of lungs. Therefore, the projection method is used to suppress non-parallel directions for vertex searching. In addition, a new target boundary deformation function is designed so that target points are found near a corresponding initial mesh and the mesh adaptation is stopped correctly. The new distance function is considered from two aspects: the region distance and the surface distance. The region distance makes the deformation of the mesh as close as possible to the target boundary. The surface distance ensures smoothness of the deformed mesh.

4) A robust 3D shape recognition approach iteratively integrates 3D representative shape prior and discriminative appearance prior together to derive patient-specific initial shapes to target boundary based on the label reconstruction by DAD [8] and threshold segmentation [9]. The reconstructed label value near the target surface is combined with the normal-based and GVF-based search strategy. Therefore, this integration can combine discriminative appearance priori and representative shape priori, in order to deal with lung tumors to accurately recognize local details and solve the efficiency problem of the original appearance dictionary learning method.

II. Related Work

Many previous methods have been proposed for automatic pathological lung segmentation. An overall review of different algorithms have been presented by [10]–[12], including their feasibility and shortcomings in the case of the most common lung abnormalities. Prasad $et\ al.$ [13] proposed an adaptive thresholding based pathological lung segmentation method. The threshold value was determined by a comparison of the curvature of the lung boundary to that of the ribs. Although their method improved over conventional thresholding techniques significantly, it produced large errors in lesions lying inside the pulmonary lobe. Wang $et\ al.$ [14]–[16] proposed a two-stage method for the segmentation of lungs with interstitial disease. Sluimer $et\ al.$ [17] proposed a registration approach for the segmentation of pathological lungs. Their method achieved high accuracy when the lung tissue was only affected by minimal to moderate pathology, but failed in many pathological cases. Shi $et\ al.$ [18] leveraged low-rank and sparse decomposition theory for robust pathological lung segmentation in CT images. Although the authors delivered promising results for severely pathological cases, the computing time reached 145 min and needed to be reduced to clinically acceptable range.

Recently, the deformable segmentation model has been applied in the field of the medical imaging computing. In order to locate target boundary in three dimensions adequately, boundary candidates were searched along the normal vector of the vertex at equidistant positions for each vertex on a subspace shape model [9]. The normal vector only takes into account the initial shape information. For many local details with relatively large curvature, the deviation of the mesh vertex often occurs. This gives a high requirement for the subsequent reconstruction algorithm. Ecabert $et\ al.$ [19], [20] utilized a projection-based approach to suppress non-parallel directions for vertex searching. This alleviated the deficiency of the normal vector of the vertex. Nevertheless, the final mesh was mainly based on the normals of the mesh vertices. Kiaei $et\ al.$ [21] improved a snake algorithm to deform through the guidance of the gradient vector flow field instead of the gradient field. Nevertheless, except for adjusting parameters, no priori knowledge about the shape can be expressed. Traditional model-based deformable approaches, such as active shape model [22] and its series of extensions [23]–[25], are able to combine low-level appearance and high-level shape in a unified framework. However, it is difficult to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging

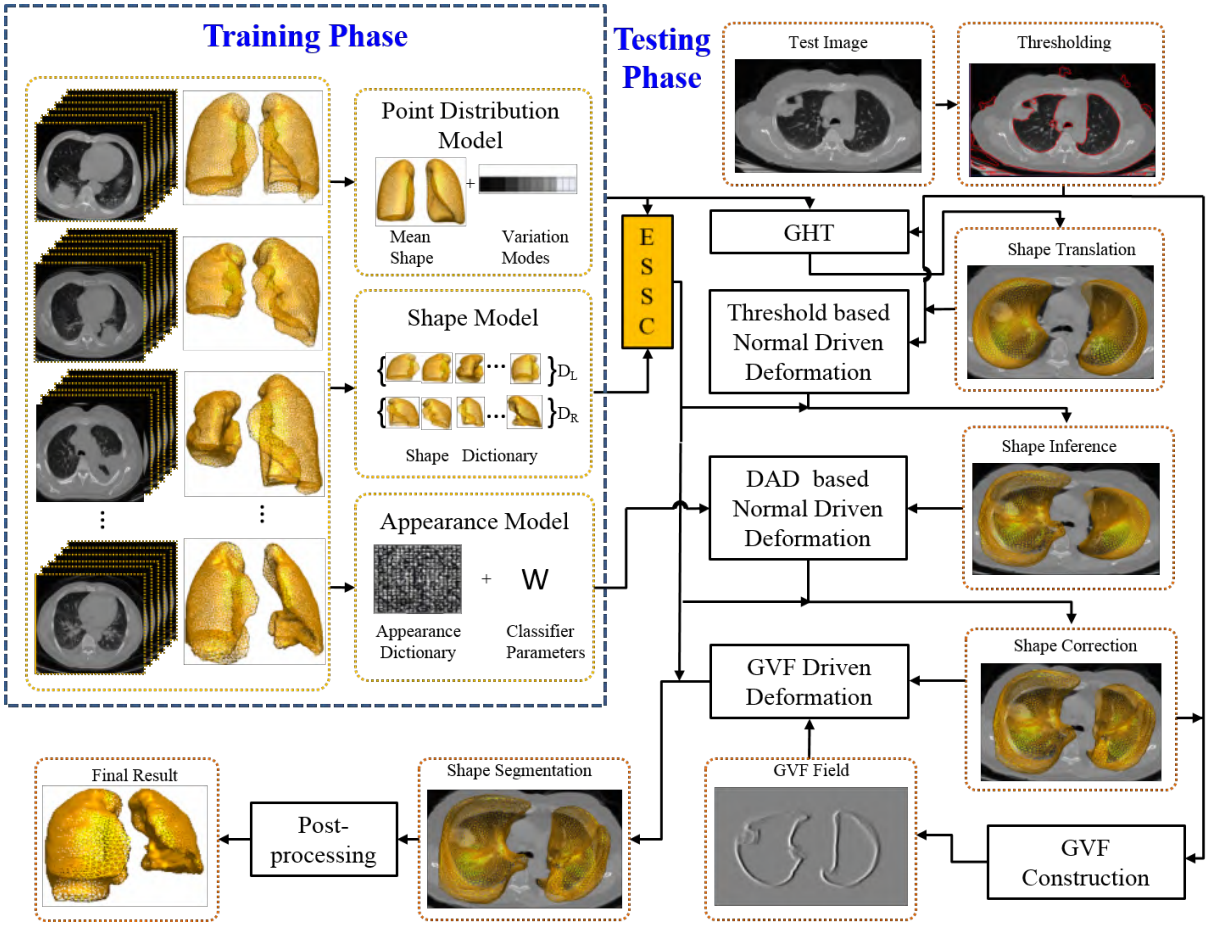


Fig. 2. The framework of the proposed method.

preserve local details. In addition, the appearance model is assumed to fit the intensities as Gaussian distributions, which do not hold in many applications [4].

To address these challenges in a unified framework, Zhang *et al*. [3] proposed a combined method with landmark detection and shape inference. However, it might lead to image occlusion or landmarks missing. To solve this problem, they modeled reconstruction error as a sparse vector. However, it was difficult to effectively fit the shape modes especially for local details with relatively large curvature. Farhangi *et al*. [26] proposed to segment 3D lung nodules by active contour based SSC algorithm. To solve a large scale sparse optimization problem, Zhang *et al*. [5] performed a sparse shape composition on learned dictionary instead of using the original database. However, it tended to neglect the local details. A set of local SSC models [**?**], [4], [5] were proposed to describe the shape in a segment-to-segment manner and build shape models on them independently. Regardless of the above improvements or in combination with other algorithms [**?**], the model had been not fundamentally optimized.

In addition, deep learning [27] has been demonstrated to be a powerful tool in medical image segmentation. One of the most popular methods for semantic segmentation was U-Net [28]. Md Zahangir *et al*. [29] proposed a Recurrent Residual convolutional Neural Network (RRCNN) for lung lesion segmentation, which utilized U-Net, Residual Network and RCNN. Rodney LaLonde *et al*. [30] developed a convolutional-deconvolutional capsule network (SegCaps) for pathological lung segmentation in CT scans. The proposed framework was recently introduced from the capsule network architecture. Experimentally, SegCaps not only produce slightly improved accuracies, but also contain 95.4% fewer parameters than U-Net [28] and 38.4% fewer than Tiramisu [31], when compared with state-of-the-art networks U-Net [28], [31].

III. Method

The proposed method aims to segment pathological lungs in 3D low-dose CT images using both shape and appearance dictionary. Fig.2 shows the framework of the proposed method, including training and testing phases. In the training phase, we construct three models: PDM includes mean shape and its variation modes, shape dictionary and discriminative appearance model based on dictionary learning. In the testing phase, a threshold based normal driven deformation method is utilized to obtain an initial shape. Subsequently, the initial shape is iteratively deformed and refined by ESSC with vertex normal and GVF field.

A. Model Training

Labeled binary images by clinical experts are converted into triangulated meshes to represent manual segmentation by

using marching cube algorithm [32]. Minimum description length algorithm [33] is used to obtain the corresponding relationship between all triangulated meshes for training data. A training sample mesh is randomly selected as the reference, and the rest of the meshes are aligned through similarity transformation in a three dimension Cartesian coordinate system.

The training repository of shape prior model for sparse shape composition can be represented as a matrix $D_s = [d_1, d_2, \cdots, d_i, \cdots, d_{k_s}] \in R^{n_s \times k_s}$, where d_i is three dimension Cartesian coordinate vector of the ith training mesh, k_s is the number of training meshes and n_s is the number of vertices of the mesh multiplied by the dimension. Inter-patient and inter-phase shape variability of lungs can be learned from the consistent set of training meshes using PCA. The PDM can be used to describe shape variability. With the combination of similarity transformation, the resulting PDM can be expressed as,

$$\psi_k = T^{-1}\left(\bar{\psi} + \sum_{m=1}^{k_s} \lambda_m p_m\right), \tag{1}$$

$$\bar{\psi} = \frac{1}{k_s}\sum_{m=1}^{k_s} \psi_m \tag{2}$$

where, ψ_k denotes the kth aligned triangulated mesh. T^{-1} is the inverse of similarity transformation from the original coordinate system to the registered shape coordinate system. p_m is the variation mode of obtained through PCA. λ_m is the corresponding weight for the principal mode p_m.

For each vertex in the ψ_k, the same number of 3D patches is extracted in the same neighboring positions from all training CT images to form a training 3D patch library P_l. Each 3D patch is denoted as a column vector and group all the 3D patches together as a matrix $P_l = [p_1, p_2, \cdots, p_i, \cdots, p_{n_a}] \in R^{m_a \times n_a}$. P_l denotes the training patch p_i library containing n_a patches, and m_a is the size of the 3D patch. Therefore, with the combination of manual label, the appearance dictionary can be expressed as [8], [34],

$$\arg\min_{D_a,\alpha,w} = \|P_l - D_a\alpha\|_2^2 + \lambda_1 \|L - w\alpha\|_2^2 + \lambda_2\|\alpha\|_1 \tag{3}$$

where $D_a = [a_1, a_2, \cdots, a_i, \cdots, a_{k_a}] \in R^{m_a \times k_a}$ is the learned dictionary. $\alpha \in R^{k_a \times n_a}$ is the sparse coding coefficient matrix of the input patch library, k_a is the number of dictionary elements and L is manual label corresponding to P_l. w is the learned linear predictive classifier. λ_1, λ_2 are scalar constants.

B. Eigenspace Sparse Shape Composition

Given an input shape to be refined $\psi \in R^{n_s}$, it can be approximately represented as a weighted linear combination of existing shapes D_s, and the parts which cannot be approximated are noise [3]. Mathematically, it seeks the optimal weight β_1 in a least-squares sense, leading to the following minimization problem,

$$\arg\min_{\beta_1} \|T(\psi) - T_S(D_s)\beta_1\|_2^2 \tag{4}$$

where T is similarity transformation from the deformed mesh to the average mesh calculated in the training step, and T_S is similarity transformations from the training meshes to the average shape. The refined input shape ψ is presented as D_s and transformed back by the inverse of the transformation T^{-1}.

There are two limitations in Eq.(4) according to [3]. First, the matrix D_s maybe overcomplete and the function may not have a unique solution. Second, if any linear combination can be used, the noises included in the input shape may be preserved. Thus, the classical SSC performs a sparse linear combination of training shape instances to represent a pre-aligned input shape. The specific approach is to add the constraints of $L0$ norm on the weight β_1. The minimization problem can be expressed as,

$$\begin{aligned} &\arg\min_{\beta_1} \|T(\psi) - T_S(D_s)\beta_1\|_2^2, \\ &s.t. \|\beta_1\|_0 \le K_1 \end{aligned} \tag{5}$$

where K_1 is the pre-defined sparsity and ensures that the number of nonzero elements in β_1 is smaller than K_1. Two remaining limitations of the SSC model in Eq.(5): (1) The global shape tends to be fitted only as the overall shape and local details are often neglected; (2) It will have poor performance in case where the input shapes are corrupted by gross errors or outliers due to complex lesions or sharp structures. To address these problems of the classical SSC, eigenvectors are integrated into SSC, named eigenspace sparse shape composition, so as to represent local reconstruction errors of the shape by additional sparse weight β_2 to select the learned different components in P_m. ESSC is designed according to two observations: (1) After being aligned to a unified space, a given shape can be approximated by a sparse linear combination of training shape instances; (2) If the approximated shape is refined by SSC, reconstruction errors might include gross errors from boundary refinement, especially local errors from pathological changes, sharp structures and individual variations. Therefore, ESSC is expressed as minimizing the following objective function,

$$\begin{aligned} &\arg\min_{\beta_1,\beta_2} \|T(\psi) - T_S(D_s)\beta_1 - P_m\beta_2\|_2^2, \\ &s.t. \|\beta_1\|_0 \le K_1, \|\beta_2\|_0 \le K_2 \end{aligned} \tag{6}$$

where K_2 is the sparsity of P_m. The optimization problem in Eq.(6) is in general intractable and NP-hard due to the non-convexity of $L0$ norm. Recently, it has been proved that solving this kind of problem through $L1$ norm relaxation can achieve the same recovery accuracy. Therefore, (6) can be defined as,

$$\arg\min_{\beta_1,\beta_2} \|T(\psi) - T_S(D_s)\beta_1 - P_m\beta_2\|_2^2 + \lambda_3\|\beta_1\|_1 + \lambda_4\|\beta_2\|_1 \tag{7}$$

where the first term is the shape reconstructive term with the combination with SSC and eigenvectors, the second term adds the sparsity constraint over the shape coding coefficients β_1, the third term adds the sparsity constraint over the eigenspace coding coefficient β_2, ψ is the deformed mesh, $P_m = [p_1, p_2, \cdots, p_i, \cdots, p_m]$ is the eigenvector matrix, λ_3 and λ_4 are two scalar constants.

Eq.(7) is solved by alternating minimization scheme. After ψ and D_s are registered to the average mesh $\bar{\psi}$, the first

optimization is executed using the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) [35] in the following form,

$$\arg\min_{\beta_1} \|T(\psi) - T_S(D_s)\beta_1\|_2^2 + \lambda_3\|\beta_1\|_1 \tag{8}$$

The overall shape ψ can be reconstructed from the aligned D_s in the coordinate system of the average mesh $\bar{\psi}$. However, due to pathological changes, sharp structures and individual variations, the gross errors are inevitable. This problem was tacked in [3] by adding a special error term e to the objective function. Shape dictionary D and an unit matrix M were concatenated into a matrix $[D, M]$, and then the weight $\beta = [\beta_1, e]$ was computed using FISTA. The final refined shape was constructed by $\psi' = D\beta_1$. The matrix concatenation reduced the efficiency of the algorithm tremendously. In addition, the fitted errors were not utilized to reconstruct the mesh. Local details such as small structures with large curvature tended to be neglected. Therefore, the reconstruction error e is further represented by variation modes P_m. The second optimization is to fit the local errors e by selecting different components in P_m. Mathematically, it leads to the following minimization problem,

$$\arg\min_{x_2} \|e - P_m\beta_2\|_2^2 + \lambda_4\|\beta_2\|_1 \tag{9}$$

Eq.(9) is also solved by FISTA [35]. The refined shape ψ is generated as

$$\psi = T^{-1}\left(T_S(D_s)\beta_1 + P_m\beta_2\right) \tag{10}$$

C. Shape Deformation

A curvature anisotropic diffusion filter [36] is applied to remove noise in the test image. The intensity range of the smoothed image is normalized to $[I_{min}, I_{max}] = [0, 255]$, and then a binary image is obtained by thresholding and morphological operation. Generalized Hough Transform (GHT) [37] is used to find the initial center based on the average shape model. The average mesh is moved to the center as the initial mesh. Initially, the moved average mesh is far from the target boundary. Briefly, the coarse lung is recognized by using thresholding and appearance dictionary; the deformed mesh is adapted along normal or GVF direction to the boundary of the recognized lung by DAD, and the adapted mesh is reconstructed by ESSC. The iterative process is successively employed in three alternating steps.

(1) Normal Based Lung Deformation

In order to derive patient-specific initial shape, an initialization method combining appearance reconstruction and normal search strategy is proposed. Specifically, the initialization consists of three parts: 1) threshold-based shape adaptation; 2) DAD-based shape adaptation; 3) ESSC-based initial shape inference.

1) Threshold-based shape adaptation: Initially, the moved average mesh is adapted as previous work [7], [9]. The mesh is progressively deformed along the normal of its vertices to detect the candidate boundary of the thresholded coarse lung.

2) DAD-based shape adaptation: The recognized coarse binary lung by thresholding is then corrected by the DAD based label reconstruction around the deformed mesh ψ.

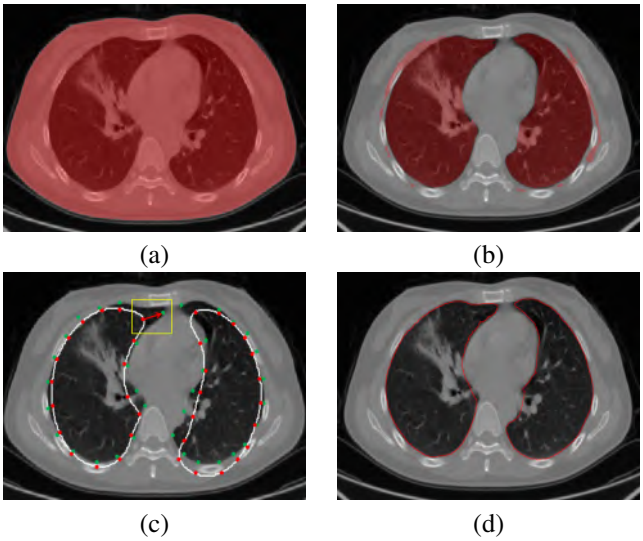


Fig. 3. Illustration of the difference between our DAD and [8]. (a) A CT image slice with the red mask representing the segmentation scope of in [8]. (b) Segmentation result by [8]. (c) Our DAD vertices initialization. White curve represents the initial surface. Yellow rectangle indicates the vertices of the 3D mesh (red points) are searched towards the target vertices (green points) along the normal direction. (d) Segmentation results after ESSC inference.

Appearance dictionary representation in testing stage can be described as

$$\arg\min_{\alpha_t} \|p_t - D_a\alpha_t\|_2^2 + \lambda_2\|\alpha_t\|_1 \tag{11}$$

where p_t denotes a patch around a vertex in the deformed mesh, $\alpha_t \in R^{k_a}$ is the sparse coding coefficient vector of the input patch p_t. The center voxel of the input patch p_t can be given a label according to

$$h = w * \alpha_t \tag{12}$$

Reconstructed label is used to guide the shape deformation as shown in Fig.3 (c). Specifically, assuming a discrete vertex v_i is in the deformed mesh ψ, and the normal of the vertex is represented by a vector n_i, the candidate position can be expressed as

$$v_{ij} = v_i + \frac{h - h_t}{|h - h_t|} j\delta n_i \tag{13}$$

where δ is the search step size, $j = 0, 1, ..., n_m$ is the search index along the normal direction. n_m is the maximum value of the searching range and h_t is the threshold value of reconstructed gray value of vertex v_i. When the label of v_{ij} is not the same as that of v_i, sampling is stopped. If $j > n_t$, and then v_i is updated by v_{ij}. n_t is a moving threshold value.

3) ESSC-based shape inference: Due to the appearance cues misleading and intrinsic defect in normal direction as shown in Fig.5 (a)-(c), there may be errors from the vertex selection. ESSC can be used for shape inference as a regularization step during deformation and employed as high level constrains to avoid getting stuck in local minima. Specifically, the deformed mesh ψ and shape dictionary D_s are registered to the average mesh $\bar{\psi}$ and get the registered mesh $T(\psi)$ and the registered dictionary $T_s(D_s)$. Then, β_1 is computed by solving Eq.(8). $T_S(D_s)\beta_1$ is considered as an overall constructed shape. The local reconstruction error is further fitted in the eigenspace by

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging

solving Eq.(9). The refined mesh ψ' is computed by solving Eq. (10) with β_1 and β_2. Iterative deformation is repeated until the distance Δd of the refined meshes between successive iterations is not larger than a given threshold value Δd_t. Eq.(8) reconstructs the shape by selecting a few globally similar training shapes but lung shape variation is inhomogeneous. In the refinement procedure, e may not have large value. However, the variation modes can be better used to capture the local shape variation.

(2) GVF Based Lung Deformation

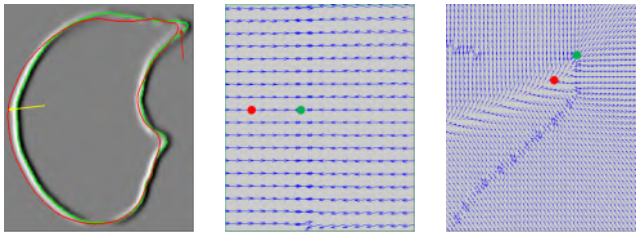


Fig. 4. The process of GVF-based shape deformation. (a) Gray scale of GVF field; (b) The locally enlarged GVF field pointed by the yellow arrow in (a); (c) The locally enlarged GVF field with large curvature pointed by the red arrow in (a). (b) and (c) show initial points are guided by GVF towards the corresponding target points (Green). Manual segmentation is shown in green curves, and initialization is shown red curves.

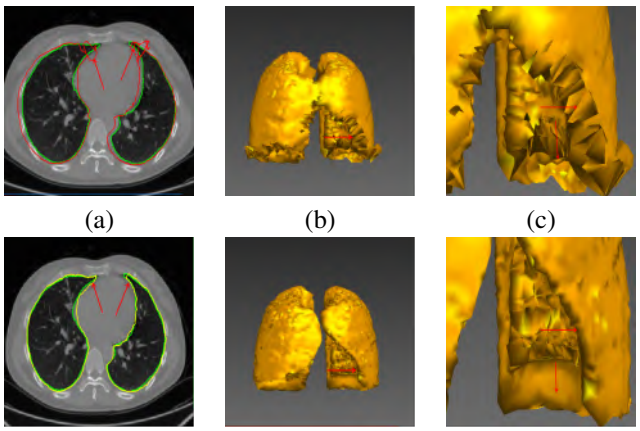


Fig. 5. Comparison of shape deformation based on normal (1st row) and GVF (2nd row). (a) Segmented lungs (red curves) by normal-based shape deformation; (b) Segmented lung surfaces of (a); (c) Locally enlarged region of (b); (d) Segmented lungs (red curves) by GVF-based shape deformation; (e) Segmented lung surfaces of (d); (f) Locally enlarged region of (e). Local detail (red arrows) detection is significantly improved by GVF-based shape deformation.

Although deformation along the normal of vertices allows the mesh to move toward target boundary. There are also two key difficulties with normal orientation. First, initial mesh is difficult to progressively and smoothly reach target boundary with large concavity, such as sharp structures. The external force is too strong to cause the boundary to overwhelm where the curvature of the shape is large. The second problem is that the normal of the vertices only take into account initial shape curvature and the moving direction of the vertices is only related to its initial position. It means that the initial vertex generally must be close to target boundary, and otherwise it will probably lead to false deformation, as shown in Fig.5(a)-(c).

In this section, a new external force based on GVF is introduced into our model. GVF force pertains rotation forces to addresses above problems. According to the previous work [6], [21], [38], an edge map can be defined as

$$G = -G_\sigma * B \tag{14}$$

where the G_σ is a 3D Gaussian function with standard deviation σ, $*$ is convolution and B is the binary image by thresholding and DAD reconstruction. GVF field vectors generally have large magnitudes only on the boundaries and point toward the boundaries (green point), as shown in Fig.4. The GVF capture range is controlled by setting the standard deviation σ. As shown in Fig.4, the improved method is used as the follows. First, the starting vertex v_i (red point) in the deformed mesh ψ has a GVF vector g_c, and g_c's angle is θ_c. Considers the point that is a neighboring point of v_i in the direction of θ_c as v_j (green point). v_j should have a GVF vector g_d nearly paralleling to that of v_i, i.e., $\theta_c \approx \theta_d$. Secondly, $\|G(v_j)\|$ represents the edge map gradient magnitude at the point v_j. Considering a new point v_j $(j = 1, 2, 3...)$ as the neighboring point v_{j-1} along the direction of v_{j-1}, v_j is successively searched if the following criterion is satisfied,

$$\theta_{v_j} \approx \theta_{v_{j-1}}, \|G(v_j)\| > \|G(v_{j-1})\| \tag{15}$$

The profile points $\{v_j^1, v_j^2 ... v_j^L\}$ are placed in regular intervals along the GVF as described above. The target point is defined by

$$v_i^{target} = \arg \max_{\{v_j^l \| l=0,1...L\}} (F_i(v_j^l) - \gamma(v_j^l - v_i)^2) \tag{16}$$

where the $F_i(\cdot)$ is the edge detector at a discrete position v_j^l. The target point is chosen on the profile where term in bracket is maximal. γ controls the weight of a penalty term, which biases the search to nearby points.

In our work, the 3D image GVF field G is used as the boundary detectors. G is projected onto the triangle normal n_i to suppress non-parallel edges. The result is passed through a sigmoid function limiting the response to maximal magnitude g_{max}.

$$G_{proj}^{limit}(v_j^l) = (n_i \cdot G(v_j^l)) \frac{g_{max}(g_{max} + \|G(v_j^l)\|)}{g_{max}^2 + \|G(v_j^l)\|^2} \tag{17}$$

The boundary detection function $F(\cdot)$can be given as

$$F_i(v_j^l) = \begin{cases} \pm G_{proj}^{limit}(v_j^l), & \text{if } H_{v_j^l} \in [h_{min}, h_{max}] \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

$\pm$ ensures the consistency of direction, which is to opposite at the edge of intersection prevent. Each $H_{v_j^l}$ denotes the reconstructed label near or across boundary by appearance dictionary learning which must full inside a corresponding acceptance interval $[h_{min}, h_{max}]$.

Although the GVF field converges to the concave boundary and the maximum deformation range is controlled in order to minimize the influence of lesions. Local detail mistakes are inevitable and ESSC is used again for shape inference

Algorithm 1 GVF Based Lung Deformation.

Require: DAD-based adaptation mesh ψ; Mean mesh $\bar{\psi}$; Principal modes P_m; Appearance dictionary D_a; Shape dictionary D_S; Classifier parameters w; Balance parameter, λ_2, λ_3, λ_4, γ, λ_e; Distance threshold value E_t; The corresponding GVF field G.

Ensure:
Adapted mesh, ψ_τ.

1: Register each training shape d_i in the shape dictionary D_S to the mean mesh $\bar{\psi}$ and get similarity transformation parameter $T_i(d_i)$ of the transformed mesh d_i;
2: Construct the registered dictionary $T_S(D_S)$;
3: $\psi_\tau \leftarrow \psi$;
4: Initialize energy $E_b \leftarrow \infty$;
5: Sample vertices along the GVF vector g_c of ψ_τ according to Eq. (15).
6: Compute the reconstructed label $H_{v_j^l}$ according to Eq.(11) and Eq.(12);
7: Update v_i^{target} according to Eq.(16) and obtain the deformed mesh ψ_τ;
8: Compute the mesh deformation region distance $E_{region}(p)$ according to Eq. (20).
9: Register the deformed mesh ψ_τ to the mean mesh $\bar{\psi}$ and get the similarity transformation parameter T and the registered mesh $T(\psi_\tau)$;
10: Compute weight β_1, β_2 of the ESSC using FISTA [35] to Eq.(8) and Eq.(9);
11: Compute the refined mesh ψ_λ according to Eq.(10);
12: Compute $E_{surface}$ according to Eq.(21);
13: Compute E according to Eq.(19);
14: $\Delta E \leftarrow E_b - E$. $E_b \leftarrow E$.
15: If $|\Delta E| < E_t$, $\psi_\tau \leftarrow \psi_\lambda$, and then go to 5; else, get ψ_τ and stop.

iteratively. In each iteration, the first step is mesh deformation by progressively detecting the candidate boundary along the GVF field direction. In the second step, the mesh by the normal initialization and variation modes is adjust to generate a subspace shape model ψ_λ and constrains the deformation of mesh ψ_τ. The deformable adaptation can be described as minimized a weighted combination

$$E = E_{region} + \lambda_e E_{surface} \tag{19}$$

where, E_{region} denotes a region term, which measures the magnitude distance between ψ_τ and ψ_λ in the GVF field. ψ_t represents mesh of the threshold image and the reconstructed label image. The distance function is defined as,

$$E_{region} = \frac{1}{N_r} \sum_p | \, \|G_\tau(p)\| - \|G_t(p)\| \, | \tag{20}$$

p represents a voxel in GVF field. N_r denotes a normalization parameter in the local region. $\|G_\tau(p)\|$ and $\|G_t(p)\|$ represent the corresponding magnitude of ψ_τ and ψ_t in the GVF field, respectively. $E_{surface}$ denotes a boundary term which measures the coordinate distance between ψ_τ and ψ_λ. The constraint of neighborhood information is considered as

$$E_{surface} = \frac{1}{N_s} \sum_{v=1}^{n_s/3} \sum_{v' \in N_v} |\psi_\tau(v) - \psi_\tau(v') - (\psi_\lambda(v) - \psi_\lambda(v'))| \tag{21}$$

N_s denotes a normalization parameter for the deformed mesh ψ_τ. N_v lists the indices of neighbor vertices of v. The iterative process between GVF-based shape adaptation and ESSC-based shape inference is described as Algorithm 1.

IV. Experiments and Results

A. Data

The proposed method was evaluated on 78 3D low-dose CT images obtained from different patients with lung tumors. The image are acquired by a GE Discovery ST16 PET-CT scanner from the top of the skull to the upper part of the femur. CT scanning parameters were 120KV voltage, 150 mA current, and 3.75mm thickness. The pathological lungs were manually labeled as ground truth by clinical experts. The leave-one-out strategy was used to test the proposed method. All the original CT image size is $512 \times 512 \times 299$, and the voxel size is $0.98mm \times 0.98mm \times 3.75mm$.

B. Evaluation

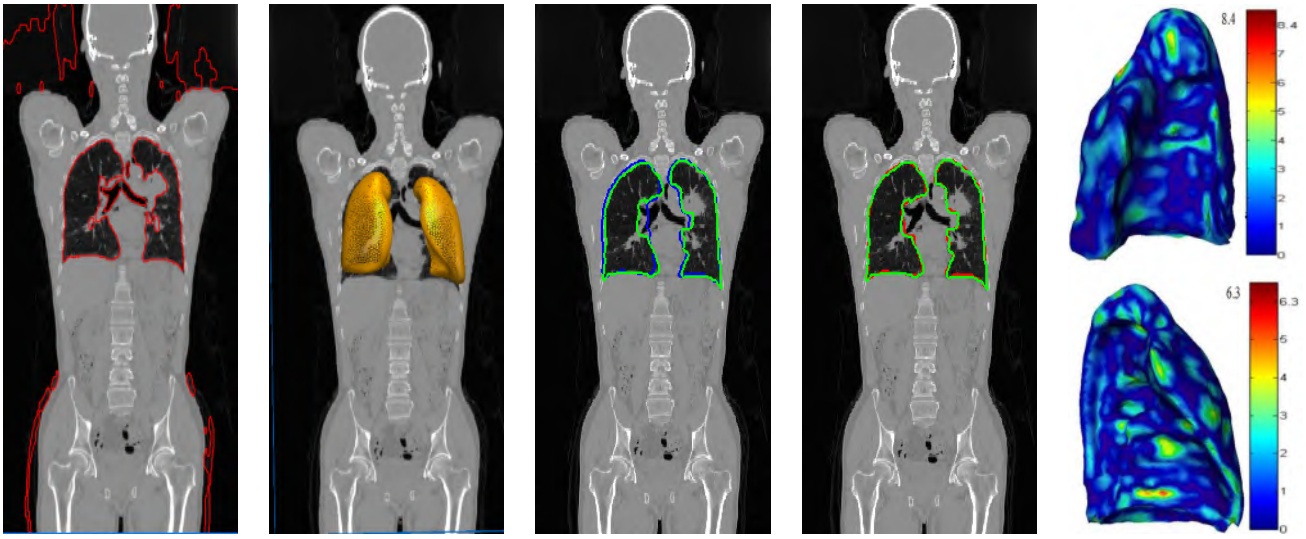
To quantitatively assess the performance of our proposed method, we compared the segmentation results with the ground truth according to the following five volume and surface based metrics: average symmetric surface distance (ASD), maximum surface distance (MSD), true positive fraction (TPF), false positive fraction (FPF) and dice similarity coefficient (DSC) [9]. Paired t-test were conducted to compare the difference in segmentation results between our method and related methods, and a p-value less than 0.05 was considered statistically significant.

C. Parameter Setting

Eq.(7) has two user-adjustable parameters λ_3 and λ_4, which are generally critical for performance and convergence. Fortunately, it is not sensitive to different images. λ_3 controls the sparsity of x_1. λ_4 controls the sparsity of x_2. To generate a sparse coefficient x_1, a large λ_3 is necessary. Otherwise, Eq.(7) degenerates to the choice of shape mode p_m through x_2. λ_4 is similar to λ_3. To test the sensitiveness of the proposed ESSC model to the two main parameters λ_3 and λ_4, we have conducted experiments on the 78 3D low-dose CT images. The left and right lungs were segmented separately, as shown in Fig.6. Both of λ_3 and λ_4 took values from 15 to 55 with an interval of 10. Therefore, fifty experiments were done. DCS in Fig.6 shows that this model is not sensitive to these two hyper-parameters. In this study, $\lambda_3 = 35$ and $\lambda_4 = 35$ were used.

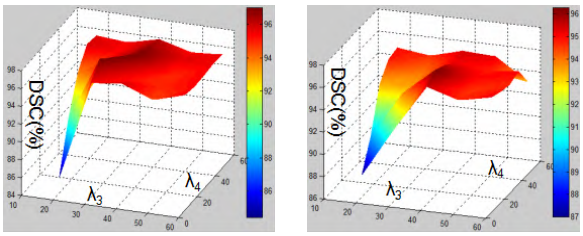
For the remaining parameters, we used empirically [7], [9] the following parameters: $n_t = 10$, $n_m = 41$, $k_s = 57$, $h_t = 0.5$, $\lambda_1 = 1$, $\lambda_2 = 0.15$, $\sigma = 0.1$, $d_t = 0.5$, $\lambda_e = 0.1$, $\gamma = 0.5$ and $E_t = 0.05$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging



(a) (b) (c) (d) (e)

Fig. 8. Lung initialization and segmentation in a CT image. (a) The red curve represents preprocessed boundaries; (b) Moved shape model (yellow surface) using an GHT based method; (c) Initialized lungs based on DAD and ESSC (blue curve); (d) The red curve represents final segmentation of the lungs, the green curve represents ground truth; (e) The surface distance of corresponding lung final segmentation to manual segmented lung surface.



(a) (b)

Fig. 6. Parameter sensitivity tests of λ_3 and λ_4. (a) Parameter setting experiments were performed on the right lung; (b) Parameter setting experiments were performed on the left lung.

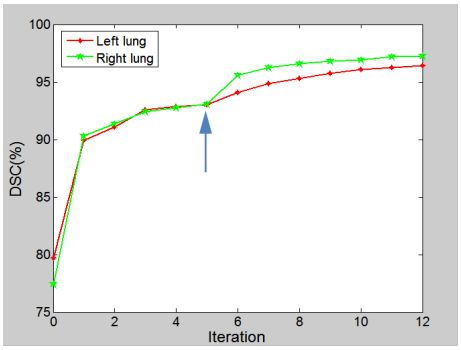


Fig. 7. The segmentation accuracy of the left and right lungs as the number of iteration increases. The blue arrow indicates that GVF begins to participate in vertex deformation. The starting point of the abscissa axis means initializing with the average shape.

D. Experimental Results

(1) Qualitative Results

Fig.8 shows the whole process using the proposed method in low-dose CT image. During the experiment, the GHT-based method can be successfully applied to localize the coarse position of the right and left lungs. Fig.8 (c) and (d) show the initial and final surface, respectively. Fig.7 shows the segmentation accuracy of the left and right lungs as the number of iterations increases, respectively.

(2) Comparison between Normal-based and GVF-based Methods

Fig.5 shows the results of deforming the vertices by using normal-based and GVF-based methods. As can be seen in Fig.5(a), it was difficult to detect those boundaries of the sharp structures in the both right and left lungs when the mesh was adapted along its normal direction. However, GVF-based method tended to smoothly progress into concave boundary. Therefore, the effects of gross error with vertices positioning will be largely eliminated in the subsequent processing steps.

Fig.9 demonstrates the initialization results of the same image by threshold-only deformation model as in [7], [9] and DAD-based deformation model, respectively. It clearly shows that, under the same shape constraints (ESSC), the initial position of the threshold-only deformation model was far away from ground truth, while the DAD based deformation model can drive the deformable shape mostly to the target boundary even with large tumors. Therefore, it can derive more patient-specific initial shapes in our segmentation framework.

To quantitatively analyze the difference of segmentation with the respective use of normal and GVF after average lung model localization. As shown in Table I, the GVF-based deformation method achieved higher accuracy than that normal-based method under both DAD and thresholding. NDAD means to detect the candidate boundary along the normal of vertices under DAD reconstructed, and NTM is under threshold and morphological operation as in [7], [9]. In addition, through the comparison between NDAD and NTM, we can also find that the shape initialization method is superior to thresholding as in [7], [9].

(2) Comparison between Different Shape Models

Fig.10 illustrates the comparative results of the lung segmentation by using shape prior models based on PCA, conventional SSC [4] and our proposed ESSC in the same

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging

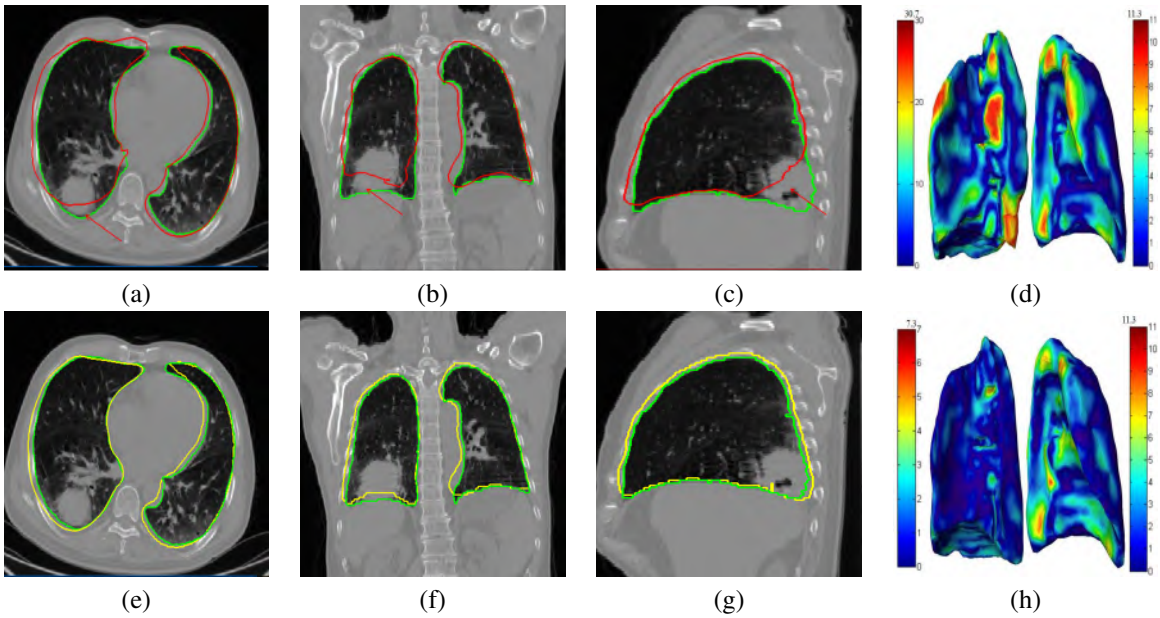


Fig. 9. Comparison of two different initialization generated by (a)-(d) threshold-based (red curves in the 1st row) and (e)-(h) DAD-based (yellow curves in the 2nd row) methods. (d) and (h) Initialized surface distance to ground truth, and the color bar maps and clips the surface distance between [0,10]. The maximal surface distance is given above the color bar.

TABLE I
QUANTITATIVE COMPARISON BETWEEN NORMAL-BASED AND GVF-BASED DEFORMATION AFTER INITIALIZE SHAPE INFERENCE (MEAN ± SD).

Method	Lung	ASD(mm)	MSD(mm)	TPF(%)	FPF(%)	DSC(%)
NTM	Left	3.79 ± 1.04	23.81 ± 6.44	94.31 ± 2.83	0.41 ± 0.12	91.31 ± 1.36
	Right	3.72 ± 0.72	23.62 ± 8.10	94.33 ± 2.49	0.51 ± 0.15	91.25 ± 1.18
NDAD	Left	3.17 ± 0.80	22.85 ± 8.76	94.91 ± 1.87	0.32 ± 0.13	93.05 ± 1.32
	Right	3.16 ± 0.49	21.82 ± 7.04	95.00 ± 1.44	0.39 ± 0.15	93.10 ± 1.17
GVF	Left	$\mathbf{1.71 \pm 0.61}$	$\mathbf{19.00 \pm 9.23}$	$\mathbf{96.89 \pm 2.17}$	$\mathbf{0.15 \pm 0.05}$	$\mathbf{96.39 \pm 1.34}$
	Right	$\mathbf{1.37 \pm 0.44}$	$\mathbf{15.21 \pm 6.39}$	$\mathbf{97.79 \pm 1.57}$	$\mathbf{0.14 \pm 0.05}$	$\mathbf{97.25 \pm 0.92}$

TABLE II
P VALUE OF QUANTITATIVE COMPARISON BETWEEN NORMAL-BASED AND GVF-BASED DEFORMATION AFTER INITIALIZE SHAPE INFERENCE.

Method	Lung	ASD	MSD	TPF	FPF	DSC
NTM vs. GVF	Left	4.79×10^{-16}	7.58×10^{-3}	1.04×10^{-6}	1.89×10^{-6}	2.31×10^{-18}
	Right	$8.61 \times 10{-}16$	9.38×10^{-9}	1.81×10^{-13}	2.38×10^{-17}	4.75×10^{-18}
NDAD vs. GVF	Left	5.26×10^{-17}	1.15×10^{-3}	6.66×10^{-5}	4.24×10^{-8}	4.49×10^{-15}
	Right	4.40×10^{-13}	1.59×10^{-7}	8.69×10^{-10}	7.98×10^{-11}	8.77×10^{-15}

TABLE III
QUANTITATIVE COMPARISON BETWEEN PCA, CONVENTIONAL SSC AND ESSC AFTER SHAPE DEFORMATION BY GVF (MEAN ± SD).

Method	Lung	ASD(mm)	MSD(mm)	TPF(%)	FPF(%)	DSC(%)
PCA	Left	3.28 ± 0.59	23.95 ± 8.13	94.98 ± 3.74	0.37 ± 0.21	92.03 ± 3.45
	Right	3.09 ± 0.82	22.59 ± 7.68	94.05 ± 1.49	0.18 ± 0.06	94.77 ± 0.94
SSC [4]	Left	2.72 ± 0.76	22.28 ± 6.94	93.86 ± 1.78	0.17 ± 0.05	94.19 ± 1.26
	Right	2.33 ± 0.93	19.47 ± 5.66	95.09 ± 1.78	0.15 ± 0.07	95.68 ± 1.04
ESSC	Left	$\mathbf{1.71 \pm 0.61}$	$\mathbf{19.00 \pm 9.23}$	$\mathbf{96.89 \pm 2.17}$	$\mathbf{0.15 \pm 0.05}$	$\mathbf{96.39 \pm 1.34}$
	Right	$\mathbf{1.37 \pm 0.44}$	$\mathbf{15.21 \pm 6.39}$	$\mathbf{97.79 \pm 1.57}$	$\mathbf{0.14 \pm 0.05}$	$\mathbf{97.25 \pm 0.92}$

TABLE IV
P VALUE OF QUANTITATIVE COMPARISON BETWEEN PCA, CONVENTIONAL SSC AND ESSC AFTER SHAPE DEFORMATION BY GVF (MEAN ± SD).

Method	Lung	ASD	MSD	TPF	FPF	DSC
PCA vs. ESSC	Left	2.77×10^{-10}	1.65×10^{-2}	6.74×10^{-7}	4.10×10^{-9}	7.69×10^{-13}
	Right	3.10×10^{-14}	5.04×10^{-7}	1.18×10^{-18}	1.04×10^{-3}	6.35×10^{-17}
SSC vs. ESSC	Left	1.12×10^{-11}	3.32×10^{-2}	4.73×10^{-11}	1.54×10^{-2}	3.08×10^{-10}
	Right	1.75×10^{-11}	5.43×10^{-5}	1.28×10^{-12}	0.33	1.63×10^{-9}

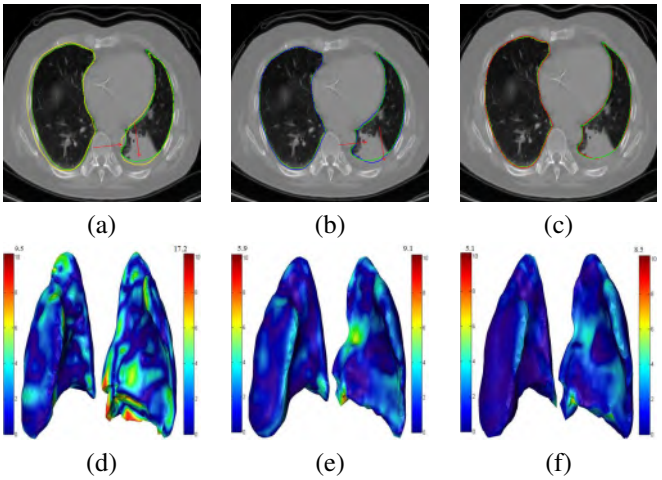


Fig. 10. Comparison between PCA, conventional SSC and ESSC shape prior model. (a) and (d) The refined shape and surface distance to ground truth using PCA, respectively. (b) and (e) The refined shape and surface distance to ground truth using conventional SSC, respectively. (a) and (d) The refined shape and surface distance to ground truth using ESSC, respectively. The color bar maps and clips the surface distance between [0,10]. The maximal surface distance is given above the color bar.

pathological case. In the experiment, only the shape prior model will be changed and other parameters remain the same, including the GVF based vertex searching strategy. Fig.10 (c) shows that ESSC detected more details than the other two models. It means that our shape prior model is able to preserve local shape details derived by appearance information even when they are not statistically significant in the shape repository.

Table III presents the quantitative comparative results by using the three different shape models. ESSC achieved the highest segmentation performance and outperformed the other two shape models when they were employed to refine the deformed shape.

(3) Comparison With Related State-of-the-art Methods

To highlight the advantage of our segmentation method, we compared it with state-of-the-art segmentation approaches: PCA and graph cut proposed by Li *et al*. [9], the shape-based composition proposed by Zhang *et al*. [3], active shape model(ASM) algorithm improved by Sun *et al*. [23], active contour method improved by Rebouccas *et al*. [39] and U-net improved by Alom *et al*. [29] and SegCaps proposed by Lalonde *et al*. [30].

The parameters for the U-net-based method were set as follows. The encoding part had 10 layers, and the decoding part had the same number of layers as the encoding part. The encoding part was down-sampled twice and the decoding part was up-sampled twice. The binary cross entropy loss was used, and the MSE and dice were calculated simultaneously. The total number of training steps was 200 and batch size was 8. The learning rate was decayed by a factor of 0.05 upon validation loss stagnation for 5000 iterations and early stopping was performed with a patience of 250000 iterations based on validation dice scores.

SegCaps was implemented using Keras with TensorFlow. The batch size was 1 to match the original U-Net. The binary cross entropy loss was used, and the MSE and dice were calculated simultaneously. The learning rate was decayed by a factor of 0.05 upon validation loss stagnation for 5000 iterations and early stopping was performed with a patience of 250000 iterations based on validation dice scores. Adam optimization was used with an initial learning rate of 0.00001.

All traditional supervised methods and our method were trained and tested using the same leave-one-out strategy. Each of 78 3D data had 130 slices with lungs, were divided three groups to train (25% for validation) the deep networks. Considered that the efficiency of deep learning was easily affected by the training samples, we augmented the training samples. The data was augmented as follows. Rotation range was 0.2, width shift range was 0.05, height shift range was 0.05, Anti-clockwise angle of shear transformation was 0.05, zoom range was 0.05, random horizontal flip and augmentation ratio was 1:30.

Fig.11 shows three different types of challenging cases. The 1st column shows the case with the blurred boundary of lungs between neighboring tissues in the low-dose CT image. The 3rd column shows the case with movement artifacts due to breathing. The 5th column shows the case with the tumors near the boundary. The 2nd, 4th and 6th columns show surface distance between segmented surfaces and ground truth.

The first row shows segmentation results were obtained by using Li's method [9]. As shown Fig.11(a), local details of the left lung was not segmented and led to undersegmentation. The movement artifact was segmented and led to oversegmentation as shown in Fig.11(c). As shown Fig.11(e), it was difficult to segment lesions connected to the boundary and also led to undersegmentation. Affected by the tumor, vertices cannot be accurately localized.

The second row shows segmentation results were obtained by using Zhang's method [3]. As shown Fig.11(g)(k), local details of the left lung was not segmented and led to under-segmentation. Muscle and pulmonary vessels were segmented as the lungs since SSC can strongly smooth the meshes. As shown Fig.11(i), the lung near the movement artifact was undersegmented. From these three cases, we can see SSC tended to be strongly constrained by its shape model and also led to undersegmentation or oversegmentation. The reconstruction algorithm can not restore the desirable surface.

The third row shows segmentation results were obtained by using Sun's method [23]. Compared to previous two methods, undersegmentation or oversegmentation was smaller. However, it might produce irregular deformation in the local details with large curvature as shown in Fig. 11 (m) and (q).

The fourth row shows segmentation results were obtained by using Rebouccas's method [39]. As shown Fig.11(s) and (u), the lung near the heart, the liver and the movement artifact were undersegmented. As shown Fig.11(w), the pathological changes due to the tumor was undersegmented.

The fifth row shows segmentation results were obtained by using Alom's method [29]. Compared with the traditional method, U-net can segment normal lungs accurately even with large curvature. As shown in Table V, U-net acquired the optimal MSD and FPF. U-net ignored the constraints of shape so that tumors and areas similar to the background intensity

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging

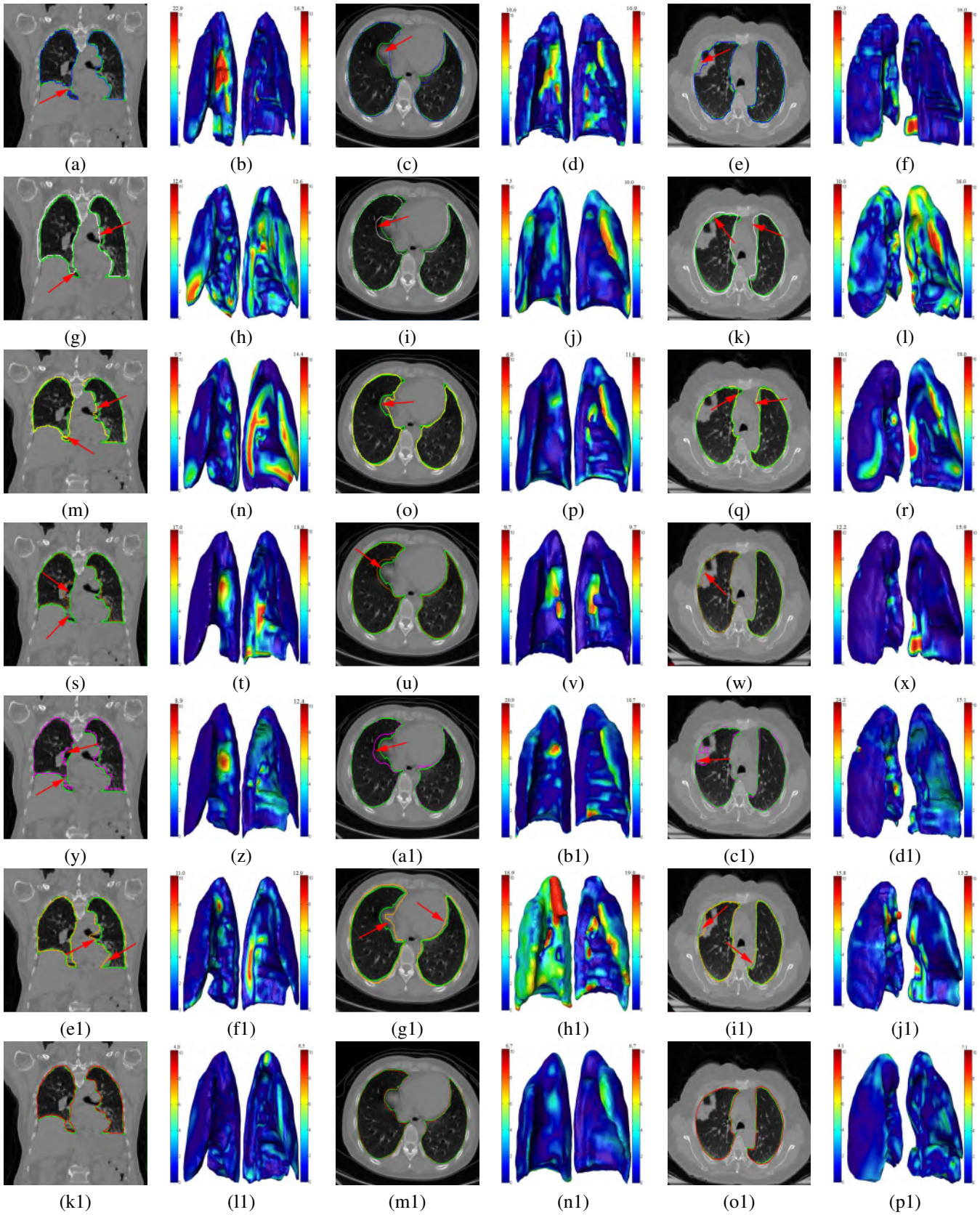


Fig. 11. Segmentation comparison between Li's method [9] (1st row), Zhang's method [3] (2nd row), Sun's method [23] (3rd row), Rebouccas's method [39] (4rd row), Alom's method [29] (5rd row), Lalonde's method [30] (6rd row) and our method (7th row). The 1st column shows the case with the blurred boundary. The 3rd column shows the case with movement artifacts. The 5th column shows the case with the tumors near the boundary. The 2nd, 4th and 6th columns show surface distance between the segmented surfaces and ground truth, and the color bar maps and clips the surface distance between [0,10]. The maximal surface distance is given above the color bar.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMI.2018.2890510, IEEE Transactions on Medical Imaging

TABLE V
COMPARISON OF THE PROPOSED METHOD TO RELATED METHODS (MEAN ± SD).

Method	Lung	ASD(mm)	MSD(mm)	TPF(%)	FPF(%)	DSC(%)
Li *et al.* [9]	Left	2.62 ± 0.91	24.97 ± 9.20	94.08 ± 3.89	0.18 ± 0.11	94.30 ± 2.25
	Right	2.54 ± 0.65	24.91 ± 7.43	93.92 ± 2.35	0.15 ± 0.12	95.20 ± 1.15
Zhang *et al.* [3]	Left	2.94 ± 0.74	26.60 ± 7.15	92.77 ± 2.56	0.17 ± 0.04	94.13 ± 1.54
	Right	2.70 ± 0.71	25.18 ± 7.50	93.78 ± 1.96	0.15 ± 0.03	95.12 ± 1.23
Sun *et al.* [23]	Left	3.18 ± 0.83	25.01 ± 7.50	94.90 ± 2.53	0.25 ± 0.13	93.42 ± 1.87
	Right	2.31 ± 0.93	19.89 ± 8.07	96.41 ± 2.83	0.21 ± 0.09	95.64 ± 1.40
Rebouccas *et al.* [39]	Left	2.75 ± 1.18	24.22 ± 7.92	94.17 ± 3.33	0.19 ± 0.04	94.12 ± 2.97
	Right	2.23 ± 1.02	19.87 ± 7.20	94.47 ± 2.47	0.19 ± 0.05	95.87 ± 2.14
Alom *et al.* [29]	Left	2.02 ± 0.67	$\mathbf{16.78 \pm 6.11}$	92.60 ± 3.22	$\mathbf{0.09 \pm 0.12}$	94.68 ± 2.79
	Right	2.25 ± 1.11	$\mathbf{14.98 \pm 6.91}$	93.59 ± 2.56	$\mathbf{0.09 \pm 0.13}$	95.62 ± 1.52
Lalonde *et al.* [30]	Left	2.91 ± 0.80	23.06 ± 5.47	96.48 ± 2.61	0.35 ± 0.15	93.63 ± 1.79
	Right	2.87 ± 0.95	24.22 ± 6.65	97.61 ± 1.99	0.49 ± 0.31	93.99 ± 1.79
Our	Left	$\mathbf{1.71 \pm 0.61}$	19.00 ± 9.23	$\mathbf{96.89 \pm 2.17}$	0.15 ± 0.05	$\mathbf{96.39 \pm 1.34}$
	Right	$\mathbf{1.37 \pm 0.44}$	15.21 ± 6.39	$\mathbf{97.79 \pm 1.57}$	0.14 ± 0.05	$\mathbf{97.25 \pm 0.92}$

TABLE VI
P VALUE COMPARISON OF THE PROPOSED METHOD TO RELATED METHODS.

Method	Lung	ASD	MSD	TPF	FPF	DSC
Li *et al.* [9] vs. our	Left	1.86×10^{-5}	8.37×10^{-4}	3.05×10^{-4}	0.054	3.16×10^{-7}
	Right	1.70×10^{-10}	6.58×10^{-11}	2.83×10^{-12}	0.50	9.58×10^{-11}
Zhang *et al.* [3] vs. our	Left	4.14×10^{-5}	2.05×10^{-3}	7.88×10^{-5}	5.52×10^{-6}	182×10^{-4}
	Right	1.86×10^{-5}	3.82×10^{-5}	2.47×10^{-6}	5.75×10^{-3}	1.31×10^{-5}
Sun *et al.* [23] vs. our	Left	7.52×10^{-12}	7.40×10^{-4}	3.27×10^{-9}	6.39×10^{-7}	7.96×10^{-12}
	Right	1.69×10^{-10}	3.33×10^{-4}	2.10×10^{-5}	2.26×10^{-6}	8.20×10^{-11}
Rebouccas *et al.* [39] vs. our	Left	1.23×10^{-6}	1.99×10^{-4}	1.92×10^{-8}	1.83×10^{-12}	6.02×10^{-6}
	Right	2.30×10^{-5}	1.26×10^{-4}	1.20×10^{-8}	3.39×10^{-9}	5.21×10^{-5}
Alom *et al.* [29] vs. our	Left	0.03	0.33	2.15×10^{-10}	1.33×10^{-9}	7.67×10^{-6}
	Right	6.48×10^{-5}	0.56	2.28×10^{-10}	2.64×10^{-7}	3.64×10^{-6}
Lalonde *et al.* [30] vs. our	Left	5.36×10^{-10}	7.54×10^{-3}	2.17×10^{-9}	2.31×10^{-9}	1.61×10^{-6}
	Right	9.21×10^{-10}	5.61×10^{-7}	2.76×10^{-5}	6.94×10^{-7}	1.18×10^{-9}

were undersegmented as shown in Fig. 11 (y) and (c1).

The sixth row shows segmentation results were obtained by using Lalonde's method [30]. As shown Fig.11(e1) and (g1), the lung near the heart, the liver and the movement artifact were oversegmented. This was in accord with Table V where FPF of the method was the largest. As shown Fig.11(i1), the pathological changes due to the tumor was undersegmented.

The last row shows segmentation results were obtained by using our method. These figures indicated that our method had a good performance these challenging cases. Table V and Table VI show the quantitatively comparative results by using five methods. Our method achieved the highest segmentation performance except for MSD and FPF, and outperformed the other six methods.

E. Running time

The proposed algorithm was implemented in c++ and test on a PC with Intel i5-3470 CPU@3.20GHz and 8GB of RAM. All deep learning methods involved in the comparison experiment were trained using the NVIDIA Tesla K40m GPU with 12G memory. The average running time of the algorithm was $310 \pm$ 29s. The average running time of the Li's algorithm was $586 \pm$ 69s. The average running time of the Zhang's algorithm was 230 ± 37s. The average running time of the Sun's algorithm was 360 ± 35s. The average running time of the Rebouccas's algorithm was 189 ± 32s. The average running time of the Alom's algorithm was 7 ± 0.2s. The average running time of the Lalonde's algorithm was 48 ± 0.2s.

V. CONCLUSION

In this paper, we have proposed an automatic segmentation scheme for pathological lung segmentation with the combination of 3D shape and appearance based prior information. After the initial mesh is adapted, DAD is employed to effectively label voxels. A refinement model integrating variation models by PCA and SSC is used to achieve both robust and accurate segmentation results. The refinement process is implemented in a uniform average space. All training shapes are registered to the average shape of the PDM using similarity transformation to construct shape dictionary for sparse shape composition in the eigenspace. On the one hand, it is helpful to reconstruct the overall lung shape by SSC. On the other hand, the local reconstruction error can be precisely fitted by the shape variation mode. The deformed mesh can be adapted to the target boundary and is further inferred by ESSC. Then, 3D GVF is also proposed to impose the gradient information of appearance and effectively drive the vertices deformation.

To verify the applicability of the proposed method, various pathological lungs were segmented in 78 3D low-dose CT images with lung tumors. The proposed method is able to deal with different challenging cases, such as the blurred boundary, movement artifacts, tumors near the boundary, sharp structures

and experimental results show the proposed method performs better than state-of-the-art lung segmentation methods.

REFERENCES

[1] W. Jia, T. Aguilera, D. Shultz, M. Gudur, D. L. Rubin, B. W. L. Jr, M. Diehn, and R. Li, "Early-stage noncsmall cell lung cancer: Quantitative imaging characteristics of 18f fluorodeoxyglucose pet/ct allow prediction of distant metastasis," *Radiology*, vol. 281, no. 1, p. 151829, 2016.

[2] S. G. Armato and W. F. Sensakovic, "Automated lung segmentation for thoracic ct impact on computer-aided diagnosis," *Academic Radiology*, vol. 11, no. 9, pp. 1011–21, 2004.

[3] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and S. Z. Xiang, "Towards robust and effective shape modeling: Sparse shape composition," *Medical Image Analysis*, vol. 16, no. 1, p. 265, 2012.

[4] Y. Shao, Y. Gao, Y. Guo, Y. Shi, X. Yang, and D. Shen, "Hierarchical lung field segmentation with joint shape and appearance sparse learning," *IEEE transactions on medical imaging*, vol. 33, no. 9, pp. 1761–1780, 2014.

[5] S. Zhang, Y. Zhan, and D. N. Metaxas, "Deformable segmentation via sparse representation and dictionary learning," *Medical Image Analysis*, vol. 16, no. 7, pp. 1385–1396, 2012.

[6] Q. Li, T. Deng, and W. Xie, "Active contours driven by divergence of gradient vector flow," *Signal Processing*, vol. 120, pp. 185–199, 2016.

[7] D. Xiang, U. Bagci, C. Jin, F. Shi, W. Zhu, J. Yao, M. Sonka, and X. Chen, "Cortexpert: A model-based method for automatic renal cortex segmentation," *Medical Image Analysis*, vol. 42, pp. 257–273, 2017.

[8] T. Tong, R. Wolz, Z. Wang, Q. Gao, K. Misawa, M. Fujiwara, K. Mori, J. V. Hajnal, and D. Rueckert, "Discriminative dictionary learning for abdominal multi-organ segmentation," *Medical image analysis*, vol. 23, no. 1, pp. 92–104, 2015.

[9] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, and D. Xiang, "Automatic liver segmentation based on shape constraints and deformable graph cut in ct images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, p. 5315, 2015.

[10] A. Mansoor, U. Bagci, B. Foster, Z. Xu, G. Z. Papadakis, L. R. Folio, J. K. Udupa, and D. J. Mollura, "Segmentation and image analysis of abnormal lungs at ct: Current approaches, challenges, and future trends," *Radiographics*, vol. 35, no. 4, pp. 1056–1076, 2015.

[11] E. M. van Rikxoort and B. van Ginneken, "Automated segmentation of pulmonary structures in thoracic computed tomography scans: a review," *Physics in Medicine & Biology*, vol. 58, no. 17, p. R187, 2013.

[12] D. Xiang, H. Tian, X. Yang, F. Shi, W. Zhu, H. Chen, and X. Chen, "Automatic segmentation of retinal layer in oct images with choroidal neovascularization," *IEEE Transactions on Image Processing*, vol. 27, no. 12, pp. 5880–5891, Dec 2018.

[13] M. N. Prasad, M. S. Brown, S. Ahmad, F. Abtin, J. Allen, I. da Costa, H. J. Kim, M. F. McNitt-Gray, and J. G. Goldin, "Automatic segmentation of lung parenchyma in the presence of diseases based on curvature of ribs," *Academic radiology*, vol. 15, no. 9, pp. 1173–1180, 2008.

[14] J. Wang, F. Li, and Q. Li, "Automated segmentation of lungs with severe interstitial lung disease in ct," *Medical physics*, vol. 36, no. 10, pp. 4592–4599, 2009.

[15] A. Mansoor, U. Bagci, Z. Xu, B. Foster, K. N. Olivier, J. M. Elinoff, A. F. Suffredini, J. K. Udupa, and D. J. Mollura, "A generic approach to pathological lung segmentation," *IEEE Transactions on Medical Imaging*, vol. 34, no. 1, p. 354, 2015.

[16] E. M. van Rikxoort, B. de Hoop, M. A. Viergever, M. Prokop, and B. van Ginneken, "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection," *Medical physics*, vol. 36, no. 7, pp. 2934–2947, 2009.

[17] I. Sluimer, M. Prokop, and B. van Ginneken, "Toward automated segmentation of the pathological lung in ct," *IEEE Transactions on Medical Imaging*, vol. 24, no. 8, pp. 1025–1038, 2005.

[18] C. Shi, Y. Cheng, J. Wang, Y. Wang, K. Mori, and S. Tamura, "Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation," *Medical image analysis*, vol. 38, pp. 30–49, 2017.

[19] O. Ecabert, J. Peters, H. Schramm, C. Lorenz, J. von Berg, M. J. Walker, M. Vembar, M. E. Olszewski, K. Subramanyan, G. Lavi *et al.*, "Automatic model-based segmentation of the heart in ct images," *IEEE transactions on medical imaging*, vol. 27, no. 9, p. 1189, 2008.

[20] F. Wenzel, C. Meyer, T. Stehle, J. Peters, S. Siemonsen, C. Thaler, L. Zagorchev, A. D. N. Initiative *et al.*, "Rapid fully automatic segmentation of subcortical brain structures by shape-constrained surface adaptation," *Medical image analysis*, vol. 46, pp. 146–161, 2018.

[21] A. A. Kiaei and H. Khotanlou, "Segmentation of medical images using mean value guided contour," *Medical Image Analysis*, vol. 40, pp. 111–132, 2017.

[22] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models‐their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[23] S. Sun, C. Bauer, and R. Beichel, "Automated 3-d segmentation of lungs with lung cancer in ct data using a novel robust active shape model approach," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 449–460, 2012.

[24] C. Santiago, J. C. Nascimento, and J. S. Marques, "2d segmentation using a robust active shape model with the em algorithm," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2592–2601, 2015.

[25] M. Sofka, J. Wetzl, N. Birkbeck, J. Zhang, T. Kohlberger, J. Kaftan, J. Declerck, and S. K. Zhou, "Multi-stage learning for robust lung segmentation in challenging ct volumes," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2011, pp. 667–674.

[26] M. M. Farhangi, H. Frigui, A. Seow, and A. A. Amini, "3-d active contour segmentation based on sparse linear combination of training shapes (scots)," *IEEE transactions on medical imaging*, vol. 36, no. 11, pp. 2239–2249, 2017.

[27] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[28] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[29] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation," *arXiv preprint arXiv:1802.06955*, 2018.

[30] R. LaLonde and U. Bagci, "Capsules for object segmentation," *arXiv preprint arXiv:1804.04241*, 2018.

[31] S. Jégou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*. IEEE, 2017, pp. 1175–1183.

[32] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," in *Conference on Computer Graphics and Interactive Techniques*, 1987, pp. 163–169.

[33] T. Heimann, I. Oguz, I. Wolf, M. Styner, and H. P. Meinzer, "Implementing the automatic generation of 3d statistical shape models with itk," *ISC/NA-MIC Workshop on Open Science at MICCAI*, 2006.

[34] T. Tong, R. Wolz, P. Coupé, J. V. Hajnal, D. Rueckert, A. D. N. Initiative *et al.*, "Segmentation of mr images via discriminative dictionary learning and sparse coding: application to hippocampus labeling," *NeuroImage*, vol. 76, pp. 11–23, 2013.

[35] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *Siam Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[36] R. T. Whitaker and X. Xue, "Variable-conductance, level-set curvature for image denoising," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 3, 2001, pp. 142–145.

[37] K. Khoshelham, "Extending generalized hough transform to detect 3d objects in laser range data," *International Society for Photogrammetry and Remote Sensing*, vol. xxxvi, 2007.

[38] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on image processing*, vol. 7, no. 3, pp. 359–369, 1998.

[39] P. P. Rebouças Filho, P. C. Cortez, A. C. da Silva Barros, V. H. C. Albuquerque, and J. M. R. Tavares, "Novel and powerful 3d adaptive crisp active contour method applied in the segmentation of ct lung images," *Medical image analysis*, vol. 35, pp. 503–516, 2017.