Spline curve deformation model with prior shapes for identifying adhesion boundaries between large lung tumors and tissues around lungs in CT images

Xin Zhang and Jie Wang
College of Electronic Information Engineering, Hebei University, Hebei Baoding 071000, China

Ying Yang
Hebei University Affiliated Hospital, Hebei Baoding 071000, China

Bing Wang[a)]
College of Mathematics and Information Science, Hebei University, Hebei Baoding 071000, China

Lixu Gu[a)]
School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200000, China

(Received 29 January 2019; revised 18 November 2019; accepted for publication 2 December 2019; published 28 January 2020)

Purpose: Automated segmentation of lung tumors attached to anatomic structures such as the chest wall or mediastinum remains a technical challenge because of the similar Hounsfield units of these structures. To address this challenge, we propose herein a spline curve deformation model that combines prior shapes to correct large spatially contiguous errors (LSCEs) in input shapes derived from image-appearance cues.The model is then used to identify the adhesion boundaries between large lung tumors and tissue around the lungs.

Methods: The deformation of the whole curve is driven by the transformation of the control points (CPs) of the spline curve, which are influenced by external and internal forces. The external force drives the model to fit the positions of the non-LSCEs of the input shapes while the internal force ensures the local similarity of the displacements of the neighboring CPs. The proposed model corrects the gross errors in the lung input shape caused by large lung tumors, where the initial lung shape for the model is inferred from the training shapes by shape group-based sparse prior information and the input lung shape is inferred by adaptive-thresholding-based segmentation followed by morphological refinement.

Results: The accuracy of the proposed model is verified by applying it to images of lungs with either moderate large-sized (ML) tumors or giant large-sized (GL) tumors. The quantitative results in terms of the averages of the dice similarity coefficient (DSC) and the Jaccard similarity index (SI) are 0.982 ± 0.006 and 0.965 ± 0.012 for segmentation of lungs adhered by ML tumors, and 0.952 ± 0.048 and 0.926 ± 0.059 for segmentation of lungs adhered by GL tumors, which give 0.943 ± 0.021 and 0.897 ± 0.041 for segmentation of the ML tumors, and 0.907 ± 0.057 and 0.888 ± 0.091 for segmentation of the GL tumors, respectively. In addition, the bidirectional Hausdorff distances are 5.7 ± 1.4 and 11.3 ± 2.5 mm for segmentation of lungs with ML and GL tumors, respectively.

Conclusions: When combined with prior shapes, the proposed spline curve deformation can deal with large spatially consecutive errors in object shapes obtained from image-appearance information. We verified this method by applying it to the segmentation of lungs with large tumors adhered to the tissue around the lungs and the large tumors. Both the qualitative and quantitative results are more accurate and repeatable than results obtained with current state-of-the-art techniques. [https://doi.org/10.1002/mp.13998]

Key words: lung tumor segmentation, shape group-based sparse prior models, spline curve deformation model, tumors attached to normal structures

1. INTRODUCTION

Accurate delineation of the lesions in computed tomography (CT) images is a fundamental and challenging problem in computer-assisted diagnosis systems. Highly accurate, robust, and automated techniques are available to segment tumors that appear as homogenous, high-intensity solid lesions with clear boundaries against a background of low-intensity lung parenchyma.[1–4] However, when tumors (especially large tumors) are adhered to the anatomical structure around the lung, such as the chest wall or mediastinum, tumor segmentation remains a technical challenge because of the similar intensity of the tumor and the normal attached organs and the irregular shape of large tumors. Figure 1 shows typical

examples of lung-boundary adhesion large tumors, where a large tumor remains confined to the damaged lung or if the ratio of the tumor area to the damaged lung area is much less than unity, it is categorized as a moderate large-sized (ML) tumor; otherwise, it is categorized as a giant large-sized (GL) tumor. In the past, numerous methods have been proposed to mainly segment on small lung masses attached to lung boundaries (e.g., juxta-pleural nodules) including region-growing or watershed followed by morphological processing, supervised segmentation models based on feature extraction and support vector machines or neural networks, and energy minimization models with prior knowledge.[5–25] With the intensive investigation of sparse representation, the sparse constraints have been used to model shape priors, which explicitly model spatially contiguous gross errors (non-Gaussian errors) in input shapes with sparse vectors and assume that these gross errors are sparse with respect to the given shape information.[26–30] However, this assumption does not hold when very large tumors are attached to lung boundaries because the non-Gaussian errors caused by these tumors are no longer sparse relative to the input lung shapes.

In this paper, we first propose a spline curve deformation model that combines prior shapes to correct large spatially contiguous errors (LSCEs) in input shapes derived from image-appearance cues. The external and internal forces drive the control points (CPs) of the model to fit the targets. The model is then used to identify the adhesion boundaries between large lung tumors and tissue around the lungs as shown in Fig. 1. Depending on tumor size, the initial shapes (prior shapes) of the deformation are automatically inferred by the sparse-shape group composition (SGC) model or by the sparse similar-shape linear combination (SLC) model.

2. MATERIALS AND METHODS

2.A. Spline curve deformation model

B-spline algorithms have become the standard for computational geometric representation because they offer the advantages of local modification, continuity, and strong convex hull to give finer shape control.[31,32] In our method, the B-spline algorithm is used to fit the transformed CPs on deformation model to a curve, which not only drives the whole curve to deform but also is able to ensure the smoothness of the whole shape, where the deformation model is to drive the CPs to match observations (e.g., the boundary points of an object) starting from a prior shape. Suppose that the input shapes containing the LSCEs (i.e., the contours of the objects derived from image data) are the simple shapes that do not intersect themselves and the possible errors on the parts of the non-LSCEs only follow the Gaussian distribution with small variances. The goal of the proposed model is to correct the LSCEs on the input shapes by deforming simple curves to fit the specified positions on the parts of the non-LSCEs of the input shapes (called target control points, TCPs), where the initial shapes of the curves are obtained from the prior shapes of the object. That is, the TCPs serve as the external force to drive the corresponding CPs on the model to match with them. We name the CPs that are expected to match the TCPs as critical control points (CCPs) and the curve segments that contain the CCPs as handles.

The energy function of the external force is denoted as:

$$E_{ext} = \left\| \mathbf{T}\mathbf{V}_c - \mathbf{V}_c^{'} \right\|_2^2, \tag{1}$$

where $\| \cdot \|_2$ is the L2 norm, sparse matrix $\mathbf{T}$ represents a transformation of all the points of the model, $\mathbf{V}_c$ and $\mathbf{V}_c^{'}$ are the column vectors of the CCPs and corresponding TCPs respectively, which are equal in length to T by adding 0 values at non-CCP and TCP points. To obtain the natural and reasonable shapes in the process of the model deformation, we limit transformation to translations, rotations and isotropic scales. Instead of measuring the $\mathbf{T}$'s distances directly, we optimize the $\mathbf{T}$ for all the CPs. In order to design a convex optimization of the energy function, we reformulate the $\mathbf{T}$ to transformation vector $\mathbf{t}$. Thus, the CP k's transformation $T_k v_k$ is substituted with $A_k t_k$, where $\mathbf{A}_k$ contains the coordinates of the CP k. We concatenate the

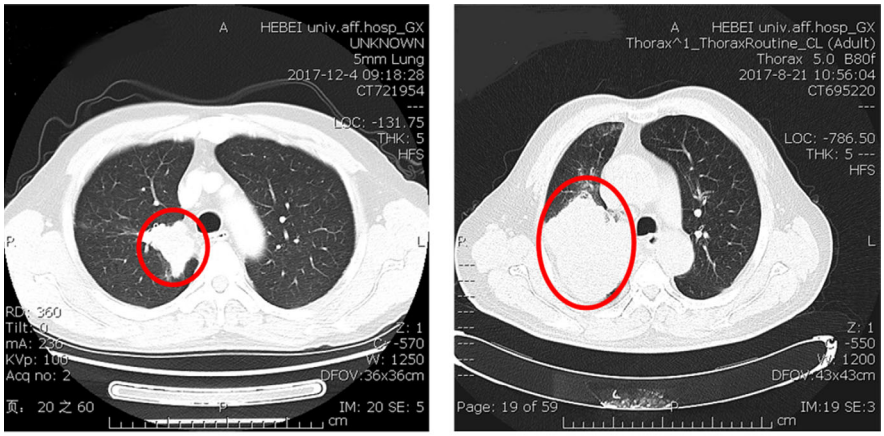

FIG. 1. Typical instances of moderate large-sized (ML) and giant large-sized (GL) lung-boundary adhesion tumors. The tumors are encircled in red: (a) ML tumor; (b) is GL tumor. [Color figure can be viewed at wileyonlinelibrary.com]

coordinates $\mathbf{A}_k$ into a sparse matrix. The matrix of the CCPs is $\mathbf{D}_c = diag(A_1, A_2, \cdots, A_m, 0, \cdots, 0)$. Then, Eq. (1) is formulated as:

$$E_{ext} = \left\|\mathbf{D}_c\mathbf{t} - \mathbf{V}_c^{'}\right\|_2^2. \tag{2}$$

Our internal force maintains the local similarity of the model by regularizing the differences of the transformation matrices between neighboring CPs. So the energy function of the internal force is:

$$E_{int} = \sum_{k\in V}\sum_{\tau\in N(k)} w_{k\tau}\|A_k t_k - A_k t_\tau\|_2^2, \tag{3}$$

where weight $w_{k\tau}$ is the strength of connection between the CPs k and τ, $N(k) \subset \mathrm{N}$ is the neighborhood structure of the CP k. In order to express the energy function of the internal force as a quadratic form of all the transformation vectors $\mathbf{t}$ like Eq. (2), we encode the matrix A_k into its Kronecker product with matrix S_k as $K_k = S_k \otimes A_k$, where S_k is a $\delta \times n$ matrix of the CP k. For each neighbor CP τ, there is one row in S_k, where the kth item is $w_{k\tau}$ and the τth item is $-w_{k\tau}$, while the other items are all zeros. The energy function of the internal is formulated as:

$$E_{int} = \|\mathbf{K}\mathbf{t}\|_2^2. \tag{4}$$

Thus, the deformation model is defined as:

$$\arg\min_{\mathbf{t}}\left\{\|\mathbf{K}\mathbf{t}\|_2^2 + \omega\left\|\mathbf{D}_c\mathbf{t} - \mathbf{V}_c^{'}\right\|_2^2\right\}, \tag{5}$$

where is a weight that controls to what extent the model curves can deform to match the TCPs. Larger leads to a better matching, but the deformed shape may not be smooth. This equation can be solved by least square minimization.

2.B. Separation of large tumors from attached organs

In this section, we present a robust separation method of large lung tumors from the attached lung-boundary organs in CT images. The proposed spline curve deformation model is employed to correct LSCEs on the input shapes of the lungs caused by the tumors, so as to pinpoint the adhesion boundaries between the tumors and normal tissues. The proposed method involves five major procedures: (a) lung contour extraction, (b) deformation model initialization, (c) handle positioning, (d) spline curve deformation, and (e) tumor segmentation. Figure 2 shows a flowchart of the proposed method. The lung contour is extracted based on image-appearance cues, like adaptive-thresholding-based segmentation followed by morphological refinement. Here even-distributed sparse points on the lung boundary are selected to represent the lung contour. Thus corner points and high curvature points were extracted as the sparse points by corresponding selectors firstly and the rest is obtained using a nearest-neighbor interpolation algorithm.[33] The initial contour of deformation model for a specific slice comes from the lung's shape prior in that slice which is generated by shape group-based sparse prior models (SGC or SLC). The whole tumors are segmented using the random walk method,[34] where target and background seeds are extracted uniformly with the help of a computer-aided system.

2.B.1. Sparse-shape group composition model

Inspired by the sparse-shape-composition (SSC) model to model shape priors,[26] we introduced the spatial relationship of multiple objects into the shape composition model to propose the SGC model.[30] In this work, the SGC model is used to obtain the prior shape of a lung with an ML adhesion tumor. We assume that at least one of the lungs is intact. Thus, the shape groups made with the left and right lungs of the same slice are applied to the model to make the LSCEs sparse with respect to the given shape groups. Thus the training repository for the group of shapes can be represented as a matrix $G = [g_1, g_2, \ldots, g_v]$, where v is the number of training samples, $g_i = \left[\left(s_i^L\right)^T\left(s_i^R\right)^T\right]^T$ is ith shape group in the training samples, which is constructed by left lung contour s_i^L and right lung contour s_i^R in serial mode. Accordingly, the input shape group is expressed as $y_g = \left[y_L^T\, y_R^T\right]^T$ in serial mode. The advantage of the serial mode is that it can directly reserve the patient information in the model and compute the optimal parameters of the model by iteratively applying the expectation step and the maximization step of the expectation-maximization algorithms proposed in Ref. [30].

2.B.2. Model based on linear combination of sparse similar-shape groups

For the contours derived from lungs with GL tumors attached, their prior shapes fail to be inferred based on the SGC model as the LSCEs corresponding to input shape groups are not sparse. Here the SLC model is proposed to overcome this problem, which uses the lung contour without boundary defect to infer the lung contour with boundary defect in the same slice.We represent the lung contour without boundary defect in a given shape group as a reference shape. First, the prior shape of the reference shape is gained by the SSC model, where the shapes that constitute a linear combination of the reference shape are called its linear combination shapes. Meanwhile, the shape groups in test database that contain the linear combination shapes are named as the sparse similar-shape groups of the reference shape. Then, the prior shape of the damaged lung in the same slice as the reference shape is gained by the linear combination of sparse similar-shape groups.

2.B.3. Automatic positioning of handles

The purpose of locating the handles on the initial contour is to generate the TCPs on the input shapes. Because the TCPs are the target points of the CCPs' displacements, we specify them as the intersections between the normal lines

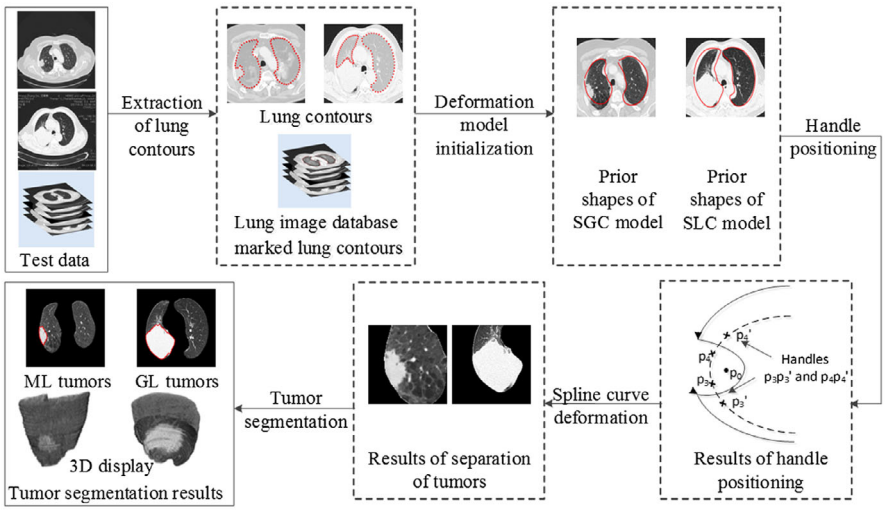

FIG. 2. Flowchart for separating large tumors from attached organs surrounding the lungs. [Color figure can be viewed at wileyonlinelibrary.com]

passing the CCPs on the handles and the input curves. Thus, one end of a handle is fixed to the CPs on the non-LSCE segments, and the other end of the handle is positioned at the intersection between the initial contour and the line that tangents to the input curve through the center point of the large error position. Figure 3 illustrates the handle-position approach. The center point p_0 is in the largest of the regions determined by set $C = A - (A \cap B)$, where sets A and B denote the regions surrounded by the initial shape and input shape, respectively. The tangent lines p_0p_1 and p_0p_2 intersect the initial shape at points p_3 and p_4, respectively. The CPs p_3' and p_4' on the non-LSCE segments are selected as the end points of the handles. Thus, the segments p_3p_3' and p_4p_4' on the initial shape are the specified handles, respectively.

3. EXPERIMENTS AND RESULTS

3.A. Experiment data and environment

To verify the proper functioning of the proposed method, we applied it to three databases. We used the LIDC and the RIDER databases from http://cancerimagingarchive.net, which are the most extensive publicly available collection of annotated CT images. We also collected lung tumor images from the local medical center database with in-plane resolution from 0.5 to 0.95 mm, slice intervals of 5.00 mm, tube voltages of 120 and 140 kVp, and currents from 40 to 340 mA. The LIDC database is used to build shape repositories with slice intervals of 1.00-5.00 mm.

In the image database, slice locations are numbered in ascending order from the top of the lung to the bottom of the lung and each of them contains at least 280 images from different cases.The slices of each test case are also sorted in the same order as the image database, which effectively reduces the search scope of the test slices in the shape repository because only images whose relative positions in the image database are similar to the relative position of the test slice in the designate case are searched to construct the sparse combination representation of the test slices. The interval between the CPs of the model is 5.5–6.0 mm, and each handle has four CCPs.We focused primarily on tumors 40–110 mm in diameter. A total of 185 tumor images were selected, 120 ML tumors and 65 GL tumors, to verify the segmentation of large tumors adhered to lung boundaries. For the ground truth, the boundaries of lungs and tumors in each CT image were marked by three radiologists from different institutions and were cross-checked to guarantee accuracy. The proposed models were implemented on a desktop PC with a Pentium Dual-Core CPU E5800@ 3.20 GHz, 4 GB RAM, a NVIDIA GeForce GT 430 GPU, with MATLAB 2009 on Windows 7.

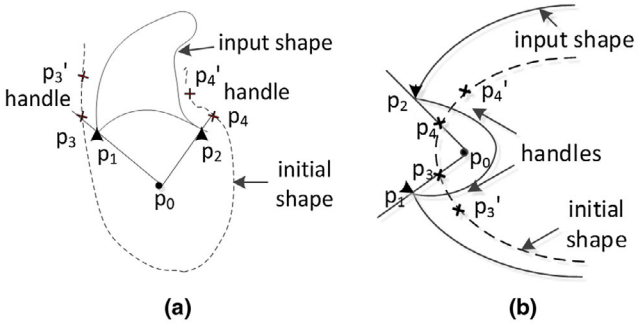

FIG. 3. Schematic illustration showing the method used to locate handles of the model. The segments p_3p_3' and p_4p_4' on the initial shape are the specified handles, respectively: (a) Case of initial shape outside input curve; (b) Case of initial shape inside input curve. [Color figure can be viewed at wileyonlinelibrary.com]

3.B. Comparison of methods and evaluation metrics

We compared the proposed spline curve deformation model with the deformation models based on the image gradient information. In addition, we compared the proposed SGC and SLC with the SSC to verify the role of the shape group in modeling prior shapes of lungs containing large

tumors. To summarize, the following methods were compared:

SGC-I: represents deformation model based on gradient information, where initial shape of the model is generated by SGC model.

SLC-I: represents deformation model based on gradient information, where initial shape of the model is generated by SLC model.

SGC-R: represents deformation model based on gradient information, where SGC model is used as a regularization step in the model.

SLC-SGC-R: represents deformation model based on gradient information, where initial shape is generated by SLC model and SGC model is used as a regularization step in the deformation model.

SGC-SCD: is the proposed spline curve deformation model, where initial shape of the model is generated by SGC model. It is used to correct lung boundaries containing ML tumors.

SLC-SCD: is the proposed spline curve deformation model, where initial shape of the model is generated by SLC model. It is used to correct lung boundaries containing GL tumors.

We use as evaluation metrics the dice similarity coefficient (*DSC*) and the Jaccard similarity index (*SI*) between binary masks to quantitatively compare different methods. We also use as metric the bidirectional Hausdorff distance (*BHD*) to quantify the dissimilarity between the sets of points on the contour for the reference gold standard and the sets of points on the contour of the object being tested.

3.C. Recognition of adhesion boundaries of large tumors and segmentation of lungs and tumors

We first verify the SGC model by using it to model the shape of a lung with an ML adhesion tumor. Figure 4 (in the third row) offers qualitative results of some inferred lung contour cases. It shows that the damaged lungs are well-reconstructed through SGC model. Figure 5 shows the results of the segmentation by four different methods of two representative examples of lungs with ML adhesion tumors. Although SGC generates better initial constraint shapes, SGC-I does not segment accurately in the presence of large weak-appearance cues. The segmentation by SGC-R is better than that by SGC-I. However, the air pipe near the lung is incorrectly included in the lung contour (see red circle). The proposed SGC-SCD model gives correct segmentation results.

Next, we test the validity of the SLC model by using it to determine the shape of a lung with a GL adhesion tumor. Figure 6 shows that SGC cannot also capture such gross errors in e, whereas SLC can approximately recover these missing lung shapes. Figure 7 shows the segmentation results of different methods applied to two representative examples. SGC-I and SGC-R poorly segment these severely damaged lungs. Although SLC provides approximate initialization results, limited improvements are achieved by using the segmentation

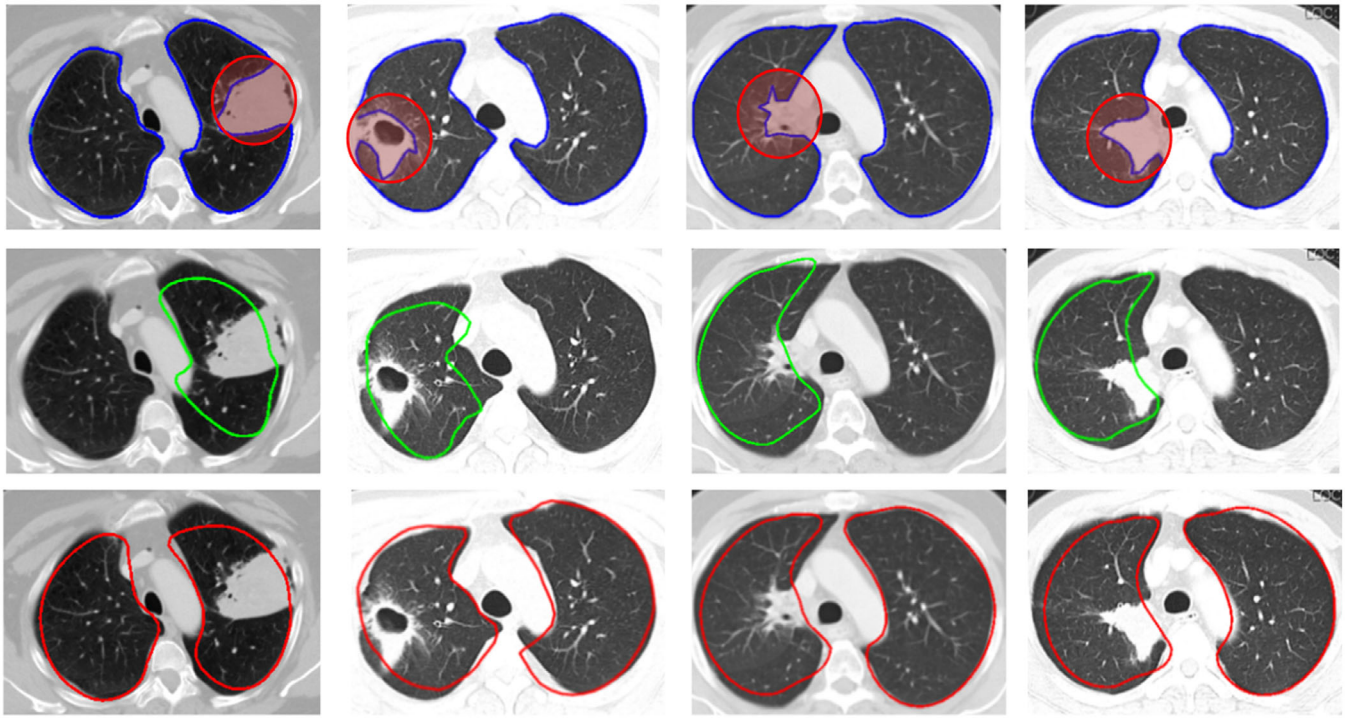

FIG. 4. Comparisons of contour inferences produced by sparse-shape group composition (SGC) and sparse-shape-composition (SSC) for lungs with moderate large-sized adhesion tumors. The first row shows the lungs segmented by using image-appearance information (blue curves). A single gross error appears in each example (red circles).The second row shows the inferred shapes (green curves) of lungs containing large tumors based on SSC. The third row shows the inferred shape of the lung (red curves) based on SGC. [Color figure can be viewed at wileyonlinelibrary.com]

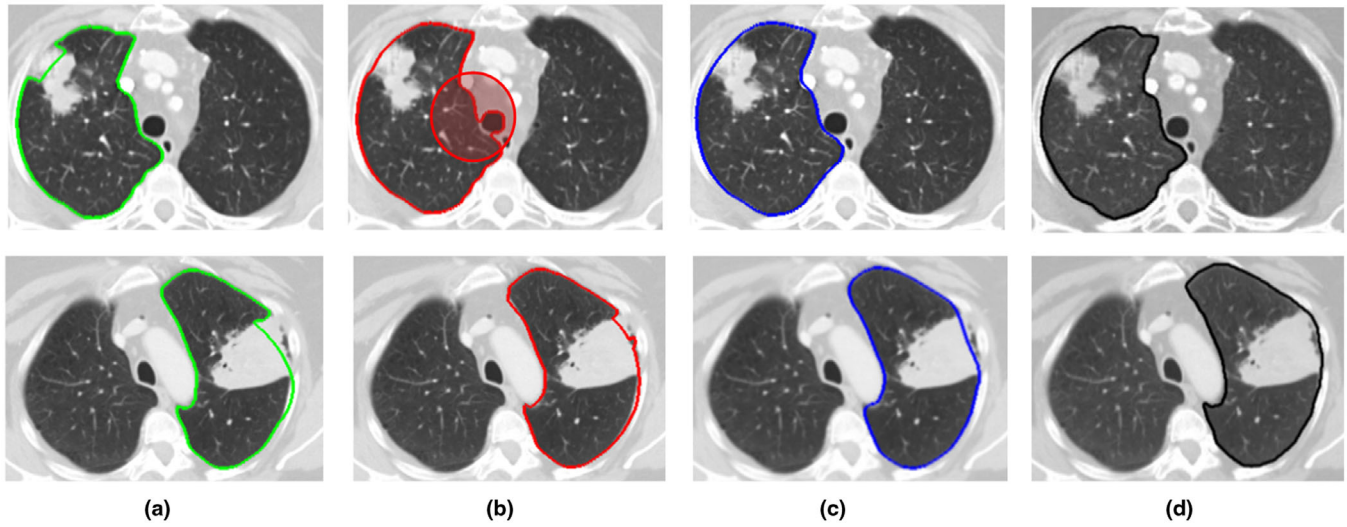

FIG. 5. Comparison of lungs with moderate large-sized adhesion tumors and segmented by using three different methods: (a) sparse-shape group composition (SGC)-I; (b) SGC-R; (c) Proposed method (SGC-SCD); (d) Ground truth. [Color figure can be viewed at wileyonlinelibrary.com]

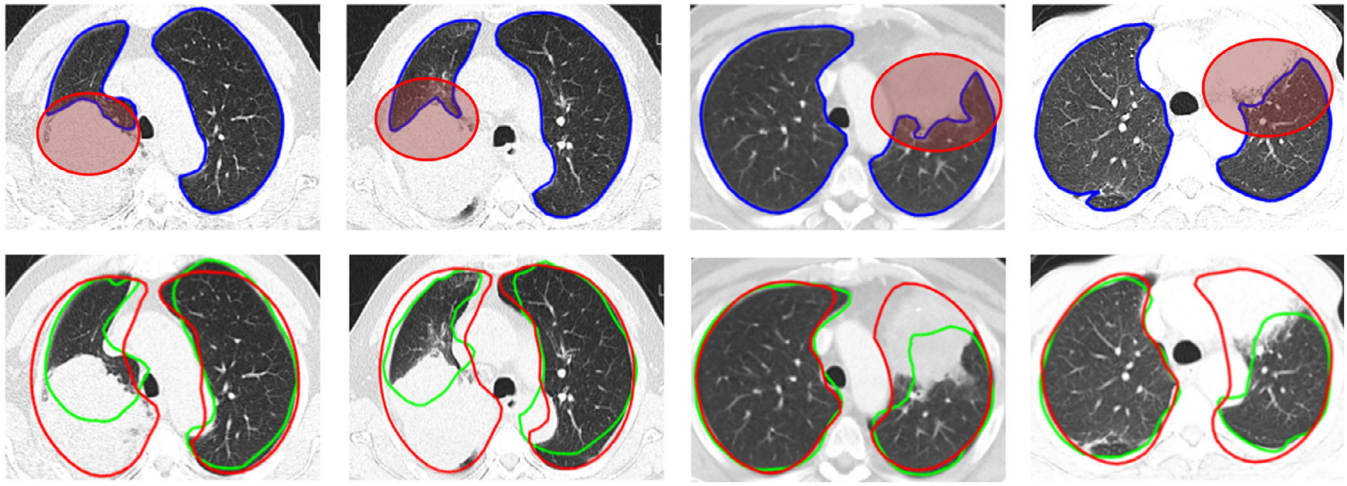

FIG. 6. Comparison of contour inference obtained by applying sparse-shape group composition (SGC) and sparse similar-shape linear combination (SLC) to lungs with GL adhesion tumors. The first row shows lungs segmented by using image-appearance information (blue curves). A single gross error appears in each example (see ellipses).The second row shows the inferred lung shapes obtained by using SGC (green curves) and SLC (red curves). [Color figure can be viewed at wileyonlinelibrary.com]

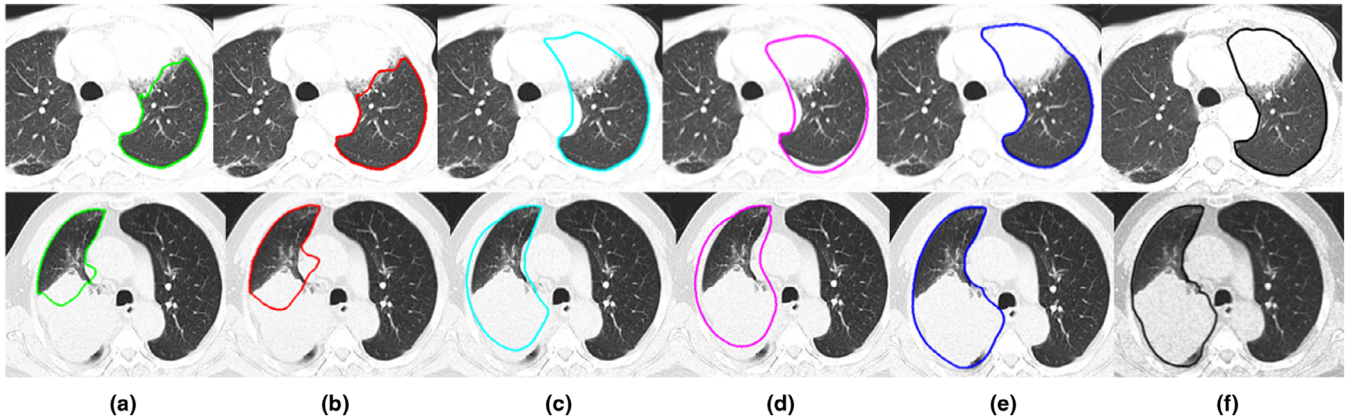

FIG. 7. Comparison of lungs with GL adhesion tumors segmented by different methods: (a) sparse-shape group composition (SGC)-I; (b) SGC-R; (c) Sparse similar-shape linear combination (SLC)-I; (d) SLC-SGC-R; (e) SLC-SCD; (f) Ground truth. [Color figure can be viewed at wileyonlinelibrary.com]

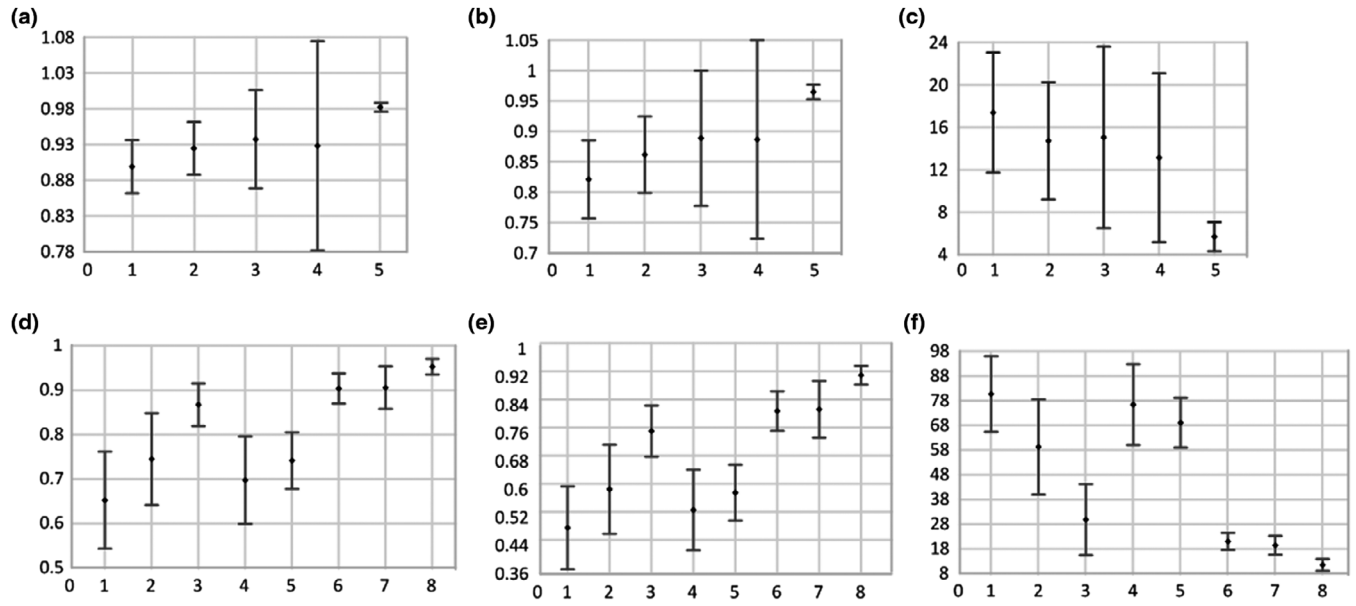

FIG. 8. Mean values and standard deviations of dice similarity coefficient (*DSC*), similarity index (*SI*), and bidirectional Hausdorff distance (*BHD*) from left or right lungs with moderate large-sized (ML) or gaint large-sized (GL) adhesion tumors. The abscissa corresponds to the different segmentation methods. For the lungs with ML tumors (top row), the methods ordered from left to right on the abscissa are sparse-shape-composition (SSC), sparse-shape group composition (SGC), SGC-I, SGC-R, and SGC-SCD. For the lungs with GL tumors (bottom row), the order from left to right on the abscissa is SSC, SGC, SLC, SGC-I, SGC-R, SLC-I, SLC-SGC-R, and SLC-SCD: (a) *DSC* of lung with ML tumors; (b) *SI* of lung with ML tumors; (c) *BHD* of lung with ML tumors; (d) *DSC* of lung with GL tumors; (e) *SI* of lung with GL tumors; (f) *BHD* of lung with GL tumors.

TABLE I. Comparison of mean dice similarity coefficient (*DSC*) for segmentation of lungs.

Methods	This work	Ref. [9]	Ref. [14]	Ref. [15]
DSC	0.972	0.968	0.964	0.766

methods based on SLC-I and SLC-SGC-R. The lung shapes provided by the proposed method are more accurate.

Figure 8 summarizes the segmentation results of the left or right lung containing ML or GL adhesion tumors in terms of different methods. The mean values and standard deviations of *DSC*, *SI*, and *BHD* of SGC-SCD method for lungs containing ML tumors are 0.982 ± 0.006, 0.965 ± 0.012, 5.7 ± 1.4 mm, respectively. The mean values and standard deviations of *DSC*, *SI*, and *BHD* of SLC-SCD method for lungs containing GL tumors are 0.952 ± 0.048, 0.926 ± 0.059, 11.3 ± 2.5 mm, respectively. Although the rest of the SGC-based methods (SGC, SGC-I, SGC-R, and SGC-SCD) for lungs containing ML tumors and SLC-based methods (SLC, SLC-I, SLC-R, and SLC-SCD) for lungs containing GL tumors also have higher segmentation scores in any evaluation metric than SSC method, the standard deviations of these method are much larger than our proposed methods.

We also compared the proposed method with other methods from the literature. Table I compares the DSC results for lungs with ML or GL tumors obtained by using the methods from Refs. 1,5,12. The results show that the proposed method ranks first out of all the methods tested.

Finally, in Fig. 9, we present typical segmentation results for different types of large tumors adhered to normal tissue around the lungs. The top row to the bottom row show the results for a right mediastinal adhesion tumor, left mediastinal adhesion tumor, right lung pleural adhesion tumor, heterogeneous right lung pleural adhesion tumor, left lung pleural adhesion tumor, right lung GL adhesion tumor, and left lung GL adhesion tumor, respectively. Table II presents the quantitative evaluations based on *DSC* and *SI* of the segmentation results for lungs, lung segments attached by tumors, and the different types of tumors are shown in Fig. 9. The results of the proposed approach for segmentation of lungs and for tumors adhered to normal tissue around lungs are encouraging.

4. DISCUSSION

We have shown that the proposed method can achieve accurate segmentation on ML and GL tumors. Its most distinctive features include (a) the adoption of spline curve deformation model to correct large spatially contiguous errors in lung shapes for separating large tumors from the adhesion lung boundaries and (b) the inference of severely damaged lung shapes that takes advantage of sparse-shape group models. Besides, the input shapes of the deformation model can be derived from image-appearance cues and the target control points driving the model to the adhesion boundaries between objects can also be obtained conveniently, so the whole segmentation model has high accuracy and low complexity compared with the traditional gradient-based deformation model.

Figure 4 shows some representative and challenging cases. The results indicate that the SGC method captures more of

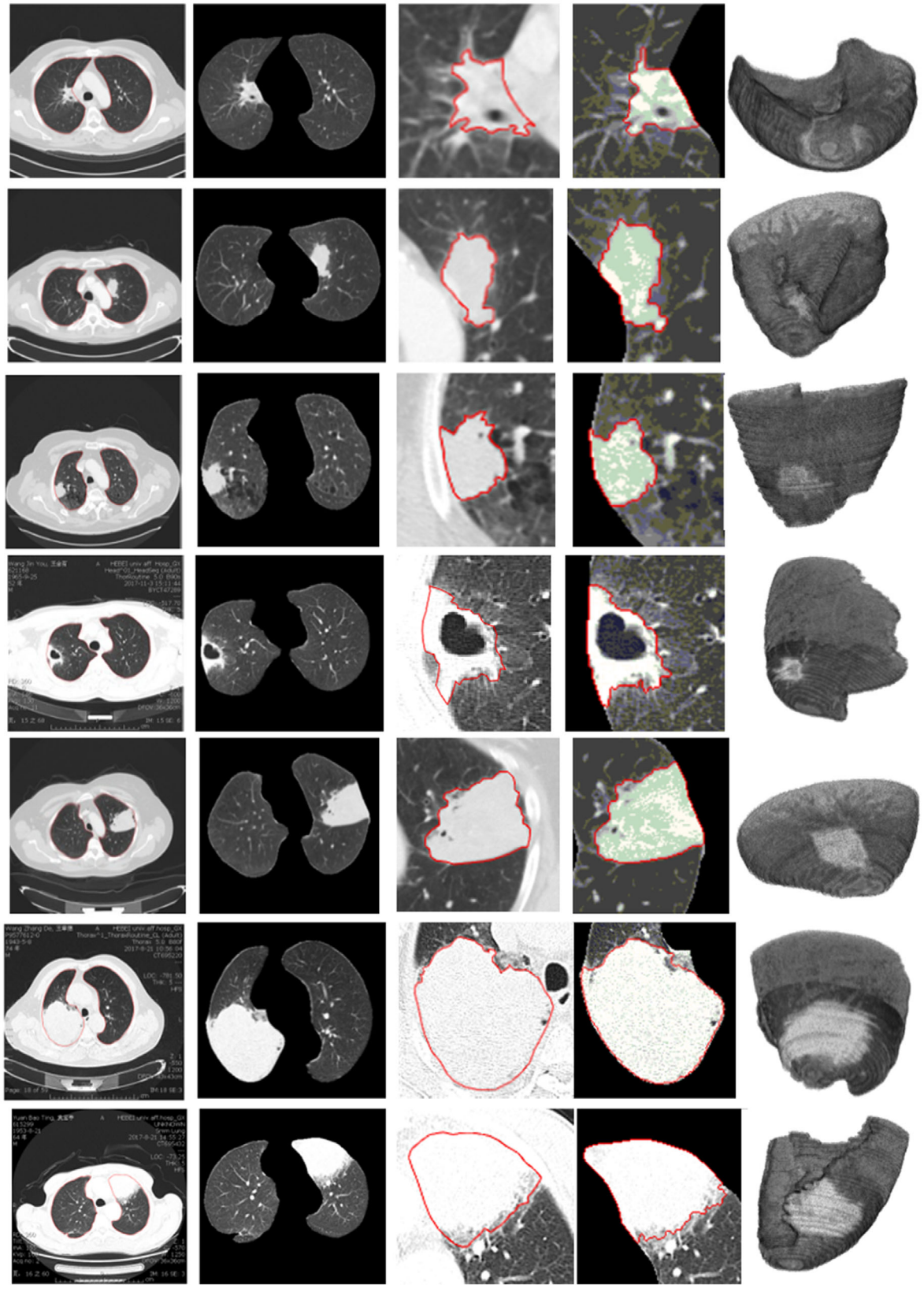

FIG. 9. Typical segmentation results for different types of large tumors adhered to normal tissue around lungs. The first column shows the ground truth for lungs (drawn by curves), the second column shows the segmentation results for damaged lungs, the third column shows the ground truth for tumors, the fourth column shows the segmentation results for tumors, and the fifth column shows the three-dimensional delineations. The top row to the bottom row correspond to a right mediastinal adhesion tumor, a left mediastinal adhesion tumor, a right lung pleural adhesion tumor, a heterogeneous right lung pleural adhesion tumor, a left lung pleural adhesion tumor, a right lung large adhesion tumor, and a left lung large adhesion tumor, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE II. Quantitative evaluation of segmentation of lungs by using dice similarity coefficient (*DSC*) and similarity index (*SI*) for lung segments attached by tumors and for different types of tumors (see Fig. 9).

	DSC			SI		
Image no.	Lung	LS	Tumor	Lung	LS	Tumor
1	0.975	0.979	0.966	0.951	0.959	0.933
2	0.978	0.985	0.934	0.956	0.971	0.875
3	0.985	0.993	0.952	0.970	0.986	0.928
4	0.976	0.972	0.919	0.953	0.946	0.850
5	0.970	0.958	0.870	0.941	0.920	0.782
6	0.963	0.957	0.940	0.938	0.921	0.887
7	0.959	0.943	0.876	0.922	0.892	0.754

Note: LS indicates lung segments attached by tumor.

the large concave errors of the boundaries in the sparse vector e and generates more reasonable shapes than does the SSC model. We attribute this result to the sparsity of the large errors relative to the whole shape group. In conventional SSC models,[20,26–29] gross errors are assumed to be sparse with respect to a single given shape. It is in lack of the ability of capturing LSCEs in input shapes.

However, for lungs with GL tumors attached, compared with SSC and SGC models, only SLC model can effectively reconstruct the missing part of the damaged lungs as shown in Fig. 6. This is because such a tumor encroaches on most of the lung and typically is not even surrounded by the lung it occupies, the segmentation method based on image-appearance information fails to find the missing part of the original lung, which provides incorrect shape cues for single shape composition models. In SLC model, the prior shape of a complete lung is a linear combination of relevant shapes obtained by SSC algorithm, while the prior shape of the incomplete lung in the same slice is a linear combination of other shapes that belong to the same shape group as these relevant shapes. Thus, the model omits the assumption of the sparsity of gross errors.

Note that *DSC* and *SI* always provide good results for all methods for lungs with ML adhesion tumors. This result is attributed to the fact that the adhesion part is relatively small compared with the whole lung. The three methods based on SGC, that is, SGC, SGC-I and SGC-R, perform better than the SCC in term of *BHD* because the spatial structure of two lungs is modeled. However, for lungs with GL adhesion tumors, the methods based on SGC produce poor results because they fail to capture the errors in the sparse vector e. The three methods based on SLC (SLC-I, SLC-SGC-R, and SLC-SCD) produce better results for these specific lungs. In terms of the three metrics, the proposed segmentation methods produce the best results. The main reason leading to these results can be explained as follows. First, the TCPs in input shapes rather than gradient information of images serve as external forces to drive deformation models to object boundaries. In the traditional energy minimization models based on gradient information,[20–25] it is usually required that the boundaries of the objects to be segmented be clear and their intensities vary suddenly in order to be captured by the model. However, the boundaries of lungs attached to large tumors are blurred, which leads to boundary leakage of segmentation results. In the proposed models, the TCPs are always automatically specified by the handles of the models in advance. There are clear external forces to drive the models to deform to the lung boundaries. The second reason is that the sparse-shape group models more effectively reconstruct the prior shapes of the severely damaged lungs than the conventional SSC model, which greatly improves the accuracy of the correction of large errors at the lung boundaries.

5. CONCLUSIONS

We propose a spline curve deformation to deal with large spatially consecutive errors in object shapes obtained from image-appearance information. The deformation of the whole curve is driven by the transformation of the control points in the model, which are influenced by external and internal forces. The proposed model is used to identify large adhesion interfaces between tumors and normal structures around lungs in chest CT images by correcting the gross errors in the lung input shapes caused by large lung tumors. The initial shape for the model is inferred from the training shapes by shape group-based sparse prior information and the input shape is obtained by adaptive-thresholding-based segmentation followed by morphological refinement. The validity of the proposed model is checked on three databases. In future research, we will focus on two issues: how to improve the extraction accuracy and efficiency of curve sections, especially for lungs with GL tumors, and how to extract more information to enhance the fault tolerance of prior shapes.

ACKNOWLEDGMENTS

This research work was partially supported by the Natural Science Foundation of Hebei Province (F2017201172) and Key Project of Hebei Education Department (ZD2018210). The authors are grateful to the clinicians in Hebei University Affiliated Hospital and Hebei Key Laboratory of Machine Learning and Computational Intelligence for their support in providing data and clinical advice.

CONFLICT OF INTEREST

The authors have no conflict to disclose.

[a)]Author to whom correspondence should be addressed. Electronic mails: wangbing@hbu.edu.cn, gulixu@sjtu.edu.cn.

REFERENCES

1. Johnson PB, Young LA, Lamichhane N, Patel V, Chinea FM, Yang F. Quantitative imaging: correlating image features with the segmentation accuracy of PET based tumor contours in the lung. *Radiother Oncol*. 2017;123:257–262.

2. Setio AAA, Traverso A, de Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. *Med Image Anal*. 2017;2017:1–13.
3. Le Y, Yang D, Zhu Y, et al. Quantitative CT analysis of pulmonary nodules for lung adenocarcinoma risk classification based on an exponential weighted grey scale angular density distribution feature. *Comput Methods Programs Biomed*. 2018;2018:141–151.
4. Javaid M, Javid M, Rehman MZU, Shah SIA. A novel approach to CAD system for the detection of lung nodules in CT images. *Comput Methods Programs Biomed*. 2016;2016:125–139.
5. Tan Y, Schwartz LH, Zhao B. Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. *Med Phys*. 2013;40:043502.
6. Nian Y, LiM Cui H, et al. Graph-based unsupervised segmentation for lung tumor CT images. *Proceedings of the IEEE International Conference on Computer and Communications*; 2017:1884–1888.
7. Keshani M, Azimifar Z, Tajeripour F, Boostani R. Lung nodule segmentation and recognition using SVM classifier and active contour modeling: a complete intelligent system. *Comput Biol Med*. 2013;43:287–300.
8. Gu Y, Kumar V, Hall LO, et al. Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach. *Pattern Recogn*. 2013;46:692–702.
9. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI;2015, vol 9351: 234–241.
10. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J.UNet++: A Nested U-Net Architecturefor Medical Image Segmentation. 4th Deep Learning in Medical Image Analysis (DLMIA 2018/ML-CDS 2018) Workshop, LNCS 11045:3–11.
11. Mullally W, Betke M, Wang J, Ko JP. Segmentation of nodules on chest computed tomography for growth assessment. *Med Phys*. 2004;31:839–848.
12. Wang J, Li F, Li Q. Automated segmentation of lungs with severe interstitial lung disease in CT. *Med Phys*. 2009;36:4592–4599.
13. Korfiatis P, Kalogeropoulou C, Karahaliou A, Kazantzi A, Skiadopoulos S, Costaridou L. Texture classification-based segmentation of lung affected by interstitial pneumonia in high-resolution CT. *Med Phys*. 2008;35:5290–5302.
14. Nakagomi K, Shimizu A, Kobatake H, Yakami M, Fujimoto K, Togashi K. Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume. *Med Image Anal*. 2013;17:62–77.
15. Zhang W, Wang X, Zhang P, Chen J. Global optimal hybrid geometric active contour for automated lung segmentation on CT images. *Comput Biol Med*. 2017;91:168–180.
16. Zhou J, Yan Z, Lasio G, et al. Automated compromised right lung segmentation method using a robust atlas-based active volume model with sparse shape composition prior in CT. *Comput Med Imaging Graph*. 2015;46:47–55.
17. Soliman A, Khalifa F, Elnakib A, et al. Accurate lungs segmentation on CT chest images by adaptive appearance-guided shape modeling. *IEEE Trans Med Imaging*. 2017;36:263–276.
18. Setio AAA, Jacobs C, Gelderblom J, et al. Automatic detection of large pulmonary solid nodules in thoracic CT images. *Med Phys*. 2015;42:5642–5653.
19. Vivanti R, Joskowicz L, Karaaslan OA, Sosna J. Automatic lung tumor segmentation with leaks removal in follow-up CT studies. *Int J Comput Assist Radiol Surg*. 2015;10:1505–1514.
20. Yu Y, Zhang S, Li K, Metaxas D, Axel L. Deformable models with sparsity constraints for cardiac motion analysis. *Med Image Anal*. 2014;18:927–937.
21. Wang B, Tian X, Wang Q, et al. Pulmonary nodule detection in CT images based on shape constraint CV model. *Med Phys*. 2015;42:1241–1254.
22. Gill G, Toews M, Beichel RR. Robust Initialization of active shape models for lung segmentation in CT scans: a feature-based atlas approach. *Int J Biomed Imaging*. 2014;13:479154.
23. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active shape models-their training and application. *Comput Vis Image Underst*. 1995;61:38–59.
24. Nahed JA, Jolly MP, Yang GZ. Robust active shape models: a robust, generic and simple automatic segmentation tool. *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*; 2006:1-8.
25. Lekadir K, Merrifield R, Yang GZ. Outlier detection and handling for robust 3-D active shape models search. *IEEE Trans Med Imaging*. 2007;26:212–222.
26. Zhang S, Zhan Y, Dewan M, Huang J, Metaxas DN, Zhou XS. Sparse shape composition: a new framework for shape prior modeling. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2011:1025–1032.
27. Zhang S, Zhan Y, Dewan M, Huang J, Metaxas DN, Zhou XS. Towards robust and effective shape modeling: sparse shape composition. *Med Image Anal*. 2012;16:265–277.
28. Zhang S, Zhan Y, Zhou Y, Uzunbas M, Metaxas DN. Shape prior modeling using sparse representation and online dictionary learning. *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*; 2012:435–442.
29. Shi C, Cheng Y, Liu F, Wang Y, Bai J, Tamura S. A hierarchical local region-based sparse shape composition for liver segmentation in CT scans. *Pattern Recogn*. 2016;50:88–106.
30. Wang B, Gu X, Fan C, et al. Sparse group composition for robust left ventricular epicardium segmentation. *Comput Med Imaging Graph*. 2015;46:56–63.
31. Figueiredo MAT, Leitão JMN, Jain AK. Unsupervised contour representation and estimation using B-splines and a minimum description length criterion. *IEEE Trans Image Process*. 2002;9:1075–1087.
32. Unser M. Splines: a perfect fit for signal and image processing. *IEEE Signal Process Mag*. 1999;16:22–38.
33. Zhong B, Ma K, Liao W. Scale-space behavior of planar-curve corners. *IEEE Trans Pattern Anal Mach Intell*. 2009;31:1517–1524.
34. Grady L. Random walks for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2006;28:1768–1783.