Learning directional relative positions between mediastinal lymph node stations and organs

David Sarrut and Simon Rit
Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044, INSA-Lyon; Université Lyon 1, F-69621 Lyon, France and Léon Bérard Cancer Center, Université de Lyon, F-69373 Lyon, France

Line Claude
Léon Bérard Cancer Center, Université de Lyon, F-69373 Lyon, France

Romulo Pinho
Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044, INSA-Lyon; Université Lyon 1, F-69621 Lyon, France and Léon Bérard Cancer Center, Université de Lyon, F-69373 Lyon, France

Graham Pitson
Department of Radiation Oncology, Andrew Love Cancer Centre, Barwon Health, Geelong 3220, Australia

Gauthier Bouilhol
Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044, INSA-Lyon; Université Lyon 1, F-69621 Lyon, France and Léon Bérard Cancer Center, Université de Lyon, F-69373 Lyon, France

Rod Lynch
Department of Radiation Oncology, Andrew Love Cancer Centre, Barwon Health, Geelong 3220, Australia

Purpose: To automatically learn directional relative positions (DRP) between mediastinal lymph node stations and anatomical organs. Those spatial relationships are used to semiautomatically segment the stations in thoracic CT images.

Methods: Fuzzy maps of DRP were automatically extracted by a learning procedure from a database composed of images with stations and anatomical structures manually segmented by consensus between experts. Spatial relationships common to all patients were retained. The segmentation of a new image used an initial rough delineation of anatomical organs and applied the DRP operators. The algorithm was tested with a leave-one-out approach on a database of 5 patients with 10 lymph stations and 30 anatomical structures each. Results were compared to expert delineations with dice similarity coefficient (DSC) and bidirectional local distance (BLD).

Results: The overall mean DSC was 66% and the mean BLD was 1.7 mm. Best matches were obtained from stations S3P or S4R while lower matches were obtained for stations 1R and 1L. On average, more than 30 spatial relationships were automatically extracted for each station.

Conclusions: This feasibility study suggests that mediastinal lymph node stations could be satisfactory segmented from thoracic CT using automatically extracted positional relationships with anatomical organs. This approach requires the anatomical structures to be initially roughly delineated. A similar approach could be applied to other sites where spatial relationships exists between anatomical structures. The complete database of the five reference cases is made publicly available. © 2014 American Association of Physicists in Medicine. [http://dx.doi.org/10.1118/1.4873677]

Key words: mediastinal lymph stations segmentation, machine learning, directional relative position

1. INTRODUCTION

In the staging and treatment of lung cancer, it is important to assess lymph node involvement. An accurate labeling of the regions of mediastinal and hilar nodes is essential as it allows for the assessment of treatment outcomes, to select appropriate therapy for patient and to perform consistent clinical trials across institutions. Accurate and consistent definitions for the nodal stations is important. In radiation therapy, there is currently no consensus on whether to include lymph node stations in the target volume and to electively irradiate uninvolved mediastinal nodal regions.[1,2]

Lymph node station labeling is performed using anatomical reference definitions know as lymph node "maps." These maps have been historically used in describing the clinical and pathological extent of lymph node metastases in lung cancer. Naruke *et al.* developed the first nodal map in Japan in the 1960s.[3] The American Thoracic Society (ATS) subsequently developed its own map. In 1997, modifications were made to the ATS map by Mountain and Dresler.[4] Mountain and Dresler's map was adopted by the American Joint Committee on Cancer (AJCC) to describe nodal metastases for lung cancer in the 6th edition of the TNM classification. Chapet *et al.*[5] from the University of Michigan published a CT atlas with definitions of the mediastinal nodal stations in 2005. The International Association for Study of Lung Cancer (IASLC) created the lung cancer staging project and in turn developed an international lung cancer database. With analysis of
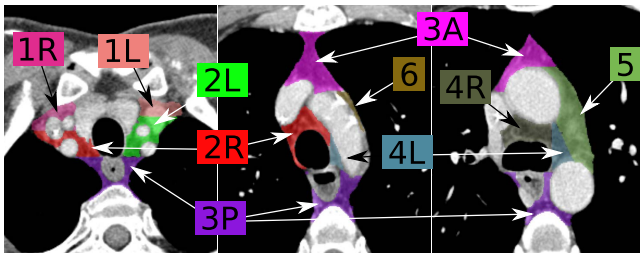
FIG. 1. Example of mediastinal lymph node stations delineated on three slices of a thoracic CT.

this database, the IASLC International Staging Committee proposed changes to the TNM staging system for lung cancer and were adopted in the new AJCC TNM 7th edition in 2010.[6]

The guidelines are composed of descriptions based on anatomical structures in the mediastinum. Each station is described with its position relative to surrounding anatomical organs such as aorta, trachea, vessels, and veins. The limits are mostly indicated via the principal patient axes LR (Left-Right), SI (Superior-Inferior), and AP (Anterior-Posterior). Nodal stations are labeled by number from their superior to inferior location, starting with the supraclavicular stations 1R and 1L (R for Right and L for Left), superior mediastinal: 2R, 2L, 3A, and 3P (A for Anterior, P for Posterior), 4R, 4L, aortic: 5 and 6, inferior mediastinal: 7-9, and hilar, lobar, and (sub)segmental: 10–14. The stations' contours are illustrated on the slices in Fig. 1. Some geometrical constructions are also defined in the guidelines to assist in delineating some nodal stations. For example, the boundary between 2R and 2L is defined by a vertical line passing tangentially along the left lateral tracheal border. Recently, Lynch *et al.*[7,8] published a CT atlas, based on the new IASLC map, analysing the stations delineated on a CT image.

Automated segmentation procedures are challenging to develop; however, they have the potential to save clinicians' time and improve segmentation consistency. The challenge is in the inter-patient anatomical and the inter-experts variability.[9] Several papers have encompassed the detection or segmentation of *nodes*;[10–13] however, none with the segmentation of mediastinal *stations*. A review of automated methods for lymph node segmentation was completed by Feuerstein *et al.*[10] Lu *et al.*[14] described a method to determine cuboid (parallelepiped) regions that encompass stations. Recently, Feuerstein *et al.*[10] proposed an interesting approach based on deformable image registration to create a probabilistic lymph node atlas from a database of segmented images. Such an atlas can be used to assign station labels to detected nodes by using Voronoi maps. Commowick *et al.*[15] also proposed an atlas-based method where reference lymph node station contours were deformed against the current patient image for head and neck lymph node station.

In this paper, we studied the manual delineation of mediastinal lymph stations on CT images by the clinicians to replicate this by an automatic process.[7,8] We investigated the relationships between stations and surrounding anatomical structures by means of their relative positioning.

2. METHOD

The method for segmenting mediastinal nodal stations consisted of two steps. In the first one, directional relative position (DRP) operators were extracted from a training dataset, which was composed of mediastinal nodal stations and anatomical structures manually delineated by clinical experts on thoracic CT images. All manual segmentations were approved by consensus. In the second step, the DRP operators were used to segment an image that was not present in the learning set. Initially, starting regions were automatically delimited for all stations by the parallelepipedic areas according to LR, SI, and AP limits. Then, this initial segmentation was progressively refined by applying the directional constraints of the DRP operators with respect to anatomical structures.

2.A. Building the fuzzy maps

Directional fuzzy maps were the basis for the construction of the DRP operators proposed in this work. They were built from a set of N_p manually segmented thoracic CT images (see Sec. 3.A). Each image contained a set $\mathcal{A}_p$ of N_A anatomical structures (AS), and a set $\mathcal{S}_p$ of N_S lymph stations (LS), with p the patient image index. Segmented AS and LS were binary images, $A(x) : \Omega_A \in \mathbb{R}^3 \to \{0, 1\}$, $S(x) : \Omega_S \in \mathbb{R}^3 \to \{0, 1\}$, with Ω_A, Ω_S the image domains, $x \in \mathbb{R}^3$ a point coordinate, 0 the background, and 1 the foreground.

In computer vision, relationships between spatial entities can be classified into topological, distance and orientation relations.[16] The inherent imprecision of those relations makes fuzzy set theory a successful mathematical model to study them. In this work, we considered fuzzy directional relations such as "left to," "right to," "anterior to," etc., or more generally, directions given by orientation angles. Several approaches have been proposed to build and analyze these relationships, e.g., the centroid method and the histogram of angles, see Ref. 16 for a review. Here, we relied on the morphological approach proposed by Ref. 17, which consists of two steps: first a fuzzy landscape is defined around a given reference object A, such that each point in this landscape corresponds to the degree of at which the spatial relation of interest (for example, "left to") is satisfied; second, an object B is placed on the fuzzy map in order to evaluate how well this object matches the referred spatial relation.

More formally, let us consider a direction α defined in 2D by an angle (two angles in polar coordinates in 3D). Given an anatomical structure A, the fuzzy map (FM) $\mathcal{F}_{\alpha,A}(x) : \Omega \in \mathbb{R}^3 \to \mathbb{R}$ expresses, for each pixel x, the degree at which x is along the direction α relative to A. According to Ref. 17, FM is defined as $\mathcal{F}_{\alpha,A}(x) = \max\left(0, 1 - \frac{2}{\pi}\beta_{\min}(x)\right)$, with Ω the domain on which it is computed. The fuzzy value at a given point x is thus a linear function of the minimal value $\beta_{\min}(x) = \min_{y \in \Omega_A}\beta(x, y)$, $\beta(x, y) = \arccos\frac{\vec{yx}.\vec{d_\alpha}}{||\vec{yx}||}$, and $\beta(x, x) = 0$, which considers angles between the unit direction
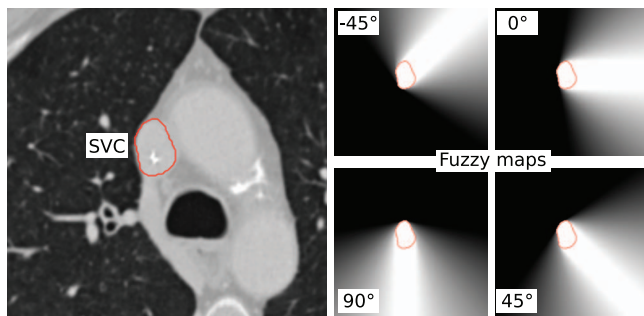
FIG. 2. Left: CT slice with the contour of the Superior Vena Cava (SVC). Right: Examples of four fuzzy maps computed from the SVC binary image for four different α values ($-45°$, $0°$, $45°$, $90°$).

$\vec{d_\alpha}$ and every point y in A. It has been shown that the presented FM can be computed by the fuzzy dilation of object A with the structuring element $\nu(x) = \max\left(0, 1 - \frac{2}{\pi}\arccos\frac{\vec{yx}.\vec{d_\alpha}}{||\vec{yx}||}\right)$. We thus used the propagation algorithm proposed in Refs. 18 and 19 with a neighborhood radius equal to three pixels.

We considered a discrete set $\mathcal{D}$ of N_α relative directions, and computed the $N_P \times N_A \times N_\alpha$ fuzzy maps for all AS in the database. Figure 2 illustrates some fuzzy maps.

2.B. Learning phase

The goal of the learning phase was to extract from the training database a set of DRP between anatomical structures and lymph stations that were common among all patients. Given a set of directions $\mathcal{D}$, only those spatial relations that were satisfied across the entire patient database were kept. For example, if we observed that the station S4R was posterior to the superior vena cava (SVC) for all patients in the database, this "posterior to" relation (represented by a corresponding DRP) was kept. Otherwise, it was not considered.

To find if a certain spatial relationship α between an anatomical structure A and a station S was common to all patients, we first needed to devise an algorithm to evaluate

the relationship of interest for a single patient. Having the fuzzy map $\mathcal{F}_{\alpha,A}$ at hand, we proposed to employ a cumulative histogram H_S to indicate whether S respects the DRP defined by $\mathcal{F}_{\alpha,A}$. The histogram $H_S(t)$ determined the number of pixels belonging to station S having a fuzzy value in $\mathcal{F}_{\alpha,A}$ lower than a threshold t. Namely, $H_S(t) = |\{x \mid S(x) = 1$ and $\mathcal{F}_{\alpha,A}(x) < t\}|$.

With this cumulative histogram, we could determine how well S satisfied the DRP defined by A and α. To achieve this, we calculated the optimal threshold $t_S = \arg\max_t H_S(t) < \varepsilon$, with ε a given tolerance expressed in number of pixels. The optimal threshold determined, in the end, the proportion of S contained in areas of high values of $\mathcal{F}_{\alpha,A}$. If the proportion was high, the relation "S is α-oriented with respect to A" was satisfied. Figure 3 illustrates this notion. When considering all patients in the database, we kept the minimal value $t_{\min}$ among the N_P optimal thresholds. If, for one patient, t_S did not exist, no $t_{\min}$ was considered. We denoted the triplet $\{A, \alpha, t_{\min}\}$, the DRP operator for a given relation between A and S. The sets $\mathcal{R}_S$ of all DRP associated with S constitute the spatial relationships between anatomical structures and stations that are common to all patients.

Note that a global rigid registration between patient images was not strictly required. However, we considered, without loss of generality, that the patients were in the same orientation according to the principal image axes. The underlying assumption of the proposed method was that the positions of stations relatively to the anatomical structures in the patient coordinate system were similar. This is in fact the assumption made in all published guidelines.

2.C. Segmentation

Once the sets $\mathcal{R}_S$ were build, they could be used to segment a new image. The algorithm (see Fig. 4), started from an initial S_0 segmentation and successively considered all spatial relationships DRP_i in $\mathcal{R}_S$. For each DRP_i operators $\{A, \alpha, t_{\min}\}$ in $\mathcal{R}_S$, a binary image F_i was build such that $F_i(x) : \Omega_S$
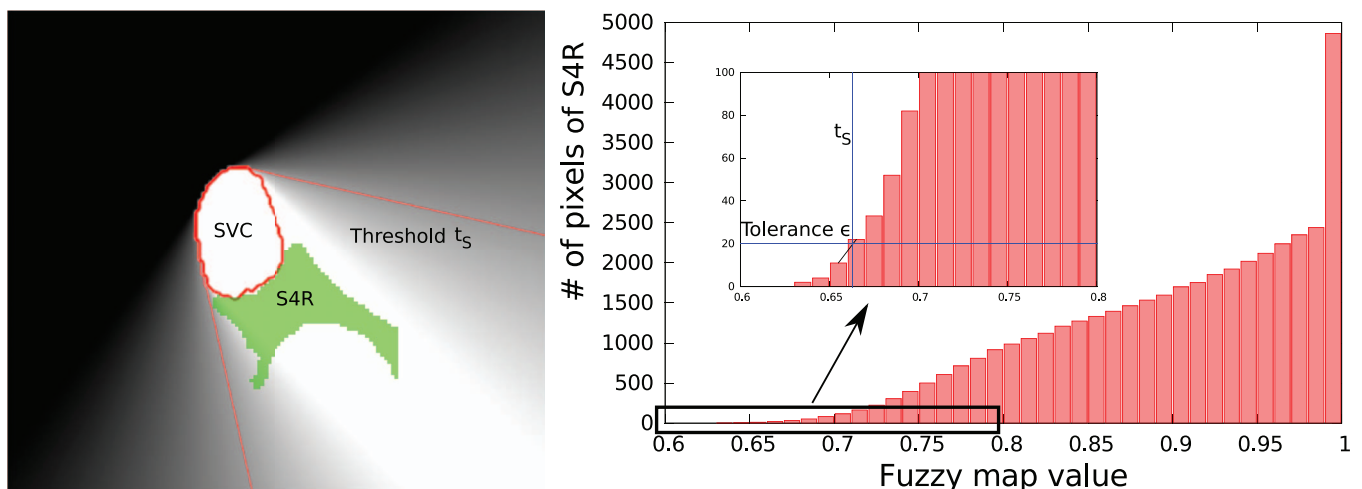

FIG. 3. Left: Example of one slice of the fuzzy map computed for the Superior Vena Cava (SVC). The station S4R is shown below. The line depicts the threshold t_S. Right: Cumulative histogram H_S computed from $\mathcal{F}_{S,\alpha=45°}$ and S4R. The optimal threshold t_S is computed according to the tolerance ε (set to 20 pixels in this example).

```
1:  procedure SEGMENT(Station S)
2:      S₀ = Compute initial station from mediastinum and organs limits
3:      for all DRPᵢ ∈ ℛ_S do
4:          Compute the binary image Fᵢ = {1 if ℱ_{α,A}(x) > t_{min}; 0 otherwise}
5:      end for
6:      The final segmentation is S = S₀ × ∏ᵢ Fᵢ
7:  end procedure
```

FIG. 4. Algorithm to segment a station according to a set of spatial relationships.

$\rightarrow [0:1] = \{1 \text{ if } \mathcal{F}_{\alpha,A}(x) > t_{\min}; 0 \text{ otherwise}\}$. The pixels in this image having value 1 were the ones considered to be "at a degree $t_{\min}$" in the direction α relative to the structure A. So every candidate pixels in S_0 that did not belong to F_i could be discarded. The final segmentation S was obtained by the product of all binary images F_i, i.e., $S = S_0 \times \prod_i F_i$. This process is illustrated Fig. 5.

The initial S_0 was built by considering pixels belonging to the mediastinum. The mediastinum was automatically extracted by first removing the lungs[20] and then applying DRP operators to areas between them. For each station, the superior and inferior limits were automatically determined as indicated in the segmentation guidelines, similarly to Ref. 14. For example, the superior limit of station S1R is the *Cricoid Cartilage*, and the inferior limit is the *Sternum*. We thus removed from the mediastinum all slices above and below both anatomical structures.

The whole procedure could be performed with 3D operators. However, in this work, we opted to use 2D operators on a slice-by-slice basis, which were then aggregated to form the 3D fuzzy maps. The cumulative histogram H_S and optimal thresholds t_S were all computed in 3D. We chose this approach because by using 3D operators, we would assume that every anatomical structures was defined in a consistent manner between patients, with similar limits, which is not always
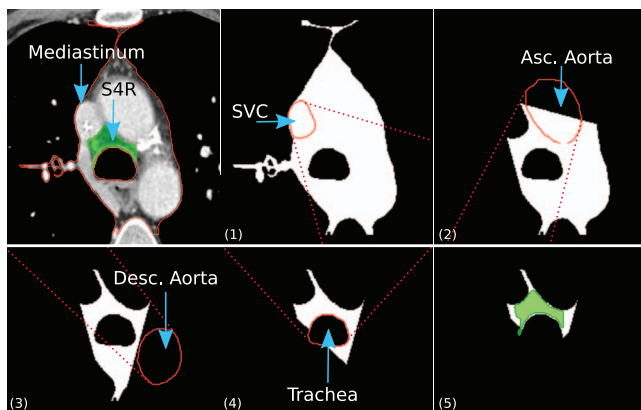

FIG. 5. Illustration of the segmentation process for the station S4R. Top left image depicts a slice of the initial CT image with the contour of the mediastinum. We also depict the targeted station S4R (considered as unknown). The other images illustrate five steps, starting from the initial S_0 mediastinum contour. A first DRP operator (1) is applied with the structure SVC and remove part of the mediastinum (2). The three other DRP operators used *Ascending Aorta*, *Descending Aorta*, and *Trachea* (2,3,4). All thresholds are shown in dotted line and correspond to the t_{min} values extracted from the learning step. The reference S4R area is displayed on the last image (5) for comparison.

possible. Moreover, the guidelines[6] were also determined on a slice-by-slice basis.

3. EXPERIMENTS AND RESULTS

3.A. Database

CT images, acquired in breath-hold, of five patients have been manually segmented by three radiation oncologists (RL, LC, and GP). The patients were being treated for lung cancer at the Centre Léon Bérard, in Lyon, France. They received an intravenous (IV) contrast agent, the inplane image resolution was in the range [0.63–0.83] mm, and the slice thickness was of [0.8–2] mm. The radiation oncologists delineated the lymph node stations from 1 to 11 following the new IASLC station definitions.[6] Segmentation was performed on a slice by slice basis, following these guidelines. Experts provided a single consensual segmentation for each station. They also provided the delineation of several mediastinal anatomical structures that were used during the contouring process of the stations. This set of structures was composed of vessels, arteries, etc. mentioned in the guidelines.[5–7] The instructions were to delineate anatomical structures only on the slices that are needed to guide the segmentation of stations, so most of the anatomical structures were not completely delineated. It is worth noting that this manual delineation of about 250 structures (16 stations and more than 30 structures per patient) was a very time consuming process,[21] estimated to 8 to 10 h per patient. The segmentation of the stations being consensus-based meant that only one delineation was available for each station preventing us from assessing inter-observer variability and to use conventional performance evaluation approaches such as STAPLE-related methods.[22, 23]

3.B. Criteria

In order to quantify the overlap between two 3D structures A and B, the Dice similarity coefficient $\text{DSC}(A, B) = \frac{2|A \cap B|}{|A|+|B|}$ was used. As proposed by Kim *et al.*,[24] we also used a robust point-to-surface distance measure named *bidirectional local distance* $\text{BLD}(A, B)$, based on the Hausdorff metric, expressed in millimeter.

3.C. First experiment

The goal of this first experiment was to evaluate the maximum performance that could be reached using the DRP concept explained above. For a given patient p, we extracted the DRP operators and used them to segment the same patient. The tolerance ε was fixed at 0.5% of the station volume. The binning of the H_S histogram was set to 100 (threshold intervals of 0.01). Three sets of orientations with increasing sizes: $N_\alpha = 4$ (0°, 90°, 180°, and 270° corresponding to the main AP and LR directions), $N_\alpha = 8$ (every 45°), and $N_\alpha = 16$ (every 22.5°) were compared. This experiment was performed for five patient images and ten stations: supraclavicular, superior mediastinal, and aortic stations (1 to 6), corresponding to stations above the Carina. Table I shows the DSC and BLD values by station averaged over the five patients.

TABLE I. Dice similarity coefficients and BLD distance averaged for the five patients for each ten stations according to three numbers of orientations. Dice criteria is expressed in percentage and BLD in mm. The mean station sizes are also indicated in cm^3.

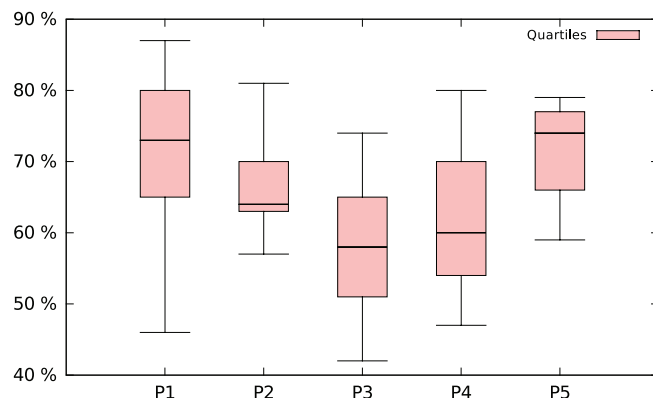|	DICE			BLD in mm			
Angles	4	8	16	4	8	16	Size in cc
Stations							
S1R	74%	77%	77%	2.0	1.7	1.7	15
S1L	71%	73%	74%	2.0	1.9	1.8	15
S2R	74%	74%	74%	1.1	1.1	1.0	11
S2L	68%	68%	69%	1.3	1.3	1.2	7
S3A	71%	73%	73%	1.4	1.3	1.3	28
S3P	83%	84%	84%	1.2	1.2	1.1	51
S4R	82%	82%	83%	0.9	0.8	0.9	7
S4L	73%	74%	74%	1.8	1.3	1.3	4
S5	87%	87%	86%	0.8	0.8	0.8	9
S6	70%	70%	70%	0.8	0.8	0.8	2
Mean	75%	76%	76%	1.3	1.2	1.2	

TABLE II. Dice similarity coefficients and BLD distances for the five experiments. Dice criteria are expressed in percentage and BLD in mm. The means by stations are also indicated.

Stations	S1R	S1L	S2R	S2L	S3A	S3P	S4R	S4L	S5	S6	Mean
DICE (%)	62%	62%	69%	57%	68%	80%	78%	61%	70%	55%	66%
BLD (in mm)	2.9	2.3	1.4	1.6	2.1	1.7	1.1	1.7	1.1	1.2	1.7

$\Delta_i = \mathrm{DSC}(S_i) - \mathrm{DSC}(S_{i-1})$, and counted the number of times Δ_i is greater (Δ^+), lower (Δ^-), and equal (Δ^0) to zero. Δ^+ (resp. Δ^-) corresponds to an improvement (resp. degradation) of the segmentation. We computed the average DSC improvement $\bar{\Delta}^+$ and degradation $\bar{\Delta}^-$.

4. DISCUSSION

In Table I, DSC ranged from about 60% to more than 80%, meaning that DRPs can retrieve a large part of the segmentation. As expected, using $N_\alpha = 8$ directions, rather than $N_\alpha = 4$ improves results (mean 76% vs 75%), but larger values ($N_\alpha = 16$) only led to marginal improvements. The smallest stations in size (S2L, S6) led to lower DSC values (around 70%) but depicted good BLD (around 1 mm). Stations S1R and S1L, with some ambiguities in the guidelines, as shown in Ref. 7, were the ones with the lowest BLD values (1.7–1.9 mm).

In Table II and Fig. 6, the mean DSC ranged from 55% (S6) to 80% (S3P), the BLD from 2.9 mm (S1R) to 1.1 mm (S4R and S5), illustrating that the learning stage managed to extract useful DRP operators. Among the stations, S3A, S4R, and S5 were the ones that yielded the best results among patients. Station S6 was the one with the lowest DSC (55%), but this was the smallest station in size and the BLD was good, around 1.2 mm. A similar trend was obtained for station 4L. Among all patients, BLD was lower than 3 mm. Stations S1R and S1L depicted the least good results, 2.3 and 2.9 mm (and 62 % DSC).

The worst DSC/BLD result was obtained with patient 3 and station S3A. The images were reviewed and it was

3.D. Second experiment

The second experiment was aimed at evaluating the possibility of detecting common DRPs across several patients and extracting stable spatial dependencies between stations and anatomical structures. A leave-one-out procedure was applied. Successively, $N_p - 1$ images were considered and the learning process described above was performed. The last image was then automatically segmented from the set of DRP operators learned from the other images. DSC and BLD were computed between reference and segmented stations. This procedure was repeated N_p times with all the images and averaged. The learning stage was performed with $N_\alpha = 8$ angles values (based on the results of the first experiment).

Table II depicts the mean DSC and BLD values for the leave-one-out procedure, for each station. Figure 6 depicts box plots for DSC and BLD values for the five experiments. Table III depicts the number of DRP operators extracted by the learning process (size of $\mathcal{R}_S$) for the five experiments. As the ground truth is known, we computed the DSC values obtained by applying each operator DRP$_i$ individually,
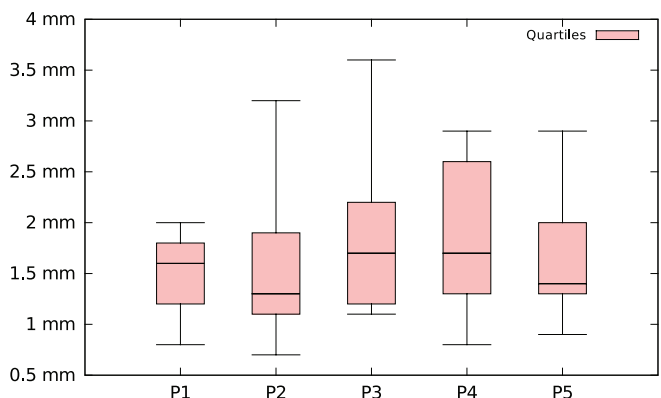

FIG. 6. Minimum, 1st quartile, median, 3rd quartile, and maximum values, per patient and among stations, for DSC (left) and BLD (right) measures, for the five experiments.

TABLE III. Analysis of the RP operators by patient. First column depicts the number of used RPs ($|\mathcal{R}|$). The column Δ^+ (resp. Δ^-) indicates the number of operators leading to positive (resp. negative) gain in DSC. The $\bar{\Delta}^+$ and $\bar{\Delta}^-$ columns, the mean DSC gain/loss. Last column depicts the number of operators that did not modify the DSC.

	$|\mathcal{R}|$	Δ^+	$\bar{\Delta}^+$	Δ^-	$\bar{\Delta}^-$	Δ^0
p1	332	132	2.2%	15	−0.3%	185
p2	335	131	2.1%	19	−1.0%	185
p3	328	131	2.4%	24	−0.8%	173
p4	331	131	2.0%	18	−1.6%	182
p5	333	132	2.3%	27	−1.0%	174
Mean	332	131	2.2%	21	−0.9%	180

found that, during the delineation by the experts, the Brachio Cephalic Vein (BCV) was not removed from station S3A while it is the case for the four others patients, as illustrated in Fig. 7. According to the guidelines,[6] only the most anterior part of the BCV should be included in station S3A, but only from the slice where the BCV divides into two parts. This situation was illustrated here to exemplify the interest of the proposed approach, not only as an automated segmentation method, but also to highlight situations where discrepancy occurs with guidelines.

In Table III, on average over the five leave-one-out procedures, 332 DRP operators were extracted (about 33 per stations). As the DRP were built according to the spatial relationships observed in the four other patients, there was no guarantee that these relationships were going to be of value for the fifth patient. Interestingly, few DRPs led to a DSC degradation (Δ^-), and in those cases, the mean DSC loss ($\bar{\Delta}^-$) was less than 1%. There were thus few situations where spatial relationships can be observed on four patients and are different for the fifth patient. More than 50% (180 over 332) of the DRP operators did not modify the DSC value. It means that in those cases, the pixels removed by the operators had already been removed by previous operators. In the other positive cases, the mean DSC gain $\bar{\Delta}^+$ by RP was of 2.2%. It is not possible to determine in advance if a DRP operator will lead to a DSC gain or loss.

One limitation of this method is that it requires the delineation of several anatomical structures prior to undertaking segmentation of the mediastinal nodal stations. The performance of the proposed method is thus linked to the accuracy of delineation of those anatomical structures. However, these
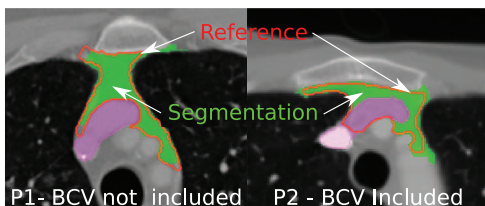

FIG. 7. The reference station S3A is depicted by its contour and the result of the automated segmentation is shown as the filled area. Part of the brachio cephalic vein (BCV) is included in the contour for patient 3 (right), while it is excluded in patient 1 (left).

anatomical structures needed to be only roughly segmented. We did not investigate precisely the impact but it is expected to be limited because only part of the structures are used. For example, consistence between patients for starting and ending slices of an organs delineation (which could be sometimes difficult for organs such as arteries or veins) is not required. What is needed is a correct delineation for the part of the structure that is adjacent to a station. Segmentation of the anatomical structures could be obtained automatically using conventional methods.

Another limitation is that the learning phase was based on five patients and this small number of patients is not sufficient to cover all anatomical variations. Increasing the number of cases in the database could reduce the number of DRP operators because of the conservative method used to set the thresholds (minimum). Among the different relationships between spatial entities, we decided here to only study directional relative positions. Other relationships, such as relationships to multiple structures ("between," "along"), adjacency notions, fuzzy distances, symmetry measures, etc., see Ref. 25, could also be integrated into the generic learning process.

Regarding the computation time, the application of each DRP was fast, a few seconds on a single core Intel Xeon CPU E5-1660 3.3GHz. The computation time is linear in the number of voxels of the stations and anatomical structures considered. The complete segmentation of a patient ranged from 2 to 7 min on a single core. No particular optimization has been made and there is no doubt that faster parallel implementations could be performed. The learning dataset is composed of $N_P \times N_S \times N_A \times N_\alpha$ fuzzy maps. For example, with $N_P = 5$ patients, $N_s = 10$ stations, $N_A = 10$ anatomical structures (in mean by stations), and $N_\alpha = 16$ orientations, it leads to a size of 8 GB (highly compressible).

5. CONCLUSION

The positional relationships between stations and anatomical structures by means of DRP operators was investigated. We showed feasibility of automatically learning those operators from a set of reference segmented images in order to provide automatic segmentation of a new image. This process intends to mimic the manual delineation based on the guidelines. It is a purely geometrical method, which provides insights into the segmentation process. To our knowledge, no other method has been published to segment mediastinal lymph stations.

Access for the research community to the database of cases used in this study is provided, in an open-data approach.[26,27] The CT images were anonymized and made available. The segmented stations and anatomical structures are also available as binary images or Dicom-RT-Struct file format. We hope it will help further research in this field: http://www.creatis.insa-lyon.fr/lymph-stations-atlas.

The proposed approach was designed specifically for autosegmenting the mediastinal lymph nodes stations. It could also be useful to help the definition of atlas. This approach is complementary to other segmentation methods based on pixel intensities. A similar approach could be applied to other

sites where spatial relationships exist between anatomical structures.

[1]A. T. Fernandes, J. Shen, J. Finlay, N. Mitra, T. Evans, J. Stevenson, C. Langer, L. Lin, S. Hahn, E. Glatstein, and R. Rengan, "Elective nodal irradiation (ENI) vs. involved field radiotherapy (IRFT) for locally advanced non-small cell lung cancer (NSCLC): A comparative analysis of toxicities and clinical outcomes," Radiother. Oncol. **95**, 178–184 (2010).

[2]C. R. Kelsey, L. B. Marks, and E. Glatstein, "Elective nodal irradiation for locally advanced non-small-cell lung cancer: Its called cancer for a reason," Int. J. Radiat. Oncol., Biol., Phys. **73**, 1291–1292 (2009).

[3]T. Naruke, K. Suemasu, and S. Ishikawa, "Lymph node mapping and curability at various levels of metastasis in resected lung cancer," J. Thorac. Cardiovasc. Surg. **76**, 832–839 (1978).

[4]C. F. Mountain and C. M. Dresler, "Regional lymph node classification for lung cancer staging," Chest **111**, 1718–1723 (1997).

[5]O. Chapet, F.-M. Kong, L. E. Quint, A. C. Chang, R. K. Ten Haken, A. Eisbruch, and J. A. Hayman, "Ct-based definition of thoracic lymph node stations: An atlas from the University of Michigan," Int. J. Radiat. Oncol., Biol., Phys. **63**, 170–178 (2005).

[6]V. W. V. V. Rusch *et al.*, "The iaslc lung cancer staging project: A proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer," J. Thorac. Oncol. **4**, 568 (2009).

[7]R. Lynch, G. Pitson, D. Ball, L. Claude, and D. Sarrut, "Computed tomographic atlas for the new international lymph node map for lung cancer: A radiation oncologist perspective," Pract. Radiat. Oncol. **3**, 54–66 (2013).

[8]G. Pitson, R. Lynch, L. Claude, and D. Sarrut, "A critique of the iaslc lymph node map: A radiation oncology perspective," J. Thorac. Oncol. **7**, 478–480 (2012).

[9]L. Kepka, K. Bujko, D. Garmol, J. Palucki, A. Zolciak-Siwinska, Z. Guzel-Szczepiorkowska, L. Pietrzak, K. Komosinska, A. Sprawka, and A. Garbaczewska, "Delineation variation of lymph node stations for treatment planning in lung cancer radiotherapy," Radiother. Oncol. **85**, 450–455 (2007).

[10]M. Feuerstein, B. Glocker, T. Kitasaka, Y. Nakamura, S. Iwano, and K. Mori, "Mediastinal atlas creation from 3-d chest computed tomography images: Application to automated detection and station mapping of lymph nodes," Med. Image Anal. **16**, 63–74 (2012).

[11]M. Feuerstein, T. Kitasaka, and K. Mori, "Adaptive branch tracing and image sharpening for airway tree extraction in 3-d chest CT," in *International Workshop on Pulmonary Image Analysis* (London, 2009).

[12]D. Maleike, M. Fabel, R. Tetzlaff, and H. V. Tengg-kobligk, "Lymph node segmentation on ct images by a shape model guided deformable surface method," in *SPIE Medical Imaging* (2008).

[13]I. Sluimer, A. Schilham, M. Prokop, B. V. Ginneken, and A. Current, "Computer analysis of computed tomography scans of the lung: A survey," IEEE Trans. Med. Imaging **25**, 385–405 (2006).

[14]K. Lu, and W. E. Higgins, "Semi-automatic central-chest lymph-node definition from 3d mdct images," in *SPIE Medical Imaging 2010: Computer-Aided Diagnosis*, edited by N. Karssemeijer and R. M. Summers (SPIE, San Diego, CA, 2010), Vol. 7624.

[15]O. Commowick, V. Gregoire, G. Malandain, and V. Grégoire, "Atlas-based delineation of lymph node levels in head and neck computed tomography images," Radiother. Oncol. **87**, 281–289 (2008).

[16]I. Bloch, "Directional relative position between objects in image processing: A comparison between fuzzy approaches," Pattern Recognit. **36**, 1563–1582 (2003).

[17]I. Bloch, "Fuzzy relative position between objects in image processing: A morphological approach," IEEE Trans. Pattern Anal. Mach. Intell. **21**, 657–664 (1999).

[18]I. Bloch, "Fuzzy relative position between objects in image processing: New definition and properties based on a morphological approach," Int. J. Uncertainty, Fuzziness Knowledge-Based Syst. **07**, 99–133 (1999), see http://www.worldscientific.com/doi/pdf/10.1142/S0218488599000088.

[19]J. Atif, O. Nempont, E. Angelini, and I. Bloch, "Integration of fuzzy mathematical morphology and fuzzy spatial relationships into ITK," in *The Insight Journal - 2005 MICCAI Open-Source Workshop* (MICCAI 2005 Palm Spring, 2005), pp. 1–5.

[20]S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images," IEEE Trans. Med. Imaging **20**, 490–498 (2001).

[21]D. Sarrut, L. Claude, S. Rit, R. Pinho, G. Pitson, and R. Lynch, "Investigating mediastinal lymph node stations segmentation on thoracic ct following experts guidelines," in *Proceedings of the First International Workshop on Image-Guidance and Multimodal Dose Planning in Radiation Therapy (MICCAI)* (miccai2012, Nice, France, 2012), pp. 1–8.

[22]S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation," IEEE Trans. Med. Imaging **23**, 903–921 (2004).

[23]O. Commowick, A. Akhondi-Asl, and S. K. Warfield, "Estimating a reference standard segmentation with spatially varying performance parameters: Local map staple," IEEE Trans. Med. Imaging **31**, 1593–1606 (2012).

[24]H. S. Kim, S. B. Park, S. S. Lo, J. I. Monroe, and J. W. Sohn, "Bidirectional local distance measure for comparing segmentations," Med. Phys. **39**, 6779–6790 (2012).

[25]I. Bloch, "Fuzzy spatial relationships for image processing and interpretation: A review," Image Vision Comput. **23**, 89–110 (2005).

[26]L. Ibanez and W. Schroeder, "The case for open science," White Paper No. 4, 2010, see http://www.kitware.com/publications/item/view/1154.

[27]P. Murray-Rust, "Open data in science," Ser. Rev. **34**, 52–64 (2008).