Automatic Segmentation of Spinal Canals in CT Images via Iterative Topology Refinement

Qian Wang, Le Lu, Dijia Wu, Noha El-Zehiry, Yefeng Zheng, Dinggang Shen*, and Kevin S. Zhou

Abstract—**Accurate segmentation of the spinal canals in computed tomography (CT) images is an important task in many related studies. In this paper, we propose an automatic segmentation method and apply it to our highly challenging image cohort that is acquired from multiple clinical sites and from the CT channel of the PET-CT scans. To this end, we adapt the interactive random-walk solvers to be a fully automatic cascaded pipeline. The automatic segmentation pipeline is initialized with robust voxelwise classification using Haar-like features and probabilistic boosting tree. Then, the topology of the spinal canal is extracted from the tentative segmentation and further refined for the subsequent random-walk solver. In particular, the refined topology leads to improved seeding voxels or boundary conditions, which allow the subsequent random-walk solver to improve the segmentation result. Therefore, by iteratively refining the spinal canal topology and cascading the random-walk solvers, satisfactory segmentation results can be acquired within only a few iterations, even for cases with scoliosis, bone fractures and lesions. Our experiments validate the capability of the proposed method with promising segmentation performance, even though the resolution and the contrast of our dataset with 110 patient cases (90 for testing and 20 for training) are low and various bone pathologies occur frequently.**

Index Terms—**Image landmark detection, image segmentation, random walk, spinal canal, topology.**

I. INTRODUCTION

MANY research and clinical studies require the automatic segmentation of the spines, which is capable of facilitating disease diagnosis, treatment, and statistical analysis/evaluation. For example, the segmentation of the spine provides spatial reference to locate and identify other neighbouring anatomical structures in abdomens and chests, thus contributing to the understanding of the full-body scan essentially [1]–[3]. In terms of image registration, the segmented spines provide important features that are helpful to the correct alignment of corresponding anatomical structures across individual subjects [4], [5]. Furthermore, it becomes easier to conduct disease-oriented analysis given the segmented topologies/shapes of the spines [6]. In the meantime, based on the segmentation of the spinal canal, the entire spinal cord can easily be delineated, making it possible to count radiotherapy dosages which are crucial to the normal functions of the nerve tracts [7]. To this end, a lot of efforts have been devoted to the segmentation of the spine and its related structures from multi-modal imaging acquisitions.

Although relatively more work in the literature focus on the segmentation of the spines in magnetic resonance (MR) images, we will investigate the automatic segmentation of the spinal canals from highly varying computed tomography (CT) images in this paper. Accurately segmenting the spinal canal facilitates the computer-aided detection process of anomalies, such as epidural masses on CT scans [8]. Another application is to enhance the bone lesion or tumor visualization by anatomy-aware attenuation correction in PET-CT imaging [9], [10]. A literature review of segmenting the spinal canals in CT images will be provided in Section II. In general, most conventional methods require user inputs to certain extent for the segmentation of the spinal canal in CT data. It is not easy to apply these semi-automatic methods to large-scale image cohorts, as human interaction is often tedious and costs high. The inconsistency among human experts also challenges the quality of the segmentation results.

Alternative automatic solutions mostly follow the *top-down* design, by recognizing related anatomical structures from the large scales to small. In this way, the spine can finally be localized with the helps from other neighboring anatomies. However, the robustness of these top-down methods are challenged if certain anatomical structure is missing in a specific image due to limited superior-inferior coverage. In particular, our patient population has a large portion of unhealthy subjects where various bone/spine pathologies can often be observed. The *top-down* strategy may not generalize well to these diseased/outlier cases. This motivates us to exploit more flexible and robust data-driven *bottom-up* method in this paper.

Meanwhile, it is worth noting that our task to segment the spinal canal is extremely difficult concerning the data, which will be handled in the method and experimental sections.

1) We acquire a total number of 110 CT scans of individual patients from eight different clinical sites for this study.

Q. Wang is with Med-X Research Institute, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: wang.qian@sjtu.edu.cn).

L. Lu is with Radiology and Imaging Science, National Institutes of Health (NIH) Clinical Center, Bethesda, MD 20892 USA (e-mail: le.lu@nih.gov).

D. Wu is with Google, Seattle, WA 98103 USA (e-mail: dijia@google.com).

N. El-Zehiry, Y. Zheng, and K. S. Zhou are with Siemens Corporate Research, Princeton, NJ 08540 USA (e-mail: noha.el-zehiry@siemens.com; yefeng.zheng@siemens.com; shaohua.zhou@siemens.com).

*D. Shen is with the Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27510 USA (e-mail: dgshen@med.unc.edu).
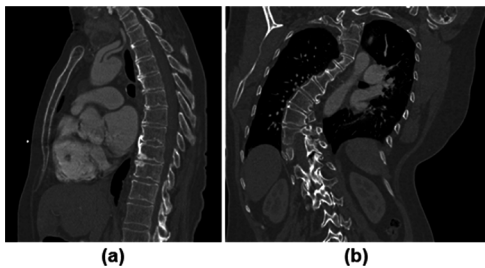
Fig. 1. The views of the spines of two individual patients. In (a) the shape of the spine is relatively normal. However, in (b), the disease (i.e., scoliosis) affects the appearance of the spine significantly by introducing a twist alongside the lateral axis. The views are sagittal and coronal in Fig. 1(a) and (b), respectively. Both exemplar images are isotropically resampled (to the voxel resolution of $2 \times 2 \times 2$ mm^3) and zoomed to the region of interest.

The cohort of the CT images is among the largest in the literature.

2) The images in our collection vary significantly in terms of the acquisition configurations, including superior-inferior coverage, sizes, spatial resolutions, etc. This variation strongly challenges the robustness of the automatic segmentation methods.

3) All images are acquired from the CT channels of the PET-CT scanners. Therefore, the quality (e.g., signal-noise-ratio, tissue contrast) of the images is relatively lower compared to regular CT acquisitions. Also, regular CT scans of the spines are often acquired with intravenous (IV) contrast. However, for our data, no contrast is involved.

4) We recruit patients and aim to verify the proposed method upon real clinical data setup. All patients are affected by diseases (common to PET-CT imaging patient population), which introduce uncontrollable impacts to the shapes/appearances of the spines. For example, typical slices of two patients are provided in Fig. 1(a) and (b), respectively. It is observable that, in Fig. 1(b), a disease (i.e., scoliosis) severely affects and then twists the spine in the lateral (left-right) axis direction, resulting in a significant abnormality.

Though facing the aforementioned technical challenges, we have successfully developed a novel method to segment the spine canals from the large-scale cohort of diverse CT images. Our automatic segmentation method relies on the iterative refinement of the topology of the spine, which provides contextual guidance to improve the segmentation of the spinal canal gradually. In particular, we describe the topology of the spine by its medial line, which can be extracted and refined from the tentative segmentation of the spinal canal. Seeding voxels are sampled according to the refined spine topology, and further fed into the random-walk solver[11] to update the segmentation of the spinal canal. By iteratively applying the above scheme, we are able to cascade several random-walk solvers, and attain highly robust and accurate segmentation of the spinal canal in CT images.

In general, the major contributions of our work consist of the following three aspects.

1) A fully automatic method is proposed for the segmentation of the spinal canal in CT image. The method is successfully validated upon one of the largest and most challenging datasets ever reported in the literature.

2) We utilize the iterative random-walk solvers to flexibly fill in the spatial occupancy of the spinal canal. The topology of the spinal canal, which is extracted and refined from the tentative segmentation, behaves as the regularized context information for the refinement of the segmentation.

3) The topology of the spinal canal is initially determined from the discriminative learning of the training images. Then, subject-specific geometric/appearance constraints are applied for the refinement of the topology of each testing subject under consideration.

The rest of this paper is organized as follows. In Section II, a literature review related to this work will be provided. Then, we will detail the proposed method in Section III and apply it to the large-scale cohort of diverse CT images in Section IV. This paper will be concluded with extended discussions in Section V.

II. RELATED WORK

A. Segmentation of the Spine

As mentioned earlier, many previous work on the segmentation of the spinal canal/cord in the literature are devoted to the modality of MR images [12]–[14], partly due to the better capability of MR imaging in rendering and differentiating soft tissues. For example, Koh *et al.* [15] utilized the gradient vector flow field [16] to segment the spinal cord for a computer-aided diagnosis system. A semi-automatic method was proposed in [17], such that the cord surface is acquired based on the manually-specified cord center line and through the active surface model with intrinsic smoothness constraints. Chen *et al.* [18] combined deformed atlas and topology preserving classification to address the segmentation of the spinal cord. Recently, Law *et al.* [19], [20] applied the gradient competition anisotropy technique to segment the spinal cord as well as to extract its center line.

Instead of the MR modality, tremendous efforts are also devoted to segmenting the spines in CT images, which is the focus of this paper as well. Several related literature reports fall into the category of semi-automatic segmentation [21]–[23], which may not be easy to be applied upon large-scale image cohorts. Alternatively, Archip *et al.* [7] presented an automatic segmentation pipeline by parsing anatomical objects in a recursive and top-down manner. Similarly, a top-down parcellation strategy was also adopted by [24], which utilized watershed and graph search to segment the spinal canal. In general, the top-down solution relies on the identification of the spine column for rough but necessary spatial reference, while the detection of all these anatomical structures can be non-trivial nevertheless.

B. Seeding and Interactive Segmentation

Interactive segmentation methods have developed rapidly in past decades and are widely applied in the area of medical image analysis. A recent survey of interactive segmentation and its applications in medical images can be found in [25]. Though not fully automatic, these methods are able to effectively conquer the difficulties challenging many automatic segmentation methods by requiring only minimal human interaction. The proposed method in this paper is also inspired by the principle of

interactive segmentation as the seeding voxels are iteratively adjusted for better segmentation. However, the fundamental difference of our method is that the interaction (i.e., the determination of the re-placed seeding voxels) is fully automatic, instead of depending on the inputs from human experts.

Many interactive segmentation methods (e.g., the graph-cut methods [26], [27]) regard the seeding voxels as the boundary condition, and solve for the segmentation in the energy-minimization style. The random-walk solver [11], for example, asks human experts to specify seeding voxels of different labels, which correspond to "spinal-canal" and "not-spinal-canal" in our case. Then, different labels are assigned to non-seeding voxels by embedding the image into a graph and utilizing intensity similarities between voxels. Human experts can further edit the placement of seeds until the desired quality of the segmentation is achieved. Although it is convenient to apply the random-walk method for interactive segmentation of the spinal canals, the interaction may cost high in terms of both expert training and lengthy processing of each image dataset.

C. Discriminative Learning and Image Segmentation

Discriminative learning is able to capture the statistical dependence of an unobserved variable (e.g., the label of the segmentation) on an observed variable (e.g., the vector of features extracted from the image) via the modeling of the posterior probability. The technique has thus become a popular choice for detecting anatomical objects (including landmarks points, lines/curves, and organs) in medical images [28]. Many supervised learning methods are applicable to the segmentation task, e.g., boosting [29] and random forest [30]. Multiple classifiers can also be combined into more complex structures (e.g., tree [31] and network [32]) for better performances. For example, in our previous work [33], voxelwise Haar-like features [34] were extracted and fed into a probabilistic-boosting-tree (PBT) classifier [31] to segment the ribs in CT images. Satisfactory results are sometimes hard to achieve in one-shot learning-based segmentation, especially concerning the complex appearance information in medical images. To this end, context information, which is modeled as the posterior probability in discriminative learning, can significantly improve the segmentation quality by feeding itself to subsequent learned classifiers [35].

III. METHOD

We treat segmenting the spinal canal as a typical binary segmentation problem. Let $p(x)$ denote the probability of the voxel x being foreground (i.e., inside the spinal canal) and $\bar{p}(x)$ the probability for being background (i.e., outside the spinal canal), respectively. In general, we have $p(x) + \bar{p}(x) = 1$ after proper normalization. The binary segmentation of the spinal canal can then be acquired by applying a confidence threshold to the map of $p(x)$ within the entire image space. For example, if $p(x) > \bar{p}(x)$, the voxel x is naturally regarded as being part of the foreground. On the contrary, if $p(x) < \bar{p}(x)$, then x belongs to the background. Note that the to-be-estimated $p(x)$ is *de-factor* a posterior probability, which is conditioned on the observation (of the features) upon the voxel x.

The probability map can be generated in different ways. For example, by providing foreground and background seeding

voxels automatically, we can acquire the probability map through the random-walk solver as in our method. The seeding voxels are determined according to the iteratively refined topology of the spinal canal. Specifically, to initiate the random-walk-based segmentation, we start from the supervised *voxelwise classification* to identify a small set of voxels, which are assigned with very high classification confidences and thus most likely to be foreground. The identified voxels act as positive seeds and are fed into the *random-walk solver* to generate a conservative binary segmentation with relatively low sensitivity but also low false-positive (FP) error. Concerning the fact that the spinal canals are generally tubular structures even though their shapes vary significantly across the patient population, both geometric and appearance constraints that are anatomically meaningful are then enforced to extract and to refine the topology of the segmented spinal canal. The refined topology, which is continuous and smooth, thus leads to the updated placement of the seeding voxels, which in turn increases the sensitivity in segmenting the spinal canal via random walk. By iteratively feeding the improved seeding voxels to the *cascaded random-walk solvers*, we have successfully built an automatic pipeline that yields satisfactory segmentation of the spinal canal. Details of our method will be provided as follows. For easy understanding, the flowchart of our method, as well as explanations to its key steps, is shown in Fig. 2.

We categorize our method as being a *bottom-up* solution, which is significantly different from the conventional *top-down* methods. Specifically, we identify a set of seeding voxels where the tentative segmentation is mostly reliable, instead of detecting the anatomical objects of much larger sizes (e.g., organs near the spine). The seeding voxels are initially detected through appearance-based discriminative learning [33] and then updated according to the (tentative) spine topology. The segmentation result, as well as the spine topology, can be improved in subsequent iterations. In this way, the segmentation of the spinal canal can accumulate or propagate from a few (seeding) voxels, until reaching satisfactory result throughout the entire image space.

A. Voxelwise Classification

In order to identify highly reliable foreground voxels as positive seeds for the random-walk solver, we turn to voxelwise classification based on discriminative learning. The classification consists of the *training* and the *testing* stages. In the *training* stage, we utilize the PBT-based classifiers to establish the dependence between the visual (appearance) features of voxels and their segmentation labels. Each new subject is then passed into the *testing* stage for determining the segmentation of its spinal canal.

In particular, we have manually annotated the medial lines of the spinal canals on the training set of 20 images. The training images are randomly selected from our dataset, while each of the eight clinical sites contributes at least one training image. The medial line is defined to connect the medial points of the spinal canal on all traverse slices. Voxels exactly along the medial lines are used for the samples of the foreground, while background samples are obtained from a constant distance away to the medial lines.
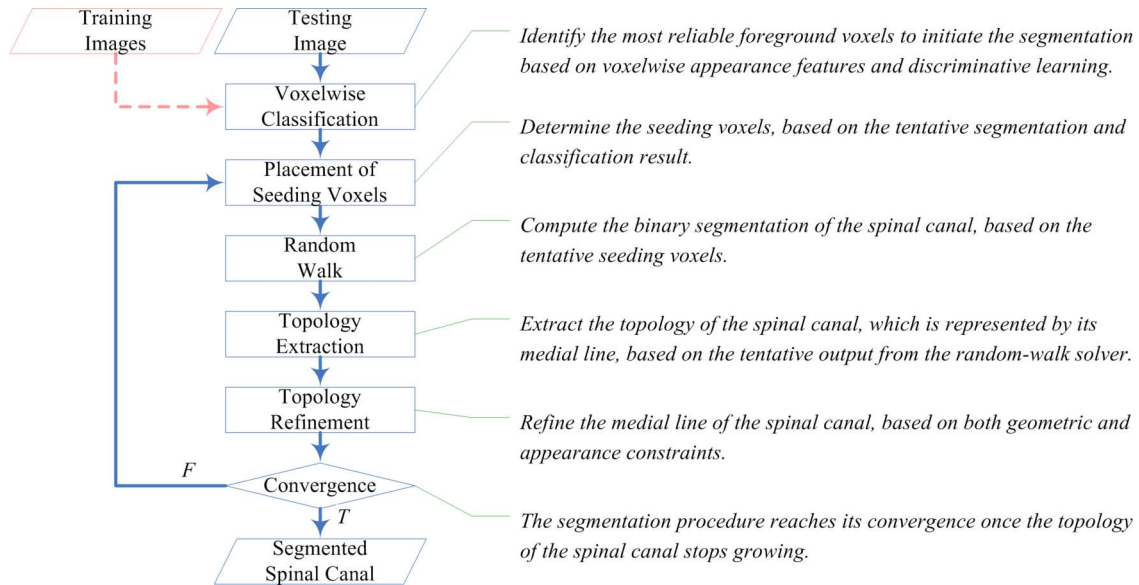
Fig. 2. The flowchart of our method.

We further use 3D Haar-like features as voxel descriptors [29], [34], [35]. The Haar-like features are computed from the neighborhood of each sample voxel under consideration, thus representing the local appearance of the voxel. In particular, we employ 10 filter templates (cf. Fig. 3) to extract Haar-like features. For each filter template, five scales (i.e., 4, 10, 16, 22, and 28 in voxels) and 11 translations (i.e., 0, ± 3, ± 6, ± 9, ± 12, and ± 15 in voxels) along each axis are used. The extraction is restricted within the block of $31 \times 31 \times 31$ voxels. In this way, we can acquire up to 91,594 Haar-like features for each voxel in the image. Note that our Haar-like features are not computed from a complete Haar wavelet band.

The PBT classifiers are then trained with AdaBoost nodes [31]. In the other word, each node of PBT corresponds to a certain AdaBoost classifier. Instead of using only one PBT classifier, we build a coarse-to-fine pyramid of PBT classifiers. That is, coarse features are used for PBT-based training/testing first, followed by features of fine scales. Totally three scales are used. In this way, we can significantly speed up our method by downsampling images in earlier stages of the classification and thus reducing the number of training/testing samples. Meanwhile, we can better exploit coarse-scale features reflecting longer range spatial context information in coarse resolution, which is usually limited by conventional Haar filters. Note that the same strategy has been successfully applied in our previous studies, e.g., bone extraction [33].

The well-performing foreground voxel confidence map (as well as the measuring color bar) with respect to a certain training subject is displayed in Fig. 4(a). However, when the classifier is applied to a new testing subject, the classification result may suffer from both false-negative (FN) and FP errors. Typical classification errors are highlighted by the red arrows in Fig. 4(b)–(c). For instance, an FP artifact is highlighted in Fig. 4(b), while Fig. 4(c) shows the discontinuity of the foreground confidence which is caused by FN errors. Since the purpose of voxelwise classification is to identify highly reliable
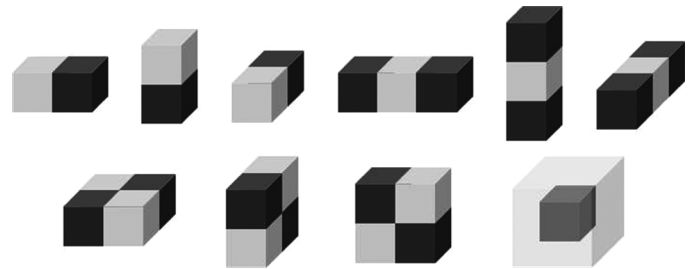

Fig. 3. There are 10 filter templates for extracting 3D Haar-like features. Each Haar-like feature is calculated as the difference between the sum of intensities inside the grey region and the the sum of intensities inside the white region.
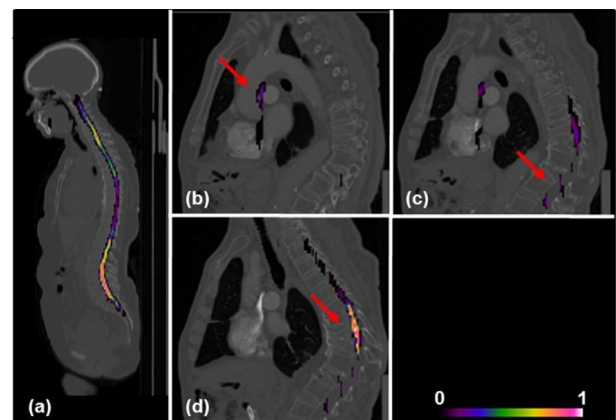

Fig. 4. Panel (a) shows the confidence map yielded by voxelwise classification on a *training* subject; panels (b)-(d) are for the voxelwise confidences of another *testing* subject. Among them, FP errors and FN errors are highlighted in (b) and (c), respectively. We use a high confidence threshold to preserve reliable foreground voxels only as in (d).

foreground voxels only (i.e., those highlighted in Fig. 4(d)), we have applied a high confidence threshold (i.e., $p(x) > 0.9$) empirically to suppress most FP errors. The binary segmentation, which will be iteratively refined by the subsequent cascaded

random-walk solvers, is thus tentatively generated after thresholding the confidence map yielded by voxelwise classification.

B. The Random-Walk Solver

Similar to the PBT-based voxelwise classification, the random-walk solver also produces voxelwise probability of being foreground/background [11]. After users have specified foreground/background seeds, the random-walk solver departs from a certain non-seeding voxel and calculates its probabilities to reach foreground and background seeds, as p_x and $\bar{p}_x$, respectively. Usually the non-seeding voxel x is assigned to foreground if $p_x > \bar{p}_x$. In terms of random walk, the image is embedded into a graph where vertices correspond to individual voxels and edges link neighboring voxels. The weight w_{xy} of the edge e_{xy}, which measures the similarity between two neighboring voxels x and y, is defined as

$$w_{xy} = \exp(-\beta(I_x - I_y)^2), \qquad (1)$$

where I_x and I_y represent intensities at two locations; β is a positive constant. Assuming that the segmentation boundaries are well aligned with intensity changes, the random-walk solver then aims to estimate p_x that minimizes the following energy function

$$E = \sum_{\forall e_{xy}} w_{xy}(p_x - p_y)^2. \qquad (2)$$

To optimize the above is equivalent to solving a Dirichlet problem with boundary conditions that are defined by the seeding voxels. Specifically, p_x is set to 1 if x is a foreground seed, and 0 for background. The calculated p_x incorporates spatial information across neighboring voxels, which differs from the independent voxelwise classification.

The probability of each voxel in random walk is associated with the paths from the voxel to seeds. Hence, p_x is dependent not only on the weights of the edges forming the path but also the length of each path. This property potentially undermines the random-walk solver, which could be sensitive to the placement of the seeding voxels. In the toy example of Fig. 5(a), there are three vertical stripes. The intensity of the middle stripe is slightly higher than two side stripe, in order to approximates the spinal canal surrounded by other tissues in CT data. We highlight certain sections of stripe boundaries in very high intensity to simulate the existence of vertebra, whose presence can also be discontinuous. Foreground seeds and background seeds are colored in red and green, respectively. The probability p_x calculated by the random-walk solver and the binary segmentation ($p_x > 0.5$) are shown in Fig. 5(b) and (c), respectively. We observe from Fig. 5(c) that the segmented foreground falls into two disconnected segments undesirably.

Though increasing the threshold on p_x and tuning β to modify the edge weights might improve the segmentation results, it would make the method too ad-hoc to be practically usable for large-scale image cohorts. On the other hand, the random-walk solver provides an interactive remedy by simply allowing human experts to place more seeds in proximity to the desired
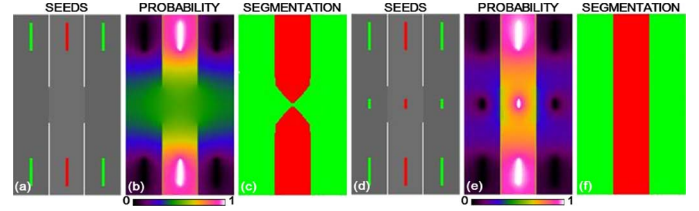

Fig. 5. With foreground seeds (in red) and background seeds (in green) (a), the calculated probabilities (b) and the corresponding binary segmentation (c) are not satisfactory. However, by manually placing more seeds (d), the segmentation results (e-f) are improved significantly.
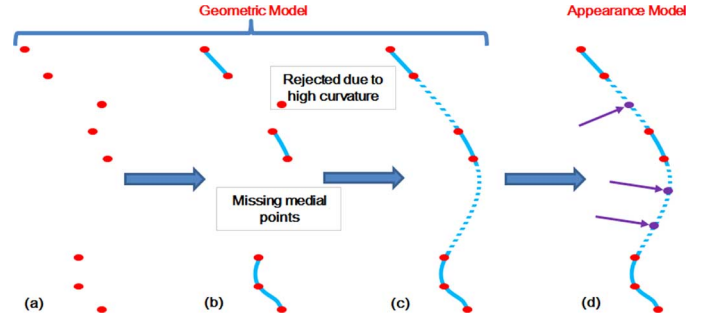

Fig. 6. Four sub-steps in refining the topology of the spinal canal include (a) estimating medial points; (b) determining medial segments; (c) calculating virtual medial segments (in dotted curves); and (d) placing more virtual medial points (in purple).

segmentation boundaries. By adding a few more seeds as in Fig. 5(d), the results clearly show better discrepancy of foreground/background and are thus much better (Fig. 5(e)–(f)). In our method, we automatically update the placement of the seeding voxels according to the topology of the spinal canal, in order to attain satisfactory segmentation in the final.

C. Spine Topology and Cascaded Random-Walk Solvers

As mentioned above, we are able to identify the *initial* seeding voxels through the voxelwise classification step. If we feed these seeding voxels to the random-walk solver in one-shot style directly, the generated binary segmentation usually breaks into several disconnected segments. The reason is similar to the case in Fig. 5(a)–(c). The initial seeding voxels are required to be highly reliable, such that they are not sufficient to cover the region of the entire spinal canal. To this end, we extract the topology of the tentatively segmented spinal canal. Then, the topology is refined for better placement of the seeding voxels, which is further utilized by the following random-walk solver. The procedure above is iteratively invoked, such that several random-walk solvers can be cascaded for the final segmentation.

1) Topology Extraction and Refinement: To acquire the complete segmentation of the spinal canal, we introduce the topology constraints. Specifically, we use the medial line of the spinal canal to represent its topology. After calculating all segments of the medial line given the tentative segmentation, we can interleave them into a single connected curve. Fig. 6 illustrates the four sub-steps to extract and to refine the topology of the spinal canal.

1) We first calculate the medial point of the foreground voxels on each transverse slice as in Fig. 6(a), based on the tentative segmentation result. The medial point is defined to have the least sum of distances to all other foreground voxels on each slice. In this way, most medial points are allocated within the spinal canal, though some of them might be affected by segmentation artifacts and are determined to be outside the spinal canal incorrectly.

2) Assuming that the medial line connects all medial points, we then connect the detected medial points into several segments in Fig. 6(b). The medial line may break into several segments since the medial points can be missing on certain traverse slices. Also, a curvature threshold (i.e., $30°$ over 4 mm physical spacing) is imposed along the connected medial line, such that those medial points incurring too sharp turns to the medial line would be rejected as outliers. The reason is that the topology of the spinal canal, which is described by the medial line, has to be smooth from the physiologic perspective.

3) With all segments computed in the above, we can interleave them by filling gaps with smooth virtual segments as represented by dotted curves in Fig. 6(c). Each virtual segment $c(s)$ minimizes $\int \|\nabla^2 c(s)\|^2 ds$ to keep the medial line smooth. Here, $s \in [0, 1]$ indicates the normalized arc-length. Note that the stationary solution to the above holds when $\nabla^4 c(s) = 0$, while the Cauchy boundary conditions are defined by both two ends of the virtual segment, as well as tangent directions at the ends. Though the numerical solution is non-unique, we apply the cubic Bézier curve for the fast estimation of the virtual segment. In particular, for a certain virtual segment, we denote its two ends as $\mathbf{P}_0$ and $\mathbf{P}_3$. An additional control point $\mathbf{P}_1$ is placed so that the direction from $\mathbf{P}_0$ to $\mathbf{P}_1$ is identical to the tangent direction at $\mathbf{P}_0$. Similarly, we can define $\mathbf{P}_2$ according to $\mathbf{P}_3$ and the associated tangent direction. We further require that the four control points are equally spaced. The virtual segment is then generated following

$$c(s) = (1-s)^3\mathbf{P}_0 + 3(1-s)^2 s\mathbf{P}_1 + 3(1-s)s^2\mathbf{P}_2 + s^3\mathbf{P}_3. \quad (3)$$

4) After predicting the virtual segment in Fig. 6(c), we finally can place more virtual medial points along the virtual segment. In addition to the subject-specific geometry constraints to keep the virtual segments smooth (cf. Fig. 6(a)–(c)), appearance model is introduced in Fig. 6(d) for the determination of qualified virtual medial points. To this end, we calculate the intensity mean and the standard deviation (STD) upon all existing medial points (red dots). The univariate Gaussian intensity model allows us to simply examine whether a new voxel is highly possible to be foreground given its intensity value. In particular, we start from both two ends of each virtual segment, and admit virtual medial point (purple dot) if its intensity is within the single STD range of the intensity model. The process to admit virtual medial points continues until reaching a disqualified candidate.

2) Placement of Seeding Voxels: After the topology of the spinal canal has been extracted and refined, we are able to pro-vide better seeds for the random-walk solver to use. All voxels along the refined medial line, including the newly admitted virtual segments, act as foreground seeds. Moreover, we qualify more voxels as foreground seeds if (1) they have been classified as foreground in previous segmentation; (2) their intensities are within the single STD range of the appearance model introduced above; and (3) they are connected to the medial line through a path passing foreground seeds only. In this way, we have inherited previous segmentation in the areas where the confidence is high, and saved computation since the random-walk solver can simply treat them as boundary conditions. Surrounding voxels with intensities higher than an empirical threshold are mostly located in the bone areas and regarded as the background seeds.

3) Cascaded Random-Walk Solvers: By repeating the refinement of the spine topology and the placement of the seeding voxels, we can cascade several random-walk solvers for the sake of the final segmentation result. Each random-walk solver in the automatic pipeline is able to improve the tentative segmentation, based on the spine topology refined from the previous segmentation. The iterative procedure terminates when the topology of the spinal canal, or the length of the medial line, becomes stable. In particular, we allow the medial line to grow at its both ends. The growth of the segmentation can stop automatically at the inferior (tail) end, since the medial line cannot penetrate the bones due to the appearance constraint (cf. Fig. 6(d)). At the superior (head) end, a maximal diameter of the segmented spinal canal is defined to stop the segmentation entering the skull. Also, it is worth noting that isolated medial segments may exist in the segmentation result, while these segments are removed from the final segmentation in our post-processing. These undesired segments are usually caused by the segmentation artifacts that are associated with leg marrows, as the legs are incorporated within the superior-inferior coverage of certain images in our dataset. An example will be discussed in the experiment section (cf. Fig. 11).

IV. EXPERIMENTAL RESULTS

We validate the proposed method using the large-scale and challenging dataset, which consists of 110 images (20 for training and 90 for testing). In general, all 110 images are collected with low voltage/dose protocols using Siemens Biograph PET-CT scanners.

- The intra-slice resolution ranges 1.2–2.0 mm. In terms of inter-slice thickness, there are 61 images of 3.0 mm and 42 images of 5.0 mm. The inter-slice thickness of the rest images is not greater than 1.0 mm (i.e., 0.6 mm, 0.75 mm, and 1.0 mm in particular). Note that all images are resampled to the isotropic spacing of 2.0 mm in our pre-processing.
- The current is 25–50 mAs. For the cases of low inter-slice thickness (i.e., ≤ 1 mm), the voltage is 120 kVp. For all other images, the voltage is 80–100 kVp.
- We have 10 cases of thoracic, 74 cases of full (including C1-L5 and partial sacrum), and 26 cases of partial full (starting from C1-C8 and ending at L1-L5).
- We visually observe pathologies for 52 out of 110 images. Specifically, there are nine patients with bone lesions (19 sclarotic and five lytic lesions in total), 29 patients of osteoarthritis (59 vertebrae in the early-moderate
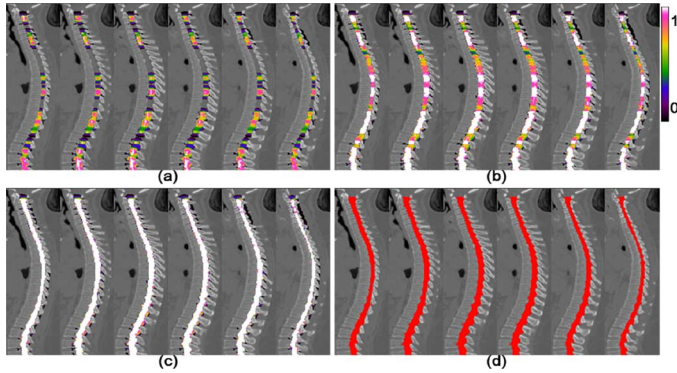
Fig. 7. Panels (a)-(c) show the foreground probability maps on six consecutive slices of a certain subject yielded after the first, second, and the final (fourth) iteration of random walk, respectively. The binary segmentation in (d) is acquired by thresholding the final probability map in (c).

stage and 60 vertebrae in the advanced stage), seven patients with compression fractures (21 vertebrae in total), 19 patients with scoliosis (two severe cases of whole spine, and 17 moderate cases of local sections), and 16 cases with stenosis (affecting 54 vertebrae in total). Among them, we identify 29 (out of 52) cases with multiple pathologies.
- All subjects are late teens to adults. The age is 52.4 ± 12.1 years in average, ranging from 19 to 76 years.
- There is no inclusion or exclusion protocol for patients with special conditions. In particular, there are two patients with metallic devices (i.e., four and six screws on the lower spine sections, respectively).

For voxelwise classification, we use the same setting with [33] in both training and testing. The implementation of the random-walk solver is contributed by [11], while default settings are applied (e.g., $\beta = 900$). Parameters and configurations are fixed across all testing subject images, for better feasibility to apply our method toward real clinical data.

A. Demonstration of the Iterative Solution

We first use an example to demonstrate the necessity of the proposed iterative solution. In Fig. 7(a), we show the foreground probability on six consecutive slices of a subject after the first iteration of random walk. Obviously the segmentation breaks into several disconnected segments, due to the limited number of seeding voxels provided by voxelwise classification. After the second iteration, however, the tentative probability map becomes much better as in Fig. 7(b). The reason is that the topology of the spinal canal, which is represented by its medial line, has been extracted and much refined. Then, better placement of the seeding voxels leads to improved segmentation via the random-walk solver. The final probability map after the fourth iteration is shown in Fig. 7(c), with the binary segmentation in Fig. 7(d). The results above demonstrate that our method can efficiently utilize the topology of the spinal canal and generate satisfactory segmentation.

B. Large-Scale Validation

Our work is among the largest-scale validations of the segmentation of the spinal canals, as totally 90 images are utilized for testing. Also, the images in our collection are highly diverse
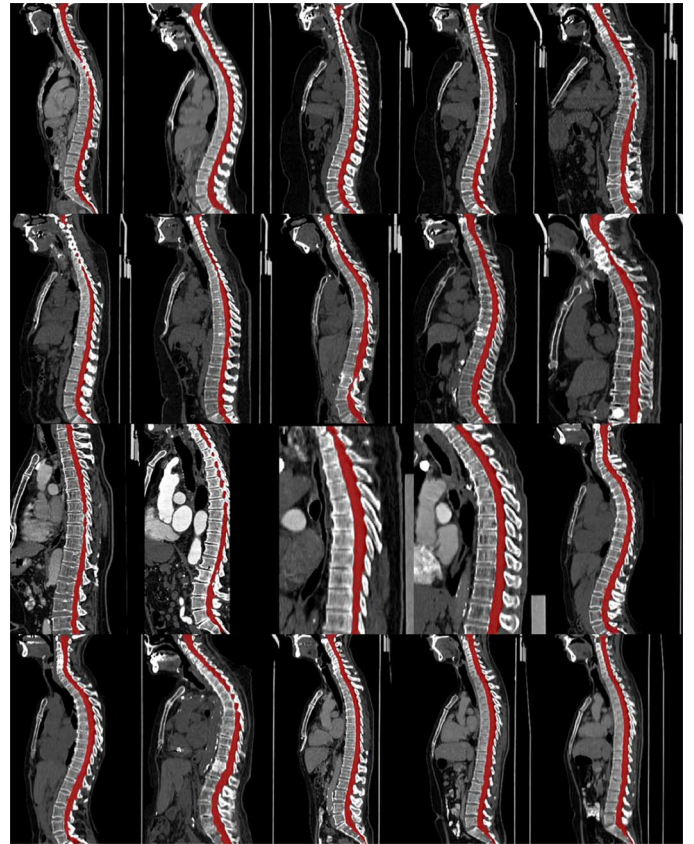

Fig. 8. Typical segmentation results of the subject images acquired from eight different clinical sites. Note that all images are isotropically resampled and zoomed to the areas of the segmented spinal canals.

in terms of their superior-inferior coverage, sizes, original spatial resolutions, etc. All images are acquired from the PET-CT scanners, implying that their quality is relatively lower compared to regular CT scans. Though facing all these difficulties, our method has successfully segmented the spinal canals on all testing cases. The results are affirmed by visually inspecting the segmentation results.

In Fig. 8, we show the typical slices (in coronal views) of the segmentation results on 20 testing patients. Note that the selected images cover all eight different clinical sites involved in this study. Although pathologies are prevalent across the entire image cohort, the appearances of the spinal canals for the selected patients are relatively normal in this figure. We observe that the segmented spinal canals colored in red are clearly reasonable. To facilitate readers' understanding, all images shown in the figure are isotropically resampled although their sizes and spatial resolutions may vary a lot. Moreover, we zoom in the areas of the segmented spinal canals in the figure, while other parts within the image are discarded.

Quantitative evaluation is important to judge the merits of the proposed method. To this end, we select 20 testing images and partly annotate their spinal canals under the traverse views. No noticeable pathologies are observed for these selected testing images. The annotation is conducted on the slices where the spinal canal is fully encircled by the neighboring bone structures. We then measure the Dice overlap ratio between the segmentation estimated by our method and the manual annotation.
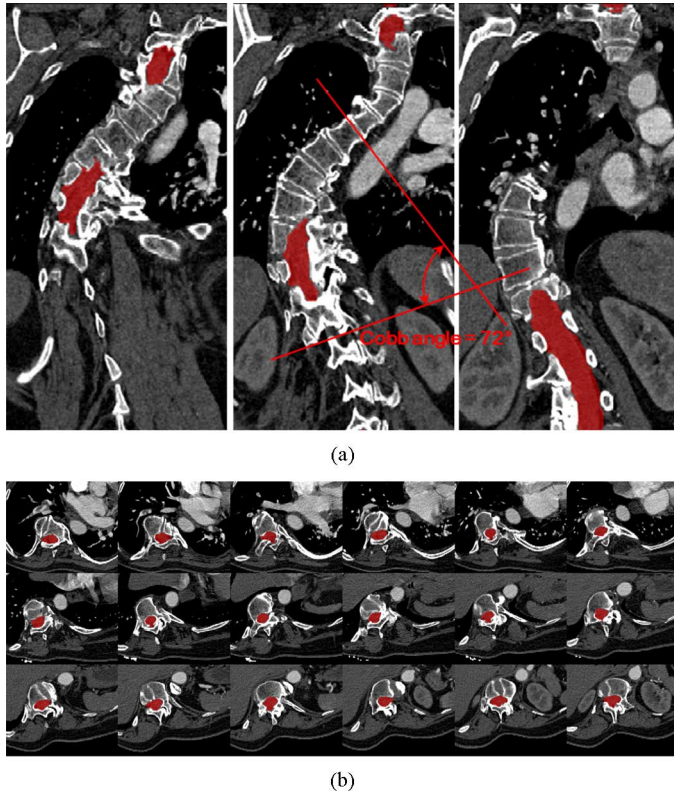
(a)

(b)

Fig. 9. The segmentation result in three consecutive coronal views (a) and 18 traverse views (b) of a difficult case. The patient is affected by the disease of scoliosis, which twists the spine in the lateral axis (left-right) orientation significantly. Note that the Cobb angle [37] for this patient reaches $72°$.

Suppose the manually annotated region of the spinal canal is T and the estimated segmentation by our method is M, the Dice overlap ratio is then computed as $2\|T \cap M\|/(\|T\| + \|M\|)$. The operator $\|\cdot\|$ calculates the volume of the respective region. The Dice overlap ratio regarding our method is $92.79 \pm 1.55\%$. The results indicate that the segmentation of the spinal canal by the proposed method is highly accurate and reliable.

Our method converges to the final segmentation within 2–5 iterations automatically for all testing subjects. Typically, it costs 20–60 seconds to segmented each subject depending on the size of the image. The time cost was measured on a DELL workstation with Intel 2.27 GHz Xeon platform and 2.75 GB RAM. We note that, with more sophisticated random-walk solver that is specifically designed for fast editing of the seeding voxels [36], the speed performance of the proposed method can be further improved.

C. Case Analysis

An extreme case is shown in Fig. 9. The patient is affected by the disease of scoliosis. Thus, an unusual twist along the lateral axis (left-right) direction is introduced to the spine. In Fig. 9(a), three consecutive coronal views of the patient is provided. We further extract 18 consecutive traverse slices near the pathological location as in Fig. 9(b). Although the topology of the spinal canal under consideration is abnormal, our method is still capable of well segmenting the whole structure. Visual inspections can confirm the success of the proposed method.
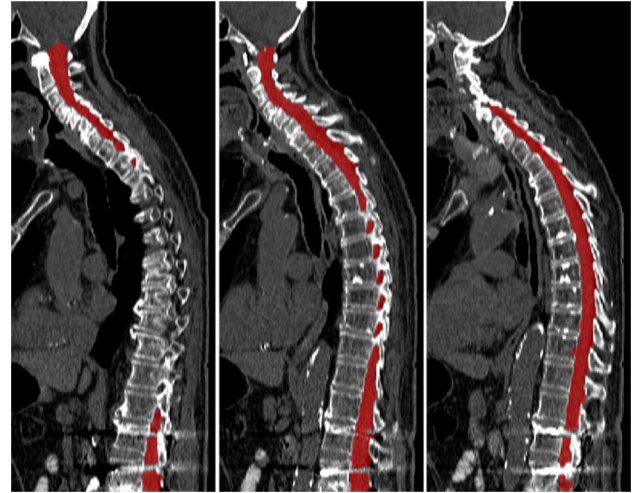

Fig. 10. The segmentation result on a patient with metal screws. The screws are clearly visible on the bottom of the figure.
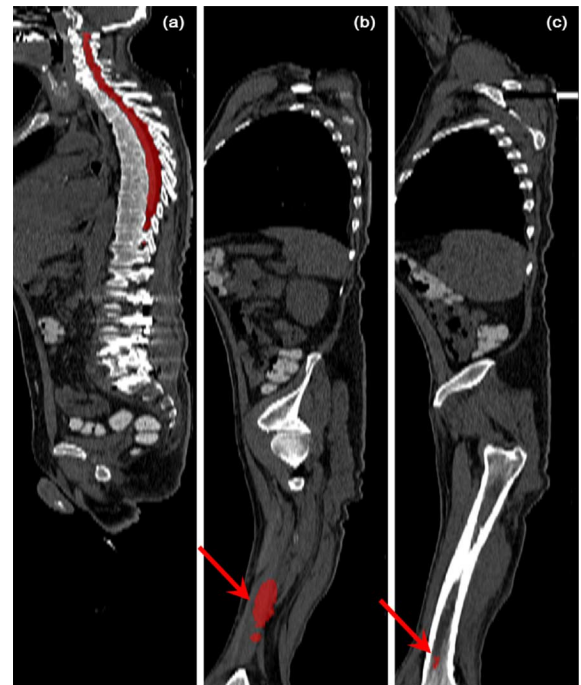

Fig. 11. Besides the spinal canal (a) in a full-body scan, false segmentation may occur especially in the areas of legs (b and c). The false segmentations can be pruned in post-processing.

Next, we show a case where metal screws are visible. From Fig. 10, we conclude that the segmentation result is still satisfactory. It is worth noting that, for most existing studies, images with metallic inplants (i.e., screws) are usually excluded. In our work, however, all images are collected from clinical routines without any special inclusion/exclusion protocol. Our method has demonstrated its capability on these difficult cases.

We use the third case to discuss a common pattern of false segmentation, which can be handled by the proposed method. In Fig. 11, we show three sagittal slices from a full-body scan. The segmented spinal canal is clearly visible in Fig. 11(a). However, as shown in Fig. 11(b) and (c), two isolated areas of false segmentation would occur if we directly thresholded the probability

TABLE I
COMPARISONS BETWEEN OUR METHOD AND STATE-OF-THE-ART METHODS

Method	Automation	Number of Testing Images	Accuracy and Robustness
Burnett et al. [21]	Semi	6	Succeed on 91% slices.
Nyúl et al. [23]	Semi	27	The mean Dice ratio is 93.66%.
Archip et al. [7]	Full	23	Succeed on 85.3% slices; fully succeed on 4.35% images.
Yao et al. [24]	Full	71	Succeed on 97.18% images.
Our work	Full	90 (another 20 for training)	Succeed on all images; the mean Dice ratio is $92.79 \pm 1.55\%$.

map yielded by the random-walk solver. The false segmentation pattern usually occur in legs, which are incorporated into the superior-inferior coverage of certain images in our dataset. To this end, these false segmentation needs be pruned in post-processing. That is, only a single connected segmentation area of the largest volume is preserved, while other isolated segmentation areas are regarded as false segmentation.

D. Comparisons

We compare the proposed method with four alternative literature reports, in order to further demonstrate the capability of our method. In general, the methods in Burnett *et al.* [21] and Nyúl *et al.* [23] belong to the semi-automatic category, which requires user input for the segmentation of the spinal canal. The other two methods, i.e., Archip *et al.* [7] and Yao *et al.* [24], are fully automatic. Note that all four methods under comparisons are not specially designed to address or to validate upon images affected by severe pathologies, while the proposed method has successfully been applied to all of our real clinical data. The comparisons between our method and all four methods are summarized one-by-one as follows.

1) Burnett et al. [21]: The testing was conducted upon only six patients, the number of which was much lower than the validation of our method (i.e, 90 testing images). The segmentation results were then reviewed by two oncologists. In overall, the segmentation of the spinal canal was successful on 91% upon all 557 slices (of six images) related to the spines. The segmentation was unsuccessful on 2% slices, while the other 7% results required further manual editing. Note that, with respect to our dataset, the proposed method has generated satisfactory segmentation results on all slices.

2) Nyúl et al. [23]: The method required minimal inputs from human experts for the localization of the spinal canal and the staring position of the segmentation task. Then, segmentation errors were quantitatively measured against the gold standard (i.e., manual segmentation) on 27 images. According to the reported data, we compute the mean Dice overlap ratio which is equal to 93.66%. Note that the score of the proposed method upon our dataset is $92.79 \pm 1.55\%$. The two scores are within the range of a single standard deviation. To this end, we argue that our method is comparable to Nyúl *et al.* in terms of the segmentation accuracy, though the latter is not fully automatic.

3) Archip et al. [7]: There were 23 images for testing the accuracy. The segmented contours of the spinal canals on all slices were examined. Averagely, the correct segmentation rate is 85.3%. In particular, the segmentation results were fully correct on only one (out of 23) image, while errors were observed for all other images. Also, it was admitted that the segmentation of the spinal canal was difficult, as the segmentation accuracy was lower than the segmentation of neighboring anatom-

ical structures (e.g., lamina). Again, we want to emphasize that the segmentation results are acceptable for all images and all slices in our experiment.

4) Yao et al. [24]: There were 71 testing images in Yao *et al.*, thus ranking the scale of its validation next to ours. The segmentation results were reported to be successful on 69 (out of 71) images, while the failures were attributed to small vertebrae and large curvature in the spine. Note that, however, our method has been successfully applied to all testing images, while our data collection is highly challenging as well.

It is worth noting that the above comparisons are based on reported results in the literature, rather than being strictly upon the same dataset and experimental setting. Therefore, it is hard to derive a rigorous conclusion to assert the superiority of individual methods. However, we believe that our method yields satisfactory segmentation results especially concerning our highly challenging dataset. In particular, we note that all four methods under comparisons could fail occasionally, indicating that their robustness can be severely challenged in certain situations. In our experiment, however, the proposed method is successfully applied upon all testing images under consideration. For easy understanding, we have also summarized the comparisons as in Table I.

V. CONCLUSION AND DISCUSSION

In this work, an automatic method to segment the spinal canals from highly varying CT images is proposed. With initial seeding voxels that are provided by PBT-based voxelwise classification, we introduce the topology constraints into the segmentation via random walk. Our iterative optimization has successfully enhanced the capability of a single random-walk solver in dealing with tubular spinal canals, in that the boundary conditions (i.e., the placement of the seeding voxels) can be iteratively improved to provide better segmentation results. Our large-scale evaluation shows that the proposed method is highly accurate and robust even if the datasets are very diverse and challenging.

Due to limited accesses to state-of-the-art methods, rigorously fair comparisons can hardly be conducted. Though we compared our method with four other methods reported in the literature, it is worth noting that the comparisons were based on different datasets that were used by individual papers. State-of-the-art methods typically reported failures when segmenting certain *outlier* images. However, it is worth noting that our method has been successfully validated on all 90 testing images. That is, no failure case has been generated through the proposed method. Concerning the challenges caused by our large-scale cohort of diverse images, we argue that our method is very robust and accurate for the segmentation of the spinal canal.

Although our method has demonstrated its robust and accurate segmentation capability upon the large-scale dataset, we are aware that the proposed method could be challenged in certain situations.

- Our images are collected from Siemens Biograph PET-CT scanners with no IV contrast. Although no inclusion/exclusion protocol has been used for the collection of the images, the variation within the dataset (i.e., patient age, pathology, scanning protocol) could still be limited. The performance of the proposed method upon other large-scale datasets is not investigated yet.
- We have assumed relatively homogeneous appearances for foreground voxels, such that individual segments of the extracted medial line can be refined into a single one. In clinical practice, however, there are certain diseases (e.g., calcified vertebral foramen) which may introduce inhomogeneous appearances into CT images of the spine. In this case, our method could possibly fail.
- Our images are acquired from adult patients only. For paediatric patients whose spine sizes may vary from adults, our method has not been validated yet.
- Our method could possibly well handle spine appearance anomalies (including lumbarisation, sacralisation, disc degeneration, bone spurs, surgical spinal-fusion, etc.). We only have very simple assumption regarding the appearance of the bones, i.e., with relatively high intensities for bone voxels especially on the boundaries. However, the exact effect of spine appearance anomalies cannot be investigated at this moment, as no such images are available in our dataset.

In general, we conclude that the proposed method can well handle our large-scale dataset, which is collected from clinical routines of multiple sites.

There are two directions in our future work. *First*, we will improve the speed performance of our method. In our current implementation particularly, the computation of each random-walk solver in the cascaded pipeline is independent. In fact, only boundary conditions of two consecutive random-walk solvers change during the segmentation process. Therefore, we would be able to reduce the redundancy in computation, e.g., by using the method reported in [36]. Second, we will apply our method to more related segmentation problems. In particular, we will probe the possibility of segmenting spinal canals/cords from magnetic resonance (MR) images via the proposed method.

REFERENCES

[1] T. Klinder *et al.*, "Automated model-based vertebra detection, identification, and segmentation in CT images," *Med. Image Anal.*, vol. 13, no. 3, pp. 471–482, 2009.

[2] K.-S. Seo, L. C. Ludeman, S.-J. Park, and J.-A. Park, "Efficient liver segmentation based on the spine," in *Advances in Information Systems*. New York: Springer, 2005, pp. 400–409.

[3] J. Ma and L. Lu, "Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model," *Comput. Vis. Image Understand.*, vol. 117, no. 9, pp. 1072–1083, 2013.

[4] S. Benameur *et al.*, "3d/2d registration and segmentation of scoliotic vertebrae using statistical models," *Comput. Med. Imag. Graph.*, vol. 27, no. 5, pp. 321–337, 2003.

[5] A. Yezzi, L. Zöllei, and T. Kapur, "A variational framework for integrating segmentation and registration through active contours," *Med. Image Anal.*, vol. 7, no. 2, pp. 171–185, 2003.

[6] M. Hardisty, L. Gordon, P. Agarwal, T. Skrinskas, and C. Whyne, "Quantitative characterization of metastatic disease in the spine. Part I. Semiautomated segmentation using atlas-based deformable registration and the level set method," *Med. Phys.*, vol. 34, no. 8, pp. 3127–3134, 2007.

[7] N. Archip, P.-J. Erard, M. Egmont-Petersen, J.-M. Haefliger, and J.-F. Germond, "A knowledge-based approach to automatic detection of the spinal cord in CT images," *IEEE Trans. Med. Imag.*, vol. 21, no. 12, pp. 1504–1516, Dec. 2002.

[8] J. Liu *et al.*, "Computer aided detection of epidural masses on computed tomography scans," *Comput. Med. Imag. Graph.*, 2014.

[9] P. J. Schleyer, T. Schaeffter, and P. K. Marsden, "The effect of inaccurate bone attenuation coefficient and segmentation on reconstructed PET images," *Nucl. Med. Commun.*, vol. 31, no. 8, pp. 708–716, 2010.

[10] J. H. Kim, J. S. Lee, I.-C. Song, and D. S. Lee, "Comparison of segmentation-based attenuation correction methods for PET/MRI: Evaluation of bone and liver standardized uptake value with oncologic PET/CT data," *J. Nucl. Med.*, vol. 53, no. 12, pp. 1878–1882, 2012.

[11] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.

[12] C. Chevrefils, F. Chériet, G. Grimard, and C.-E. Aubin, "Watershed segmentation of intervertebral disk and spinal canal from MRI images," in *Image Analysis and Recognition*. New York: Springer, 2007, pp. 1017–1027.

[13] J. Kawahara, C. McIntosh, R. Tam, and G. Hamarneh, "Globally optimal spinal cord segmentation using a minimal path in high dimensions," in *Proc. 2013 IEEE 10th Int. Symp. Biomed. Imag.*, 2013, pp. 848–851.

[14] O. Coulon *et al.*, "Quantification of spinal cord atrophy from magnetic resonance images via a b-spline active surface model," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1176–1185, 2002.

[15] J. Koh, T. Kim, V. Chaudhary, and G. Dhillon, "Automatic segmentation of the spinal cord and the dural sac in lumbar MR images using gradient vector flow field," in *Proc. Annu. Int. Conf. IEEE EMBS*, 2010, pp. 3117–3120.

[16] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[17] M. A. Horsfield *et al.*, "Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: Application in multiple sclerosis," *Neuroimage*, vol. 50, no. 2, pp. 446–455, 2010.

[18] M. Chen *et al.*, "Topology preserving automatic segmentation of the spinal cord in magnetic resonance images," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano to Macro*, 2011, pp. 1737–1740.

[19] M. W. Law, G. J. Garvin, S. Tummala, K. Tay, A. E. Leung, and S. Li, "Gradient competition anisotropy for centerline extraction and segmentation of spinal cords," in *Information Processing in Medical Imaging*. New York: Springer, 2013, pp. 49–61.

[20] M. W. Law, K. Tay, A. Leung, G. J. Garvin, and S. Li, "Intervertebral disc segmentation in MR images using anisotropic oriented flux," *Med. Image Anal.*, vol. 17, no. 1, pp. 43–61, 2013.

[21] S. S. Burnett, G. Starkschall, C. W. Stevens, and Z. Liao, "A deformable-model approach to semi-automatic segmentation of CT images demonstrated by application to the spinal canal," *Med. Phys.*, vol. 31, no. 2, pp. 251–263, 2004.

[22] G. Karangelis and S. Zimeras, "An accurate 3d segmentation method of the spinal canal applied to CT data," in *Bildverarbeitung für die Medizin 2002*. New York: Springer, 2002, pp. 370–373.

[23] L. G. Nyúl, "Method for automatically segmenting the spinal cord and canal from 3d CT images," in *Computer Analysis of Images and Patterns*. New York: Springer, 2005, pp. 456–463.

[24] J. Yao, S. D. O'Connor, and R. M. Summers, "Automated spinal column extraction and partitioning," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag., Nano to Macro*, 2006, pp. 390–393.

[25] F. Zhao and X. Xie, "An overview of interactive medical image segmentation," *Ann. BMVA*, vol. 2013, no. 7, pp. 1–22, 2013.

[26] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[27] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[28] S. K. Zhou, "Discriminative anatomy detection: Classification vs regression," *Pattern Recognit. Lett.*, vol. 43, pp. 25–38, 2014.

[29] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, vol. 1, pp. I–511.

[30] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[31] Z. Tu, "Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering," in *Proc. 10th IEEE Int. Conf. Comput. Vis.*, 2005, vol. 2, pp. 1589–1596.

[32] J. Zhang, S. K. Zhou, L. McMillan, and D. Comaniciu, "Joint real-time object detection and pose estimation using probabilistic boosting network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.

[33] D. Wu *et al.*, "A learning based deformable template matching method for automatic rib centerline extraction and labeling in CT images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 980–987.

[34] C. P. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 555–562.

[35] Z. Tu and X. Bai, "Auto-context and its application to high-level vision tasks and 3d brain image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1744–1757, Oct. 2010.

[36] S. Andrews, G. Hamarneh, and A. Saad, "Fast random walker with priors using precomputation for interactive medical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010.* New York: Springer, 2010, pp. 9–16.

[37] J. Cobb, "Outline for the study of scoliosis," *Instr. Course Lect.*, vol. 5, pp. 261–275, 1948.