Thoracic Spine Segmentation Based on CT Images

Gábor Révy, Dániel Hadházi, Gábor Hullám
Department of Measurement and Information Systems
Budapest University of Technology and Economics
Budapest, Hungary
{revy, hadhazi, hullam.gabor}@mit.bme.hu

***Abstract*—Automatic vertebrae localization and segmentation in computed tomography (CT) are fundamental for computer-aided detection (CADe) and computer-aided diagnosis (CADx) systems. However, they remain challenging due to the high variation in spinal anatomy among patients. In this paper, we propose a simple, model-free approach for automatic CT vertebrae localization and segmentation. The segmentation pipeline consists of 3 stages. In the first stage the center line of the spinal cord is estimated using convolution. In the second stage a baseline segmentation of the spine is created using morphological reconstruction and other classical image processing algorithms. Finally, the baseline spine segmentation is refined by limiting its boundaries using simple heuristics based on expert knowledge. We evaluated our method on the COVID-19 subdataset of the CTSpine1K dataset. Our solution achieved a dice coefficient of 0.8160±0.0432 (mean±std) and an intersection over union of 0.6914±0.0618 for spine segmentation. The experimental results have demonstrated the feasibility of the proposed method in a real environment.**

***Index Terms*—spine segmentation, CT, image processing, expert system**

I. Introduction

The segmentation of the rib cage can be an important stage in computer-aided detection (CADe) and computer-aided diagnosis (CADx) systems. In some cases, the spine, as part of the rib cage, may even need to be segmented separately. Current solutions for dealing with this task generally use explicit models of some kind. However, to create sufficiently robust models, large amounts of well-labelled heterogeneous data are required. This is not always available. Another drawback of these solutions is that they focus on segmenting the vertebrae individually, thus require a manually defined bounding box, which makes these algorithms only semi-automatic. Furthermore, the individual labeling is not necessary for all applications. In this work, we propose a fully-automatic, explicit model-free algorithm for spine segmentation that utilizes medical expert knowledge.

II. Related work

The segmentation of the spine is a currently active area of research. These solutions usually utilize a model to create an accurate segmentation. The two main approaches are the neural network and the statistical model-based solutions. There are several studies [1]–[4] investigating the use of U-Net [5] for spine segmentation. Cheng *et al.* [6] designed a two-stage Dense-U-Net for automatic CT vertebrae localization and segmentation. In the first step the centroid of the vertebrae was localized on the 2D slices using a 2D-Dense-U-Net. Based on the centroids, a 3D-Dense-U-Net is utilized to segment the vertebrae within the region-of-interests that are detected during the first stage. Vania *et al.* [7] combined a convolutional neural network (CNN) and a fully convolutional network (FCN). They also utilize class redundancy to improve the segmentation results. This means that the area around the spine is masked with a different value in the training data, forcing the model to accurately distinguish between the spine and its surroundings.

To determine a statistical shape model (SSM), Dryden *et al.* [8] utilized labeled landmark points of the examined shape (body). This model is fitted to the detected landmark points in case of the actual problem by registration. Benameur *et al.* [9] use a 2.5D statistical registration model from biplanar radiographic images for the 3D reconstruction of the vertebrae of a scoliotic spine.

Khandelwal *et al.* [10] designed a two-step pipeline for vertebrae segmentation utilizing the active shape model algorithm [11], which is a type of statistical shape analysis. In the first step of their pipeline the entire spine is segmented as a single surface object using region-based geometric flows. This step requires a seed point identified by the user. In the second step, the geometry of the vertebrae is taken into account during the segmentation of the individual vertebrae. They utilize a shape prior that is fitted based on the density values of the surrounding area of the spine.

III. Method

The segmentation consists of 3 main steps. First, the center of the spinal cord is estimated on each axial slice. Then, morphological reconstruction is applied to create a baseline segmentation of the spine. Finally, this segmentation is refined to obtain the final segmentation.

A. Spinal cord center estimation

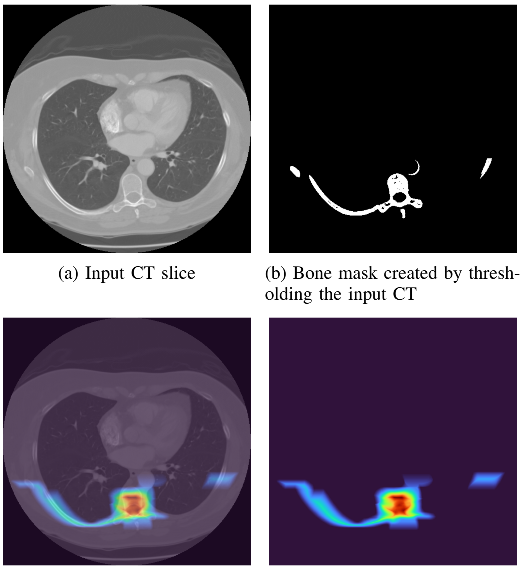
To estimate the center of the spinal cord in the axial slices, first, the Hounsfield unit (HU) values (see Figure 1a) are thresholded to highlight the values above bone density. We set this threshold to 150 HU to ensure that the bone values are included in the segmentation even in the presence of noise. You can see the resulting segmentation mask in Figure 1b.

This research was funded by the National Research, Development, and Innovation Fund of Hungary under Grant TKP2021-EGA-02.

DOI 10.3311/minisy2023-007

Although with this solution, the thresholding is robust, if contrast material is present, the heart and other blood vessels are also highlighted. To address this issue, only the posterior half of the body is taken into account on the axial slices. Another problem caused by the presence of the contrast agent concerns the descending aorta. The descending aorta runs close to the spine thus its segmentation might merge with the segmentation of the spine. This problem is solved by removing the segmentation of the descending aorta from the resulting mask. To create the segmentation of the descending aorta a Hough transform-based method [12] is utilized.

The center of the spinal cord is localized using convolution. The mask created by thresholding is convolved with a disk kernel on each axial slice. The diameter of the kernel is set to match the size of an average vertebra. The result of the convolution step is shown in Figure 1d.



(a) Input CT slice

(b) Bone mask created by thresholding the input CT

(c) The spine localization map with overlay

(d) The result of the convolution on which the center of the spine is localized

Fig. 1: The steps of estimating the center of the spinal cord. LIDC_0018 [13]

To make the localization more robust the sagittal slice that contains the largest segmented area is selected. This is expected to select the sagittal slice at the center of the spine. This sagittal slice defines a column in the axial slices, with a band around it in which we look for the center. In this band, on each slice, the point with the largest value obtained from the convolution is selected. The resulting points are smoothed by polynomial fitting using the RANSAC [14] method with L1 norm fidelity term, which makes it robust to outliers. The resulting cord center is shown in Figure 2.

B. Baseline spine segmentation

The segmentation created by thresholding the density values and the resulting estimated spinal cord center points are utilized to create a baseline spine segmentation. First, rays

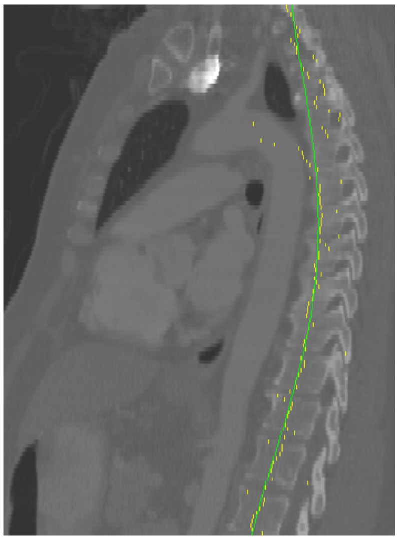


Fig. 2: The estimated spinal cord center points (yellow) and the fitted polynomial (green) projected onto the center sagittal slice determined by the algorithm. LIDC_0018 [13]

are cast from the center points to the anterior and posterior directions. The points are recorded where the rays intersect the segmentation. Points far from the center are discarded. The remaining points serve as seed points for morphological reconstruction [15], which is also based on the mask produced by thresholding (bone mask). The morphological reconstruction is performed on each slice separately. At this point, the voxels belonging to the spinous process might not be included in the reconstructed mask, since they might belong to a component other than the seed points of the reconstruction on the bone mask. This can be solved by dilating the reconstructed mask across the slices, masking it with the bone mask (using binary *AND* operation) and applying another morphological reconstruction. This solution had two drawbacks. The first problem is that the segmentation of the descending aorta is not perfect and the edge of the descending aorta might be included in the bone mask. Therefore, the dilation across the slices combined with the second reconstruction may result in a segmentation that flows into the unsegmented part of the descending aorta and into the posterior part of the heart. This is solved by applying the dilation across the slices only in the inferior direction (i.e. away from the head). This means that the segmentation of a slice is affected only by the segmentation of the slices in superior direction (i.e. towards the head). This way the imprecise segmentation of the descending aorta cannot flow into the superior slices. The other problem is similar: part of the descending aorta and other small vessels next to the vertebral body may be included in the bone mask. This can lead to the inclusion of unnecessary objects in the segmentation after the second morphological reconstruction is applied. To overcome this phenomenon the iteration number of the morphological reconstruction is limited in the anterior direction from the center point. An example of the resulting

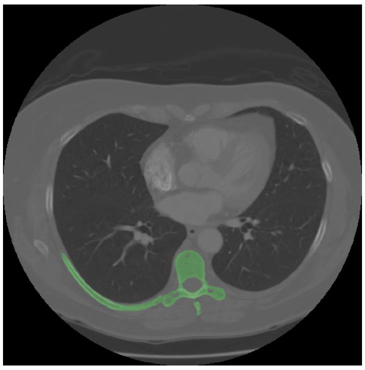


Fig. 3: The resulting baseline spine segmentation. This segmentation is refined by subsequent steps to reduce the segmentation of the ribs. LIDC_0018 [13]

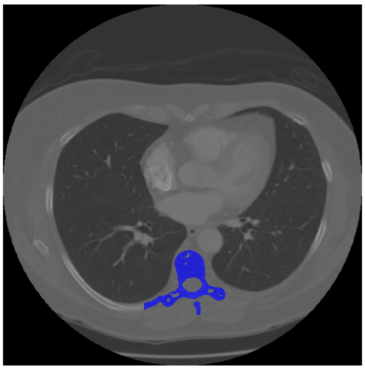


Fig. 4: Example for a resulting spine segmentation. LIDC_0018 [13]

segmentation can be seen in Figure 3.

C. Limiting the boundaries of the segmentation

As it can be seen in Figure 3, the segmentation resulting in the previous step includes unnecessary parts: typically the ribs are included in the segmentation. The performed steps described in this section aim to reduce this type of error. These algorithms utilize expert medical knowledge to determine bounding lines for the segmentation mask on the axial slices.

We start with the reconstructed mask and the spinal cord center from the previous step to create these lines. In the axial slices, this reconstructed mask is used to determine the top of the vertebral body. Several columns of this mask are selected in a band around the estimated center of the spinal cord. The upper part of the segmentation in this band is considered the most anterior part of the vertebral body.

Maximum intensity projection is performed perpendicular to the sagittal slices, followed by thresholding to determine the end of the spinous process. This can be obtained from the segmentation mask directly. Since the end of the spinous process can be found further back than the ribs, the most posterior point of the resulting segmentation is considered to be the end of the spinous process.

In the next step, the cumulative sum of the number of segmentation pixels across the columns in each axial slice is first calculated, then the median of the cumulative sum is used to calculate a more accurate center column of the vertebral body.

The upper third of the vertebral body is used to measure the width of the vertebral body in each slice, which can be derived from the top of the vertebral body and the end of the spinous process. The calculated width of the vertebral body allows us to estimate the maximal width of the whole vertebra. To precisely separate the ribs from the vertebrae a model would be required. Here, we utilize a simple heuristics that ensures that the segmentation includes the transversal process and cuts off a significant part of the segmentation of the ribs. The edges of the vertebral body are moved by its width in both directions. This way two lines are defined from which the segmentation is discarded outwards. Figure 4 shows an example of the resulting segmentation.

D. Evaluation

To evaluate the performance of the proposed segmentation algorithm the COVID-19 [16], [17] subdataset of the CTSpine1K [18] dataset was used. The CT scans can be retrieved from TCIA [19]. The COVID-19 dataset consists of chest CT scans of full thoracic view from 632 patients with COVID-19 infections. 20 of these were selected and annotated in the CTSpine1K dataset. To select CT scans with full thoracic view for evaluation was crucial, since the segmentation pipeline relies on certain anatomical features. Furthermore, the CT scans with reduced field of view are usually obtained to segment the vertebrae individually and in such cases the localisation of the spine is not required.

Two metrics were used to indicate the performance of the algorithm:

- Intersection over union (IoU) is the ratio of the intersection and union of the segmentation mask defined by the algorithm (A) and the manual labeling (B):

$$\mathrm{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

- The Dice similarity coefficient (DSC) equals twice the intersection of the segmentation mask volumes divided by the sum of the volumes:

$$\mathrm{DSC}(A, B) = \frac{2|A \cap B|}{|A| + |B|} = \frac{2|A \cap B|}{|A \cup B| + |A \cap B|}$$

Table I shows the results of the evaluation.

	IoU	DSC
avg	0.6914	0.8160
std	0.0618	0.0432
min	0.5654	0.7223

TABLE I: Evaluation results of the proposed segmentation algorithm.

Table II shows the evaluation results compared to methods introduced in the related work section. The comparison shows that the accuracy of the presented method is at the bottom of the range. However, it is worth noting, that [3] used only 9 scans for training and 1 scan for evaluating their neural network. Furthermore, [1] and [2] evaluated their model on the training dataset. In summary, most of these models achieve a high accuracy based on a relatively small dataset of samples with similar characteristics. However, their application to datasets with different properties may require additional steps, such as the retraining and fine-tuning of models. In contrast, our method is more robust, it does not require a labeled training dataset, and thus no retraining step is needed.

paper	IoU	DSC	method
[1]	0.7228	0.8477	U-Net + ASPP
[2]	0.9433	0.9708	3D Dense-U-Net Network
[3]	0.9193	0.9580	RAR-U-Net
[6]	0.911	0.953±0.014	two-stage Dense-U-Net
[7]	0.9336	0.9428	CNN
[10]	-	0.9236	Geometric Flows + Shape Priors
ours	0.6914	0.8160±0.0432	classical image processing

TABLE II: Comparison of results with related algorithms.

IV. Conclusion

In this paper, we have proposed an explicit model-free segmentation technique for spinal segmentation. The segmentation system uses a CT scan of full thoracic field of view as its input. Classical image processing algorithms that exploit medical expertise were utilized. Our approach has the advantage of not requiring an explicit model and it is fully automatic. This method can be applied when a fast and simple segmentation is preferred and it is not critical that the ribs do not completely separate from the spine on the segmentation. In such cases, however, when individual vertebral segmentation is required, a model-based approach is necessary.

Acknowledgement

The authors would like to thank the Multi-National NIH Consortium for CT AI in COVID-19. The authors acknowledge the National Cancer Institute and the Foundation for the National Institutes of Health, and their critical role in the creation of the free publicly available [20] LIDC/IDRI Database [13] used in this study. This research was funded by the National Research, Development, and Innovation Fund of Hungary under Grant TKP2021-EGA-02.

References

[1] Z. Yang, L. Chen, T. Fu, Z. Yin, and F. Yang, “Spine image segmentation based on u-net and atrous spatial pyramid pooling,” in *Journal of Physics: Conference Series*, vol. 2209, no. 1. IOP Publishing, 2022, p. 012020.

[2] M. Kolařík, R. Burget, V. Uher, K. Říha, and M. K. Dutta, “Optimized high resolution 3d dense-u-net network for brain and spine segmentation,” *Applied Sciences*, vol. 9, no. 3, p. 404, 2019.

[3] Z. Wang, Z. Zhang, and I. Voiculescu, “Rar-u-net: a residual encoder to attention decoder by residual connections framework for spine segmentation under noisy labels,” in *2021 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2021, pp. 21–25.

[4] C. Buerger, J. von Berg, A. Franz, T. Klinder, C. Lorenz, and M. Lenga, “Combining deep learning and model-based segmentation for labeled spine ct segmentation,” in *Medical Imaging 2020: Image Processing*, vol. 11313. SPIE, 2020, pp. 307–314.

[5] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[6] P. Cheng, Y. Yang, H. Yu, and Y. He, “Automatic vertebrae localization and segmentation in ct with a two-stage dense-u-net,” *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021.

[7] M. Vania, D. Mureja, and D. Lee, “Automatic spine segmentation from ct images using convolutional neural network via redundant generation of class labels,” *Journal of Computational Design and Engineering*, vol. 6, no. 2, pp. 224–232, 2019.

[8] I. L. Dryden and K. V. Mardia, *Statistical shape analysis: with applications in R*. John Wiley & Sons, 2016, vol. 995.

[9] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise, “3d/2d registration and segmentation of scoliotic vertebrae using statistical models,” *Computerized Medical Imaging and Graphics*, vol. 27, no. 5, pp. 321–337, 2003.

[10] P. Khandelwal, D. L. Collins, and K. Siddiqi, “Spine and individual vertebrae segmentation in computed tomography images using geometric flows and shape priors,” *Frontiers in Computer Science*, p. 66, 2021.

[11] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, “Active shape models-their training and application,” *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[12] U. Kurkure, O. C. Avila-Montes, and I. A. Kakadiaris, “Automated segmentation of thoracic aorta in non-contrast ct images,” in *2008 5th IEEE Int. Symp. on Biomed. Imaging*. IEEE, 2008, pp. 29–32.

[13] Armato III, Samuel G., G. McLennan, L. Bidaut *et al.*, “Data from lidc-idri,” 2015. [Online]. Available: https://wiki.cancerimagingarchive.net/x/rgAe

[14] H. Cantzler, “Random sample consensus (ransac),” *Institute for Perception, Action and Behaviour, Division of Informatics, University of Edinburgh*, vol. 3, 1981.

[15] K. Robinson and P. F. Whelan, “Efficient morphological reconstruction: a downhill filter,” *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1759–1767, 2004.

[16] P. An, S. Xu, S. A. Harmon, E. B. Turkbey *et al.*, “Ct images in covid-19,” 2020. [Online]. Available: https://wiki.cancerimagingarchive.net/x/o5QvB

[17] S. A. Harmon *et al.*, “Artificial intelligence for the detection of covid-19 pneumonia on chest ct using multinational datasets,” *Nature communications*, vol. 11, no. 1, pp. 1–7, 2020.

[18] Y. Deng, C. Wang, Y. Hui, Q. Li, J. Li, S. Luo, M. Sun, Q. Quan, S. Yang, Y. Hao *et al.*, “Ctspine1k: A large-scale dataset for spinal vertebrae segmentation in computed tomography,” *arXiv e-prints*, pp. arXiv–2105, 2021.

[19] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle *et al.*, “The cancer imaging archive (tcia): maintaining and operating a public information repository,” *Journal of digital imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.

[20] K. Clark, B. Vendt *et al.*, “The cancer imaging archive (TCIA): Maintaining and operating a public information repository,” *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1045–1057, Jul. 2013. [Online]. Available: https://doi.org/10.1007/s10278-013-9622-7