



# Deep Learning based Segmentation of Lumbar Vertebrae from CT Images

Rens Janssens and Guoyan Zheng

Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland  
guoyan.zheng@istb.unibe.ch

## Abstract

We present a method to address the challenging problem of automatic segmentation of lumbar vertebrae from CT images acquired with varying fields of view. Our method is based on cascaded 3D Fully Convolutional Networks (FCNs) consisting of a localization FCN and a segmentation FCN. More specifically, in the first step we train a regression 3D FCN (we call it “LocalizationNet”) to find the bounding box of the lumbar region. After that, a 3D U-net like FCN (we call it “SegmentationNet”) is then developed, which after training, can perform a pixel-wise multi-class segmentation to map a cropped lumbar region volumetric data to its volume-wise labels. Evaluated on publicly available datasets, our method achieved an average Dice coefficient of  $95.77 \pm 0.81\%$  and an average symmetric surface distance of  $0.37 \pm 0.06$  mm.

## 1 Introduction

An accurate segmentation of individual vertebrae from CT images are important for many clinical applications. After segmentation, it is possible to determine the shape and condition of individual vertebrae. Segmentation can also assist early diagnosis, surgical planning and locating spinal pathologies like degenerative disorders, deformations, trauma, tumors and fractures. Most computer-assisted diagnosis and planning systems are based on manual segmentation performed by physicians. The disadvantage of manual segmentation is that it is time consuming and the results are not really reproducible because the image interpretations by humans may vary significantly across interpreters.

In this paper, we address the challenging problem of automatic segmentation of lumbar vertebrae from 3D CT images acquired with varying fields of view (FOV), which is usually solved with a two-stage method consisting of a localization stage followed by a segmentation stage [1]. The localization aims to identify each lumbar vertebra, where segmentation handles the problem of producing binary labeling of a given 3D image. For vertebra localization, there exist both semi-automatic methods and fully automatic methods [5]. For vertebra segmentation, both 2D image-based methods and 3D image-based methods are introduced before. These methods can be roughly classified as statistical shape model or atlas based methods [5], graph theory (GT) based methods [5], and deep learning-based methods [4].

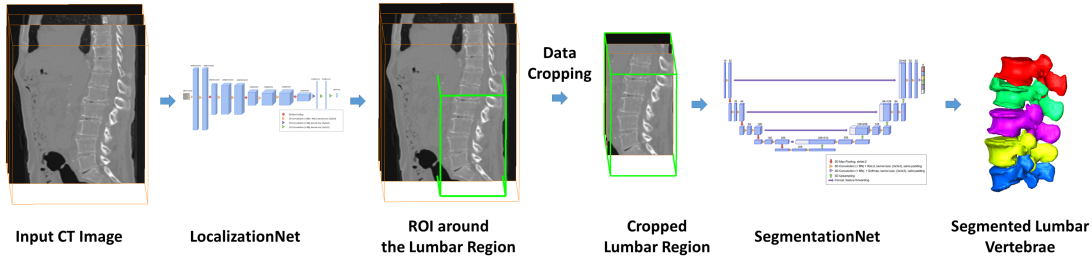


Figure 1: A schematic view of the present approach.

## 2 Materials and Methods

### 2.1 Methods

We propose a method to automatically segment lumbar vertebrae from 3D CT images using cascaded 3D FCNs consisting of a localization FCN and a segmentation FCN (see Fig. 1 for an overview). More specifically, in the first step we train a regression 3D FCN (the LocalizationNet in Fig. 1) to find a bounding box which defines the region of interest (ROI) of the lumbar region. After that, a 3D U-net like FCN (the SegmentationNet in Fig. 1) is then developed, which after training, can perform a pixel-wise multi-class segmentation to map a cropped lumbar region volumetric data to its volume-wise labels.

We formulate the lumbar region localization as a multi-variate regression problem. We use a rectangular box to represent the Region of Interest (ROI) of the lumbar region for each data, which can be represented by two diagonal corners of the rectangular box. The target of the regression is then the two relative displacement vectors between a reference voxel inside the volume and the two diagonal corners. In this study, we used Canny edge detector to select voxels with high edge responses as the reference voxels in both training and testing stages. We further propose to directly regress the target from a 3D patch sampled around a reference voxel using a deep FCN as shown in Fig. 1.

The estimated corners of the ROI will allow us to extract the lumbar region from an input CT data. The goal of this stage is then to segment each individual lumbar vertebra from the cropped data. To this end, we developed a segmentation net to conduct multi-class segmentation of the cropped lumbar region data. A 3D U-net like FCN was adopted here for our purpose. It consists of two parts, i.e., the encoder part (contracting path) and the decoder part (expansive path). The encoder part focuses on analysis and feature representation learning from the input data while the decoder part generates segmentation results, relying on the learned features from the encoder part. Shortcut connections are established between layers of equal resolution in the encoder and decoder paths.

### 2.2 Experimental Design

We obtained 15 spine CT images with ground truth segmentation from the MICCAI 2016 xVertSeg challenge<sup>1</sup>. We conducted a leave-three-out cross-validation study to evaluate the performance of the present method. More specifically, each time we randomly take 3 out of the 15 CT data as the test data and the remaining 12 CT data as the training data. The

<sup>1</sup>One can find details about the xVertSeg challenge at: <http://lit.fe.uni-lj.si/xVertSeg/>

Table 1: Segmentation results of the 5-fold cross validation on the xVertSeg challenge dataset.

	DC (%)	JC (%)	HD (mm)	ASSD (mm)
L1	96.14 ± 0.50	92.56 ± 0.91	2.97 ± 1.63	0.34 ± 0.03
L2	95.83 ± 0.72	92.01 ± 1.32	4.38 ± 2.43	0.35 ± 0.06
L3	95.89 ± 0.63	92.11 ± 1.15	4.64 ± 3.03	0.36 ± 0.04
L4	95.35 ± 0.89	91.12 ± 1.61	5.30 ± 3.33	0.41 ± 0.05
L5	95.66 ± 1.01	91.69 ± 1.80	4.34 ± 1.44	0.40 ± 0.07
Lumbar	95.77 ± 0.81	91.90 ± 1.48	4.32 ± 2.60	0.37 ± 0.06

process was repeated for 5 folds. In each fold, the segmented results of the test data were compared with the associated ground truth segmentation. For each vertebra in a test CT data, we evaluate the Average Symmetric Surface Distance (ASSD) and Hausdorff Distance (HD) between the surface models extracted from different segmentation as well as the volume overlap measurements including Dice Coefficient (DC) and Jaccard Coefficient (JC).

### 3 Results

Quantitative segmentation results of the cross validation study is shown in Table 1, where the results on each individual vertebra as well as on the entire lumbar region are presented. Our approach achieves a mean DC of  $95.77 \pm 0.81\%$  and a mean ASSD of  $0.37 \pm 0.06\text{mm}$  on the entire lumbar region. After training, it took on average about 79 seconds to finish the segmentation of one test CT data.

### 4 Discussions

The results achieved by our method are comparable to those by the state-of-the-art methods, though direct comparison of different methods is difficult as not all of them are evaluated on the same dataset. For example, when evaluating their method on 50 healthy lumbar vertebrae from 10 spinal CT data, Ibragimov et al. [2] reported a mean DC of 93.6%. On the same lumbar vertebral dataset, Korez et al. [3] reported a mean DC of 95.3%. In contrast, even evaluated on a challenging dataset with fractured vertebrae and varying FOVs, our method achieved a mean DC of 95.77%. In comparison to the method introduced in [4], our approach also achieved superior results. A mean DC of 92.7% was reported in [4] while our approach achieved a mean DC of 95.77%. As both methods are based on deep learning techniques, one possible explanation is that they have used different localization (multi-layered perceptron vs. our 3D regression FCN) and segmentation (2D U-net like FCN vs. our 3D segmentation FCN) methods from ours.

### References

- [1] C. Chu, D.L. Belavy, G. Armbrecht, M. Bansmann, D. Felsenberg, and G. Zheng. Fully automatic localization and segmentation of 3d vertebral bodies from ct/mr images via a learning-based method. *PLOS ONE*, 10(11):e0143327, 2015.
- [2] B. Ibragimov, B. Likar, F. PERNUS, and T. Vrtovec. Shape representation for efficient landmark-based segmentation in 3-d. *Medical Imaging, IEEE Transactions on*, 33(4):861–874, 2014.

- [3] R Korez, B Ibragimov, B Likar, F Pernus, and T Vrtovec. A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation. *IEEE Trans Med Imaging*, 34:1649–1662, 2015.
- [4] A. Sekuboyina, A. Valentinitich, J.S. Kirschke, and Menze B.H. A localisation-segmentation approach for multi-label annotation of lumbar vertebrae using deep nets. *arXiv:1703.04347*, 2017.
- [5] J Yao, J.E. Burns, D. Forsberg, and et al. A multi-center milestone study of clinical vertebral ct segmentation. *Computerized Medical Imaging and Graphics*, 49:16–28, 2016.