
End-to-End Fully Automatic Segmentation of Vertebrae in Spinal X-Ray Images

Abdullah-Al-Zubaer Imran^{2,1}, Chao Huang¹, Hui Tang¹, Wei Fan¹,
Kenneth M.C. Cheung³, Michael To³, Zhen Qian¹, Demetri Terzopoulos^{2,4}

¹Tencent Medical AI Lab, Palo Alto, CA, USA

²University of California, Los Angeles, CA, USA

³The University of Hong Kong, China

⁴VoxelCloud, Inc., Los Angeles, CA, USA

Abstract

Accurate vertebral identification and labeling is essential in image-guided spinal disease diagnosis and treatment planning. Unfortunately, spinal assessments traditionally rely on tedious and time-consuming manual measurement subject to observer variability. In particular, the measurement of scoliosis requires the segmentation of individual vertebrae, but an automatic method that can accurately identify and segment vertebrae is unavailable in the literature. We introduce an end-to-end fully automatic segmentation method that leverages a carefully-adjusted U-Net model with progressive side outputs in order to provide reliable segmentations of the vertebrae associated with scoliosis measurement. Our experimental results with X-ray images of scoliosis patients indicate that our method, which achieves an average Dice score of 0.993, promises to be an effective tool in the identification and labeling of vertebrae for the reliable estimation of scoliosis.

1 Introduction

Conventional spine image analysis tasks, particularly for the measurement of scoliosis, an abnormal lateral curve to the vertebral column, involve tedious manual labor subject to imprecision and variability. In addition to the landmark detection approach [4], vertebrae segmentation methods are available in literature. Existing computer-aided vertebral segmentation methods rely on manual interaction [7], hand-crafted feature engineering [10], patch-based approaches that lose full spatial context [8, 1], limited scope that fails to consider all the required vertebrae at once [6], etc. As a departure from previous vertebral segmentation approaches, we introduce a method that involves no manual intervention and is end-to-end, avoiding any pre-/post-processing steps.

2 Method

Following the progressive side-outputs in segmentation model [2, 5, 3], we propose a progressive U-Net with some careful adjustments in the U-Net. As in a U-Net [9], our model has an encoder and a decoder with skip connections. In each encoder layer, two 3×3 convolutions are followed by instance normalization, ReLU activation, and a 2×2 max-pooling. A dropout of 0.25 is applied in every encoder and decoder stage of the network. We generate side-outputs in every stage of the decoder. Progressively adding one side-output to the next improves the segmentation performance compared to collecting the output from the final decoder stage [3]. The progressive side-outputs also ensure that micro-structure is preserved at all levels of the decoder.

Table 1: Performance comparison of the vertebrae segmentation models

Model	DI	SSIM	HD	F1
UD	0.970	0.961	5.246	0.896
UX	0.956	0.955	6.767	0.868
PUD	0.993	0.966	4.597	0.919
PUX	0.993	0.970	4.677	0.922
PUDA	0.992	0.966	4.181	0.913
PUXA	0.992	0.967	4.956	0.911

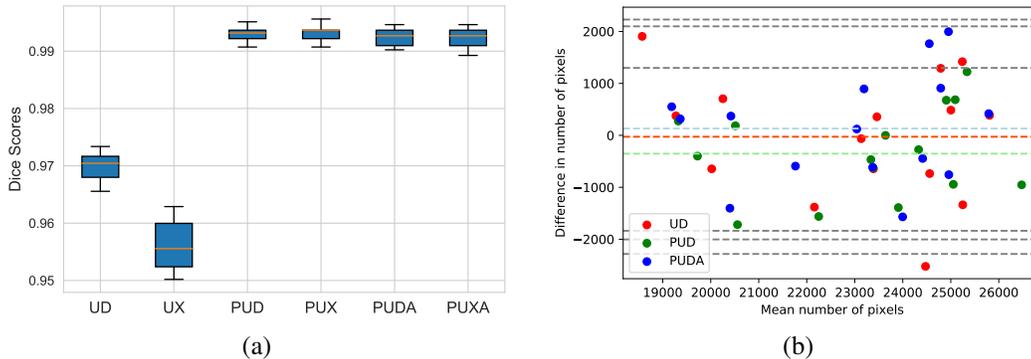


Figure 1: (a) Whisker-Boxplots of all the six models showing consistent performance of the proposed model with varying losses and data settings; (b) Bland-Altman plots show better agreement of the proposed model among the models with Dice loss in segmenting 18 vertebrae from the scoliotic X-ray test set.

3 Implementation Details

We used a dataset of 100 high-resolution spine X-ray images of children with evidence of various extents of scoliosis. The dataset contains manual annotation of 18 relevant vertebrae: C7, T1–T12, and L1–L5. We split the dataset into training (80), testing (15), and validation (5) sets. All the images were resized and normalized to $1024 \times 512 \times 1$ before feeding them to the networks. All models were trained on the training set while their performances were evaluated on the testing set. The validation set was used for hyper-parameter tuning and model selection. We used the Adam optimizer with adaptive learning rate, starting with an initial rate of 0.01 and decreasing 10 times after every 20 epochs.

As baselines, we used a regular U-Net model with binary cross-entropy (XE) and Dice loss functions. Moreover, to verify the robustness of our model, we experimented with training data augmentation (varying rotation, flipping, contrast, Gaussian noise, etc.). For simplicity, we refer to the baseline model as UD (UNet with Dice loss) and UX (UNet with XE loss), and to our model as PUD (Progressive UNet with Dice loss), PUX (Progressive UNet with XE loss), PUDA (Progressive UNet with Dice loss and data augmentation), and PUXA (Progressive UNet with XE loss and data augmentation).

4 Experimental Results

Our experimental results, based on both qualitative and quantitative evaluations, confirm the superiority of our Progressive UNet model. It consistently provides improved segmentation with varying losses (Dice and XE) and data settings (with or without augmentation). For evaluation, we used the Dice index (DI), structural similarity index (SSIM), average Hausdorff distance (HD), and F1 score (F1), as well as the qualitative visualization of masks and boundaries. In all four quantitative measures, our models achieve better scores than the baseline models (Table 1). The superiority of the proposed models is further confirmed with the visualization of boxplots and Bland-Altman plots (Figure 1). Visualizations of the segmented vertebrae boundaries show better delineation of individual vertebrae merely from the binary segmentation masks (Figure 2).

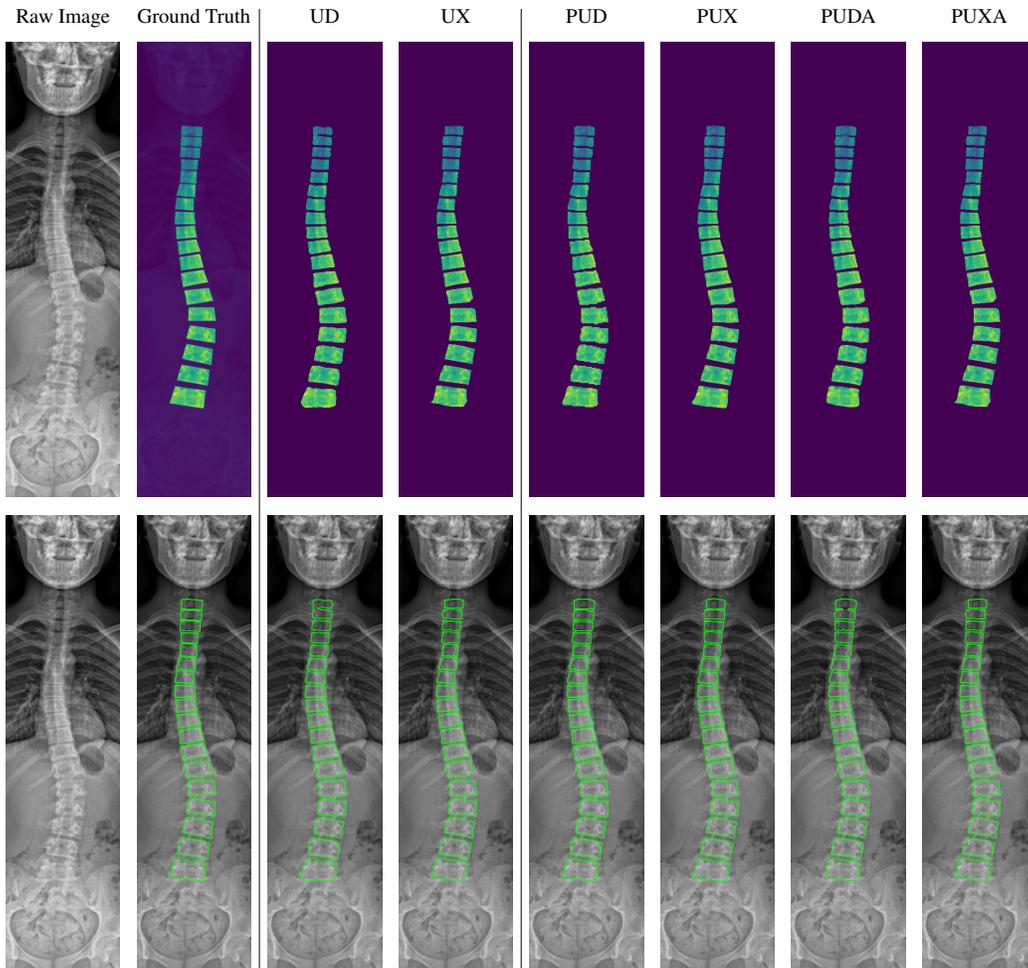


Figure 2: Visualization of the predicted vertebrae segmentation masks (top) and boundaries (bottom) in a spinal X-ray shows consistent improvement by our proposed PU model over the baseline model.

5 Conclusions

Accurate and reliable segmentation of vertebrae is a pre-requisite for the effective measurement of scoliosis. We proposed an end-to-end fully automatic model that can accurately segment the vertebrae in X-Ray images of scoliotic patients.

References

- [1] Horng, M.-H., Kuok, C.-P., Fu, M.-J., Lin, C.-J. and Sun, Y.-N. [2019], ‘Cobb angle measurement of spine from X-ray images using convolutional neural network’, *Computational and Mathematical Methods in Medicine* **2019**.
- [2] Imran, A.-A.-Z., Hatamizadeh, A., Ananth, S. P., Ding, X., Tajbakhsh, N. and Terzopoulos, D. [2019], ‘Fast and automatic segmentation of pulmonary lobes from chest ct using a progressive dense v-network’, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* .
- [3] Imran, A.-A.-Z., Hatamizadeh, A., Ananth, S. P., Ding, X., Terzopoulos, D. and Tajbakhsh, N. [2018], Automatic segmentation of pulmonary lobes using a progressive dense V-network, in ‘Deep Learning in Medical Image Analysis’, Vol. 11045 of *Lecture Notes in Computer Science*, Springer, pp. 282–290.

- [4] Imran, A.-A.-Z., Huang, C., Tang, H., Fan, W., Cheung, K. M., To, M., Qian, Z. and Terzopoulos, D. [2019], Bipartite distance for shape-aware landmark detection in spinal X-ray images, *in* 'Medical Imaging Meets NeurIPS Workshop', Vancouver, Canada.
- [5] Imran, A.-A.-Z. and Terzopoulos, D. [2019], Semi-supervised multi-task learning with chest x-ray images, *in* 'MICCAI Machine Learning in Medical Imaging', Vol. 11861, p. 151.
- [6] Lessmann, N., van Ginneken, B., de Jong, P. A. and Išgum, I. [2019], 'Iterative fully convolutional neural networks for automatic vertebra segmentation and identification', *Medical Image Analysis* **53**, 142–155.
- [7] Mateusiak, M. and Mikolajczyk, K. [2019], Semi-automatic spine segmentation method of CT data, *in* R. Szewczyk, J. Krejsa, M. Nowicki and A. Ostaszewska-Lizewska, eds, 'Mechatronics 2019: Recent Advances Towards Industry 4.0'.
- [8] Qadri, S. F., Zhao, Z., Ai, D., Ahmad, M. and Wang, Y. [2019], Vertebrae segmentation via stacked sparse autoencoder from computed tomography images, *in* '11th International Conference on Digital Image Processing (ICDIP 2019)', Vol. 11179 of *Proceedings of the SPIE*, pp. 4K:160–166.
- [9] Ronneberger, O., Fischer, P. and Brox, T. [2015], U-net: Convolutional networks for biomedical image segmentation, *in* 'International Conference on Medical Image Computing and Computer-Assisted Intervention', Springer, pp. 234–241.
- [10] Taghizadeh, E., Terrier, A., Becce, F., Farron, A. and Büchler, P. [2019], 'Automated CT bone segmentation using statistical shape modelling and local template matching', *Computer Methods in Biomechanics and Biomedical Engineering* **0**(0), 1–8.