Scoliosis Measurement on DXA Scans Using a Combined Deep Learning and Spinal Geometry Approach

Emmanuelle Bourigault [1] EMMANUELLE@ROBOTS.OX.AC.UK
Amir Jamaludin [1] AMIRJ@ROBOTS.OX.AC.UK
Timor Kadir [2] TIMOR.KADIR@PLEXALIS.COM
Andrew Zisserman [1] AZ@ROBOTS.OX.AC.UK
[1] *Visual Geometry Group, Department of Engineering Science, University of Oxford*
[2] *Plexalis Ltd*

Editors: Under Review for MIDL 2022

Abstract

We propose improvements to an automated method for scoliosis measurement. Our main novelty is the use of a spline to better model the curve of the spine, and we employ pseudo-labelling to re-train the segmentation step to mitigate the domain gap when adapting to a new dataset. We obtain promising results with a good fit of our smoothed curve to approximate the spinal midpoints in severe scoliosis cases, and obtain good agreement against human ground-truth. This work is relevant for improving the severity grading of scoliosis and potentially aiding in the treatment management of scoliosis.

Keywords: Scoliosis measurement, image segmentation, deep learning, computational geometry

1. Introduction

Scoliosis is a spinal deformation affecting between 2 to 3% of the population worldwide. Most automated methods for scoliosis measurement focus on Adolescent Idiopathic Scoliosis (AIS) but in this work our aim is to measure scoliosis in adults using DXA scans in the UK Biobank. We use the work of (Jamaludin et al., 2019), which showed curvature is a reliable method to measure scoliosis, as a base and propose several adjustments. Note that the model built by (Jamaludin et al., 2019) was validated on a completely different dataset, ALSPAC (The Avon Longitudinal Study of Parents and Children), from the UK Biobank with 2 main differences: (i) ALSPAC scans are lower in resolution (5× lower) compared to the UK Biobank, and (ii) ALSPAC participants are all adolescents while the UK Biobank is predominantly made up of adult participants. We make two contributions: (i) spline geometric representation and curvature measurement of the spine, rather than the mid-point of the spine segmentation used in (Jamaludin et al., 2019); and (ii) pseudo-labelling the segmentation to address the domain gap between ALSPAC and UK Biobank DXA scans.

2. Methods

2.1. Dataset

The dataset comprises 48,384 whole body DXA scans from the UK Biobank split into 80:10:10 for training, validation and testing. Among the test set, 308 UK Biobank DXA scans have been annotated with the maximum modified Ferguson angle using the DXA scoliosis method (DSM) as outlined in (Taylor et al., 2013). We use this test set to evaluate the association between our predicted curvature and manually annotated angle.

© 2022 E.B. , A.J. , T.K. & A.Z. .

2.2. Pseudo-labelling

The purpose of our pseudo-labelling is to bootstrap from the ALSPAC model to the higher resolution UK Biobank images. The initial labels are obtained as follows: UK Biobank DXA scans are downsampled to the ALSPAC resolution, the ALSPAC model is applied and segmentation masks are upsampled. We first train our model on 80% of the training set and incrementally add 10% of the training data with each iteration. Further, the generated pseudo-labels are combined with the original labels to re-train the model.

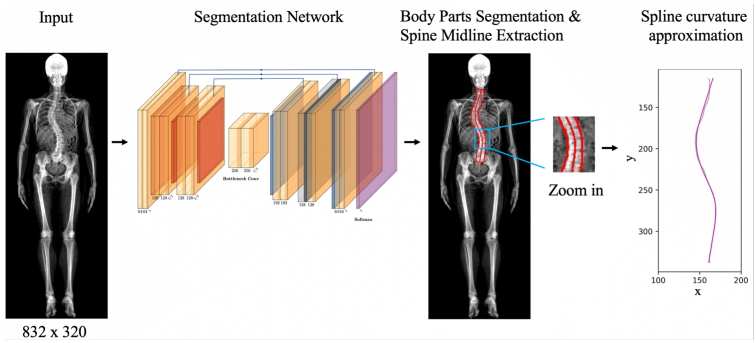

Figure 1: Pipeline of the automated scoliosis measurement method

2.3. Segmentation Task

We use as a baseline the ALSPAC model with input resolution (416 x 128). Then, we train our higher resolution segmentation model which is a U-Net like architecture (Ronneberger et al., 2015) depicted in Figure 1. We feed into the network the normalised UK Biobank DXA scan as tensor ($1 \times 832 \times 320$). The network attempts to predict the associated ground-truth masks for 6 body parts: head, spine, pelvis, cavity, left leg and right leg. The decoder outputs probability maps of size (6 x 832 x 320). At each row of the spine probability map, the weighted arithmetic mean of the probability and the indices of the scores is calculated to be the predicted midpoint.

2.4. Spline Fitting Task and Curvature Measurement

The extracted spinal midpoints from the spine probability maps can be slightly noisy especially near cases with unclear boundaries of the spine (see Figure 1). Hence, we propose a smoothed piecewise cubic spline that locally approximates the spinal midpoints and effectively handles the noise in the data. At each segment, a cubic polynomial is used. Constraints on continuity and smoothness at the knots are added. Piecewise cubic splines are determined by minimising the sum of the weighted squared residuals between the fitted spline and data points (Ezhov et al., 2018). The number of interior knots, k, controls the goodness of fit. If k is too small, the spline will be too smooth, if k is too large, the spline will capture the noise.

We compute curvature at each point along the curve with the traditional formula for plane curves. Then, we compare the maximum curvature with human annotated angles. We assume that the maximum curvature is proportional to the maximum angle.

3. Results

3.1. Segmentation Task

The Root Mean Squared Error (RMSE) is computed as the root mean squared nearest Euclidian distance between predicted to manual contour points. The lower the RMSE, the

better is model performance. Table 1 shows the improvement in segmentation accuracy (Intersection over Union (IoU) = 0.94, RMSE = 0.83) with pseudo-labelling compared to the baseline model (IoU = 0.89, RMSE = 1.18). The IoU increases by 2% while the RMSE decreases by 0.12 at sub-pixel level precision for the iteration of the pseudo-labelling.

Models	**IoU**	**RMSE**
ALSPAC model on the UK Biobank	0.86	1.80
Model trained on UK Biobank with pseudo-labels no iteration	0.89	1.18
Model trained on UK Biobank with pseudo-labels with iterations	0.94	0.83

Table 1: Table of model segmentation performance with pseudo labelling

3.2. Curvature to Angle Correspondence

Comparison of maximum curvature to maximum human annotated angle suggests a good agreement (Pearson's $r = 0.92$) with a tendency for our method to underestimate the angle for angles below 20° (see Figure 2).

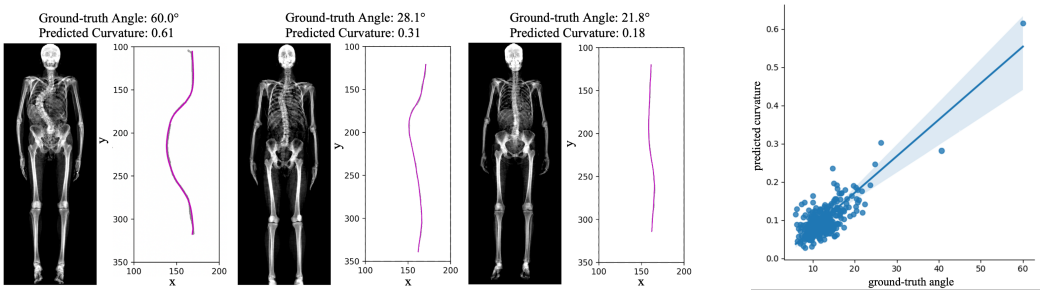

Figure 2: Comparison between predicted curvature and human annotated angles

4. Conclusion

We have implemented a model for the measurement of scoliosis by leveraging one of the largest DXA datasets available. The benefit of our pseudo-labelling in training is 2-fold, it improves the accuracy of the segmented spine compared to our baseline model, and it overcomes the low availability of human annotated data sets. The reduced computation required to produce the piecewise cubic spline curves is undoubtedly advantageous compared to traditional polynomial curve fitting. This work has implications for improving scoliosis detection and severity measurements in adolescents and adults.

References

N. Ezhov, F. Neitzel, and S. Petrovic. Spline approximation, part 1: Basic methodology. *Journal of Applied Geodesy*, 12:139–155, 2018.

A. Jamaludin, T. Kadir, E. Clark, and A. Zisserman. Predicting spine geometry and scoliosis from DXA scans. In *MICCAI Workshop: MSKI*, 2019.

O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241. Springer International Publishing, 2015.

H. Taylor, I. Harding, J. Hutchinson, I. Nelson, A. Blom, J. Tobias, and E.M. Clark. Identifying scoliosis in population-based cohorts: Development and validation of a novel method based on total-body dual-energy x-ray absorptiometric scans. *Calcif. Tissue Int.*, 92:539–547, 2013.